Anonymous Referee #1

Received and published: 2 October 2014

We would like to thank the referee for a set of very thoughtful comments which eventually helped to improve the quality of the manuscript. In the following, we will respond to each comment line-by-line.

General comments

This study examines the performance of three well-known dust emission schemes in a box model environment constrained by observations from the DO4Models field campaign in Botswana during July-October 2011. Modeling the atmospheric desert dust cycle has improved the past two decades but the remaining challenges and uncertainties are still significant. The most important aspect of this problem is the lack of detailed measurements of the dust vertical flux emitted from the soil in a variety of arid areas worldwide. Such measurements could help the modelers improve the emission parameterization schemes and take the leap towards more robust simulations and predictions of the dust cycle. The paper is well written and concise. In terms of scientific significance, I believe that the knowledge gained from the sensitivity model experiments as they are compared to measurements is an important addition to current knowledge and practices. My comments concern mostly some obscure parts of the manuscript where clarifications are necessary to improve the quality of the narrative. The only omission I found in this work was a discussion on the errors in the measurements that might have influenced the results. I am in favor of publishing this paper with Geoscientific Model Development, after carefully addressing the minor comments that follow.

Since the originally submitted version of the manuscript was heavily relying on the input of a companion paper by Wiggs et al. regarding the treatment of measurement errors and uncertainty, we are happy to report that the Background chapter has been entirely rewritten, covering all relevant aspects of measurement errors and potential implications. Moreover, ambiguities and/or unsupported claims have been removed or reformulated, particularly in the heavily revised Chapters 4.2 and 4.3.

Note that all figures have been re-plotted to reflect very late changes in the soil size distribution data which also led to a change of the values in Table 1. For the soil moisture correction, gravimetric soil moisture content ("GSMC") rather than volumetric soil moisture content ("VSMC") - as indicated in the first version of the manuscript - has been used throughout.

Specific comments

Introduction
1. Page 5740, line 25: "from a remote sensing"
2. Page 5741, line 17: "In this paper we report..."

Amended.

Background
1. Why did the authors choose Sua Pan for the field campaign? What characteristics of the soil surface make this location unique and at the same time common to a large number of soil types in desert areas worldwide? In other words, are the conclusions transferrable to other arid regions as well or they are limited to the specific crusted type of surface? Please include the information in the text.

It's mainly because Sua Pan is situated away from major North African dust sources which tend to have high background dust aerosol concentrations. Down south, background concentrations are no major concern which is why we think it's an ideal place to study the emission process (for more information, see Chapter 2 in the revised manuscript, lines 149ff**).
2. Page 5743, line 21: "anemometers at heights of".

Amended.

3. Page 5743, lines 18-29: The description and differences between the measuring sites is not clear in the text (i.e. the dust sensors are included in all 11 sites?). Please make a distinct description of the AWS and MET sites as to the differences between them.

More detail has been added.

Box model development
1. The correction factor 2.61 used in the MB95 scheme was set to 1 in later publications (Marticorena et al. 1997, Laurent et al. 2006, Darmenova et al. 2009). Why did the authors choose to include the value that originates from the experiment of White (1979)?

We actually used 1.0 as correction factor as well. We have added more detail to make that clear in the manuscript (lines 331-335**)

2. Page 5747, line 21: How do you calculate the grain size velocity ws?

Information added (function of particle mass, diameter and the drag coefficient in consideration of different possible Reynolds regimes).

3. What is the difference between minimally and fully disturbed soil size distributions? This is important to put the discussion on the findings in the right context.

Described in more detail in the manuscript now (see Chapter 2, lines 188ff**): [Surface sediment] was used in "wet" fully dispersed mode (assumed to represent the dust in suspension), and also in "dry" minimally dispersed mode using an air dispersion unit (which maintains and measures any particle agglomerates which might be assumed to comprise the saltation flux).

4. Table 2 is confusing. Sensitivity experiments (a) to (d) occurred for all model setups? If yes, then I suggest reordering the rows by leaving the experiments a-d last and inclusive of all setups.

Amended and further simplified.

