**Atmospheric
Chemistry
and Physics**

# Sensitivity tests for an ensemble Kalman filter for aerosol assimilation

**N. A. J. Schutgens**[1], **T. Miyoshi**[2], **T. Takemura**[3], **and T. Nakajima**[1]

[1]Center for Climate System Research, University of Tokyo, Japan
[2]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, USA
[3]Research Institute for Applied Mechanics, Kyushu University, Japan

**Abstract.** We present sensitivity tests for a global aerosol assimilation system utilizing AERONET observations of AOT (aerosol optical thickness) and AAE (aerosol Ångström exponent). The assimilation system employs an ensemble Kalman filter which requires tuning of three numerical parameters: ensemble size $n_{ens}$, local patch size $n_{patch}$ and inflation factor $\rho$. In addition, experiments are performed to test the impact of various implementations of the system. For instance, we use a different prescription of the emission ensemble or a different combination of observations.

The various experiments are compared against one-another and against independent AERONET and MODIS/Aqua observations. The assimilation leads to significant improvements in modelled AOT and AAE fields. Moreover remaining errors are mostly random while they are mostly systematic for an experiment without assimilation. In addition, these results do not depend much on our parameter or design choices.

It appears that the value of the local patch size has by far the biggest impact on the assimilation, which has sufficiently converged for an ensemble size of $n_{ens} = 20$. Assimilating AOT and AAE is clearly preferential to assimilating AOT at two different wavelengths. In contrast, initial conditions or a description of aerosol beyond two modes (coarse and fine) have only little effect.

We also discuss the use of the ensemble spread as an error estimate of the analysed AOT and AAE fields. We show that a very common prescription of the emission ensemble (independent random modification in each grid cell) can have trouble generating sufficient spread in the forecast ensemble.

## 1 Introduction

Given the perceived inadequacies in current aerosol modelling (Textor et al., 2006, 2007) it is important to develop a framework for improving aerosol models. One way forward is to introduce more sophisticated physics and chemistry into these models (Ghan and Schwartz, 2007), but another approach is to develop aerosol assimilation systems. These assimilation systems would serve several purposes. At first, they would merely combine information from models and observations to arrive at an improved estimate of the aerosol fields. Next, they could be used to estimate various parameters employed in these models (for instance, emission maps). Finally, they could be used to assess model errors.

Aerosol assimilation is a fairly new field that has sofar focused mostly on the first application mentioned in the previous paragraph. Researchers have either assimilated satellite observations (Collins et al., 2001; Yu et al., 2003; Generoso et al., 2007; Zhang et al., 2008; Niu et al., 2008; Benedetti et al., 2009; Sekiyama et al., 2009) or ground observations (Yumimoto et al., 2008; Niu et al., 2008; Lin et al., 2008; Tombette et al., 2009). Those observations were either AOT (and AAE) (Collins et al., 2001; Yu et al., 2003; Generoso et al., 2007; Zhang et al., 2008; Benedetti et al., 2009), LI-DAR extinction or attenuation profiles (Yumimoto et al., 2008; Sekiyama et al., 2009) or surface mass loads (Lin et al., 2008; Tombette et al., 2009).

Among the previously mentioned papers, only Lin et al. (2008) and Sekiyama et al. (2009) used ensemble Kalman filters (Evensen, 1994; Houtekamer and Mitchell, 1998; Whitaker and Hamill, 2002). Recently, we have developed a global aerosol assimilation system (Schutgens et al., 2009, henceforth paper I) using the Local Ensemble Transform Kalman filter (Hunt et al., 2007; Miyoshi and Yamane, 2007; Szunyogh et al., 2008). In an ensemble Kalman filter, an

ensemble of model simulations is used to conveniently represent the model prediction covariant. Among the advantages of ensemble Kalman filters over other approaches are: ease of implementation and realistic flow-dependent model prediction covariance, see also Kalnay et al. (2007). Also, ensemble Kalman filters provide an error estimate of the analysis in the form of the ensemble spread. However, ensemble Kalman filters require tuning of a few (generic) numerical parameters (e.g. the ensemble size and inflation factor) through validation of multiple assimilation experiments. As far as we know, no study of this sort has been published. Lin et al. (2008) mentions an ensemble size of 50 and a value of the inflation paramter from 1 to 640. Sekiyama et al. (2009) used an ensemble size of 20 and an inflation factor of 1.1. Neither reports an attempt at exploring the parameter space of these or other variables.

In this paper, we will attempt to tune several numerical parameters important to our assimilation system by varying their values in experiments and comparing the results to independent data. In addition, we have experimented with different versions of the assimilation system to see how results depend on basic assumptions about the observations, the initial conditions and the number of aerosol modes used.

In Sect. 2 we will describe the assimilation system as used in paper I and the current paper. The range of the sensitivity experiments we performed will be described in Sect. 3, while the experiments will be compared to each other in Sect. 4. Sections 5 and 6 provide independent validation of the experiments with either AERONET or MODIS/Aqua observations. The effect of assimilating AAE observations (in addition to AOT) is shown in Sect. 7. The impact of assimilation on ensemble spread is discussed in Sect. 8. A summary of the paper and its conclusions can be found in the Conclusions.

## 2   The aerosol assimilation system

In paper I we introduced a new global aerosol assimilation system and tested it using AERONET observations. Here we will briefly review its components and some conclusions of paper I. The system consists of a forward calculation of global aerosol transport for 3 h, followed by an assimilation of observations. The assimilation yields an improved estimate of the aerosol distribution, which is then carried forward for another three hours. This sequence is repeated for as long as necessary.

Aerosol transport is calculated by the Spectral Radiation-Transport Model for Aerosol Species (SPRINTARS v3.54) (Takemura et al., 2000, 2002, 2005), an AGCM[1] that includes aerosol physics for four major species: sulfate, carbon, seasalt and mineral dust. SPRINTARS is used at T42 resolution with $20\sigma$-layers in the vertical. Meteorologi-

cal fields are nudged towards NCEP/NCAR[2] reanalysis data (Kalnay et al., 1996). The implemented aerosol physics is quite sophisticated and includes emission (and a limited sulphur chemical cycle), transport, wet and dry deposition processes as well as feedbacks on clouds and the radiative balance. Emission of sulfate and carbon is based on emission maps derived from various datasets. The emission of sea salt and mineral dust depends on parametrisations that include notably windspeed (see Takemura et al. (2000) for more details).

The assimilation system is the Local Ensemble Transform Kalman filter (LETKF) (Hunt et al., 2007; Miyoshi and Yamane, 2007; Szunyogh et al., 2008) that employs an ensemble of model calculations to represent the model prediction covariance. In our case, this ensemble consists of simulations by the same version of SPRINTARS but with different (randomly modified) emissions. Although initial mixing ratios of aerosols were also perturbed, this had little effect due to the short residence times of aerosol in the atmosphere. Since we will only assimilate AERONET data (Holben et al., 1998, see also Fig. 24) no variation of sea salt emission is included. LETKF compares the model state at some time to available observations and adjusts this state accordingly, taking estimates of both model prediction error and observational error into account. During the assimilation, SPRINTARS four major aerosol species are represented by two types, a fine (carbons and sulfate) and a coarse aerosol mode (seasalt and dust).

The assimilated observations are AOT at 675 nm and AAE derived from 440 and 870 nm, as observed by the AERONET network in July 2005 (direct sun algorithm, version 2, lev. 2.0 data). These observations have been averaged over two hours to increase their representativeness at the T42 resolution of SPRINTARS. Special care was taken in estimating the observational error which results from two independent contributions: a retrieval error (we assumed 0.015 for AOT) and a representation error (5−10% for AOT, according to our study in paper I). Zhang et al. (2008) suggest an extra error contribution to the observational error because of likely errors in the observation operator (due to assumptions on particle sizes, refractive indices and hygroscopic growth). Such errors are, however, likely to introduce biases and not random deviations and should not be included in observational error estimates. We accept that this will lead to biases in aerosol mixing ratios, that will likely reduce with further research.

In paper I, we validated our assimilation system by comparing its results to independent AERONET, SKYNET and MODIS observations. The general conclusion was that the system was capable of substantially improving aerosol simulation of AOT and AAE. Our efforts focussed on the feasability and impact of AOT assimilation. It was found, however, that in regions strongly influenced by desert storms, AAE

---

[1]Atmospheric General Circulation Model

[2]National Centers for Environmental Prediction and National Center for Atmospheric Research, located in the USA

**Table 1.** Assimilation experiments used in this paper.

| id | $n_{reg}$ | $n_{ens}$ | $n_{patch}$ | $\rho$ | special | comments |
|---|---|---|---|---|---|---|
| std | | 1 | | | standard SPRINTARS | different emissions |
| R1E2free | 1 | 20 | | | no assimilation | |
| R1E1P4 | 1 | 10 | 4 | 1.1 | | |
| R1E2P2 | 1 | 20 | 2 | 1.1 | | |
| R1E2P4 | 1 | 20 | 4 | 1.1 | | baseline |
| R1E2P6 | 1 | 20 | 6 | 1.1 | | |
| R1E4P4 | 1 | 40 | 4 | 1.1 | | |
| R1E2P4_$\rho$03 | 1 | 20 | 4 | 1.03 | | |
| R128E2free | 1 | 20 | | | no assimilation | |
| R128E1P4 | 128 | 10 | 4 | 1.1 | | |
| R128E2P2 | 128 | 20 | 2 | 1.1 | | |
| R128E2P4 | 128 | 20 | 4 | 1.1 | | |
| R128E2P6 | 128 | 20 | 6 | 1.1 | | $\lambda < 0$ |
| R128E4P4 | 128 | 40 | 4 | 1.1 | | baseline |
| R128E2P4_$\rho$03 | 128 | 20 | 4 | 1.03 | | |
| R128E2P4_$\rho$20 | 128 | 20 | 4 | 1.20 | | $\lambda < 0$ |
| R128E2P4_$\rho$30 | 128 | 20 | 4 | 1.30 | | crashed |
| R128E2P4_2AOT | 128 | 20 | 4 | 1.1 | AOT at 440, 870 nm | $\lambda < 0$ |
| R128E2P4_IC | 128 | 20 | 4 | 1.1 | ensemble spunup | |
| R128E2P4_3modes | 128 | 20 | 4 | 1.1 | SU, CA, coarse aerosol modes | $\lambda < 0$ |

assimilation is vital to obtain a correct distribution of fine and coarse aerosols. We also considered the effect of emission levels on the assimilation, by scaling these differently for three major SPRINTARS aerosol types (sulfate, carbon and minreal dust). We found that these emission levels do not greatly affect the assimilation. Of course, if there are locally periods with no or few observations, no assimilation will occur and large differences between experiments with various emission levels will happen.

## 3 The sensitivity experiments

In this section, we will describe which sensisitivity experiments were performed and why. The why is important, not only to understand some of the results later on but also because it is impossible to explore the full parameter space available to us with limited computer resources. A list of all experiments used in this paper can be found in Table 1. Our baseline experiments (to which we compare the other experiments) used ensemble size $n_{ens} = 20$, local patch size $n_{patch} = 4$ and inflation factor $\rho = 1.1$. Note that in paper I, we used $n_{ens} = 40$ and considered two different scalings for the emission maps (called *E1* and *E2*). Here we only use *E2* where the ensemble mean emisssion of sulfate, carbon and dust is multiplied with factors of 0.5, 2 and 0.5 respectively.

