Review on the manuscript
« A new process-based and scale-aware desert dust emission scheme for global climate models – Part II: evaluation in the Community Earth System Model (CESM2)”

Submitted to Atmospheric Chemistry and Physics

(https://doi.org/10.5194/egusphere-2023-823)

By Danny M. Leung, Jasper F. Kok, Longlei Li, Natalie M. Mahowald, David M. Lawrence, Simone Tilmes, Erik Kluzek, Martina Klose, and Carlos Pérez García-Pando

General comment :

The objective of this paper is to evaluate of the improvement of the dust emission parameterizations of the Earth Systems Model (ESM) CESMS2 described in a companion paper, by comparison with available synthesis of measurements and with the previous parameterizations included in this ESM.

In its present state, thee paper is not totally self-sufficient since the modifications of the dust emission scheme are only briefly recalled and in a very factual way. In particular, the main arguments that justify the selection of a given parametrisation or the values of some key parameters are sometimes missing and only few elements concerning their possible limitations are given. Some sections need further clarification or verifications for consistency with the companion paper. The reader can understand that including this evaluation in the Part I paper would have make it much too dense and too long but the writing of this section should be improved.

The work presented in this paper results for several modification of the dust emission scheme that have been included in the CESM2 Earth Systems Model. These modification are mainly related to processes that include soil and surface parameters : the soil size distribution described here by the soil median diameter, the drag partition due to the presence of rocks and vegetation and the soil moisture that both modulate the erosion thresholds. Some of these parameterizations are not new so the main contribution of this work is their tuning, which largely depends on the used input data used for the simulations. The inclusion at the global scale of a parameterization of the intermittency of the saltation, and the differentiation between the fluid and the impact threshold, are two original aspects. The sensitivity of these two factors to the input wind speeds and soil moistures would have deserve more discussion since it has significant impact on the final results. To take into account the limitation due to the coarse spatial resolution of global model, a correction factor is used to scale the dust emissions computed at low spatial resolution to the dust emissions computed at a higher spatial resolution. This correction is also dependent on the initial input data set for both the surface properties (soils, land use, roughness) and the meteorological parameter but the associated uncertainties are not discussed. In addition, this up-scaling approach does not account for smaller scale dynamical features, such as meso-scale convection producing “haboobs” are not properly represented in global scale. This limitation and the consequences on the final scores of the simulations should, at least, be mentioned in the manuscript.

The two companion papers together represent a valuable piece of work regarding the examination of the different steps involved in the dust emission modelling and the possibility (or not) to improve the representation of the physical processes. The final configuration evaluated in this paper, provides a significant increase of the agreement with observation compared to the two dust schemes used for comparison. However, the new scheme contains several tuning parameters that allows the compensation of uncertainties. In fact, the general approach adopted in this work assign almost all the uncertainty of the simulated dust emissions to the surface parameters. This implies that the final configuration of the dust emission scheme

is not independent of the meteorological input fields used for the simulation. Except at the very end of the manuscript, there is almost no mention of the uncertainties on the main drivers of the dust sources and fluxes : the wind velocity and the soil moisture. This can be a relevant approach to tune a model that will then be used in the same configuration than described in the manuscript. But the relevant of the new dust scheme for othet models is thus questionable. It is even questionable for the CESM2 model without nudging to MERRA meteorological reanalysis.
Finally, if the objective of this work is to improve the performed of the CESM2 model in the presented configuration (nudged to MERRA configuration, and thus as close as possible to present day available observations) the approach is relevant, but if the intention is to recommend these modifications as more generic parametrization, it is very questionable. In this context, the objective of "examining how including more Aeolian processes can benefit to dust modeling performance" is not fully reached. I thus encourage the authors to define more precisely their scientific objectives and to adapt their conclusions accordingly.

I would thus recommend the publication of this manuscript after some revisions based on these general comments and on the detailed comments described below.

Detailed comments :

Abstract :
Line 28 : the sentence "incorporating a realistic soil size distribution " should be replace by "incorporating a simplified soil size representation", since the soil size distribution is represented by a median diameter and not a distribution and that a unique diameter is used.

