A Hybrid Kalman Filter Algorithm for Large-Scale Atmospheric Chemistry Data Assimilation

R. G. HANEA

Delft University of Technology, Delft, and Netherlands Environmental Assessment Agency, National Institute of Public Health and the Environment (RIVM), Bilthoven, Netherlands

G. J. M. VELDERS

Netherlands Environmental Assessment Agency, National Institute of Public Health and the Environment (RIVM), Bilthoven, Netherlands

A. J. SEGERS

Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands

M. VERLAAN AND A. W. HEEMINK

Delft University of Technology, Delft, Netherlands

(Manuscript received 5 April 2005, in final form 24 February 2006)

ABSTRACT

In the past, a number of algorithms have been introduced to solve data assimilation problems for large-scale applications. Here, several Kalman filters, coupled to the European Operational Smog (EUROS) atmospheric chemistry transport model, are used to estimate the ozone concentrations in the boundary layer above Europe. Two Kalman filter algorithms, the reduced-rank square root (RRSQRT) and the ensemble Kalman filter (EnKF), were implemented in a prior study. To combine the best features of these two filters, a hybrid filter was constructed by making use of the reduced-rank approximation of the covariance matrix as a variance reducer for the EnKF. This hybrid algorithm, complementary orthogonal subspace filter for efficient ensembles (COFFEE), is coupled to the EUROS model. The performance of all algorithms is compared in terms of residual errors and number of EUROS model evaluations. The COFFEE results score somewhere between the EnKF and RRSQRT results for less than approximately 30 model evaluations. For more than approximately 30 model evaluations, the COFFEE results are, in all cases, better than the EnKF and RRSQRT results. The results of the COFFEE simulations with more than about 60 model evaluations proved to be significantly better than all the EnKF and RRSQRT simulations (even better than those with 100 and 200 modes or ensemble members). The performance of both the EnKF- and RRSQRT-type filters is affected by the nonlinear properties of the model and observation operator, because both rely on linearization to some extent. To further study this aspect, several measures of nonlinearity were calculated and linked with the performance of these algorithms.

1. Introduction

The goal of data assimilation is to produce model predictions that agree as well as possible with the measurements and are consistent with the physical data provided by a numerical model. Different approaches for data assimilation have been used in atmospheric applications. In the four-dimensional variational data assimilation approach, a cost function for the difference between the model and measurements is minimized (see Elbern et al. 1997). A practical disadvantage of this method is the necessity of writing the adjoint model, which is not straightforward for a complex and large-scale problem. Another approach to data assimilation is sequential data assimilation. Although, the Kalman fil-

Corresponding author address: R. G. Hanea, Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD, P.O. Box 5031, Delft 2600 GA, Netherlands.
E-mail: r.g.hanea@ewi.tudelft.nl

DOI: 10.1175/MWR3269.1

ter can be shown to be optimal for linear problems, there are two major disadvantages: first, the atmospheric transport models have a very large number of elements in the state vector; second, the real-life applications are not linear. A classical filter approach with a full covariance matrix is impossible to use in this cases. In a prior study (Hanea et al. 2004), two low-ranking Kalman filter approaches were used to assimilate the measurements into the European Operational Smog (EUROS) regional air pollution model: the reduced rank square root Kalman filter (RRSQRT) and the ensemble Kalman filter (EnKF). The assimilation showed more accurate results for boundary layer ozone than those obtained with the EUROS model.

