Impacts of mesh refinement on the simulation of a long-range transported extreme dust storm over East Asia in the global variable-resolution model (iAMAS v2.6.3)

Yibo Xue[1], Chun Zhao[1,2,3,4,5], Jiawang Feng[1], Xin Qu[1], Zihan Xia[1], Zining Yang[1], Xiaoxiao Zhang[6], Qike Yang[1], Gudongze Li[1], Ziyu Zhang[1], Jianxi Zhang[1]

Correspondence: Chun Zhao (chunzhao@ustc.edu.cn)

We sincerely thank the three anonymous reviewers for their supportive and thoughtful remarks.

Anonymous Referee #1

General comments

In this study, titled "Impacts of mesh refinement on the simulation of a long-range transported extreme dust storm over East Asia in the global variable-resolution model (iAMAS v2.6.3)" by Xue et al., the authors conducted three experiments, with the highest model resolution increased from 50 km to 16 km to 4 km, to simulate an extreme dust storm event in East Asia. The authors analyzed the impact of model resolution on simulated dust aerosol properties and processes.

Major comments

Question 1: In section 2.2, Dust emission scheme, the authors could highlight the factors affected by model resolution.
Reply: We have revised Section 2.2 to clarify the factors through which horizontal resolution affects dust emission **(see Section 2.2 on Page 8, Lines 206-210)**. The same S04 formulation and parameter settings are used in all three experiments; therefore, horizontal resolution does not directly modify the dust-emission parameterization. Instead, mesh refinement affects dust emission through resolution-dependent changes in the factors controlling emission, including topography, land-surface heterogeneity, and near-surface meteorological forcing. In particular, these changes influence the spatial variability of surface friction velocity and the frequency and magnitude of its exceedance above the saltation threshold, thereby affecting the magnitude and spatial distribution of simulated dust emissions.

Question 2: Line 179-180: Add more information about the dust source function (e.g., definition, range, physical meaning) and model resolution in Figure 1. Same for Figure 1 title. Do variable-resolution models use the same dust source function, with 50 km resolution?
Reply: We have revised Section 2.2 and the caption of Fig. 1 to provide the definition, physical meaning, value range, and horizontal resolution of the dust source function **(see Section 2.2 on Pages 8-9, Lines 216-223, and Fig. 1)**. The source function is defined as a dimensionless weighting factor representing the relative availability of erodible sediments and the associated dust-emission potential under otherwise identical forcing conditions (Ginoux et al., 2001). The source-function values in the dataset used here range from 0 to 0.38, with larger values indicating greater dust-source potential. The original source-function dataset has a horizontal resolution of approximately 0.7°, and the same dataset is mapped onto the three model meshes rather than being independently prescribed for each configuration. The mapped source-function fields remain within the original value range of 0-0.38. We have also added the following reference:

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

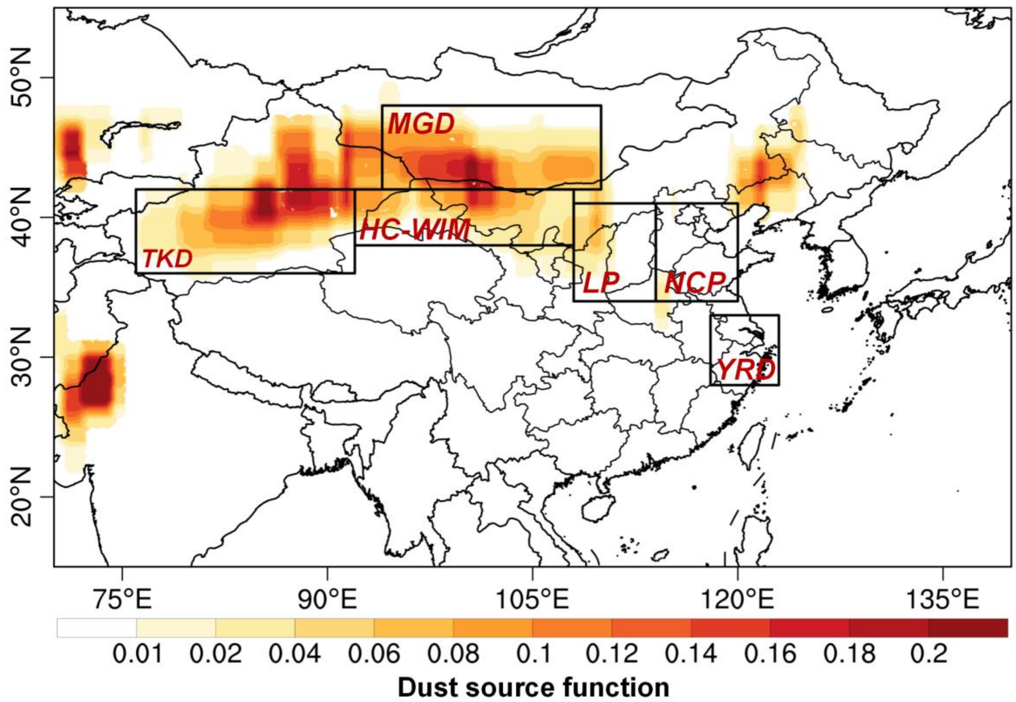

Fig. 1. Spatial distribution of the dust source function used in all three mesh configurations. This dimensionless field represents relative dust-source potential, with values ranging from 0 to 0.38 in the dataset used here. The source-function dataset has a horizontal resolution of approximately 0.7°. The boxes denote the major dust source and downwind regions analyzed in this study.

Question 3: Line 304-306: Since the resolution of ERA5 data is 0.25-degree and lower than that of variable-resolution models, how does the interpolation, from low to high resolution, affect the simulation results? Also, "Meteorological initial conditions and surface boundary fields are obtained from the ERA5 reanalysis at 6-hourly intervals", does the model read initial conditions every 6 hour?

Reply: We have revised Section 2.4 to clarify the influence of interpolating the 0.25° ERA5 fields onto the refined iAMAS meshes and to distinguish the initialization and reinitialization of the meteorological fields from the 6-hourly updating of surface boundary fields (**see Section 2.4 on Pages 13-14, Lines 347-374**). The ERA5 atmospheric fields are interpolated onto each native iAMAS mesh using the same preprocessing procedure, thereby providing a consistent large-scale initial atmospheric state for all three experiments. This interpolation does not create independent meteorological information at spatial scales finer than the original 0.25° ERA5 resolution. Therefore, although the ERA5 fields are represented on the 16 and 4 km meshes, their initial spatial variability remains constrained by the original ERA5 resolution, with limited representation of local gradients and mesoscale structures. During subsequent model integration, dynamically consistent finer-scale features can develop through the model dynamics and the resolution-dependent representation of topography, land-surface heterogeneity, and physical processes. Thus, the differences among the simulations reflect the combined effects of mesh-dependent initialization, subsequent model evolution, and the representation of atmospheric and surface processes. We have also corrected the inaccurate statement that "Meteorological initial conditions and surface boundary fields are obtained from the ERA5 reanalysis at 6-hourly intervals." The model does not read or replace the meteorological initial conditions every 6 h. The meteorological fields are initialized from ERA5 at 00:00 on 8 March and reinitialized once at 00:00 on 12 March, whereas the surface boundary fields are updated from ERA5 every 6 h. During the reinitialization on 12 March, the chemical tracers, including dust, are retained from the end of the spin-up period to ensure their continuity.

Question 4: Figure 2: The grid lines and differences in grid sizes in (a)-(c) are not clear.

Reply: We have revised and redrawn Fig. 2 (**see Fig. 2**). We added enlarged views covering the refined and transition regions to clearly illustrate the changes in mesh spacing and facilitate readers' understanding.

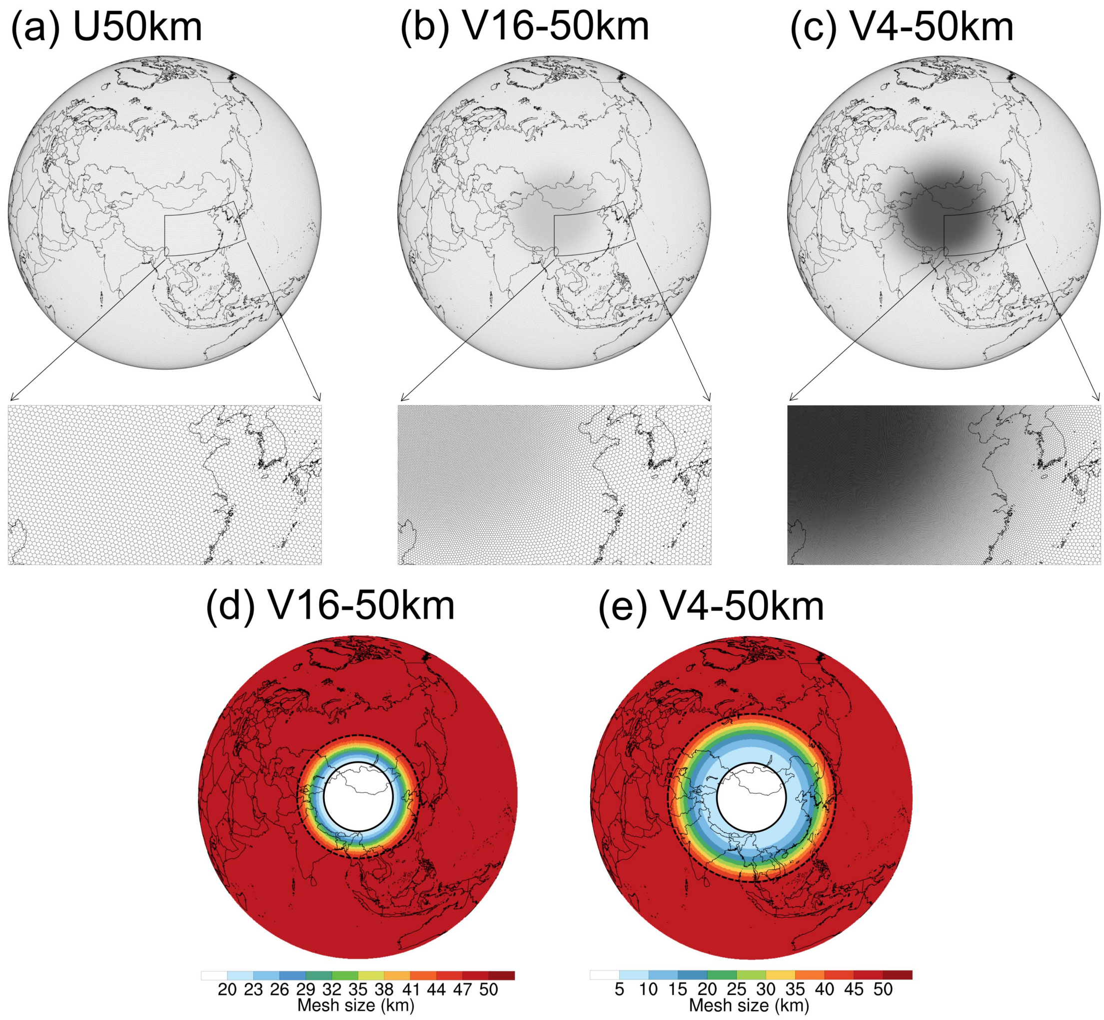

Fig. 2. (a) Global quasi-uniform mesh of 50 km (U50 km); (b, c) Global variable-resolution meshes of 16-50 km (V16 km) and 4-50 km (V4 km), refined over the major dust source regions in East Asia; (d, e) Spatial distributions of mesh spacing (cell size) for V16 km and V4 km. The black solid circles denote the refinement regions, and the black dashed circles

indicate the outermost boundary of the refinement transition zone.

Question 5: When evaluating the simulated large-scale circulation and wind speed, the authors could quantify the model performance using spatial correlation coefficients, scatterplot, root mean square error, etc. This can provide evidence for the descriptions such as “generally well reproduced”, “reasonably represented”, “comparable spatial pattern”, “broadly consistent” (Line 365-370). For 10-m wind speed, besides temporal correlation coefficients in Table 2, the scatterplot (like Figure 6) could be used to show how well the models capture the extreme wind speed in dust source regions.

Reply: We have revised the manuscript and added quantitative evaluations of the simulated large-scale circulation to support qualitative descriptions such as “generally well reproduced”, “reasonably represented”, “comparable spatial pattern”, and “broadly consistent” (**see Section 3.1.1 on Page 16, Lines 437-443; Table 2 and Table S1**). We have also added a scatterplot comparing simulated and observed daily mean 10 m wind speeds over the dust source regions and expanded the corresponding quantitative analysis in the manuscript **(see Section 3.1.2 on Pages 18-19, Lines 501-506; Fig. S5)**. The quantitative results support that all three simulations reproduce the broad synoptic-scale circulation, with spatial pattern correlations of 0.89-0.91 for mean sea level pressure and 0.97-0.98 for the 500 hPa geopotential height. The representation of the 850 hPa wind field generally shows closer agreement with ERA5 with mesh refinement: the wind-speed correlation increases from 0.69 in U50 km to 0.76 in V4 km, while the RMSE decreases from 3.69 to 3.29 m s^{-1}, and the corresponding vector correlation increases from 0.78 to 0.84 (Tables 2 and S1). For daily mean 10 m wind speed, the pooled station–day correlation increases from 0.64 in U50 km to 0.71 in V16 km and 0.75 in V4 km, while the NMB changes from +5.1% to –0.7% and –4.1%, respectively (Fig. S5). These results provide quantitative support for the descriptions of the simulated mid- and lower-tropospheric circulation and near-surface wind distribution.

Table 2. Evaluation of simulated mean sea level pressure, 500 hPa geopotential height, and 850 hPa wind speed against ERA5 during 13-16 March 2021.