The first parameter we will explore is the region size $n_{reg}$ which defines the size of region within the global grid in which the same random modification factor (per member) is applied to the emission inventories. Standard practice is to set this prameter $n_{reg} = 1$, so each gridpoint has its own, independent modification. Our experiments lead us to believe this is not necessarily an optimal choice, so we also considered $n_{reg} = 128$. In effect, this means that the same random modification (but still different for each ensemble member) is used throughout the whole grid. This will serve to increase the spread within our ensemble (see also Sect. 8). The choice $n_{reg} = 128$ should not be seen as overly restrictive. First, the emission inventories for sulfate and carbon are to a large extent based on inventories for their gaseous precursors. With $n_{reg} = 128$ we effectively create an ensemble of inventories derived with different conversion factors gas-aerosol. Second, in LETKF the analysis is done locally. Although some continuity is imposed, the analysis at two grid points separated by a distance of $2n_{patch}$ is mostly independent (allowing the analysis to favour different conversion factors, so to say).

The second parameter we will explore is the ensemble size $n_{ens}$, which of course governs the accuracy with which the model prediction covariant error is evaluated. Higher $n_{ens}$ means higher accuracy but also higher CPU requirements. Consequently we would like $n_{ens}$ as small as possible.

The third parameter is the local patch size $n_{\text{patch}}$. During the assimilation, only observations within a rectangular box of size $2n_{\text{patch}} + 1$, centered at the analysed gridpoint, will be considered (the rectangular box is not a fundamental limitation of LETKF). Observations outside this box are considered uncorrelated with the analysed gridpoint. In this way, the effect of noise in the ensemble statistics due to its limited size are mitigated. $n_{\text{patch}}$ is therefore related to $n_{\text{ens}}$. In general larger ensembles will allow larger $n_{\text{patch}}$, which means more observations will be considered for the analyses of a gridpoint. Consequently, we would like $n_{\text{patch}}$ as large as possible. Closely linked to the patch size is the horizontal correlation length, which governs the observation localization. We set this length scale to half the local patch size, so 2 in case of $n_{\text{patch}} = 4$. (In this case, the rectangular box has a size of 2500 km at the equator). Note that some ensemble Kalman filters do not use observation localization, but covariance localization instead.

The fourth parameter will be the inflation factor $\rho$. In ensemble Kalman filters, the forecast model covariance is often multiplied with a factor $\rho > 1$ for two reasons. One is that for a limited ensemble ($n_{\text{ens}} < \infty$), the ensemble spread will typically be underestimated and $\rho > 1$ partially corrects for this. The other reason is more fundamental: if the ensemble becomes unrepresentative of the real model prediction error, the assimilation will surely fail. Inflation is a simple technique to force the ensemble to include a sufficiently large part of phase space and not collapse unto an irrelevant subspace (Rodgers, 2000).

These four parameters are generic parameters, in that all ensemble Kalman filters, whether they assimilate aerosols or something quite different, use them. Note that covariance localization and observation localization both define a spatial correlation length scale for the assimilation system.

Finally, we have conducted four experiments where we did not modify the numerical parameters of LETKF, but introduced more fundamental changes in the way the assimilation works. These experiments were often devised after we had studied the previous ones, to resolve some questions we had. Since we usually assimilate observations of AOT and AAE, we decided to conduct one experiment (R128E2P4_2AOT) in which only AOT at multiple wavelengths (440 and 870 nm) were assimilated. In principle, these observations have the same information content as a single AOT and AAE, and we expected similar results. We will later show that this is only partly true. Also, we usually analyse a fine and a coarse mode of aerosol during assimilation. The fine mode consists of both sulfate and carbon aerosol whose scattering properties show different dependance on wavelength. Consequently their contribution to AAE is quite different (see also Fig. 1 in paper I). Therefore, we conducted one experiment (R128E2P4_3modes) where we analysed three modes: sulfate, carbon and coarse aerosol. We also considered the impact of initial conditions on the assimilation. We usually start the ensemble calculation from the same initial conditions but

**Table 2.** Geographic locations.

| Location | longitude | latitude |
|---|---|---|
| North America | 140 W–50 W | 20 N–65 N |
| Europe & North Africa | 25 W–65 E | 00 N–65 N |

with randomly perturbed aerosol mass mixing ratios. The first assimilation occurs after three hours of SPRINTARS calculation. But one might argue it is better to allow the ensemble to evolve freely (i.e. without assimilation) for a longer time, maybe the aerosol residence timescale, so that the correct model covariances can develop. We will show that this is rather inconsequential for the resulting ensemble mean but does have some positive effect on the ensemble spread (R128E2P4_IC).

We will analyse the results of our experiments (conducted for July 2005) in various ways. First, we will compare relevant experiments amongst themselves, focussing our efforts on two parts of the globe where a dense AERONET network is in place, North America and Europe & North Africa (including the Arabian pensinsula), see Table 2. For these areas, we calculate time- or area-averaged differences in AOT and AAE for various experiments. In particular we will use relative RMS differences of two fields $f_1$ and $f_2$ ($\sqrt{\langle (f_2 - f_1)^2 / f_1^2 \rangle}$) since the experiments without assimilation show mostly a bias versus the baseline, while the other assimilation experiments show mostly random deviations. Time-averaged differences are calculated over the period 8–27 July, to exclude potential initial effects.

Second, we will compare selected experiments to observations of AOT and AAE from six AERONET sites that did not yield observations for the assimilation. This constitutes a validation with independent observations. Thirdly, we will compare spatial distributions of AOT as observed by MODIS/Aqua in North America or Europe & North Africa to selected experiments. This also constitutes a validation with independent observations, although we caution against placing too much faith in the MODIS observations over land (we will return to this point later).

Finally, we will consider the ensemble spread and how it is affected by assimilation and various parameter choices.

## 4 Comparison among the experiments

### 4.1 Experiments for region size $n_{\text{reg}} = 128$

We start by comparing experiments with varying ensemble size $n_{\text{ens}}$ and patch size $n_{\text{patch}}$ for $n_{\text{reg}} = 128$ to our baseline experiment R128E2P4. The area-averaged differences with other experiments are shown in Fig. 1. It is obvious that assimilation has a strong impact on simulated AOT, but we also see that there are differences in the assimilated AOT due to
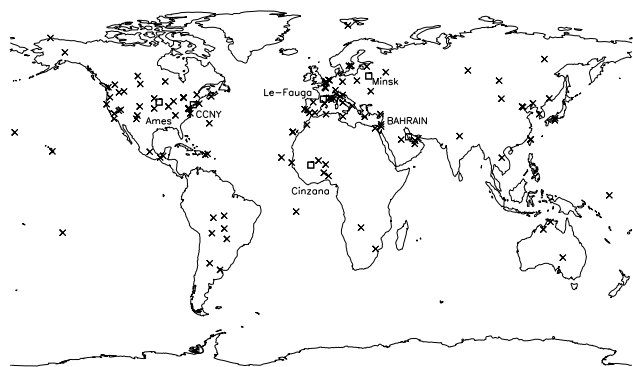
**Fig. 1.** Location of all surface sites used in this study. Crosses: AERONET sites used for assimilation; blocks: AERONET sites used for validation; The names of the validation sites are also indicated.

parameter choices. However, the assimilation results converge unmistakenly as $n_{ens}$ increases, as is to be expected. For North America, it would appear that even a small ensemble of $n_{ens} = 10$ may be sufficient. We will later show that also for Europe & North Africa, $n_{ens} = 10$ gives satisfying results as long as one stays close to AERONET sites. The impact of $n_{patch}$ is more significant, and, more importantly does not show convergence (this we did not expect anyway). As a matter of fact, the R128E2P6 experiment suffered an instability (unrealistic increase in AOT during assimilation) at some point, and so will not be considered any further (it is possible this instability would not happen for a larger ensemble). For North America, the differences between the baseline and $n_{patch} = 2$ seem to grow in time, making it especially important to fix this parameter. Note that such a phenomenon is not seen for the ensemble size.

In Fig. 2 we compare time-averages of the difference between R128E2free, R128E1P4, R128E4P4 and R128E2P2 with the baseline for Europe & North Africa. We see, first, that the effect of assimilation is felt throughout the larger part of the domain. Furthermore, no matter the choice of parameters, similar AOT fields result. Nevertheless, the effect of the parameter value $n_{patch}$ is very pronounced in certain areas, like West Africa for Europe & North Africa. Similar conclusions can be drawn for North America (since the results for North America are very similar for different $n_{reg}$, we will postpone discussion until the next subsection). Finally, we note that the effect of parameter choices is only pronounced away from the AERONET sites, in areas not covered by the network, or near sites which do not contribute a lot of observations due to cloudiness or malfunction (many sites in the north-west of North America and the very north of Europe & North Africa).
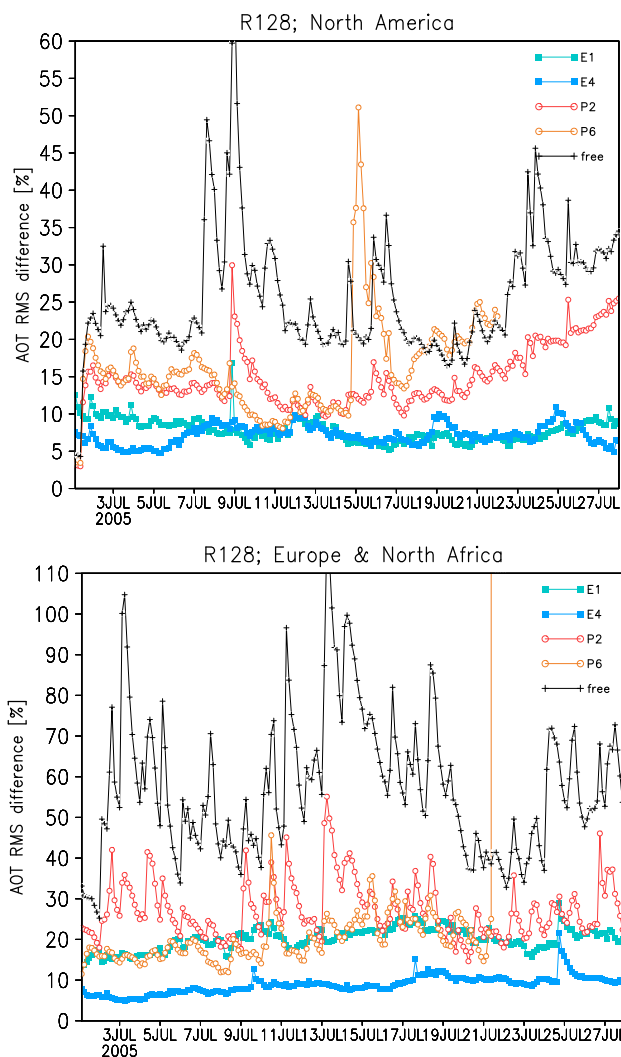


**Fig. 2.** Area-averaged differences [%] in AOT, for $n_{reg} = 128$ experiments. The R128E2P4 experiment is used as the reference. The coloured lines refer to experiments for various ensemble sizes ($E$) and patch sizes ($P$). Notice how the assimilation converges as ensemble size increases. Also notice this is not the case for patch size, which has overall a bigger effect on the results.