Introduction :

Line 79-90 : In this section, the limitations of GCM et ASM due to their coarse spatial resolution is discussed but without any link with the small-scale meteorological processes and phenomena that are poorly or not represented in global models such as mesoscale convective systems, sea breezes, low level jets ..

Part 2. Description of CESMS and its default dust scheme

Line 64-65 : I would suggest to replace "interparticle forces due to soil microbes" by "interparticle forces involved in soil crusting"

Line 145-149 : Please replace "uses a global optimal diameter of Dp=75µm for the threshold calculation" by " uses a unique diameter Dp=75µm that is optimum for wind erosion since it corresponds to the lowest threshold"

Line 129-135 : The authors argued on the necessity to distinguish the fluid and the impact thresholds whose values differ by 20%. Is the precision of the modelled wind speeds higher than 20% ?

Line 155-156 : Kok, et al. (2014b) tested three values of a (0.5, 1 and 2) but the presented results are for a =1, the values that gives the best agreement with observations.

Part 3. Modification of the CESM2 dust emission scheme

Line 250-254: the description of the "erodibility" and its link with moisture is not clear and U^*_{st0} is not defined

Line 261 and 262 : What means "bridging the gap between" scales and how if the value of 0.5 determined ?

Line 283 : As a matter of fact, using γ=1.65 10^{-4} kg s^{-2} instead of the value of 3. 10^{-4} kg s^{-2} recommended by Shao and Lu (2000) gives almost the same minimum threshold than IW82. So numerically, the change should not be very sensitive.

Line 265-270 : If the threshold of VAI is set to 1 instead of 0.3, it automatically increases the extend of the dust sources. How much does it increase the extent of semi-arid areas ? Since the total emissions and the global AOD are tuned, how does it impact the spatial distribution of the simulated emissions?

Line 290-291 : It is not clear how the value of a=2 has been selected. Has the value of a=1 being maintained for the K14 simulation as in Kok et al. (2014b) ?

Line 298-299 : The conclusions from Martin and Kok (2019) are based on results obtained for soils with median diameter ranging from 0.3 to 0.5 µm in coastal dunes areas and poorly represent the dynamics of fine particles due to sampling limitations. And the conclusion of equal susceptibility concerns particles in the $D_i/D_{50,bed}$ size classes. It is constant with the much lower median diameter considered in this study as a reasonable approximation (127µm) and more generally to desert soils ?

Line 301 : Please, indicate the number of data sets used investigate the link between PSD and soil texture.

Line 303-304 : The author argued that the range of minimum fluid threshold (0.204 to 0.298 m.s^{-1}) that correspond to the range of D_{med} in the soil data set (40 to 250 µm) can be considered as negligible. But this difference is comparable to the difference between the fluid and the impact threshold (Line 129-131) and the authors argue on the necessity to account for it. This is not consistent. An acceptable argument is that the available data sets does not allow to define precisely the soil size distribution and thus the minimum wind threshold velocity in the arid and semi-arid areas.

Line 307 : Finally, what is the value of the minimum U_{*t} corresponding to Dp=127 µm ? How different is it from the threshold for Dp=75 µm used for Z03 ? The minimum threshold used in Kok et al. (2014) is much lower (0.16 m.s^{-1}). How does it impact the difference in the dust simulations with K14 and the new scheme ? Is there any compensation with the coefficient of tuning of the moisture impact (a=1 in K14 and a=2 in this study) ?

Line 339 : According to Marticorena and Bergametti (1995) z_{0s} =D_p/30 and not 2D_p/30.
In Marticorena and Bergametti (1995), X was set to 10 cm, while a value of 10 m is used here. A short justification of this choice would be welcome regarding the different values used in the literature (X=10 cm in Marticorena and Bergametti (1995) for small roughness elements and low roughness density ; Y=12,225 cm in Mac Kinnon et al. (2004) for vegetated surfaces with relatively large and high shrubs and high roughness density; X=40 cm in Pierre et al. (2014) on

millet fields and savannah). And why not using different X for the different land cover categories ?