Experiment	MSLP (hPa)			Z500 (gpm)			WS850 (m s^{-1})		
	r	RMSE	MB	*r*	RMSE	MB	*r*	RMSE	MB
U50 km	0.89	4.38	+1.84	0.97	27.80	−1.95	0.69	3.69	+0.75
V16 km	0.90	4.07	+1.82	0.98	24.15	−5.95	0.73	3.41	+0.44
V4 km	0.91	4.18	+2.02	0.98	23.60	−8.12	0.76	3.29	+0.36

Notes: *r* denotes the spatial pattern correlation coefficient. MSLP, Z500, and WS850 denote mean sea level pressure, 500 hPa geopotential height, and 850 hPa wind speed, respectively. The values are averages of the daily evaluation metrics calculated over 75−140°E and 33−57°N during 13-16 March 2021.

Table S1. Evaluation of the simulated 850 hPa wind vectors and wind components against ERA5 during 13-16 March 2021.

Experiment	Vector	U850 (m s^{-1})		V850 (m s^{-1})	
	r	RMSE	MB	RMSE	MB
U50 km	0.78	3.22	+0.63	3.72	−0.07
V16 km	0.82	2.92	+0.44	3.30	0.01
V4 km	0.84	2.81	+0.37	3.09	−0.01

Notes: Vector r denotes the area-weighted centered vector pattern correlation coefficient calculated jointly from the 850 hPa zonal and meridional wind components. U850 and V850 denote the 850 hPa zonal and meridional wind components, respectively. All values are averages of the daily evaluation metrics calculated during 13-16 March 2021.

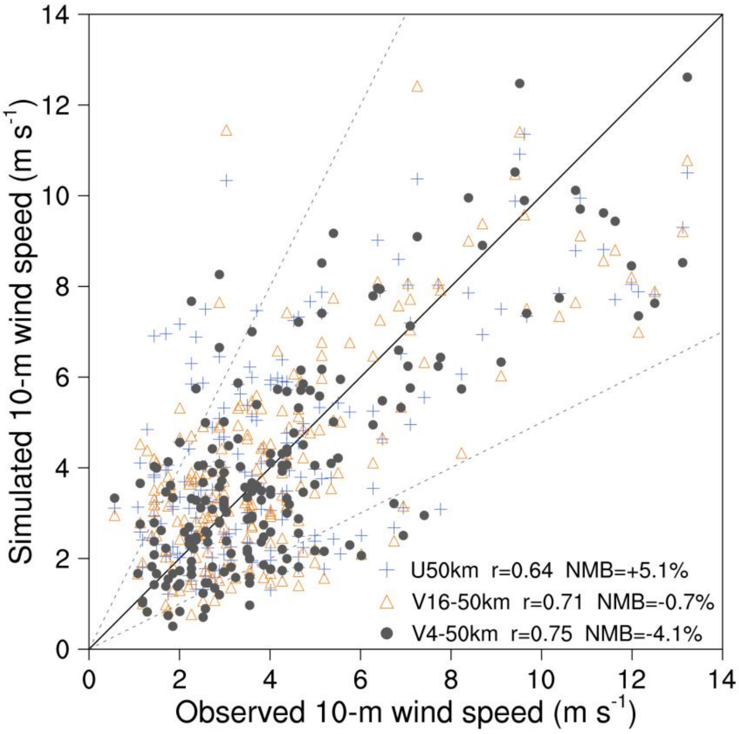

Fig. S5. Scatterplots of simulated versus observed daily mean 10-m wind speeds at collocated NOAA surface stations over the combined dust source regions during 13-16 March 2021. The solid line indicates 1:1 agreement, whereas the dashed lines mark the factor-of-two bounds. The correlation coefficient (r) and normalized mean bias (NMB) are calculated from all pooled station–day samples for each experiment.

Question 6: Line 367: For mean sea level pressure in TKD region, ERA5 shows a low pressure and wind convergence, but models do not capture this, especially during March 15-16. Since TKD is a key source of dust, this bias may affect dust transport. Are the model results regridded to ERA5 grids? If so, the authors could show the differences between simulations and ERA5 in supplementary information.

Reply: We have revised the inaccurate description and redrawn Figs. 3 and S4 after regridding the model fields onto the 0.25° ERA5 comparison grid. We have also added daily model-minus-ERA5 differences in mean sea level pressure and 850 hPa wind vectors in Fig. S2 (**see Section 3.1.1 on Pages 16-17, Lines 443-450; Fig. 3, S2 and S4**). The results show that, although the simulations reproduce the broad synoptic-scale circulation, they do not fully capture the low-pressure system and associated low-level wind convergence over the Taklimakan Desert (TKD), particularly on 15-16 March. We acknowledge this regional discrepancy and clarify that it may introduce uncertainty into the simulated local wind forcing, dust emission, and initial transport from this important source region.

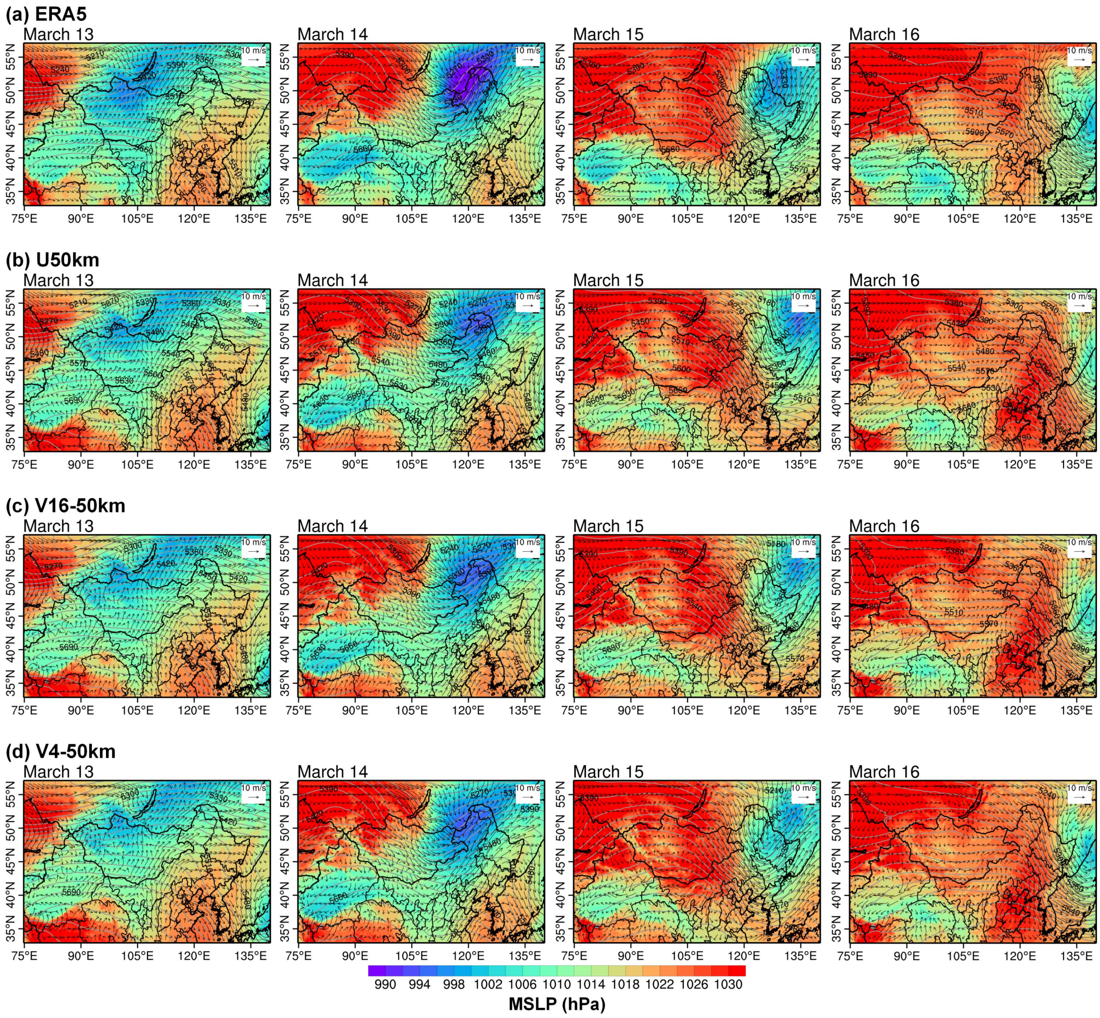

Fig. 3. Daily mean sea level pressure (shading, hPa), 500 hPa geopotential height (contours, gpm), and 850 hPa wind vectors for ERA5 and iAMAS simulations (U50 km, V16 km, and V4 km) during 13-16 March 2021. All model fields were interpolated onto the 0.25° ERA5 grid for comparison.

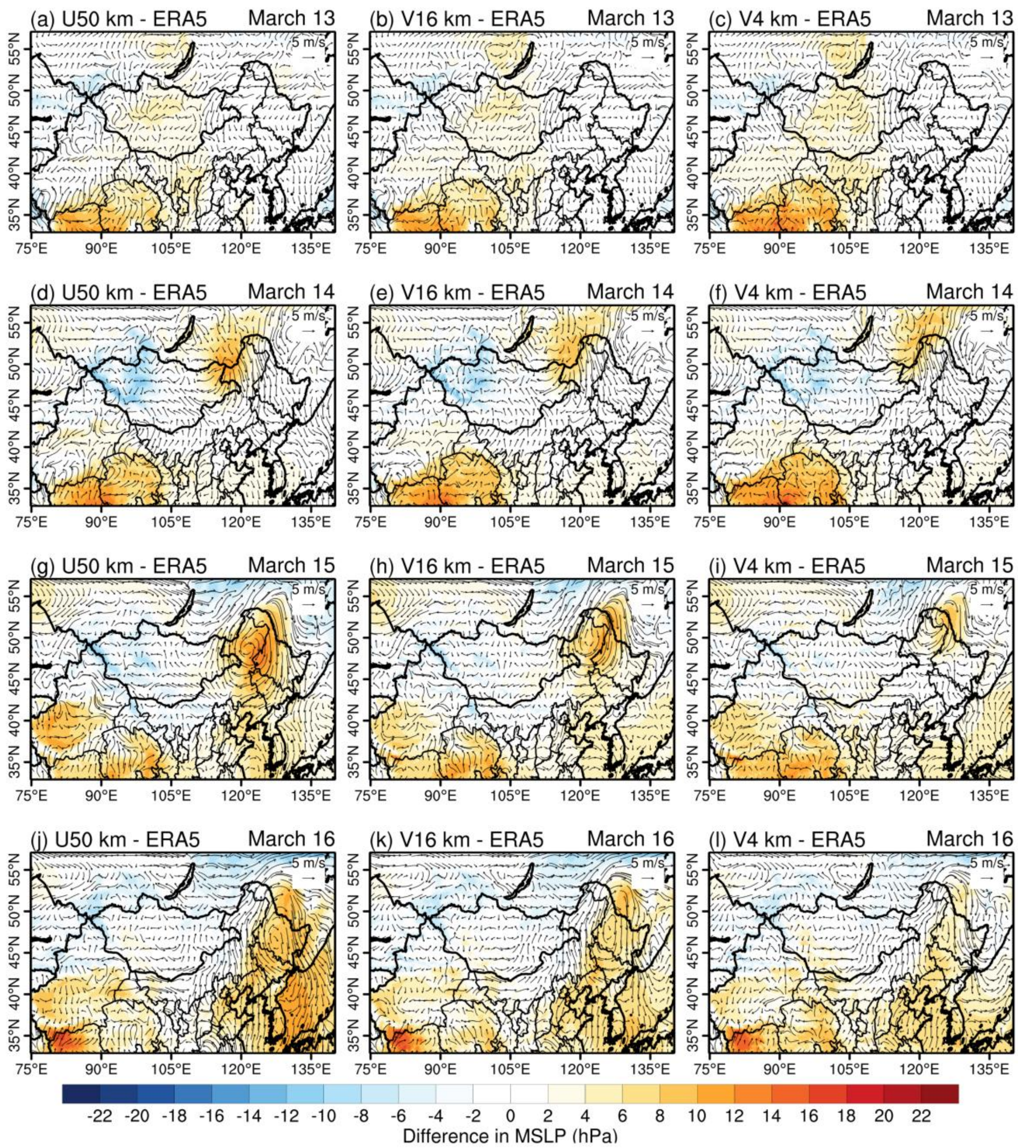

Fig. S2. Daily differences in MSLP (shading; model minus ERA5, hPa) and 850 hPa wind vectors (vectors; model minus ERA5, m s^{-1}) for the U50 km, V16 km, and V4 km simulations during 13-16 March 2021. All model fields were interpolated onto the native 0.25° ERA5 grid before calculating the differences.

