Response to reviewers of: “Assimilation of MODIS Dark Target and Deep Blue observations in the dust aerosol component of NMMB/BSC-CTM version 1.0”

Enza Di Tomaso, Nick A. J. Schutgens, Oriol Jorba, and Carlos Pérez García-Pando

General Response

We wish to thank the reviewers for their interest in our paper, for their constructive comments and useful suggestions that lead to an improved manuscript.

We first note that our model has been recently renamed, after the reviewing process of another GMDD paper on the development of the model gas-phase chemistry component (doi:10.5194/gmd-2016-141). We have substituted in the revised version of the manuscript (also in the title) the name NMMB/BSC-CTM with the new name NMMB-MONARCH, where MONARCH stands for "Multiscale Online Nonhydrostatic AtmospheRe CHemistry model".

In the following, we report our answers to general and specific comments of the reviewers. Reviewers’ questions and comments are shown in bold-italic, our answers appear in standard type.

Specific Response (Anonymous Reviewer #1)

Answer to general comment:

1. Overall, I think the figures need to be reworked. They are quite small and the font on the labels is too small to read on most of them, making it harder to evaluate the results.

We have reworked all the figures following the reviewer's suggestion.

Answer to specific comments:

1. Page 4, Line 98-100, what about the UK Met Office? They assimilate dust AOD in their unified model.

We say at page 18, lines 599-600, of the discussion paper that the UK Met Office assimilates MODIS Deep Blue over desert. Following your comment, we have now added it also in the introduction of the revised manuscript at page 3, lines 93-94.

2. Page 4, Lines 107-109, limited spatial correlations have been shown in some studies, depending on what they are sampling, but do you suspect this would be much longer for big dust transport events, especially coming off of the Sahara over the Atlantic ocean? This looks like it would be the case based on MODIS observations.

Correlations in dust AOD are automatically set by the ensemble. However, we limit their use because we use a small patch size. The reviewer has a good point. Spatial correlations in long-range dust transport (in particular for dust fine mode) deserve further investigation. We have added a comment about this in the introduction of the revised manuscript at page 4, lines 99-102.

3. Page 5, I think it would be helpful to include units on the variables in the vertical dust mass flux equation. Also, what value do you use for C? This is constant globally? Is the source mode coefficient how you distribute the mass among the size bins? What threshold do you typically use for the friction velocity (when not perturbed for the ensembles)?

The units for the vertical dust mass flux equation have been added at page 5, lines 145-151 of the revised manuscript. The variables for which we have not specified the units are unit-less. We have added also the value used for the constant *C* (0.768) at line 154, which is an updated value for the global constant estimated in paragraph 4.2.1 of Pérez et al. (2011). *C* is a tuning factor for the total vertical flux. As we say in the discussion paper (page 5, lines 146-148), the total vertical flux mass is distributed among the dust transport bins according to a specific dust distribution at sources derived from D'Almeida (1987) which assumes that the vertical dust flux is size distributed according to three lognormal background source modes. In few words, the mass is distributed at sources among the 8 transport bins according to the coefficients represented with the red thick (red dashed, in the revised version) line in Figure 1. The threshold on the friction velocity is not fixed, but dynamically estimated as function of different soil characteristics, among which soil water content (please see equation 3 to 5 in Pérez et al., 2011). We have now specified also the latter in the revised manuscript at page 5, lines 135-136.

4. Page 5, Line 165, what were the main sources of uncertainty identified in the evaluation efforts?

Emissions are identified as one of the sources of uncertainty. Quoting the referred evaluation studies, for example it has been detected an overestimation of the Bodélé emissions and the under estimation of the Mali/Mauritania border emissions. There are however evaluation sites, like Solar Village, where it has not been determined if errors in AOD are due to an inaccurate source prescription or to the inability of the model to reproduce the associated meteorology, or, sites like Izaña, where instead errors in AOD could be due to the deficiency in the model representation of the steep orography. Also, the model tends to overestimate the very low background concentrations far away from sources, which hints towards an overestimation of the smallest dust particles due to either inaccuracies in the size distribution of the emissions, vertical transport and/or removal.

5. Page 6, to clarify your data assimilation approach, you might want to mention some specifics about the 4-dimensional extension of the LETKF and why you chose to use the extension since you are assimilating observations regularly over 6 hour intervals with the NRL MODIS product. Do you expect to incorporate observations in the future that are asynchronous? Did you test at all the performance of the LETKF versus LETKF with the extension? This would be interesting.

We have not made the test without the 4D extension. It could be interesting to perform it, however this would require resources (and more I/O time). We are indeed assimilating asynchronous observations with some degree of approximation: the analysis is calculated once a day with observation slices of 6 hours; simulated observations and background departures are calculated at each time slot (every 6 hours) using the background ensemble for that time. As explained at page 11, lines 350-355, of the discussion paper we concatenate observation vectors and background matrices and use the same analysis equation used for synchronous observations. In this, we are following the 4-dimensional extension of the LETKF described in Section 4 of Hunt et al. (2007). What we have not done, which, quoting Hunt could be of advantage, is to take into account the timing of the observations when deciding which of them to use in a given local analysis.

6. Page 7, Line 210-214, what units do you use to define the distance in the localization and what localization factor do you use? It's hard to tell from this how much localization is used.

The distance in the localisation function is calculated in the grid space. We have specified this now in the revised manuscript at page 7, line 209. As we wrote at page 11, line 362, of the discussion paper in the description of the experiment setup, we have set the horizontal localization factor to the value 1, which means that after 2 grid points this function is very close to zero.

7. Page 7, Ensemble perturbations in the vertical flux, You are perturbing the distribution of dust emissions among the size bins, but the total mass flux is held fixed? Are the perturbations that you show in Figure 1 the same for all locations or does this vary by grid or region? It might be good to change the solid red line with the error bars in Figure 1 to make it easier to distinguish from the ensemble perturbed lines. Maybe to a dotted or dashed line? Also, it might be useful to show somewhere what sizes the bins correspond to.

Yes, we are perturbing the distribution of dust emission among the size bins, but the total mass flux is not held constant. As we wrote at page 7, line 223-225 of the discussion paper, the model ensemble is created perturbing the vertical flux of dust in each of the eight dust bins. This is equivalent to perturbing the total vertical flux as well as its size distribution at sources.

The source perturbations are constant in time and space as we wrote at page 8, lines 238-239 of the discussion paper.

We have modified the style of the solid red line in Figure 1 (a dashed line in the revised version) to make it more visible, as the reviewer suggested.

We have added the information about bin sizes in the ensemble perturbation section at page 7, lines 220-221 of the revised manuscript and in the caption of Figure 1.

8. Page 7, Ensemble perturbations in the threshold friction velocity perturbation. Again, do the random perturbations vary with location or are the same perturbations applied everywhere? This matters as it will determine your covariances and how an observation spatially impacts your model state.

Also these perturbations are spatial and temporally constant. We have now specified at page 7, line 236, of the revised manuscript that what we wrote at page 8, lines 238-239, of the discussion paper refers to both types of source perturbations.

9. Page 8, Lines 237-239, if the structure of your source perturbations is temporally and spatially constant, you are essentially specifying your background covariances, much in the way a variational approach operates. As you mention, this is the first stage of development, so I think that's a reasonable first means for generating the ensemble and will probably help you do well near source regions, but you may have problems for transport events.

The spin-up period for the ensemble ensures that perturbations applied at the sources propagate everywhere and dynamically create covariance structures due to the different size distribution, emissions, but also due to observation localization and limited patch size for the local analysis. The background covariances therefore are not constant. However, implementing spatially (and temporally) varying perturbations should be tested in the future in case it can better represents model uncertainty.

10. Page 8-9. MODIS Dark Target, I would increase the size of Figure 2, it's too small to see. It would probably also be useful to see some sort of summary of the observations over the experimental time period, perhaps a data count to see where your simulations are being constrained or a mean of your observations. Also, I'm concerned about using over-land AE as a filter for dust. It's been shown that this product is pretty binary (see Levy et al. 2010) and more problematic for coarse mode aerosol than fine. Have you checked to make sure you aren't getting other aerosol in there, like biomass burning aerosol? Perhaps this could be contributing to some of the bias that you are seeing.

We have increased the size and resolution of Figure 2.

We have added a plot of observation counts over the experiment period (Figure 4 of the revised manuscript).

AE over land has indeed considerable uncertain. To overcome this shortcoming in our dust filter we have added a quality control on the assimilated observations based on normalized first-guess departures that rejects observations that are too far from the background (as we say at page 9, lines 304-306 of the discussion paper). We are aware that this is only a temporary solution until we will run the assimilation with a complete aerosol model.

11. Page 10, Numerical Setup. The Control is the exact same model as the ensemble free run, the only difference in the ensemble free run is you have perturbed dust emissions (either in the distribution in the bins or threshold friction velocity) and the control is a single run?

Yes, that is correct.

12. Page 11, Lines 358-360. I suspect your insensitivity to ensemble size is a result of how you are generating the ensembles themselves (you are sampling from a specified distribution) and also maybe you are heavily localizing (can't tell without units though). This will likely change as you add other perturbations to your system and you may find that you need a much larger ensemble as 12-24 members is quite small.

This hypothesis will have to be tested. There is evidence that the number of ensemble members (if not too small) does not matter too much as long as the model is kept close to a reanalysis.

13. Page 11, Lines 365-367. Did including vertical localization make much of a difference for AOD assimilation? I'm not sure if you tested it without, but I would think that this wouldn't have much impact for a column-integrated observation.

Yes, vertical localization does not have much impact without any vertical observational constrain. Since our vertical localization is using the background sensitivity in the vertical, it is equivalent to distribute the mass increments according to the model background vertical profile.

14. Page 12, your use of error as observation minus model is a bit confusing to me. The bias for example would have a negative value when the model is biased high. Typically, you would use your estimator (model) minus the expected value (observation). I would suggest flipping this so that your bias maps in Figure 11 and stat tables/bar graphs don't confuse the reader into thinking the model is biased low when the opposite is true.

We have changed our convention for the model field error which was defined at page 12, line 407, of the discussion paper, now defined with an opposite sign in the revised manuscript at page 13, line 409. We have changed accordingly the sign of the bias in Table 3, Figure 12 to 16 of the revised manuscript.

15. Page 13, Section 7.1 I think it would be beneficial in Figure 5 to also show the difference between the DA experiments and your ensemble free run (or control). The difference between the DT+DB simulation and the free run is pretty clear, but harder to see with the DT run. Also, I assume this is dust AOD only? If so, you should probably put that in the Figure caption and mention that in the text as well (Page 13, line 444). Are these differences persistent over the entire simulation since you only show one month?

We have added the difference plots between the DA experiments and the ensemble free run (bottom panels of Figure 6 of the revised manuscript).

Yes, Figure 5 refers to dust AOD only, we have specified this in the text (page 14, line 448) and figure caption (Figure 6) of the revised manuscript.

The difference between the experiments vary during the different months according to differences in the dust emissions and transport over time, but the conclusions stay valid.

16. Figure 6, Does the DT simulation's coefficient of variation look similar to the DT+DB? If so, you might want to mention that in the text. If they are different, you should probably show both. Also, does the mean AOD change much with the different perturbation schemes (Figure 6 and 7)?

The DT simulation's coefficient of variation shows higher values in the Northern Hemisphere compared to the DT+DB simulation. These differences are due to less observational constraint over land when DB is not used. We have added the plot (central panel of Figure 7 of the revised manuscript), as suggested, and comments in the manuscript at page 14, line 457 of the revised manuscript.

17. In Figure 6, I'm surprised that you have considerable spread in places that I wouldn't expect, like near the poles in the Southern hemisphere. Are the ensemble members being inflated as part of the data assimilation?

No, we have not used inflation. Please note that the plot shows a normalized spread. A considerable normalized spread is expected in the Southern Hemisphere (SH): there the values of dust AOD are quite small, while the ensemble members show differences among them due the perturbation of emissions in the SH sources in South America, Africa and Australia. Prompt by this question, we have double checked how the ensemble spread evolves from zero on the first day of the spin-up, to values greater than zero only close to the sources in the first days of the spin-up, to finally propagate everywhere in the SH by the end of the spin-up period. We have also added a comment in the text about this at page 14, lines 458-461 since it is a point that raised questions by both reviewers.

18. Page 14, Lines 460-462. This sentence implies the more spread the better since you'll just push towards the observations. However, your goal is to really have sufficient spread that represents the uncertainty in the system. Have you tried to determine whether or not the spread that you are generating is representative of the uncertainty?

We have modified the sentence since, as the reviewer correctly pointed out, it implied that the more spread the better, adding the text at page 14, lines 469-477 of the revised manuscript. We have calculated the ratio between prior total spread and RMSE and found that our ensemble configuration is under representing uncertainty. As stated in other part of the manuscript, other perturbations should be tested in the future.

19. Figure 8, I would remove the color bars here for each subplot to save space and increase the individual plot size and font size (same for all the figures). I also wonder if you increase the number of bins in your scatterplot, whether the asymmetry that you talk about would be more apparent.

We have used one colour bar for all the sub-plots, increased font size, figure resolution, and also the number of bins in the scatter plots in Figure 9 of the revised manuscript.

20. Figure 9, The analysis increments that you are showing are in dust AOD? If so, you should add that to the figure caption or labels.

Yes, the analysis increments are in dust AOD. We have specified it now in the figure caption (Figure 10 of the revised manuscript) and in the text at page 15, line 496.

21. Page 15, Section 7.2 For AERONET sites in transport regions, such as La Parguera, it looks like the dust AOD has decreased with data assimilation compared to the control. However, the analysis increments shown in Figure 9 show an increase in AOD. Perhaps the prior state has decreased so much with the near-source corrections that the increase observed over the oceans still produces an AOD at sites impacted by transport that is still less than the control? I'm curious what you found with that.

Yes, it is as the reviewer writes: as a result of the near-source corrections, the overall AOD in La Parguera is less than the control. The analysis increments show local changes while the AOD found in La Parguera is a result of both the local analysis corrections plus the mass transported from Africa which is affected by other local analysis corrections (reduction of mass over Sahara).

22. Figure 11, I would put one colorbar at the bottom of each column of figures then maybe add one label at the top of each column (Control, DA-NRL, DA-NRL-DB) and add one label on the y-axis for each row (Bias, RMSE, Corr, FRGE). That way you can increase the size of each map and make the labels larger. Also, it's so small that it is impossible to see any difference in the circle sizes and there is no reference to use to determine what number of samples the circle size corresponds to.

We have used one colour bar for each sub-plot row (as columns represent different statistics with different colour bars), and added one label for each column (experiment) and row (statistics) in Figure 12 of the revised manuscript. We have also added one row for the number of observation samples per stations to have a clearer reference, increased font size and resolution.

23. Figure 14 and 15 need to be fixed, the labels are way too small to be able to read. It makes it hard to evaluate your forecast results.

We have increased font size, figure resolution and, to save space, we have removed the bars relative to the standard deviation (SD), in Figure 15 and 16 of the revised manuscript, as it can be derived from the bias and RMSE statistics.

24. I wonder if you might want to show in your statistics bar graphs the average dust AOD as well to give some context to how large the errors really are and maybe considering adding error bars (maybe through bootstrapping) to your statistics to test if the differences are statistically significant.

We have added in the plots of the bar graphs in Figure 13 to 16 of the revised manuscript the average value of AOD for the observations used for validation (indicating the number in one of the upper corner of the plot) and specified this in the figure captions. We have also specified in the text at page 17, lines 554-557 and lines 578-579, whether the results for the correlation are statistically significant.

Comments on technical corrections:

1. Page 3, Line 67, change "to different model inter-comparison" to "in different model inter-comparison"

We have changed it accordingly, thanks.

2. Page 3, Line 73, saying the community resorted to data assimilation makes it sound kind of negative. Maybe you could say something like...because of these large uncertainties, the atmospheric composition community has begun to make use of data assimilation for better characterizing and predicting.…

We have changed it accordingly, thanks.

3. Page 3, Line 79, you might want to cite the Sessions et al. 2015 paper after the sentence where you mention that assimilation of aerosol observations is now operational at many forecasting centers.

We have added it. Thanks for the suggestion.

4. Page 10, Line 325. You should probably cite the AERONET uncertainty

We have added it. Thanks for the suggestion.

5. As a suggestion on your equations, you may want to go through and make sure the variables are consistent across equations. For example, in equation 4 the size bins 1 through 8 are indicated with a b while in equation 1 they are indicated with a

k. Later k refers to ensemble members. This might confuse the reader. Also, it would be useful to include units with your variables.

We have changed the variable for the size bins in equation 1 of the revised manuscript to have consistency and avoid confusion with the letter used for the ensemble members. We have added the units also for variables used for the AOD operator, unless they are unit-less, at page 8, line 242, of the revised manuscript.

6. Page 13, regions for validation (Lines 430-440). I think in Figure 4 it would be good to list the regions associated with each box. You can probably just put this in the figure caption and say which color box goes with which region, to tie the map to Table 2.

We have added it as suggested, thanks.

7. For Figures 5,6,7, the colorbars are the same on the different subplots within each figure, so I would only show the colorbar once to save space and make the maps larger. They are too small to see clearly.

