This is an interesting paper that uses wind-tunnel experiments to put forth the hypothesis that the renewal of fine particles in a soil’s top layer is critical to sustaining dust emissions. This is an appealing hypothesis and this process is currently missing from models. They test this hypothesis using a series of wind tunnel measurements, which seem well designed. This article thus has the potential to be an important contribution.

However, there are several major issues with the article. Paramount is a major deficiency in how the results and discussion are presented. Almost throughout the “Results and Analysis” and “Conclusions” sections, the authors present hypotheses, of which their (otherwise very interesting) data are merely suggestive, as facts. Words such as “show” and “demonstrate” are used abundantly. This is not appropriate considering the level of evidence the authors present, and the enormous complexity of dust emissions. I will give a few (of many) examples of this below. The authors need to completely
rewrite these sections. In particular, they should split up the “Results” and “Discussion” sections, to make it clear what are indisputable facts from their experiments, and what is their interpretation of these facts.

In addition, there are some major scientific issues:

- The saltation bombardment section has major issues, which I’ll list below:
  * “c0 reflects the fraction of effective saltators, namely, grains available for saltation at a given friction velocity”. This is inconsistent with Owen (1964), and also with the paper’s own Eq. 18, where c0 is linked to the terminal velocity.
  * P. 7, lines 10-11: I’m not aware of any measurements supporting the idea that the number of available saltators depends on the (theoretical) thresholds for individual particles. Rather, when saltation is initiated, the splashing process can mobilize particles of a wide range of sizes (e.g., Rice et al., 1995). The authors should either provide experimental evidence for their viewpoint, or note the opposing view (even if they do not adopt it).
  * P. 7, lines 13-14: Did the authors directly measure what particles constituted the saltators? If not, this is interpretation, yet as presented as fact.
  * P. 7, lines 14-15: This similarly is interpretation presented as fact. The saltation flux depends on many closely coupled and complex processes. Linking a change in the flux to any one parameter (the fraction of effective saltators in this case) without directly measuring it is speculative. That’s fine to do in the discussion section, but should be presented as such.
  * For the fitting with equation (15), how was u^t obtained? Was it fit as well? And how was v_t calculated?
  * The use of Eq. (16) – (19) is very interesting. However, the procedure here is very unclear to me, and might have some scientific flaws. My primary concern is that the parameters in the u^t relation seem to be fit to the measurements, such that Eqs. (16) –
(19) have, as far as I can tell, three tunable parameters (proportionality constant (c0?), r, and An). Since the data they fit to are only four data points, these fits are statistically not that meaningful (only 1 degree of freedom). Thus the conclusion that “the above method gives a more accurate estimate of Q than Equation (15)” needs to be put on a more solid statistical basis.

* Related to the above comment, please provide the fitted u*t(d) relationships for the three soils so that the reader can judge whether they are reasonable. This is necessary to judge whether the visually good agreement is due to a good description of the physics, or because of a sufficient number of tuning parameters. You could provide these fits in a supplement to the paper.

- Sections 4.3 and 4.4: A central argument of the authors here is that the dust supply for aerodynamic entrainment is maintained by the intense sand flux for S2 and S3, but not for S1, which has lower sand flux at a given u*. However, the authors should compare apples to apples here and thus compare data with similar sand fluxes, for instance u* = 0.37 m/s for S1 and u* = 0.23 m/s for S3. The S1 data point shows a large dust flux decrease during the first minutes, whereas the S3 data point does not. This is not explained by their hypothesis, and should be clarified.

- I found section 4.4 very difficult to follow. Please use paragraphs in this section and make sure that the text flows smoothly. More importantly, this section again uses many interpretations of the data and would benefit enormously from separation into a results (facts) section and a discussion (interpretations and hypotheses) section. As it is written, I cannot sufficiently judge the scientific merit of this section.

- Section 4.5 suffers from similar issues as the other sections, with many hypotheses presented as though they were measured experimentally (line 9-11 “Due to the neglect of the supply-limiting effect and of the variation of bombardment efficiency, all three models underestimated the dust flux at low friction velocity, but slightly overestimated at high friction velocity”; line 14-15 “With the increase of u*, the bombardment
efficiency decreases because of changed surface property due to intrusive sand particles.”; line 18-19 “S04 appears to perform somewhat better than the others due to improved treatment for saltation bombardment and aggregates disintegration.”; line 21-22 “This shows that threshold friction velocity u*t represents different properties of the soil surface in the Owen model and the GP88 model.”)

Other comments:

- Please make line numbers continuous in revised article to make the review easier.

- In the literature I'm familiar with, the term “supply limited” is generally used to refer to a lack of supply of saltators, not a lack of supply of fine soil particles. The authors should clarify this point.

- Line 31-32, p. 1: Why do differences in dust emission after disturbing a soil indicate the importance of aerodynamic entrainment? This should be clarified or removed.

- Sections 2.2 and 2.3: While the authors cannot be expected to compare their data against every single dust emission model, they should at least mention the other ones (e.g., Marticorena and Bergametti (1995); Alfaro and Gomes (2001); Kok et al. (2014)).

- Eq. (9): What is the averaging time for u(z)?

- Line 15, p.5: This statement on sonification requires justification. For instance, the impact of saltating particles can chip and break them, which does not occur during sonification. Therefore, whereas sonification disaggregates particles, won’t grinding result in the wearing down of individual (disaggregated) particles, thereby changing the size distribution?

- Please add a brief discussion whether the use of the gradient method is reasonable for your experiment. Compared to field measurements, your fetch is very small (a few meters, compared to 100s or 1000s of meters in the field). You partially compensated for this by moving your dust sensors close to the ground, but can you expect dust to be well-mixed (and thus follow a logarithmic profile) at only a few meters of fetch? How
will this affect your results?

- Section 3.3: was the wind flow seeded with particles in your experiments? If not, do you expect your sand flux to be saturated? The results of Shao and Raupach (1992) suggest that you need more length than the 8 m of your set-up.

- P. 7: please define d1 and d2 in Eq. 16. Also, the last d should be d_s

- P. 7: Please provide the value of the particle-to-air density

- In general, how exactly is the fitting performed? What quantity is minimized? Given that the data spans several orders of magnitude, it makes most sense to me to minimize the squared distance in log space, not in linear space (as the authors seem to have done).

- P. 8, line 3: does this refer to radius or diameter? Does this mean that the reported dust fluxes are limited to D (or r) < 15 um? Please clarify.

- P. 8, line 10-15: There are a lot of hypotheses used here to interpret the data in terms of arising from either aerodynamic entrainment or saltation bombardment, and whether or not the dust supply was limited. These factors were not measured directly, so these interpretations should be presented conservatively, rather than as statements of facts.

- P. 10: The scaling of aerodynamic entrainment with $u^*$ to the 10th power seems a bit extreme. Can you put uncertainty bounds on this result? How does this compare against other literature measurements such as Shao et al. (1993) and Loosmore and Hunt (2000)? What could explain the differences? Also, since you did not actually measure just aerodynamic entrainment (saltation was always present, as far as I understand), this conclusion should be more conservative.

- P. 11: “Supply limit is the major reason to restrict dust emission.” This statement illustrates the main problem with the paper in its present form. Your measurements do not show this because you did not directly measure the supply limitations. You are merely hypothesizing this based on other measurements. I think it’s a reasonable
hypothesis, but needs to be presented as such, and not as a fact or hard conclusion. This problem is persistent throughout the entire paper.