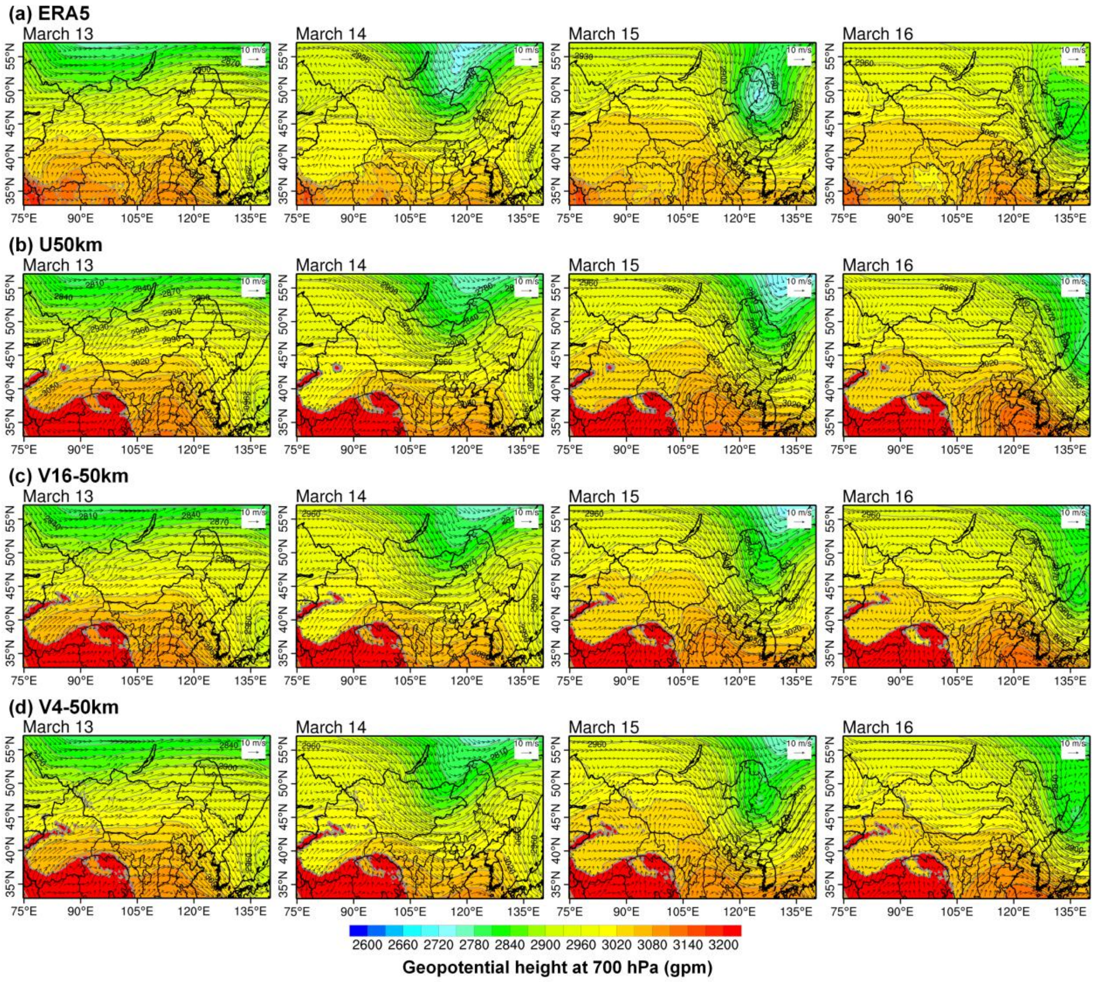

Fig. S4. Daily mean geopotential height at 700 hPa (shading, gpm) and 700 hPa wind vectors for ERA5 and iAMAS simulations (U50 km, V16 km, and V4 km) during 13-16 March 2021. All model fields were interpolated onto the native 0.25° ERA5 grid for comparison.

Question 7: Line 381-382: Figure 3 does not show the information about "terrain transitions", and the audiences might not be familiar with the definitions of Hexi Corridor/Mongolian Plateau. The authors could overlay the winds on topography and highlight the regions with significant differences in winds if possible.

Reply: We have revised the manuscript by removing the description of terrain-related wind variations that was not directly supported by the original Fig. 3 and by adding Fig. S3, in which the daily mean 850 hPa winds from ERA5 and the U50 km, V16 km, and V4 km simulations are overlaid on a common terrain-height field during 13-16 March 2021 **(see Section 3.1.1 on Page 17, Lines 457-461; Fig. S3)**. The newly added terrain-overlaid wind fields provide the topographic context for evaluating the spatial differences in the simulated low-level circulation. The revised analysis shows that all simulations reproduce the broad evolution of the 850 hPa circulation, while some localized differences in wind magnitude and direction occur among the simulations, particularly over northern China during 14-16 March. These variations are mainly located near regions with strong elevation gradients, as illustrated in Fig. S3.

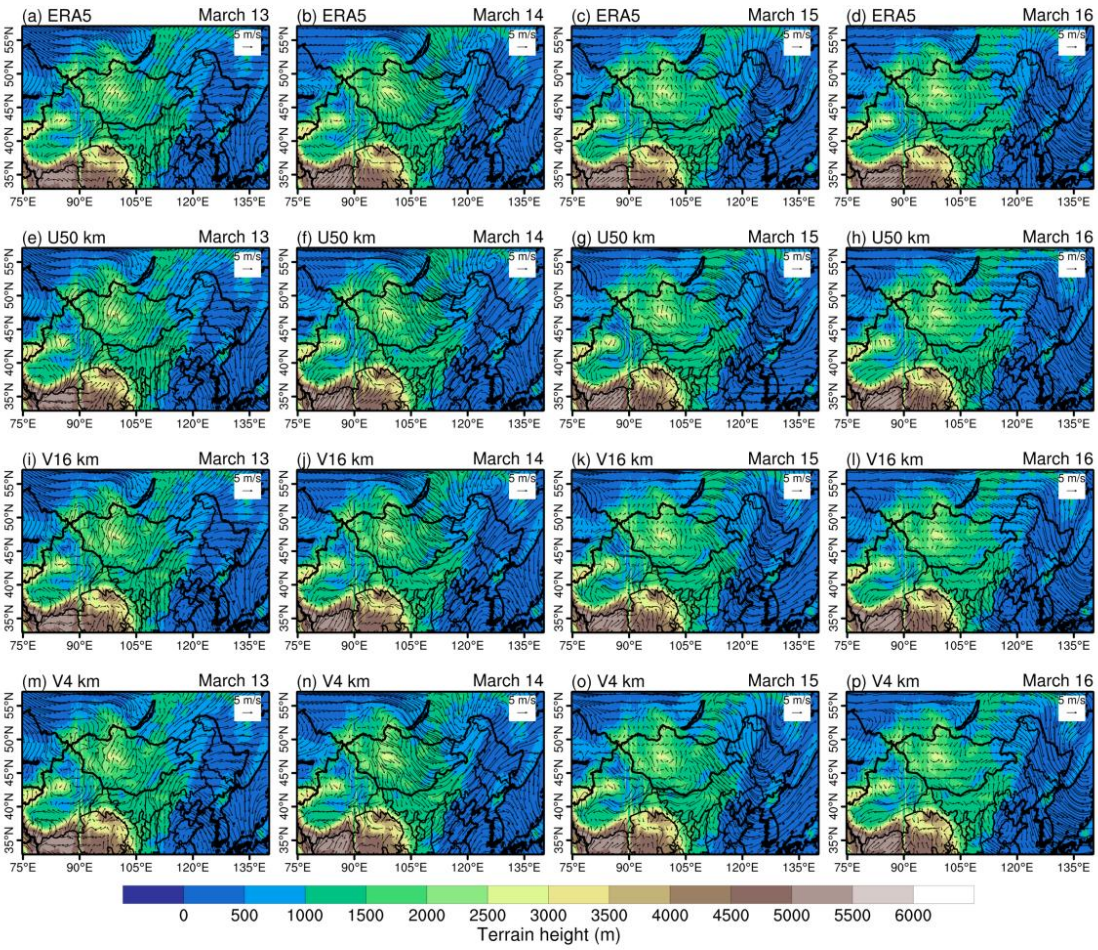

Fig. S3. Terrain height (shading; m) and daily mean 850 hPa winds (vectors; m s^{-1}) from ERA5 and the U50, V16, and V4 simulations during 13-16 March 2021.

Question 8: Line 463: What do the authors mean by "more spatially diffuse"? The dust in U50 goes farther from source to downwind regions? On March 16 (Figure 5), over the downwind Shandong Peninsula, V16 and V4 show larger concentrations (>500 ug/m^3) than U50 does (200-500 ug/m^3). This suggests that dust in V16 and V4 goes farther. Where are "elevated background concentrations" (Line 465-466) located at? For the "positive mean bias presented in Fig. 6" (Line 466-467), most of the observed concentrations are below 400 ug/m^3 while simulated concentrations are 200-1300 ug/m^3, so, I guess these points are associated with stations in northwestern regions in Figure 6 where all three models show higher bias in concentrations.
Reply: We have revised the manuscript and removed the inaccurate descriptions "more spatially diffuse", "elevated background concentrations ", and "positive mean bias presented in Fig. 6" (**see Section 3.1.3 on Page 20, Lines 544-554**). The revised text now describes the differences among the simulations in terms of the continuity, intensity, and spatial organization of the transported dust plume. In particular, on 16 March, the V16 km and V4 km simulations produce higher PM_{10} concentrations over the downwind Shandong Peninsula than the U50 km simulation, indicating stronger dust transport toward this region in the refined simulations.

Question 9: Line 489-492: How about the effects of dry deposition near dust source regions on dust concentrations/transport?
Reply: We have revised the manuscript to explicitly consider the role of dry deposition in the resolution-dependent differences in dust concentrations and transport. In Section 3.1.3, we now clarify that the simulated differences in dust transport may result from the combined effects of the spatiotemporal evolution of dust emissions, circulation-related wind changes, dry deposition, and wet scavenging **(see Section 3.1.3 on Page 21, Lines 569-571).** In addition, dry deposition over the source and near-source regions is quantified and discussed in Section 3.2.1 and Table S7 **(see Section 3.2.1 on Page 26, Lines 726-730; Table S7**). The results show that dry deposition varies with mesh refinement. Compared with U50 km, the V4 km simulation produces higher total dry deposition over the analyzed source regions, with particularly large increases over the MGD and HC–WIM. This enhanced near-source removal may remove part of the additionally emitted dust before the remaining airborne dust is transported downstream, and may therefore influence the resolution-dependent differences in downstream dust concentrations and transport.

Table S7. Total dust dry deposition over different regions during 13-17 March 2021.

Regions*	U50 km (Tg)	V16 km (Tg)	V4 km (Tg)	V16 km Relative Change (%)**	V4 km Relative Change (%)**
TKD	12.269	10.360	15.663	–15.56%	27.66%
MGD	34.551	39.980	64.754	15.71%	87.42%
HC–WIM	18.956	22.252	35.103	17.39%	85.18%
LP	2.432	2.717	3.723	11.72%	53.08%

* The locations and spatial extents of the analyzed regions are shown in **Fig. 1**. ** Relative change (%) is calculated relative to the U50 km simulation.

Question 10: Section 3.1.4: The authors could compare the simulated spatial distribution of AOD with satellite observations/reanalysis data. Does the model output include dust optical depth? If so, the authors could evaluate it using observation/reanalysis data such as MERRA2. What other aerosol species are considered in the model? The biases in anthropogenic aerosols could affect the model performance in simulating AOD in the downwind regions of dust.

Reply: We have revised the manuscript and added Fig. S6, which compares the daily spatial distributions of MODIS AOD with the U50 km, V16 km, and V4 km simulations, together with Table S2, which summarizes the spatial correlations and biases of the 14-18 March mean AOD fields **(see Section 3.1.4 on Page 23, Lines 626-634; Fig. S6, Table S2)**. For the satellite reference, we used the MODIS Terra+Aqua Multi-Angle Implementation of Atmospheric Correction (MAIAC) aerosol product (MCD19A2, Version 6.1), which provides daily gridded AOD at 0.55 μm at approximately 1-km spatial resolution. Only quality-controlled valid AOD retrievals were retained **(see Section 2.5 on Page 16, Lines 423-426)**. For the model–satellite comparison, MODIS AOD was spatially mapped to the corresponding model grid cells, and simulated AOD was sampled using the same grid-cell-by-day availability as MODIS during 14-18 March 2021. MODIS and simulated AOD were then averaged over the matched valid dates at each grid cell, and the spatial correlation coefficient, RMSE, and mean bias were calculated from the resulting paired multi-day mean fields. As shown in Table S2, the spatial correlations range from 0.62 to 0.66, indicating moderate spatial correspondence between the simulations and MODIS retrievals, while positive biases remain in all three simulations. Regarding the aerosol composition, only mineral dust aerosols are simulated in the present configuration. Thus, the reported model AOD represents the optical contribution of simulated dust, although the output variable is not separately labelled as dust optical depth. We agree that anthropogenic aerosols may contribute to the observed AOD and may therefore affect the model–observation comparison, particularly over downwind regions. The effects of anthropogenic aerosols on the simulated AOD will be further investigated in future work.

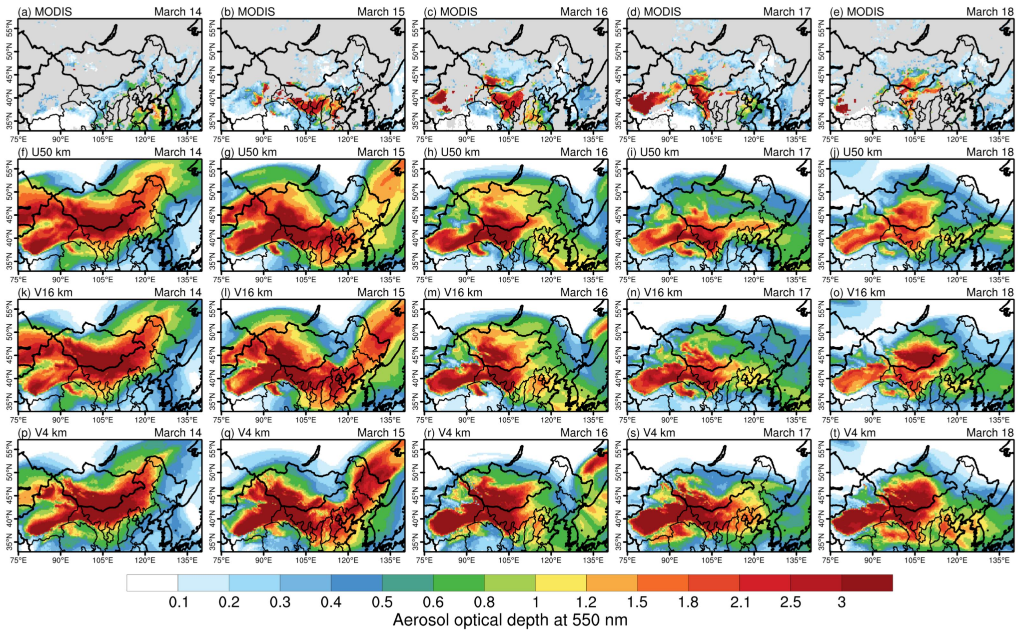

Fig. S6. Daily spatial distributions of MODIS-retrieved and simulated AOD at 550 nm during 14–18 March 2021.

Table S2. Spatial evaluation of the simulated 14-18 March 2021 mean AOD fields against MODIS satellite retrievals.

