Atmos. Chem. Phys. Discuss., 15, C13115–C13123, 2016
www.atmos-chem-phys-discuss.net/15/C13115/2016/

Atmospheric Chemistry and Physics
Discussions

Interactive comment on "Forecasting the North African dust outbreak towards Europe in April 2011: a model intercomparison" *by* N. Huneeus et al.

N. Huneeus et al.

nhuneeus@dgf.uchile.cl

Received and published: 11 March 2016

We very thankful for the various comments and suggestions that contributed to improve the quality of the paper. References to pages and lines correspond to the ACPD manuscript.

R: While the reason for choosing the 5 selected models remains unclear, it is certainly a worthwhile exercise in light of the absence of such studies. Sure, it would be desirable to have a full-blown model inter-comparison exercise with all state-of-the-art models available; the paper provides a useful framework for future such studies. Ideally, the authors can provide the community with stringent guidelines as to how

a quasi-operational model validation exercise should look like. For example, given that there already exists an operational forecast evaluation project within the SDS-WAS framework (http://sdswas.aemet.es/forecast-products/forecast-evaluation/model-evaluation-metrics), it seems fairly straight-forward to extend this effort beyond the current setup (perhaps introducing sub-regions to facilitate dust event evaluation). Binietoglou et al 2015 could be added in this context as well.

A: The conclusions section was modified and specific aspects future study should focus on were highlighted. Suggesting stringent guidelines for model validation is difficult since it depends crucially on a dense network of quality-controlled observations over remote desert regions. Where such data are available, techniques from the present work should be replicated, particularly in other dust source regions and for other dust events. In any case, a serious limitation in exhaustive validation of dust models is the availability of observations other than AOD. Surface concentration used in this study are not available on a routine basis, they need to be derived from PM10 measurements first. Further, only a limited number of meteorological sounding stations exist in northern Africa, and most of them (except one) are on the borders of the dust source regions. Also, only a few lidar instruments are available in northern Africa and again mostly on the borders of the source regions. While such observations are available over Europe, they are urgently needed over northern Africa given the many processes involved.

R: Equally desirable, yet beyond the scope of this study, would be an extension of this validation exercise to different types of dust events. In particular, it would be interesting to see whether there are systematic forecast model biases with regard to the breakdown of the low-level jet or is the forecast skill sufficient to predict convectively triggered haboobs with some lead time. Admittedly, the latter depends on the model resolution and might not work with the selected set of models (or at least not at the chosen horizontal model resolution) to start with, but it would be worth putting such suggestions for followup work in the discussion/conclusion section. Also, a method to quantify the impact of imprecise forecast of synoptic conditions upon the dust emission flux would help to detect the key aspects of future work. Based on my own work with the HadGEM3 model at 12x12km grid size, the surface winds are very well reproduced (compared with direct observations at 10m height) even when allowing for considerable lead-time (unsurprisingly, the MetUM used in this study shows similarly good results for all lead times). This suggests that future work should focus on improving the emission schemes, which is something I wish the authors of this paper could confirm.

A: We thank the reviewer for this comment and we fully agree that this validation exercise should be extended to different types of dust events, but also to different dust source regions like mentioned in the previous statement. In fact the SDS-WAS NAMEE is starting a project aiming to assess the model performance to predict an intense dust event in Iran (haboob). We have modified the conclusion sections and added a new paragraph stating the needs of future studies aiming to evaluate the performance of dust models.

Specific Comments:

R: p.26666, lines 4/5: A short justification or explanation why those 5 (and only those 5) models have been chosen for the analysis would be desirable.

A: The reasons are of a practical nature and not scientific one. At the beginning of the intercomparison project, an invitation was send to a large number of modeling groups (>5) and all those that responded to our invitation by simulating the period of interest and submitting their model outputs are included in the paper. We believe this not to be relevant for the understanding of the paper and results of the study and therefore choose not to include an explanation in the manuscript.

R: p.26667, lines 5/6: The orange dots in Fig 1 are really hard to identify. I suggest to put all station information in a separate plot in order to facilitate identification.

A: Changed as suggested.

R: p.26667, lines 17-26: MODIS AOD is also biased towards the time of satellite passage. Do you account for this potential source of error when you validate the model results? If not, how much of an on the results of the analysis could it have? Appropriate reference needed.

A: To explore the impact of temporal sampling on our comparison we have computed the AOD as the average of the fields at 12 and 15 UTC. This should reduce potential biases due to temporal sampling. This information was added to the manuscript (and the figure caption) to clarify it to the reader. We found that the bias is even larger when the average of 12 and 15 UTC is used. In general the main features of the spatial distribution of AOD is the same, it is the magnitude, in particular in region with maximum AOD, which was reduced when the daily average is considered. We note however that we conduct a qualitative analysis against MODIS so that we had chosen daily means initially. In order to minimize sampling errors further, we have now replaced the data with the more precise temporal matching; the impact on our results is relatively small and the key findings are the same.

