



Ensemble filter based estimation of spatially distributed parameters in a mesoscale dust model: experiments with simulated and real data

V. M. Khade¹, J. A. Hansen², J. S. Reid², and D. L. Westphal²

¹University Corporation for Atmospheric Research, Visiting Scientist Program, Boulder, CO 80307, USA

²Naval Research Laboratory, Monterey, CA 93943, USA

Correspondence to: V. Khade (vikram@geos.tamu.edu)

Received: 13 September 2012 – Published in Atmos. Chem. Phys. Discuss.: 5 November 2012

Revised: 26 February 2013 – Accepted: 4 March 2013 – Published: 27 March 2013

Abstract. The ensemble adjustment Kalman filter (EAKF) is used to estimate the erodibility fraction parameter field in a coupled meteorology and dust aerosol model (Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS)) over the Sahara desert. Erodibility is often employed as the key parameter to map dust source. It is used along with surface winds (or surface wind stress) to calculate dust emissions. Using the Saharan desert as a test bed, a perfect model Observation System Simulation Experiments (OSSEs) with 40 ensemble members, and observations of aerosol optical depth (AOD), the EAKF is shown to recover correct values of erodibility at about 80 % of the points in the domain. It is found that dust advected from upstream grid points acts as noise and complicates erodibility estimation. It is also found that the rate of convergence is significantly impacted by the structure of the initial distribution of erodibility estimates; isotropic initial distributions exhibit slow convergence, while initial distributions with geographically localized structure converge more quickly. Experiments using observations of Deep Blue AOD retrievals from the MODIS satellite sensor result in erodibility estimates that are considerably lower than the values used operationally. Verification shows that the use of the tuned erodibility field results in better predictions of AOD over the west Sahara and the Arabian Peninsula.

models inaccurate. To obtain the best initial conditions possible, estimation techniques (e.g. data assimilation) are used to combine the state estimates given by the model and those given by the observations. There are a multitude of data assimilation (DA) techniques used in the geophysical community. The first truly operational DA systems have been based on relatively simple 2-D variational techniques Zhang et al. (2008). Apart from 2-D techniques, 4-D variational techniques have been implemented in both research and quasi-operational modes (Wang et al., 2001; Uno et al., 2008; Benedetti et al., 2009). Dubovik et al. (2008) have implemented an inversion technique to retrieve global aerosol source. Perhaps the most promising development for broad applications, however, has been in the application of ensemble-based techniques to not only estimate the state but also to tune aerosol source functions (Lin et al., 2008a, b; Schutgens et al., 2010; Sekiyama et al., 2010; Yumimoto and Takemura, 2011; Huneus et al., 2012). Recently Schutgens et al. (2012) have developed an ensemble Kalman smoother to estimate aerosol emissions.

Ensemble-based assimilation forms an important class of data assimilation methodologies. Ensemble-based DA was introduced into atmospheric and oceanic sciences by Evensen (1994) and Houtekamer and Mitchell (1998). Since then the scientific community has actively researched the theory and practices of ensemble-based data assimilation. The theoretical development includes different formulations of the ensemble filter (Bishop et al., 2001; Tipett et al., 2003; Zupanski, 2005; Hodyss, 2012) and their intercomparisons (Lawson and Hansen, 2004; Lei et al., 2010). Ensemble-based DA has been applied to an entire gamut of atmospheric (Majumdar et al., 2002; Whitaker et al., 2004) and

1 Introduction

Uncertainty in initial conditions, incorrect boundary conditions, and model inadequacies render forecasts of the atmosphere generated using numerical weather prediction (NWP)

oceanographic problems. The performance of ensemble DA in mesoscale models has been investigated by Dirren et al. (2007) by using radio soundings and aircraft observations in the Weather Research and Forecasting Model. Wang et al. (2008) have explored a hybrid DA technique using the WRF model over the North American domain with radiosonde observations. Szunyogh et al. (2008) showed that a global analysis and forecast can be efficiently produced using the parallelized local ensemble transform Kalman filter. Keppenne and Rienecker (2002) have designed and implemented a parallelized multivariate ensemble Kalman filter in an ocean model in the Pacific domain using sparse temperature data.

Apart from incorrect initial conditions, imperfections in model parametrizations are also responsible for inaccurate forecasts. The technique of ensemble-based parameter estimation (Annan et al., 2005) has been employed by numerous researchers as a means of attempting to reduce model error. Ensemble-based parameter tuning, apart from state estimation, is becoming increasingly popular in the estimation community. The ensemble Kalman filter was employed in Aksoy et al. (2006) to estimate multiple parameters in a sea-breeze model. The EnKF was used in Hacker and Snyder (2005) for PBL state estimation by assimilating simulated surface mesonet observations. That work concluded that the PBL state can be effectively constrained by surface observations, thereby reducing forecast errors. The moisture availability parameter was also correctly estimated. Encouraged by these results Hacker and Rostkier-Edelstein (2007) implemented the EnKF to estimate the PBL profiles using real surface observations. It was found that the error could be reduced by up to 85 % compared to the case when data are not assimilated. Model imperfection not only results in significant forecast errors but also distorts the estimates of model predictability (Khade and Hansen, 2004).

The previously mentioned success with ensemble DA methods is suggestive of a number of aerosol-related problems. Aerosol modeling and estimation of uncertainties in its emission and transport is an important subset of atmospheric sciences (Cakmur et al., 2006; Cooke and Wilson, 1996; Lavoue, et al., 2000; de Meij, et al., 2006; Textor, et al., 2007). Already skill improvement in aerosol loadings by ensemble DA techniques is well documented (aforementioned (Lin et al., 2008a, b; Schutgens et al., 2010; Sekiyama, et al., 2010; Yumimoto and Takemura 2011)). A second area of great promise is application to model parameterization problems. Perhaps greatest of these are aerosol source functions, which are widely known to have high uncertainties and often drive significant divergence between aerosol modeling systems. Given the relative simplicity of chemical transformational processes associated with dust relative to other species, as well as its strong, clear and intercontinental signal in remote sensing data sets, dust is an ideal species to examine how ensemble data assimilation can impact not only aerosol loading, but other model parameterizations such as source

functions. Indeed, while commonly used dust models often converge in observables such as bulk regional aerosol optical depth (AOD), there is considerable divergence in lifecycle processes and budgets (Huneeus, et al., 2011).

In this study we perform a series of studies to examine the application of ensemble-based methods to improve model simulations of dust production. Throughout this work Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS) is used in the North Africa/Saharan domain. In this work,

- the ensemble adjustment Kalman filter (Anderson, 2001; Karspeck and Anderson, 2007) is employed within the DART framework (Anderson et al., 2009; Whitcomb, 2008),
- the aerosol state (AOD and the dust concentration) and parameters (erodibility as a proxy for source region) related to dust production are estimated by assimilating observations of AOD, and
- estimation experiments with both simulated and real satellite observations are performed.

In this work the *state* includes the meteorological state and the aerosol state. The meteorological state is temperature, three components of wind speed, and humidity. The aerosol state is dust concentration and AOD. The augmented state is the erodibility. This paper is organized as follows. The model is described in Sect. 2. The tuning experiments using simulated observations are presented in Sects. 3, 4 and 5. Section 3 describes the setup of the simulated data tuning experiments. Section 3 also discusses the tuning of erodibility at a particular grid point in detail. The tuning of erodibility over the whole domain is discussed in Sect. 4. In this section the perturbations in the erodibility at each grid point are assumed to be independent. The case of correlated perturbations in erodibility is considered in Sect. 5. The tuning experiments with real satellite data are described in Sect. 6. The tuned erodibility is used to run verification experiments whose results are presented in Sect. 6. The conclusions of this work are summarized in Sect. 7.

2 COAMPS Mesoscale aerosol model

The meteorological community, over the years, has developed many mesoscale models (e.g., WRF (Skamarock et al., 2005) for researching and forecasting weather phenomenon. COAMPS (Hodur, 1997; Chen et al., 2003) is a mesoscale model used to simulate various atmospheric (Doyle and Bond, 2001) and oceanographic phenomenon (May et al., 2011). It is used not only for basic research, but operationally by the US Navy. The atmospheric model is non-hydrostatic and fully compressible. It allows for nested grids in which the resolution can increase up to a few meters. COAMPS employs staggered horizontal and vertical grids

with a terrain-following sigma Z system for the vertical coordinate. COAMPS includes advanced parameterizations for subgrid scale mixing, radiation, cumulus parameterization, and explicit moist physics. In this work COAMPS is run with 30 vertical levels. The highest model level is at 31 km. Throughout this work the model uses a resolution of 81 km in the horizontal.

Both the research and operational versions of COAMPS includes a dust module to model the generation, transport and physical effects of aerosols particles, including their size and physical transformations (Liu et al., 2003, 2007). The module includes simulation of sinks such as sedimentation, dry deposition and wet removal. The integration of the aerosol module provides outputs of various quantities like mass loading, size distribution, optical depth etc. COAMPS can be used for research purposes.

The details of the COAMPS dust aerosol model are as follows. The vertical dust flux F at a particular grid point (i, j) is given by Westphal et al. (1988) as

$$F_{i,j} = k \times \alpha_{i,j} \times u_{*i,j}^4, \quad (1)$$

where the subscript i, j denotes the latitude and longitude index, respectively.

$$k = 1.42 \times 10^{-5}$$

$\alpha_{i,j}$ is the erodibility, $u_{*i,j}$ is the friction velocity in m s^{-1} .

The dust is generated only if $u_{*i,j} > u_{*ti,j}$, where $u_{*ti,j}$ is a threshold friction velocity.

The amount of dust mobilized depends upon the transfer of (atmospheric) momentum to the earth's surface. This transfer of momentum is proportional to the surface stress τ (Gill, 1982). The friction velocity u_* is related to the surface stress through $u_* = \sqrt{\tau/\rho}$ where ρ is the density. Using theory and experimentation it is shown that the dust flux is proportional to the fourth power of the friction velocity (Gillette and Passi, 1988; Nickling and Gillies, 1993). This proportionality forms the basis of Eq. (1). The dust is mobilized because the surface wind erodes the land surface. Different land surfaces have different susceptibilities to erosion by wind. The susceptibility basically depends on the type of soil covering the land surface. For example, a land surface covered by thick vegetation is less susceptible to erosion than one covered with loose and disturbed soil. The production of dust requires a threshold friction velocity to be reached before dust particles can be lifted from the surface. This threshold friction velocity is represented by u_{*t} . At a given grid point dust is not mobilized for $u_* < u_{*t}$. The values of u_{*t} for various land types have been estimated using field experiment data and laboratory experimentation (Gillette and Passi, 1988). Various modeling studies (Westphal et al., 1988; Liu et al., 2007) use a value of $u_{*t} = 0.6 \text{ m s}^{-1}$ for all land types for simplicity.

Given a particular model grid box, the whole grid box need not be covered by erodible land. Therefore, even if u_* exceeds u_{*t} , only a part of the grid box that is erodible may

emit dust. This is quantified by the erodibility $\alpha_{i,j}$ in Eq. (1) as a spatial weighting function. The erodibility gives the fraction of the grid box covered by dust. At each grid point the erodibility has a value between 0 (no emission) and 1.0 (all emission). In this work α is used to denote the erodibility vector whose components $(\alpha_{i,j})$ are the erodibilities at various grid points. Accurate forecasts of dust production and transport depend critically on an accurate map of erodibility (Liu et al., 2007). The value of the constant k in Eq. (1) is taken from Westphal et al. (1987), which in turn was motivated by Gillette (1981). This value of k is the slope of the linear fit to the scatter plot of experimentally obtained flux data for various values of friction velocity. Since the current study focuses on satellite data assimilation, we model only the actively optical and transportable dust with an assumed diameter of 2 microns for microphysical purposes.

The amount of dust in the atmosphere is quantified by the dust concentration (c_m) in $\mu\text{g m}^{-3}$. The AOD is another measure of the amount of dust. The AOD at a particular grid point i, j is obtained by vertically integrating the dust light extinction over the atmospheric column, which is simply defined here as the mass concentration times a mass extinction efficiency (a_e) taken as $0.5 \text{ m}^2 \text{ g}^{-1}$.

$$\text{AOD}_{i,j} = \int (a_e c_{m,i,j}) dz,$$

where z is the height. Hence, here we are assuming that AOD is linearly proportional to total mass concentration. In reality, dynamics of dust particle size, especially large particles near sources can be quite complicated. However, for the purpose of this work this assumption is valid because we want to tune the dust emitting areas to the first order.

The dust generated at various locations in the domain is mixed vertically and advected horizontally. The dust in the atmosphere at a given grid point and vertical level is due to local generation and that advected from upstream areas. The share of the advected and local dust in the total dust depends on meteorological conditions, specifically the wind field. The amount of local dust depends on the erodibility and friction velocity at that grid point. It is possible that for a particular grid point at a particular time at some vertical levels the advected dust dominates, while at other levels the local dust constitutes the major portion of the dust. In general the total dust contains contributions from local production and dust transported from other areas. Since AOD is the vertical integral of dust concentration, the total AOD at a grid point has contribution from local and transported dust.