Experiment	r	RMSE	Mean bias
U50 km	0.63	0.85	0.49
V16 km	0.62	0.96	0.56
V4 km	0.66	1.09	0.70

Notes: MODIS and simulated AOD were averaged at each grid cell over 14-18 March 2021 using only dates with valid MODIS retrievals at that grid cell. The spatial correlation coefficient (*r*), RMSE, and mean bias were then calculated from the resulting multi-day mean fields.

Question 11: Line 583-585: What does "the strongest" mean? It refers to magnitude or spatial range? What do the authors mean by "consistent enhancement"?

Reply: We have revised the manuscript by removing the ambiguous phrases "the strongest" and "consistent enhancement" (**see Section 3.2.1 on Page 25, Lines 708-709**). The revised text now states that the difference maps show spatially extensive positive emission anomalies over the southern Mongolian Plateau, thereby describing only the spatial pattern supported by the figure.

Question 12: Line 643-645: Both the model resolution and topography can affect the low-level circulation. Therefore, to separate these two factors, the authors may need to use higher-resolution model with coarse-resolution topography. In TKD (40°N, 90°E), the topography differences are small (<10m, Figure 9 d, e), but the differences in wind friction field are relatively large (-0.2~-0.3, Figure 10 d, e).

Reply: We agree that both horizontal model resolution and topographic representation can affect the low-level circulation and friction velocity, and that their individual contributions cannot be strictly separated in the present experiments. We have therefore removed the original discussion that directly linked terrain differences to changes in low-level circulation and friction velocity and revised the corresponding interpretation (**see Sect. 3.2.2 on Page 28, Lines 778-783**). In particular, we note that relatively large friction-velocity differences occur over parts of the TKD where the corresponding terrain-height differences are small, indicating that the friction-velocity response cannot be explained by terrain representation alone. We further clarify that resolution-dependent changes in atmospheric dynamics, boundary-layer processes, and the grid-scale representation of land-surface properties may also contribute to these differences. However, the present experiments do not allow these individual effects to be quantified separately. Dedicated sensitivity experiments, such as using a high-resolution atmospheric mesh with coarse-resolution topography, would be required to separate the effects of atmospheric resolution and topographic representation, and such experiments will be considered in future work.

Question 13: Figure 12: With mesh refinement, dust burden decreases in the region with large topography gradient (30–40°N, 105°E). This may be related to higher topography blocking dust transport. In northeastern China (40–45°N, 110–130°E), V16km shows decreases in dust burden, but V4km shows increase. Which model has model performance regarding this? The authors may compare the simulation results with observations/reanalysis data of dust burden or optical depth.
Reply: We have revised the corresponding discussion and added a regional comparison of the simulated 550 nm AOD with MODIS observations over northeastern China (40–45°N, 110–130°E) during 14–18 March 2021 **(see Sect. 3.2.3 on Page 31, Lines 864-869; Table S9)**. The statistics were calculated from the 14-18 March mean fields using collocated grid cells and dates with valid MODIS retrievals. Table S9 shows positive AOD biases in all three simulations. Among the three configurations, V16 km exhibits the highest spatial correlation (r = 0.65) and the lowest RMSE and MB (0.64 and +0.48, respectively). In comparison, U50 km has r = 0.51, RMSE = 0.77, and MB = +0.59, while V4 km has r = 0.58, RMSE = 0.72, and MB = +0.55. Thus, among the two refined simulations, V16 km shows closer agreement with the observed regional AOD than V4 km over northeastern China for this comparison period and observational sample.

Table S9. Regional evaluation of simulated 550 nm AOD against MODIS over northeastern China during 14-18 March 2021.

Experiment	r	RMSE	MB
U50 km	0.51	0.77	+0.59
V16 km	0.65	0.64	+0.48
V4 km	0.58	0.72	+0.55

Notes: Northeastern China is defined as 40–45°N and 110–130°E. Statistics were calculated from the 14-18 March mean fields using collocated grid cells and dates with valid MODIS retrievals.

Question 14: Figure 13: The authors may change the range of colorbar to show values below 10 mg/m2/day, because the current figure shows precipitation decrease (Figure 13c, blue color) but wet scavenging does not (Figure 13f, white color) as the changes are less than 10 mg/m2/day. Changing colorbar could be more consistent with Table 5 which shows wet scavenging in NCP decreases (from 6.733 to 1.109 mg/m2/day) and support the conclusion "precipitation differences … modulate dust removal" (Line 755-757).

Reply: We have revised the color range of Fig. 13 to resolve wet-scavenging changes smaller than 10 mg m^{-2} day^{-1} (**see Fig. 13**). In the revised Fig. 13f, the decrease in dust wet scavenging over the NCP is now more clearly visible and is consistent with Table 6, which shows a reduction from 6.733 to 1.109 mg m^{-2} day^{-1}. This revision provides clearer support for our conclusion that resolution-dependent precipitation differences modulate regional dust removal.

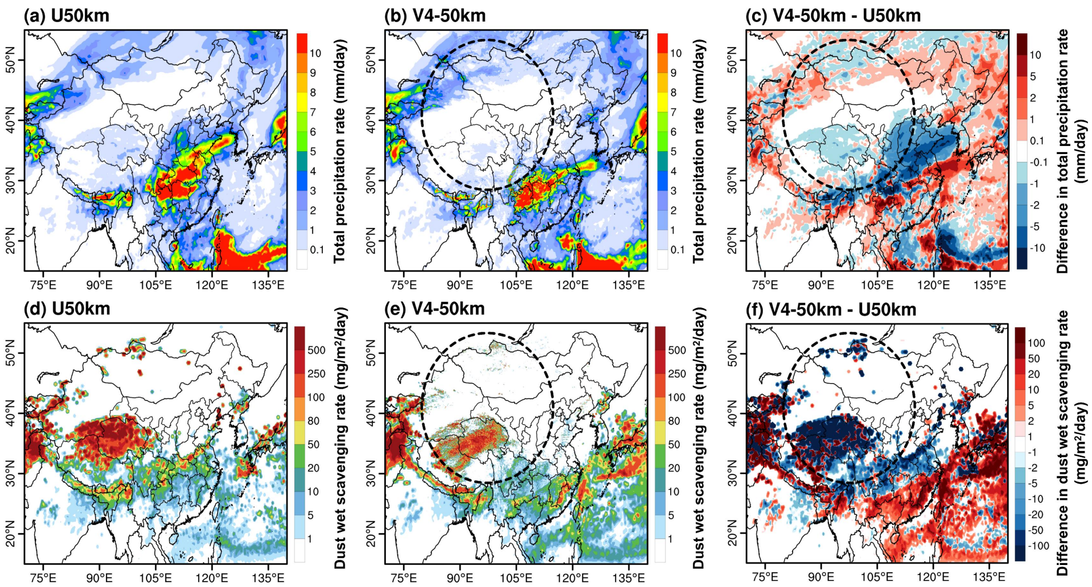

Fig. 13. (a–b) Spatial distributions of period-mean total precipitation rate (mm day^{-1}) from the U50 km and V4 km simulations during 13-17 March 2021. (c) Difference in period-mean total precipitation rate between V4 km and U50 km. (d–e) Spatial distributions of period-mean dust wet scavenging rate (mg m^{-2} day^{-1}) from the U50 km and V4 km simulations during 13-17 March 2021. (f) Difference in period-mean dust wet scavenging rate between V4 km and U50 km. The dashed outline denotes the refinement region.

Question 15: Line 764-767: What do the authors mean by "more consistent"? How to quantify it? Do you mean the contribution from emission changes to dust burden changes is larger than that from wet removal changes? Also, in these source regions, what are the changes in dust dry deposition as the model resolution increases?

Reply: We have deleted the ambiguous statement "more consistent" and revised the corresponding discussion by adding an analysis of dry deposition over the source regions **(see Sect. 3.2.4 on Page 33, Lines 926-931; Table S7)**. The present diagnostics do not allow us to quantitatively separate the relative contributions of emission changes and wet removal changes to the dust-burden response. We have therefore revised the text to avoid implying that one process contributes more strongly than the other. The dry-deposition results further show that the response is region dependent. Over the TKD, total dry deposition changes non-monotonically with mesh refinement, whereas it increases with mesh refinement over the MGD, HC–WIM, and LP (Table S7). Together with the concurrent changes in wet scavenging, these results indicate that the dust-burden response reflects the combined influence of resolution-dependent changes in emissions, transport, and removal processes.

Table S7. Total dust dry deposition over different regions during 13-17 March 2021.

Regions*	U50 km (Tg)	V16 km (Tg)	V4 km (Tg)	V16 km Relative Change (%)**	V4 km Relative Change (%)**
TKD	12.269	10.360	15.663	–15.56%	27.66%
MGD	34.551	39.980	64.754	15.71%	87.42%
HC–WIM	18.956	22.252	35.103	17.39%	85.18%
LP	2.432	2.717	3.723	11.72%	53.08%

* The locations and spatial extents of the analyzed regions are shown in **Fig. 1**. ** Relative change (%) is calculated relative to the U50 km simulation.

Question 16: Line 776: Besides "precipitation structure", how about the effects of model resolution on clouds? How does the in-cloud and below-cloud wet scavenging change as model resolution increases?

Reply: We agree that resolution-dependent changes in cloud properties, as well as in-cloud and below-cloud scavenging, may also affect dust wet removal. We have therefore revised the corresponding discussion to avoid attributing the differences in wet scavenging solely to changes in precipitation structure **(see Sect. 3.2.4 on Pages 33-34, Lines 939-947)**. The present model diagnostics allow us to evaluate the total wet-scavenging response, but do not provide sufficient information to separately quantify changes in cloud properties or the individual contributions of in-cloud and below-cloud scavenging. Therefore, we interpret the stronger wet scavenging over the YRD in V4 km more cautiously as an increase in diagnosed total wet scavenging that occurs alongside changes in precipitation. We have also explicitly acknowledged this limitation in the revised manuscript and added the following statement: "Future work should examine resolution-dependent changes in cloud properties and separately quantify the contributions of in-cloud and below-cloud scavenging to dust wet removal."

Anonymous Referee #2

General comments

This manuscript investigates the impacts of mesh refinement on the simulation of an extreme East Asian dust storm using a global variable-resolution model. The study is interesting and provides useful insights into how increasing resolution affects near-surface winds, dust emission, and long-range transport. I have several comments that may help improve the manuscript.

Major comments

Question 1: Contributions of dust-emitting area and emission intensity. The authors have already discussed the changes in saltation occurrence frequency and emission intensity, which helps explain the enhanced dust emission at finer resolution. However, it would be useful to further quantify the contributions from active dust-emitting area and emission intensity per unit area, and directly compare these quantities among U50 km, V16 km, and V4 km. This would provide a clearer explanation for the substantial increase in total dust emission with mesh refinement.

Reply: We have added a quantitative comparison of the active dust-emitting area fraction and the mean emission intensity over active grid cells for the major source regions **(see Sect. 3.2.1 on Page 25, Lines 693-702; Table S5)**. Relative to U50 km, the active-area fraction decreases by 7.73-28.04% in V16 km and 11.74-46.49% in V4 km, whereas the corresponding emission intensity over active grid cells increases by 29.50-127.37% and 57.81-241.14%, respectively. The increase in total dust emission with mesh refinement therefore occurs despite a reduction in the active emitting area and is accompanied by substantially stronger emission intensity within the active source areas.

Table S5. Time-mean active dust-emitting area fraction and time-mean emission intensity over active grid cells in the major dust source regions during 13-17 March 2021.

Region	Experiment	Active-area fraction (%)	Emission intensity over active area (μg m^{-2} s^{-1})	Active-area change relative to U50 (%)	Intensity change relative to U50 (%)
TKD	U50 km	41.34	348.43	–	–
	V16 km	35.01	507.87	−15.31	+45.76
	V4 km	31.02	549.84	−24.96	+57.81
MGD	U50 km	53.32	759.70	–	–
	V16 km	49.20	1014.32	−7.73	+33.52
	V4 km	47.06	1255.29	−11.74	+65.23
HC–WIM	U50 km	59.13	255.53	–	–
	V16 km	52.97	330.91	−10.42	+29.50
	V4 km	47.91	423.21	−18.98	+65.62
LP	U50 km	6.99	77.57	–	–
	V16 km	5.03	176.37	−28.04	+127.37
	V4 km	3.74	264.62	−46.49	+241.14

Notes: Active emitting grid cells are defined as grid cells with positive dust emission flux at each hourly time step. The active-area fraction is first calculated at each time step as the area of active emitting grid cells divided by the total area of the corresponding region and is then averaged over 13-17 March 2021. The emission intensity is first calculated at each time step as the area-weighted mean dust emission flux over active emitting grid cells and is then averaged over the same period. Relative changes are calculated with respect to the U50 km simulation.

Question 2: PDF of friction velocity. The manuscript suggests that the increase in dust emission is mainly associated with stronger extreme friction velocity rather than more frequent saltation. I suggest providing the probability density distributions (PDFs) of friction velocity for the three resolutions, with the threshold friction velocity indicated, to clearly show how mesh refinement modifies the high-u* tail.

Reply: We have added area-weighted probability density functions (PDFs) of the normalized friction velocity u^*/u_t^* for the three mesh configurations over the major dust source regions, with $u^*/u_t^*=1$ explicitly indicating the saltation threshold (**see Sect. 3.2.2 on Page 29, Lines 805-813; Fig. S8**). This analysis allows the below-threshold range and the upper tail of the normalized friction-velocity distribution to be directly compared among the three resolutions. The finer-mesh simulations generally show enhanced probability density over the below-threshold range, consistent with the reduced saltation occurrence frequency shown in Fig. 11. Meanwhile, modest enhancements in the far-right tail are evident in several source regions, particularly the MGD and HC–WIM. These results show that the enhanced dust emission with mesh refinement cannot be explained simply by more frequent saltation-favorable conditions and that large normalized friction-velocity exceedances can still occur even as the overall saltation frequency decreases.