### 4.2 Experiments for region size $n_{reg} = 1$

Next we perform a similar analysis for the experiments with region size $n_{reg} = 1$. In Fig. 3 we show the area-averaged differences between for experiments that where compared to the baseline R1E2P4. In this case, the experiment for $n_{patch} = 6$ completed without problems. Note again the convergence of the experiments with increasing ensemble size. Note also that for Europe & North Africa, the impact of ensemble size and patch size seems reduced compared to the experiments for $n_{reg} = 128$. Likely this is because the difference between the free run R1E2free and the baseline R1E2P4 is smaller to start with (compare to Fig. 1).
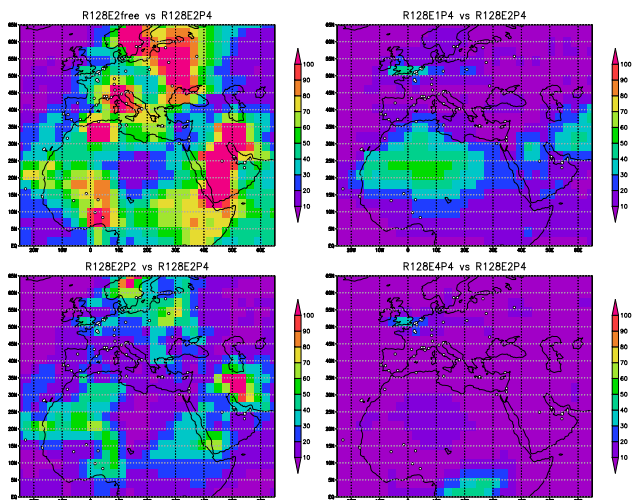
**Fig. 3.** Time-averaged differences [%] in AOT, for $n_{reg} = 128$ experiments. The black & white dots represent AERONET sites.

We also present time-averaged differences for North America for several experiments with $n_{reg} = 1$, in Fig. 4. Again, we see that differences due to varying $n_{patch}$ are most pronounced away from the AERONET sites. Experiment R1E4P4 is not shown as the differences are typically below 20%. For Europe & North Africa, the differences are located in similar areas as for $n_{reg} = 128$ but smaller.

### 4.3 Effect of region size

Later, we will show that $n_{patch} = 4$ is a decent choice for the patch size. For this value, what can we say about the effect of region size on the assimilation? Figure 5 shows the time-averaged differences between R1E4P4 and R128E4P4 (the area-averaged differences are remarkably constant, after a week or so). For North America, these differences are on the order of the differences for the $n_{ens} = 10$ experiments discussed earlier. For Europe & North Africa, $n_{reg}$ has a similar impact as patch size. Again we see that the largest differences are away from the AERONET sites. Clearly, it is not possible to decide in favour of either of these two experiments based on this comparison, but they seem not to differ too much in ensemble mean AOT.

### 4.4 Effect of inflation

For $n_{reg} = 128, n_{ens} = 20$ and $n_{patch} = 4$ we conducted experiments with $\rho = 1.03, 1.10$ (the baseline), $1.20$ and $1.30$ (The experiment with $\rho = 1.03$ was also repeated for $n_{reg} = 1$). The experiment for $\rho = 1.30$ developed an instability: the solution of the Kalman equation contained unrealistically large mixing ratios. In Fig. 6 we show the area-averaged differences between the remaining experiments and the baseline. The experiment with $\rho = 1.20$ resulted in quite unrealistic AOT after some time. For Europe & North Africa,
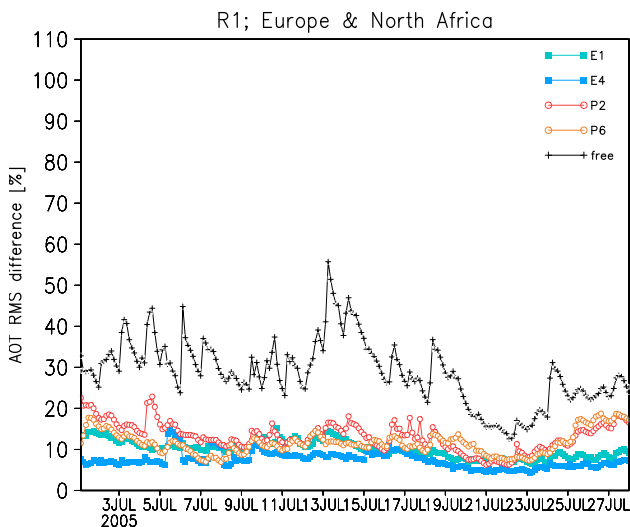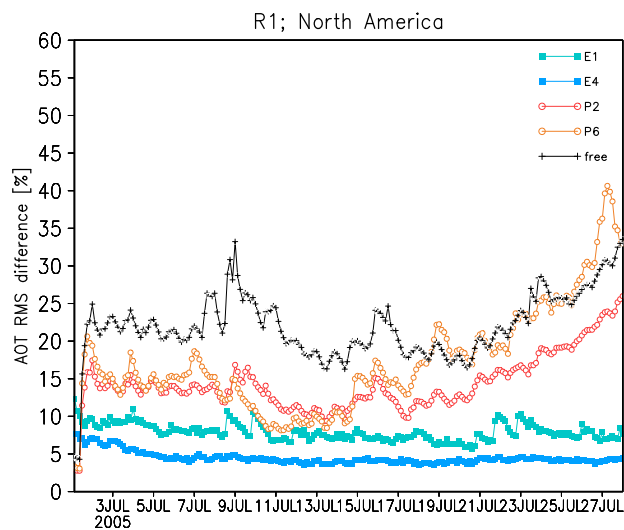


**Fig. 4.** Area-averaged differences [%] in AOT, for $n_{reg} = 1$ experiments. The R1E2P4 experiment is used as the reference. The coloured lines refer to experiments for various ensemble sizes ($E$) and patch sizes ($P$). Notice how the assimilation converges as ensemble size increases. Also notice this is not the case for patch size, which has overall a bigger effect on the results.

R128E2P4_$\rho$03 and R128E2P4 are actually very close and the time-averaged differences are very small. For North America we see larger differences that, moreover, grow in time. Since inflation is done by multiplying the model covariance matrix every time assimilation is performed, growing differences for varying $\rho$ is not really surprising. From Fig. 7 it is clear these differences are located in West Canada, just as in the case of the patch size (see Sect. 4.1). They are likely an edge-effect (note there are no AERONET sites further North).
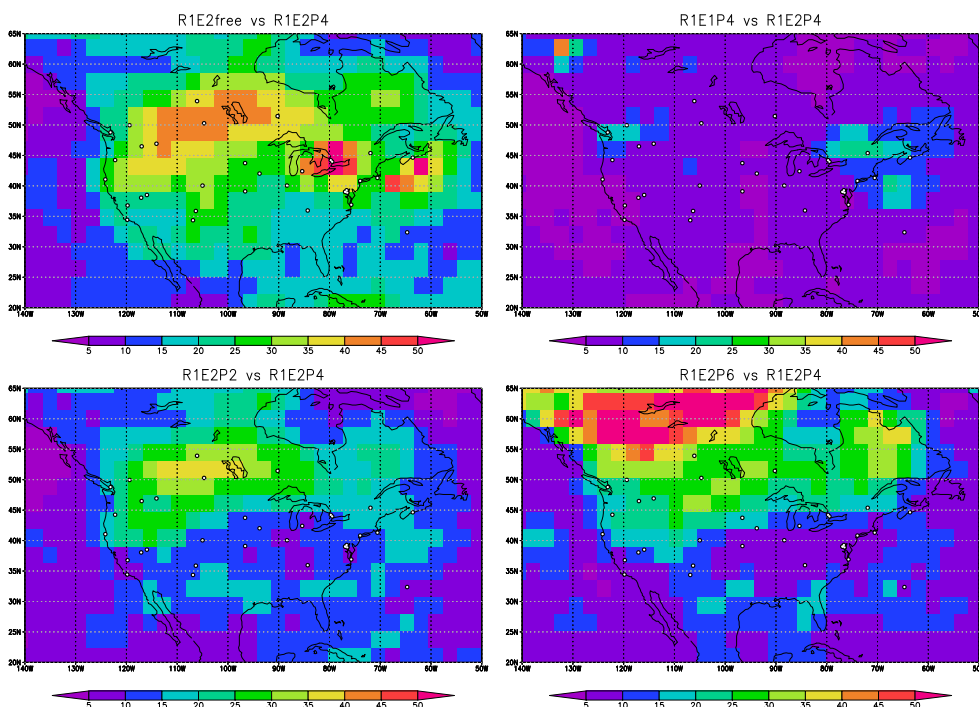
**Fig. 5.** Time-averaged differences [%] in AOT, for $n_{\mathrm{reg}} = 1$ experiments. The black & white dots represent AERONET sites.
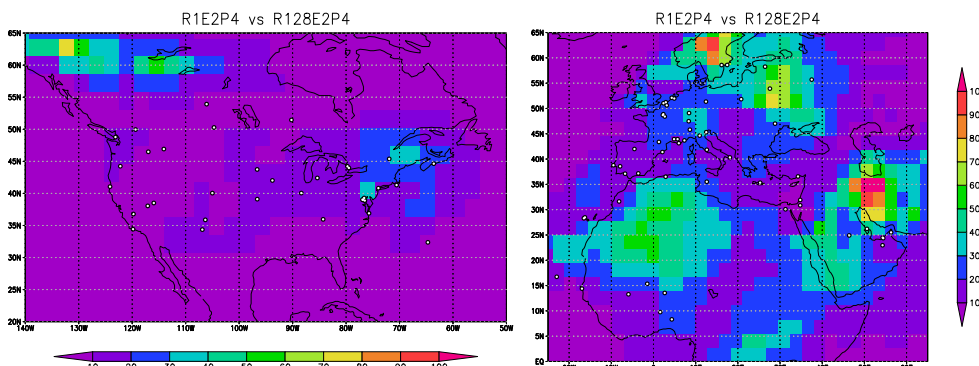


**Fig. 6.** Time-averaged differences [%] in AOT, for $n_{\mathrm{ens}} = 20$, $n_{\mathrm{patch}} = 4$ experiments. The black & white dots represent AERONET sites.
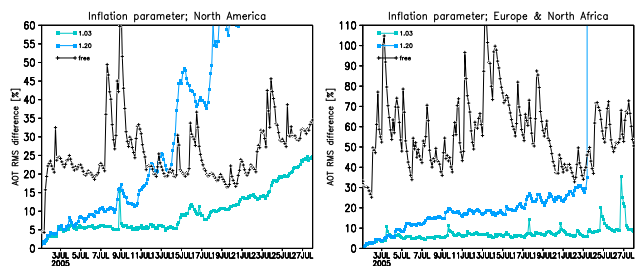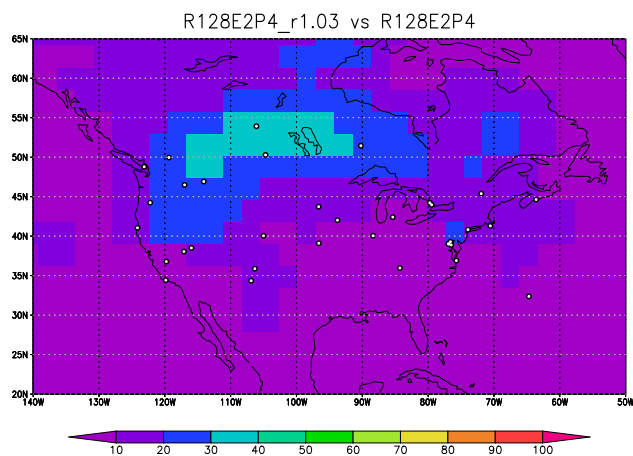


**Fig. 7.** Area-averaged differences [%] in AOT, for various $\rho$ experiments. The R128E2P4 experiments is used as the reference.