The EnKF, introduced by Evensen (1994), has been used successfully in many applications (see Evensen and van Leeuwen 1996; Houtekamer and Mitchell 1998; Canizares 1999). The EnKF is based on the representation of the probability density of the state estimate by a finite number N (N being much smaller than the number of elements in the original state vector) of randomly generated system states (also called ensemble members). The standard deviation of the errors in the state estimate is statistical and only converges slowly with the number of ensemble members. The RRSQRT approach was introduced by Cohn and Todling (1995, 1996) and Verlaan and Heemink (1995, 1997). A RRSQRT filter expresses the error covariance in a small number of modes by projecting the full covariance matrix into a smaller subspace formed by the first q largest eigenvectors of the covariance matrix. To compensate for an increase of the numbers of modes caused by system noise, the RRSQRT filter reduces the representation of the covariance matrix only to the leading eigenvectors. Because of the reduction, the covariance matrix is always underestimated; although this bias reduces with the number of modes. As a result, the algorithm is sensitive to filter-divergence problems. The RRSQRT approach can be viewed as an EnKF for which the modes are not chosen randomly, but in the direction of the largest eigenvectors. In both cases the number of modes represents a measure of the storage and computation time required by the filter, and should be as low as possible, while still providing a satisfactory accuracy.

The complementary orthogonal subspace filter for efficient ensembles (COFFEE) algorithm was introduced by Heemink et al. (2001) and the performance of this algorithm was tested on a simple problem. A twin experiment was performed for a 2D advection diffusion equation to obtain insight into the computational and convergence aspects. From the results, the COFFEE filter could be concluded to be more efficient. The COFFEE algorithm is a combination of the ensemble-type and square root–type algorithms and it can be interpreted as a variance reducer for the EnKF.

In this paper the COFFEE algorithm was coupled with the EUROS model. Contrary to Heemink et al. (2001), in this study, a real-life, large-scale atmospheric chemistry model with a complex ozone chemistry scheme was used, with the grid covering the whole of Europe (Hanea et al. 2004). The complexity of the model and the large-scale properties make this application very interesting from the point of view of the performance of the low-ranking Kalman filter algorithms.

A description of the hybrid algorithms specified above is given in section 2. The EUROS model is presented in section 3, and the results of the assimilation process are presented and explained in section 4 using the nonlinearity measure. Section 5 presents the conclusions.

2. Hybrid filters

a. Kalman filter for a nonlinear system

This section will briefly review standard extended Kalman filtering theory (Jazwinski 1970), mainly to introduce the notations (Ide et al. 1997) that will be used in this study. Consider the system of a general, nonlinear stochastic model $\mathcal{M}$ and the following observations:

$$\mathbf{x}^f(t_k) = \mathcal{M}[\mathbf{x}^f(t_{k-1})] + \mathbf{w}(t_{k-1}), \tag{1}$$

$$\mathbf{y}^o(t_k) = \mathbf{H}(t_k)\mathbf{x}^f(t_k) + \mathbf{v}(t_k), \tag{2}$$

where $\mathbf{x}^f(t_k) \in \mathcal{R}^n$ denotes the forecast of the system state at time t_k and $\mathcal{M}[\mathbf{x}^f(t_k)]$ represents one time step of the model. A normal distributed system noise $\mathbf{w}(t_k) \in \mathcal{R}^n$, $\mathbf{w}(t_k) \sim N(0, \mathbf{Q})$ is introduced to take the uncertainties in the model into account. The vector $\mathbf{y}^o \in \mathcal{R}^r$ represents the measurements, which are supposed to be a sum of a linear observation of the state, represented by the operator $\mathbf{H}$. The observation noise is represented by $\mathbf{v}(t_k) \sim N(0, \mathbf{R})$. To obtain an optimal estimate we need to combine the measurement taken from the actual system and modeled by Eq. (2) with the information given by the system model [Eq. (1)]. The forecast state at time t_k, denoted by $\mathbf{x}^f(t_k)$, is the forecast from observation time t_{k-1} to observation time t_k by the following equations:

$$\mathbf{x}^f(t_k) = \mathcal{M}[\mathbf{x}^a(t_{k-1})], \tag{3}$$

$$\mathbf{P}^f(t_k) = \mathbf{M}(t_{k-1})\mathbf{P}^a(t_{k-1})\mathbf{M}(t_{k-1})^{\mathrm{T}} + \mathbf{Q}(t_{k-1}), \tag{4}$$