Line 357-375 : From the description given here, the drag partition due to vegetation and to rocks appear as represented independently : there is no possibility to represent a surface covered by rocks and vegetation. Since a data base allows to estimate the fractional area of rocky surface and the fractional area for vegetated surfaces, it is not clear why and average F_{eff} is needed. The dust fluxes could have been computed for each type of surface and then summed up at the grid cell scale.

Line 380-392 : The inclusion of a parameterization that account for the intermittency of the saltation is relevant in terms of physical process. The parameterization from Comola et al. 2019 allows to account for the occurrence of saltation when the wind friction velocity range between the fluid and the impact threshold whose values differs by about 20%. This raises again the question of the accuracy of the simulated wind speed that is the input parameter for the computation of the wind friction velocity. Indeed in the paper by Comola, the intermittency factor is estimated either from in-situ measurements of the wind friction velocity either using an historical 10m wind speed data bases re-scaled to the in-situ measurements.
It is not explicitly mentioned in section 3.4 but the application of this parametrization also requires the estimation of the Monin-Obukov length from the meteorological model output which also questions the precision of the estimation.

Line 403-405 : Following the previous comment, it appears that decreasing the minimum threshold by using the impact threshold and the intermittency parametrization increase unrealistically high emission in marginal source region. The emissions are thus lowered down by decreasing K. The authors should examine the possibility that the modelled winds are not correct and may exceed the minimum threshold too often. Another source of uncertainty is the fact that the intermittency parametrization was established on a site with a defined wind speed distribution and thus may not be directly applicable in other places with a meteorological context.

Line 404-405. The author mention that following Leung et al (2022), they “cap K at 5”. But in Leung et al (2022) the comments of equation 13a and 13b state that K is limited to 3.

Line 421-422 : It is not clear what Leung et al. (2022) demonstrated on the scale of the dust emission.

Line 437-450 : The capability of GCM and ESM that use spatial resolution of the order of 1° is a key question. This was questioned in particular for regions with complex topography such as the region of the Bodele Depression and addressed by regional modelling.
The upscaling method proposed here consists in weighted the dust emissions computed at coarse resolution by the one computed at a higher resolution, taking advantage of the fact that the input data are available at this scale. This allows to account for the smoothing of the input parameters but it does not account for small scale meteorological features such as meso-scale convective systems that can dominated the dust emissions in some regions like the Sahel. This limitation should be stated and discussed.
It is not clear why the dust emissions computed at the two spatial resolutions are normalized to the total emission to perform the scaling. The normalization implies that if dust emissions are increased in one place, it will be reduced in another. As a matter of fact in figure 1, large areas

in the Sahel, South Africa and Australia have correction factors lower than 1. Why not scaling directly the intensity of the computed emissions?

Final comment on section 3 : At almost all stages of the improvement of the dust emission scheme, adjustable parameters are involved that have been fixed to obtain realistic simulations. But the criteria used to evaluate what is realistic are generally not described. If it is the level of agreement with the observations presented in the following section, it should be clearly stated.

Part 4. Observational and reanalysis datasets for evaluating the dust emission scheme

Line 503-504 : Please precise that the mean AOD used by Kok et al. (2021) based on the data set Adebiyi et al (2020) (but not shown in the paper) are estimated from reanalysis that used satellite products. It would be nice to have the values for the 15 dusty regions in the supplementary material.

Line 523-524 : The overlaying of the AERONET values on figure 5 is unreadable on such small figure. Please add the values and the location of the stations in the supplementary material. In addition, according to the AERONET guidelines for data use and publication, when publishing data from 'many' sites, a general acknowledgement is expected. Citing the appropriate key AERONET papers as well as citing relevant manuscripts pertaining to previously published site data is also expected.

Line 527-539 : Except for the temporal resolution, the MIDAS data set does not seems very different by construction than the combined data sets of Ridley et al. (2016) and Adebiyi et al (2020) used by Kok et al. (2021). Was the consistency of the MODAI data set and the global and seasonal average used by Kok et al. (2021) ?

