Dear Adrian and co-authors,

I made below a few comments related to your submitted manuscript as I found several cases of mis-representation or inaccurate description of the model and data I have been developing with my collaborators. Hopefully, you will find them useful to improve the manuscript.

Regards,

Paul Ginoux
Thanks for taking the time to make comments on our manuscript. We are a little surprised that none of your comments relate to the achievements described in our manuscript.

We did not intend to mis-represent or inaccurately describe dust emission modelling. We make no special case about the model and data that you have been developing. Our results show an inherent bias when calibrating a dust emission model to dust in the atmosphere and illustrate this using dust optical depth.

We answer each of your remarks below using bold, indented bullet points.

Regards,

Adrian and co-authors

Lines 29-31: “Many of the traditional dust emission models (TEM) assume that the Earth’s land surface is devoid of vegetation, then adjust the dust emission using a vegetation cover complement, and finally calibrate the magnitude of simulated emissions to dust in the atmosphere”
continuity equations. Emission of dust in numerical models depends on the discretization of surface winds. The surface winds are inferred from the pressure level wind vectors derived by solving numerically the momentum equations. The numerical discretization of these equations will be affected by the numerical resolution. Obviously higher resolution will resolve sharp topographic variations with stronger downslope winds. On the other hand, flat terrain without roughness elements using low resolution will generate stronger gustiness. So, changing model resolution has a non-trivial effect on surface winds. Concerning dust emission, the flux depends on the cubic power of surface winds (see Equation 2), which will amplify wind bias related to model resolution. This implies that “tuning” dust emission is a required method to simulate scale-aware tracer with numerical model. This is also true for other tracers, such as sea salt emission from the oceans.

- This comment refers to the italicised excerpt from our manuscript (Lines 29-31) which states clearly that the calibration of traditional modelled dust emission magnitude is to dust in the atmosphere. It is evident from many publications (Zender et al., 2003; Woodward, 2001; Tegen et al., 2002) that the dust emission magnitude is adjusted by a calibration factor. That factor is obtained by converting dust emission magnitude to optical properties that are then compared with measurements of aerosol optical properties. In our manuscript, we show the discrepancy between satellite observed dust emission and dust optical depth in the spatio-temporal frequency of occurrence. Consequently, we explain that dust models calibrated to dust in the atmosphere do not require fidelity in dust emission modelling.

Lines 82-83: “The common approach to modeling dust emission in ESMs uses globally constant values of aerodynamic roughness length (z0), which are static over time and fixes R(z0) ≈ 0.91.”

This is incorrect. In ESMs the momentum roughness length is calculated at every time steps and in every grid cells as a function of terrain variations, vegetation cover, snow cover, etc.
This comment refers to the italicised excerpt from our manuscript (Lines 82-83) which should describe one common approach in ESM dust emission described in the Appendix of our manuscript.

We accept that traditional dust models (TDMs) include roughness lengths but those roughness lengths values may not represent the heterogeneous vegetation that occurs within large (up to 50 km pixels).

We accept that traditional dust models allow those roughness lengths to vary in time / spatially due to snow, vegetation etc. What is held constant, uniformly and globally in traditional dust emission models is the roughness length of each surface type. The roughness length of bare ground never changes, but the ground may not be bare all year. The same is true of desert shrubland which may not be bare everywhere within the land cover type. So in many dust-prone parts of the world, especially those that are bare soil, the roughness length never changes in TDMs. This is the roughness length that the mid-layer (resolved) wind speed is solved for.

With that wind, some TDMs that we are familiar with construct a wind friction speed appropriate for erodible surfaces based on the assumption that erodible surfaces are bare (constant roughness length). That wind friction speed is then used to deflate dust. This "dust friction speed" need not be the same as the friction speed for turbulent fluxes.

It is evident from this discussion alone that the implementation of TDMs varies model-by-model. There is commonality between models as illustrated in the dust emission schemes illustrated in our Appendix.

In response to this comment we will revise our manuscript to ensure that we refer to “one” of the common approaches in dust emission modelling. As we stated in our opening paragraph, our work is not specifically focused on the model and data that Paul and his team has been developing. Our aim is to raise awareness of weaknesses in modelling that provide opportunities for model development for the benefit of the community consistent with the aims this journal’s scope.
“The common approach to modelling dust emission in ESMs... This emission is then reduced by a function of vegetation cover and ultimately ‘tuned’ down to match observed in the atmosphere.”

I am unaware of any ESMs who have implemented dust emission as described. I can certainly speak for NASA and GFDL models (Ginoux et al., 2001; Evans et al., 2016).

- The italicised excerpt (above) refers to a common approach in ESM dust emission which is described in the Appendix of our manuscript. For clarity, our description of the method is that which once (and possibly is still) used in CESM and E3SM.