Results and discussion
1. Figures 2-4 contain a large number of sub-plots that make the reading of the figure very difficult. The plots are an important part for this work. I would suggest to either make the subplots bigger or cut each figure in two.

We are afraid that this is the only point we have concerns with. While certainly filled with plenty of information, we can't really see how splitting the panels into several sub-plots would make it easier to interpret. We do admit, that we have seen much smaller plots elsewhere and are therefore a bit surprised that Figs. 2-4 cause issues. While we are certainly able to cut the plots, we would like to ask the editor to make a final judgment. For the time being, we have left them as they were before.

2. Figure 1 could be improved from the indication of the measuring sites in their respective location.

Fig 1 has been re-done completely. We hope that helps to facilitate the reading.

3. A discussion on the possible sampling errors/uncertainties in the measured quantities is essential. Inclusion of quantitative values is desirable, if possible.

As pointed out earlier, Chapter 2 now contains several paragraphs which are dealing with measurement uncertainties in all detail.

##

Anonymous Referee #2

Received and published: 3 October 2014

We would like to thank the referee for a set of very thoughtful comments which eventually helped to improve the quality of the manuscript. In the following, we will respond to each comment line-by-line.

General:

The authors developed a box model framework for three different dust emission modules. This work exposes the shortcomings of all these three modules in reproducing measured dust emission fluxes from a field campaign in Botswana in 2011. The authors show that the simulated horizontal (vertical) fluxes are several orders (one order) of magnitude too high. They ascribe the differences to crusted surfaces that are not represented in the emission schemes and to dust entrainment and conclude that both processes should be included in future emission schemes.
The emission is the first step of the entire dust cycle and a good reproduction of the atmospheric dust load and its deposition in weather and climate models crucially depends on the dust emission. The here presented box models of different emission schemes and their systematic investigation will help to understand the critical processes that need to be captured by models in order to simulate the dust cycle in a realistic way. The description of the box model development is in general comprehensible but needs some clarification (see below).
My major criticism refers to the comparison with the measurements. Although the field campaign might provide one of the best data sets in terms of horizontal and vertical mass fluxes, emission fluxes were only measured on very few (five?) days. How representative are these measurements? How about measurement errors? One needs to be careful in drawing general inferences from these measurements about the overall performance of the dust emission schemes. Furthermore, there is the question on the representativity of the field sites. The authors mention that the crusted surface at Sua Pan "can be found in many dust source regions" (P5743) (where?) but on the other hand, that the soil combustion is different to "many other desert soil samples" (P5764). Please include a discussion of this in sections 4 and/or 5.
My recommendation is to publish this paper in GMD after addressing the following comments. To make the paper fit better within the scope of this journal, the focus should be little more on the development of the box models and on potential future applications/extensions while section 4 needs to be condensed (see major comment below).

Since the originally submitted version of the manuscript was heavily relying on the input of a companion paper by Wiggs et al. regarding the treatment of measurement errors and uncertainty, many very reasonable questions have been asked which need to be addressed indeed. Due to the fact that said paper is still not available, the Background chapter has been rewritten entirely, covering all relevant questions raised by the referee. Some issues are a result of unfortunate phrasing or insufficient detail in the original manuscript which has led to minor confusion. We have either added more context or reformulated in a more concise fashion in these cases in order to resolve the problems.

Major comment:

*Figs. 5, 7 and 8: The conclusions drawn from these figures mainly refer to the dependency of u*thr on the correction schemes in general, on the moisture and on the roughness. In total 18 panels are too many for this analysis. The question is: What is the "full range of observed u*thr values" (P5764, L12) and what do the emission schemes simulate? In my opinion, the observed u*thr values can be read from Fig. 7 for different conditions with regard to soil moisture and roughness. The median for the different clusters (Obs z0>1cm, . . ., Obs vsmc=0-3%) would represent a good estimate for the upper limit of u*thr under these different conditions. For the emission schemes u*thr can be calculated for all the different conditions explicitly, i.e., the u*-value at the leftmost end of the lines in these figures.*
*The same could be calculated for all sites separately (Figs. 5 and 8) and compared with the respective u*thr of the various emission schemes and all experiments. The comparison of these values of u*thr can then be done in a much more concise way without showing 12 panels in Figs. 5 and 8. Fig. 7 can stay as it is, as it gives a nice overview on how the observed and simulated fluxes depend on u*, z0 and w.*
The distinction of different soil classes as given in Figs. 5 and 8 can be skipped as it not really discussed in the text.
Some of the minor comments below might become superfluous when the figures are changed and section 4 is partly rewritten by addressing this comment.