We have used one colour bar for the plots sharing the same one in Figure 6, 7, 8 of the revised manuscript and made the maps larger and at a higher resolution.

8. In the caption for Figure 10, you should mention that this is the analysis AOD and not the prior.

We have added it as suggested, thanks.

9. Page 17, Line 567-568. This sentence isn't very clear. You are referring to the Sahara? Better temporal evolution, reflected by the increase in correlation with AERONET over time?

We have rewritten the sentence at page 18, lines 593-595. We were referring to the SubSahel and ShortAtl regions where the correlation degrades after day 1.

Specific Response (Anonymous Reviewer #2)

Answer to general comments:

1. The abstract is quite prolixity. Abstracts should include only important information.

We have reduced the abstract as suggested, thanks.

2. I cannot read some figures due to poor resolution and small labels. The authors should re-draw the figures.

We have reworked all the figures following the reviewer' s suggestion making them bigger, bigger fonts and at a higher resolution.

Answers to specific comments:

1. Page 6, line 187: The authors use 100 km as the cut-off (localization) length. How

do you estimate this values? For example, Rubin et al. (2016) and Yumimoto and Takemura (2011) used more longer length (1000 and about 600 km).

Rubin et al., Atmos. Chem. Phys., 16, 3927-3951, 2016, doi:10.5194/acp-16-3927-2016
Yumimoto and Takemura, Geophys Res. Lett., 38, L21802, doi:10.1029/2011GL049258

Please see our answer to the specific comment n.6 of reviewer#1. Our cut-off length is hence longer than 100km and in the range of the values used in the studies mentioned by the reviewer. We have added the references suggested to put it in the context of other studies.

2. Page 7, line 212:
"h" is already used in line 209. Use another character to represent horizontal localization factor.

We have changed the letter for the horizontal localization factor at page 7, lines 208 and 210 to avoid confusion. Thanks for spotting it out.

3. Section 3.0:
Ensemble-based methods usually use inflation methods. Does this system use any inflation method?

We have not used inflation in the experiments described in the manuscript. We take the reviewer's question as a suggestion for our future tests when with other perturbations, in case the ensemble spread is not representing well enough model uncertainty.

4. Figure 1:
Can you add ensemble mean of the vertical flux in the figure?

We have added a line (dash-dotted) in Figure 1 for the mean of the ensemble perturbations.

5. Page 8, line 237-239:
You use AOT (optical column amount) as the observational constraint. How does the system adjust 3D mass concentration fields of dust bins from the 2D observational constraint?

This is explained at page 11, lines 365-367 of the discussion paper.

6. Page 9, line 274:
Do you consider error in AE? AE over the land may have much large uncertainty than ocean. Can you separate the dust-dominant condition correctly over the land?

We appreciate this concern and in fact use a quality control on the observations. Please see our answer to specific comment n. 10 of reviewer #1.

7. Page 9, line 276:
Coverage and observation time of MODIS do not correspond to those of OMI (particularly for AOTs from Terra satellite). How do you derive the AOTs under dust-dominant condition when there is no OMI observation corresponding to? You do

not use MODIS measurements from Terra satellite?

We use both Terra and Aqua, and, as we wrote at page 8, section 4.1, of the discussion paper, we have used only Level 3, daily, products. When there is no OMI observation, data are not selected for assimilation.

8. Page 11, Line 344:
The authors extend the system to 4D-LETKF. What are the merits of the extension instead of sequential assimilation? You assimilate AOD with 6-hour interval. I read literature suggests that 4-dimentional methods (smoothers) have advantages in assimilating observation with fine temporal resolution comparing with 3-dimentional methods (filters). However, the 6-hour interval is not so short (actually longer for 4D-LETKF). Addition to this, the main purpose of this study is improving of forecasting with assimilation. Why do you choice smoother for this purpose rather than filter? Did you try the 3D-LETKF? Did you find that the 4D-LETKF is superior to the 3D-LETKF in forecast performance?

Please see our answer to specific comment n. 10 of reviewer #1. It is true however that the literature suggests that it the 4D extension has merits for temporal resolutions finer than the resolution we have used, hence it would be worth testing in any future 3D extension.

9. Page 11, Line 344:
Do you introduce temporal localization? The assimilation window (24 hours) is too long to examine assimilation without the temporal localization.

No, we have not tested temporal localization for this system. Thank you for the suggestion. Tests by the authors with the LETKF on a different model system have shown no significant difference with a 12 hour window.

10. Page 11, line 365:
The authors use the vertical localization. What are the merit of that for assimilating vertically integrated observations?

This feature has been implemented to have a system that can handle also the assimilation of profiles, and it is not having impact with integrated observations. Please see our answer to specific comment n. 13 of reviewer #1.

11. Figure 6:
I think this figure shows ensemble spread of dust AOD. Why the spread exhibits much large value all over the Southern hemisphere?

Please note that the plot shows a normalized spread, which is expected to be considerable on the in the Southern Hemisphere. For more details, please see our answer to specific comment n. 17 of reviewer #1.

12. Figure 10:
Could you adjust the vertical axis of panels? For example, AOD values at Lecce_University are too small to plot with vertical axis of 0.0-4.0. Could you add MODIS-measured AOD on the panels? It would be good to see difference (error) in MODIS AOD.

We prefer to use the same vertical axis for the different validation sites to have the different ranges of AOD values that we are validating in the different regional domains of Figure 4 of the discussion paper, close and far from sources, visually clear.

We have added MODIS AOD from the set of assimilated observations in the time-series of Figure 11 of the revised manuscript. Note, however, that these satellite observations are not an independent reference of validation for the analysis, nor are entirely representative of the observational constraint used to calculate the analysis in the given station location, without taking into account the localisation function and observation uncertainty of all the observation in the local patch around the station location.

13. Page 14, line 456:
'Top' should be left. 'bottom' should be right.

We have change it accordingly, thanks.

14. Page 14, line 460:
The higher spread does not mean the better spread (background error covariance). If you used the larger perturbation, you'd obtain the higher spread.

We have modified the sentence following the good point that the reviewer had made. For more details, please see our answer to specific comment n. 18 of reviewer #1.

15. Figure 11:
Do you compare model result with AERONET observation in daily average? hourly average? or monthly value?

We use the closest AERONET value in a +/- 30 minute interval from the model time step, and we use only one value without doing any averaging. We have specified this in the section Methodology for the evaluation of the simulations at page 13, lines 425-426 of the revised manuscript.

16. Figure 9:
There are some regions where the DA-NRL-DB shows opposite increment from the DA-NRL. For example, the DA-NRL-DB obtains negative increment around Somalia Peninsula. However the DA-NRL shows positive one. Does this mean there is biasses between the Dark-target and the Deep Blue AODs?

Yes, as the reviewer says, this could be due to unresolved conflicting biases between the two types of retrievals. We have added this comment also in the revised manuscript at page 15, lines 503-505.

Manuscript prepared for Geosci. Model Dev.
with version 2015/09/17 7.94 Copernicus papers of the LaTeX class copernicus.cls.
Date: 10 February 2017

Assimilation of MODIS Dark Target and Deep Blue observations in the dust aerosol component of ~~NMMB/BSC-CTM~~ NMMB-MONARCH version 1.0

Enza Di Tomaso[1], Nick A. J. Schutgens[2], Oriol Jorba[1], and Carlos Pérez García-Pando[3,4]

[1]Earth Sciences Department, Barcelona Supercomputing Center, Spain
[2]Atmospheric, Oceanic and Planetary Physics, University of Oxford, United Kingdom
[3]NASA Goddard Institute for Space Studies, New York, USA
[4]Department of Applied Physics and Applied Math, Columbia University, New York, USA

Correspondence to: Enza Di Tomaso (enza.ditomaso@bsc.es)

Abstract. A data assimilation system has been developed for the chemical transport forecast model ~~NMMB/BSC-CTM~~NMMB-MONARCH, with a focus on mineral dust, a prominent type of aerosol. ~~Before this work, the system did not have an aerosol data assimilation capability and dust was produced uniquely from model estimated surface emission fluxes. As emissions are recognized as a major factor limiting the accuracy of dust modelling, remote sensing observations from satellites have been used to improve the description of the atmospheric dust load in the model.~~ An ensemble-based Kalman filter technique (namely the Local Ensemble Transform Kalman Filter - LETKF) has been utilized to optimally combine model background and satellite retrievals. Our implementation of the ensemble is based on known uncertainties in the physical parametrizations of the dust emission scheme. ~~We have considered for assimilation satellite Aerosol Optical Depth (AOD) at 550 nm retrieved from measurements of top-of-atmosphere reflectances by the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor on-board the NASA Aqua and Terra satellites, after applying a mineral dust filter. In particular we have assimilated two MODIS Level 3 AOD products: the U.S. Naval Research Laboratory (NRL) and University of North Dakota AOD, which is available over land and ocean, with the exclusion of bright reflective surfaces and is based on the MODIS Dark Target Collection 5 Level 2 product, and the MODIS Deep Blue Collection 6 AOD, which is available over land including bright arid surfaces, such as deserts. Data assimilation experiments using the LETKF scheme have been evaluated against observations from the Aerosol Robotic Network (AERONET) of ground-based stations and against MODIS satellite retrievals.~~ Experiments showed that MODIS AOD retrievals using the Dark Target algorithm can help ~~NMMB/BSC-CTM~~ NMMB-MONARCH to better characterize atmospheric dust. This is particularly true for the analysis of the dust outflow in the Sahel region and over the African Atlantic coast. The ~~additional assimilation of~~ assimilation of MODIS AOD retrievals based on the Deep Blue algorithm has a further positive impact in the analysis downwind from the strongest dust sources of Sahara and in the Arabian peninsula. An

analysis-initialized forecast performs better (lower forecast error and higher correlation with observations) than a standard forecast, with the exception of underestimating dust in the long-range Atlantic transport and degradation of the temporal evolution of dust in some regions after day 1. Particularly relevant is the improved forecast over Sahara throughout the forecast range thanks to the assimilation of Deep Blue retrievals over areas not easily covered by other observational datasets.

The present study on mineral dust is a first step towards data assimilation with a complete aerosol chemical transport model that includes multiple aerosol species.

1 Introduction

Among the different aerosol species, mineral dust is one of the main components of the atmospheric aerosol loading and is of great interest for a variety of reasons. Mineral dust plays an important role in the earth's energy balance and has a relevant impact on economical activities, on the ecosystem, on health, as well as on weather and climate (Knippertz and Stuut, 2014). The strong dust storms occurring near emission sources constitute a big hazard to air traffic and road transport as they can reduce the visibility down to few meters. However dust does not affect only local economies: because of its transport over thousands of kilometres, it has an impact from local to global scales. Dust deposition provides nutrients to continental and marine ecosystems. African dust for example has a role in fertilization of the Amazon rainforest (Yu et al., 2015), while dust deposition over oceans has implication on sea biogeochemistry as the iron contained in the dust particles is a nutrient for phytoplankton, whose photosynthetic activity in turn affects the carbon cycle (Jickels et al., 2005). Dust has health implications both close and far from sources. For example, studies have shown the usefulness of dust aerosol climatologies to predict part of the year-to-year variability of the seasonal incidence of meningitis in Niger (Pérez García-Pando et al., 2014), while particulate matter measurements taken in areas far from sources show that Saharan dust outbreaks have adverse effects of cardiovascular and respiratory conditions (Mallone et al., 2011; Morman and Plumlee, 2013; Pandolfi at al., 2014). Mineral particles perturb the earth-atmosphere's radiation budget through their interaction with the short-wave radiation, through scattering and absorption, and long-wave radiation, through absorption and re-emission. Due to this redistribution of the energy, dust aerosols can have a strong impact on atmospheric processes at short (weather) and long (climate) term periods while they can affect atmospheric circulations at large spatial scales (e.g. Asian monsoons; Lau et al. (2006)). Furthermore, this can generate feedback processes since changes in weather and climate can in turn lead to changes in the dust cycle.

Different types of ground-based (e.g., Kim et al., 2011; Pey at al., 2013) and space-borne (e.g., Kaufman et al., 2005; Luo et al., 2015) observations have been utilized to describe the variability of atmospheric dust. However, due to either insufficient spatial representativeness or accuracy, the spatio-temporal features of dust aerosols are not fully captured by the current observing system.

Neither do models accurately describe atmospheric and surface dust concentrations (Huneeus et al., 2011). High uncertainties are also in our knowledge of the optical and micro-physical properties of dust, and in our representation of its vertical structure. The latter has implication on the radiation's budget and transport. On the other hand, an accurate quantification of dust's spatial and temporal distribution is key in correctly characterizing the effect that it has on the earth's energy balance, as well as in improving the skill of forecasting its concentrations in the atmosphere as well as in forecasting the weather (Pérez García-Pando at al., 2006; Grini et al., 2006; Chaboureau et al., 2011).

Regional and global centres, predicting the most important aerosol species or dust only, participate ~~to~~ in different model inter-comparison initiatives like the Aerosol Comparisons between Observations and Models (AeroCom; Tsigaridis et al., 2007) project, the International Cooperative for Aerosol Prediction (ICAP; Sessions et al., 2015) initiative and the WMO Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS; Terradellas et al., 2015). Multi-model ensemble spreads give an indication of large uncertainties in the modelling schemes and confirm the need of a better characterization of aerosols. Relatively recently because of these large uncertainties, the atmospheric composition community has ~~resorted to~~ begun to make use of data assimilation (DA) for a better characterization and prediction of atmospheric constituents such as aerosols and trace gases (Bocquet et al., 2015). Though their dynamic is mainly driven by forcings such as emissions, recent studies showed that the usage of observations through data assimilation has improved significantly the accuracy of short-term forecast and the global monitoring of both aerosols and trace gases (Benedetti et al., 2009; Elbern and Schmidt, 2001). Since the first experiments in the early 2000s, the assimilation of aerosol observations is now operational in some of the main aerosol forecasting centres (Sessions et al., 2015) . Zhang et al. (2014) have highlighted in particular the importance of a combined assimilation of satellite products for aerosol forecast.

The Earth Sciences Department of the Barcelona Supercomputing Centre (ES-BSC) is implementing a gas-aerosol module able to predict atmospheric composition at different spatial and temporal scales within the state-of-art meteorological model NMMB (Non-hydrostatic Multi-scale Model on the B grid; Janjic and Gall, 2012). This modelling system is known as the ~~NMMB/BSC-Chemical Transport Model (NMMB/BSC-CTM~~Multiscale Online Nonhydrostatic AtmospheRe CHemistry mode (NMMB-MONARCH). We report here on the extension of ~~NMMB/BSC-CTM~~ NMMB-MONARCH with a data assimilation functionality using satellite aerosol optical depth. ~~NMMB/BSC-CTM~~ NMMB-MONARCH version 1.0, as in ~~Pérez García-Pando et al. (2011)~~ Pérez García-Pando et al. (2011, where the model was previously named NMMB considers dust only but other aerosols are being implemented (Spada et al., in prep). The focus of this work on mineral dust is justified by the operational services provided by the ~~NMMB/BSC-CTM~~NMMB-MONARCH. This model provides an operational dust forecast for the Barcelona Dust Forecast Centre under an initiative of the World Meteorological Organization. It participates in the multi-model dust ensemble of the aforementioned ICAP initiative, providing daily global dust forecast up to 120 hours. It also provides daily regional forecast up to 60 hours to the WMO SDS-WAS system. Before this work, the system did not have an aerosol data assimilation capability and dust was produced uniquely from model estimated surface emission fluxes. The present study on mineral dust is a first step towards data assimilation with a complete aerosol chemical transport model that includes multiple aerosol species (not only dust but also seasalt, sulphate and organics).

Previous studies of assimilation of dust aerosol only have been conducted for the Chinese Unified Atmospheric Chemistry Environment - Dust (CUACE/Dust) forecast system (Niu et al., 2008; Wang and Niu, 2013). These studies have used variational data assimilation techniques (3D-Var) which require, in their most practical implementation, pre-calculated and constant in time model error structures. Alternatively, ensemble-based techniques use flow-dependent model error amplitudes and structures which evolve during forecast and, in theory, should be able to capture better instabilities in the background flow (Evensen, 1994; Kalnay et al., 2007). Dust AOD is currently assimilated at the UK Met Office with a hybrid variational data assimilation technique (hybrid 4D-Var).

In this work we present the coupling of ~~NMMB/BSC-CTM~~ NMMB-MONARCH with an ensemble-based technique known as Local Ensemble Transform Kalman Filter (LETKF; Hunt et al., 2007; Miyoshi and Yamane, 2007). The LETKF scheme has shown to be particularly suitable for the assimilation of aerosol information since it has been observed by Anderson et al. (2003) , Shinozuka and Redemann (2011), and Schutgens et al. (2013) that aerosol fields have limited spatial correlations. Long-range transport of dust could be an exception to this. Since detailed studies of spatial correlation length scales for dust long-range transport are still missing in the literature, in this work we assume that what has been derived (limited spatial correlations) in general for aerosols is valid for dust. The utility of ensemble-based techniques for global aerosol simulations has been shown in previous studies (Schutgens et al., 2010a; Sekiyama et al., 2010; Rubin et al., 2016; and more recently Yumimoto et al., 2016). The main novelty in our study is the creation of the ensemble, our implementation is based on known uncertainties in the physical parametrizations of the sophisticated dust emission scheme used by the ~~NMMB/BSC-CTM~~ NMMB-MONARCH model, as well as in the use of observations particular relevant for dust applications, like MODIS Deep Blue.