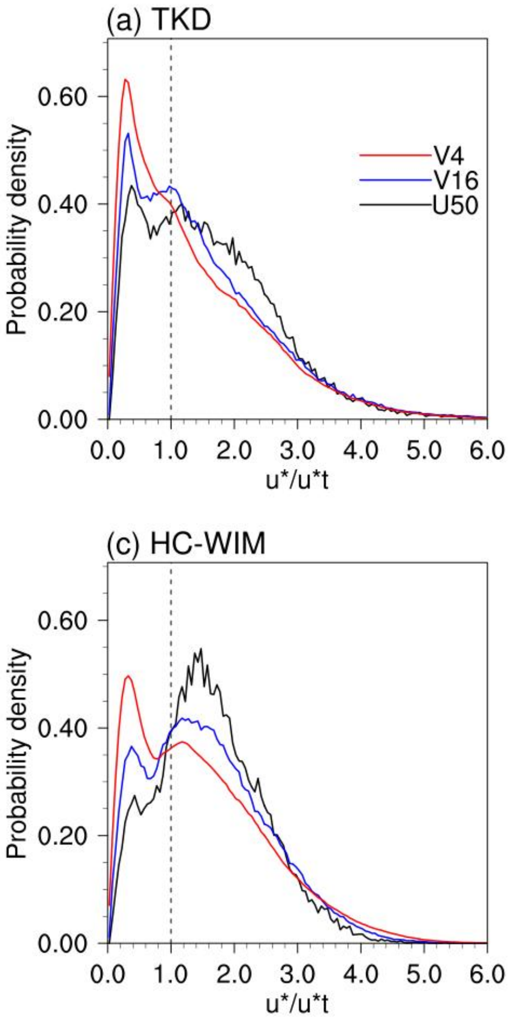

Fig. S8. Area-weighted probability density functions of the normalized friction velocity u^*/u_t^* over the (a) TKD, (b) MGD, (c) HC–WIM, and (d) LP during 13-17 March 2021 for the U50 km, V16 km, and V4 km simulations. The vertical dashed line at u^*/u_t^*=1 denotes the saltation threshold.

Question 3: Does the resolution effect converge? Dust emission continuously increases from U50 km to V16 km and V4 km. This raises an important question: would dust emission continue to increase if the model resolution were further refined? The authors do not necessarily need to conduct additional simulations, but the possible resolution limit or convergence of dust emission should be discussed.

Reply: We have added a discussion of the resolution dependence of dust emission to assess whether the increase with mesh refinement shows any indication of convergence (**see Sect. 3.2.1 on Page 25, Lines 686-692**). Specifically, we compared the incremental changes between successive mesh configurations. Over East Asia, total dust emission increases by 35.24% from U50 km to V16 km, whereas the increase from V16 km to V4 km is reduced to 14.59%. This smaller increment suggests a reduced sensitivity of dust emission to further mesh refinement at finer resolution. However, the present three mesh configurations are insufficient to determine whether this behavior represents the onset of resolution convergence.

Determining whether dust emission approaches a resolution-converged regime would require additional simulations at resolutions finer than V4 km.

Question 4: Reduced dust loading around the Qaidam Basin. Figure 12 shows decreased dust mass loading around the Qaidam Basin in the refined simulations. The reason for this regional decrease is unclear. Please clarify whether it is related to changes in local emission, terrain-modulated transport, or deposition/removal processes.

Reply: We have added a quantitative comparison of local dust emission, dry removal, and wet scavenging over the Qaidam Basin region to clarify the processes associated with the localized decrease in dust mass loading **(see Sect. 3.2.3 on Pages 30-31, Lines 856-864; Table S8)**. Specifically, from V16 km to V4 km, the area-weighted mean dust emission flux decreases by approximately 14%, indicating reduced local dust emission that may partly contribute to the lower dust loading. In contrast, both dry removal and wet scavenging also decrease in V4 km, indicating that the reduced dust loading cannot be attributed to enhanced local removal. These results suggest that the localized decrease in dust mass loading is associated partly with reduced local emission, while resolution-dependent changes in dust transport over the complex terrain surrounding the Qaidam Basin may also contribute.

Table S8. Area-weighted mean dust emission and removal diagnostics over the Qaidam Basin region during 13–17 March 2021.

Process	U50 km	V16 km	V4 km	V16 vs. U50 (%)	V4 vs. V16 (%)	V4 vs. U50 (%)
Dust emission flux (μg m^{-2} s^{-1})	53.604	65.508	56.572	+22.21	−13.64	+5.54
Dry-removal flux (μg m^{-2} s^{-1})	31.690	35.151	21.114	+10.92	−39.93	−33.37
Wet-scavenging rate (mg m^{-2} day^{-1})	492.376	411.431	154.461	−16.44	−62.46	−68.63

Notes: The Qaidam Basin analysis region is defined as 90–100°E and 35–39.5°N. All quantities are area-weighted mean values over the analysis region during 13-17 March 2021. Dry removal includes dry deposition and gravitational settling, whereas wet scavenging includes both large-scale and convective scavenging components. Relative changes are calculated between the corresponding mesh configurations.

Minor comments

Question 1: Figure 4. Figure 4 contains too many wind-field panels, making detailed comparison difficult. I suggest selecting one or two representative days and enlarging the panels so that the differences among observations and the three simulations can be more clearly examined.

Reply: We have revised Fig. 4 by retaining only the results for 14-15 March 2021 and enlarging the corresponding panels to facilitate a clearer comparison between the

NOAA surface-station observations and the three simulations **(see Fig. 4)**. These two days were selected as representative periods during the development of the dust event. The revised figure provides a more focused view of the spatial differences in daily mean 10-m wind speed among the U50 km, V16 km, and V4 km simulations and the collocated observations.

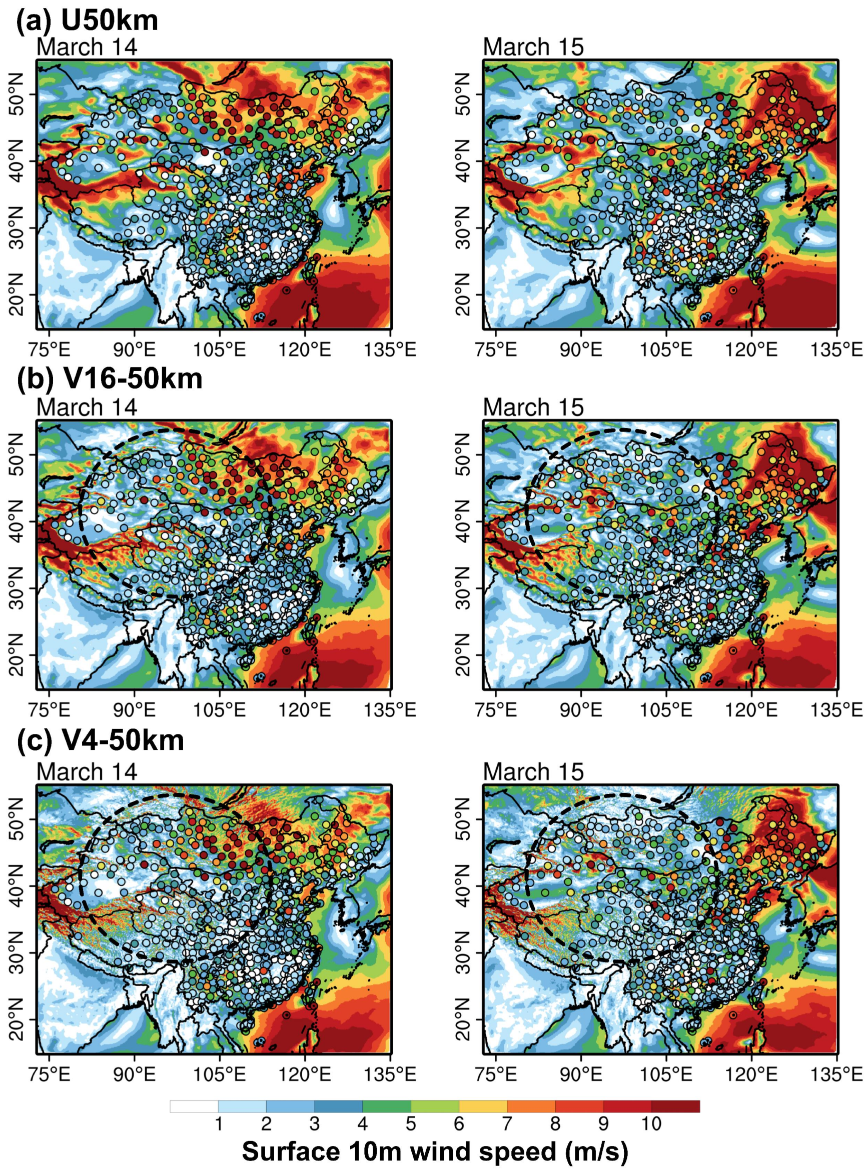

Fig. 4. Spatial distribution of daily mean 10-m wind speed (shading, m s^{-1}) from the iAMAS simulations (U50 km, V16 km, and V4 km), with collocated NOAA surface-station observations (453 stations) overlaid as colored markers, during 14-15 March 2021. These two days are selected as representative periods during the development of the dust event. The dashed outline denotes the refinement region.

Question 2: Figure 5. The current design makes it difficult to distinguish station observations from the background PM_{10} shading, particularly over regions with high concentrations. Please redesign the figure using clearer station symbols, outlines, or another approach to improve the distinction between observations and simulations.

Reply: We have revised Fig. 5 to improve the visual distinction between the surface PM_{10} observations and the simulated PM_{10} fields **(see Fig. 5)**. Specifically, we reduced the size of the observation-station markers to minimize overlap with the background PM_{10} shading and to make the spatial distribution of both the observations and simulations more clearly distinguishable, particularly over regions with high PM_{10} concentrations.

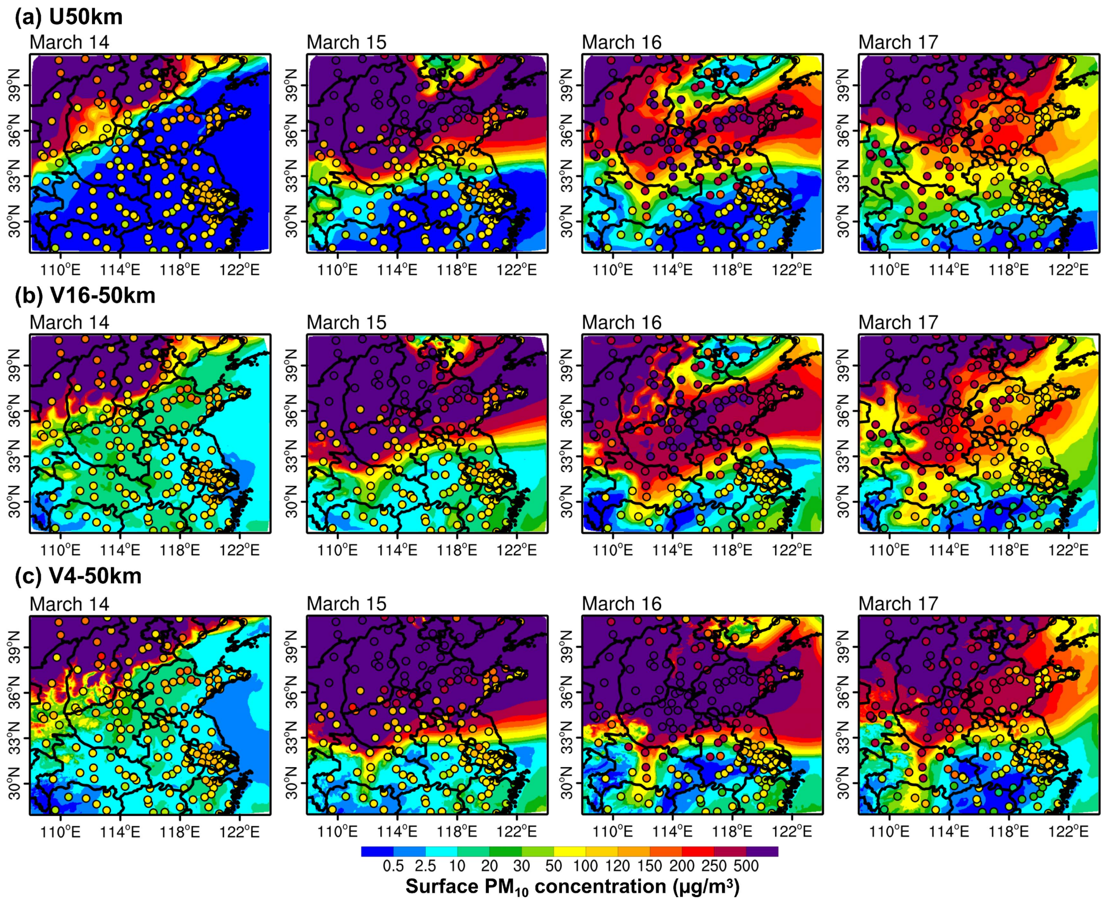

Fig. 5. Spatial distribution of daily mean surface PM_{10} concentrations over eastern China from the iAMAS simulations during 14-17 March 2021. Shading denotes simulated PM_{10} concentrations from the U50 km, V16 km, and V4 km configurations, and colored markers denote CNEMC observations at 192 collocated monitoring stations.

Anonymous Referee #3

General Comments

This manuscript investigates the impacts of mesh refinement on the simulation of an extreme East Asian dust storm using the variable-resolution iAMAS model. While the topic is relevant and the study presents some interesting results, I have some concerns about the current manuscript. Important aspects of the model implementation and experimental design are not sufficiently documented, and the mechanistic interpretation is not adequately supported by the analyses presented. In several places, diagnostic or statistical relationships are interpreted as physical mechanisms without clear process-level attribution. Moreover, the conclusions are largely drawn from a single event, raising concerns about their robustness and general applicability.