The AOD at a particular grid point can be expressed as

$$\text{AOD}_{i,j} = \text{AOD}_{i,j}^{\text{local}} + \text{AOD}_{i,j}^{\text{transport}}.$$

The transported AOD is due to dust that is produced in upstream regions and advected by winds. The local generation is given by the dust flux $F_{i,j}$. Therefore,

$$\text{AOD}_{i,j} = \int \int (a_e F_{i,j}) da dt + \text{AOD}_{i,j}^{\text{transport}}.$$

The horizontal area element is represented by da and the time step is given by dt . Substituting for the dust flux from Eq. (1), $AOD_{i,j}$ can be expressed as

$$AOD_{i,j} = \int \int \left(a_{ek} \times \alpha_{i,j} \times u_{*i,j}^4 \right) da dt + AOD_{i,j}^{\text{transport}}. \quad (2)$$

Though the sink term is not mentioned in these equations, the actual model calculates the removal of the aerosol. The dust that is emitted locally but not advected away is included in the transport term. Equation (2) decomposes the total AOD into the local and advected component. This decomposition is central to the understanding of the tuning of erodibility as will be evident in Sect. 4. The erodibility plays an important role in the calculation of AOD. The determination of the value of α at various locations on the earth is a formidable task. Many researchers (Westphal et al., 1988; Tegen and Fung, 1994; Park and In, 2003; Walker et al., 2009) have made significant efforts to produce maps of α for important dust producing regions of the earth. The efforts made by these researchers involve the analyses of different types of landforms and the variation of their properties with season etc. These efforts involve the visual inspection of atlases and also observations of AODs.

In the current work we aim to use satellite observations of AOD to estimate α in the North African region by employing an ensemble Kalman filter based estimation approach. Note that the satellite observations of the total AOD, that is the left hand side of Eq. (2) are available. Observations of local and transported AOD are not separately available. In the next section we describe the estimation experiments with simulated AOD data.

3 Observation System Simulation Experiment (OSSE)

The ultimate objective of this work is to improve the forecasts of AOD over the Sahara by tuning α using satellite observations of AOD. However, prior to performing experiments with satellite data, an Observation System Simulation Experiment (OSSE) is performed. OSSEs are important tools to assess the amenability of a model to tuning. The OSSE uses simulated observations drawn from the *perfect model experiment*. A particular set of values of erodibility are defined to be correct (or perfect). Observations of AOD are drawn from a model run using the defined correct values of erodibility. An imperfect model is defined by choosing values of α different from the perfect model values.

The meteorological boundary and initial conditions are obtained from Navy Operational Global Atmospheric Prediction System (NOGAPS) global model (Hogan and Rosmond, 1991). Ensemble analysis boundary conditions are used every 6h. These ensemble analysis are obtained by the local Ensemble transform technique (McLay, et al., 2010). The ensemble analysis is used as initial conditions so that each ensemble member is a different realization of meteorology.

Since each ensemble members corresponds to a different realization of initial and boundary conditions the advection (that is wind) is different for each ensemble member. The resulting spread in the boundary layer wind is of the order of 0.7 m s^{-1} . For the lateral dust boundary conditions we have assumed that dust does not enter the domain, which is quite large. For the period of our study there is no dust storm east of the Arabian Peninsula. So these dust boundary conditions approximately hold. This approximation may impact the estimation results in the real data experiments but it does not impact the OSSE results in any way.

The perfect model experiment uses a particular ensemble member of meteorology. The AOD *observations* from the assumed perfect model are assimilated into the imperfect model and the following question is posed: Are the perfect values of erodibility recovered by the ensemble-based tuning? Because one has defined the imperfect model to be different from perfect model only in α , the OSSE represents the best-case scenario. Compared to nature the model has many errors apart from imperfect α . Hence, if the perfect α values are not recovered in the OSSE, then α will certainly not be tuned correctly using real data.

The OSSE is run using the meteorology of June/July 2009, corresponding to the well-known peak in the Saharan dust production and its westward transport into the subtropical Atlantic Ocean. Throughout this work the meteorological state is not estimated. The boundary conditions used contain observational information. The operational values of α over the Sahara domain are shown in Fig. 1a and are used in the operational run of COAMPS. These values are used here for the perfect model run. An index of the frequency with which the threshold friction velocity is achieved in June/July 2009 at 12:00 Z is shown in Fig. 1b. This index is the fraction (expressed in percent) of the total days in June/July 2009 period that u_{*t} exceeded 0.6 m s^{-1} at 12:00 Z. The observations of AOD are used to update the AOD, dust concentration, and the erodibility map. Note that the dust concentration is a three-dimensional field. Operationally, a threshold value of $u_{*t} = 0.6$ is used. However, using the threshold value of $u_{*t} = 0.6$ in the OSSE would tune values only in high friction velocity regions, thus complicating the interpretation of OSSE results. Therefore, in the OSSE a value of $u_{*t} = 0.0$ is used.

Throughout this work an ensemble size of $N = 40$ is used. Each ensemble member has a different initial value of α . At each grid point the ensemble for $\alpha_{i,j}$ is obtained by sampling 40 ensemble members from a Gaussian distribution $\xi(0.25, 0.25)$ where 0.25 is the mean and the standard deviation (spread). The ensemble members with negative values are set equal to 0.01. This distribution defines the initial guess. The initial guess is also called the prior or *background*. The maps of the mean and standard deviation of this initial guess are shown in Fig. 1c and d, respectively. The model is spun up by integrating ensemble members for 60h starting at 00:00 Z, 10 June 2009. The first DA cycle is implemented

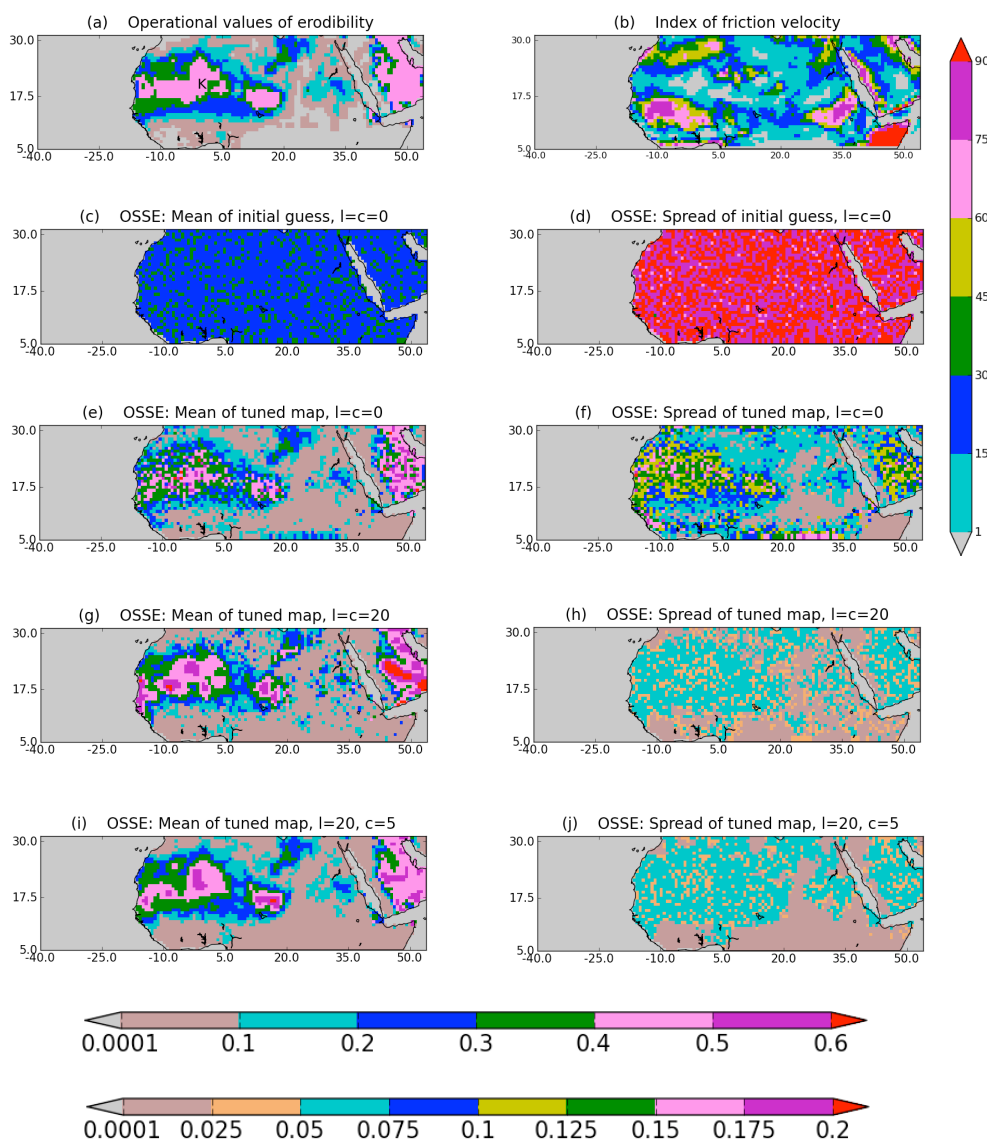


Fig. 1. Shows the mean and standard deviation of erodibility (α) map, except for (b). The vertical colorbar applies to (b). The upper horizontal colorbar applies to the left column. The lower horizontal colorbar applies to the right column (except for (b)). The mean and standard deviations are those for the OSSE. (a) These operational values are used as the perfect values of erodibility (α) in the OSSE. (b) An index of the strength of friction velocity during June/July 2009. The color gives the fraction (expressed as percentage) of the total number (48) of times the friction velocity exceeds a value of 0.6 m s^{-1} . For example, in the Horn of Africa the friction velocity is very strong, exceeding 0.6 m s^{-1} at all times during the 48 cycles. See the friction velocity (u_*) (not shown) looks different at each update cycle. (c) and (d): mean and standard deviation of the initial guess of α map, respectively, for the tuning experiment described in Sects. 3, 4 and 5. (e) and (f): mean and standard deviation of the map (after 48 update cycles), respectively, for the tuning experiment described in Sect. 4. (g) and (h): mean and standard deviation of the tuned α map, respectively, for the tuning experiment with correlation length scale $l = 20$ and cutoff radius $c = 20$ described in Sect. 5. (i) and (j): mean and standard deviation of the tuned α map, respectively, for the tuning experiment with correlation length scale $l = 5$ and cutoff radius $c = 5$ described in Sect. 5.

at 12:00 Z, 12 June 2009. The DA cycling frequency is 24 h. That is, the DA cycle (update) is implemented at 12:00 Z, every day. This frequency for update is chosen because real satellite data are available at 12:00 Z every day. The OSSE is run for 48 days, ending on 18 July 2009 at 12:00 Z, so that there are 48 update cycles. Only the AOD is observed. The

dust concentration and erodibility are not observed. In this work meteorological observations are not assimilated. The observational error is set to 10 % of the mean AOD observation.

Data assimilation experiments with only aerosol state (dust concentration and AOD) estimation were performed.

The forecasts did not improve with aerosol state estimation alone. The results from this experiment show that improving the initial conditions in the dust concentration (c_m) is not important. This is because the source and sinks of dust are strong over a 24 h period and play the key role in deciding the forecast. In other words, over a 24 h period the sources dominate the dust transport especially over areas of strong dust generation.

The theory underlying parameter estimation is the same as that of state estimation. Therefore, the state is augmented by the parameters (α) and data are assimilated. However, the dynamical equations to integrate the parameters in time for an untuned model are not known, which gives rise to some problems. The next section describes these problems and the methodology used in this work to address these problems.

3.1 Spread in α

The meteorological state variables (temperature, the three components of wind speed and humidity) have dynamics evolution equations that are used to integrate these variables forward in time between consecutive updates. However, in the imperfect model the parameter α (which is tuned in this work) does not have a dynamics evolution equation. At a particular grid point we use the following equation to step forward each ensemble member of $\alpha_{i,j}$

$$\alpha_{ij}^k(t+1) = \alpha_{ij}^k(t),$$

where k denotes a particular ensemble member and t denotes time. This is equivalent to using $d\alpha/dt = 0$ as the dynamical equation. If the true α is constant in time, this equation is exact for a tuned system but is inexact for an untuned model.

The dynamic evolution equation for α , used in this work, gives rise to another issue – that of spread in α . Theory of data assimilation states that each time data assimilation updates α , the spread in α must decrease or remain constant. The smaller the spread in α , the less impact observations in succeeding update cycles will have. Because $d\alpha/dt = 0$, the prior spread at a particular update is simply the posterior spread at the last update cycle. This problem is addressed in this work by using conditional inflation (Aksoy et al., 2006). If the posterior spread in $\alpha_{i,j}$ falls below a particular threshold value ($\Delta\alpha_{th}$), the posterior perturbations in $\alpha_{i,j}$ are scaled so that the spread is equal to a particular fixed value ($\Delta\alpha_{fix}$).

$$\alpha_{ij}^k(t+1) = \alpha_{ij}^k(t) + \Delta\alpha_{fix} \left(\alpha_{ij}^k(t) - \tilde{\alpha}_{ij}(t) \right),$$

where $\tilde{\alpha}_{i,j}$ is the mean erodibility.