### 4.5 Various other experiments

We now come to an interesting section, where we will compare several experiments in which we didn't change the numerical parameters of the assimilation system, but some basic assumptions on how it should work. The nature of these experiments has already been descibed in quite some detail before (Sect. 3). Here we only present the area- and time-averaged differences with the baseline R128E2P4. Area-averaged differences of AOT are shown in Fig. 8. The R128E2P4_2AOT experiment suffered an instability (about which later more). The R128E2P4_3modes experiment, on the other hand, is very similar to the baseline (so no time-averaged differences are shown). The most remarkable
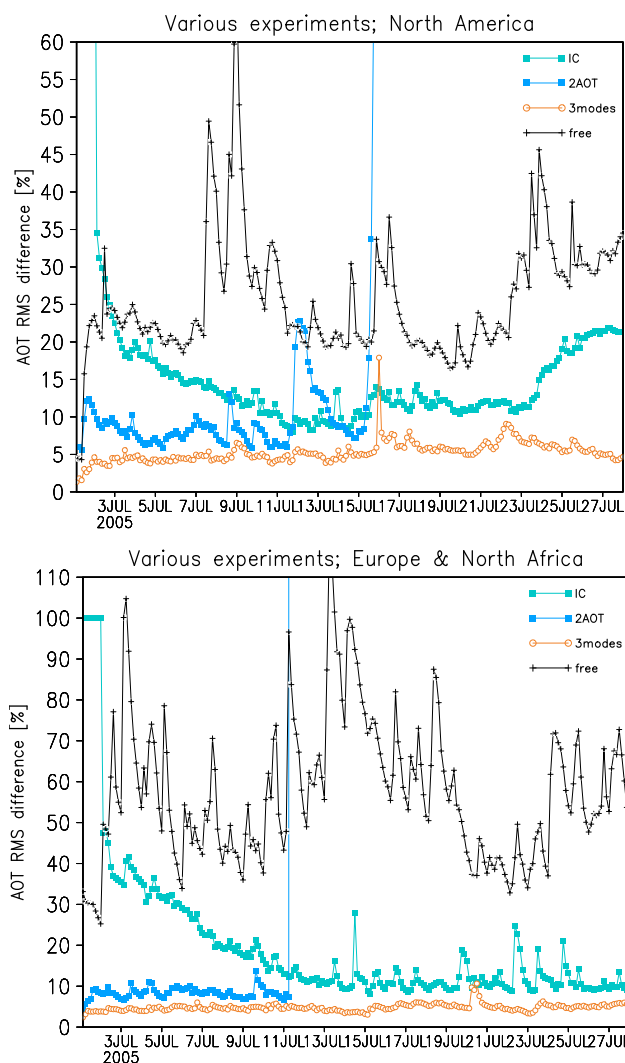
**Fig. 8.** Time-averaged differences [%] in AOT, for $\rho = 1.03$ and 1.10 experiments. The black & white dots represent AERONET sites.



**Fig. 9.** Area-averaged differences [%] in AOT, for various experiments. The R128E2P4 experiment is used as the reference.

experiment is R128E2P4_IC which initially shows very large differences but then rapdily converges unto the baseline. This is most obvious for Europe & North Africa, but it can also be seen for North America (before 25 July, after that we see an increase of differences that we have seen for several other experiments as well). The reason for this convergence is twofold: 1) from 2 July onward both experiments use identical emission maps and the residence time of aerosol in the free atmosphere is short (~ 1 week at most); 2) assimilation forces the experiments closer. The differences in AAE behave similarly to those in AOT and are not shown.

The time-averaged differences of R128E2P4_IC for Europe & North Africa do not reveal anything interesting, but for North America (see Fig. 9) there are localized differences, again in West Canada.

The instability in the R128E2P4_2AOT experiment was quite unexpected: in principle both it and the baseline experiment use identical information. This is borne out by the initial development of both experiments which is very similar, especially when compared to the other experiments. The instability occurs during the assimilation when the analysed mixing ratios for coarse aerosol become locally unphysically large. Although the model transports these large aerosol loads without problems, subsequent assimilation phases trigger additional instabilities in the fine mode. The instability dissappears when a larger ensemble ($n_{ens} = 40$) is used, and seems to point to an issue with the accuracy of the model covariant. We will later show that even for this larger ensemble size, the resulting AAE for two AERONET sites is inferior to the baseline experiment. From the limited experiments we have at our disposal, we surmise that convergence of the filter depends not only on the numerical parameters introduced in Sect 3, but also on the type of assimilated observations. The main difference between these two sets of observations are the correlations (see Fig. 10) among the observations. Due to the non-linear transformation between AOT and AE, it is

expected that even for an exact covariant (i.e. infinite ensemble size) the two experiments will differ, as the ensemble mean observations represent slightly different atmospheric state vectors.

**4.6 Preliminary conclusions**

From our results so far, a few conclusions may be drawn. The first is that an ensemble size $n_{ens} = 20$ seems to be sufficient for an accurate assimilation. The second is that, although other parameters ($n_{reg}$ and $n_{patch}$) affect the results more strongly than $n_{ens}$, spatial patterns in assimilated AOT are fairly similar, especially when compared to the free run of the ensemble. It would appear that, especially close to the AERONET sites, the exact value of $n_{reg}$ and $n_{patch}$ or $\rho$ is not very important. Thirdly, spinning-up the ensemble prior to the assimilation or allowing more freedom in the
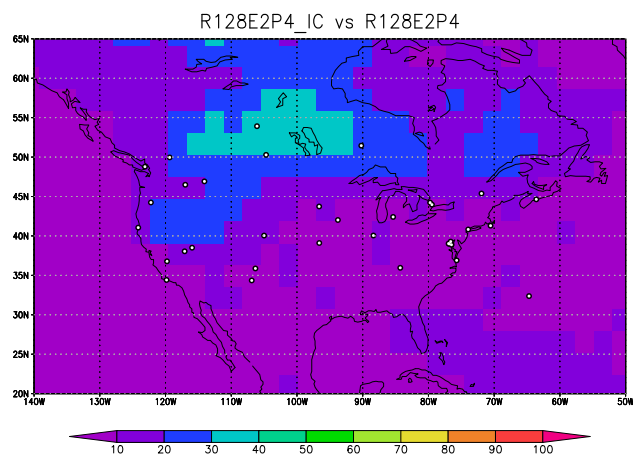
**Fig. 10.** Time-averaged differences [%] in AOT, for the spinup experiment. The black & white dots represent AERONET sites.
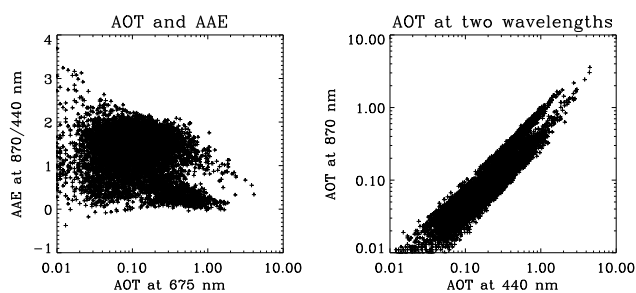


**Fig. 11.** Correlations among co-located observations for the R128E2P4 and R128E2P4_2AOT experiments. Unsurprisingly, strong correlations exist when using AOT at two wavelengths.

representation of AAE (R128E2P4_3modes) really seems to have no big impact. Finally, it appears prudent to choose decorrelated co-located observations over correlated.

## 5 Comparison with AERONET observations

In this section, we will compare results of the various assimilation experiments to AERONET observations at individual sites. We will use several sites around the world (Ames, CCNY, Le Fauga, Minsk, Cinzana, BAHRAIN) that did not provide observations for the assimilation. This comparison therefore uses independent data and can be considered a validation. The same sites, see Fig. 24, were also used in paper I and were chosen as they are in proximity to other AERONET sites that provided plentiful observations to the assimilation (we do not consider the Darwin site here for reasons explained in paper I, Karlsruhe was excluded as it has only observations during two weeks of July 2005).

In Fig. 11 we show AOT observed at these sites and simulated in various experiments. It is obvious that for these sites, the impact of numerical parameter values and other choices

is rather minimal. This is certainly so, if we take the temporal variation in AERONET observations into account. In all cases, the assimilation improves on the standard experiment. We remind the reader that the assimilation experiments use different emission maps than the standard SPRINTARS simulation (see also paper I).

For $n_{\mathrm{reg}} = 1$, both ensemble size and patch size appears to have little impact on the assimilation at these sites. Although in general, R1E1P4 and R1E2P2 deviate a bit from R1E2P4, R1E2P6 and R1E4P4 (that are all very similar) the differences are usually not significant. An exception is the Cinzana site, for which we show AOT and AAE in Fig. 12. Here it seems clear that the better experiment is R1E4P4, and that in particular R1E2P2 gives inferior results.

For $n_{\mathrm{reg}} = 128$, we similarly find that the numerical parameter values of $n_{\mathrm{ens}}$ and $n_{\mathrm{patch}}$ do not affect the assimilation much, except at Cinzana, see Fig. 13. The only firm conclusion that one can draw is that R128E1P4 is clearly inferior to the other experiments. Notice also how R128E2P6 yields similar results to R128E2P4, until just before it experiences the instability.

A comparison for two independent AERONET sites of assimilation results as a function of $n_{\mathrm{reg}}$ is shown in Fig. 14. Although the differences are never very big, it seems that $n_{\mathrm{reg}} = 128$ nevertheless yields better results, especially at the Cinzana site. We do note a strange sudden change in AOT at the Ames site for R128E4P4 (its counterpart R1E4P4 is much smoother), however this may actually be realistic given the rapid changes in AERONET AOT at this time. For the other sites, AOT is more or less similar.