- As described above (Para 7) we will clarify in our revised manuscript that we refer to one common approach to modelling dust emission and will cease where appropriate, using definite articles. This we hope will provide a more balanced representation of TDMs that may do things slightly differently than TDMs with which we are more familiar.

“The $u_s$ is obtained directly from $\omega_n$, the normalised and rescaled shadow (1-albedo), enabling an albedo-based dust emission model (AEM; see Appendix for full description of the implementation)”
Do I read correctly that you are scaling the friction velocity using 3 parameters with an exponential function of $\omega_{ns}$? Am I right that you will have to rescale $\omega_{ns}$ for any other satellite instruments with different viewing angles or radiometric characteristics? Is this not a global tuning?

- Our albedo-based method was developed from a series of ground-based bi-directional spectral reflectance experiments in the laboratory and in the field (Chappell et al., 2005, 2006; Chappell et al., 2007). Over the last ten years or so, the approach has shifted from spectral reflectance to waveband independent albedo focused on land surface structure (Chappell et al., 2010). Our approach was shown to be similar to the widely accepted sheltering concept and its implementation (Raupach, 1992; Raupach et al., 1993). We established the bases for how the albedo-based approach could improve drag partition, sediment transport and hence dust emission (Chappell and Webb, 2016; Chappell et al., 2019). Since then we have confirmed the structural information content and demonstrated its scale invariance from ground-based measurements of albedo to MODIS-based estimates (Chappell et al., 2018; Ziegler et al., 2020).

- The shadow (1-albedo) is normalized to provide structural information for an area and is directly related to the wind friction velocity (of the area) measured in wind tunnel experiments (Chappell and Webb, 2016). To enable this grain-scale areal relation to be used at any other areal scale requires a measurement of albedo. Albedo scales linearly with increasing area in contrast to the wind friction velocity. To explicitly tackle sub-grid scale heterogeneity (Raupach and Lu, 2004) albedo measurements (or indeed any area-weighted estimates) must be rescaled before being used with the same calibration.

Line 128-132: “Evans et al., 2016”
The characterization of GFDL model (Evans et al. 2016) is not correct. You may want to read the paper. We are not using E=1-Av. The bare surface is calculated using an exponential function of the canopy (LAI) and stems, twigs, litters (SAI). The dust emission is calculated in each land tiles (primary, secondary vegetation, pasture and cropland) independently. Then the flux of dust is passing through a flux-exchanger into the atmosphere while a flux down from turbulence and settling is going in the land model. The latest ESM4 includes also tiles from fires and rangeland, in addition for taking into account slopes (Dunne et al., 2020; Horowitz et al., 2020). I will disagree with you when calling such detailed and consistent modeling of dust cycle a “crude model representation”

- The full sentence (quoted below from the manuscript at lines 128-132) reveals that the citation of Evans et al. (2016) is one of several which correctly cite dust models which use NDVI.

“This E is used in some ESMs so that leaf area index (LAI) or satellite ‘greenness’ observations e.g., normalized difference vegetation indices (NDVI) can be used as a surrogate of the land surface fraction occupied by green vegetation (Evans et al., 2016; Galloza et al., 2018; Zender et al., 2003a; Sellar et al., 2019).”

- The exponential relation between bare area and LAI (or NDVI) is an empirical representation of the drag partition which does not account for the physical aerodynamics that fast wind speeds are influenced less by a given roughness than slow wind speeds (Raupach et al., 1993; Chappell and Webb, 2016; Mayaud and Webb, 2017; Webb et al., 2014; Walter, 2012). Consequently, changing vegetation cover (in planform) is not itself sufficient to represent a change in sediment transport and dust emission.

- We acknowledge that vegetation indices can be used to obtain a first order correction to the erodible surface area, whether through thresholds and ramps (as CESM once did) or through more sophisticated canopy schemes like GFDL.
We note that the full sentence also refers to some ESMs and as we have described (in our opening paragraph) there is no attempt in our manuscript to comment on model and data that Paul and his team has been developing.

Lines 133-135: “When the TEMs are applied in dust-climate ESMs it is assumed that this parameterization is adequate for climate projections. In contrast, the albedo-based scheme for sediment flux and dust emission (AEM; Eqs. 3, 4 & 5) represents the drag partition physics without pre-tuning to a fixed land surface condition, without the need for $E$, and thereby removes these additional sources of uncertainty.”

The main point of using climate model has been missed here. Despite their approximations, ESMs simulate the different Earth’s climate systems consistently over time using different projection scenarios. While the proposed used of $\omega_{ns}$ (the normalized and rescaled shadow) is considered fixed (beyond MODIS period), ignoring vegetation and land use changes. The AEM technique is inadequate for future or past climates.