In this work the threshold value used is $\Delta\alpha_{th} = 0.05$ and $\Delta\alpha_{fix} = 0.05$. These values are chosen after experimentation with different values. If the mean of posterior $\alpha_{i,j}$ is close to the limits of $\alpha_{i,j}$ (0 and 1), then a different strategy is employed. If the mean is less than 0.05 or more than 0.095, the

spread is set equal to 0.015. Also, if the posterior mean decreases below 0.03 (increases above 0.97) it is reset to 0.03 (0.97). One more issue is the risk of unphysical values of the posterior parameter ensemble. Negative values of $\alpha_{i,j}$ are physically meaningless and it is possible that an update results in negative values for some members of the posterior $\alpha_{i,j}$. In this work such ensemble members with negative values are set equal to 0.01.

Before presenting the results of tuning over the whole domain, in the next section the tuning of $\alpha_{i,j}$ (at a single grid point) in the OSSE is explained.

3.2 Tuning at a grid point

It is instructive to consider the tuning of $\alpha_{i,j}$ at a single grid point, allowing the illumination of various issues involved in ensemble-based parameter estimation. The full three-dimensional COAMPS model is run with assimilation of simulated AOD data every 24 h. In the experiment described in this section the AOD is observed at all grid points. However, each $\alpha_{i,j}$ is updated using AOD observation only at that grid point i, j . This is because in this experiment the cutoff radius is set equal to zero (Hamill, et al., 2001).

The tuning of $\alpha_{i,j}$ at point K (Fig. 1a) as the update cycles proceed is shown in Fig. 2a. It shows the mean and standard deviation of the $\alpha_{i,j}$ estimate as the update cycles proceed. The red line shows the truth; that is, the operational value of $\alpha_{i,j}$ at this point K. The mean and standard deviation of the initial guess is 0.3 and 0.2, respectively. As the update cycles proceed the estimate of $\alpha_{i,j}$ approaches the correct value. For this grid point, by 20 cycles the correct value is recovered. The estimates of AOD are shown in Fig. 2d.

In this example the $\alpha_{i,j}$ update uses AOD observations only at the same grid point, the mean of the posterior (or update) at any update cycle is given by

$$\alpha_{up} = \alpha_{prior} + \frac{\text{cov}(\alpha_{prior}, \text{AOD}_{prior})}{\text{var}(\text{AOD}_{prior}) + \text{var}(\text{AOD}_{obs})} [\text{AOD}_{obs} - \text{AOD}_{prior}] \quad (3)$$

In this equation the subscript i, j is not used. All the quantities in this equation are at grid point K. This equation is the Kalman equation for parameter estimation and is presented here to clarify the role of various quantities in the estimation. In this equation AOD_{obs} represents the AOD observation. The observational error variance is given by $\text{var}(\text{AOD}_{obs})$. The other terms are calculated from the short-term ensemble forecast (the prior). The covariance $\text{cov}(\alpha_{prior}, \text{AOD}_{prior})$ plays an important role in the update equation. This covariance exists because of the relation between AOD_{prior} and α_{prior} , which is given by Eq. (2). The short-term ensemble forecast gives 40 different AOD realizations. Each realization of AOD_{prior} corresponds to a particular realization of local variables, non-local variables and meteorology. The local (at point K) variables are AOD_{prior} and u_* . The non-local variables are α and u_* at regions that are upstream of point K. The meteorological variable of interest is wind because

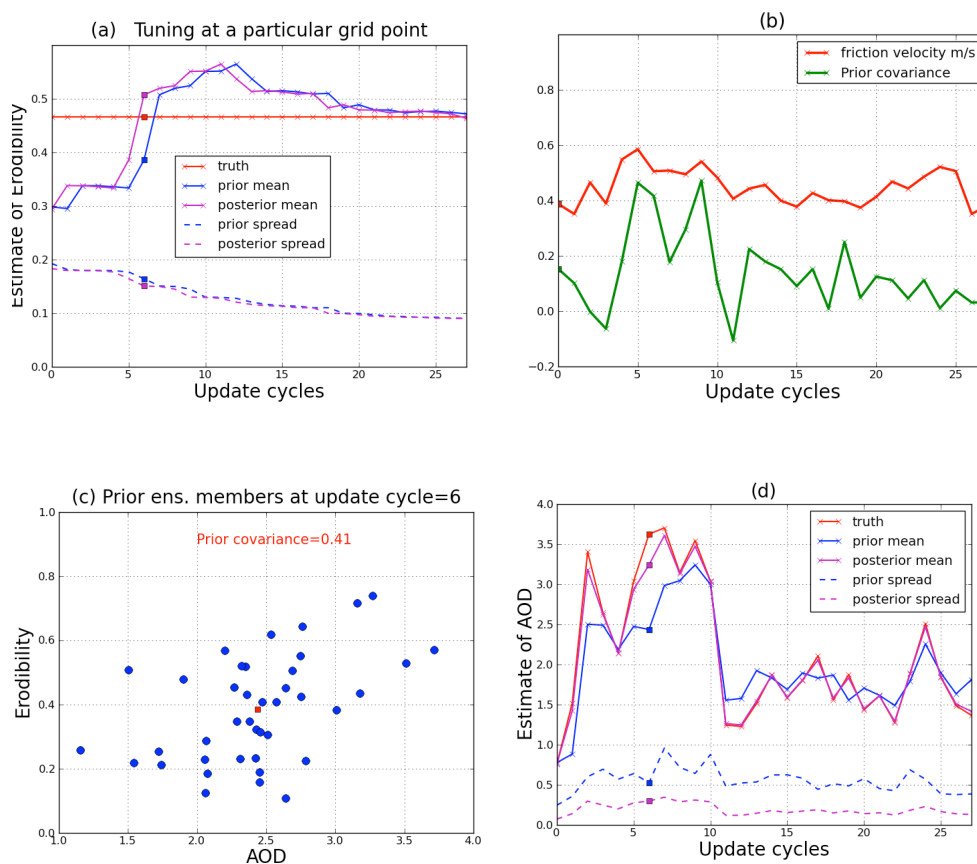


Fig. 2. Tuning of at a particular grid point marked K in Fig. 1a. The estimates of various quantities at this grid point are shown. See Sect. 3.2 for discussion. **(a)** Estimates of α (means and spreads). **(b)** The time series of covariance between α and AOD is shown in green. **(c)** The prior ensemble members at update cycle 6. This update cycle is marked with squares in **(a)**, **(b)** and **(d)**. The red square shows the mean of the ensemble. **(d)** Estimates of AOD (means and spreads).

it advects dust from upstream regions. Therefore, the uncertainty in the $\text{AOD}_{\text{prior}}$ ensemble is due to the uncertainties in local α , local u_* , upstream α , upstream u_* , and winds. From Eq. (2) the uncertainty in prior AOD can be written as

$$\text{Var}(\text{AOD}_{\text{prior}}) = \text{Var}(\text{AOD}_{\text{prior}}^{\text{local}}) + \text{Var}(\text{AOD}_{\text{prior}}^{\text{transport}}). \quad (4)$$

The contribution of uncertainty in local variables is contained in the first term on the right-hand side (rhs) of Eq. (4). The contribution from non-local variables and winds is given by the second term on the rhs of Eq. (4). Out of the total spread of $\text{AOD}_{\text{prior}}$ only a part is correlated with the α_{prior} . This part is the first term of Eq. (4). The remaining spread is due to that in the transported AOD given by the second term that acts as advective additive noise. Given a particular magnitude of advective noise, the strength of the covariance between $\text{AOD}_{\text{prior}}$ and α_{prior} depends on the magnitude of the friction velocity. The time series of u_* and $\text{cov}(\alpha_{\text{prior}}, \text{AOD}_{\text{prior}})$ is shown in Fig. 2b. The covariance (scaled up by a factor of 10) is shown by the green curve. The covariance tends to be higher for higher values of u_* . This is because stronger local generation helps the covariance signal to rise above advective

noise. The prior AOD and α ensembles at update cycle 6 are shown in Fig. 2c. The update cycle 6 is marked on each of the curves in Fig. 2a, b and d by squares. The red square in Fig. 2c shows the mean of the prior ensemble. The spread in the prior AOD ensemble in Fig. 2c includes spread due to local α and additive noise. A finite-size ensemble is used to estimate the true covariance. Because of the small size of the ensemble, it is expected that the ensemble estimate of covariances will not match the true covariance. Such covariances are termed spurious. Spurious covariance is basically an inaccurate estimate of the true covariance due to sampling errors. For example (Fig. 2b), the negative covariance values at update cycle 4 and 11 are spurious. The finite (small) size of the ensemble (40 in this work) is the reason for these spurious covariances.

The estimate of covariance obtained using the ensemble plays a central role in deciding the quality of tuning. Spurious covariance can seriously hamper successful tuning, and since they are unavoidable it is important to properly account for them. This issue of spurious covariances is especially important when AOD observations at many different spatial

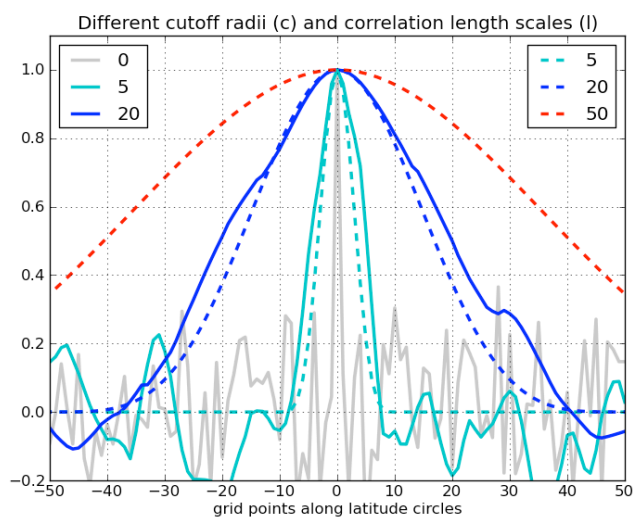


Fig. 3. Illustration of cutoff radius and correlation functions with different length scales. The length scales are specified in units of grid points. The dashed lines show the Gaspari–Cohn localization functions with different cutoff length scales (c). The solid lines are correlation functions. For example, the solid blue line shows the correlation between α at the grid point marked 0 on the x-axis and the neighboring grid points for $l = 20$.

locations are available and used in the estimation. In principle, α at a single grid point is updated using all available observations of AOD. The covariance between α and AOD at the observed grid point determines the weight given to the innovation in the calculation of the increment. This covariance is calculated using the ensemble. If a very large ensemble size is used, the ensemble covariance is more accurate. With a small ensemble size the estimated covariance tends to be inaccurate, especially if the true covariance is small. The true covariance with a grid point geographically far away tends to be smaller, and hence the estimated covariance should be trusted less for far away grid points. The concept of a cutoff radius or a localization radius is widely used in ensemble-based filtering work to address the problem of spurious covariances (Hamill, et al., 2001). The cutoff radius, c , dictates the distance over which observations are used to calculate the correction. This is achieved by defining a (localization) function that decays as one moves away from the grid point being tuned. The Gaspari–Cohn function (Gaspari and Cohn, 1999) is used in this work. The width of this function is governed by the value of c . In the current work, we will run experiments with various values of c . The functions corresponding to the values of c used in this work are shown in Fig. 3 as dashed curves for a single grid point. This grid point is marked 0 on the x-axis. The dashed cyan curve shows the Gaspari–Cohn function corresponding to $c = 5$ grid points. In this work the distance is mentioned in units of grid points. The horizontal resolution used in this work is 81 km (5 grid points is equal to 400 km).

The functions peak at the grid point marked 0. This is the grid point where α is being tuned. The Gaspari–Cohn function value at a particular grid point is used as a multiplicative factor to decrease the weight given to the observation in correcting the α at that grid point. The functions are shown in one dimension along a latitude circle in Fig. 3. However, actual functions are defined in two dimensions. Practically, observations at all grid points more than a distance of $2c$ from the point of interest do not have any impact on the corrections.

In the next section the results of tuning α (all over the domain) is described.

4 Tuning with uncorrelated α perturbations in OSSE

In the last section it was assumed that observations of AOD are available at all points in the domain, whereas in reality actual satellite observations are available for many (but not all) locations. At any given update cycle the satellite observations are sparse. This sparseness of satellite observations is mimicked in the OSSE by observing AOD at 20 % of grid points in the domain. These 20 % grid points are randomly chosen at each update cycle. In the satellite observations, however, the sparse regions need not change randomly with time. The observational error is set equal to 10 % of the mean AOD observation. This observational error is motivated by AOD satellite data whose error is at least 10 % of mean observation.