R: p.26668, section 2.3: I would suggest to introduce the MERRA reanalysis here as well (as you are using its wind data). Could be put into the model section as well. A short paragraph of known issues with reanalysis data in general and MERRA in particular should also be added. NCEP as well as ERA40/Interim reanalysis considerably overestimate nighttime wind speeds and underestimate higher wind speeds in general (e.g. Haustein et al 2012; Largeron et al. 2015; more to come soon from Engelstaedter et al. in Review).

A: Simulating the diurnal cycle of winds remains a challenge due to the necessity of physical parameterizations of sub-grid scale processes (e.g., Fiedler et al., 2013, Heinold et al., 2013). An evaluation of the climatology of near-surface winds from MERRA can be found in Fiedler et al. (2015). A general discussion of the statistical evaluation of winds from re-analysis is beyond the scope of this work, so that we chose a brief statement that points to other studies. In the revised manuscript, MERRA was introduced in the model section and a statement was added about limitations of reanalysis and reads as follows: "In addition to these five models, we use the Modern-Era Retrospective Analysis for Research and Application (MERRA) from the National Aeronautics and Space Administration (NASA; Rienecker et al., 2011) to evaluate the model performance in reproducing the synoptic-scale conditions of the event. Near-surface winds from MERRA are shown for completeness. A discussion of limitations of winds from re-analysis can be found elsewhere (e.g., Menut, 2008; Fiedler et al., 2013, 2015, Largeron et al., 2015)."

R: p.26672, section 4.1 and p.26673/74, section 4.2: What is the main reason that the MetUM overestimates the dust emission flux and the surface concentrations so consistently (a feature which is also apparent in the operational forecast)? Is the preferential source map (based on topography) switched on in their operational model setup? I recommend to add a paragraph in the discussion section that deals with this noticeable problem in this model. Ideally, it can be established what the likely cause for the overestimation is (e.g. strong tuning due to poor parameterisation of deposition). I note that the emission/deposition ratio is briefly mentioned at p.26680, lines 20-24. Perhaps this is where the discussion fits best.

A: The operational NWP dust configuration of the MetUM uses a simplified two-bin scheme for dust emission. This might over-simplify the complex nature of the dust-emission size-distribution but is necessary in an operational high-resolution global model. To better understand the discrepancies between AOD and emission flux between the different models, an insight into the different dust size distributions would be needed. AOD is the key parameter by which the model is operationally evaluated and the global emission flux is tuned to give a good evaluation in AOD. This is an interesting outcome of the paper and highlights that other dust variables (such as surface concentration) should be evaluated routinely against observations. Please refer to lines 11-15 of page 26682.

R: p.26673, line 23: NNMB âĂŤ> NMMB

A: Changed as suggested.

R: p.26674, lines 5/6: Again, it would help to have a short discussion of the potential causes for the large range of model outcomes wrt emission flux in the corresponding section.

A: We have added the following paragraph in the discussion section:" A difference in emission of the order of a factor of ten is observed between the models (Fig. 6). The individual reasons for the model differences are unknown, but potential sources for differences are discussed in the following. One potential reason for different emission, are the model-dependent emission parameterizations with different particle size distributions. ECMWF/MACC has a size distribution with particles of up to 20 μm in diameter whereas the other four models have maximum sizes of 10 μm (Table 1). However, ECMWF/MACC has the smallest emission. Even for the three models with the same number of bins and the same size distribution (NNMB/BSC-Dust, BSC-DREAM8b and DREAM8-NMME) large emission differences exist pointing to the importance of other aspects. Furthermore, previous studies have shown that dust-emitting winds differ amongst models and can be attributed to the representation of atmospheric processes (e.g., Fiedler et al., 2015). Future studies should examine the detailed differences in winds and size distribution of the emissions, including aspects of model resolution that is crucial to represent different atmospheric processes. Deposition (and its size distribution) should also be examined further in future studies given its importance in model performance to simulate dust concentration and AOD."

R: p.26676, lines 18/19: Are there any known issues with BSC-DREAM8b (e.g. with regard to the PBL or soil moisture scheme) that could be causing such discrepancies? Could be revisited in the discussion section.