Line 540-549 : Please add, in the supplementary material, the mean values, the location of the stations and the time period for which PM and deposition measurements are available. The relevant references should be given in order to properly credit the measurements producers/owners.

Part 5. Model evaluation

Line 557-569 : This section is a little bit confusing. Should the reader understand that there is a double tuning, one to reach a global annual DAOD of 0.03 and another to reach an annual mean DOAD of 0.1 in average for the 15 dusty regions ? Or only the second one ? In this case, the sentence “in this section we, we thus also scale our dust simulations with a global tuning factor ..” should be removed. It is not clear why the Z03 simulation cannot be scaled in the same manner. This is a clear advantage for the simulations with the K14 and the new dust scheme proposed here and this may impact the scores of the comparison with observations.

Section 5.1 The section on the simulated dust emissions mainly comments the differences in the spatial distribution between the different dust emission schemes. The global annual dust emissions should be given. Even if there is no absolute reference, there is an expected range for the dust emission simulated with global models to refer to. The dust emissions simulated with Z03 looks very different from the one reported in the paper by Zender et al. 2003, which are not as discontinuous as the one presented here. Why is there such a difference ?
It could be comfortable for the reader to mention the countries in which the mentioned desert areas are located.

Line 584-585 : It is not clear why the capability of the K14 scheme to simulate major sources in the Sahara and the Arabian Peninsula and some emissions in the USA and Asian desert is attributed to the soil erodibility only. The wind fields obviously plays a role too.

Line 592-593 : It is the first time that a possible bias in the input meteorological field is mentioned ! Why should it impact the K14 scheme only if the three emission schemes use the same meteorological inputs ?

Line 595-599 : At this stage of the manuscript, there is no objective argument to sustain the fact that the simulated emissions are over-supressed in the USA and the Thar desert or overestimated in the Tibetan Plateau.

Section 5.2 : In the following the simulated DAOD are regionally averaged AOD from Ridley and to the MIDAS data set for the spatial distribution and the daily scores. The two data sets should have been compared in section 4, and if they bring the same information, as the correlation, slope and RMSE between the two data sets suggest, then the most useful of the two should be used in the following.

Line 633-634 : How much is the DAO on the Strzelecki Desert based on MIDAS ? Is it significant ?

Line 644 : The end of the sentence is not clear. Do the authors mean that the input soil moisture in overestimated in the Taklamakan and the Thar Desert ?

Line 647-648 : Are the higher DAOD levels simulated over semi-arid regions in better agreement with MIDAS ?

Line 654-655 : Once again, if the data set from Ridley et al. (2016) and MIDAS are consistent with each other, why not using MIDAS only ?

Line 660 : Part of the difference obtained with between the Z03 dust scheme and the two other can be due to the tuning of the AOD over the dusty region for the two other dust schemes. Has this effect been evaluated ?

Line 680-685 : The authors argue that bias in the wet and dry deposition may explain the underestimation of the lower regional AOD. But since the deposition parametrization and the precipitation fields are the same for the three simulations, this would affect almost similarly the three simulations. This is not consistent with the fact that the new dust scheme seems to reproduce low DOAD quite correctly and not the others. This argument could be sustained by examining the level of agreement of the three schemes in transport region only.

Line 702-705 : Once again this comparison should have been shown in section 4.3 since the MIDAS data set seems to be considered as a “best guess” used as the main reference to evaluate the simulations.

Line 719-720 : Of course evaluating a model against local in-situ measurements point and regionally averaged reanalysis can yield different conclusion. At least because of the difference in the number of measurements in the different dusty regions used for the comparison.

Line 740-743. How have been selected the daily DAOD threshold value (0.25) ? What is meant by "smoothed out" ? What is the final sampling rate in number of days for the 4 years period ? The argument that regions containing less dust signal are less credible is quite weak. Is't it a requirement of dust emission scheme not to produce dust emissions in such areas ? It has been the main problem of many dust emissions scheme for a long time.