We begin with an experiment in which the cutoff radius is set to zero. This means that $\alpha_{i,j}$ at any given grid point uses only the AOD observation at that grid point. The mean and standard deviation of the initial guess α is shown in Fig. 1c, d. The perturbations in α in initial guess are uncorrelated. The result of this experiment; that is, the ensemble mean of the tuned α (after 48 cycles) over the domain is shown in Fig. 1e. The uncertainty in this estimated mean is given by the standard deviation in the ensemble which is shown in Fig. 1f. Because this experiment is an OSSE we know the perfect values of erodibility at every grid point which is shown in Fig. 1a.

If the ensemble estimation worked correctly, then the tuned values in Fig. 1e should match those in Fig. 1a. Comparing Fig. 1e to Fig. 1a, it is clear that the ensemble-based tuning is able to recover the perfect values of α to a large extent. The estimation is especially successful over the Arabian Peninsula and parts of the domain where true value of α is small. The assimilation of data constrains the tuned values quite well, in that the standard deviation in the tuned ensemble decreases in Fig. 1f compared to Fig. 1d. Some areas (like the Horn of Africa) have a very small spread (less than 0.025). This is because the mean in these areas decreases below 0.05 and the spread in such areas is set equal to 0.015 (see Sect. 3.1).

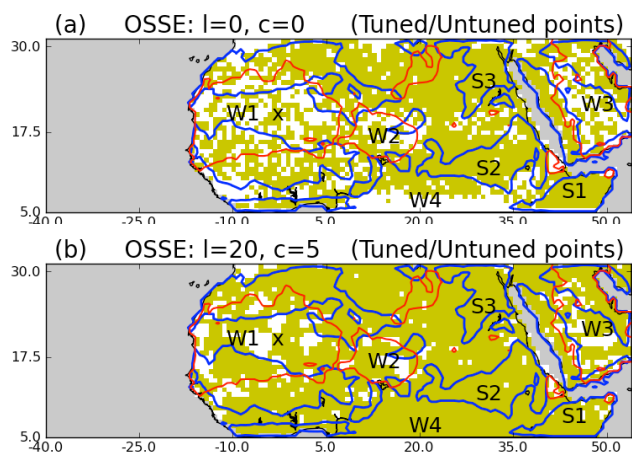


Fig. 4. Yellow color points show the (correctly) tuned grid points. The white or clear grid points show the untuned points. The (correctly) tuned point lies within 0.05 absolute units of the perfect value at a given grid point. The blue colored contours enclose areas with strong friction velocity (Fig. 1b). The red contours enclose areas with high erodibility (Fig. 1a). (a) Shows the quality of tuning for the tuning experiment corresponding to Fig. 1e. (b) Shows the quality of tuning for the tuning experiment corresponding to Fig. 1i.

The success of the tuning experiment is further quantified by comparing the tuned α value at each grid point to the true α value at that grid point. α at a particular grid point is deemed to be (correctly) tuned if its tuned value lies within 0.05 of the true value at that grid point. Otherwise it is deemed to be untuned. This criterion of 0.05 (in absolute units of α) is an arbitrary choice. This criterion is used throughout this work to determine the quality of tuning. The distribution of the tuned and untuned points over the domain is shown in Fig. 4a. The grid points colored with yellow are those tuned successfully. The white grid points are the untuned grid points. The blue contours enclose areas with high friction velocity. These contours correspond to areas in which the friction velocity is above 0.6 m s^{-1} at least 20 % of times (Fig. 1b). The red contours enclose areas with high true erodibility (more than 0.25). Some of the areas with high friction velocity are marked S1, S2 and S3. Notice that in these areas the tuning is successful. In the Horn of Africa (S1) almost all the grid points are successfully tuned. Recall that the friction velocity gives rise to the covariance signal. Consequently areas with strong friction velocity tend to be tuned well. Some areas with weak friction velocity are marked W1, W2, W3 and W4. These areas tend to be poorly tuned. Consider area W1. Note that W1 is an area of weak friction velocity sandwiched between areas of high friction velocity on its north and south. Not only does it have a weak signal but also high advection noise because it lies in an area of high erodibility (it is enclosed by the red contour). As pointed out in Sect. 3.2, the advection noise is additive noise. The combination of low friction velocity (small local signal

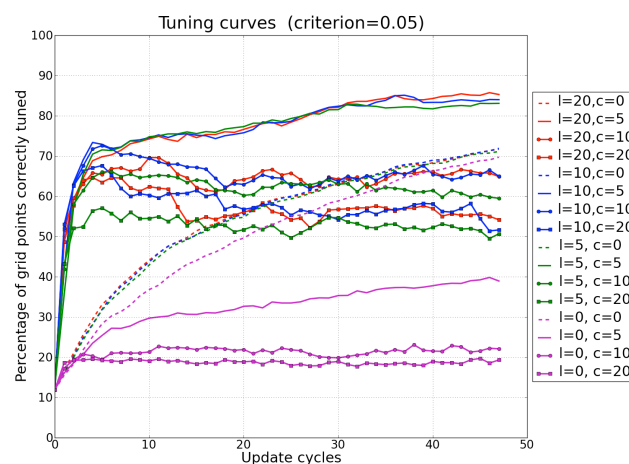


Fig. 5. The tuning curves for OSSE experiments with different values of l and c . Each curve shows the percentage of grid points tuned correctly as the update cycles proceed. The values of l and c are specified in grid point units. A distance of 10 grid points corresponds to 800 km. For this problem a correlation length scale of 5 grid points (400 km) gives the best results. See Sects. 4 and 5 for discussion.

of AOD), and large amounts of advected AOD makes it difficult to correctly estimate the erodibility parameters. The situation is similar with area W3 in the Arabian Peninsula and W2 in the center of the domain. The area W4 in the south of the Sahara has weak friction velocity and low erodibility.

Using this criterion of 0.05, the number of grid points successfully tuned is counted and expressed as percentage of the total number of grid points. This percentage is shown in Fig. 5 as the dashed magenta curve. By the end of 48 update cycles, about 70 % of the points in the domain are correctly tuned.

The experiment discussed above (dashed magenta curve in Fig. 5) used a value of $c=0$. However, using a cutoff radius of 0 prohibits the update at any grid point from using observations in adjoining areas. To assess the impact of using more observations, tuning experiments are run with non-zero values of c . The solid magenta (squares) curve in Fig. 5 shows the percentage of grid points tuned correctly for an experiment with uncorrelated initial α perturbations and a cutoff radius equal to 20 grid points. Clearly, the results degrade compared to the $c=0$ experiment (dashed magenta curve) even though the update of α at any given grid point uses more observations in $c=20$ experiment than that in $c=0$ experiment. Recall that apart from the observation of AOD, a good estimate of the covariance between α at the point being updated and AOD at the location of observation is also important for correct tuning. Apparently, in the $c=20$ tuning experiment the ensemble does not correctly estimate the covariance between α at any given grid point being updated and the neighboring location where the observation is available. This leads to the degradation of the tuning because along

with many observations, the update uses many bad covariances. The reason for the bad covariances is a combination of the effect of advective noise and the small size of the ensemble. The covariance between α at a particular point and AOD at another point is partly controlled by the correlation between α perturbations at these two points. For the experiment described in this section, the updates do not result in correlating the α perturbations; that is, the initially uncorrelated α perturbations remain uncorrelated at the end of all the update cycles. In the current experiment the perturbations are uncorrelated, and hence dust generated by all the points within the neighborhood of point of interest contributes to the advective noise. The small ensemble finds it difficult to capture the local signal due to this advective noise resulting in spurious covariances. Hence, including the observations of AOD from neighboring grid points degrades the tuning rather than improving it. The solid magenta curve shows the result of the experiment with $c = 5$. Its performance is intermediate between $c = 0$ and $c = 20$. The experiment with $c = 10$ (magenta circles) gives almost the same result as that with $c = 20$.

However, one would like to use as many observations as possible by setting $c > 0$. The main hurdle to using $c > 0$ is the advective noise. What can be done to address this problem? A possible solution to this problem is to correlate the perturbations in neighboring α , thereby reducing the advective noise. Also, the results of the experiments in this section suggest that the assimilation of observations does not impose a correlation structure in the α field. That is, the observations are unable to recover the correlation structure (if any) between initially uncorrelated α . Can the assimilation of observations recover the correlation structure if the α perturbations are initially correlated? The next section considers the issue of initially correlated α perturbations.

5 Tuning with correlated α perturbations in OSSE

In this section the initial perturbations of α are spatially correlated. Some examples of the correlation functions between the α perturbations are shown in Fig. 3. The point marked 0 on the x-axis is the point of interest. The solid blue curve gives the correlation between the α perturbations at point 0 and that at various neighboring points along the latitude circle corresponding to a correlation length scale of $l = 20$ grid points. The standard deviation of this correlation function is $l = 20$. This correlation is constructed by first sampling from (uncorrelated) $\xi(0.25, 0.25)$ and then constructing a spatially smoothed perturbation for each ensemble member separately. These weights are chosen proportional to a two-dimensional Gaussian function with standard deviation of $l = 20$. The cyan curve shows the correlation function for $l = 5$. The gray curve shows the correlation function for $l = 0$; that is, independent perturbations. The correlation function of any grid point in the experiment described in the Sect. 4 looks like the gray curve. The term *correlation function* will

imply correlations between α perturbations (between two grid points).

The red (squares) curve in Fig. 5 shows the tuning curve for an experiment with correlation length scale $l = 20$ and cutoff radius $c = 20$. The initial mean and standard deviation for this experiment is shown in Fig. 1c and 1d, which is the same as that for the $l = 0$ experiment described in Sect. 4. The initial guess for the magenta curves and red (squares) curve in Fig. 5 is the same, except that for the red curve the initial α perturbations are correlated over a length scale of 20 grid points. The correlation function of any grid point in the domain for $l = c = 20$ experiment looks like the solid blue curve in Fig. 3. The red (squares) curve in Fig. 5 shows that the tuning is successful for about the first 5 update cycles and there after degrades. The reason for this degradation can be understood by considering the correlation function at a particular grid point as the update cycles proceed. The correlation function at a particular grid point (marked x in area W1 in Fig. 4) is shown in Fig. 6 as the solid green curve. The number in each panel indicates the update cycle. At the initial time a correlation length scale of $l = 20$ is imposed. The green curve is the correlation between the α perturbations at the point marked 0 on the x-axis and that at the neighboring grid points around the latitude circle. The dashed yellow line shows the localization function corresponding to cutoff radius $c = 20$. The dashed black curve in each panel shows a Gaussian with length scale of 5 grid point for reference. At each update cycle AOD data are assimilated and all these α perturbations are updated, thereby modifying the correlation of α with surrounding points. As the update cycles proceed, the correlation function narrows down, as seen in the successive panels in Fig. 6. In fact, it converges towards a function with a length scale of about 5 grid points as seen in the last few update cycles. The parameter estimation results in a length scale of α perturbations of approximately 5 grid points, but the localization is allowing information from much further away to impact the local estimate of α . The correlations with points further away than 5 grid points tend to be bad, and hence as the updates proceed the red (squares) curve in Fig. 5 degrades.

The correlation functions for many grid points at various locations are inspected, and it is found that the correlation length converges to about 5 grid points. A new tuning experiment is run with $l = 5$, $c = 5$. The result of this experiment is shown in Fig. 5 as the solid green curve. Clearly, $l = 5$, $c = 5$ performs far better than $l = c = 0$ and $l = c = 20$. Another experiment is run with $l = 20$, $c = 5$. The tuning curve for this experiment is shown by the solid red curve in Fig. 5. The tuning for $l = 20$, $c = 5$ is as good as that for $l = 5$, $c = 5$. This is because for the $l = 20$ experiment as the update cycles proceed the correlation function narrows to about 5 grid points. Also, because $c = 5$, effectively only observations within a radius of about 5 grid points are used to update α at any grid point.