The ~~NMMB/BSC-CTM~~ NMMB-MONARCH chemical transport model is described in more detail in Section 2, with emphasis on its dust module. A description of the data assimilation scheme and of the assimilated observations follows respectively in Section 3 and Section 4. We report then in Section 5 about the characteristics of the simulations that we have run, in Section 6 about the evaluation methodology that we have followed, and in Section 7 about the evaluation results of our simulation experiments. The final section concludes the paper with a summary of this development, the main results achieved, and future perspectives.

2 The ~~NMMB/BSC Chemical Transport Model~~ NMMB-MONARCH model and its mineral dust component

The ES-BSC is implementing a new gas-aerosol module within the NMMB meteorological model from the Unites States National Centers for Environmental Prediction (NCEP). The new modelling system is known as the ~~NMMB/BSC-CTM (Pérez García-Pando et al., 2011; Jorba et al., 2012; Spada et al., 2013)~~ NMMB-MONAI (Pérez García-Pando et al., 2011; Jorba et al., 2012; Spada et al., 2013; Badia et al., 2016, where it was previously named NMMB/B and is developed in collaboration with NCEP and other research institutions. The chemistry (aerosols included) and meteorology are fully on-line integrated. ~~NMMB/BSC-CTM~~ NMMB-MONARCH is able to work with a wide range of spatial scales thanks to its unified non-hydrostatic dynamical core, keeping consistent parametrizations at different spatial and temporal scales. Furthermore, the dynamical core and the coupled modules are computationally highly efficient satisfying current and projected operational requirements. The rest of this section will briefly describe some characteristics of the dust component of the ~~NMMB/BSC-CTM~~NMMB-MONARCH, with particular focus on the emission scheme.

The dust emission scheme implemented in the ~~NMMB/BSC-CTM~~ NMMB-MONARCH follows the empirical relationship of Marticorena and Bergametti (1995) and Marticorena et al. (1995) according to which the vertical dust flux is proportional to the horizontal sand flux. The horizontal to vertical flux ratio reflects the availability of dust in four soil populations (clay, silt, fine/medium sand, and coarse sand) (Tegen et al., 2002). The horizontal sand flux is modelled as the flux of the saltating particles H simulated according to White (1979) and proportional to the third power of the wind friction velocity. A soil moisture-dependent threshold on the friction velocity determines the velocity above which the soil particles begin to move in horizontal saltation flux. This threshold is dynamically estimated according to soil characteristics. Soil moisture effects are included following Fecan et al. (1999) through the soil moisture correction parameter included in the expression for the threshold on the friction velocity. A sectional approach is used for the transport of dust particles, i.e. the dust size distribution is decomposed in small size bins. More exactly, dust is modelled using eight dust size bins varying from 0.1 to 10 microns, and, within each transport bin, dust is assumed to have a time-invariant lognormal distribution (Zender et al., 2003). The total vertical flux mass is distributed among the dust transport bins according to a specific dust distribution at sources. ~~NMMB/BSC-CTM~~ NMMB-MONARCH uses a distribution over sources derived from D'Almeida (1987) which assumes that the vertical dust flux is size distributed according to three lognormal background source modes. More explicitly, the dust vertical mass flux ~~F_k~~ F_b $[kg\,s^{-1}\,m^{-2}]$ in a given transport bin ~~k~~ b at each grid cell is given by

$$F_{kb} = C\,S\,(1-V)\,\alpha\,H \sum_{i=0}^{3} m_i\,M_{i,k\,i,b} \qquad kb = 1,\ldots,8 \tag{1}$$

where S is a source erodibility factor defined on bare ground surfaces, representing the probability to have accumulated sediments in the given grid cell that are potential dust sources; $(1-V)$ is the grid's fraction of bare soil; α $[m^{-1}]$ is the horizontal to vertical flux ratio calculated for four soil populations classes (clay, silt, fine/medium sand, and coarse sand); H $[kg\,s^{-1}\,m^{-1}]$ is the horizontal sand flux; $M_{i,k}$ is the mass fraction of background source mode i carried in a transport bin k calculated following Zender et al. (2003), and weighted by specific background source mode coefficient m_i; and C is a global tuning factor empirically set to 0.768, which is meant to compensate for the uncertainty associated with the different component of F_k. More details about the above formulation of dust emission can be found in Pérez García-Pando et al. (2011).

The mineral dust module has been extensively evaluated in studies at global and regional scales (Pérez García-Pando et al., 2011; Haustein et al., 2012; Huneeus et al., 2011, 2016), showing that its evaluation scores lie in the upper range of the AEROCOM model evaluation performance scores. However, these evaluation efforts confirmed, similarly to other modelling systems, different sources of uncertainty in the ~~NMMB/BSC-CTM~~ NMMB-MONARCH dust modelling.

3 The data assimilation scheme

We have coupled the ~~NMMB/BSC-CTM~~ NMMB-MONARCH with the LETKF scheme (Hunt et al., 2007; Miyoshi and Yamane, 2007; Schutgens et al., 2010a; Schutgens et al., 2013) with four-dimensional extension as described in Hunt et al. (2007), in order to estimate optimal initial conditions for our dust model. The overall scheme implements an iterative approach consisting in a forecast step and state estimation step. The state estimation step combines information from mineral dust observations and a prior first-guess, or background from our model. A short-term forecast is used as background information. The background incorporates information from past observations, therefore the analysis is estimated using both current and past observations. LETKF is a development of the ensemble-based transform Kalman filter (ETKF; Bishop et al., 2001) and of the local ensemble Kalman filter (LEKF; Ott et al. , 2004), and is particularly suited to high-performance computing applications. A very attractive feature of an ensemble-based technique is the use of a flow-dependent background error covariance, which is derived from the ensemble of model states at the assimilation time, and evolves during forecast. At any given time in fact the state estimate is represented by an ensemble of system states and its uncertainty is derived from the ensemble. LETKF has the advantageous feature that it applies localization, i.e. it performs the analysis locally (at each grid point only observations within a certain distance are assimilated), allowing the global analysis to explore a much higher-dimensional space than the subspace spanned by the ensemble (whose dimensionality is limited by the number of ensemble members). Localization also reduces the effect of spurious long-range covariances, effectively eliminating them after a given distance. This is particularly suitable for the assimilation of aerosol information since, as mentioned in the introduction, it has been observed that aerosol fields have limited spatial correlations ($\sim$100 km). Schutgens et al. (2010a, b) have already shown the positive impact of assimilating aerosol ground station observations using a LETKF assimilation system for the SPRINTARS model, while Sekiyama et al. (2010) used it to assimilated CALIOP vertical profiles in the MASINGAR model and Dai et al. (2013) used it to ingest MODIS observations in the NICAM-SPRINTARS model.

Here we discuss the basic concepts behind the LETKF algorithm, a more detailed description of the scheme can be found in Hunt et al. (2007). Consider a state vector $\boldsymbol{x}$ of the dynamic variables of a system (for our application this is dust mass mixing ratios). The mean analysis increment at a grid point is estimated as a linear combination of the background ensemble perturbations $\mathbf{X}^b$:

$$\bar{\boldsymbol{x}}^a = \bar{\boldsymbol{x}}^b + \mathbf{X}^b \mathbf{w} \tag{2}$$

where we use the superscripts a and b to denote respectively the analysis and background state vector, and where the ith column of the matrix $\mathbf{X}^b$ is $\boldsymbol{x}^{b(i)} - \bar{\boldsymbol{x}}^b$, $\{i = 1, 2, \ldots, k\}$ with k ensemble members, i.e. the difference between the ith ensemble forecast $\boldsymbol{x}^{b(i)}$ and the ensemble forecast mean $\bar{\boldsymbol{x}}^b$. $\mathbf{w}$ is termed the "weight" matrix specifying what linear combination of the background ensemble perturbations is added to the background mean to obtain the analysis ensemble. The "weight" matrix is given by

$$\mathbf{w} = [\mathbf{Y}^b \mathbf{R}^{-1} \mathbf{Y}^b + (k-1)\mathbf{I}]^{-1} \mathbf{Y}^b \mathbf{R}^{-1} (\boldsymbol{y}^o - \bar{\boldsymbol{y}}^b) \tag{3}$$

where $\mathbf{Y}^b$ is the background ensemble perturbation matrix in observation space (or background observation ensemble perturbation matrix), $\mathbf{R}$ is the observation error covariance matrix which we assume diagonal, $\mathbf{I}$ is the identity matrix, $\boldsymbol{y}^o$ is the vector of observations and $\bar{\boldsymbol{y}}^b$ is the mean background observation ensemble. The background observation ensemble is obtained applying the observation operator $h(\cdot)$ to the ensemble forecast members $\boldsymbol{x}^{b(i)}$, i.e. $\boldsymbol{y}^{b(i)} = h(\boldsymbol{x}^{b(i)})$.

LETKF uses R-localization, i.e. the localization is performed in the observation error covariance matrix, making the influence of an observation on the analysis decay gradually toward zero as the distance from the analysis location increases. To achieve this, the observation error is divided by a distance dependent function that decays to zero with increasing distance: $\cancel{e^{-\frac{dist^2}{h^2}}}$ $e^{-\frac{dist^2}{l^2}}$, where $dist$ is the distance in the grid space between an observation and the model grid in which the analysis is calculated, and $\cancel{h}$ l is horizontal localisation factor.

3.1 Ensemble perturbations

We run the data assimilation scheme under an imperfect model scenario assumption: each ensemble member is run with a different perturbation of uncertain model parameters in the dust emission scheme. In dust modelling, the emission source term is a particularly large contributor to model error (Huneeus et al., 2011). In the case of ~~NMMB/BSC-CTM~~ NMMB-MONARCH one of the component to the uncertainty in the emission term has been identified for example in the vertical flux distribution at sources (Gama et al., 2016).

The model ensemble is created perturbing the vertical flux of dust in each of the eight dust bins. As described in Section 2, ~~NMMB/BSC-CTM~~ NMMB-MONARCH follows a sectional approach for dust, i.e. the size distribution is decomposed in small size bins that from bin 1 to bin 8 go from 0.1 to 10 μm with division intervals at 0.18, 0.3, 0.6, 1, 1.8, 3, and 6 μm. This is equivalent to perturbing the total vertical flux as well as its size distribution at sources. The perturbations are extracted imposing some physical constraint: correlated noise is used across the bins so that noise correlation decreases with increased difference of the normalized cubic radius among the bins; the noise is applied multiplicatively and has mean 1 and standard deviation of 30% of the unperturbed value in each bin; and the final distribution is unimodal. Figure 1 shows how the vertical flux is perturbed in our ensemble simulations. Additionally, we have perturbed the threshold friction velocity for dust emission by extracting a multiplicative random factor from a normal distribution with mean 1 and spread 0.4. This considers the uncertainty of the model with respect to both surface winds and soil humidity. At low resolution, model surface winds are typically underestimated over dust sources. Also, the model uses the scheme of Fecan et al. (1999) to calculate the increase of the threshold friction velocity with soil humidity, which is typically overestimated in arid regions (Haustein et al., 2015). The spin-up period for the ensemble ensures that perturbations applied at the sources propagate everywhere in the globe. For this reason at this first stage of development of our ensemble system we did not consider necessary a combined meteorology and source perturbation. The structure of our source perturbations, for both types of perturbations, is temporally and spatially constant.

3.2 Observation operator

Our state vector is the dust mass mixing ratios. Therefore an observation operator is needed to map the ensemble mean state vector into the observation space. The simulated AOD at wavelength λ is calculated at a given observation location according to the following linear operator:

$$AOD_{\lambda} = \sum_{b=1}^{8} \frac{3}{4\rho_b r_b} M_b Q_{\lambda b} \tag{4}$$

where ρ_b [$kg m^{-3}$] is the particle mass density, r_b [m] is the effective radius, M_b [$kg m^{-2}$]is the dust column mass loading calculated from each dust bin mixing ratio, and $Q_{\lambda b}$ is the extinction efficiency factor which is calculated for using the Mie scattering theory assuming dust spherical, non soluble particles, and, within a bin, a lognormal distribution for dust with geometric radius of 0.2986 μm and standard deviation of 2.0.

When using in the state vector the total mass mixing ratio, as we will explain in Section 5, an ensemble averaged extinction efficiency is calculated as in Schutgens et al. (2010b) as an average of the extinction efficiency of the individual bins weighted by the bin mixing ratios.

Hereafter, when we will use the term AOD without specifying the wavelength, we imply that we refer to aerosol optical depth at a wavelength of 550 nm, which is the most commonly reported value in the literature.

4 Observational data

4.1 MODIS Dark Target and OMI

We consider for assimilation the MODIS Level 3 AOD product produced by the U.S. NRL and University of North Dakota (~~(~~hereafter called NRL MODIS). The NRL MODIS product is produced for the purpose of assimilation into aerosol transport models (Zhang and Reid, 2006; Hyer et al., 2010; Shi et al., 2011), post-processing the Level 2 MODIS Dark Target product from the so-called Collection 5 (Remer et al., 2008; Levy et al., 2007a, b), and is available both over land and ocean. The MODIS Level 2 product is an average of the 1 km by 1 km retrievals (at nominal resolution) generated by the Dark Target algorithm applied to top-of-atmosphere reflectances observed by the MODIS sensor on-board of the NASA polar-orbiting satellites Terra and Aqua. The NRL MODIS Level 3 product is filtered and corrected in order to eliminate outliers and gross systematic biases, spatially aggregated into a 1 by 1 degree mesh in order to avoid the assimilation of sub-grid features, and an error is estimated for each observation. The product is generated every six hours at 0, 6, 12, 18 UTC and is based on MODIS Level 2 observations in a 6 hour interval around those times. The retrieval errors estimated by NRL/University of North Dakota were used for the observation errors. They include the instrumental error variance and the spatial representation error variance. Following the approach in Zhang et al. (2008), we assume uncorrelated observation errors. These observations pertain to the total atmospheric aerosol column, not just to dust AOD. The selection of observations in dust-dominated conditions is performed using retrievals of Ångström Exponent (AE) from the original MODIS Level 3 product (Hubanks et al., 2008), for information about the size of the particles, and using retrievals of Aerosol Absorbing Index (AAI) from the Ozone Monitoring Instrument (OMI) sensor (Torres et al., 2007), for information about the absorption characteristics of the particles. Ångström Exponent (AE) values are based on quality assurance-weighted 470 and 660nm optical depths over land, and 550 and 865nm optical depths over sea. Observations are selected when daily MODIS Aqua or Terra products contain a value for AE<0.75 and daily OMI products contain value for AAI>1.5. Figure 2 shows an example for the NRL MODIS Level 3 product for a day of May 2007 after the filter for dust-dominated conditions is applied.

4.2 MODIS Deep Blue

The MODIS Dark Target product does not provide information over very bright reflective surfaces, including deserts, as the retrieval algorithm assumes low surface albedo. We consider the assimilation of MODIS Deep Blue Level 3 daily AOD product from Collection 6 whose algorithm retrieves

AOD also over bright arid land surfaces, such as deserts. The Collection 6 product covers all cloud-free and snow-free surfaces, and can be potentially very useful for mineral dust applications as it is able to provide observational constraint close to dust sources. The Deep Blue algorithm uses top-of-atmosphere reflectances at 412 and 470 nm. In the presence of heavy dust load also the reflectance at 650 nm is used. The algorithm exploits the fact that, over most surfaces, darker surface and stronger aerosol signal is seen in the blue wavelength range than at longer wavelengths. The quality of the MODIS Deep Blue AOD product is improved in Collection 6 compared to Collection 5, as work of Sayer et al. (2014), based on Level 2 retrievals, showed. Similar findings, for the northern African and Middle East deserts, were reported by Gkikas et al. (2015b), who used Level 3 retrievals over the period 2002-2014.

We have applied to this product the same filter for dust-dominated conditions described in Section 4.1. In addition we have masked out Level 3 retrievals obtained with less than 30 Level 2 retrievals, since Gkikas et al. (2015a) showed that the agreement between MODIS-AERONET is improved when the sub-pixel spatial representativeness is increased. The MODIS Deep Blue observations are not corrected for possible systematic biases, however, we are aware that for future applications we should address any possible bias in the product. It is important to notice that the Level 3 product is an aggregation of Level 2 retrievals that is produced using the highest quality retrievals (i.e. retrievals with quality assurance flag value 3). Furthermore, we have applied a quality control on all the assimilated observations based ~~of~~ on normalized first-guess departures. As proxy for the normalization factor, we have used the standard deviation of first-guess departures.