Next, we turn to the $n_{\mathrm{reg}} = 128, n_{\mathrm{ens}} = 20, n_{\mathrm{patch}} = 4$ experiments for various values of the inflation parameter $\rho$. In general, this parameter has little impact on assimilation, with a few noteworthly exceptions. In Fig. 15 we show AOT at Cinzana and Le Fauga. The similarity between the experiments is readily seen. However, it would appear that for $\rho = 1.20$, some episodes (Cinzana on 7 July and Le Fauga on 16 July) are better simulated. It is possible that the larger inflation here partially overcomes limitations due to SPRINTARS model errors. Unfortunately, the $\rho = 1.20$ experiment later on develops an instability with unrealistically high AOT.

Finally, we discuss the various other experiments we conducted: R128E2P4_IC, R128E2P4_3modes and R128E2P4_2AOT. Once more, we are struck by how little impact these choices seem to have on the results. As before, the Cinzana site provides the starkest contrasts. In Fig. 16 we see some initial AOT differences for experiments R128E2P4_IC and R128E2P4_2AOT, while R128E2P4_3modes is indistinguishable from the baseline R128E2P4. Of course, R128E2P4_2AOT developed an instability at some point, but before that AOT is very similar to the other experiments. Interestingly, this is not the case for AAE. It would seem that the baseline experiment yields a better AAE than R128E2P4_2AOT, and this remains true
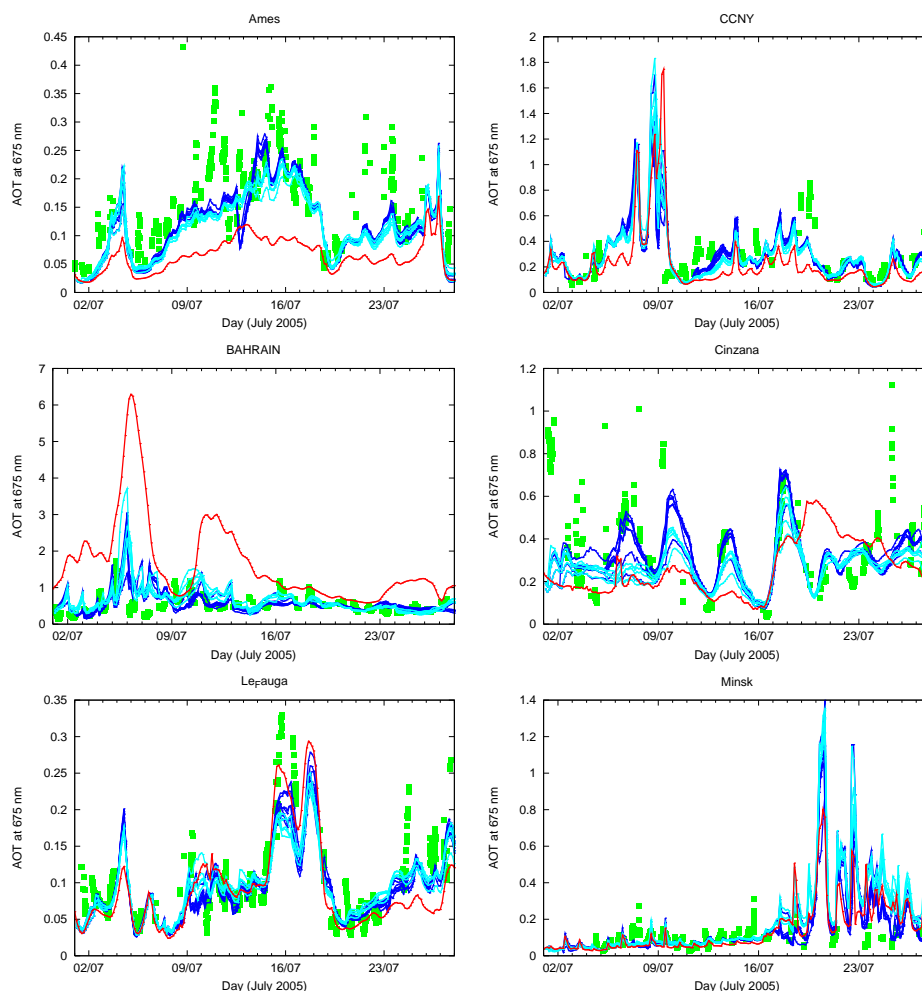
**Fig. 12.** AOT and AAE at selected AERONET sites for various experiments. In red the standard SPRINTARS simulation. In dark blue, experiments for $n_{\text{reg}} = 128$, in light blue, experiments for $n_{\text{reg}} = 1$. Also shown are actual observations (green squares).

even if we double the ensemble size (and remove the instability) in the latter experiment.

### 5.1 Preliminary conclusions

Comparison of the assimilation experiments with independent AERONET observations shows that the analysis is very robust and does not depend sensitively on parameter choices. The validation for the Cinzana site suggests that the region size affects the minimum ensemble size ($n_{\text{ens}} = 40$ for $n_{\text{reg}} = 1$ but $n_{\text{ens}} = 20, 40$ for $n_{\text{reg}} = 128$). Again, the patch size seems to have the bigger impact, with results from the Cinzana site suggesting $n_{\text{patch}} = 4$ or 6 is required. Inflation factors do not seem to matter much, as long as $\rho < 1.2$. Due to the short residence time of aerosol, spinning up the ensemble without assimilation seems not to have much effect. It is preferable to assimilate AAE instead of multiple AOTs, since it allows for a stable assimilation and more accurate AAE fields.
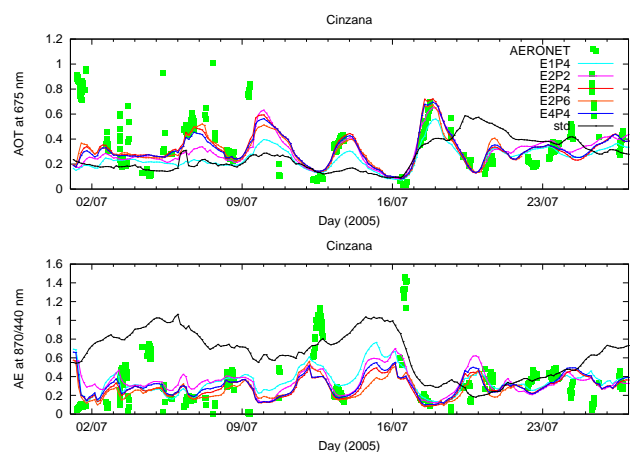
## 6 Comparison with MODIS/Aqua

We now attempt to compare our assimilation experiments to MODIS/Aqua (coll. 5) observations at AOT at 550 nm. Not all our experiments calculated AOT at 550 nm, so we decided to use AOT at 675 nm and extrapolate logarithmically to 550 nm, using the AAE for 870/440 nm. Incurred errors are very small. A much bigger issue is how exactly to compare MODIS observations to these experiments. Both the temporal and spatial sampling are very different for MODIS and SPRINTARS. In addition, MODIS observations are not always available due to cloudiness, or may have been contaminated by small clouds. Finally it is known to be very difficult to retrieve satellite AOT over land (we will mention some problematic cases later on).

We would prefer to compare our experiments to MODIS observations at different times and see whether the assimilation improves on the evolution of spatial patterns in AOT. Although we have attempted this, not much could be learned
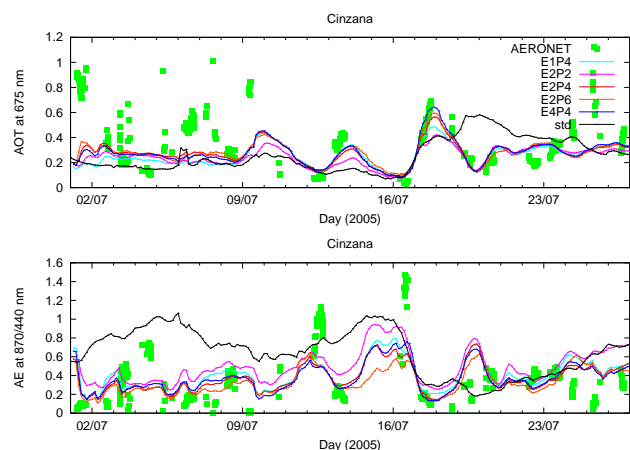
**Fig. 13.** AOT and AAE at Cinzana for various experiments with $n_{\text{reg}} = 1$. Also shown are actual observations (green squares).
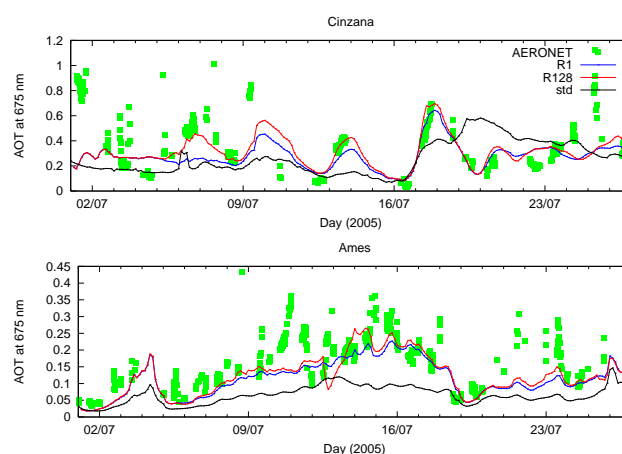
**Fig. 15.** AOT at Cinzana and Ames for experiments R1E4P4 and R128E4P4. Also shown are actual observations (green squares).

**Fig. 14.** AOT and AAE at Cinzana for various experiments with $n_{\text{reg}} = 128$. Also shown are actual observations (green squares).

**Fig. 16.** AOT at Cinzana and Le Fauga for experiments R1E4P4-$\rho$1.03, R128E4P4 and R128E4P4-$\rho$1.20. Also shown are actual observations (green squares).
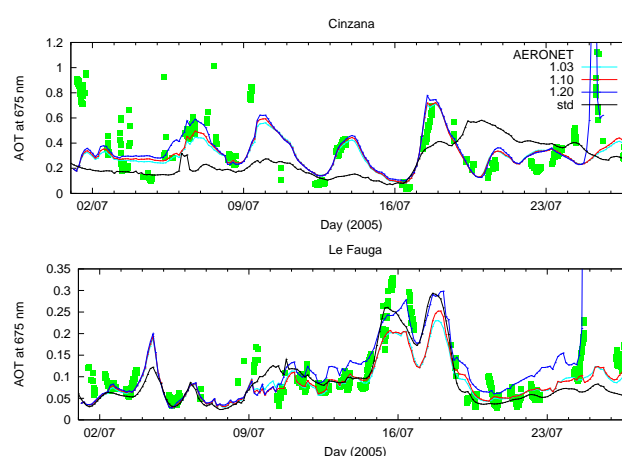
from it for two reasons. One reason is that our experiments are very similar, demanding a highly reliable MODIS AOT to make a meaningful comparison. The other reason is that more often than not cloudiness prevented MODIS from succsesfully retrieving AOT at exactly those locations where the simulations differed most.

In the end, we decided to average AOT for both MODIS and the experiments over 8–27 July. Note that the experiments provide AOT every three hours, while MODIS only once a day. To deal with potential skewing of our average due to cloudiness, we demanded that at least 50% of the days should have a valid observation for the average to be calculated.