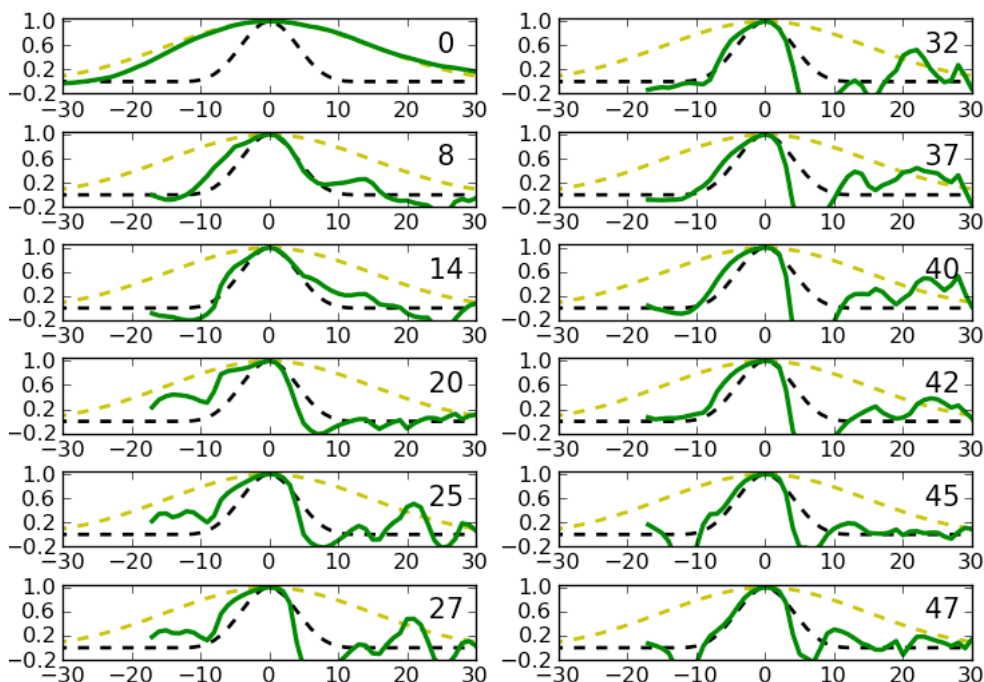


Fig. 6. Evolution of the correlation function at a particular grid point (marked x in area W1 in Fig. 4) for the experiment with $l = c = 20$. Each panel is for a different update cycle. The numbers inside each panel show the update cycle.

The tuned map at the last update cycle for $l = 20$, $c = 5$ experiment is shown in Fig. 1i. The tuned map corresponding to the $l = c = 20$ experiment (solid red (squares) curve in Fig. 5) is shown in Fig. 1g. Clearly, the tuned map in Fig. 1i recovers the perfect map shown in Fig. 1a more accurately than does the $l = c = 0$ experiment (Fig. 1e) or the $l = c = 20$ experiment (Fig. 1g). Comparing Fig. 1f and Fig. 1j the estimate from the $l = 20$, $c = 5$ experiment is constrained better than $l = c = 0$ experiment as can be inferred from the lower values of spread in Fig. 1j.

The spatial distribution of tuned points for $l = 20$, $c = 5$ experiment is shown in Fig. 4b. Comparing this figure with Fig. 4a correlating perturbations and using more observations leads to tuning gains in high advection/low friction velocity regions like W1, W2, W3 and W4. This strengthens, to some extent, our hypothesis that correlating perturbations leads to an improvement in the covariance estimates. This improvement in the signal (covariance) can be considered to be an *effective* decrease in advection noise. The reduction in the degrees of freedom (because of correlations) increases the impact of observations, thereby improving the tuning. It appears that for this particular problem, on an average over the domain, an emergent correlation length scale is about 5 grid points (400 km). Imposing a correlation function of $l = 5$ is leading to better covariances. This does not mean that advection mainly happens over a length scale of 5 grid points. Advection most probably is taking place over longer length

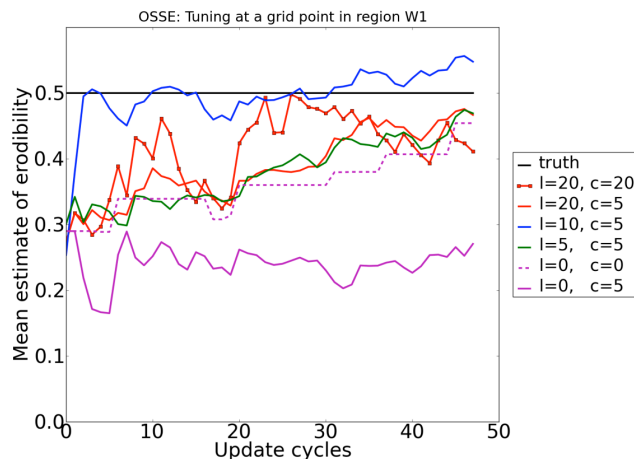


Fig. 7. Comparison between tuning for different values of l and c at a particular point. This point is marked x in the W1 area in Fig. 4. The evolution of the correlation function at this point is shown in Fig. 6.

scale. However, the linear signal due to advection survives only over a length scale of $l = 5$ grid points.

Various other experiments with different values of l and c are run to further investigate the interplay between correlation length and cutoff radius. The red and blue curves correspond to experiments with correlation length scales $l = 20$ and $l = 10$ (800 km), respectively. The behavior of the red (circles) curve ($l = 20$, $c = 10$) is similar to that of red

(squares) curve. For $l = 20$, $c = 10$, the correlation function narrows down and converges to about 5 grid points, similar to the case of $l = c = 20$. However, the degradation is not as much as the $l = c = 20$ because only observations within radius of $c = 10$ grid points are being assimilated. The amount of bad covariances being used is less in the $c=10$ experiment compared to the $c=20$ experiment. The experiment with $l = 10$, $c = 5$ (solid green curve) gives a result comparable with $l = 5$, $c = 5$ and $l = 20$, $c = 5$. This shows that if $l > c$, then the correlation length is effectively $l = c$ as far as the data assimilation is concerned. As the update cycles proceed, l converges to 5 grid points. Before this convergence happens since $c < l$, the observational information beyond a distance of c is not used. For the $l = 10$ experiments with $c = 10$ (blue circles) and $c = 20$ (blue squares), the behavior is similar to the $l = 20$ experiment with similar values of c . This shows that $l > 5$ is too broad for this problem and if c is specified longer than 5, then the data assimilation narrows the correlation function to 5 grid points. Lastly, consider the dashed curves that show results for $c = 0$ for various values of l . These curves approximately overlap, showing that it is futile to correlate perturbations without using observations in the neighborhood. The curves with $c > = l$, for $c > 5$, show that using observations outside the correlated area degrades the tuning, which is because of inaccurate covariances. Figure 7 shows the tuning at a particular point marked x in W1 area Fig. 4. The dashed magenta line uses observations only at the same grid point, and hence the updates take place only when data are available at that grid point. Though the solid magenta curve has access to more observations, the covariance estimates are not good enough because the perturbations are not correlated. The red curve (squares) uses observations over a length scale of $c = 20$, while as the updates progress the correlation narrows to 5 grid points. Consequently, the estimate does not converge towards the perfect value of α at this grid point very well. The solid green, blue and red curves converge smoothly because the correlation is over a scale of 5 grid points. These curves have access to more observational information and improved signal because of correlation.

The results from all these experiments suggest that the observations are able to uncover the correlation scale between neighboring α field, provided the initial α perturbations are correlated over a broad length scale. This correlation scale for this problem is about 5 grid points. As seen in Sect. 4, if the initial α perturbations are uncorrelated, the observations are not able to impose a correlation structure as the updates proceed.

The sensitivity of the OSSE tuning results to ensemble size was found by running experiments with smaller ensemble sizes. As noted, for an ensemble size of $N = 40$, about 85 % of the grid points are tuned for the $l = 20$, $c = 5$ experiment (solid red curve in Fig. 5). This percentage decreases to 75 %, 60 % and 45 % for an ensemble size of 20, 10 and 5, respectively.

Though results from the OSSE experiments are not guaranteed to hold for experiments with real data, they do provide valuable insights into the tuning of erodibility. They show that under ideal circumstances the erodibility is amenable to tuning, given realistic observational coverage and errors. Ideal circumstances mean that the only model error is imperfect values of erodibility. Even so it provides confidence in the tuning methodology to proceed with experiments with real data. The next section describes the tuning experiments with real satellite data.

6 Real data

In this section the tuning experiments with satellite data are described in Sect. 6.1. In Sect. 6.2 the estimated map of erodibility is verified using satellite data.

6.1 Tuning

MODIS Deep Blue data (Remer et al., 2005; Hsu et al., 2004, 2006; Shi et al., 2011) are used for the experiments with real data. The satellite data are averaged over a box of 3 grid points (about 240 km) to obtain super observations. The errors in the observations could be correlated. The averaging serves to decorrelate these errors. The super observations are assimilated into the COAMPS model using the ensemble-based tuning methodology. Here the observational error is set equal to $0.15 + 10\%$ AOD units, but realistically the errors for some locations can be considerably greater (Shi et al., 2011). Shi et al. (2011) also found that lower values of AOD observations tend to have higher relative uncertainty than higher values. Incorporating 0.15 AOD units in the observational error assigns high errors to observations below 0.15. The tuning experiment for the real data runs from 12 June 2009 to 8 July 2009. MODIS satellite data are assimilated every 24 h at 12:00 Z. In total the tuning experiment uses 28 update cycles. The period from 8 July 2009 to 30 July 2009 is used for verification. The threshold friction velocity is set to 0.6 m s. It has to be noted that the experiment with real data is completely separate from the OSSE experiment described in Sects. 3, 4 and 5.

The operational values of α (Fig. 1a) are used as the mean of the initial guess. The ensemble perturbations in α are correlated over a length scale of 5 grid points. The standard deviation of α at each grid point is set equal to 0.25. The negative ensemble members are set equal to 0.01. The mean and standard deviation of this initial guess are shown in Fig. 8a and b, respectively. The standard deviation in some areas in Fig. 8b is lower than 0.25. This is because in these areas the mean of the initial guess has low values, and therefore the ensemble members below the value of zero are set equal to 0.01. This decreases the standard deviation below 0.25. In this experiment the correlation cutoff radius is set equal to 5. After 28 update cycles the mean of tuned α converges to values shown

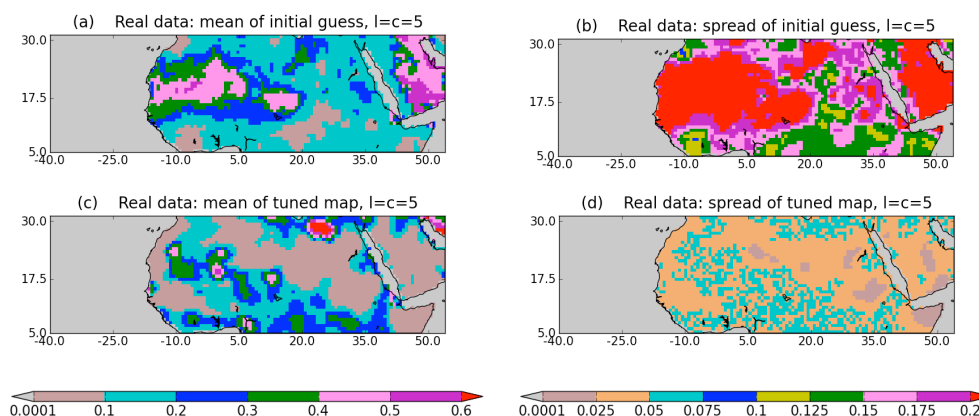


Fig. 8. The tuning experiment with real satellite data. The left colorbar is for (a) and (c). The right colorbar is for the standard deviations shown in (c) and (d). (a) The operational values of α are used as mean of initial guess. (b) Standard deviation of initial guess α are is set equal to 0.25. (c) The mean of tuned values after 28 update cycles. (d) The standard deviation of tuned values after 28 updates cycles.

in Fig. 8c. The standard deviation in the mean of these tuned values is shown in Fig. 8d.

The estimates of α as a function of the update cycles at four different grid points is shown in Fig. 9. For the grid point in panel (a) the estimate decreases from 0.3 to about 0.05. The convergence is not smooth, but clearly the estimation corrects a bias in the first guess in the downward direction. Between cycles 10 and 28 the mean wiggles between 0.05 and 0.1 rather than staying at a constant value. This is because the estimated erodibility can compensate for other errors in the model like those in threshold velocity and advection. Similarly in panel (b) the estimation corrects the erodibility in an upward direction but does not remain constant. Panel (c) shows a case where the erodibility has clearly not converged. In panel (d) the estimate appears to converge between updates 10 and 15, but undergoes large variation after update 20. The estimation curves shown in these panels are representative of many locations in the domain. The assumption that the model is imperfect only in the erodibility is too simplistic. There are many other imperfections in the model. The estimate of erodibility inadvertently corrects for these imperfections. The imperfections in threshold velocity and near surface wind would have the highest impact on the estimate of erodibility because these control the dust flux. The friction velocity depends on the 10 m wind. It is possible that the estimation corrects α to account for imperfection in the 10 m wind. Therefore, one has to exercise caution while interpreting the tuned map of erodibility.