This is actually the only point to which we would object to some extent. Given that we have condensed the information of literally thousands of plots into 3 figures (Figs. 5, 7 and 8), we do think that they are all relevant as each of them conveys the gist of the analysis in a different way. We are aware that all 18 panels refer to the dependency of u*thr on moisture and roughness, however, this is the part of the emission schemes that can be analyzed in all detail with our field data, whereas the implicit sand transport models are more binary in that they either do or do not match the flux estimates based on observational data (Sensit particle counter and DustTrak). We did reduce the number of panels in Fig 8, but we argue that the reader might wish to see the "true" variability as represented by field observations (Fig. 8) in addition to Fig. 7 (which focuses on the impact of moisture and roughness in much more general terms). We may consider merging Figs 5 and 8 into one single figure if we are urged to do so, though.

We also think that it is beneficial for the reader to know what the emission threshold of the individual size categories is (thin lines in Fig. 5 and dashed lines in Fig. 8) as opposed to the sum of all bins represented by the bold line in Figs. 5 and 8. We have made the distinction clearer in the figure capations.

The associated text in Chapters 4.2 and 4.3 has been thoroughly revised with particular regard to the points raised by the referee. For example, the confusing part which referred to the "reproduction of the full range of observed u*thr values" has been taken out and been reformulated in a shorter and more concise way (lines 922-937** in the revised manuscript).

Note that all figures have been re-plotted to reflect very late changes in the soil size distribution data which also led to a change of the values in Table 1. For the soil moisture correction, gravimetric soil moisture content ("GSMC") rather than volumetric soil moisture content ("VSMC") - as indicated in the first version of the manuscript - has been used throughout.

Minor comments:

- P5741, L17: What does "undisturbed" mean here?

Dry processed, i.e. particle aggregate size is maintained (or size distribution for that matter).

*- Equations in general: Please be more precise in the description of all the variables in the equations. Some examples: How is u*dry defined in Eq. 1? It is not the one from Eq. 7,I guess, as this one is "adjusted" (P5748, L23); What is the definition of w_s in Eq. 5?; What is rho and g in Eqs. 7 and 6?*

Checked and amended wherever needed.

- P5748, L11f: Please clarify the units of the fluxes and of alpha. HFLUX is given in g/m/s and VFLUX in g/m2/s (Figs. 2-4). Is alpha defined as HFLUX/VFLUX, as I assume from "horizontal-to-vertical-mass-flux-ratio? Then the unit of alpha is m or cm but not cm-1.

Has indeed not always been consistent. Thanks for pointing that out. Changed accordingly.

- Section 3.4: The first paragraph confuses me (reference to Sect. 4.2 should be 4.3, I guess, and 4.2 is the "second step"). It should be skipped here as it is repeated in the beginning of section 4 where it is placed better.

Amended.

- P5753, L10: Just remind the reader what is done in these experiments: "...and 5a, i.e., all correction schemes are switched on, using the schemes MB95, SH04 and AF01.". Such a reminder could be included again later, e.g. on P5757, L5f.

Added where possible.

- P5753, L24f: According to Table 2, (b) and (c) are mixed up here. - P5754, L7: Which "box model components"?

Amended and more detail added.

- P5755, L6ff: The "peak shear velocities" are the same in Figs. 2-4 as this are the measured ones. What is meant here is the temporal agreement between observed and simulated fluxes, right? Otherwise, "particularly for MB95" would make no sense.

This part was confusing indeed. Reformulated to make clear what we intend to show. In fact, we wanted to highlight the agreement between observed shear stress (u*) and resulting model emission flux when the model was driven with observed wind data.