A study by Sayer et al. (2014) shows that highest quality data have an absolute uncertainty of approximately $(0.086 + 0.56AOD_{550})/AMF$, where AMF is the geometric air mass factor with a typical AMF value of 2.8. We have used this quantification of the uncertainty for the Level 3 product. Furthermore, we have defined the representation component of the error as the standard deviation of the values used in the aggregated product. Although a more accurate treatment for the representation error could be envisaged following for example the approach of van Leeuwen (2014), we deem small the impact that our approximation has on the analysis. Figure 3 shows an example for the MODIS Deep Blue Collection 6 Level 3 product for a day of May 2007 after the filter for dust-dominated conditions is applied.

The number of MODIS Deep Blue and Dark Target observations used over the experimental period is shown in Figure 4.

4.3 AERONET

For validation purposes we have used observations from the ground-based stations of the global Aerosol Robotic Network (AERONET; Holben et al., 1998) of direct-sun photometers. These observations have not been assimilated in our test simulations. In particular, we have used their retrievals of column-integrated aerosol optical depth from direct-sun photometric measurements. The retrievals are obtained observing the extinction of direct solar radiation due to the presence of aerosols in the atmosphere. For this reason AERONET retrievals are not available under cloudy sky conditions and during night-time. These observations suffer of a relatively sparse spatial coverage but are very valuable for validation purposes as their uncertainty on these retrievals is estimated to be between 0.01 and 0.02 (Eck et al., 1999) . Several studies have in fact used the AERONET data for validation purposes, or for the correction of biases in satellite measurements (Zhang and Reid, 2006; Hyer et al., 2010). We considered cloud-screened and quality-assured (Level 2.0) direct-sun AOD retrievals between 440 and 870 nm. AERONET AOD at 550 nm was obtained using the Ångström law.

5 Numerical simulation set up

We have run a set of different experiments (listed in Table 1): a control experiment to produce a 5-day forecast (hereafter called Control experiment) with the same operational configuration (but at a coarser resolution) and version that provides daily global forecast to the aforementioned ICAP multi-model ensemble, and which is initialized for dust from the previous day 24 hour forecast (FC+24). Assimilation experiments were run with NRL MODIS AOD (hereafter called DA-NRL experiment) and with NRL MODIS AOD and MODIS Deep Blue AOD (hereafter called DA-NRL-DB experiment) with a preprocessing to the observations as described in Section 4. Additionally, we have run also free ensemble simulations without assimilating any observation (hereafter called ENS-free-run). We have also run a 5-day forecast experiment initialized from the analysis produced by the DA-NRL-DB experiment (hereafter called AN-initialized experiment) in order to evaluate the impact of the analysis on the forecast. The Control experiment was run for five months in the spring/summer period of 2007 (from 1 April to 31 August 2007) starting from a cold start for dust and with a spin up period of one month (April 2007) which is excluded from the analysis of the results. Also the ensemble is spun up before data assimilation is applied.

We use a 24-hour assimilation window and observations are considered for assimilation at four time slots within the window, at 0, 6, 12 and 18 UTC. The system uses as first-guess a 1-day forecast with output every 6 hours. Simulated observation and background departures are calculated at each time slot. The time slots are exactly the ones corresponding to the times in which NRL MODIS AOD observations are available. We are using the LETKF implementation with a four-dimensional extension as described in Hunt et al. (2007). The state vector comprises of the mixing ratio at all the time slots considered and so does the observation AOD vector. Background observation means $\bar{\mathbf{y}}_j$ and perturbation matrices $\mathbf{Y}_j$ are formed at the various time slots j when the observations are available. They are then vertically concatenated to form a combined background observation mean $\bar{\mathbf{y}}$ and perturbation matrix $\mathbf{Y}$. $\bar{\mathbf{y}}$ and $\mathbf{Y}$ are used for the calculation of a weight matrix $\mathbf{w}$ using the standard LETKF, i.e. we calculate a single **w** based on all innovations throughout the day. This same **w** is then applied to the state vector at different times throughout the assimilation window.

We have tuned different aspects of the data assimilation system including testing the number of ensemble members, different perturbations of the ensemble, and a different state vector for the control variables. Using 24 ensemble members did not produce a significant impact on the dust analysis compared to the use of 12 ensemble members. This could be explained with our setting of a localization factor which makes the influence of an observation on the analysis decay gradually toward zero as the distance from the analysis location increases. We have set the horizontal localization factor to the value 1 in all the data assimilation experiments. This means that after 2 grid points the localization function is very close to zero. The value chosen is in the range of the ones used in previous studies such as Rubin et al. (2016) and Yumimoto and Takemura (2011) . Covariance localization in fact effectively eliminates background spatial correlations after a certain distance, and might have solved possible sampling errors introduced by the low dimensionality of the 12 member ensemble compared to the 24 member ensemble. We also apply vertical localization following Miyoshi and Yamane (2007) approach of localizing the error covariance vertically for radiance assimilation. The observation error is divided by the square of the model AOD ~~normalised~~ normalized sensitivity function.

We have tested the usage of different perturbations of the dust emission scheme: a perturbation of the mass vertical flux per dust bin, or a the perturbation of both the mass vertical flux and the threshold on the wind friction velocity. As we show in the next section, the latter configuration was deemed better as it spans a larger space of possible system states.

We have tested two different options for the state vector: a control variable consisting of the mixing ratio of eight individual dust bins or the total dust mixing ratio defined as the sum of the eight dust bins at each grid point and for all the vertical levels. In the latter case the mixing ratios for the individual dust bin after data assimilation are determined from the background, i.e. from their relative fractions before assimilation. The observation operator is calculated using the original mixing ratio following the approach for the observation operator in Schutgens et al. (2010b). The tests that we have performed show that representing individually the bins in the state vector does not have any significant impact on the analysis, while it increases the computational cost of the assimilation compared to using the total mixing ratio. Moreover, the use of a bulk approach is common in systems assimilating total AOD values as the observations are not able to fully constrain the individual bin profiles. We should note that this choice of state vector makes still meaningful the creation of the ensemble perturbing the vertical flux for the individual bins, as this allows us to express in the background the uncertainty in the size distribution at sources, and to span a larger space of possible system states.

In the next section we show the results of assimilating NRL MODIS NRL and MODIS Deep Blue observations using 12 ensemble members obtained perturbing the mass vertical flux per bin at sources together with the threshold on the wind friction velocity, as described in Section 3.1, and using the total dust mixing ratio as analysis variable in the state vector. All simulations were run on a global domain with 40 hybrid pressure-σ layers, 5 hPa top pressure, and a horizontal resolution of 2.8 by 2 degree. The NCEP final analysis at 1 by 1 degree at 0 UTC were used to initialize the meteorology at every forecast run.

6 Methodology for the evaluation of the simulations

The evaluation of the simulations is done in three stages: (a) an internal check of the data assimilation system; (b) evaluation of the analysis using as reference independent observations; (c) evaluation of a 5-day forecast with and without analysis initialization using as reference independent observations.

The consistency of the data assimilation system is checked through considerations on statistics of the ensemble, of simulation departures from assimilated observations, and of the temporal mean of assimilation increments. The ensemble mean and the coefficient of variation for the ensemble are calculated with and without data assimilation. The coefficient of variation is defined as ratio of the standard deviation of the ensemble to the ensemble mean. Additionally, statistics for first-guess (FG) and analysis (AN) departures are calculated, where departures are defined as difference between assimilated observations and simulations (first-guess or analysis), while mean increments are defined as temporal mean of differences between analysis and first-guess at the different time slots of the assimilation window.

The evaluation of analysis and forecast with respect to independent observations are performed in terms of statistics of model field errors e_i from observations, where ~~$e_i = o_i - m_i$~~$e_i = m_i - o_i$, with index i indicating an instance of observation o_i and where m_i is the model field in observation space, bi-linearly interpolated at the observation location. We consider the root mean square error ($RMSE$), the mean error ($BIAS$), the standard deviation of the error (SD), the fraction gross error ($FRGE$), and the correlation coefficient ($CORR$) of the model AOD compared to either quality-assured (Level 2.0) AERONET or to satellite retrievals. The $FRGE = \frac{2}{n} \sum_{i=0}^{n} \mid \frac{o_i - m_i}{o_i + m_i} \mid$ is added to the most widely used set of statistics for the error as it behaves symmetrically with respect to under and over estimation without emphasizing the outliers, and is normalized to the sum of observation and simulation values. The SD of the error, though it can be derived from the other statistics, is also reported so to make more explicit the changes in the bias-free mean square error and aid the interpretation of the evaluation results. The above set of evaluation statistics are calculated for measurements from individual ground-based stations, groups of stations, regional domains observed by satellite sensors, and globally.

For AERONET AOD measurements dust-dominated conditions are identified using the approach of Basart et al. (2009) as follows: AOD is classified as 'Dust' AOD when the associated AE<0.75; we set 'Dust' AOD to 0 when the associated AE>1.3; we identify a mixed aerosol type when the associated $0.75<AE<1.3$. The latter values are excluded from the validation. We use the AERONET AOD value closest to the model time step and within a ± 30 minute interval. For satellite AOD retrievals we use the set of satellite observations quality controlled and filtered for dust-dominated conditions used in the assimilation step. We use these satellite observations to validate uniquely the forecast range following the assimilation window. We show the forecast evaluation statistics corresponding to measurements and simulations at 12 UTC only, so that they refer to an approximate equal number of pair of observations and model simulated values at each forecast lead time that we are considering. Hence a smaller number of AERONET observations (at 12 UTC only) are used to verify the forecast compared to the ones used in the evaluation of the analysis.

We have identified eight regions of interest for the validation purposes in our study period, namely: Long Atlantic transport (LongAtl), Short Atlantic transport (ShortAtl), Sub-Sahel (SubSahel), Sahara (Sahara), Extended Mediterranean (ExtMediter), Middle East (MiddleEast), Central Asia (CenAsia), East Asia (EastAsia). These names do not necessary correspond to the conventional names of exact geographical locations but are meant to identify regional domains in a convenient way according to dust intrusions and to group observational stations. Most of regional domains contain ground-based stations with a minimum number of observations during the study period (stations with less than 30 'Dust' observations are discarded), with the exception of Central and East Asia. The ground-based stations are listed in Table 2, and shown in the map in Figure 5 together with regional domains used for the validation of the experiments either against ground-based or satellite observations.

7 Evaluation results

7.1 Ensemble, departure and increment statistics

We compare here the dust fields in the Control, ENS-free-run, DA-NRL and DA-NRL-DB experiment in terms of mean values and, when applicable, ensemble spread. Figure 6 shows the dust AOD values averaged over a month of the study period for the four above experiments, and the difference in AOD between the data assimilation experiments and the ENS-free-run. By visual inspection it can be noticed that the ensemble mean of the ENS-free-run experiment compares well with the Control experiment, which suggests that the ensemble perturbations are ~~not altering~~ altering only at a small extent the model mean state as defined by a standard run. The analysis clearly shows conspicuous changes in the dust field compared to the Control experiment or the ENS-free-run. Figure 7 shows the coefficient of variation for AOD in the ENS-free-run and ~~DA-NRL-DB experiment~~the data assimilation experiments. Data assimilation clearly lowers the values of the coefficient of variation in the regions where observations are present, with values lower for the DA-NRL-DB than for the DA-NRL experiment, which indicates a reduction of the ensemble spread due to the assimilated observations. The ~~ensemble~~ high values of the coefficient of variation in the Southern Hemisphere, with or without data assimilation, are due to the perturbation of the dust sources present in the south part of the globe. These values are not negligible due to differences among the ensemble members normalized to small dust AOD values. The ensemble of Figure 7 (and Figure 6) is created perturbing the emitted mass vertical flux for each dust bin and the threshold on the friction velocity generating dust horizontal flux. Creating the ensemble without perturbing the threshold on the friction velocity produces a reduced spread. See Figure 8 for this second configuration of the ensemble with ~~ensemble mean and~~ coefficient of variation for the ENS-free-run in the ~~top panels, and~~ left panel and for the experiment with data assimilation in the ~~bottom panels~~right panel. Perturbing the threshold on the friction velocity has an impact on the spread also outside source regions because, as explained earlier, the spin-up period for the ensemble ensures that perturbations applied at the sources propagate everywhere. ~~A smaller spread~~ Furthermore this ensemble configuration better represents model uncertainty since the ratio of the prior total spread (the square root of the sum of the ensemble background variance and the observation error variance) to the prior RMSE (of the ensemble background against NRL MODIS and MODIS Deep Blue observations) is closer to 1 compared to when no perturbation is applied to the threshold on the friction velocity. It should be noted, however, that this chosen ensemble configuration is under representing uncertainty since it has a prior total spread smaller than the RMSE (ratio equal to 0.82). Other better perturbations should to be tested for a future implementation since an under representation of the background uncertainty might translate to giving a lower weight to the observations with respect to the background~~, therefore the ensemble with higher spread is used in the simulations described in the rest of the paper.~~ .

We evaluate in the rest of this section the assimilation experiments in terms of statistics of the departures of the analysis and first-guess from the assimilated satellite observations. Figure 9 shows for May to August 2007 first guess dust AOD (on the left panels) and analysis dust AOD (on the right panels) versus observations for the DA-NRL and DA-NRL-DB experiment. The departure statistics with respect to the two sets of observations that we have assimilated are in Table 3. In both experiments a smaller scatter and a higher correlation coefficient for the analysis indicate that the assimilation improves the agreement with observations and hence a positive sanity check of the data assimilation system. The asymmetric behaviour of all the analysis scatter plots suggests that the system is more successful in correcting too high AOD values than correcting too low AOD values, which could be due to the fact that usually we have larger observation errors and smaller ensemble spread for low AOD values. The BIAS is significant smaller than the RMSE and the RMSE improves in the analysis over the forecast. The issue of a higher BIAS in the analysis departures compared to the first-guess departures has been identified in other assimilation system (see Benedetti et al. (2009), Section 4) and might be attributed to the fact that AOD is a positive definite variable, as this provides a deviation from the Gaussianity condition in the prior which is assumed in the analysis step. A solution to this problem worth investigating in the future would consist in applying a transformation of the state variables into new variables which present Gaussian features, a procedure known as Gaussian anamorphosis (Amezcua and Van Leeuwen, 2014).

Figure 10 shows global maps of mean dust AOD analysis increments, i.e. the monthly-averaged difference between analysis and short-term forecast, respectively in the case in which only NRL MODIS AOD observations are assimilated and in the case in which also MODIS Deep Blue AOD observations are assimilated. Both experiments show non-zero systematic increments which are to be interpreted as systematic corrections that these sets of observations are making, in particular removing mass close to sources and, to a lesser extent, adding mass in the outflow. The spatial distribution of the increments highlights the role that MODIS Deep Blue observations play in particular over the Sahara dust sources. There are some regions where the two data assimilation experiments show opposite increments. This could be due to unresolved conflicting biases between the two types of MODIS retrievals.

7.2 Validation of the analysis

We perform in this section a validation of the dust fields simulated either with or without data assimilation through a comparison with observations from ground-based stations that have not been assimilated for May to August 2007. We calculate the statistics for individual stations and for groups of stations. Figure 11 shows the time-series of dust AOD values for May to August 2007 for the Control experiment (blue), for the analysis of the DA-NRL (green) ~~,~~ and of the DA-NRL-DB (red) experiment, and for AERONET observations in dust-dominated conditions (black) at six locations within the different regional domains of Figure 5, which are in the proximity of dust sources (Tamanrasset in Algeria), affected by short-range dust transport (Dakar in Senegal, Ilorin in Nigeria, and Hamim in the United Arab Emirates), or affected by long-range dust transport in Europe (Lecce in Italy), and across the Atlantic (La Parguera in Puerto Rico). For reference, also the MODIS AOD observations from the assimilated dataset (NRL and Deep Blue) which are at the closest distance and within a 2 degree radius from the location of the AERONET station are included in the time-series (magenta circles). Note, however, that these latter observations are not an independent reference for validation of the analyses, nor are entirely representative of the observational constraint used to calculate the analysis in the given station location. The time-series show an overestimation in the Control experiment of the optical depth near the sources, and to a smaller extent in the transport which clearly suggests that the model tends to overestimating dust emissions. The current calibration for model version 1.0 has the shortcoming to accurately capture long-range transport at the expenses of an overestimation over the sources. This overestimation is reduced with data assimilation. By a first eyeball inspection, the AOD simulation variance is reduced by data assimilation and is more in accordance with the AOD observation variance.

Maps in Figure 12 show results of validation statistics calculated for the full study period at each AERONET station for the three experiments performed. These maps allow us to appreciate the strongest features of the three simulations at individual AERONET stations and how those stations are representative of the regional domains that we have identified. The Control experiment shows that the strongest BIAS and highest RMSE are in the sub-Sahel region. The BIAS indicates that the model systematically over-predicts AOD in that region. The highest FRGE are in the long transport over the Atlantic or Europe as expected in areas of low AOD values. The correlation between model and observation values is in general lower near source areas than in outflow regions. This could be due to the too coarse model resolution not able to follow as good as the observations the dynamic of the dust field near source areas. The assimilation of MODIS NRL observations decreases some of the strongest biases in particular in the dust outflow regions in Sahel and over the African Atlantic coast, which is reflected in a reduced FRGE and RMSE, and is associated with improved correlation. The assimilation of the MODIS Deep Blue observations additionally to the NRL MODIS observations is of further benefit: it reduces the BIAS and RMSE downwind from the strongest dust sources of Sahara. It is also relevant to notice that the additional assimilation of MODIS Deep Blue observations improves the correlation over the above areas and in the Arabian peninsula.

