

We would like to thank the referees for their careful reading of our manuscript. We are most grateful for the comments, constructive criticism and very useful suggestions received on how to improve the paper. Please find attached our detailed answers to the questions and a new version of the paper, which we hope satisfactorily to address the points raised during the discussion. The referee's comments are presented in black, and our responses in blue.

Reply to Reviewer 1: RC 1

General Comments:

This study presents the evaluation of WRF-Chem simulations in dust events across a broad region covering North Africa, Middle East, Central Asia and Europe during 2021. It focuses on the comparison of the performance of three dust emission schemes (GOCART, AFWA, and UoC) and two land surface models (Noah-MP and CLM). The study has shown some noticeable differences in dust simulations among the three dust emission schemes and between the two land surface models. It also summarizes some key differences and similarities. This study could provide some important information on the applicability of the dust emission schemes and land surface models in the simulation and forecast of dust events, and thus it could potentially serve as a useful reference for the relevant community. However, due to unclear statements and unreasonable interpretation in some key results, I suggest denial of the manuscript for its further improvement.

Reply: We sincerely thank the reviewer for the careful assessment of our manuscript and for recognizing the relevance of this study to the dust modelling community. We acknowledge the reviewer's concerns regarding the clarity and interpretation of some of our results. We have carefully addressed all these points and revised the manuscript, accordingly, clarifying the interpretation of the results and better defining the scope and contribution of the study. We have carefully outlined the detailed responses to each comment and the corresponding changes below.

Major comments:

- **Dust belt (in the title and main text):** Generally speaking, the dust belt should consist of a larger region. In addition to the region shown in the

study, it can also include western North Africa and adjacent oceans as well as northern China and Mongolia located in East Asia. I think the study region is misinterpreted as dust belt.

We thank the reviewer for the comment. We have replaced the term “dust belt” with the appropriate names for the regions included in this study throughout the manuscript, and changed the title to reflect this:

“Evaluation of WRF-Chem (version 4.7.1) aeolian dust emission and land surface models in North Africa, Middle East, Central Asia and Europe.”

- **MIDAS dust AOD:** It is used as a reference to evaluate the model performance. Although it is a great dataset which largely enhances our understating, I think it should be noted that it is subject to large uncertainty. It is partly based on MERRA-2 which uses the GOCART dust emission scheme. I think this is one reason why WRF-Chem with GOCART scheme resembles MIDAS DOD (dust optical depth). For AFWA, soil erodibility factor is also used as GOCART, which can partly explain the similarity between GOCART and AFWA.

In fact, in MERRA-2, biases in DOD would lead to the biases in the ratio of DOD to total AOD. This would further lead to the biases in the estimation of DOD in MIDAS, which combines MODIS AOD with MERRA-2 ratio of DOD to total AOD (according to Equation in Line 223). If MERRA-2 underestimates the dust concentrations (especially the peak dust concentrations during strong dust events as some previous studies have shown), MIDAS DOD would be underestimated too. Therefore, compared to MIDAS DOD, WRF-Chem with UoC scheme shows significant **overestimation**, but this is not necessarily true. Indeed, compared to AERONET observations, WRF-Chem with UoC scheme shows **underestimation**, which is inconsistent with overestimation when compared to MIDAS DOD. Given the uncertainty in MERRA-2, it is not appropriate to use MIDAS DOD to have a quantitative conclusion. If the authors want to use MIDAS DOD to derive a quantitative conclusion, I think it should be combined with AERONET AOD to generate a new dataset for a consistent conclusion.

We thank the reviewer for this insight regarding the uncertainties associated with MIDAS DOD arising from MERRA-2 DOD. This is also highlighted in

comments by other reviewers. We would like to provide additional context regarding the conceptual framework underlying the development of the MIDAS dataset. The methodology used to quantify the uncertainty of MIDAS DOD is described in detail in Sect. 3.2 (Gkikas et al., 2021), while (Gkikas et al., 2022) further demonstrate how the grid-cell-level MIDAS DOD uncertainty propagates into the regional DOD estimates (Sect. 3). Specifically, the DOD uncertainty is estimated by combining the MODIS AOD retrieval uncertainty, which depends on the applied retrieval algorithm and the underlying surface type, with the uncertainty of the MERRA-2 dust fraction. The latter is quantified using the satellite-based LIVAS dataset as an independent reference dataset, which is not involved in the development of the MIDAS product.