- P5755, L23: I would say even 4 orders of magnitude (10E3 against 10E7).

Agreed! Amended.

- P5755, L24ff: I disagree. There are black dots (=observed vertical fluxes) in Fig. 2f and 2j. My conclusion would be that the soil was not too wet for dust emission (observed HFLUX and VFLUX > 0) but that the moisture threshold in the model was exceeded, inhibiting dust emission. Please clarify.

We agree! The text has been changed to reflect your objection.

- P5756, L4. "Fig. 5" needs to be Fig. 3!

Amended.

- P5756, L6f: I do not see any drop in soil moisture in Fig. 3b. From the caption I assume that the soil moisture is above 0.05 kg kg-1 when no data are shown. Is this true? Please be more precise in the caption.

There was a slow decline in soil moisture at site B3 (figure captions complemented). No soil moisture data mean that we do not have observational data indeed for that period of time (DOY 272 at site B3).

- P5756, L7-14: This is over-interpreted in my opinion. Both schemes produce extremely unrealistic horizontal fluxes (at most sites very much too often and too strong). I think the only conclusion from these figures can be that sometimes the one scheme performs better (MB95 at site I04) and sometimes the other (SH04 at site J11). I do not see the "advantage" of SH04 for site D10: The VFLUX is almost the same as with MB95 but the The VFLUX is almost the same as with MB95 but the HFLUX is even worse than with MB95. The sentence in line 12 could be skipped as none of the schemes produces anything at L05. The question arises if Fig. 3 could be reduced to show only the panels for sites I04 and J11.

We agree that there was some over-interpretation of the results on our side, with a few conclusive statements that are not entirely supported by the data. The paragraph has been reformulated in order to avoid confusion and to increase the robustness of our statements.

- P5756, L15f: I do not understand why there are differences in the left panels of Fig. 3 and Fig 4. Reading Table 2, I would expect the same HFLUX for experiments 4 and 5. The left panels of Fig 4. could be skipped then. Please clarify.

In fact, the size distribution in the box model is different for exps. 4 and 5. Exp 4 uses 4 parent soil bins while exp 5 uses a sub-bin size distribution with more than 80 size bins. Hence minor differences can occur, albeit unwanted. Fortunately, the difference is indeed only minor. We have highlighted this issue in the text (lines 648ff) which should clarify matters.

- P5756, L17f: Point (1) kind of disagrees with the sentence in line 10f "Modelled emission frequency...".

Reformulated.

- P5756, L21f: I disagree that the "opposite is true..." which would mean that simulated fluxes are lower than the observed ones. At the few days with observed VFLUX at sites B03, I04 and J11, the simulated VFLUX fits quite well. But still there are many many days with no VFLUX observed but simulated, meaning an overestimation. Please correct this.

Agreed! Corrected due to slightly ambiguous interpretation of the data on our part.

- Section 4.2 can be drastically shortened in my opinion. The conclusion from these two pages of text is that the threshold shear velocity strongly depends on the moisture and roughness.

Shortened and reformulated in order to be much more concise.

- P5757, L16: Please mention that "emission fluxes" always refers to the vertical fluxes. This was sometimes confusing me when reading the paper for the first time.

Amended where possible.

- P5757, L18ff: How can one conclude on "soil and surface features" from this figures?

Paragraph deleted.

- P5757, L28: "sand transport models"

Amended.

- P5758, L4: Observed values are not limited, better: "u never exceeds 085...".*

Amended.

- P5759, L28: How can I see this from Fig. 6?

Paragraph is completely reformulated (lines 757-772**).

- P5760. L6: One can hardly read values of alpha from Fig. 6. See my recommendation for Fig. 6 below.

While it is hard to translate the plotted results into a corresponding number, we tend to think that the most intuitive way of presenting the data is to put them in a flux-flux diagram. We did provide the associated range of alpha values in order to simplify the interpretation (see also below).

- P5760, L14: What is meant with "direct entrainment"? From the surroundings? Would this be included when running a full 3D model allowing for dust transport from surrounding grid boxes?