The chart plots for the validation statistics calculated for all the AERONET stations considered (hereafter called global statistics) and for stations grouped according to regional domains of interest are respectively in Figure 13 and Figure 14. Global statistics show that assimilation produces in general a better representation of dust concentrations in the atmosphere, and that the assimilation of Deep Blue retrievals has a positive impact over the assimilation of Dark Target retrievals only.

When considering the regional domains, the assimilation of NRL MODIS AOD has a positive impact on the quality of the analysis everywhere, with the only exception of a slightly increase of RMSE in the Middle East region. This positive impact is more pronounced in the short Atlantic transport and in the sub-Sahel region. The additional assimilation of MODIS Deep Blue AOD has a considerable positive impact in the Sahara, sub-Sahel and Middle East regions, and neutral or slightly detrimental in the rest of the transport, in particular in the long range Atlantic transport. The correlations for the global domain and for all the regional domains are highly statistically significant with the exception of the Sahara region (in the Control and DA-NRL experiments only) where number of observations is smaller than other regional domains.

It should be noted, however, when interpreting the above statistics that the validation against AERONET observations ~~introduces~~ may introduce significant errors when comparing a global model grid-box against a point observation (Schutgens et al., 2016).

7.3 Validation of the forecast

We have validated the forecast up to 5 days ahead initialized at 0 UTC from either the control experiment or an analysis (from DA-NRL-DB). We have calculated for May to August 2007 the errors for the forecast at 12, 36, 60, 84, 108 hours (hereafter indicated as FC+12, FC+36, FC+60, FC+84, FC+108) with respect to either AERONET observations or satellite observations. As mentioned when describing our evaluation methodology, we use as reference the set of satellite observations from the Dark Target and Deep Blue algorithm ingested in the assimilation step, i.e quality-controlled and filtered for dust-dominated conditions. They are used only to validate the forecast range following the assimilation window. As expected, all the validation statistics worsen with increased forecast step in both experiments (see Figure 15 for global statistics). The impact of initializing the model with a dust analysis is positive in the first day. The analysis produces a better forecast in terms of BIAS and RMSE (and also SD of the error) up to FC+108, and a better correlation in the first day. The correlation is slightly lower from FC+36 onwards. The conclusions drawn by validating against AERONET or satellite observations are equivalent. Results calculated for regional domains (Figure 16) show that the Control experiment tends to overestimate AOD everywhere with the exception of central and east Asia. This suggests an overestimation in particular of the Sahara emissions which is consistent with the bias found in the analysis and which is maintained during the forecast. The correlations for the global domain and for all the regional domains, at all forecast lead times, are highly statistically significant. Initializing the 0 UTC forecast with the DA-NRL-DB dust analysis reduces the overestimation compared to satellite retrievals in the first day of the forecast consistently with the improvement observed in the analysis in the previous section. However, this produces an underestimation of AOD in the long-range Atlantic transport during all the forecast lead times, which, because of the relatively small AOD values in that area, is reflected in particular in the FGRE. Although there is an overestimation of AOD, there is a better agreement of the temporal evolution in that region. The underestimation of AOD in the Atlantic transport might be due to too strong deposition which affects in particular the long-range transport, and in the standard run is compensated by an overestimation over the sources. As said earlier, a shortcoming of the current model calibration is to capture well the long-range transport at the expenses of an overestimation over the sources, which data assimilation reduces. To identify the exact cause for it will require, however, further investigation together with a better adjustment of the current model parameters. With the exception of this underestimation of AOD across the Atlantic, all the error statistics and correlation coefficients are improved in the first day of the forecast in all the regional domains. The error of the analysis-initialized forecast is lower ~~for more than 4 days into the forecast~~also in the rest of the forecast range (up to 5 days), though, after day 1, the ~~temporal evolution is less in agreement (lower correlation)~~ correlation with satellite observations in some regions (SubSahel and ShortAtl) ~~, compared to~~ is lower for the analysis-initialized forecast than for a standard forecast. It is particularly relevant to notice that the dust forecast over Sahara is improved for all the statistics and throughout the forecast range.

8 Conclusions

We have developed a data assimilation system for the ~~NMMB/BSC-CTM~~ NMMB-MONARCH model version 1.0, which considers dust only, while other aerosols are being implemented. We have coupled the ~~NMMB/BSC-CTM~~ NMMB-MONARCH with an ensemble-based data assimilation technique known as LETKF. For this purpose we have created a forecast ensemble based on known uncertainties in the physical parametrizations of the mineral dust emission scheme. We have processed satellite aerosol optical depth retrievals for assimilation with a dust filter. Due to the presence of other aerosols in the selection of dust-dominated conditions, uncertainties might have been introduced in our assimilation process. It should be noted however that the identification of dust-dominated conditions is performed in this study as a proof of concept to demonstrate the potential of using data assimilation in ~~NMMB/BSC-CTM~~NMMB-MONARCH, and will not be strictly necessary in a future model upgrade including all the major aerosol species. Still, efforts towards aerosol speciation could continue to be pursued when assimilating information about total aerosol optical properties. In this respect, operational centres currently rely merely on model background to distribute assimilation increments among the different aerosol species.

Assimilation experiments showed that aerosol optical depth retrieved with the Dark Target algorithm can help ~~NMMB/BSC-CTM~~ NMMB-MONARCH to better characterize atmospheric dust. This is particularly true for the analysis of the dust outflow in the Sahel region and over the African Atlantic coast. The additional assimilation of Deep Blue retrievals has a further positive impact in the analysis downwind from the strongest dust sources of Sahara and in the Arabian peninsula.

An analysis-initialized forecast performs better (lower forecast error and higher correlation) than a standard forecast everywhere in the first day of the forecast. The only exception to this is an underestimation of the forecast of AOD in the long-range Atlantic transport. The error of the analysis-initialized forecast is lower also in the rest of the forecast range (up to 5 days), though, after day 1, in sub-Sahel and short Atlantic transport the temporal evolution of dust is less in agreement with independent observations, compared to a standard forecast. Particularly relevant is the improved forecast over Sahara throughout the forecast range thanks to the assimilation of Deep Blue retrievals over areas not easily covered by other observational datasets. To the best of our knowledge, this is the first study quantifying the benefit of assimilating MODIS Deep Blue from Collection 6 specifically for mineral dust simulations. This product is currently operationally assimilated by the UK Met Office who consider only Deep Blue observations over desert, and by the European Centre for Medium-Range Weather Forecasts.

In our future implementation of the forecast ensemble, we plan to exploit spatial patterns of variation in model parameter uncertainty, for example source-dependent uncertainties, as well as uncertainties in the deposition term. A better representations of uncertainties in dust emission flux inherently will help the representation of uncertainties in other parts of the dust cycle. A recent study by Rubin et al. (2016) shows that, for their system, a combined meteorology and aerosol source ensembles are necessary to produce sufficient spread in outflow regions. Notwithstanding that their conclusion might be system-dependent, we will be take into account their results in our future studies.

9 Code availability

Copies of the code are readily available upon request from the corresponding authors.

Acknowledgements. This work was funded by the SEV-2011-00067 grant of the Severo Ochoa Program awarded by Spanish Government, the CGL-2013- 46736-R grant of the Spanish Ministry of Economy and Competitiveness, and the ACTRIS Research Infrastructure Project of the European Union's Horizon 2020 research and innovation programme under grant agreement n. 654169. The authors thank all the Principal Investigators and their staff for establishing and maintaining the AERONET sites, NRL-University of North Dakota for the MODIS AOD L3 product, the MODIS and OMI mission scientists and associated NASA personnel for the production of the AOD, AAI and AE data used in this investigation. ~~All the simulations have been run on the MareNostrum supercomputer hosted by BSC.~~ The authors thankfully acknowledge the computer resources at MareNostrum and the technical support provided by the Barcelona Supercomputing Center (RES-AECT-2015-1-0007). AXA Professor Carlos Pérez García-Pando acknowledges long-term support from the AXA Research Fund.

References

Amezcua, J. and van Leeuwen, P. J.: Gaussian anamorphosis in the analysis step of the EnKF: a joint state-variable/observation approach, Tellus, 66A, 23493, 2014.

Anderson, T. L., Charlson, R. J., Winker, D. M., Ogren, J. A., and Holmèn K.: Mesoscale Variations of Tropospheric Aerosols, J. Atmos. Sci., 60(1), 119-136, 2003.

Badia, A., Jorba, O., Voulgarakis, A., Dabdub, D., Pérez García-Pando, C., Hilboll, A., Gonçalves, M., and Janjic, Z.: Description and evaluation of the Multiscale Online Nonhydrostatic AtmospheRe CHemistry model (NMMB-MONARCH) version 1.0: gas-phase chemistry at global scale, Geosci. Model Dev., 10, 1-30, 2017.

Basart, S., Pérez García-Pando, C., Cuevas, E., Baldasano, J. M., and Gobbi, G. P.: Aerosol characterization in Northern Africa, Northeastern Atlantic, Mediterranean Basin and Middle East from direct-sun AERONET observations, Atmos. Chem. Phys., 9, 8265-8282, doi:10.5194/acp-9-8265-2009, 2009.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons, A. J., and Suttie, M.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, J. Geophys. Res., 114, D13205, doi:10.1029/2008JD011115, 2009.

Bishop, C. H, Etherton, B. J., Majumdar, S. J.: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects, Mon. Wea. Rev., 129, 2001.

Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Žabkar, R., Carmichael, G. R., Flemming, J., Inness, A., Pagowski, M., Pérez Camaño, J. L., Saide, P. E., San Jose, R., Sofiev, M., Vira, J. Baklanov, A. Carnevale, C., Grell, G., and Seigneur, C.: Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models, Atmos. Chem. Phys., 15, 5325-5358, 2015.

Chaboureau, J.-P., Richard, E., Pinty, J.-P., Flamant, C., Di Girolamo, P., Kiemle, C., Behrendt, A., Chepfer, H., Chiriaco, M., and Wulfmeyer, V.: Long-range transport of Saharan dust and its radiative impact on precipitation forecast: a case study during the Convective and Orographically-induced Precipitation Study (COPS), Q. J. Roy. Meteor. Soc., 137, 236-251, doi: 10.1002/qj.719, 2011.

Dai, T., Schutgens, N. A. J., Goto, D., Shi, G.,and Nakajima, T.: Improvement of aerosol optical properties modeling over Eastern Asia with MODIS AOD assimilation in a global non-hydrostatic icosahedral aerosol transport model, Environmental pollution, 195, 2014.

D'Almeida, D. A.: On the variability of desert aerosol radiative characteristics, J. Geophys. Res., 92, 1987.

Jickels, T. D., et al.: Global Iron Connections Between Desert Dust, Ocean Biogeochemistry, and Climate, Science, 308, 2005.

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A.,O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, Journal of Geophysical Research, 104, 31333-31349, 1999.

Elbern, H. and Schmidt, H.: Ozone episode analysis by four-dimensional variational chemistry data assimilation, Journal of Geophysical Research, 106, 2001.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10143-10162, 1994.

Fecan, F., Marticorena, B., and Bergametti, G.: Parametrization of the increase of the aeolian erosion threshold wind friction velocity due to soil moisture for arid and semi-arid areas, Ann. Geo-phys., 17, 1999.

Gama C., Basart, S., Baldasano, J. M., Pio, C., Borrego, C., and Tchepel, O.: Assessing the size distribution of dust emissions in the NMMB/BSC-Dust model, 8th International Workshop on Sand/Duststorms and Associated Dustfall, Lisbon, Portugal, 2-3 May 2016, 50, 2016.

Gkikas, A., Basart, S., Hatzianastassiou, N., Marinou, E., Amiridis, V., Kazadzis, S., Pey, J., Querol, X., Jorba, O., Gassó, S., and Baldasano, J. M.: Mediterranean desert dust outbreaks and their vertical structure based on remote sensing data, Atmos. Chem. Phys. Discuss. 15, 27675-27748, doi:10.5194/acpd-15-27675-2015, 2015a.

Gkikas, A., Basart, S., Korras-Carraca, M., Papadimas, C., Hatzianastassiou, N., Sayer, A., Hsu, C., and Baldasano, J. M.: Intercomparison of MODIS-Aqua C051 and C006 Level 3 Deep Blue AOD and Ångström exponent retrievals over the Sahara desert and the Arabian Peninsula during the period 2002-2014, Geophysical Research Abstracts, 17, 2015b.

Grini, A., Thulet, P., and Gomes, L.: Dusty weather forecasts using the MesoNH mesoscale atmospheric model, J. Geophys. Res., 111, 2156-2202, doi:10.1029/2005JD007007, 2006.

Haustein, K., Pérez García-Pando, C., Baldasano, J. M., Jorba, O., Basart, S., Miller, R. L., Janjic, Z., Black, T., Nickovic, S., Todd, M. C., Washington, R., Müller, D., Tesche, M., Weinzierl, B., Esselborn, M., and Schladitz, A.: Atmospheric dust modeling from meso to global scales with the online NMMB/BSC-Dust model - Part 2, Chem. Phys., 12, 2933-2958, 2012.

Haustein, K., Washington, R., King, J., Wiggs, G., Thomas, D. S. G., Eckardt , F. D., Bryant, R. G., and Menut, L.: Testing the performance of state-of-the-art dust emission schemes using DO4Models field data, Geosci. Model Dev., 8, 341-362, 2015.

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET-A federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1998.

Huneeus, N. et al., Global dust model intercomparison in AeroCom phase I, Atmos. Chem. Phys., 11, 2011.

Huneeus, N., Basart, S., Fiedler, S., Morcrette, J.-J., Benedetti, A., Mulcahy, J., Terradellas, E., Pérez García-Pando, C., Pejanovic, G., Nickovic, S., Arsenovic, P., Schulz, M., Cuevas, E., Baldasano, J. M., Pey, J., Remy, S., and Cvetkovic, B.: Forecasting the northern African dust outbreak towards Europe in April 2011: a model intercomparison, Atmos. Chem. Phys., 16, 4967-4986, doi:10.5194/acp-16-4967-2016, 2016.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, Physica D, 230, 2007.

Hubanks, P. A., King, M. D., Platnick, S. A., and Pincus, R. A.: MODIS Atmosphere L3 Gridded Product Algorithm Theoretical Basis Document, MODIS Algorithm Theoretical Basis Document No. ATBD-MOD-30 for Level-3 Global Gridded Atmosphere Products (08 D3, 08E3, 08M3), available at: http://modis.gsfc.nasa.gov/data/atbd/atbd_mod30.pdf, last access: 6 October 2015, 2008.

Hyer, E. J., Reid, J. S., and Zhang, J.: An over-land aerosol optical depth data set for data assimilation by filtering, correction, and aggregation of MODIS collection 5 optical depth retrievals, Atmos. Mea. Tech., 4, 379-408, 2011.

Janjic, Z. I. and Gall, R.: Scientific documentation of the NCEP nonhydrostatic multiscale model on the B grid (NMMB). Part 1 Dynamics, Technical Report, NCAR/TN-489+STR, 2012.

Jorba, O. Dabdub, D., Blaszczak-Boxe, C., Pérez García-Pando, C., Janjic, Z., Baldasano, J. M., Spada, M., Badia, A., and Gonçalves, M.: Potential significance of photoexcited NO2 on global air quality with the NMMB/BSC chemical transport model, J. Geophys. Res., 117, 2012.

Kalnay, E., Li, H., Miyoshi, T., Yang, S.-C., and Ballabrera-Poy, J.: 4D-Var or ensemble Kalman filter?, Tellus, 59A, 758-773, 2007.

Kaufman, Y. J., Koren, I., Remer, L. A., Tanre, D., Ginoux, P., and Fan, S.: Dust transport and deposition observed from the Terra-Moderate Resolution Imaging Spectroradiometer spacecraft over the Atlantic Ocean, J. Geophys. Res., 110, 2005.

Kim, D., Chin, M., Yu, H., Eck, T. F., Sinyuk, A., Smirnov, A., and Holben, B. N.: Dust optical properties over North Africa and Arabian Peninsula derived from the AERONET dataset, Atmos. Chem. Phys., 11, 10733-10741, doi:10.5194/acp-11-10733-2011, 2011.

Knippertz, P. and Stuut, J.-B. W. (Eds.): Mineral Dust: A Key Player in the Earth System, Springer Science, Dordrecht, Netherlands, 509 pp., 2014.

Lau, K. M., Kim, M. K., and Kim, K. M.: Asian summer monsoon anomalies induced by direct forcing: the role of the Tibetan plateau, Clim. Dynam., 26, 855-864, 2006.

Levy, R. C., Remer, L. A., and Dubovik, O.: Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land, J. Geophys. Res.-Atmos., 112, D13210, doi:10.1029/2006JD007815, 2007a.

Levy, R. C., Remer, L. A., Mattoo, S., Vermote, E. F., and Kaufman, Y. J.: Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance, J. Geophys. Res.-Atmos., 112, D13211, doi:10.1029/2006JD007811, 2007b.