In addition, there is the contribution of non-dust aerosol species to the total AOD, which directly affects the dust fraction estimated by MERRA-2. MERRA-2 assimilates a variety of aerosol-related observations (see Table 2 in (Randles et al., 2017)), while the partitioning of the assimilation increments among the different aerosol species is determined by the model. In (Gkikas et al., 2021), a comprehensive evaluation of the MERRA-2 dust fraction (MDF) against the corresponding DOD-to-AOD ratios derived from the independent LIVAS dataset is presented in Sect. 4.1. In brief, this evaluation demonstrates that MDF satisfactorily reproduces the DOD-to-AOD ratio over the study region considered in the present work.

Moreover, the MIDAS DOD product was independently evaluated against AERONET retrievals in (Gkikas et al., 2021). For this purpose, the AERONET observations were appropriately screened to identify dust-dominated conditions and thereby provide a suitable reference for the evaluation of DOD (Sect. 4.2 of (Gkikas et al., 2021)). The assessment showed very good agreement between the satellite-derived MIDAS DOD and the ground-based AERONET retrievals, both at the global scale and at the station level, with satisfactory performance found for the majority of the investigated AERONET sites.

When comparing our WRF results to AERONET, the applied AOD and Ångström exponent thresholds may not fully eliminate the contribution of non-dust aerosol species, meaning that a residual contribution from other aerosol types may still be present. We also note that for the MIDAS–WRF collocation, only one satellite overpass is available at a given location and day, whereas for the

AERONET–WRF collocation, multiple ground-based observations may be available within the course of the day and are therefore used for the model evaluation. These may lead to differences between the MIDAS and AERONET comparisons, in addition to the inherent uncertainty highlighted by the referee.

We have revised the manuscript to discuss these uncertainties and the potential biases arising from the use of GOCART dust emissions by both MERRA-2 and our WRF-Chem simulations. The discussion in Section 5 has also been extended to include the following: *“The evaluation of our WRF-Chem configurations with MIDAS is subject to several uncertainties as it is based on MODIS-Aqua AOD and MERRA-2 dust-to-total AOD ratio, whose individual biases can propagate into the derived DOD (Gkikas et al., 2021). Although MERRA-2 assimilates observed total AOD, its aerosol speciation is not directly constrained and may depend on underlying aerosol model and associated parameterisations (Buchard et al., 2017). Nevertheless, the MIDAS DOD product has been evaluated against AERONET observations under dust-dominated conditions (Gkikas et al., 2021), supporting its suitability as a reference dataset for DOD evaluation. A further consideration is that the MERRA-2 aerosols fields used to derive MIDAS are generated using the GOCART aerosol module (Randles et al., 2017) and this overlap with the GOCART as well as the AFWA (which is based on GOCART) modules of the WRF-Chem model can contribute to the closer agreement between the modelled and MIDAS data. The scaling factors derived from MIDAS therefore inherit uncertainties associated with the dataset.”*

- **EMEP:** some stations which are located quite far from the source regions. Although the authors have developed some methods to rule out the stations in the build-up region, it might be still possible to be affected by transported particles from anthropogenic sources nearby. In fact, for the selected stations, the coarse particle concentrations at some moments are quite low (below 1-5 $\mu\text{g}/\text{m}^3$). I suspect that some of these particles are not dust particles. I understand it is challenging to detect dust events (especially weak dust events), but it might be more reliable to use a high threshold.

Reply: We thank the reviewer for this suggestion. We have now selected EMEP stations closer to the dust emission regions by limiting to stations below a

latitude of 45N and applied a concentration threshold of $>10\mu\text{g}/\text{m}^3$ to more reliably isolate dust events. Further analysis of this comparison through statistical metrics using all available EMEP stations across the domain, is presented in Fig. 7.

- **Soil moisture and surface wind:** These two variables are evaluated in a broad region, with **no** specific focus on the dust emission regions. I do not think the general analysis shown in Section 4.2 could be applied to dust emission regions.