We recognise and appreciate the experience that Paul’s comment provides. We agree that ESMs play a critical role in understanding Earth’s systems and that they must be designed not to rely on observations.

Our scheme is albedo-based from any source or scale as evident in our recent investigation of scale invariance using both ground-based albedometers and satellite-retrieved albedo (Ziegler et al., 2020).

Since albedo is one of the main drivers of energetic change in ESMs there is considerable potential for ESM-based albedo (not MODIS-based albedo) to drive the albedo-based dust emission model which will enable multiple internal consistencies in the modelling of the energy and dust cycles.
- In the excerpt above there is no mention that albedo is taken from MODIS.

Lines 145-149: “To understand the extent to which AOD estimates the spatial variation in dust emission magnitude and frequency we calculated the probability of dust occurrence modeled by the dust optical depth (DOD>0.2) using the criteria established previously (Ginoux et al., 2012). We note the stated limitations of DOD to be largely restricted to bright land surfaces in the visible wavebands which implies reduced performance over areas where vegetation is present.”

This sentence contains several misunderstandings of our latest method developed with my co-authors to derive DOD. In our 2012, we used the collection 5.1 of MODIS Deep Blue (DB), which provided aerosol products over bright surfaces. Since 2013 Collection 6 MODIS DB aerosol products have been extended to cover most (without snow or cloud cover) land surfaces (Sayer et al., 2014). All subsequent papers deriving DOD is using Collection 6.1 MODIS DB (e.g. Pu and Ginoux, 2017, 2018a, 2018b, 2020; Yu and Ginoux, 2021). A second update is the method to calculate DOD. Since Pu and Ginoux (2017), DOD is calculated using a quadratic function of aerosol optical depth (AOD) and the single scattering albedo (SSA). In our 2012 paper, DOD is calculated using an on/off switch depending on the value of the Angstrom Exponent (AE). Then a threshold is applied to detect the highest frequency to correspond to actual dust sources. The method has been compared to independent geomorphological data over the Chihuahuan desert (Baddock et al., 2016) to prove that MODIS DB DOD is able to successfully detect high-resolution dust sources. It will be necessary to add a note the text stating that you are referring to an old dataset long replaced by thoroughly validated values using the latest MODIS aerosol products. Preferably, you replace the sentence by referring to more recent thoroughly validated values using the latest MODIS aerosol products.

- Satellite observed dust emission point sources (DPS) have an uncertainty of around +/-2 km (Kandakji et al., 2020) due to the phase difference between timing of dust emission and availability of the imagery. At the point scale, this unexplained variance is reduced by aggregating information at a larger scale. This is well-established in geostatistical literature used across multiple disciplines and the issue of incompatible scales is summarised (Gotway and Young, 2002). We reduce the unexplained spatial variance by aggregating the DPS data.
The excerpt above describes clearly that we used the established criteria (Ginoux et al., 2012). We then used those criteria with monthly MODIS Deep Blue (MOD08 M3 V6.1) data to establish dust optical depth across 1-degree grid boxes. Those grid boxes coincide with the occurrence of satellite observed dust emission point sources (DPS). These DPS data were aggregated to months across the same 1-degree grid boxes to provide compatibility with the DOD data. We note that these DPS data occur mainly in bright land surfaces which include vegetation.

In our revised manuscript, we are happy to modify the description of our method to ensure that it is balanced. For clarity, we have used the latest MODIS Deep Blue Collection 6 dataset and apply the established criteria (Ginoux et al., 2012).

In response to Paul’s query about the method, we looked up the recent methods (Pu and Ginoux, 2017) which cites (Ginoux et al., 2012) for their method. Apart from some small alternations to the Angstrom component, we could not find any no difference between our method, what we described in the manuscript and the method described recently (Pu and Ginoux, 2017).

Line 149-150: "To calculate DOD, we used wavebands available from monthly Moderate Resolution Imaging Spectroradiometer (MODIS; MOD08 M3 V6.1) at a 1-degree pixel resolution (Platnick, 2015)"

These are monthly gridded products at 1 x1 degree resolution from 1 km daily pixels. Using coarse resolution monthly aerosol products is totally inadequate to compare with point source dust plumes.

As described above (para 20) there are compelling reasons to use aggregated DPS data. Consequently, we aggregated the DPS data over space. For compatibility with the monthly, 1 degree grid we aggregated the DPS to that same resolution.
In our assessment, frequency of occurrence from either the model, or the DPS observations are described by the occurrence (modelled or observed) of dust emission at any of the DPS pixels within the 1° grid cell. i.e., should one DPS produce dust, that cell is considered on (FoO = 1) for that day.