Considering the tuned map (Fig. 8c), on an average in the west Sahara and the Arabian Peninsula the parameter estimation results in lower values of α compared to the operational values (Fig. 1a). It is possible that the estimation decreases the erodibility in these areas to correct for a positive bias in the friction velocity. In the south Sahara region (between latitude 5 and 10° N) α converge to higher values. During the tuning process the ocean values of α are set equal to zero

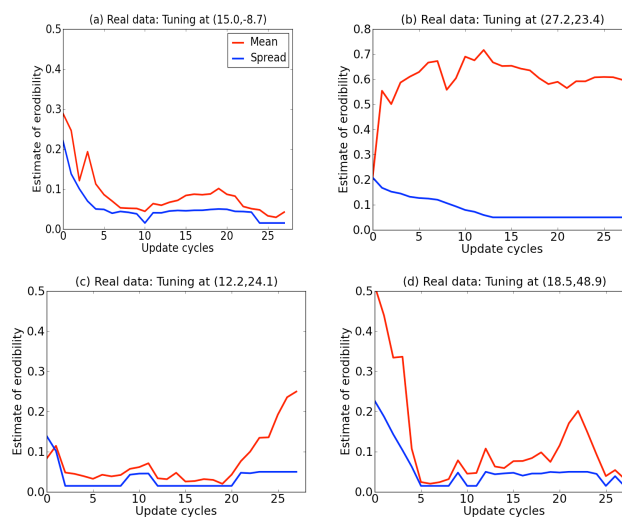


Fig. 9. Each panel shows the mean and spread of the erodibility at 4 different grid points for the real data experiment. The latitude and longitude of the point is mentioned in the title of each panel.

within the model. Comparing (Fig. 8) panel (b) and (d) it is evident that the standard deviation in the mean of the tuned values decreases to about 0.05 compared to the initial guess standard deviation.

Since one does not know the real erodibility, the tuned map has to be assessed indirectly by verifying forecasts of AOD. The next section describes such a verification experiment.

6.2 Verification

The performance of the tuned map (Fig. 8c) in forecasting AOD is compared to that of the operational map. The verification is done over a period independent of the tuning period. Recall that the tuning experiment for the real data runs from 12 June 2009 to 8 July 2009 (28 days). The period from

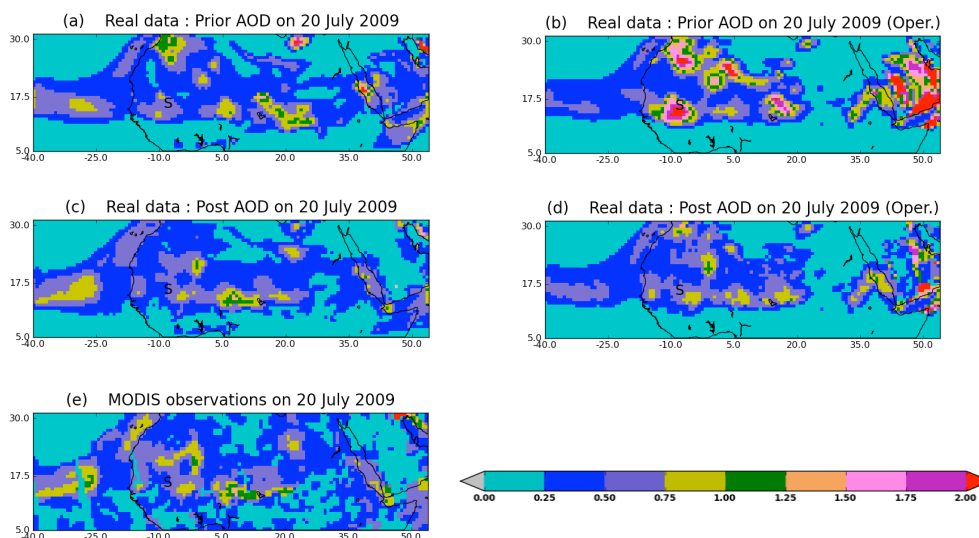


Fig. 10. Mean AOD estimates on 20 July 2009 at 12:00Z. (a) and (b): priors using the tuned (Fig. 8c) and the operational maps (Fig. 1a), respectively. (c) and (d): posteriors obtained by assimilating satellite data into the priors shown in (a) and (b), respectively. (e) Satellite data.

8 July 2009 to 30 July 2009 (19 days) is used for verification. Two separate data assimilation (verification) experiments are run over the verification period. In these experiments only the dust concentration and AOD fields are estimated. The erodibility parameter map is held fixed. The first DA experiment uses the operational erodibility map (Fig. 1a) and the second uses the tuned erodibility map (Fig. 8c). The same MODIS observations are assimilated in each of these experiments. For each experiment we have access to analysis ensemble on 19 different days. For each experiment, 24, 48, 72 and 96 h ensemble forecast is launched from each of these analysis ensembles. Consequently, for each of the two experiments we have 19 different forecast ensemble means. The MODIS observations at the respective days are used to verify the forecast means in each experiment. For a given day MODIS observations are used to verify the forecast launched from the last day, but this data are also assimilated to generate the posterior. This is not a problem because 24 h is long enough for the dust generation and transport to render the forecast almost independent of the initial conditions. The source of dust, that is the erodibility values, plays a dominating role in deciding the spatial distribution of dust over the 24 h period.

Consider the verifications of these two experiments on a particular day. Figure 10 shows the mean estimates of AOD on 20 July, 2009. Panel (e) shows the satellite observations of AOD at 12:00Z, 20 July 2009. The right side panels ((b) and (d)) corresponds to the operational experiments. Panel (a) and (c) shows the estimates from the tuned experiment. The same MODIS data are assimilated into each of these experiments. The prior shown in panel (a) is the mean of the 24 h ensemble forecast launched starting from the posterior AOD on 19 July 2009 for the tuned experiment. This forecast for the operational experiment is shown in panel (b).

Comparing panel (a) and (b) to (e), the tuned forecast agrees with the observations more than the operational forecasts. Panel (c) shows the posterior AOD field corresponding to the prior in panel (a). Panel (d) shows the posterior corresponding to panel (b). The same data (panel (e)) are assimilated into the tuned and operational priors to obtain posteriors in panels (c) and (d), and hence these posteriors are similar. These posteriors are used as initial conditions to launch the next 24 h ensemble forecasts. These forecasts (priors) valid at 12:00Z, 21 July are shown in panels (a) and (b) in Fig. 11. The satellite observations on 21 July 2009 are shown in Fig. 11c. The tuned forecast (panel (a)) matches better with the observations (panel (c)) than does the operational forecast (panel (b)). Note that these tuned and operational forecasts used similar initial conditions in AOD, which are given by panel (c) and (d) of Fig. 10. In spite of these similar initial conditions, the operational forecast (prior) is different from the tuned forecast on 21 July with operational forecasts giving higher AOD values on 21 July. Note that the same meteorology is used in both the operational and tuned experiments. The only difference between the tuned and operational experiments is the different maps of erodibility. Therefore, the difference between these forecasts is due to different values in the erodibility maps. The lower values of AOD in tuned forecasts are attributable to lower values of erodibility in the tuned map (Fig. 8c) compared to the operational map (Fig. 1a).

Both on 20 and 21 July, the tuned forecasts give lower values of AOD, thus resulting in better verifications compared to the operational forecasts. The comparison of verifications over the 19 days is performed by using *mean absolute error*, which is calculated as follows. At each particular grid point, for each 24 hour lead time the absolute difference between

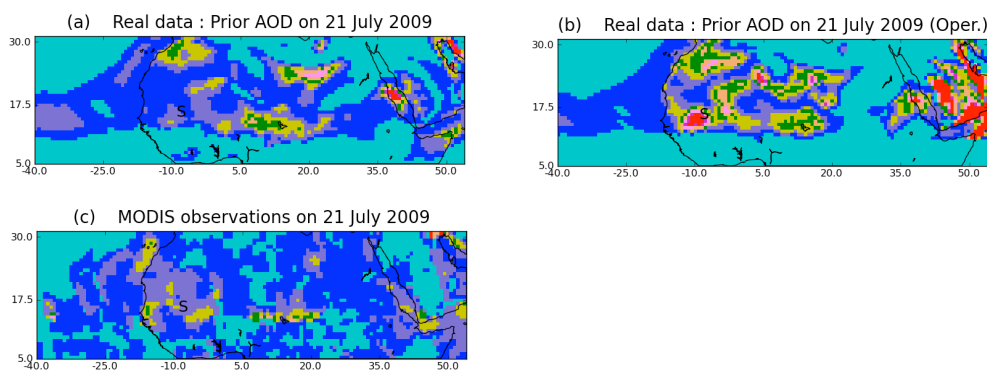


Fig. 11. Mean AOD estimates on 21 July 2009 at 12:00Z. See the colorbar in Fig. 10. (a) and (b) are the forecasts launched from the tuned and operational posteriors on 20 July, respectively. (c) Satellite data.

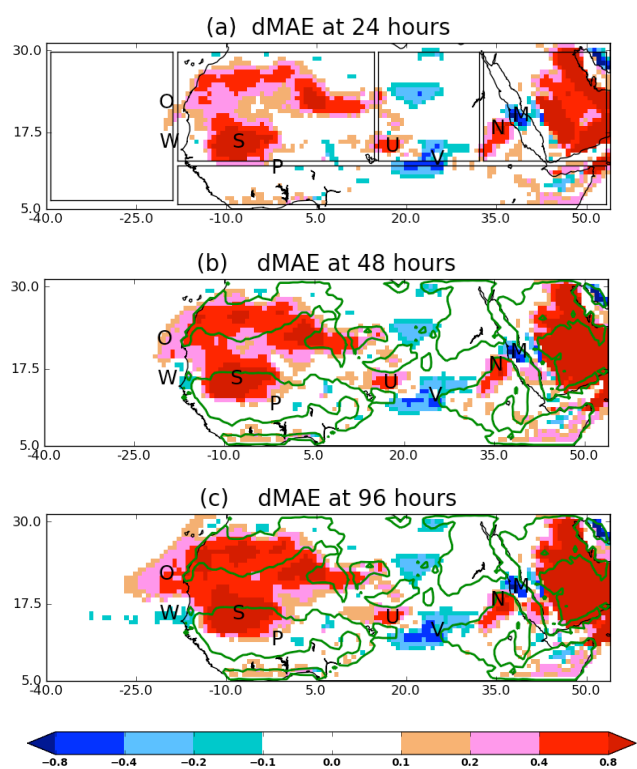


Fig. 12. Result of the verification experiment. The green contours enclose areas of strong friction velocity.

the mean operational forecast and the MODIS observation is calculated.

$$\varepsilon_0 = |AOD_{op} - AOD_{obs}|$$

Then at each grid point, the average of ε_0 over different forecasts is the mean absolute error for the operational model. Similarly, the absolute difference between the mean tuned forecast (AOD_{tu}) and observation is calculated.

$$\varepsilon_{tu} = |AOD_{tu} - AOD_{obs}|$$

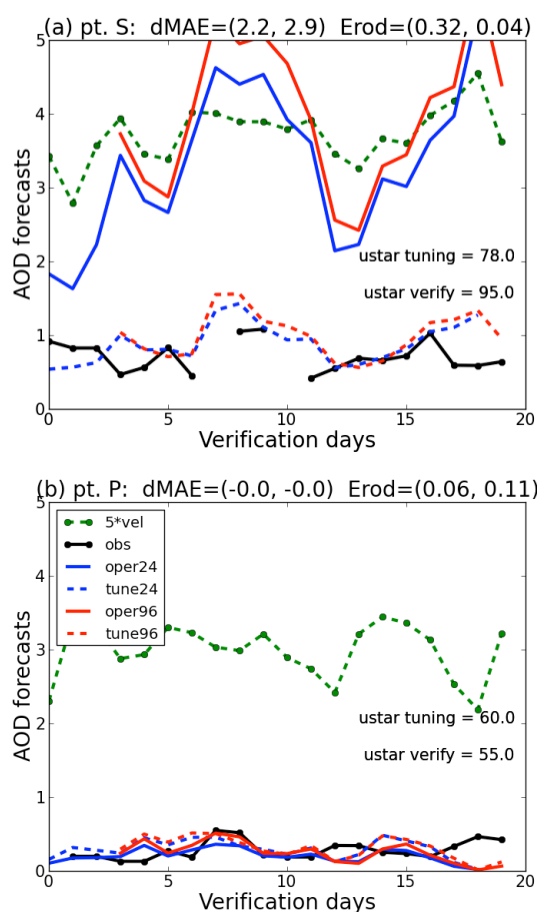


Fig. 13. AOD verifications at grid points S and P in Fig. 12. Legend in (b) applies to (a).