Luo, T., Wang, Z., Zhang, D., Liu, X., Wang, Y., and Yuan, R.: Global dust distribution from improved thin dust layer detection using A-train satellite lidar observations, Geophys. Res.Lett., 42, 620-628, doi:10.1002/2014GL062111, 2015.

Mallone, S., Stafoggia, M., Faustini, A., Gobbi, G. P., Marconi, A., and Forastiere, F.: Saharan dust and associations between particulate matter and daily mortality in Rome, Italy. Environmental health perspectives, 119(10), 1409, 2011.

Marticorena, B. and Bergametti, G.: Modeling the atmospheric dust cycle: 1. design of a soil-derived dust emission scheme, J. Geo-phys. Res., 100, 1995.

Marticorena, B., Bergametti, G., Aumont, B., Callot, Y., N'Doumé, C., and Legrand, M.: Modeling the atmospheric dust cycle 2. simulation of Saharan dust sources, J. Geophys. Res., 102, 4387-4404, doi:10.1029/96JD02964, 1997.

Miyoshi T. and Yamane, S., Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution, Mon. Wea. Rev., 135, 2007.

Morman, S. A. and Plumlee, G. S.: The role of airborne mineral dusts in human disease, Aeolian Research, 9, 203-212, 2013.

Niu, T., Gong, S. L., Zhu, G. F., Liu, H. L., Hu, X. Q., Zhou, C. H., and Wang,Y. Q.: Data assimilation of dust aerosol observations for the CUACE/dust forecasting system, Atmos. Chem. Phys., 8, 2008.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A.: A local ensemble Kalman filter for atmospheric data assimilation, Tellus A 56, 2004.

Pandolfi, M., Tobias, A., Alastuey, A., Sunyer, J., Schwartz, J., Lorente, J., Pey, J., and Querol, X.: Effect of atmospheric mixing layer depth variations on urban air quality and daily mortality during Saharan dust outbreaks, Sci. Total Environ., 1, 2014.

Pérez García-Pando, C., Nickovic, S., Pejanovic, G., Baldasano, J. M., and Özsoy, E.: Interactive dust-radiation modeling: a step to improve weather forecasts, J. Geophys. Res., 111, 2006.

Pérez García-Pando, C., Haustein, K., Janjic, Z., Jorba, O., Huneeus, N., Baldasano, J. M.,T. Black, T., Basart, S., Nickovic, S., Miller, R. L., Perlwitz, J. P., Schulz, M., and Thomsonet, M.: Atmospheric dust modeling from meso to global scales with the online NMMB/BSC-Dust model - Part 1, Atmos. Chem. Phys., 11, 2011.

Pérez García-Pando, C., Stanton, M. C., Diggle, P. J., Trzaska, S., Miller, R. L., Perlwitz, J. P., Baldasano, J. M., Cuevas, E., Ceccato, P., Yaka, P., and Thomson, M. C.: Soil Dust Aerosols and Wind as Predictors of Seasonal Meningitis Incidence in Niger, Environ. Health Perspect., 112 2014.

Pey, J., Querol, X., Alastuey, A., Forastiere, F., and Stafoggia, M.: African dust outbreaks over the Mediterranean Basin during 2001-2011: PM10 concentrations, phenomenology and trends, and its relation with synoptic and mesoscale meteorology, Atmos. Chem. Phys., 13, 1395-1410, doi:10.5194/acp-13-1395-2013, 2013.

Remer, L. A., Kleidman, R. G., Levy, R. C., Kaufman, Y. J., Tanré, D., Mattoo, S., Martins, J. V., Ichoku, C., Koren, I., Yu, H. and Holben, B. N.: Global aerosol climatology from the MODIS satellite sensors, J. Geophys. Res.-Atmos., 113, D14S07, doi:10.1029/2007JD009661, 2008.

Rubin, J. I., Reid, J. S., Hansen, J. A., Anderson, J. L., Collins, N., Hoar, T. J., Hogan, T., Lynch, P., McLay, J., Reynolds, C. A., Sessions, W. R., Westphal, D. L., and Zhang J.: Development of the Ensemble Navy Aerosol Analysis Prediction System (ENAAPS) and its application of the Data Assimilation Research Testbed (DART) in support of aerosol forecasting, Atmos. Chem. Phys., 16, 2016.

Sayer, A. M., Munchak, L. A., Hsu, N. C., Levy, R. C., Bettenhausen, C., and Jeong, M.-J.: MODIS Collection 6 aerosol products: Comparison between Aqua's e-Deep Blue, Dark Target, and "merged" data sets, and usage recommendations, J. Geo-phys. Res. Atmos., 119, 13, 2014.

Sessions, W. R., Reid, J. S., Benedetti, A., Colarco, P. R., da Silva, A., Lu, S., Sekiyama, T., Tanaka, T. Y., Baldasano, J. M., Basart, S., Brooks, M. E., Eck, T. F., Iredell, M., Hansen, J. A., Jorba, O. C., Juang, H.- M. H., Lynch, P., Morcrette, J.-J., Moorthi, S., Mulcahy, J., Pradhan, Y., Razinger, M., Sampson, C. B., Wang, J., and Westphal, D. L.: Development towards a global operational aerosol consensus: basic climatological characteristics of the International Cooperative for Aerosol Prediction Multi-Model Ensemble (ICAP-MME), Atmos. Chem. Phys., 15, 335-362, doi:10.5194/acp-15-335-2015, 2015.

Shinozuka Y. and Redemann, J.: Horizontal variability of aerosol optical depth observed during the ARCTAS airborne experiment, Atmos. Chem. Phys., 11, 2011.

Schutgens N. A. J., Miyoshi, T., Takemura, T., and Nakajima T.: Applying an ensemble Kalman filter to the assimilation of AERONET observations in a global aerosol transport model, Atmos. Chem. Phys., 10, 2010.

Schutgens N. A. J.,Miyoshi, T.,Takemura, T., and Nakajima1 T.: Sensitivity tests for an ensemble Kalman filter for aerosol assimilation, Atmos. Chem. Phys., 10, 2010.

Schutgens, N.; Nakata, M.; Nakajima, T. Estimating Aerosol Emissions by Assimilating Remote Sensing Observations into a Global Transport Model. Remote Sens. 2012, 4, 3528-3543.

Schutgens N. A. J., Nakata, M., and Nakajima, T.: Validation and empirical correction of MODIS AOT and AE over ocean, Atmos. Meas. Tech., 6, 2013.

Schutgens, N. A. J., Gryspeerdt, E., Weigum, N., Tsyro, S., Goto, D., Schulz, M., and Stier, P.: Will a perfect model agree with perfect observations? The impact of spatial sampling, Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2015-973, in review, 2016.

Shi Y., Zhang, J., Reid, J. S., Holben, B., Hyer, E. J., and Curtis, C.: An analysis of the collection 5 MODIS over-ocean aerosol optical depth product for its implication in aerosol assimilation, Atmos. Chem. Phys., 11, 2011.

Spada M., Jorba, O., Pérez García-Pando, C., Janjic, Z., and Baldasano, J. M.: Modeling and evaluation of the global sea-salt aerosol distribution: sensitivity to size-resolved and sea-surface temperature dependent emission schemes, Atmos. Chem. Phys., 13, 2013.

Spada, M., Jorba, O., Pérez García-Pando, C., Tsigaridis, K., Soares, J., and Janjic, Z.: Global aerosols in the online multiscale NMMB/BSC Chemical Transport Model, in prep.

Sekiyama T. T., Tanaka, T. Y., Shimizu, A., and Miyoshi, T.: Data assimilation of CALIPSO aerosol observations, Atmos. Chem. Phys., 10, 2010.

Tegen, I., Harrison, S. P., Kohfeld, K., Prentice, I. C., Coe, M., and Heimann, M.: Impact of vegetation and preferential source areas on global dust aerosol: results from a model study, J. Geophys. Res., 107, 4576, doi:10.1029/2001JD000963, 2002.

Terradellas, E., Nickovic, S., and X. Zhang: Airborne dust: a hazard to human health, environment and society. WMO Bulletin, 64(2), 42-46, 2015.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P., and Levelt, P.: Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview, J. Geophys. Res., 112, 2007.

Tsigaridis, K., Daskalakis, N., Kanakidou, M., Adams, P. J., Artaxo, P., Bahadur, R., Balkanski, Y., Bauer, S. E., Bellouin, N., Benedetti, A., Bergman, T., Berntsen, T. K., Beukes, J. P., Bian, H., Carslaw, K. S., Chin, M., Curci, G., Diehl, T., Easter, R. C., Ghan, S. J., Gong, S. L., Hodzic, A., Hoyle, C. R., Iversen, T., Jathar, S., Jimenez, J. L., Kaiser, J. W., Kirkevåg, A., Koch, D., Kokkola, H., Lee, Y. H, Lin, G., Liu, X., Luo, G., Ma, X., Mann, G. W., Mihalopoulos, N., Morcrette, J.-J., Müller, J.-F., Myhre, G., Myriokefalitakis, S., Ng, N. L., O'Donnell, D., Penner, J. E., Pozzoli, L., Pringle, K. J., Russell, L. M., Schulz, M., Sciare, J., Seland, Ø., Shindell, D. T., Sillman, S., Skeie, R. B., Spracklen, D., Stavrakou, T., Steenrod, S. D., Takemura, T., Tiitta, P., Tilmes, S., Tost, H., van Noije, T., van Zyl, P. G., von Salzen, K., Yu, F., Wang, Z., Wang, Z., Zaveri, R. A., Zhang, H., Zhang, K., Zhang, Q., and Zhang, X.: The AeroCom evaluation and intercomparison of organic aerosol in global models, Atmos. Chem. Phys., 14, 10845-10895, doi:10.5194/acp-14-10845-2014, 2014.

van Leeuwen, P. J.: Representation errors and retrievals in linear and nonlinear data assimilation, Q. J. R. Meteorol. Soc., 114, 2014.

Wang, H. and Niu, T.: Sensitivity studies of aerosol data assimilation and direct radiative feedbacks in modeling dust aerosols, Atmospheric Environment, 64, 2013.

White, B.R.: Soil transport by winds on Mars, J. Geophys. Res., 84, 1979.

Yu, H., Chin, M., Yuan, T., Bian, H., Remer, L. A., Prospero, J. M., Omar, A., Winker, D., Yang, Y., Zhang, Y., Zhang, Z., and Zhao, C.: The fertilizing role of African dust in the Amazon rainforest:A first multiyear assessment based on data from Cloud-Aerosol Lidar and Infrared Path finder Satellite Observations, Geophys. Res. Lett., 42, 2015.

Yumimoto, K. and Takemura, T.: Direct radiative effect of aerosols estimated using ensemble?based data assimilation in a global aerosol climate model, Geophys. Res. Lett., 28, L21802, 2011.

Yumimoto, K., Nagao, T. M., Kikuchi, M., Sekiyama, T. T., Murakami, H., Tanaka, T. Y., Ogi, A., Irie, H., Khatri, P., Okumura, H., Arai, K., Morino, I., Uchino, O., and Maki, T.: Aerosol data assimilation using data from Himawari-8, a next-generation geostationary meteorological satellite, Geophys. Res. Lett., 42, 2016.

Zender, C. S., Bian, H., and Newman, D., Mineral Dust Entrainment and Deposition (DEAD) model: Description and 1990s dust climatology, J. Geophys. Res., 108, 2003.

Zhang J. and Reid, J. S.: MODIS aerosol product analysis for data assimilation: assessment of Level 2 aerosol optical thickness retrievals, J. Geophys. Res.-Atmos., 111, 2006.

Zhang J., Reid, J. S., Westphal, D. L., Baker, N. L., and Hyer, E. J.: A system for operational aerosol optical depth data assimilation over global oceans, J. Geophys. Res., 113, 2008.

Zhang, J., Campbell, J. R., Hyer, E. J., Reid, J. S., Westphal, D. L., and Johnson, R. S.: Evaluating the impact of multisensor data assimilation on a global aerosol particle transport model, J. Geophys. Res. Atmos., 119, 4674-4689, 2014.

Table 1. Characteristics of the simulation runs.

Experiment name	Ensemble configuration	Dust initial conditions at 0 UTC on day 1	Spin-up period	Dust initial conditions at 0 UTC after day 1
Control	No	Cold start	1 month	FC+24 from previous day run
ENS-free-run	Yes	Warm start from Control	11 days	FC+24 of the individual members from previous day run
DA-NRL	Yes	Warm start from ENS-Free-run	None	Analysis at 0 UTC of the individual members from previous day DA cycle
DA-NRL-DB	Yes	Warm start from ENS-Free-run	None	Analysis at 0 UTC of the individual members from previous day DA cycle
AN-initialized	No	Warm start from Control	None	Ensemble mean analysis from DA-NRL-DB

Table 2. Regional domains and respective groups of AERONET stations used for validation purposes

Regional domain (short name)	AERONET stations
Long Atlantic transport (LongAtl)	La_Parguera, White_Sands_HELSTF, Univ_of_Houston
Short Atlantic transport (ShortAtl)	Capo_Verde, Dakar, La_Laguna
Sub-Sahel (SubSahel)	IER_Cinzana, Banizoumbou, Ilorin, Agoufou
Sahara (Sahara)	Tamanrasset_INM
Extended Mediterranean (ExtMediter)	Saada, FORTH_CRETE, Lecce_University, Rome_Tor_Vergata
	Villefranche, Avignon, Evora, Barcelona, Granada
Middle East (MiddleEast)	SEDE_BOKER, Solar Village, Hamim
Central Asia (CenAsia)	*None*
East Asia (EastAsia)	*None*

Table 3. Statistics of departures of first guess and analysis from assimilated observations, calculated for May to August 2007.

Experiment (departures)	Observations	BIAS	RMSE	CORR	FRGE	SD
DA-NRL (FG)	NRL	~~0.074~~ -0.074	0.37	0.59	0.66	0.36
DA-NRL (AN)	NRL	~~0.118~~ -0.118	0.27	0.75	0.54	0.24
DA-NRL-DB (FG)	NRL	~~0.160~~ -0.160	0.35	0.58	0.70	0.31
DA-NRL-DB (AN)	NRL	~~0.169~~ -0.169	0.29	0.72	0.61	0.24
DA-NRL-DB (FG)	DB	~~0.001~~ -0.001	0.35	0.40	0.49	0.35
DA-NRL-DB (AN)	DB	~~0.075~~ -0.075	0.23	0.64	0.35	0.22

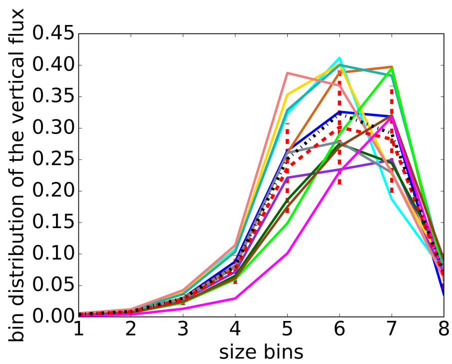

Figure 1. Distribution of the mass vertical flux at sources across the eight dust transport bins for the different ensemble members in different colours, where the bin sizes from bin 1 to bin 8 go from 0.1 to 10 μm with division intervals at 0.18, 0.3, 0.6, 1, 1.8, 3, and 6 μm. The distribution derived from D'Almeida (1987), and used in the standard forecast, is ~~in thick~~ the dashed red line, with horizontal bars indicating the standard deviation of the noise used to create the perturbations. The mean of the ensemble perturbations is the dash-dotted line.

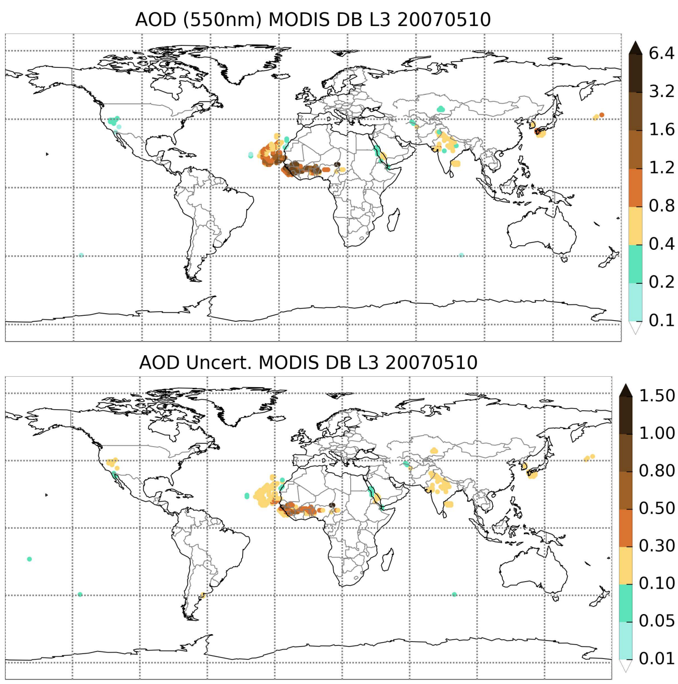

Figure 2. Aerosol optical depth (~~left~~top) and its associated observation error (~~right~~bottom) for May 10 2007 for the NRL MODIS Level 3 product after the application of a filter for dust-dominated conditions.