In its current form, I do not think the manuscript provides sufficient evidence to support several of its main conclusions. Substantial revisions are required before the manuscript can be considered for publication. In particular, the authors need to clearly address the methodological uncertainties, strengthen the mechanistic attribution, and better establish the robustness and novelty of the results. The major issues listed below should be fully addressed in a revised manuscript.

Major Comments

Question 1: The resolution dependence of dust emissions and its novelty need to be better addressed (Sects. 2.2 and 3.2.2): The nonlinear and threshold dependence of dust emission on surface wind speed, and the resulting sensitivity of dust emissions to model resolution, have been well documented in previous studies (e.g., Ridley et al., 2013; Meng et al., 2021). In particular, Ridley et al. (2013) demonstrated strong resolution dependence of dust emissions in GEOS-Chem due to the nonlinear response of dust emission to resolved surface winds, while Meng et al. (2021) developed grid-independent high-resolution dust emissions to reduce this artificial resolution dependence. These studies are directly relevant to the central issue of this manuscript but are not discussed. The authors should compare the magnitude and behavior of the resolution sensitivity found here with previous studies and clarify what new physical insight is provided by the variable-resolution framework, rather than presenting the resolution dependence itself as a new mechanism.

Reply: We agree that the resolution dependence of dust emissions associated with the nonlinear and threshold-dependent response of dust emission schemes to surface winds has been well documented, particularly by Ridley et al. (2013) and Meng et al. (2021). We have therefore added discussions of these studies and clarified the distinction between the established resolution sensitivity of dust emissions and the additional insights provided by the present variable-resolution simulations **(see Introduction on Page 4, Lines 83-92; Sect. 2.2 on Page 8, Lines 200-206; Sect. 3.2.2 on Pages 27-28, Lines 820-840)**. Specifically, we have incorporated the findings of Ridley et al. (2013), who demonstrated the resolution dependence of dust emissions in GEOS-Chem associated with the representation of high-wind events at different resolutions, and Meng et al. (2021), who developed a grid-independent high-resolution dust emission approach to reduce artificial resolution sensitivity. We now explicitly clarify that the purpose of this study is not to establish resolution dependence itself, but to examine how regional mesh refinement within a global variable-resolution framework modifies the resolved conditions controlling dust emission and the subsequent dust lifecycle. In Sect. 3.2.2, we further compare our results with these previous studies and clarify the physical insight obtained for the March 2021 event. The results show that mesh refinement does not simply increase the frequency of saltation occurrence. Instead, over the major dust source regions, the area-averaged saltation frequency decreases while extreme friction-velocity exceedances become stronger, indicating a shift toward less frequent but more intense erosion-favorable conditions. We also compare this response with Feng et al. (2023), who reported reduced annual-mean dust emissions over East Asia with increasing resolution using a GOCART formulation driven by 10-m wind speed. This comparison indicates that the sign and magnitude of the resolution response can depend on how mesh refinement modifies the wind variables controlling the respective nonlinear emission formulations. We have also added the following references:

Ridley, D. A., Heald, C. L., Pierce, J. R., and Evans, M. J.: Toward resolution-independent dust emissions in global models: Impacts on the seasonal and spatial distribution of dust, Geophys. Res. Lett., 40, 2873-2877, https://doi.org/10.1002/grl.50409, 2013.

Meng, J., Martin, R. V., Ginoux, P., Hammer, M., Sulprizio, M. P., Ridley, D. A., and van Donkelaar, A.: Grid-independent high-resolution dust emissions (v1.0) for chemical transport models: application to GEOS-Chem (12.5.0), Geosci. Model Dev., 14, 4249-4260, https://doi.org/10.5194/gmd-14-4249-2021, 2021.

Question 2: The opposite resolution response compared with Feng et al. (2023) deserves careful explanation (Introduction and Sect. 3.2.2): Feng et al. (2023), using a closely related global variable-resolution modeling framework, reported decreased dust emissions over East Asia with mesh refinement, attributed to reduce near-surface winds associated with better-resolved complex terrain. In contrast, the present study finds a substantial increase in dust emissions with increasing resolution and attributes this to stronger local wind/friction-velocity extremes. Why does mesh refinement produce opposite responses in these two studies? Is the difference related to synoptic conditions, season, source regions, resolution range, or other aspects of the model configuration? This apparent contradiction should be explicitly discussed and may provide important insight into when and why mesh refinement increases or decreases dust emissions.

Reply: We have added discussions of the contrasting resolution responses between Feng et al. (2023) and the present study **(see Introduction on Page 4, Lines 92-98; Sect. 3.2.2 on Page 30, Lines 830-840)**. In the Introduction, we now clarify that the response of dust emissions to mesh refinement is not necessarily uniform and cite Feng et al. (2023) as an example in which refinement reduced dust emissions over

East Asia. In Sect. 3.2.2, we further discuss the physical and methodological differences between the two studies. Feng et al. (2023) used a GOCART-based emission formulation driven by 10-m wind speed and found that mesh refinement weakened regional near-surface winds and reduced the frequency of emission events, in association with better-resolved topographic complexity, leading to lower dust emissions. In contrast, the S04 scheme used in this study is controlled by threshold friction velocity and saltation flux. Our results show that mesh refinement reduces the area-averaged saltation frequency but strengthens extreme friction-velocity exceedances. The opposite emission responses therefore likely reflect differences in how mesh refinement modifies the wind variables controlling the respective nonlinear emission formulations, rather than a universally positive or negative effect of increasing resolution. We also note that Feng et al. (2023) mainly examined an annual East Asian response using U120 km and V16 km meshes, whereas the present study focuses on an extreme March 2021 dust event using U50 km, V16 km, and V4 km meshes. Differences in synoptic conditions, season, source regions, and resolution range may therefore also contribute to the contrast.

Question 3: S04 Scheme, land surface properties, and dust source function details: Since this study focuses primarily on dust simulation, a more detailed explanation of the S04 dust emission scheme and its key differences from S11 should be provided in the text. Furthermore, while the physical emission processes are described, the manuscript should elaborate on how land surface properties (e.g., soil moisture, vegetation cover, snow cover) modulate dust emissions. Lastly, please clarify the definition of the "dust source function" used in the model. Does it specifically represent soil erodibility?

Reply: We have added discussions of the S04 dust emission scheme, the role of land-surface properties in regulating dust emission, and the definition and physical meaning of the dust source function to provide a more complete description of the dust emission parameterization used in this study **(see Sect. 2.2 on Pages 8-9, Lines 195-223)**. Specifically, we clarify that S04 retains a more explicit representation of saltation bombardment and aggregate disintegration than the more simplified S11 formulation. We have also added a description of how land-surface conditions regulate dust emission by modifying the threshold conditions for saltation initiation. Soil moisture increases particle cohesion and suppresses dust release, whereas vegetation and snow cover reduce the availability of exposed erodible surfaces and thereby limit saltation activity. Furthermore, we clarify that the dust source function is a dimensionless weighting factor representing the relative availability of erodible sediments and the associated dust-emission potential under otherwise identical forcing conditions. It therefore represents a relative indicator of dust source potential rather than a direct quantitative measure of soil erodibility. We have also added its value range, horizontal resolution, and clarified that the same source-function dataset is applied in all three mesh configurations.

Question 4: The aerosol spin-up period appears too short (Sect. 2.4, around Lines

304–315): The simulation starts on 8 March, providing only about four days of aerosol spin-up before the meteorological reinitialization on 12 March. Four days may be insufficient, particularly for the smaller dust size bins with relatively long atmospheric lifetimes. The authors should either use a longer spin-up period (e.g., 10 days to two weeks), provide a sensitivity test, or provide sufficient evidence demonstrating that a four-day spin-up is adequate for the aerosol fields considered in this study.

Reply: We have added a discussion of the purpose and implementation of the aerosol spin-up period, as well as the continuity of chemical fields across the meteorological reinitialization **(see Sect. 2.4 on Pages 13-14, Lines 355-374; Fig. S1)**. We clarify that the four-day spin-up period is designed to allow the dust and chemical fields to dynamically evolve from the prescribed initial conditions before the target event, rather than to achieve a long-term aerosol climatological equilibrium. During this period, dust emission, transport, and removal processes are continuously simulated, allowing the aerosol and chemical fields to adjust to the evolving meteorological conditions. The same initial chemical conditions and spin-up procedure are applied to all three mesh configurations, providing a consistent basis for the resolution-sensitivity comparison. Importantly, the meteorological reinitialization on 12 March does not restart the aerosol fields. The prognostic chemical tracers, including dust, are retained from the end of the spin-up integration and continue to evolve through the analyzed event. As shown in Fig. S1, the regional dust burden evolves smoothly during the 12 h before and after the meteorological reinitialization, with no apparent discontinuity or abnormal change in any of the three mesh configurations. This indicates that the adopted spin-up and restart procedure provides a continuous dynamically evolved dust state for the subsequent event simulation, without introducing an apparent artificial disturbance at the transition.

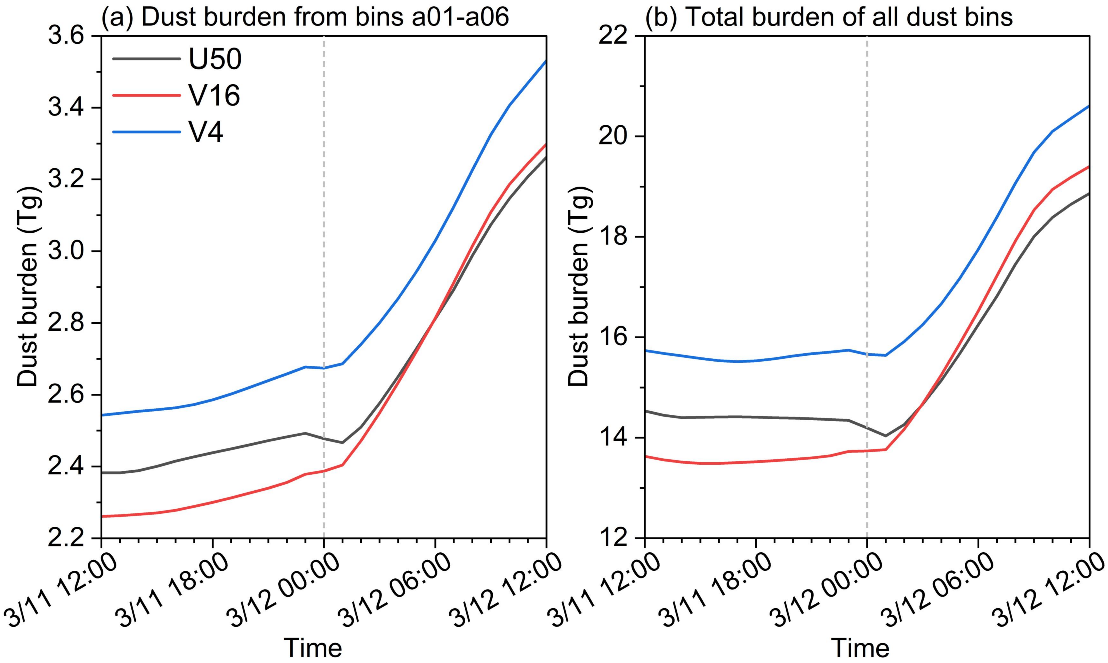

Fig. S1. Temporal evolution of regional dust burden during the 12 h before and after the meteorological reinitialization at 00:00 on 12 March 2021 for the U50 km, V16 km, and V4 km simulations. (a) Dust burden from bins a01–a06, corresponding to dust size bins 1–6 in Table 1; (b) Total dust burden from all bins a01–a08, corresponding to dust size bins 1–8 in Table 1. The vertical dashed line denotes the meteorological reinitialization time.

Question 5: The meteorological reinitialization procedure needs further clarification and justification (Sect. 2.4, around Lines 308–315): The meteorological fields are reinitialized on 12 March while the chemical tracers are retained. The authors explain that this approach is intended to reduce meteorological drift while maintaining continuity of the chemical fields. However, retaining the aerosol fields while replacing the meteorological state may introduce inconsistency between the two. The authors should clarify how the reinitialization is technically implemented and, more importantly, demonstrate or provide evidence that this procedure does not introduce significant discontinuities or affect the subsequent dust emission, transport, and deposition during the analyzed event.

Reply: We have clarified the technical implementation of the meteorological reinitialization procedure and added an evaluation of its potential impact on dust continuity **(see Sect. 2.4 on Page 14, Lines 372-374; Fig. S1)**. Specifically, the meteorological fields were replaced by the ERA5-based fields at 00:00 on 12 March 2021, while the prognostic chemical tracers, including dust aerosol fields, were directly carried forward from the preceding spin-up simulation without reinitialization. As shown in Fig. S1, the regional dust burden exhibits a smooth temporal evolution during the 12 h before and after the reinitialization for all three mesh configurations, with no apparent discontinuity or abnormal change after the meteorological fields were updated. This result indicates that the reinitialization procedure does not introduce an apparent artificial discontinuity into the prognostic dust fields. Following the reinitialization, dust emission, transport, and removal processes continue to be calculated online in response to the updated meteorological conditions.

Question 6: Aerosol–meteorology interactions need to be clarified and separated from the resolution effect (Sects. 2.4 and 3.2.4): It is unclear whether the direct and indirect effects of aerosols, particularly dust, are activated in these experiments. This needs to be explicitly stated. If aerosol–radiation and/or aerosol–cloud interactions are included, the interpretation becomes considerably more complicated because mesh refinement changes dust emissions and distributions, which can in turn modify radiation, clouds, precipitation, and other meteorological fields. The meteorological differences among U50, V16, and V4 may therefore reflect both (1) the direct effect of mesh refinement on meteorology and (2) aerosol–meteorology feedbacks induced by resolution-dependent dust emissions and distributions. These contributions need to be distinguished. If aerosol feedbacks are active, sensitivity experiments with these feedbacks turned off, or other quantitative analyses, would be needed to determine their relative contributions.
Reply: We have added explicit descriptions of the aerosol–meteorology coupling used in the experiments and revised the related discussion to distinguish the effects of mesh configuration from aerosol feedbacks **(see Sect. 2.4 on Page 15, Lines 389-390; Sect. 3.2.4 on Pages 33-34, Lines 939-947).** Specifically, aerosol–radiation feedback is turned off in all U50 km, V16 km, and V4 km experiments, and dust is not coupled to cloud microphysics as cloud condensation nuclei or ice nuclei. Therefore, the resolution-dependent changes in dust emissions and distributions analyzed in this study do not feed back to the simulated meteorological fields through dust radiative effects or dust-induced ice nucleation. Dust transport, dry deposition, and wet scavenging remain active and respond to the simulated meteorological conditions. We have also revised Sect. 3.2.4 to avoid implying a two-way dust–cloud–precipitation feedback. The discussion now focuses specifically on how differences in precipitation and atmospheric transport among the mesh configurations affect dust wet removal and the downstream dust distribution. In particular, the previous wording concerning "dust–cloud–precipitation interactions" has been replaced by "precipitation-related dust wet-removal processes". We further note that future work should examine resolution-dependent changes in precipitation characteristics and separately quantify the contributions of in-cloud and below-cloud scavenging to dust wet removal.