At each grid point, the average of ε_{tu} over different forecasts is the mean absolute error for the tuned model. At each grid point, the operational and tuned mean absolute errors are used to calculate the metric *difference mean absolute error* (dMAE),

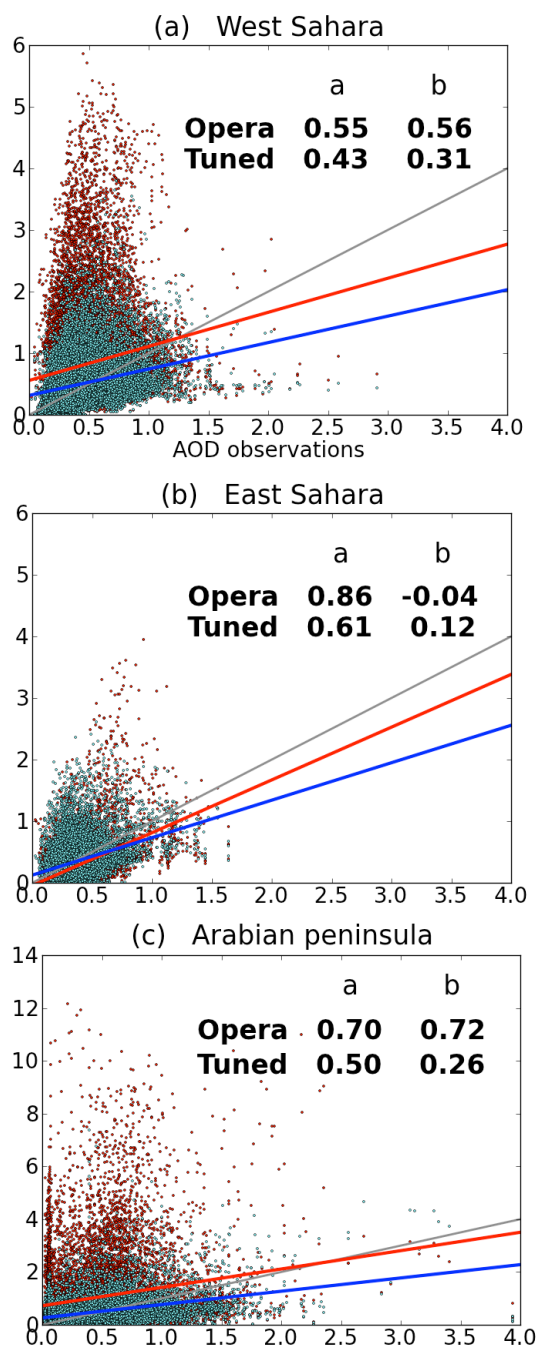


Fig. 14. The scatter plots of forecast AOD (24 h lead time) and AOD observations. (a) corresponds to the box containing point S in Fig. 12a. (b) corresponds to the box containing point U in Fig. 12a. (c) corresponds to the box containing point N in Fig. 12a.

$$\text{dMAE} = \varepsilon_{\text{op}} - \varepsilon_{\text{tu}}$$

This metric is a simple and convenient way to quantify the comparative performance of the operational and tuned models in forecasting the AOD, at each grid point. If $\text{dMAE} > 0$ it means that the operational model errs more than the tuned

model in forecasting the AOD. If $\text{dMAE} > 0$ at a particular grid point, then the tuned model outperforms the operational model. On the other hand, if $\text{dMAE} < 0$ it means that the operational map performs better at that grid point.

The dMAE corresponding to the tuned map in Fig. 8c is shown in Fig. 12, with contours of high friction velocity overlaid. Figure 12a shows the dMAE calculated for the forecast lead time of 24 h. The tuned map outperforms the operational map largely in the west Sahara and Arabian Peninsula regions. The tuned map gives better forecasts than the operational map to some extent in the Horn of Africa. In most of the other regions the dMAE is within -0.1 and $+0.1$, indicating that the tuned and operational forecast are almost similar. There are a couple of pockets near central Sahara where the tuned map gives degraded performance. These areas are blue in color. Panels (b) and (c) in Fig. 12 show the dMAE for longer lead times of 48 and 96 h, respectively. Comparing panels (a), (b) and (c) it is clear that broadly the pattern of areas where the tuned model outperforms the operational model are similar for all lead times. However, comparing the red areas in the vicinity of point S in panels (a) and (b) the tuned model performs better over a larger region for the 96 h forecast compared to the 24 h forecast. Also, the magnitude of improvement of the tuned model is higher for longer lead time in this area. This is also true in the Arabian Peninsula. An important dMAE feature that develops with longer lead times is in the vicinity of points O and W off the coast of Africa. The red color near point O in panel (c) indicates that the tuned model gives a better forecast at 96 h, whereas the tuned model is as good as the operational model in this area at 24 h. In Fig. 13 the relative performance of the tuned and operational models is further probed by inspecting the AOD forecasts at two of the points marked in Fig. 12.

The time series of AOD forecasts at point S are shown in Fig. 13a. The black curve shows the AOD observations. The dashed green curve shows u_* scaled by a factor of 5, during the verification period. Clearly, u_* is above the threshold level of 0.6 m s for almost all the verification times. The solid curves show the operational forecasts at lead times of 24 and 96 h. The dashed curves show the tuned forecasts. The title of the panel shows the value of operational and tuned erodibility. At point S the operational α is 0.32 and the tuned α is 0.04. The title of the panel also shows the dMAE at 24 and 96 h, which is 2.2 and 2.9, respectively. The number 78.0 and 95.0 shown in the panel are the percentage of times when the u_* exceeds the threshold value during the tuning and verification periods, respectively. So at this point out of the total number of cycles (28) in the tuning period, u_* exceeds the threshold value 78 % of times. This point is an example of a grid point where u_* is very strong both during the tuning and verification periods. Because the signal is strong during the tuning period, this point is *tuned correctly*, decreasing the value from 0.32 to 0.04. The phrase *tuned correctly* should be carefully interpreted. We do not know the values of erodibility in nature. Because the (tuned) forecasted AOD

matches well with the observations at this point, we draw the conclusion that the tuned value is correct. The dashed blue curve matches well with the observations while the operational forecast (blue curve) is too high. The low value of tuned AOD can be directly attributed to the lower value of tuned erodibility at point S. Note that the tuned AOD not only has a smaller bias (with respect to the observations) compared to the operational AOD values but also a smaller standard deviation. In both the tuned and operational model the 96 h forecast is higher than the 24 h forecast. This suggests that there is some accumulation of dust over the 96 h. This accumulation seems to be more for the operational than the tuned model as the separation between red and blue curves is larger for the operational model. This accumulation might be because in the operational model the production is more because of higher erodibility of 0.32 (compared to 0.04). The higher dMAE of 2.9 at 96 h compared to 2.2 at 24 h means that the operational model errs more (compared to the tuned model) at 96 h than at 24 h. Note that both the operational and tuned forecasts follow the variations in the friction velocity (green curve).

In Fig. 12 consider the white area to the lower right of point S, around the point marked P. The verification for this point is shown in Fig. 13b. At this point, u_* exceeds the threshold value for about half the time during both tuning (60 %) and verification (55 %) periods. This point has low value of erodibility in the operational map. Because the operational values are *correct* the tuning methodology does not change this value much. The inference that these values are *correct* is drawn from the fact that at point P both operational and tuned models do (equally) well in predicting the observations.

The inspection of the forecasts' time series and observations shown in Fig. 13 suggests that the positive dMAE in Fig. 12 is because the tuned model AOD has a lower bias compared to the operational model AOD.

The biases in the tuned and operational models for various regions are shown in Fig. 14. These various regions are marked by boxes in Fig. 12a. The panel (a) shows the scatter plot of AOD observations versus the forecast AOD in the west Sahara region. The red dots show the operational AOD. The grey line is a reference line with zero bias and slope equal to unity. The red line is the linear fit to the red dots. This linear fit is given by the equation $AOD_{op} = a \times AOD_{obs} + b$

The regression coefficients a and b (which is the bias) are shown in the upper right corner of the panel. The blue line is the linear fit to the cyan dots, which shows the tuned AOD. The tuned model reduces the bias from 0.56 to 0.31. Though the tuned model on an average overestimates the AOD by 0.31 it decreases the bias by 0.25 (compared to the operational model) which is a substantial improvement.

The panel (b) shows the scatter plot for the east Saharan region. In this region the tuned model increases the bias (0.12). The operational model has a small negative bias of -0.04 .

Panel (c) shows that the tuned model decreases the bias by 0.46 from 0.72 to 0.26 in the Arabian Peninsula.

The decrease in bias in the west Sahara and the Arabian Peninsula contributes towards the positive dMAE in these regions. This decrease in bias is due to the downward correction of the tuned erodibility values (compared to the operational values) in these regions. In the south Saharan region the operational bias is -0.09 (result not shown). The tuned model changes this bias to 0.07. This might be due to upward correction in tuned values in the south Saharan region. The positive bias of 0.12 in the east Saharan region might be due to the increased advection from the south Saharan region. In the Atlantic region the biases in the tuned and operational model are comparable (results not shown).

Consider Fig. 12c. The red areas coincide with areas with high operational erodibility. In these areas (west Sahara and the Arabian Peninsula) the operational α was corrected by the tuning to a lower value. In the (white) areas other than west Sahara and the Arabian Peninsula, both operational and tuned maps perform equally well. The improvement of forecasts in west Sahara and the Arabian Peninsula is because tuning leads to a better model of dust generation by decreasing the erodibility. However, an improvement in the dust generation over the red areas does not impact the forecasts in the other areas. This means that the effect of tuning is localized in space. This might be because of the model error in dust transport. Both the tuned and operational models used the same meteorology and therefore the same winds. These might be different from the winds in nature. Both the tuned and operational model suffer from the model error in meteorology. Because of this model error in dust transport, the improved dust generation in the red areas might not necessarily improve dust forecasts in other areas. In this work meteorological variables are not estimated.

At almost all the points in the domain (two of which are shown in Fig. 13), the 96 h forecasts are higher than the 24 h forecasts, for the tuned and the operational model separately. The higher AOD at longer lead time points to accumulation of AOD either from local production or from upstream transport. The higher AOD at 96 h explains the larger coverage of the red area in the Sahara in Fig. 12c compared to Fig. 12a. Because both the models use the same meteorology and sinks, the higher operational AOD at 96 h is due to higher production in the upstream areas. This higher production is due to the higher operational erodibility.

The tuned and operational forecasts were compared to climatology, and it was found that neither were able to outperform the climatology.

7 Conclusions and further work

This work establishes the importance of the correlation structure while performing ensemble-based estimation of spatially extended parameters. The EAKF (ensemble adjustment

Kalman filter) is implemented in an OSSE in which the only model error is due to imperfect erodibility map. The results of OSSE experiments showed that these parameters could be successfully tuned, given observations of AOD, if an appropriate correlation structure is applied to the initial perturbations of erodibility and a supporting localization radius is applied. The tuning results are inferior if a correlation length is not imposed on the initial perturbations of erodibility.

The methodology used involves running experiments with different correlation length scales imposed on the prior parameter perturbations and also different cutoff radii. A priori one does not know the appropriate length scale that should be used for the correlation structure. If an initial correlation structure is not imposed, the data assimilation does not induce a correlation among the erodibility perturbations. However, if a long correlation length scale is imposed along with a long cutoff length scale, then at each grid point the correlation length scale converges towards the appropriate correlation length scale. It is found that the appropriate length scale for this problem is 5 grid points, which is about 400 km. The tuning experiment with a correlation length scale of 5 grid points resulted in the best tuning in OSSE. This technique of specifying a long length scale and allowing the filter to identify the appropriate value is introduced in this work.

It is very important to choose an appropriate value for the cutoff radius. If the cutoff radius is larger than the correlation length scale, then the tuning degrades. This is because then the update uses covariances from regions that are uncorrelated in erodibility.

The tuning methodology is implemented with MODIS satellite data. In general, it is found that the operational model overestimates observed AOD. The ensemble-based tuning correctly identifies this high bias and corrects the operational erodibility maps to lower values. The tuned values are especially low compared to the operational values in the west Sahara and the Arabian Peninsula regions. Verification experiments show that the tuned forecasts are more in agreement with the MODIS observations than the operational forecasts. In the west Sahara and the Arabian Peninsula areas the gain in forecast accuracy due to tuning is as high as 1.5 AOD units.

However, the tuned map obtained with MODIS data should be considered only as demonstration of feasibility of estimation of erodibility. This is because a number of simplistic assumptions about the model are made in this work. These include the assumption that the model is perfect apart from imperfection in erodibility. But the advection in the model is imperfect. Also, the value of 0.6 used for the threshold velocity might be different from that in nature. The friction velocity strongly depends on the 10 m winds and any error in model meteorology renders the 10 m winds erroneous. The low estimated values in west Sahara and the Arabian Peninsula could be because the estimation is correcting for a bias in friction velocity and threshold velocity by correcting erodibility. In this work we have assumed that dust aerosol

is the only aerosol over the North African domain. This assumption does not hold in reality. We have also assumed that the dust aerosol has only one size. The model resolution is also quite coarse. Further, we have assumed that the aerosols do not feed back on the meteorology.

While these assumptions do not impact the OSSE results, they do impact the estimate of erodibility obtained with real data. Therefore, the real data results presented in this work should be considered as a first step towards obtaining ensemble-based estimates of erodibility.

There is a wide scope to improve the erodibility estimates presented in this work. The meteorological state is not estimated in this work. Meteorological observations could be assimilated to estimate this state, which can potentially correct the errors in transport, thereby improving the 10 m winds. It is important to estimate the threshold velocity and the sinks of dust apart from the erodibility. An estimate of erodibility for realistic applications could be obtained by running the estimation experiment for many different years. It would be interesting to compare such an estimate with that obtained from 4DVAR.

It would be exciting to investigate whether this technique of determining the appropriate length scale is effective in other parameter estimation problems. In the current problem the true parameter map is static in that it does not change with time. It would be interesting to test this methodology in a problem where the true parameter map evolves in time.

Acknowledgements. COAMPS is a registered trademark of the Naval Research Laboratory. This work was carried out under ONR project N0001412WX20258. We thank ONR for the same.