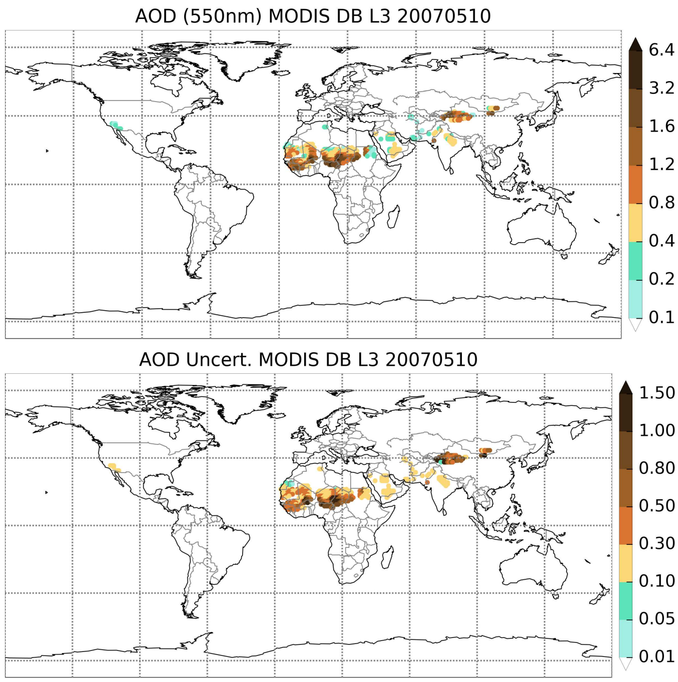

Figure 3. Aerosol optical depth (~~left~~top) and its associated observation error (~~right~~bottom) for May 10 2007 for the MODIS Deep Blue Collection 6 Level 3 product after the application of a filter for dust-dominated conditions.

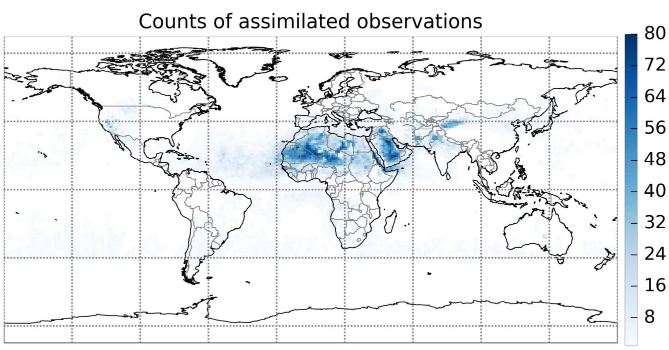

Figure 4. ~~Map~~ Number of ~~AERONET stations~~ NRL MODIS and ~~of the different regional domains used for validation purposes.~~MODIS Deep Blue Level 3 observations assimilated between May and August 2007.

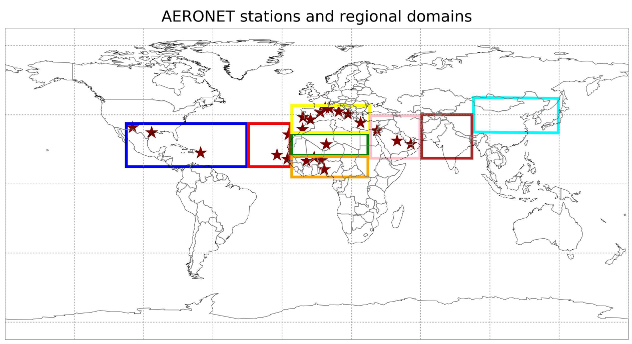

Figure 5. Map of AERONET stations and of the different regional domains used for validation purposes. The regional domains are indicated with different colours: Long Atlantic transport (LongAtl) in blue, Short Atlantic transport (ShortAtl) in red, Sub-Sahel (SubSahel) in orange, Sahara (Sahara) in green, Extended Mediterranean (ExtMediter) in yellow, Middle East (MiddleEast) in pink, Central Asia (CenAsia) in granada, and East Asia (EastAsia) in cyan.

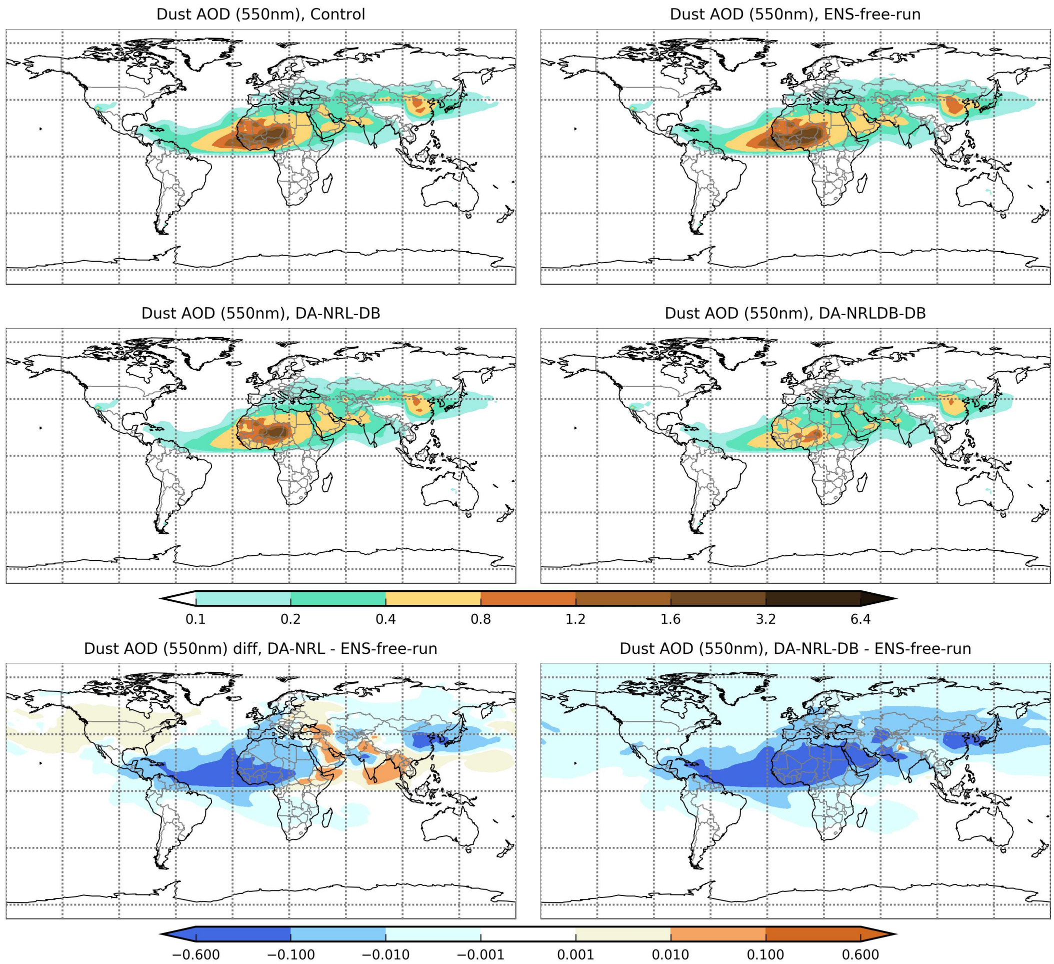

Figure 6. ~~Aerosol~~ Dust optical depth averaged for the month of May 2007 for the Control (top left), ENS-free-run (top right), DA-NRL (~~bottom~~ centre left), DA-NRL-DB (centre right) experiment, and dust optical depth difference between the DA-NRL (bottom left), DA-NRL-DB (bottom right) and the ENS-free-run experiment.

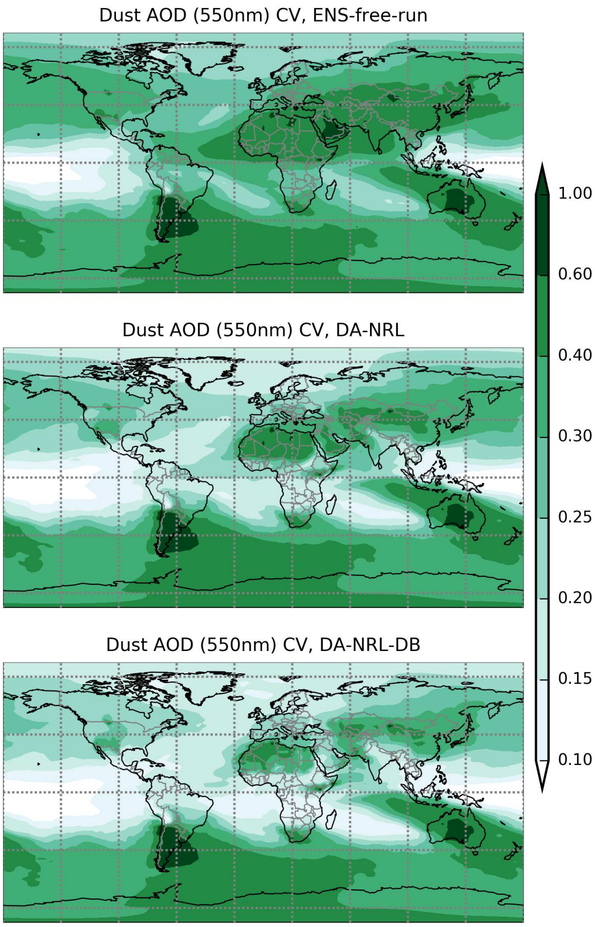

Figure 7. Coefficient of variation for the month of May 2007 for the ENS-free-run (~~left~~top), DA-NRL (centre) and DA-NRL-DB (~~right~~bottom) experiment, when the ensemble is created perturbing the emitted mass vertical flux for each dust bin and the threshold on the friction velocity generating dust horizontal flux.

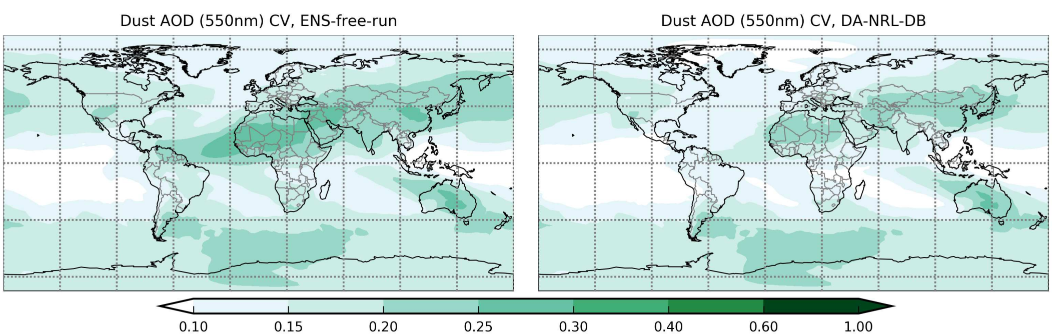

Figure 8. Coefficient of variation for the month of May 2007 for the ENS-free-run (left) and DA-NRL-DB (right) experiment, when the ensemble is created perturbing the emitted mass vertical flux for each dust bin.

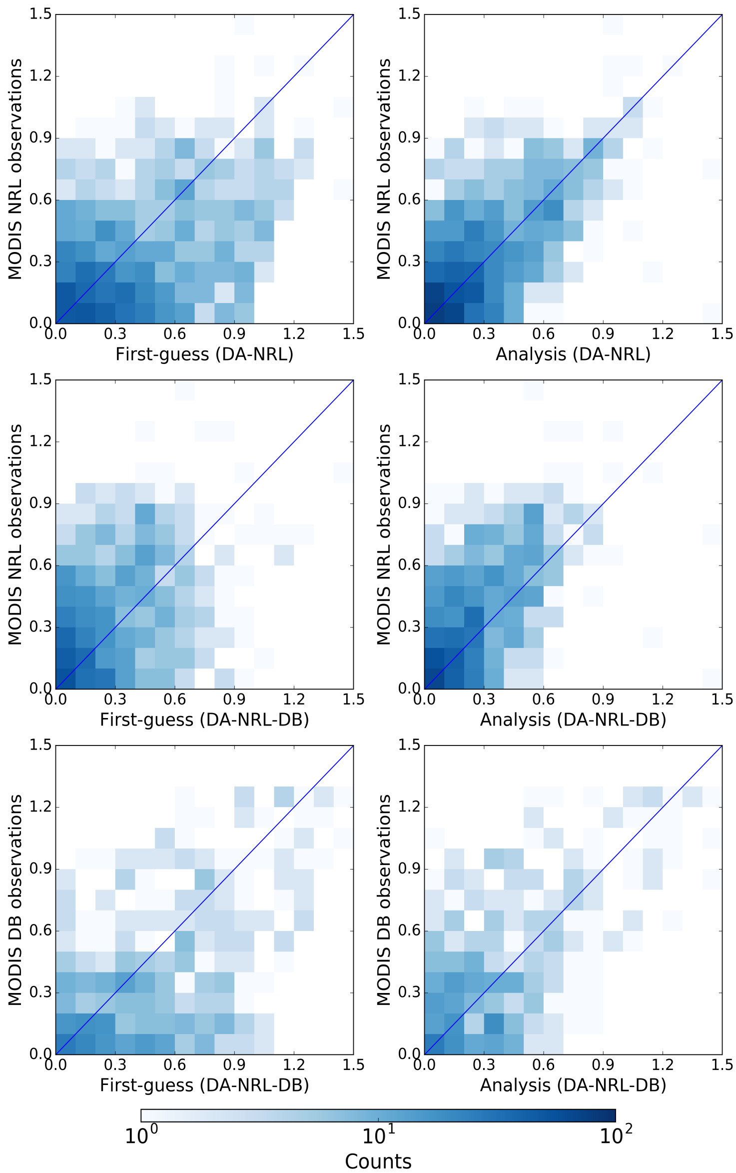

Figure 9. Binned scatter plots of the counts of the logarithm of assimilated observations and first-guess (left plot) and analysis (right plot) for the DA-NRL experiment (top row) and DA-NRL-DB experiment (central and bottom rows), calculated for May to August 2007. A logarithmic scale is used for the counts.

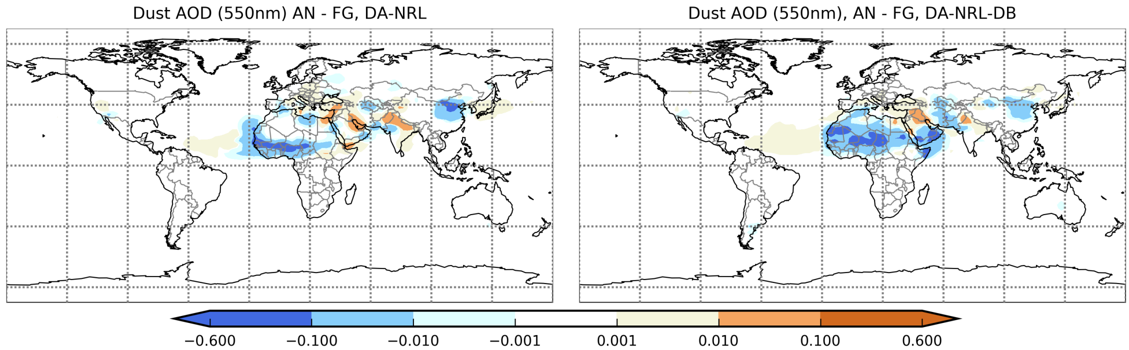

Figure 10. ~~Analysis mean~~ Mean dust AOD analysis increments for May to August 2007 at 12 UTC for the DA-NRL experiment (left) and for the DA-NRL-DB experiment (right).