Question 7: The mechanistic attribution remains incomplete (Sect. 3.2): The diagnostic analyses of terrain, surface winds, friction velocity, dust emissions, transport, and wet scavenging are useful, but they do not fully isolate the contributions of individual processes. Since these processes change simultaneously with mesh refinement, their concurrent changes alone are not sufficient to establish causality or quantify their relative contributions. Additional process-level analyses or sensitivity experiments would strengthen the mechanistic attribution. For example, using the same dust emissions across different-resolution simulations could help separate emission-related effects from those of transport and removal. A quantitative dust budget would also help distinguish the contributions of emission, transport, dry deposition, and wet deposition. If such analyses are not feasible, the mechanistic conclusions should be appropriately qualified.

Reply: We have revised the mechanistic interpretation throughout Sect. 3.2 to more clearly distinguish the process-level diagnostics from a quantitative attribution of individual contributions **(see Sect. 3.2 on Pages 23-34, Lines 647-947)**. In particular, we now explicitly state at the beginning of Sect. 3.2 that dust emission, transport, and removal processes vary concurrently among the mesh configurations and that the following diagnostics are intended to provide a process-based interpretation of their coupled responses, rather than to quantitatively isolate the contribution of each individual process. Accordingly, we have revised several statements in Sects. 3.2.1–3.2.4 that previously implied stronger causal or relative attribution. The friction-velocity analysis is now interpreted in terms of the threshold-based and nonlinear response of the S04 emission scheme, while the discussion of regional dust loading emphasizes the combined responses of emission, transport, and removal rather than attributing the loading differences to a single process. Similarly, for the YRD, enhanced wet scavenging is described as a process that may contribute to the reduced downstream dust burden rather than as a quantitatively isolated cause. These revisions retain the process-level diagnostic evidence while limiting the mechanistic interpretation to what can be supported by the present experimental design. A quantitative separation of the relative contributions of emission, transport, dry deposition, and wet deposition would require dedicated process-isolation experiments and dust-budget analyses, which are identified as directions for future work.

Question 8: The conclusions rely on a single dust event and their robustness needs to be demonstrated (Sects. 3.2 and 4): The proposed mechanisms are derived entirely from the March 2021 event. However, the response of surface winds and dust emissions to mesh refinement may depend strongly on synoptic conditions, source regions, season, and background meteorology. This concern is particularly relevant given the opposite resolution response reported by Feng et al. (2023). I strongly recommend adding several dust events under different meteorological conditions to test whether the proposed mechanisms are robust. If additional cases are not feasible, the conclusions should be substantially narrowed and explicitly presented as findings specific to this event.

Reply: We have added explicit statements to clarify the event-specific scope of the diagnosed resolution responses and have correspondingly narrowed the mechanistic interpretation and conclusions to avoid generalizing the findings beyond the March 2021 dust event **(see Sect. 3.2 on Pages 23-34, Lines 647-947; Sect. 4 on Pages 34-36, Lines 950-1015)**. Specifically, at the beginning of Sect. 3.2, we now state that the diagnosed responses are specific to the March 2021 event and should not be interpreted as a universal response of dust emission and transport to mesh refinement under all meteorological conditions. The discussion of the friction-velocity and dust-emission response has also been explicitly framed in the context of this event. In particular, the decrease in area-averaged saltation frequency together with enhanced extreme friction-velocity exceedances is now presented as an event-specific response rather than as a general consequence of mesh refinement. This interpretation is also consistent with our comparison with Feng et al. (2023), which further illustrates that the direction of the dust-emission response to mesh refinement may differ across model formulations and meteorological settings. We have likewise revised Sect. 4 to limit the broader conclusions to the present event and explicitly state that the process responses identified here should be regarded as specific to the March 2021 event, because the relative roles of terrain-induced wind variability, circulation-related transport pathways, and precipitation-related removal may depend on the prevailing synoptic conditions and regional hydrometeorological environment. Accordingly, assessment across additional dust events and different synoptic conditions is needed before the robustness and broader applicability of these resolution responses can be established.

Question 9: The observational constraints remain limited for a single-case study (Sects. 2.5 and 3.1): The authors acknowledge the limited AERONET coverage (23 sites). However, the evaluation relies largely on event-mean AOD, which provides limited information on the timing and evolution of the dust plume. PM_{10} is also not a direct measurement of dust. Given that the conclusions are based on a single event, stronger observational constraints would be helpful. The authors should consider evaluating the temporal evolution of the event and, if available, include satellite observations or measurements of the vertical dust distribution to better constrain the simulated dust transport. As highlighted in the Introduction, a key strength of the iAMAS global variable-resolution model is its seamless global-to-regional multi-scale capability. However, the current manuscript focuses mainly on source and near-source regions (e.g., Gobi and Taklimakan). To fully evaluate the model's capability in capturing long-range transboundary transport, additional validation over downwind remote regions should be included. Incorporating available surface PM_{10} networks and AERONET AOD data in South Korea and Japan would significantly strengthen the manuscript.
Reply: We agree that the observational evaluation based solely on event-mean AOD provides limited information on the temporal evolution and long-range transport of the dust plume. However, the available AERONET observations have limitations in temporal continuity and spatial coverage during this extreme dust event, which limits a detailed assessment of the temporal evolution using AERONET alone. To strengthen the observational constraints, we have expanded the AERONET evaluation by including additional downwind observations from South Korea in Fig. 7, which extends the validation from source and near-source regions to remote transport regions and provides further constraints on the simulated long-range dust transport. In addition, we have incorporated daily MODIS AOD retrievals to evaluate the spatial evolution and transport pathway of the dust plume during 14-18 March 2021. The spatial correlation coefficient (r), RMSE, and mean bias between the simulated and MODIS AOD fields are further calculated to quantitatively assess the correspondence between the simulated and satellite-retrieved AOD distributions **(see Section 3.1.4 on**

Pages 22-23, Lines 619-634; Fig. 7, Fig. S6 and Table S2). The results show that the simulations reproduce some of the broad spatial features and downstream evolution of the dust plume, with moderate spatial correlations against MODIS retrievals (r = 0.62-0.66). Meanwhile, the additional South Korea AERONET sites provide further observational constraints on the simulated column aerosol loading over downwind regions during the dust event.

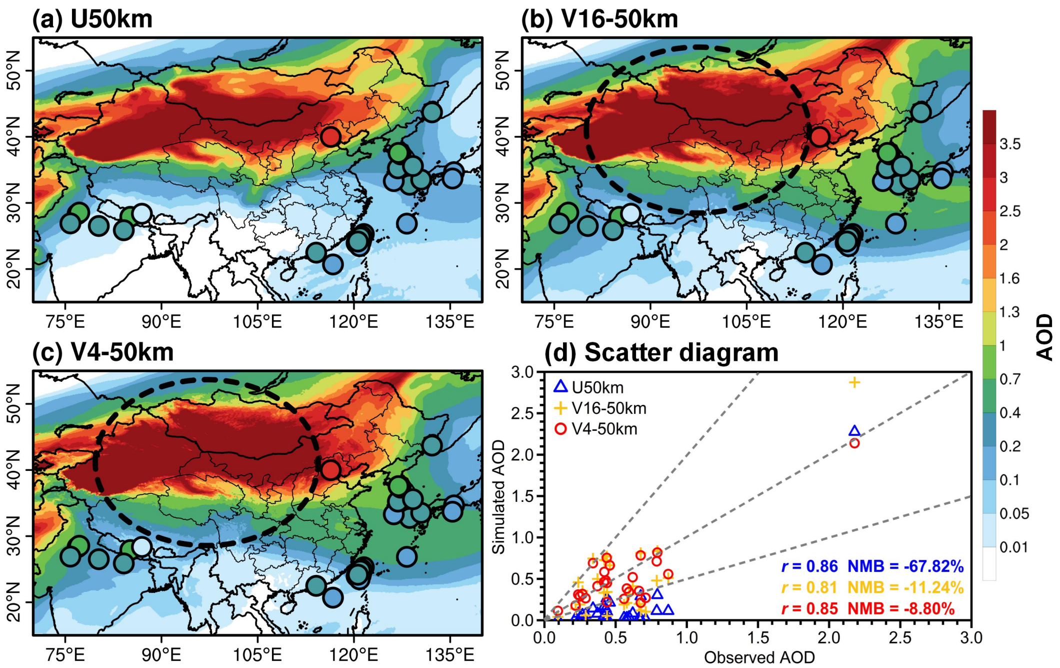

Fig. 7. (a–c) Spatial distribution of mean AOD at 500 nm averaged over 13-18 March 2021 from the iAMAS simulations using the U50 km, V16 km, and V4 km mesh configurations, with collocated AERONET observations overlaid as markers. (d) Scatterplot of simulated versus observed period-mean AOD at 500 nm at the collocated AERONET sites, with the correlation coefficient (*r*) and normalized mean bias (NMB) reported for each configuration. The dashed outline denotes the refinement region.

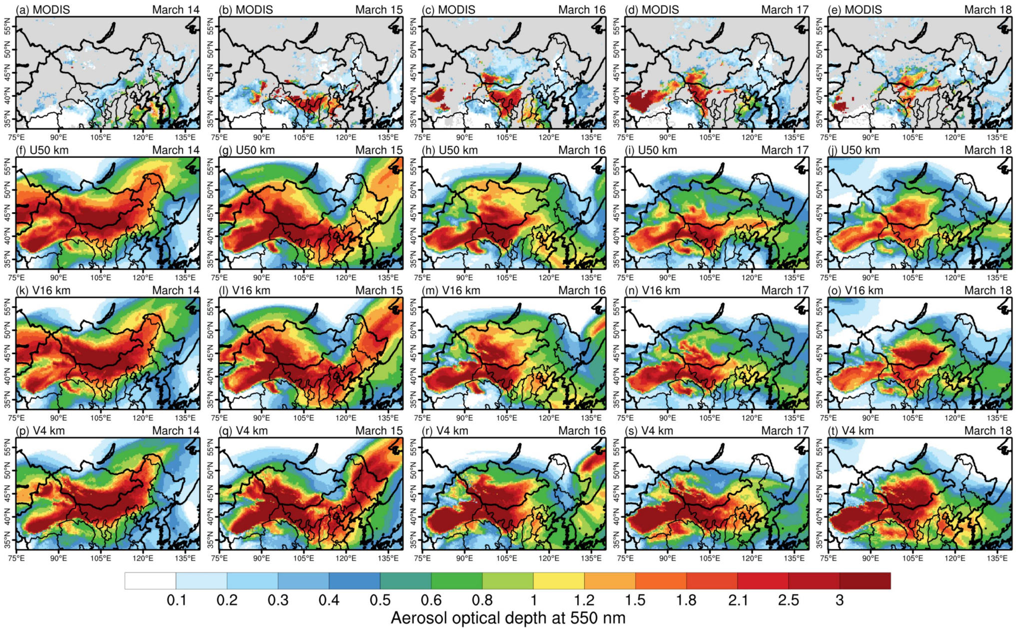

Fig. S6. Daily spatial distributions of MODIS-retrieved and simulated AOD at 550 nm during 14-18 March 2021.

Table S2. Spatial evaluation of the simulated 14-18 March 2021 mean AOD fields against MODIS satellite retrievals.

Experiment	r	RMSE	Mean bias
U50 km	0.63	0.85	0.49
V16 km	0.62	0.96	0.56
V4 km	0.66	1.09	0.70

Notes: MODIS and simulated AOD were averaged at each grid cell over 14-18 March 2021 using only dates with valid MODIS retrievals at that grid cell. The spatial correlation coefficient (*r*), RMSE, and mean bias were then calculated from the resulting multi-day mean fields.

Question 10: Reconciling PM_{10} and AOD validations across mesh resolutions: There is an apparent discrepancy in model performance between surface PM_{10} and column AOD. Why does U50 km show the lowest AOD values while significantly overestimating surface PM_{10}? Conversely, V4 km shows good agreement for both metrics. Did the different horizontal resolutions affect the vertical mixing or size distribution of dust particles? Please provide a more detailed explanation connecting the spatial plume structure, vertical distribution, and AOD/PM_{10} mechanics. Additionally, if spatial plume distribution strongly drives AOD accuracy, AERONET sites nearer to the source regions should be included in the analysis if available.

Reply: We have added a discussion to clarify the different responses of surface PM_{10} and column AOD to mesh refinement and introduced additional diagnostics of the vertical dust distribution and size-related characteristics **(see Sect. 3.1.4 on Page 23, Lines 636-645; Table S3)**. We now clarify that surface PM_{10} primarily reflects local near-surface particle mass, whereas AOD represents vertically integrated aerosol extinction and is therefore more directly related to the column aerosol loading and optical properties. The additional diagnostics show that mesh refinement modifies the vertical dust distribution, with a larger fraction of dust below 1 km and a smaller fraction above 2 km in the refined simulations, whereas the column fine-dust mass fraction changes only modestly among the three mesh configurations. These results indicate that the relatively high surface PM_{10} but low AOD in U50 km may not be primarily caused by stronger near-surface dust confinement or by a substantial resolution-dependent shift in particle-size composition. We have therefore revised Sect. 3.1.4 to interpret PM_{10} and AOD as complementary constraints on different aspects of the simulated dust field. Regarding AERONET observations near the major dust source regions, valid measurements remain limited during this event, as now explicitly noted in Sect. 3.1.4.