The DOD data describe the probability of any part of the 1° cell producing DOD > 0.2 consistent with the DPS FoO. If the reviewer is arguing that DOD data are ‘totally inadequate’ to verify emission from dust point sources, why do many studies use DOD results to therefore describe the spatial variability in dust emission (Ginoux et al., 2012: Pu and Ginoux 2017). The purpose of this analysis was to demonstrate this very point, highlighting the inadequacy of DOD by using actual dust observations to verify simulated dust emission.

Consistent with previous studies using DOD to display spatial variability, we show the discrepancy between spatial variation in DOD and dust emission observed using satellites.

That discrepancy between DOD and observed dust emission does not occur when DOD is compared with simulated dust emission.

The cause of that discrepancy is that DOD measures dust in the atmosphere which is not directly related to dust emission.

The difference with MODIS DB DOD from my team is that it is 0.1 x 0.1 degree twice-daily
products, which was shown to be appropriate to detect tiny dust sources (Baddock et al., 2016).

- **We make no comparison with the work done by your team comparing dust in the atmosphere at 10 km pixels with satellite observed dust emission point source (DPS) data. In our approach we tackled explicitly the incompatibility of spatial scales (Gotway and Young, 2002) by upscaling DPS data to 1 degree pixels consistent with the DOD data and simulated dust emissions.**

It is unclear how you get these DOD. Citing Platnick, 2015 is not helpful. What MODIS Level 3 products are you using? Are you using Dark Target or Deep Blue algorithm? Maybe you are using a blend of the two algorithms? DOD is not part of these products, so a description on how you obtain DOD from MODIS would be helpful, especially that you assimilate it to Ginoux et al. (2012) which is outdated.

- **As described in the manuscript and in our response above (para 9) we used the established criteria (Ginoux et al., 2012) with monthly MODIS Deep Blue (MOD08 M3 V6.1) data to establish dust optical depth across 1-degree grid boxes. In the revised manuscript we will include the specific wavebands to make this clearer.**

Line 154: “We also provided a theoretical basis for TEMs formulation to be incorrect.”
I commented earlier that your description of TEMs formulation is mostly incorrect.

- We described above (para 9) that the original manuscript erred by implying that all TEMs are formulated in the same way when in fact there is a spectrum of implementations for each process. The revised manuscript will avoid this overgeneralization when using indefinite articles e.g., "many TEMs". Nevertheless, it is important to give example process formulations and in those cases we have double-checked and the references previously given for TEM process descriptions are correct. We cannot find in your above text the argument by which our description in the Introduction of the manuscript of the theoretical basis for TEMs formulation is incorrect. The basis for the incorrect formulation is already published (Webb et al., 2020) and here we are extending that work to dust emission modelling.

Line 166: “using MODIS data at 250 m spatial resolution with visible to thermal infrared wavebands”

This is incorrect. Only bands 1 and 2 are provided at 250 m. Bands 3 to 7 are at 500 m resolution. Bands 8 to 36 are at 1 km resolution. Red is band 2, blue is band 3 and green band 4. Deep blue is band 8 or 1 km pixel. Then 10 x 10 pixels are aggregated to provide 10x10 km Level 2 daily aerosol products.

- To clarify this statement we will remove from the revised manuscript mention of “...at 250 m...”.
In Baddock et al. (2016), we used DOD>0.75 over the Chihuahuan desert, but Pu et al. (2020) used 2 threshold values (0.2 and 0.02) depending on the continent. Choosing a DOD threshold should be adapted to the objectives of the study but...

- The DOD data are describing dust in the atmosphere and the DPS data are describing dust emission. The discrepancy between DOD data and DPS data is not impacted by the choice of threshold.

... using gridded 1-degree monthly products is too coarse spatially and temporally to study dust sources.

- We explained above (para 20-29) that we ensured that the gridded 1-degree monthly DOD is at the same scale (space and time) as the DPS data. In this approach we tackle explicitly the incompatibility of spatial scales. Our methodology explains and our results demonstrate that these 1-degree monthly products are not too coarse to study dust sources.

Lines 248-249: “the TEM is driven by wind speed attenuated by aerodynamic roughness which is fixed over space and static over time,...”
If the surface conditions don't change (no snow, no vegetation or land use changes) this will be true, but most ESMs (or TEMs) include such changes when resolving the boundary layer properties.

- **We have described in our responses above that there are at least two perspectives on this issue and we think that they converge. We will revise our manuscript using definite articles and describe some of the different ways in which surface conditions are represented in dust emission models.**

References cited in our responses


Chappell, A., Strong, C., McTainsh, G., and Leys, J.: Detecting induced in situ erodibility of


Pu, B. and Ginoux, P.: Projection of American dustiness in the late 21st century due to climate change, Scientific Reports, 7, 5553, 10.1038/s41598-017-05431-9, 2017.


Walter, B.: Wind Tunnel Studies of Shear Stress Partitioning in Live Plant Canopies,