A: The BSC-DREAM8b model (which includes the regional hydrostatic model, ETA) uses a step-like representation of mountains in the z-vertical coordinate. The rest of the participating models (NMMB/BSC-Dust, MetUM, MACC and DREAM8-NMME) include a sigma coordinate model. The advantage of the step-like mountains is that the coordinate surfaces are quasi-horizontal. However, the representation of the physical processes in the surface layer and the planetary boundary layer (PBL) is a problem. If one wants to represent these processes in a reasonably uniform way throughout the integration domain, including both low-lying and elevated terrain, an approximately equidistant spacing of the vertical levels is required in the lower few kilometers of the atmosphere. However, the vertical resolution needed in order to achieve this goal is still too high. This was indeed one of the major problems in the process of developing the physical package for the “Eta” model (Janjic, 2001). The hydrostatic Meso model with the step-mountains (“ETA coordinate”) produces reasonable synoptic scale meteorological guidance. The blocking by the step-mountains is able to depict reasonably well the synoptic scale flow around the obstacles. Another problem possibly related to the mountain representation is that the Eta Model using the step-mountains could produce precipitation too far down on the slopes of major orographic obstacles (Staudenmeier and Mittelstadt, 1998, Janjic 1998).

R: p.26678, lines 4/5: See earlier comment on MERRA uncertainties

A: The following sentence was added: “Largeron et al. (2015) attributed the overestimation of night-time surface winds of different reanalysis (MERRA one of them) to be linked to overestimation of the turbulent diffusion of the nocturnal dry stable surface layer. This is a common problem of state-of-the-art re-analysis products (Sandu et al., 2013) that can affect dust emission (Fiedler et al., 2013).”.

R: p. 26680, lines 18/19: Recent findings (e.g. from Allen et al. 2013; Ryder et al. 2013) suggest that larger dust particles can indeed be found in higher levels of the atmosphere, suggesting that omission of larger particles (or their treatment in terms of deposition, respectively) in models is a potential source of error.

A: We thank the reviewer for this comment. We have included the following sentence in the manuscript: “However, observations taken during the Fennec project (Washington

et al., 2013) suggest the presence of large particles in higher levels (Allen et al., 2013; Ryder et al., 2013). This could indicate potential dust deposition further away from the source as illustrated by the models and highlights the role of large particles in removal processes as a potential source of errors”.

R: p.26682, lines 23-27: Could go into the conclusions.

A: Changed as suggested.

R: On a more general note, as alluded to in my general comments already, what would be most useful for the modelling community to have is a quantification of the impact imprecise capturing of synoptic conditions in general and surface wind speeds in particular would have upon the resulting model emission flux. Or in other words, we need an assessment which tells us what spatial model resolution is required to reproduce observed wind speeds (and wind gusts) good enough to exclude it as a major source of error when it comes to testing the performance of the individual components of the dust scheme in the model. I do think this study can already provide some clues in that regard (albeit not in a strictly quantified analytical sense) which is why I would appreciate a slightly more in depth discussion of this crucial subject. If the authors don't feel comfortable to go out on a limb on that, I would recommend to put it at least as a major short term research goal in the conclusion section in order to draw the readers attention to what appears to be the most pressing issue (in my humble opinion that is).

A: Simulating winds for dust emission remains challenging due to shortcomings in atmospheric model components, such as the parameterization of convection and the planetary boundary layer (e.g., Fiedler et al., 2013, Heinold et al., 2013). It is unlikely that increasing the spatial resolution alone would solve the problems, since subgrid scale processes will still be needed for representing processes at smaller scales. Moreover, a general recommendation of a horizontal resolution would be difficult, since models might behave differently. We have changed the manuscript at several points to address present uncertainties in our understanding of dust-emitting winds and outlined in the conclusions what we believe is amongst the most pressing issues. Please also refer to our comments aloft.

R: p.26683, lines 1-6: Repetition of what has already been said in the discussion section (âĂŤ> delete)

A: Remove as suggested.

R: The conclusions are generally a bit too repetitive wrt the previous discussion section. While I tend to structure things the same way myself, the conclusion section should focus more on the impact/repercussions of the findings/results which have been discussed before. For example, the topic of separating meteorological/synoptic and dust cycle parameterisation related problems would fit the conclusion section perfectly. This goes along with an outlook of follow up research of this particular paper and suggestions where future research on the subject should focus on in general. Therefore I recommend to overhaul (and shorten) the conclusion section as recommended. I am convinced that it can help to wrap up this otherwise very well written and well thought-out paper in a neat and concise fashion.

A: We are thankful for this comment/suggestion. We have shortened the conclusions and included a paragraph with suggestions for future studies addressing the performance of dust models.

Interactive comment on Atmos. Chem. Phys. Discuss., 15, 26661, 2015.