In Fig. 17 we show results over North America for MODIS and the experiments std, R128E2P4 and R1E2P6. First, we'd like to note that AOT over the ocean may not be accurate at all, since we have no observations there to assimilate, and

moreover did not create an ensemble emission map for sea salt. Limiting ourselves to AOT over land and high AOT over the ocean (supposedly outflows), we see that assimilation definitely improves AOT. Over the east coast, Central USA, over Hudson bay and Southern Califonia and Baya California AOT is raised to levels more in line with MODIS AOT. The big outflow in front of the coast of Nova Scotia and Newfoundland can unfortunately not be compared to MODIS data due to cloudiness.

We also see two conspicuous differences between the experiments and MODIS. First, there appears to be a "trough" in simulated AOT at and to the south of the AERONET stations at Boulder and Sevilleta. MODIS does not show these low values. On the contrary, just to the South-East of Sevilleta there appears to be an area of high AOT. Just to the west of the trough, over Nevada, MODIS also shows elevated
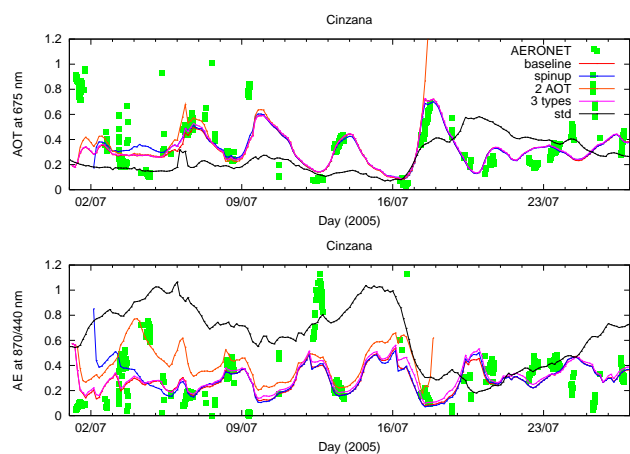
**Fig. 17.** AOT and AAE at Cinzana for various experiments. Also shown are actual observations (green squares).

AOT. Yet when we compare AERONET and MODIS observations for sites like Railroad valley, Sevilleta and Boulder, we see that MODIS greatly overestimates AOT (see Fig. 18). We have not inspected all other sites, but a fair number of them anyway (Bratts Lake, GSFC, Halifax, KONZA EDC, MD Science centre, Missoula, MVCO, Prospect Hill, SERC, Sioux Falls, Walter Branch, Waskesiu). It seems that for sites in Canada and East America MODIS and AERONET generally agree (although there are instances where MODIS overestimates AOT). Finally, the main difference between R128E2P4 and R1E2P6 is the increased AOT that the latter experiment shows in the North West of North America, note that this is also were we saw the largest deviations among experiments in Sect. 4.

A comparison for Europe & North Africa can be seen in Fig. 19. The large dust storms over the Arabian pensinsula are greatly decreased due to assimilation. In Africa, the center of dust AOT shifts from the center in the standard SPRINTARS run to the west, although the intensity differs for R128E2P4 and R1E4P6. MODIS seems to favour the latter experiment. There are a few isolated dust storms visible in MODIS AOT that are absent in all simulations (over the Suez canal, over Niger & Chad and off the West coast near Dakar). They are likely very local phenomena and would require a denser network of AERONET sites to be correctly simulated (in 2009 there are more sites in North Africa than in 2005). The large pollution over North Europe can unfortunately not be verified due to cloudiness. It appears the end of July when many sites in Northern Europe did not yield observations. Consequently this is essentially a free run of the ensemble, be it with initial conditions determined from previous assimilations. The high MODIS AOT for Spain (higher than the analysis) is rather odd, since most Spanish AERONET sites show reasonable agreement with MODIS (the exception is Palencia). There are however instances when MODIS strongly overestimates AOT and we suspect this influences

the mean greatly. Finally, just as for North America, we found a couple of sites where MODIS systematically over-estimates AOT: Palencia and SEDE-BOKER.

## 6.1 Preliminary conclusions

The assimilation clearly improves spatial distributions of AOT, however due to insufficiencies in MODIS observations and the high spatial variability in observed aerosol, it is difficult to prefer one experiment over another. Over Africa, far away from the AERONET sites, the $n_{\text{reg}} = 1$ experiments appear to agree more with the observations than the $n_{\text{reg}} = 128$ experiments.

## 7 Does assimilation lead to improved AAE?

In paper I, we showed that the inclusion of AAE observations potentially has a big impact on resulting AAE fields. This was shown by comparing assimilation experiments without and with AAE as observation (in addition to AOT). At the desert sites Cinzana and BAHRAIN, we saw a significant redistribution of aerosol species to better match the independent AERONET observations. However, for European and American sites, the effect of including AAE was not so obvious. Here we revisit this issue again.

For these American and European sites, the standard simulation often provided a better approximation of AAE than the assimilation experiments. This suggested that assimilating AAE did not positively affect the results. However, one has to keep in mind that the standard simulation and these assimilation experiments were conducted with different emission maps. So it is more reasonable to compare AAE from the assimilation experiments to AAE from the free run (these experiments use the same emission maps). Figure 20 shows AAE for various experiments and we can see that not only at BAHRAIN and Cinzana assimilation improves on AAE. Also for Ames and CCNY there seem to be small improvements. Only for Le Fauga and Minsk is it debatable whether AAE has improved due to assimilation. Note that for the last four sites, AOT is generally low $< 0.1$ and we can consequently expect high errors ($> 0.3$) in both AAE used for the assimilation (from nearby sites) and in the independently observed AAE.

To appreciate the impact of AAE assimilation we show the time-averaged difference in AAE for two experiments from paper I, *A1E2* and *A2E2* (the first experiment only assimilated AOT while the second experiment assimilated both AOT and AAE.). This difference may be interpreted as the impact of assimilating AAE observations on the simulated AAE (Fig. 21). We see that AAE is reduced in the major desert areas, their outflows but also over e.g. continental America.
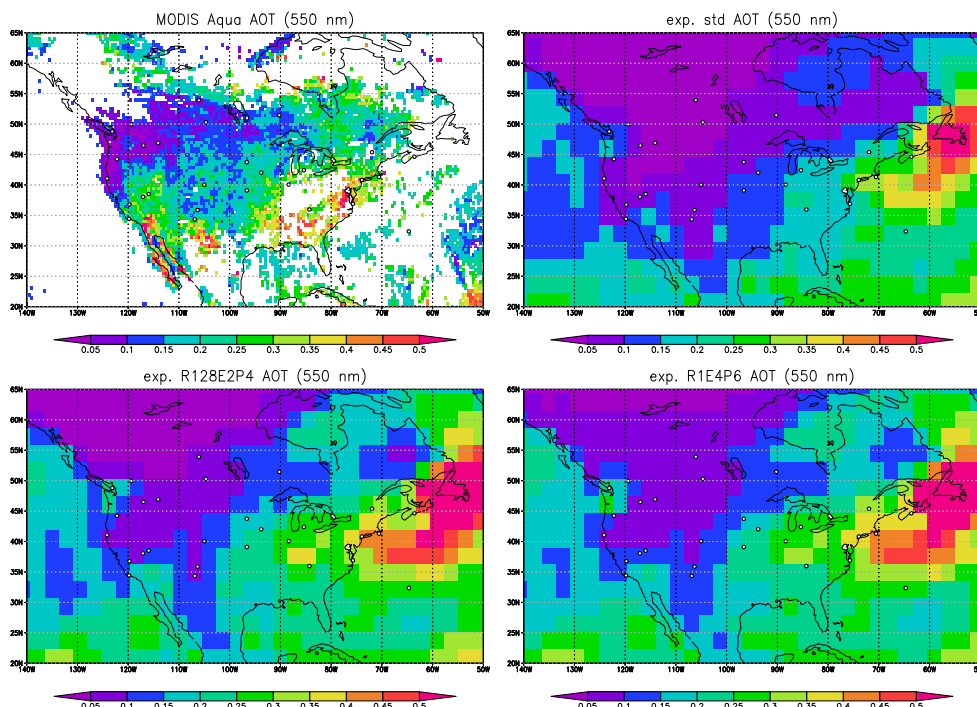
**Fig. 18.** AOT over North America for MODIS Aqua observations and several experiments.
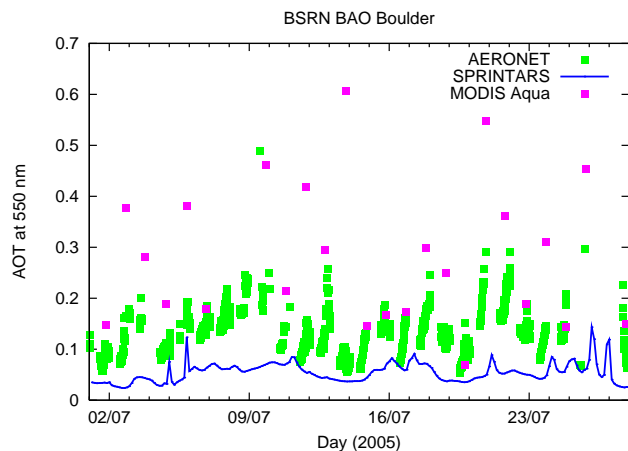


**Fig. 19.** Observed AOT at Boulder for AERONET and MODIS Aqua. Also shown is the standard SPRINTARS' AOT. MODIS clearly overestimates AOT for this site.

## 8 Ensemble spread and the impact of assimilation

Now that we have compared the ensemble mean of AOT and AAE, we will discuss the ensemble spread. Initially, this spread is due to the assumed spread in initial conditions, but as time progresses the spread in emission scenarios becomes the dominant factor. Carbon, sulfate and dust emissions were independently perturbed by multiplying the original emission with a random factor drawn from log-normal distribu-

tions (to ensure positivity) as shown in Fig. 23, (comparison of the *A2E1* and *A2E2* experiments in paper I suggests the non-Gaussian nature does not negatively affect the assimilation). More-over, when observations are assimilated, this will also affect the ensemble spread.

Since our ensemble simulation allows us to calculate a flow-dependent model prediction covariant, the ensemble spread may be interpreted as an error estimate of the analysis, given uncertainties in emission scenarios and observations (assimilation schemes that assume the covariant may not be used in this way). Since we currently have no way of estimating emission uncertainties, it seems best to assume that the emission ensemble yields significantly larger spreads in AOT than the typical observational errors of AERONET. Upon assimilation, this ensemble spread should then decrease, indicating the increased accuracy of the simulated AOT fields.