Table S3. Vertical and size-related characteristics of simulated dust over the MGD, NCP, and YRD on 13–18 March 2021.

Region	Experiment	Dust mass below 1 km AGL (%)	Dust mass above 2 km AGL (%)	Column fine-dust fraction (%)
MGD	U50 km	42.36	38.61	16.48
	V16 km	45.59	34.00	16.26
	V4 km	62.04	17.96	16.51
NCP	U50 km	9.97	75.22	16.78
	V16 km	22.42	61.32	17.17
	V4 km	29.57	54.31	18.33
YRD	U50 km	6.60	86.69	17.07
	V16 km	14.77	76.03	17.51
	V4 km	15.28	69.63	17.88

Notes: AGL denotes above ground level. Fine dust refers to particles with diameters ≤ 2.5 μm.

Question 11: Improved agreement with PM_{10} and AOD does not necessarily imply more realistic dust emissions (Sects. 3.1.3–3.1.4 and 3.2.1): The manuscript should avoid implying that improved PM_{10}/AOD agreement validates the larger dust emission flux. Better agreement of V4 with observed PM_{10} and AOD does not directly demonstrate that the larger dust emission flux is more realistic, because PM_{10} and AOD are also controlled by transport, vertical mixing, and dry and wet deposition. The authors should therefore avoid using improved PM_{10}/AOD performance as direct validation of the increased dust emission flux. The discussion should clearly distinguish among the resolution sensitivity of dust emissions, improvement in simulated atmospheric dust concentrations/loading, and accuracy of the emission flux itself.

Reply: We agree with the reviewer that improved agreement with observed PM_{10} and AOD does not directly demonstrate that the larger simulated dust emission flux is more realistic. We have therefore reviewed the relevant discussion and revised the manuscript to avoid using the improved PM_{10}/AOD performance of the refined simulations as validation of the increased dust emission flux **(see Sect. 3.1.3-3.1.4**

and 3.2.1 on Pages 19-27, Lines 534-751). In the revised text, we explicitly distinguish between the resolution sensitivity of simulated dust emissions and the agreement of simulated atmospheric dust concentrations and column aerosol loading with observations, without interpreting the latter as direct validation of the emission flux itself. Specifically, the PM_{10} and AOD evaluations in Sects. 3.1.3-3.1.4 are now interpreted only as observational constraints on the simulated atmospheric dust distribution. The larger dust emissions in the refined simulations are discussed separately in Sect. 3.2.1 as a resolution-dependent model response. We have also clarified that the differences in simulated PM_{10} and AOD among the mesh configurations likely reflect the combined responses of dust emission, transport, and removal processes to mesh refinement, rather than changes in dust emission magnitude alone.

Question 12: The attribution of the YRD dust reduction to enhanced precipitation and wet scavenging needs stronger evidence (Sect. 3.2.4): The manuscript attributes the reduced dust loading over the YRD in V4 primarily to enhanced precipitation and wet scavenging. However, dust emissions, dynamical transport, PBL mixing, convection, gravitational settling, and dry deposition all change simultaneously among the experiments. These processes are also closely linked to the meteorological fields, which themselves are affected by the change in model resolution. Therefore, the concurrent increase in wet scavenging and decrease in dust loading does not, by itself, establish a causal relationship. A quantitative regional dust budget separating emission, transport convergence/divergence, dry deposition, and wet deposition would provide much stronger support for this conclusion. The authors should quantify the relative contributions of these processes to the dust-loading differences over the YRD rather than inferring the dominant mechanism mainly from their spatial and temporal correspondence.
Reply: We agree with the reviewer that the concurrent enhancement of wet scavenging and reduction in dust loading over the YRD does not, by itself, demonstrate that wet scavenging is the dominant cause of the reduced dust loading, because transport and other removal processes also vary among the mesh configurations. We have therefore revised Sect. 3.2.4 to avoid attributing the YRD dust-loading reduction primarily to enhanced precipitation and wet scavenging and to qualify the corresponding mechanistic interpretation **(see Sect. 3.2.4 on Pages 32-34, Lines 907-947).** Specifically, the revised text now describes the enhanced wet scavenging over the YRD in V4 km as occurring concurrently with the reduced regional dust loading. The reduced dust loading is interpreted as a combined response of emission, transport, and removal processes, with enhanced wet scavenging in V4 km identified as one possible contributing pathway rather than as the dominant mechanism. We have also revised the conclusion of this section to state that mesh refinement can affect precipitation-related dust wet-removal processes, rather than implying that enhanced wet scavenging alone explains the YRD dust-loading reduction. Because the present experiments do not quantitatively isolate the relative contributions of emission, transport, dry deposition, and wet deposition, we now

identify dedicated process-isolation experiments and regional dust-budget analyses as necessary for such quantitative attribution in future work.

Minor Comments

Question 1: Section 2.4: Please provide the total number of grid cells and, if available, the computational cost for U50, V16, and V4.
Reply: We have added the total number of grid cells for the three mesh configurations in Sect. 2.4 **(see Sect. 2.4 on Page 13, Lines 343-345)**. Specifically, the U50 km, V16 km, and V4 km configurations contain 237430, 268888, and 858645 grid cells globally, respectively. At present, an accurate quantitative estimate of the computational cost is not available for the three configurations.

Question 2: Section 2.5 and 3.1: Please clarify how the surface wind, PM_{10}, and AERONET observations are collocated with the model grids, particularly given the different grid spacing among the three experiments.
Reply: We have added a description of the collocation method used for the station-based evaluations **(see Sect. 2.5 on Page 16, Lines 421-423)**. Specifically, for each surface wind, PM_{10}, and AERONET observation site, the corresponding model value is extracted from the geographically nearest model grid cell in each native mesh configuration. This procedure is applied independently to the U50 km, V16 km, and V4 km experiments to account for their different grid spacing.

Question 3: Section 3.2.2: Please clarify the temporal resolution used to calculate u^*, u_t^*, the threshold-exceedance frequency, and S_3.
Reply: We have added a clarification of the temporal resolution used for these diagnostics **(see Sect. 3.2.2 on Page 28, Lines 791-793)**. Specifically, the calculations use hourly u^* and u_t^* fields during 13-17 March 2021, and the threshold-exceedance frequency and S_3 are calculated from the corresponding hourly samples.

Question 4: Section 2.5 and 3.1.3: PM_{10} is used as a proxy for dust. Please briefly discuss the possible contribution of non-dust aerosols to the observed PM_{10} during the event.
Reply: We have added a brief discussion of the possible contribution of non-dust aerosols to the observed PM_{10} during the event **(see Sect. 2.5 on Page 15, Lines 402-409)**. Specifically, we now clarify that although the elevated PM_{10} concentrations during this strong dust event were strongly influenced by mineral dust and therefore provide a useful observational constraint for dust model evaluation, PM_{10} is not a dust-specific measurement and may also contain contributions from anthropogenic and other non-dust aerosol components. Accordingly, observed PM_{10} is interpreted as an approximate observational constraint on near-surface dust conditions rather than as a direct measurement of dust alone.

Question 5: Section 3.1: Please report the number of observations used in each statistical comparison, particularly for the regional PM_{10} and surface-wind evaluations.

Reply: We have added the number of observation stations used in the evaluations to the caption of Table 3 and the captions of Figs. 4 and 5 in Sect. 3.1 **(see the caption of Table 3 and the captions of Figs. 4 and 5)**. Specifically, Table 3 reports 13, 25, 16, 18, 28, and 19 NOAA stations for the regional wind evaluations over the TKD, MGD, HC–WIM, LP, NCP, and YRD, respectively. Fig. 4 now indicates that 453 NOAA surface stations are used for the 10-m wind comparison, and Fig. 5 specifies that 192 collocated CNEMC monitoring stations are used for the surface PM_{10} evaluation.

Question 6: Lines 60–62: The sentence describing regional model limitations is oversimplified. Please expand it to provide a more precise description, such as highlighting the "limited representation of lateral boundary conditions in regional models".

Reply: We have expanded the description of the limitations of regional models to provide a more precise explanation of the role of lateral boundary conditions **(see Introduction on Page 4, Lines 76-83)**. Specifically, we now clarify that regional models rely on prescribed lateral boundary conditions from larger-scale models, and that the spatial and temporal resolution of these boundary conditions may constrain the representation of meteorological and aerosol conditions entering the regional domain, particularly for long-range dust transport. We also note that adequately representing dust source regions and long-range transport pathways often requires large regional domains, which can substantially increase the computational cost of high-resolution simulations.

Question 7: Lines 76–78: The description of regional model constraints is too simplified. Regional models are limited in dust simulations not only due to their strong dependency on global driver models, but also because running large domains at high resolution to cover long-range transport pathways incurs substantial computational costs.

Reply: We have expanded the discussion of regional-model constraints to provide a more complete description of the limitations relevant to long-range dust simulations **(see Introduction on Page 4, Lines 76-83)**. Specifically, we now clarify that regional models rely on prescribed lateral boundary conditions from larger-scale models, whose spatial and temporal resolution may constrain the representation of meteorological and aerosol conditions entering the regional domain. We also note that adequately representing both dust source regions and long-range transport pathways often requires sufficiently large regional domains, which can substantially increase the computational cost of high-resolution simulations.

Question 8: Lines 121–123: Please provide a clearer rationale for selecting the March 2021 event. How severe was this dust storm compared to other major East Asian dust events or observation campaign periods?

Reply: We have added a clearer description of the severity of the March 2021 dust storm to provide a stronger rationale for selecting this event **(see Introduction on Page 6, Lines 140-143)**. During this event, the dust plume covered more than 3.8×10^6 km^2 over China, and the daily regional-mean dust optical depth reached the highest or close-to-highest levels observed over the past two decades in northern China (Gui et al., 2022). We have also added the following reference:

Gui, K., Yao, W., Che, H., An, L., Zheng, Y., Li, L., Zhao, H., Zhang, L., Zhong, J., Wang, Y., and Zhang, X.: Record-breaking dust loading during two mega dust storm events over northern China in March 2021: aerosol optical and radiative properties and meteorological drivers, Atmospheric Chemistry and Physics, 22, 7905–7932, https://doi.org/10.5194/acp-22-7905-2022, 2022.

Question 9: Lines 424–450: Have you evaluated the model against hourly surface wind speed observations? In high and complex terrain, comparing daily mean wind speed is insufficient because the model does not use daily winds dynamically, and sub-daily peak winds drive threshold exceedance. How do the normalized mean biases (NMBs) in wind speed quantitatively impact dust emission amounts across the different experiments? Providing a quantitative assessment of emission variability relative to hourly wind speed biases would help readers better understand the underlying sensitivity.

Reply: We have added an evaluation of 3-hourly surface wind speeds against collocated NOAA observations and a quantitative comparison between wind-speed biases and dust emission responses among the three mesh configurations **(see Sect. 3.2.1 on Pages 24-25, Lines 676-686)**. Specifically, Table S4 summarizes the 3-hourly wind-speed normalized mean biases (NMBs), the biases of the 95th-percentile wind speeds representing strong-wind conditions, and the corresponding dust emission amounts for each mesh configuration. The results show that mesh refinement generally reduces the magnitude of the 3-hourly wind-speed NMB over the major dust source regions while dust emissions increase. For example, over the MGD region, the 3-hourly wind-speed NMB changes from –16.97% in U50 km to –11.91% in V4 km, accompanied by an increase in dust emission from 140.841 Tg to 232.920 Tg (+65.38%). The P95 wind-speed biases, however, show region-dependent responses and do not decrease uniformly with mesh refinement. Thus, the wind-speed bias changes should not be interpreted as a direct quantitative explanation for the emission differences among the simulations. Rather, these diagnostics provide a sub-daily observational assessment of the resolved wind field, while the dust-emission response also depends on the nonlinear threshold behavior of friction velocity and saltation in the S04 scheme.

Table S4. Comparison of 3-hourly surface wind-speed biases, extreme wind representation, and dust emission responses under different mesh configurations during 13-17 March 2021.

Region	Mesh	3-hourly wind NMB (%)	P95 wind bias (%)	Dust emission (Tg)	Emission change relative to U50 (%)
TKD	U50 km	+61.17	+45.08	72.401	−
	V16 km	+48.86	+31.35	102.021	+40.91
	V4 km	+33.76	+34.97	105.412	+45.59
MGD	U50 km	−16.97	−17.36	140.841	−
	V16 km	−15.24	−13.87	194.115	+37.83
	V4 km	−11.91	−5.96	232.920	+65.38
HC–WIM	U50 km	+10.39	+9.39	33.760	−
	V16 km	+2.95	+11.50	44.497	+31.80
	V4 km	+2.05	+12.49	52.172	+54.54
LP	U50 km	+24.21	+14.81	1.625	−
	V16 km	+21.25	+38.97	1.911	+17.60
	V4 km	−3.14	+17.76	2.642	+62.58

Notes: Wind-speed statistics are calculated using collocated NOAA 3-hourly surface observations. The P95 wind bias represents the relative difference between simulated and observed 95th-percentile 3-hourly wind speeds. Dust emission changes are calculated relative to the U50 km simulation.

Question 10: Please use terms such as "improvement" with caution, as the response to increasing resolution is not always monotonic across U50, V16, and V4.

Reply: We agree with the reviewer that the impacts of increasing resolution are not necessarily monotonic across U50 km, V16 km, and V4 km. We have therefore carefully checked and revised the descriptions in the Results section, especially those related to model evaluation and the impacts of mesh refinement. The wording has been adjusted to avoid implying that higher resolution always leads to improved performance and to emphasize that the response to mesh refinement depends on the evaluated process and region **(see Sect. 3 on Pages 16-34, Lines 427-947)**.