In this paper the authors present an approach to verifying dust emission model performance by considering how the model(s) compares to satellite observations of dust emission point sources (DPS). For this purpose they consider dichotomous (presence or absence) comparisons of near-global DPS observations (from MODIS and SEVIRI, for North America, Northern and Southern Africa, the Middle East and Central Asia, and Australia) and simulated emissions from their albedo-based dust emission model (AEM). The model appears to overestimate the number of dust events by an order of magnitude, while the DPS database indicates that dust emission events are actually quite rare in space and time, considering the vast size of some of the deserts. The authors use this dichotomous information to understand the conditions under which the model performance is influenced by the representation of the wind fields, the assumption of infinite sediment supply, and the threshold wind friction velocity $u_{t,s}$. There is value to this approach to validating dust emission models using observations that represent (as closely as possible) the dust emission events that actually occurred.

- We found your summary of our work to be consistent with our intent. Perhaps just one point of clarity. The frequency of occurrence evident in the dust emission point source (DPS) data is governed by the very large number of pixels where dust emission can occur but also their opportunity to occur in time. Both these conditions for actual dust emission are constrained by the entrainment threshold and the availability of sediment. Therefore, DPS data over space and time record dust emission for comparison with dust emission models.

I am curious about this repeated claim in the manuscript that “The aim here is to demonstrate an alternative to comparing dust emission models to atmospheric dust” (line 79). As I see it, the satellite dust emission point sources are themselves interpretations of observations of atmospheric dust, back-tracking atmospheric dust in the satellite imagery for the purpose of identifying these emission points. Hence it seems to me that the authors are themselves still comparing the dust emission models to atmospheric dust.
suppose there is a distinction here between comparisons with the retrievals of atmospheric dust transport (implicitly or explicitly being criticised), as described by the dust AOD, and with the derivations of dust emission from atmospheric observations taken shortly afterwards (apparently a more robust method).

- Whilst both approaches described here, and in the manuscript, have their origins in optical satellite remote sensing, each approach is measuring a property very different from each other. In the most straight-forward situation an image is used to pin-point the source of dust emission; that is a dust emission occurrence in space and time quantified by coordinates \((x, y, z)\) in latitude, longitude and time (as shown in the figure below from Lee et al., 2009).

- “Fig. 1. A MODIS image (sensor: Aqua) of the region during the dust storm; image obtained from: http://visibleearth.nasa.gov/view_rec.php?id=19043. The image used has a pixel size of 250 m. Political boundaries, cities and points to identify dust sources were added by the authors. The source points were identified on an enlarged version of this image, with greater detail than shown here” (taken from Lee et al., 2009).

- In the example given above, manual inspection of the image permits the observer to identify only the point of origin (white circle) of a visible [atmospheric] dust plume. In contrast, measurements of Dust Optical Depth (DOD) are performed automatically, with limited opportunity to differentiate source of emission from adjacent pixels of transported dust. For that same image, DOD (above a threshold) reduces the information content to presence / absence of dust in the atmosphere, including areas downwind. These dichotomous values include dust emission, but also dust that may be suspended for days (depending on particle size and wind patterns). Consequently, dust emission source represents a diminutive contribution to all observed dust pixels, with the majority of DOD pixels representing transported atmospheric dust.

Now I actually agree that it might be more robust to try to pinpoint the dust emission sources and to compare those with the dust emission model, but I think that it would very much be worthy of clarification and further discussion on the distinction between the two types of satellite observations of atmospheric dust which are being considered here.

- This approach to pinpointing dust emission source has been used and has been evident in the literature for >10 years. With the support of the community working with these data, we have collated most of the published studies. In our manuscript, we used more than two pages to describe these DPS data and how they are retrieved. We then need to describe the innovation within a modelling framework to get the most value out of these data.
Considering your perspective, in the revised manuscript we will consider how best in the space available to explain the methodology.

I think it would also be worthwhile to clarify the final sentence of the abstract ("... dust emission models should not be evaluated against atmospheric dust") to indicate that this relates to the dust AOD, assuming that this is what you mean.

- **Thanks for pointing that out. We will clarify that point in the revised manuscript. Note that when we pin-point the dust emission point source it is related to but not a description of dust in the atmosphere.**

This points to a general criticism that I have, that while this does seem to me to be a worthwhile paper, describing a useful method of comparing models with dust emission sources identified by satellite sources, I do feel that it is let down at times by over-generalised statements such as the one described in the above paragraph. Another such statement is "These results demonstrate that there is no reasonable basis to calibrate model performance through an adjustment to a fixed global $u_{ts}$" (lines 482-484). If this is an implicit criticism of some model dust emission schemes, is this not something of a straw man argument? I would have thought that it has been quite widely accepted for many years that $u_{ts}$ is dependent on the soil size distributions and roughness lengths, which vary worldwide. I suppose however that there is some value to confirming that this should not be done, as has been done here.

- **Most, if not all, dust emission models assume that the entrainment threshold is fixed and static over time. This assumption in dust emission modelling has been evident since dust emission models were developed more than two decades ago. Our statement quoted above is evident from our results and it is a timely reminder that dust emission results and interpretations are not based on robust dust emission modelling but on calibration with atmospheric dust. With this new framework for understanding large scale dust emission patterns and timing we have new insight to tackle these long-standing assumptions.**

**Specific comments**

- **p.5, lines 156-157: a minor point, but as written this sentence implies to me that the red beam is associated with the 8.7 µm channel, while the green beam is associated with the 12.0 µm channel. It is of course the other way round.**

- **Thanks, in the revised manuscript we will improve this description.**

- **p.7, Eq.6: presumably in these equations it is $(1-A_s)$ multiplied by $(1-A_f)$. As written, this equation implies to me that no snow = no dust!**

- **Thanks, in the revised manuscript we will improve this description.**

- **Figure 3: it might make sense here to fit the y-axis to 0-1 (as has been done for the x-axis, 0-0.6), it would make it slightly more intuitive to read.**

- **Agreed, we will do that in the revised manuscript.**

- **p.13, line 392: “Taklamakan (Kazakhstan)”. Presumably you mean China.**

- **Thanks, in the revised manuscript we will correct this mistake.**
References used in our responses