Edited by: J. Quaas

References

- Aksoy, A., Zhang, F., and Nielsen-Gammon, J. W.: Ensemble-based simultaneous state and parameter estimation in a two-dimensional sea-breeze model, *Mon. Weather Rev.*, 134, 2951–2970, 2006.
- Anderson, J. L.: An Ensemble Adjustment Kalman filter for data assimilation, *Mon. Weather Rev.*, 129, 2884–2902, 2001.
- Anderson, J. L., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., and Avellano, A.: The Data Assimilation Testbed: A Community Facility, *B. Am. Meteor. Soc.*, 90, 1283–1296, 2009.
- Annan, J. D., Hargreaves, J. C., Edwards, N. R., and Marsh, R.: Parameter estimation in an intermediate complexity Earth System Model using an Ensemble Kalman filter, *Ocean Model.*, 8, 135–154, 2005.
- 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 Center 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., and Majumdar, S. J.: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects, *Mon. Weather Rev.*, 129, 420–436, 2001.
- Cakmur, R. V., Miller, R. L., Perlwitz, J. P., Geogdzhayev, I. V., Ginoux, P., Koch, D., Kohfeld, K. E., Tegen, I., and Zender, C. S.: Constraining the magnitude of the global dust cycle by minimizing the difference between a model and observations. *J. Geophys. Res.*, 111, D06207, doi:10.1029/2005JD005791, 2006.
- Chen, S., Cummings, J., Doyle, J., Hodur, R. H., Holt, T., Liou, C., and Liu, M.: COAMPS 3.0 Model Description-General Theory and Equations, Monterey: NRL Tech Note NRL/PU/7500-0-3-448, Naval Research Laboratory, 2003.
- Cooke, W. F. and Wilson, J. J. N.: A global black carbon aerosol model. *J. Geophys. Res.*, 104, 22137–2162, 1996.
- de Meij, A., Krol, M., Dentener, F., Vignati, E., Cuvelier, C., and Thunis, P.: The sensitivity of aerosol in Europe to two different emission inventories and temporal distribution of emissions, *Atmos. Chem. Phys.*, 6, 4287–4309, doi:10.5194/acp-6-4287-2006, 2006.
- Dirren, S. R., Torn, R. D., and Hakim, G. J.: A data assimilation case-study using a limited-area Ensemble Kalman Filter, *Mon. Weather Rev.*, 135, 1455–1473, 2007.
- Doyle, J. D. and Bond, N. A.: Research Aircraft Observations and Numerical Simulations of a Warm Front Approaching Vancouver Island, *Mon. Weather Rev.*, 129, 978–998, 2001.
- Dubovik, O., Lapyonok, T., Kaufman, Y. J., Chin, M., Ginoux, P., Kahn, R. A., and Sinyuk, A.: Retrieving global aerosol sources from satellites using inverse modeling, *Atmos. Chem. Phys.*, 8, 209–250, doi:10.5194/acp-8-209-2008, 2008.
- 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.
- Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, *Q. J. Roy. Meteor. Soc.*, 125, 723–757, 1999.
- Gill, A. E.: *Atmosphere-Ocean dynamics*, Academic Press, USA, 318–330, 1982.
- Gillette, D. A.: Production of dust that may be carried great distances, edited by: Pewe, T. L., in: *Desert dust: Origin, characteristics and effect on man*, Boulder, Geological Society of America, 11–26, 1981.
- Gillette, D. A. and Passi, R.: Modeling dust emission caused by wind erosion. *J. Geophys. Res.*, 93, 14233–14242, 1988.
- Hacker, J. P. and Rostkier-Edelstein, D.: PBL state estimation with surface observations, a column model and Ensemble filter, *Mon. Weather Rev.*, 135, 2958–2972, 2007.
- Hacker, J. P. and Snyder, C.: Ensemble Kalman filter assimilation of fixed screen-height observations in a parameterized PBL, *Mon. Weather Rev.*, 133, 3260–3275, 2005.
- Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter, *Mon. Weather Rev.*, 129, 2776–2790, 2001.
- Hodur, R. M.: The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS), *Mon. Weather Rev.*, 125, 1414–1430, 1997.
- Hodyss, D.: Accounting for Skewness in Ensemble Data Assimilation, *Mon. Weather Rev.*, 140, 2346–2358, 2012.
- Hogan, T. F. and Rosmond, T. E.: The description of the Navy operational global atmospheric prediction system's spectral forecast model, *Mon. Weather Rev.*, 119, 1786–1815, 1991.
- Houtekamer, P. L. and Mitchell, H. L.: Data Assimilation using an Ensemble Kalman Filter Technique, *Mon. Weather Rev.*, 126, 796–811, 1998.
- Hsu, N. C., Tsay, S.-C., King, M. D., and Herman, J. R.: Aerosol properties over bright-reflecting source regions, *IEEE Trans. Geosci. Remote Sens.*, 42, 557–569, 2004.
- Hsu, N. C., Tsay, S.-C., King, M. D., and Herman, J. R.: Deep Blue Retrievals of Asian Aerosol Properties During ACE-Asia, *IEEE Trans. Geosci. Remote Sens.*, 44, 3180–3195, 2006.
- Huneus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Prospero, J., Kinne, S., Bauer, S., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Fillmore, D., Ghan, S., Ginoux, P., Grini, A., Horowitz, L., Koch, D., Krol, M. C., Landing, W., Liu, X., Mahowald, N., Miller, R., Morcrette, J.-J., Myhre, G., Penner, J., Perlwitz, J., Stier, P., Takemura, T., and Zender, C. S.: Global dust model intercomparison in AeroCom phase I, *Atmos. Chem. Phys.*, 11, 7781–7816, doi:10.5194/acp-11-7781-2011, 2011.
- Huneus, N., Chevallier, F., and Boucher, O.: Estimating aerosol emissions by assimilating observed aerosol optical depth in a global aerosol model, *Atmos. Chem. Phys.*, 12, 4585–4606, doi:10.5194/acp-12-4585-2012, 2012.
- Karspeck, A. R. and Anderson, J. L.: Experimental implementation of an Ensemble Adjustment Filter for an intermediate ENSO model, *J. Climate*, 20, 4638–4658, 2007.
- Keppenne, C. L. and Rienecker, M. M.: Initial testing of a massively parallel Ensemble Kalman Filter with the Poseidon Isopycnal ocean general circulation model, *Mon. Weather Rev.*, 130, 2951–2965, 2002.
- Khade, V. M. and Hansen, J. A.: State dependent predictability: impact of uncertainty dynamics, uncertainty structure and model inadequacies, *Nonlin. Processes Geophys.*, 11, 351–362, doi:10.5194/npg-11-351-2004, 2004.
- Lavoue, D., Liousse, C., and Cachier, H.: Modeling of carbonaceous particles emitted by boreal and temperate wildfires at northern latitudes, *J. Geophys. Res.*, 105, 26871–26890, 2000.
- Lawson, W. G. and Hansen, J. A.: Implications of stochastic and deterministic filters as Ensemble-Based Data Assimilation Methods in Varying Regimes of Error Growth, *Mon. Weather Rev.*, 132, 1966–1981, 2004.
- Lei, J., Bickel, P., and Snyder, C.: Comparison of ensemble Kalman filters under non-Gaussianity, *Mon. Weather Rev.*, 138, 1293–1306, 2010.
- 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, 2008a.
- Lin, C., Zhu, J., and Wang, Z.: Model bias correction for dust storm forecast using Ensemble Kalman Filter, *J. Geophys. Res.*, 113, doi:10.1029/2007JD009498, 2008b.
- Liu, M., Westphal, D. L., Wang, S., Shimizu, A., Sugimoto, N., Zhou, J., and Chen, Y.: A high resolution numerical study of the Asian dust storms of April 2001, *J. Geophys. Res.*, 108, 8653, doi:10.1029/2002JD003178, 2003.
- Liu, M., Westphal, D. L., Walker, A. L., Holt, T. R., Richardson, K. A., and Miller, S. D.: COAMPS real-time dust storm forecasting during operation Iraqi freedom, *Weather Forecast.*, 22, 192–206, 2007.

- Majumdar, S. J., Bishop, C. H., and Ehterton, B. J.: Adaptive sampling with the Ensemble Transform Kalman Filter. Part II : Field program implementaion, *Mon. Weather Rev.*, 130, 1356–1369, 2002.
- May, P. W., Doyle, J. D., Pullen, J. D., and David, L. T.: Two-Way Coupled Atmosphere-Ocean Modeling of the PhilEx Intensive Observational Period, *Oceanography*, 24, 48–57, 2011.
- McLay, J., Bishop, C. H., and Reynolds, C. A.: A local formulation of the Ensemble Transform (ET) analysis perturbation scheme, *Wea. Forecast.*, 25, 985–993, 2010.
- Nickling, W. G. and Gillies, J. A.: Dust emission and transport in Mali, West Africa, *Sedimentology*, 40, 859–868, 1993.
- Park, S.-U. and In, H.-J.: Parametrization of dust emission for the simulation of the yellow sand (Asian dust) event observed in March 2002, *J. Geophys. Res.*, 108, 4618, doi:10.1029/2003JD003484, 2003.
- Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R. C., Kleidman, R. G., Eck, T. F., Vermote E., and Holben, B. N.: The MODIS Aerosol Algorithm, Products, and Validation, *J. Atmos. Sci.*, 62, 947–973, 2005.
- 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.
- Schutgens, N. A. J., Nakata, M., and Nakajima, T.: Estimating aerosol emissions by assimilating remote sensing observations into a Global transport model, *Remote Sens.*, 4, 3528–3543, 2012.
- 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.
- Shi, Y., Zhang, J., Reid, J. S., Hyer, E. J., Eck, T. F., Holben, B. N., and Kahn, R. A.: A critical examination of spatial biases between MODIS and MISR aerosol products – application for potential AERONET deployment, *Atmos. Meas. Tech.*, 4, 2823–2836, doi:10.5194/amt-4-2823-2011, 2011.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the advanced Research WRF Version 2, Boulder, NCAR Technical Note, 2005.
- Szunyogh, I., Kostelich, E. J., Gyarmati, G., Kalnay, E., Hunt, B. R., Ott, E., Satterfield, E., and Yorke, J. A.: A Local Ensemble Transform Kalman Filter data assimilation system for the NCEP global model, *Tellus*, 60, 113–130, 2008.
- Tegen, I. and Fung, I.: Modeling of mineral dust in the atmosphere: Sources, transport and optical thickness, *J. Geophys. Res.*, 99, 22897–22914. doi:10.1029/94JD01928, 1994.
- Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., Bernsten, 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., Kirkevåg, 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.
- Tipett, M. K., Anderson, J. L., and Bishop, C. H.: Ensemble Square Root Filters, *Mon. Weather Rev.*, 131, 1485–1490, 2003.
- Uno, I., Yumimoto, K., Shimizu, A., Hara, Y., Sugimoto, N., Wang, Z., Liu, Z., and Winker, D. M.: 3D structure of Asian dust transport revealed by CALIPSO lidar and a 4DVAR dust model, *Geophys. Res. Lett.*, 35, L06803, doi:10.1029/2007GL032329, 2008.
- Walker, A. L., Liu, M., Miller, S. D., Richardson, K. A., and Westphal, D. L.: Development of a dust source database for mesoscale forecasting in southwest Asia, *J. Geophys. Res.*, 114, D18207, doi:10.1029/2008JD011541, 2009.
- Wang, K.-Y., Larry, D. J., Shallcross, D. E., Hall, S. M., and Pyle, J. A.: A review on the use of the adjoint method in four-dimensional atmospheric-chemistry data assimilation, *Q. J. Roy. Meteor. Soc.*, 127, 2181–2204, 2001.
- Wang, X., Barker, D. M., Snyder, C. and Hamill, T. M.: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part II : Real observation experiments, *Mon. Weather Rev.*, 136, 5132–5147, 2008.
- Westphal, D. L., Toon, O. B., and Carlson, T. N.: A two-dimensional investigation of the dynamics and microphysics of Saharan dust storms, *J. Geophys. Res.*, 92, 3027–3049, 1987.
- Westphal, D. L., Toon, O. B., and Carlson, T. N.: A case study of mobilization and transport of Saharan dust, *J. Atmos. Sci.*, 45, 2145–2175, 1988.
- Whitaker, J. S., Compo, G. P., Wei, X., and Hamill, T. M.: Reanalysis without radiosondes using Ensemble Data Assimilation, *Mon. Weather Rev.*, 132, 1190–1200, 2004.
- Whitcomb, T. R.: Tropical observability and predictability, Cambridge, Massachusetts Institute of Technology, USA, 2008.
- 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.*, 38, L21802, doi:10.1029/2011GL049258, 2011.
- 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, D10208, p. doi:10.1029/2007JD009065, 2008.
- Zupanski, M.: Maximum Likelihood Ensemble Filter: Theoretical Aspects, *Mon. Weather Rev.*, 133, 1710–1726, 2005.