A table with the global mean and the regional mean pearson coefficients for the 15 dusty regions would be more convenient and readable than Figure 6. The differences between the schemes could be discussed from this table. The same suggestion applies to figure 7.

Line 843-844 : The very poor correlation between the dust emissions and U* for the Z03 scheme in figure S8 is extremely surprising regarding equation 3. But from the sentence on line 843-844, it seems that the correlation is computed even when the dust emission fluxes are null, which should be the major part of the time. This obviously impact strongly the correlation and makes the result confusing. A correlation restricted to the situations where F>0 and an indication of the frequency of dust emissions should be more relevant to fully understand the sensitivity of the schemes to the meteorological input data.

Figures 8 and 9 : On the lower panel, the color of the points represents the bias, which is already illustrated by the distance to the 1:1 line. Instead of the bias, using symbols to highlight the different regions of measurements (like on the previous figures) would be more useful and informative. This would also be more consistent with the comments of these figures.

Line 864 : It is not clear what "due to the emphasis by the Z03 source mask" means.

Line 885-869 : The first sentence says that the simulated depositions fluxes are of the order of 10^{-4} kg m^{-2} yr^{-1} or lower and the measurements order of magnitude of 10^{-4}, which is quite consistent. But the second sentence suggests a strong bias in the simulations compared to measurements. Is there a mistake in the numbers?

Line 893-895 : To sustain the argument of an overestimated tropical wet scavenging of dust did you check that wet deposition dominates total deposition downwind of the Sahara ?

Section 5.5 : Since the methodology to up-scale the dust emissions was described in section 3, like all the other change in the dust scheme, it is very surprising to understand in the section 5.5 that this was not included in the simulations and in the deep comparison with observations of the previous sections ! In addition, this section shows that the improvement is quite modest. I would thus suggest to remove this section and to use the complete scheme (including the K correction) in the comparisons of section 5.1 to 5.4. The improvement brought by the scaling can be commented among the other factors.

Line 914-921 : The comments on the effect of this scaling raise again the question of the relevance of a "relative" scaling. As suspected, when the emissions are increased in some regions, they automatically decrease in others. Is it really what is expected from an upscaling method that is supposed to better represent high local dust events ?

Line 941-945 : Of course, the difference in the number of AERONET stations for the different dust sources regions has some implications. This is not specific to the evaluation of the scaling method. It could have been mentioned in section 4.2.

Part 6. Discussion and Conclusions

Line 978-980 : This is a very strong conclusion. It would have been interesting to quantify how much of the improvement in the level of agreement with observations is due to drag partition and to intermittency, like it has been done for K. Concerning the drag partition, it would have been interesting to evaluate the sensitivity to the roughness data set.

Line 1010 : The sentence is ambiguous : isn't it realistic to have extremely low soil moisture over hyper-arid areas ?

Line 1016-1034 : This section could be removed or largely reduced if the comparison with observations includes all the changes in the dust emission schemes including the K coefficients.

Line 2033-2024 : The fact that this "new process-based emission scheme will still work in ESM across different resolution and in regionally refined models" is very questionable regarding the number of parameters that have been adjusted and that compensate most of the uncertainties on the meteorological input data.

Line 1039-1041 : This is the first mention to a bias due to the meteorological forcing. It should have been mentioned and stated in the presentation of the results. The same remarks apply to the possible overestimation of soil moisture or the incorrect estimation of the roughness length in the Taklamakan (line 1052).

Line 1060 : Concerning the improvement of the dust emission physics, the uncertainties due to the process themselves and those due to the input data should be distinguished. As an example, Zender et al. (2003) argued that they used the drag partition scheme from Marticorena and Bergametti (1995) but since they used a unique global value of Z0 it has almost no impact on the simulated dust emissions. But at this period there was no regional or global data sets of aeolian roughness length available. The use of a unique global Dp to characterize the soil size distribution in sources in this work is another good example.

Line 1095-1115 : These factors are quite far from the focus of the paper and could be removed.

Line 1118-1120 : Evaluating the temporal variability of the dust simulations is much more common in regional modelling and could inspire global modellers.