Reply: We have included a new Figure (Fig. 6) to represent the regional behaviour of the model against ERA5 Land. We also present seasonal metrics for soil moisture and surface wind in the Supplement. We have extended Section 4.2 to include regional statistics to better interpret the behaviour of these surface parameters over specific dust emission regions:

“We extend the analysis of surface emission drivers to assess their regional variability across the major dust-source regions in our domain; Western North Africa, Eastern North Africa, Sahel and Middle East and central Asia as described in (Kok et al., 2021). Figure 6 presents statistical evaluation of the surface parameters against ERA5 Land. With prominent regional differences in, we see relatively small biases over North Africa, while larger errors and variability occur over the Middle East, Central Asia and Sahel. Near the sources, the model generally overestimates soil moisture indicating a wetter surface, which can subsequently effect dust emission and mobilisation resulting in reduced dust concentration and AOD. The Sahel exhibits a systematic negative MBE signifying WRF-Chem tends to simulate drier soil in this region. Surface wind speed is seen to better represented over North Africa, notably for CLM4 while Noah-MP shows subtle overestimation. Strong positive biases are evident over the Sahel, Middle East and Central Asia.”

- **Line 115:** 0.1 to 10 μm in diameter: I think it is in radius.

Corrected. We thank the reviewer for their diligence.

Specific comments:

Line 5: year-long: please also mention the year here.

The year has been added.

Lines 30-31: within the atmospheric model: deleted?

Correction-> Changed to: with atmospheric models

Line 34: models: not clear.

Correction-> Changed to: Land surface models

Line 40: for the GOCART aerosol model: deleted?

Corrected.

Line 42: to model: changed to “to simulate”.

Correction-> Changed to: “*simulate*”

Line 63: in both cases: not clear.

Changed to-> “*the above studies*” to clarify.

Line 67: the dust scheme: which scheme?

Correction-> Changed to: “*the UoC dust scheme*”

Line 70: land surface: changed to “land surface model”?

Corrected.

Line 73: that captures: changed to “that is affected by”?

Corrected.

Line 76: The sentence is too long.

The sentence is split into two to aid with readability.

Line 78: in the broader spatio-temporal domain affected by the dust belt: not clear.

Changed in the text to make more clear: “*We evaluate the model against satellite-derived products over the whole emission domain,*

and determine the values for the empirical emissions scale factors that could be applied to simulation domains affected by the major dust source regions.”

Lines 111-112: based on topographic source identification: Moved to after "Ginoux et al. (2001)"?

Corrected.

Line 123: dust particles: changed to “soil particles”.

Corrected.

Line 133: mesoscale features: What it means is not clear to me.

Changed to make meaning more clear: “[...] the AFWA scheme may have an advantage at scales of the order of 10 km due to reduced sensitivity to limited soil and vegetation data.”

Line 139: m is the mass: m is the mass of saltating particle.

Corrected.

Line 146: Shao (2004): changed to “(Shao, 2004)”.

Changed.

Line 148: deleted “, (Shao, 2004)”.

Removed the comma.

Line 151: where σ_p represents the mass ratio between free dust particles and aggregated dust: I do not think so. Probably not use σ_p here, and directly use $p_m(d_i)/p_f(d_i)$ as above.

We have made the changes as suggested by the reviewer.

Line 153: The grammar is not correct.

Correction-> Changed to: “All dust emission schemes use the erodible fraction field from the WRF-Chem pre-processor as input, which is a map representing the proportion of each grid cell susceptible to wind erosion for different soil classes (sand, silt, and clay), as shown in Fig 1.”

Line 163: Niu et al. (2011): changed to “(Niu et al., 2011)”.

Corrected.

Lines 166-167: The grammar is not correct.

Corrected.

Line 169: which consists of.

Corrected.

Lines 176-177: It is not clear to me. I think only soil moisture at the top surface layer affects dust emission.

The lines were removed to avoid confusion for the reader. All relevant details are available in the cited literature describing the land surface models.

Line 178: Sec. 4.1.2: Sec. 3.1?

Corrected.

Line 179: fiels?

Corrected.

Lines 184-188: It is not clear to me.

We primarily selected level 2 data wherever available as it is quality controlled, however for cases where level 2 data was not available the next quality level was chosen. Rephrased in the text to: *“For maximum data coverage across the modelled period and to minimize unscreened or uncalibrated raw measurements, the Level 2.0 time-series was chosen for all the stations in our domain (Fig. 1)) that are not affected by anthropogenic emissions (see details of this filtering in Sect. {4.1.2}). Due to the sparse availability of Level 2.0 data at certain locations, for maximum data coverage across the modelled period, the dataset was gap-filled using the next highest available quality levels”* to make it clear.