Direct entrainment refers to vertical emission flux without associated saltation flux. It's clarified in the text.

- P5761, L10: Skip "experiment" after "(JADE)".

Amended.

- P5762, L2f: The values for w in Fig. 7 are 1, 5 and 10%. Why not 1, 8, 16 to represent the observed range better?

Changed to 6 (or <6% for what it's worth), 11 and 16% to represent the observed range better.

- P5762, L18f: I would say about 50% of the red dots have a gray circle (w>6%) and values of u<0.4 m/s. This is not "occasionally".

Agreed! Reformulated or taken out, respectively.

*- P5764, L3: How can the observed u*thr be read from Fig. 8? This question is more general and can be asked for all figures 5, 7 and 8. Please give an explanation.*

We didn't measure dust fluxes below a certain shear speed threshold (e.g. site I4 = red dots in Fig 8) which, presumably, indicates a lower emission boundary which we refer to as observed u*thr. Arguably, there are low measured fluxes at low shear velocities, but they are potentially unreliable (see next point).

- P5764, L13: Should it be "fluxes < 0.001 mg m-2 s-1"? But why are these observations "questionable"?

As outlined in the manuscript (more details added now), the small sample size and the large measurement uncertainties suggest artefactual behavior. While we can't rule out that it is wind induced vertical flux which becomes suspended for a longer period of time, it is more likely an artefact of local dust devils of which we have observed quite a few (picking up considerable amount of dust while producing only extremely localized wind gusts which we are unlikely to catch). See also lines 873ff **).

- P5766, L12f: The conclusion could be that it might be a worthwhile effort to incorporate a sub-grid scale emission scheme in climate or NWP models.

We tend to think that a stochastic emission scheme could be useful in this regard. Also, we might want to solve the crust problem first as it relates to the emission scheme's saltation flux formulation which would inevitably be part of any sub-grid scheme for the time being. However, we agree that it is desirable to have a sub-grid scale emission scheme in place. We will discuss the sub-grid variability in an upcoming paper in more detail.

Figures:
- Fig. 1: What is the difference between the green and the pink region? Both are labeled the same in the caption, so they should have the same color.

Figure 1 has been completely re-done, including figure caption which now contains all the relevant information.

- Figs. 2-4: Please mention in the caption that u and w refer to the right ordinates in the respective panels.*

Amended.

- Fig. 6: It would be better to plot alpha directly (against HFLUX). Then one could actually read the values of alpha from the figure. In the present form one needs to read VFLUX and HFLUX of a specific dot and divide them to get alpha. Why is the time color-coded here; is it important? I recommend to plot all sites in one panel with different colors for single sites. Values for Exp 1a might be skipped as alpha is constant.

As mentioned before, we argue that it is more intuitive to plot the data in a flux-flux diagram and provide the associated alpha values to facilitate the interpretation. However, we are willing to reconsider our decision if we are presented with a more persuasive argument, which is not to say that we completely disagree with the referee on the matter. The time is color-coded to give a flavor of how the change in observed moisture and roughness over time influences the modelled (and the observed) dust fluxes. Since we regard it as another interesting piece of information for the reader, we would prefer to leave it as is. If it weren't for the temporal change, we agree that everything could be put into one single plot.

- Fig. 7: "vsmc" never appears in the text and should be changed to "w" in the legend. There is no legend entry for the dark red dots. The caption says "black and dark grey open circles". What is the difference between them? I cannot distinguish between them in the figure.

It is changed to "w" for the sake of consistency. Also the legend now contains all the 4 categories assigned to observational data. Note that the value range has changed according to your earlier suggestion (now in line with the observed range). Values below w=6% are bundled together due to the fact that the box model moisture correction scheme became insensitive to values of w < 6%.

- Fig. 8: The differentiation for the soil types is not necessary here as they are not discussed in the text. Furthermore, I cannot see any thin grey lines in the figure.

See reasoning regarding plotting of sub-soil types at the beginning. Caption has been changed referring to the dashed lines now.

** The page numbers refer to the revised manuscript which I uploaded as supplement.