In Fig. 22 the top panels show the relative spread in AOT for the free runs R1E2free and R128E2free, averaged over the period 8–27 July. We see that the different variations of emission scenarios cause different standard deviations in AOT, with $n_{\mathrm{reg}} = 128$ having a substantially larger spread (for comparison, the estimated observational error in AOT is 5–10%). This relative spread depends mostly on the spread within the emission ensemble. Notice also that over ocean, the spread is very small as we do not perturb the emission of sea salt (over remote oceans the spread in AOT is not zero as the ensemble members will have slightly different wind-fields).
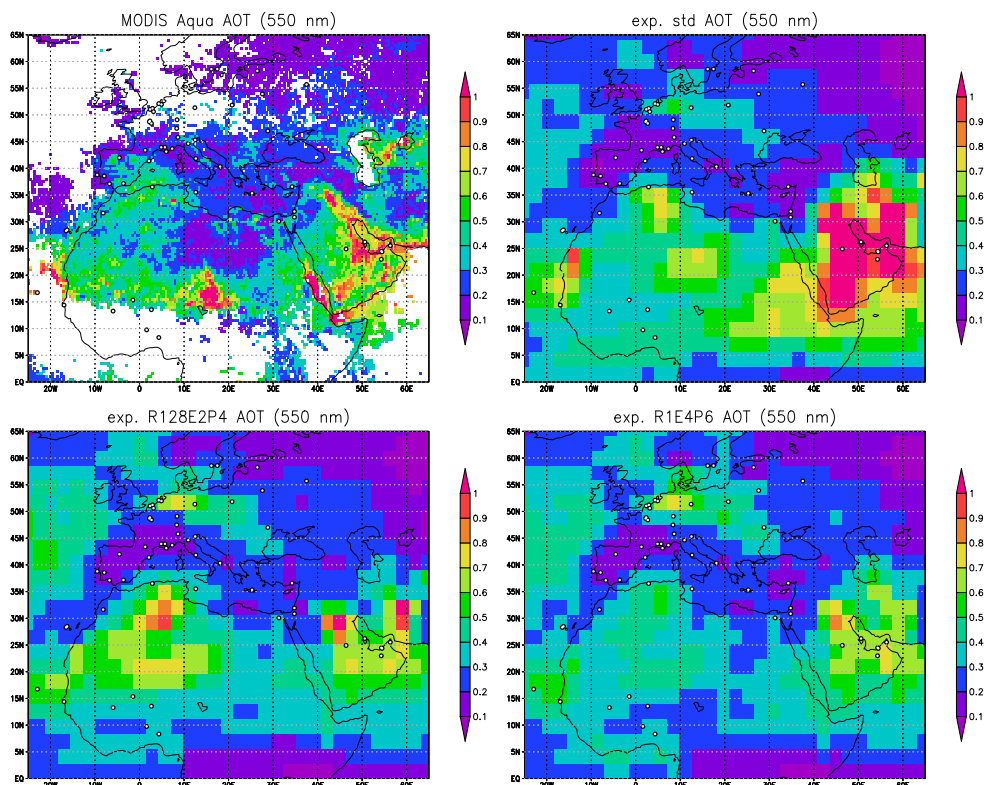
**Fig. 20.** AOT over Europe & North Africa for MODIS Aqua observations and several experiments.

The bottom panels in Fig. 22 show the ratio of ensemble spread of AOT for the baseline and free run experiments. This can be interpreted as the change in ensemble spread due to assimilation. In the right panel ($n_{reg} = 128$), assimilation tends to decrease the spread around AERONET sites and the outflows asscoiated with those regions. Over ocean, the spread does not change, since we do not assimilate observations there. Only over Alaska and Canada do we see an increase in AOT spread due to assimilation. In both the R128E2P4_$\rho$03 and R128E2P4_IC experiments this increase of spread does not occur. Smaller inflation factors naturally lead to smaller spread, while the experiment with spun-up ensemble only includes spread due to the emission ensemble and not due to different initial conditions. It must also be noted that all three experiments R128E2P4, R128E2P4_$\rho$03 and R128E2P4_IC are very similar in their prediction of where the spread significantly decreases.

In the lower left hand panel ($n_{reg} = 1$) we see large areas where the spread increases due to assimilation. Again, this is due to inflation, and the increase strongly reduces for $\rho = 1.03$ (but stays above 1). Also, the area where the spread significantly decreases is rather small. For $n_{reg} = 1$, the original spread in the ensemble (R1E2free) is simply too small to allow much improvement due to assimilation of observations! This conclusion does not change if we increase the spread in the emission ensemble (see Fig. 23) from 1 to 2.

Interestingly, the deviations between assimilation experiments and their baseline usually are on the order of the baseline ensemble spread or much smaller. In this sense, the assimilation experiments can be considered identical, again suggesting robustness of our results.

However, if we want to use the assimilation system to assess not only AOT and AAE but also their errors, it is important that the ensemble spread should not increase due to e.g. inflation. Assimilating (good) observations should improve our knowledge of the aerosol fields, that is: decrease the ensemble spread. For $n_{reg} = 128$, this would suggest using a low inflation parameter $\rho = 1.03$ and a spin-up of the ensemble prior to assimilation.

## 9   Conclusions

In this paper, we discuss a variety of sensitivity experiments for the global aerosol assimilation system which we first described in paper I (Schutgens et al., 2009). These sensitivity experiments are necessary to establish the robustness of the assimilation result and to tune certain numerical parameters that govern the efficiency and/or accuracy of our assimilation scheme. To our knowledge, this is the first time such sensitivity experiments are discussed in the context of aerosol assimilation.
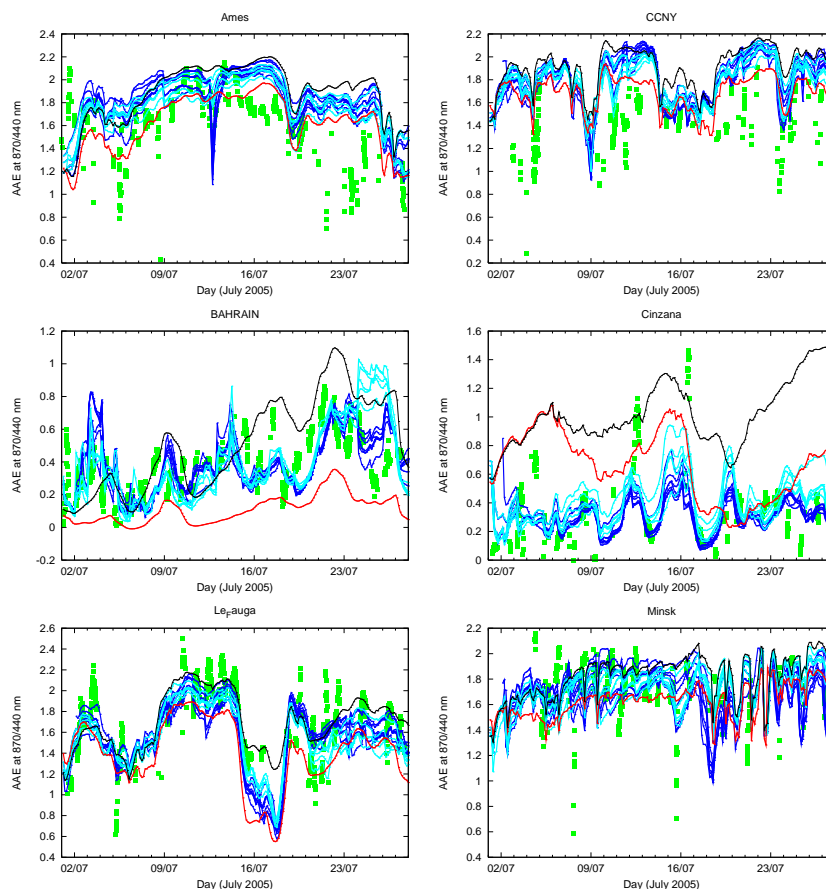
**Fig. 21.** AAE at selected AERONET sites for various experiments. In red the standard SPRINTARS simulation. In dark blue, experiments for $n_\text{reg} = 128$, in light blue, experiments for $n_\text{reg} = 1$. The free ensemble run for $n_\text{ens} = 128$ is shown in black. Also shown are actual observations (green squares).
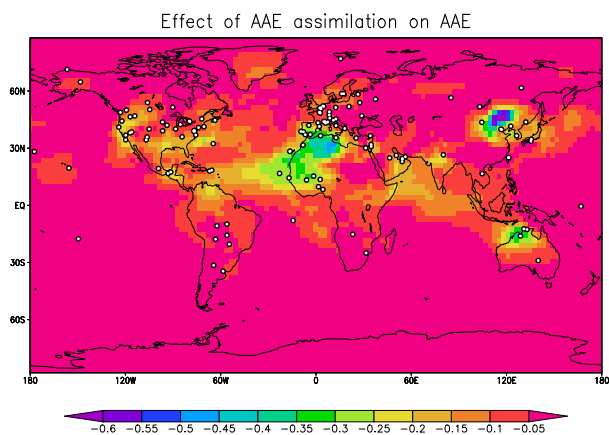


**Fig. 22.** Effect of assimilating AAE observations on AAE simulation. Shown is the time-averaged (8–24 July) difference between the experiments *A1E2* and *A2E2* from paper I.

Our assimilation system is based on the global aerosol transport model SPRINTARS and the Local Ensemble Transform Kalman filter (LETKF). In the present paper it uses AERONET observations, but we have also used observations from MODIS and the SKYNET and CSHNET ground networks. At the present, we prefer to focus on the AERONET observations as it is the most reliable, accurate and world-wide observation set.

The very first conclusion to be drawn is how little these (parameter) choices seem to affect the resulting aerosol fields. Naturally, different values lead to different analyses but the relative differences are only very large when AOT is very small or when one considers areas far away from the observational network. For these areas, it seems more reasonable to call our aerosol fields forecasts than analyses, since AOT at these places is only affected through analyses performed somewhere else.