Lines 214-215: AOD 550 nm MODerate resolution Imaging Spectroradiometer

Corrected.

215 (MODIS) satellite products?

Correction-> Changed to: *“products”*

Lines 235-238: Use the unit of cm as the previous paragraph?

Changed to cm as suggested by the reviewer.

Lines 241-242: What do the differences mean here (difference between different layers)? Please clarify.

The manuscript text was simplified for clarity.

Line 256: Table 2: It appears before Table 1.

Table 1 was replaced in the revised version with a listing for all stations (in Appendix and Supplement).

Line 274: pockets?

Correction-> Changed to: "*areas*"

Figure 2: What does SF mean?

We understand the abbreviation has not been introduced before, and we have included it in the figure caption of Fig 2.

Figure 2: middle: It is better to use modeled dust AOD minus a reference data
Changed.

Figure 4: Table 1: Why are only one value of MBE, MAE, or NME shown in Table 1?

The average values across all stations were initially included in Table 1 for brevity. In the revised version, all statistical metric values have been added to Appendix Tables A1 and A2 and Supplementary Table S1.

Figure 5: The range of soil moisture is too large for dust emission region.

Fig. 5 represents the full range of possible soil moisture over the entire domain and is not restricted to dust emission regions, as some higher soil moisture values occur over parts of our domain. We have now included in the manuscript additional analysis of the performance of the model specifically in the dust emission regions (Fig.6), as suggested by the reviewers.

Line 332: East Asia: changed to "Central Asia".

Corrected.

Lines 396-398: It is not clear.

Rephrased to make meaning clearer.

Section 6: Conclusions: It is a bit too long. Please shorten it.

The Conclusions section has been shortened to make more succinct.

Line 469: the coarse dust consistency of surface concentrations: not clear.

Rephrased to make meaning clearer.

Citation: <https://doi.org/10.5194/egusphere-2026-1900-RC1>

References:

Buchard, V., Randles, C. A., Da Silva, A. M., Darmenov, A., Colarco, P. R., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A. J., Ziemba, L. D., & Yu, H. (2017). The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part II: Evaluation and Case Studies. *Journal of Climate*, 30(17), 6851–6872. <https://doi.org/10.1175/JCLI-D-16-0613.1>

Gkikas, A., Proestakis, E., Amiridis, V., Kazadzis, S., Di Tomaso, E., Marinou, E., Hatzianastassiou, N., Kok, J. F., & García-Pando, C. P. (2022). Quantification of the dust optical depth across spatiotemporal scales with the MIDAS global dataset (2003–2017). *Atmospheric Chemistry and Physics*, 22(5), 3553–3578. <https://doi.org/10.5194/acp-22-3553-2022>

Gkikas, A., Proestakis, E., Amiridis, V., Kazadzis, S., Di Tomaso, E., Tsekeri, A., Marinou, E., Hatzianastassiou, N., & Pérez García-Pando, C. (2021). ModIs Dust AeroSol (MIDAS): A global fine-resolution dust optical depth data set. *Atmospheric Measurement Techniques*, 14(1), 309–334. <https://doi.org/10.5194/amt-14-309-2021>

Kok, J. F., Adebisi, A. A., Albani, S., Balkanski, Y., Checa-Garcia, R., Chin, M., Colarco, P. R., Hamilton, D. S., Huang, Y., Ito, A., Klose, M., Leung, D. M., Li, L., Mahowald, N. M., Miller, R. L., Obiso, V., Pérez García-Pando, C., Rocha-Lima, A., Wan, J. S., & Whicker, C. A. (2021). Improved representation of the global dust cycle using observational constraints on dust properties and abundance. *Atmospheric Chemistry and Physics*, 21(10), 8127–8167. <https://doi.org/10.5194/acp-21-8127-2021>

Randles, C. A., Da Silva, A. M., Buchard, V., Colarco, P. R., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., Shinozuka, Y., & Flynn, C. J. (2017). The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part I: System Description and Data Assimilation Evaluation. *Journal of Climate*, 30(17), 6823–6850. <https://doi.org/10.1175/JCLI-D-16-0609.1>