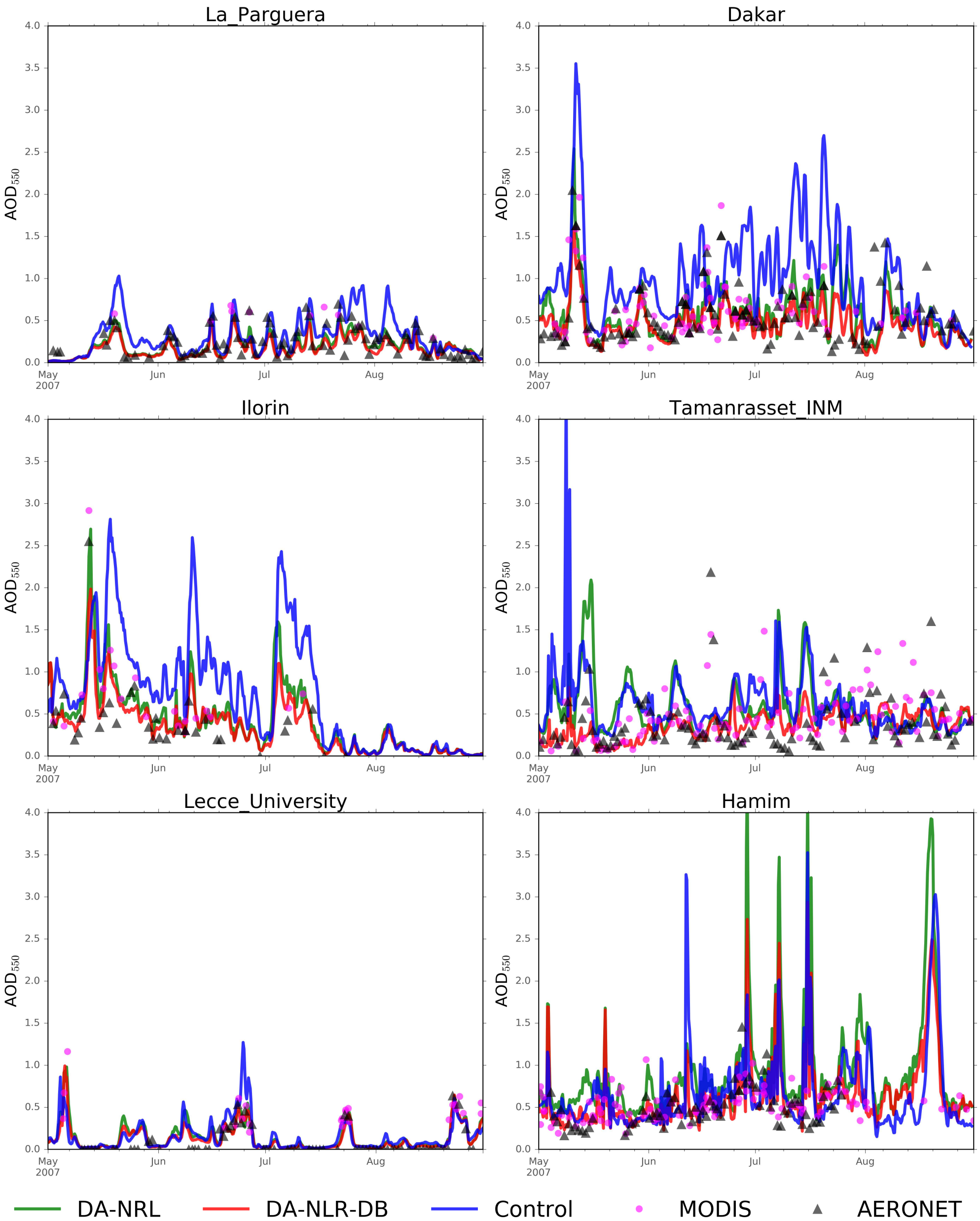

Figure 11. Time-series of AOD values for May and August 2007 in La Parguera (top left), Dakar (top right), Ilorin (centre left), Tamanrasset INM (centre right), Lecce University (bottom left), and Hamim (bottom right) for Control (blue), DA-NRL (green), DA-NRL-DB (red) experiment, for MODIS AOD (NRL and DB; magenta circles), and for AERONET AOD ~~in dust-dominated conditions~~ (black triangles) in dust-dominated conditions. Analysis values are used for the data assimilation experiments.

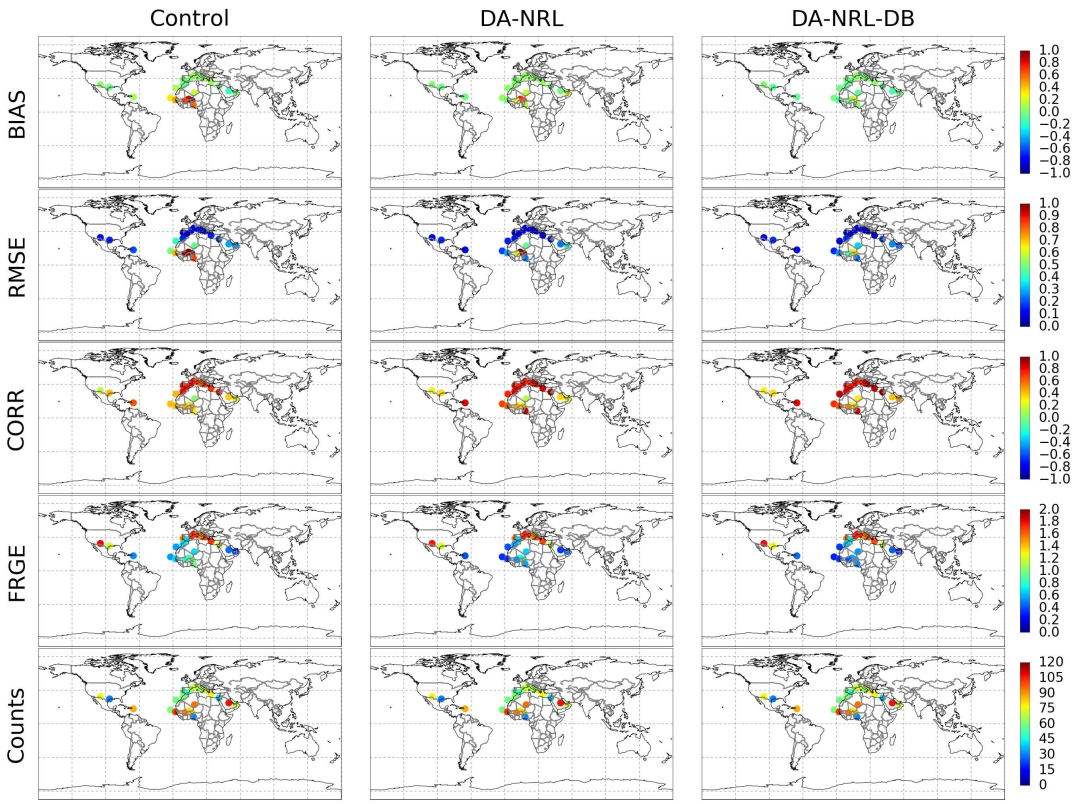

Figure 12. Maps of validation statistics: BIAS, RMSE, CORR, FRGE for the Control (left), DA-NRL (centre) and DA-NRL-DB (right) experiment calculated against AERONET AOD for a selection of stations providing observations during the study period (May to August 2007). ~~The size~~ Maps of the ~~circles is proportional to~~ observation counts used for validation are shown in the ~~number of the available samples~~bottom row.

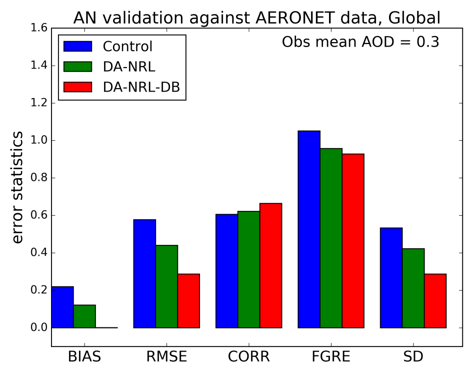

Figure 13. BIAS, RMSE, CORR, FRGE and SD for the Control experiment, for the experiment assimilating MODIS NRL observations (DA-NRL) and for the experiment assimilating MODIS NRL and MODIS Deep Blue observations (DA-NRL-DB) calculated against AERONET observations for all the stations in Figure 5. The dust mean AOD for the observations used for validation during the experiment period is also reported.

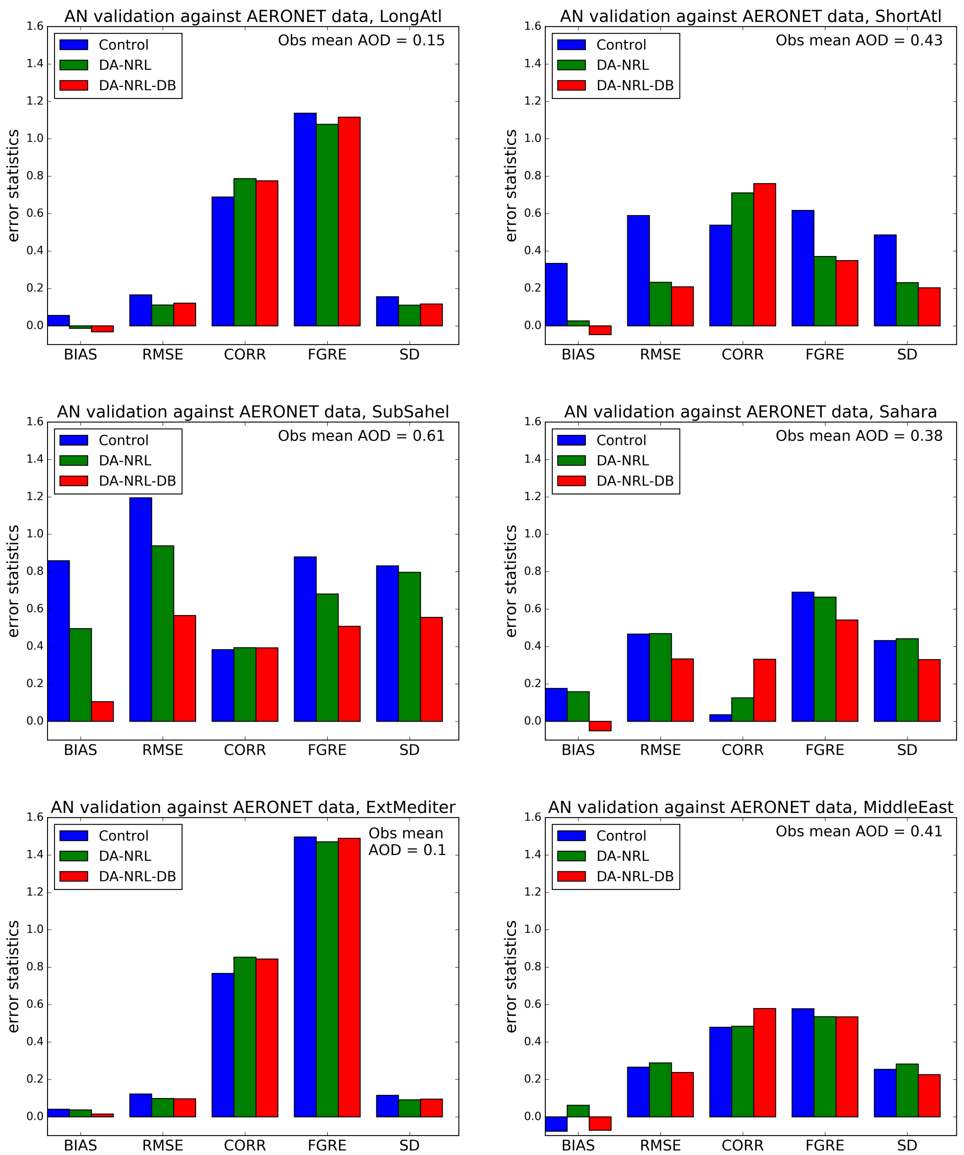

Figure 14. BIAS, RMSE, CORR, FRGE and SD for the Control experiment, the DA-NRL experiments and the DA-NRL-DB experiment calculated against AERONET observations for groups of stations within the regional domains in Figure 5. The dust mean AOD for the observations used for validation during the experiment period is also reported.

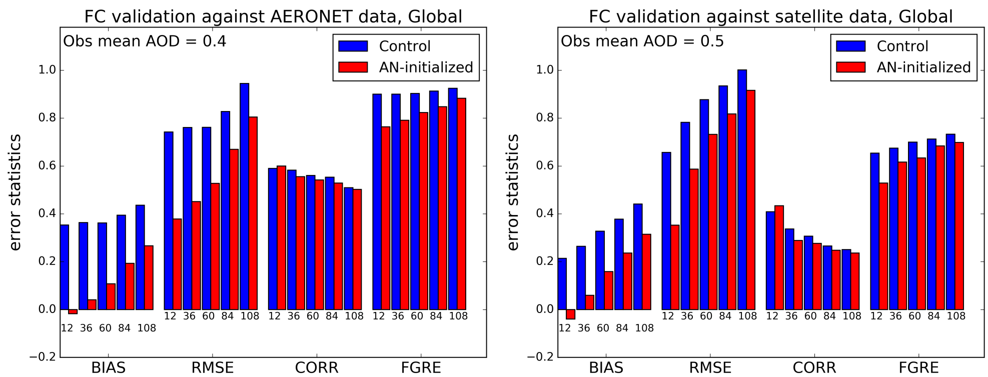

Figure 15. BIAS, RMSE, CORR, ~~FRGE~~ and ~~SD~~ FRGE for the forecast at ~~12~~12, ~~36~~36, ~~60~~60, ~~84~~ 84 and ~~108~~ 108 hours of the Control (blue) and AN-initialized (red) experiment, i.e. the experiment initialized with the DA-NRL-DB analysis, calculated against AERONET observations (left) and against global satellite retrievals, both NRL MODIS and MODIS Deep Blues, (right) filtered for dust-dominated conditions. The AERONET stations are the ones in Figure 5. The dust mean AOD for the observations used to validate the 12 hour forecast during the experiment period is also reported.

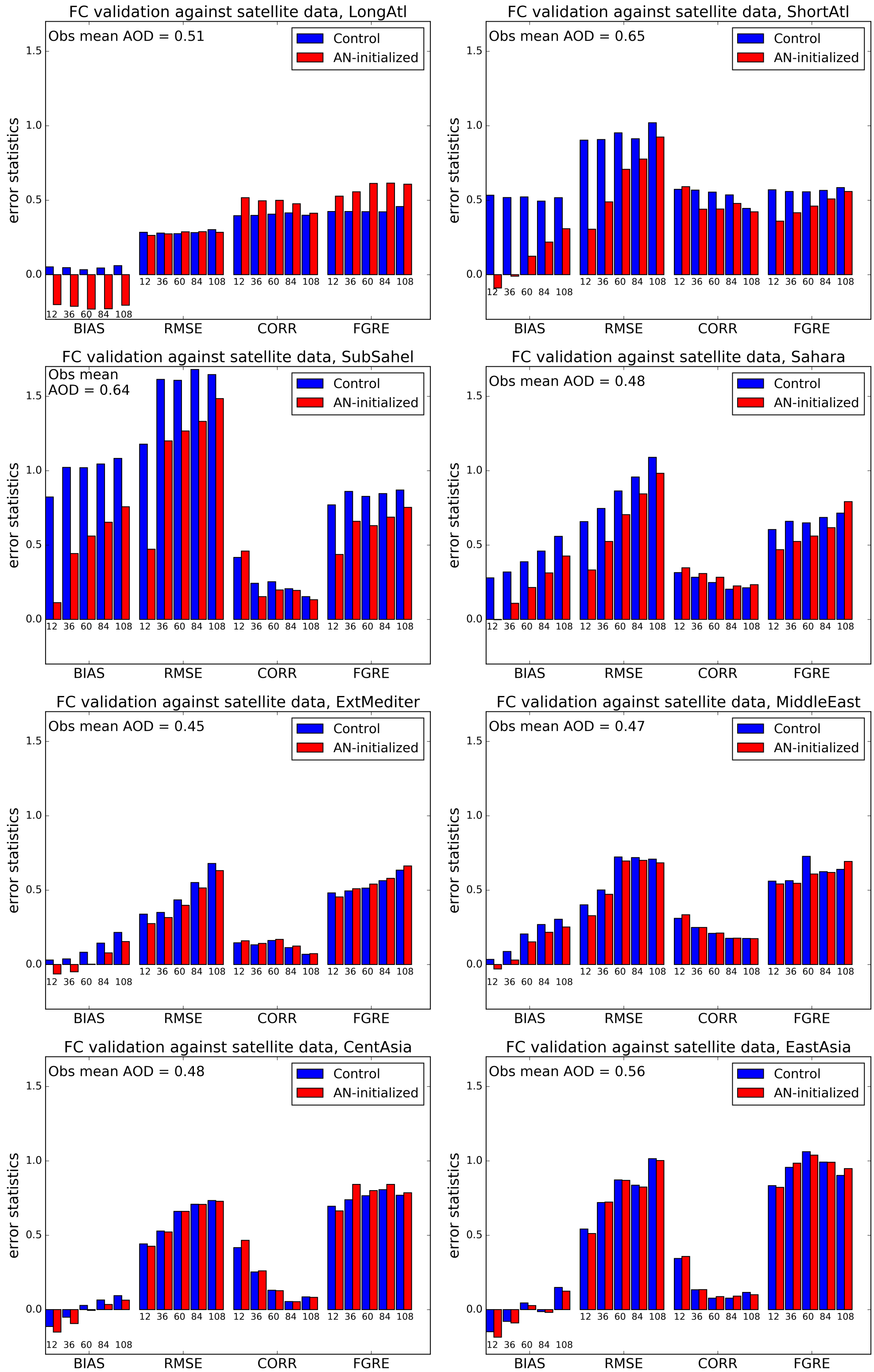

Figure 16. ~~BIAS, RMSE, CORR, FRGE and SD for~~ As the ~~forecast at 12, 36, 60, 84 and 108 hours~~ right panel of Figure 15 but for the ~~Control (blue) and AN-initialized (red) experiment, i.e. the experiment initialized with the DA-NRL-DB analysis, calculated in~~ different regional domains ~~against satellite retrievals (both NRL MODIS and MODIS Deep Blue) filtered for dust-dominated conditions.~~ of Figure 5.