The second conclusion is that, among the parameter values we explored, the optimal choice is a region size $n_\text{reg} = 128$, an ensemble size $n_\text{ens} = 20$ , a local patch size $n_\text{patch} = 4$ and an inflation factor $\rho = 1.03$. For these values, fast and stable
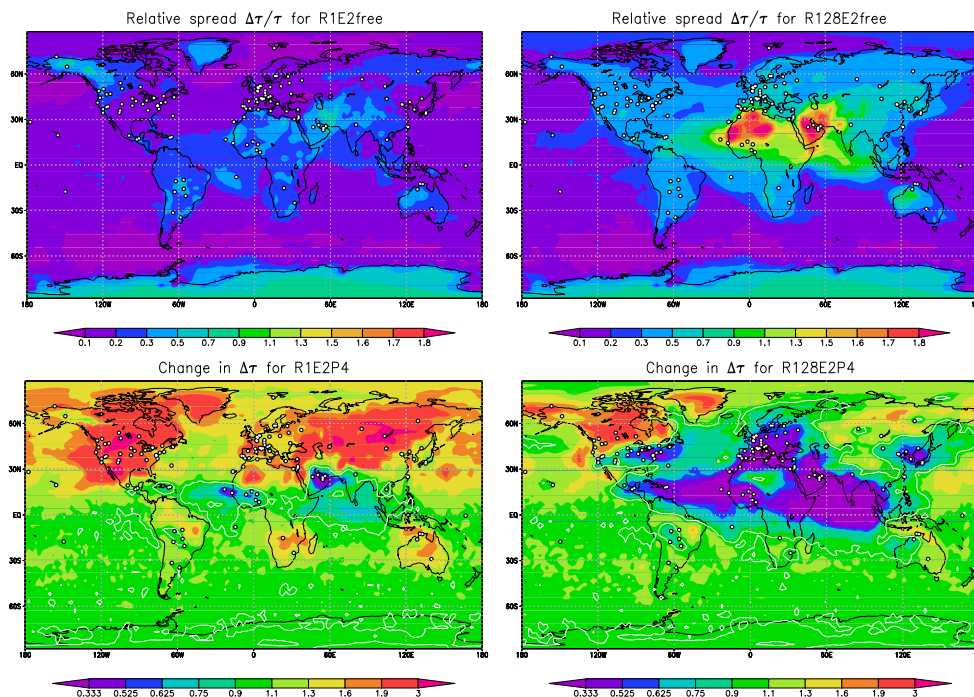
**Fig. 23.** Top panels show the relative ensemble spread in AOT for the $n_{reg} = 1$ and 128 free run experiments. Bottom panels shows the ratio of ensemble spread in AOT for the baseline and free run experiments. The white line is the contour of value 1 (indentical spread).
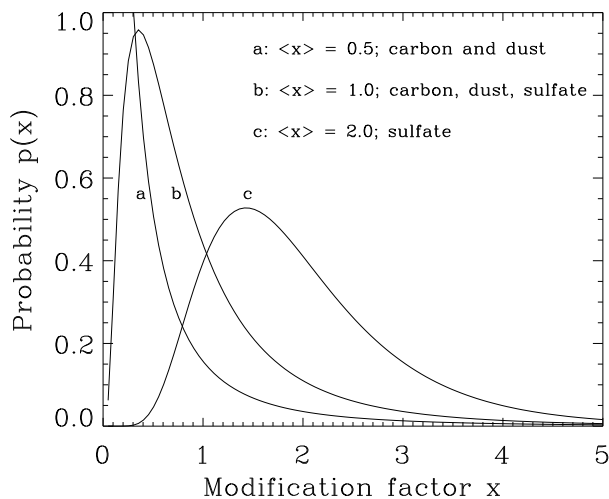


**Fig. 24.** The lognormal distributions that yield the random factors to modify SPRINTARS standard emission scenarios. In all cases the standard deviation is 1. Distributions **(a)** and **(c)** are used in this paper. In paper I, **(b)** was also used.

assimilation experiments whose results compare well to independent observations are possible. Also, the spread in the ensemble may then be used as an indication of the analysis error.

At reduced accuracy, assimilation experiments for $n_{ens} = 10$ seem feasible (although the results wil not have converged properly). This of course significantly affects required computer resources (at $n_{ens} = 20$ the forward calculation of the ensemble takes up the majority of our computational resources). It is of significant interest that the local patch size actually has a bigger impact on the assimilation than ensemble size.

We also suggest that a commonly used emission ensemble, wherein the emission is independently and randomly modified in each grid cell ($n_{reg} = 1$), generates too small an AOT spread in the free run experiment (it is often comparable to observational accuracies). Consequently the uncertainty in AOT due to uncertainties in emissions will be underestimated. Therefore, we suggest a different emission ensemble ($n_{reg} = 128$) which assumes perfect correlation of the modification factors for the grid cells. We would also like to mention that (unpublished) experiments, where we assimilated MODIS observations and validated results with AERONET sites worldwide, showed that $n_{reg} = 128$ is the better choice.

Our tests also revealed that observational datasets with identical information (AOT and AAE versus AOT at two different wavelengths) do not necessarily yield the same results. It seems prudent to use observations that are not highly correlated. Furthermore, describing the analysed aerosol with more than two modes (fine and coarse) when only two independent pieces of information (AOT and AAE) are assimilated does not yield better results. Finally, it seems that initial

conditions have little effect on the analysis beyond a time-scale of about a week.

To what extent may our results be generalised to other ensemble assimilation systems for aerosol? Clearly, this question can only be adequately answered through comparative studies. But here we wish to present some arguments that the present paper provides results generic enough to guide the construction of similar assimilation systems. First, the numerical parameters discussed in this paper are sufficiently generic, allowing that localization is either performed in model space (covariant localization) or observation space (observation localization, this paper). Second, although different models may differ in their AOT fields, the structure of their covariant matrices (under perturbation of emissions) is dominated by transport and should be fairly similar. Third, although the spatial and temporal sampling of the observations may also affect the numerical parameters, recent experiments with MODIS observations suggests otherwise. Finally, unlike weather prediction models, aerosol models do not exhibit chaotic behaviour (if the meteorology is fixed). Hence also the overall robustness of our results for different parameter values.

During our study a dilemma became very obvious: if one assimilates reliable AERONET data, the validation is hampered by lack of sufficiently accurate and widely available independent data. On the other hand, when one assimilates satellite data, one is faced with less reliable observations for the assimilation while being able to use all AERONET sites for validation. Especially over land, MODIS observations were sometimes widely off mark with respect to AERONET, frustrating our attempts at validation. Assimilating such observations would pose a great challenge indeed and require, at the very last, a thorough pre-assimilation vetting of those observations, see e.g. Zhang and Reid (2006). We are currently in the proceess of developing such vetting procedures for MODIS. At the same time, we hope that newer satellites (such as GOSAT) that utilize UV wavelengths (where surface albedo is much reduced) for aerosol retrievals over land will make assimilation of aerosol over land much easier.

## References

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J., Kinne, S., Mangold, A., Razinger, M., Simmons, A., 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.

Collins, W., Rasch, P., Eaton, B., Khattatov, B., and Lamarque, J.-F.: Simulating aerosols using a chemical transport model with assimilation of satellite aerosol retrievals: methodology for INDOEX, J. Geophys. Res., 106, 7313–7336, 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.

Generoso, S., Bréon, F., Chevallier, F., Balkanski, Y., Schulz, M., and Bey, I.: Assimilation of POLDER aerosol optical thickness into the LMDz-INCA model: implications for the artic aerosol burden, J. Geophys. Res., 112, D02311, doi:10.1029/2005JD006954, 2007.

Ghan, S. and Schwartz, S.: Aerosol properties and processes, B. Am. Meterorol. Soc., 88, 1059–1083, doi:10.1175/BAMS-88-7-1059, 2007.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A.,Kaufman, Y., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.:AERONET – A federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, 1998.

Houtekamer, P. and Mitchell, H.: Data assimilation using an ensemble Kalman filter technique, Mon. Weather Rev., 126, 796–811, 1998.

Hunt, B., Kostelich, E., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: a Local Ensemble Transfom Kalman Filter, Physica D, 230, 112–126, 2007.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, B. Am. Meterorol. Soc., 77, 437–471, 1996.

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

Lin, C., Wang, Z., and Zhu, J.: An Ensemble Kalman Filter for severe dust storm data assimilation over China, Atmos. Chem. Phys., 8, 2975–2983, doi:10.5194/acp-8-2975-2008, 2008.

Miyoshi, T. and Yamane, S.: Local Ensemble Transform Kalman filtering with an AGCM at a T159/L48 resolution, Mon. Weather Rev., 135, 3841–3861, 2007.

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, 3473–3482, doi:10.5194/acp-8-3473-2008, 2008.

Rodgers, C.: Inverse methods for atmospheric sounding: theory and practice, vol. 2 of Atmospheric, Oceanic and planetary physics, World Scientific, 2000.

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, 2561–2576, doi:10.5194/acp-10-2561-

2010, 2010.

Sekiyama, T. T., Tanaka, T. Y., Shimizu, A., and Miyoshi, T.: Data assimilation of CALIPSO aerosol observations, Atmos. Chem. Phys., 10, 39–49, doi:10.5194/acp-10-39-2010, 2010.

Szunyogh, I., Kostelich, E., Gyarmati, G., Kalnay, E., Hunt, B., Ott, E., Satterfield, E., and Yorke, J.: A local ensemble transform Kalman filter data assimilation system for the NCEP global model, Tellus, 60A, 113–130, doi:10.1111/j.1600-0870.2007.00274.x, 2008.

Takemura, T., Okamoto, H., Maruyama, Y., Numaguti, A., Higurashi, A., and Nakajima, T.: Global three-dimensional simulation of aerosol optical thickness distribution of various origins, J. Geophys. Res., 105, 17853–17873, 2000.

Takemura, T., Nakajima, T., Dubovik, O., Holben, B., and Kinne, S.: Single-scattering albedo and radiative forcing of various aerosol species with a global three-dimensional model, J. Clim., 15, 333–352, 2002.

Takemura, T., Nozawa, T., Emori, S., Nakajima, T., and Nakajima, T.: Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model, J. Geophys. Res., 110, D02202, doi:10.1029/2004JD005029, 2005.

Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Feichter, H., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Horowitz, L., Huang, P., Isaksen, I., Iversen, I., Kloster, S., Koch, D., Kirkevåg, A., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Liu, X., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, Ø., Stier, P., Takemura, T., and Tie, X.: Analysis and quantification of the diversities of aerosol life cycles within AeroCom, Atmos. Chem. Phys., 6, 1777–1813, doi:10.5194/acp-6-1777-2006, 2006.

Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Chin, M., Dentener, F., Diehl, T., Feichter, J., Fillmore, D., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Horowitz, L., Huang, P., Isaksen, I. S. A., Iversen, T., Kloster, S., Koch, D., Kirkevå g, A., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Liu, X., Montanaro, V., Myhre, G., Penner, J. E., Pitari, G., Reddy, M. S., Seland, Ø., Stier, P., Takemura, T., and Tie, X.: The effect of harmonized emissions on aerosol properties in global models – an AeroCom experiment, Atmos. Chem. Phys., 7, 4489–4501, doi:10.5194/acp-7-4489-2007, 2007.

Tombette, M., Mallet, V., and Sportisse, B.: PM$_{10}$ data assimilation over Europe with the optimal interpolation method, Atmos. Chem. Phys., 9, 57–70, doi:10.5194/acp-9-57-2009, 2009.

Whitaker, J. and Hamill, T.: Ensemble data assimillation without perturbed observations, Mon. Weather Rev., 130, 1913–1924, 2002.

Yu, H., Dickinson, R., Chin, M., Kaufman, Y., Holben, B., Geogdzhayev, I., and Mishchenko, M.: Annual cycle of global distributions of aerosol optical depth from integration of MODIS retrievals and GOCART model simulations, J. Geophys. Res., 108(D3), 4128, doi:10.1029/2002JD002717, 2003.

Yumimoto, K., Uno, I., Sugimoto, N., Shimizu, A., Liu, Z., and Winker, D. M.: Adjoint inversion modeling of Asian dust emission using lidar observations, Atmos. Chem. Phys., 8, 2869–2884, doi:10.5194/acp-7-4489-2007,2008.

Zhang, J. and Reid, J.: MODIS aerosol product analysis for data assimilation: assessment of over-ocean level 2 aerosol optical thickness retrieval, J. Geophys. Res., 111, D22207, doi:10.1029/2005JD006898, 2006.

Zhang, J., Reid, J. S., Westphal, D., Baker, N., and Hyer, E.: A system for operational aerosol optical depth data assimilation over global oceans, J. Geophys. Res., 113, D10208, doi:1029/2007JD009065, 2008.