where

$$\mathbf{M}(t_k)_{i,j} = \frac{\partial \mathcal{M}_i[\mathbf{x}^f(t_k)]}{\partial \mathbf{x}_j^f(t_k)} \tag{5}$$

is the tangent linear model, and $\mathbf{P}^f(t_k)$ is the forecast error covariance. The superscript f means forecast, superscript a means analyzed, and superscript o means observed. At time t_k, observation $\mathbf{y}^o(t_k)$ is available and the estimate is updated by the analysis step

$$\mathbf{x}^a(t_k) = \mathbf{x}^f(t_k) + \mathbf{K}[\mathbf{y}^o(t_k) - \mathbf{H}(t_k)\mathbf{x}^f(t_k)], \tag{6}$$

$$\mathbf{P}^a(t_k) = \mathbf{P}^f(t_k) - \mathbf{K}(t_k)\mathbf{H}(t_k)\mathbf{P}^f(t_k), \tag{7}$$

where

$$\mathbf{K}(t_k) = \mathbf{P}^f(t_k)\mathbf{H}(t_k)^{\mathrm{T}}[\mathbf{H}(t_k)\mathbf{P}^f(t_k)\mathbf{H}(t_k)^{\mathrm{T}} + \mathbf{R}(t_k)]^{-1} \tag{8}$$

is the minimum variance gain. The final result of the Kalman filter using the minimum variance gain is a time series of estimates for mean and covariance of the true state.

Direct use of these equations in an atmospheric chemistry transport model with nonlinear processes and with a very large grid will cause three problems: 1) an ill-conditioned matrix $\mathbf{P}$, 2) a long computation time, and 3) the use of the tangent linear model as an approximation for the original model, which is a source of numerical errors. Therefore, in a prior study, (Hanea et al. 2004), two low-ranking algorithms, RRSQRT and EnKF, were used instead of the full Kalman filter approach.

b. Low-ranking Kalman filters

1) ENKF

The EnKF, introduced by Evensen (1994) and later clarified by Burgers et al. (1998), has been successfully implemented and used in different types of applications. The EnKF is a Monte Carlo approach based on the representation of the probability density of the state estimate by a finite number of randomly generated system states. For the initial state estimate x_0, the uncertainty is expressed by an ensemble $\boldsymbol{\xi}_i^a$, $i = 1, \ldots, N$ of randomly generated states. The ensemble members are propagated from one time step to another using the original model operator (no tangent linear system operator is needed)

$$\boldsymbol{\xi}_i^f(t_k) = \mathcal{M}[\boldsymbol{\xi}_i^a(t_{k-1})] + \mathbf{w}_i(t_{k-1}), \tag{9}$$

with $\mathbf{w}_i(t_k)$ realizations of the noise process $\mathbf{w}(t_k)$, that is, noise is added to the most uncertain parts of the model to estimate the covariance between observations and the model variables. The ensemble mean

$$\bar{\mathbf{x}}^f(t_k) = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\xi}_i^f(t_k) \tag{10}$$

represents the state estimate at time t_k. Using this estimate the error covariance can be computed as

$$\mathbf{E}^{f,\mathrm{EN}}(t_k) = [\boldsymbol{\xi}_1^f(t_k) - \bar{\mathbf{x}}^f(t_k), \ldots, \boldsymbol{\xi}_N^f(t_k) - \bar{\mathbf{x}}^f(t_k)], \tag{11}$$

$$\mathbf{P}^f(t_k) = \frac{1}{N-1}\mathbf{E}^{f,\mathrm{EN}}(t_k)[\mathbf{E}^{f,\mathrm{EN}}(t_k)]^{\mathrm{T}}. \tag{12}$$

With the error covariance $\mathbf{P}^f$ calculated, the Kalman gain $\mathbf{K}(t_k)$ is obtained from Eq. (8) and the update equations for the analyzed ensemble are

$$\boldsymbol{\xi}_i^a(t_k) = \boldsymbol{\xi}_i^f(t_k) + \mathbf{K}(t_k)[\mathbf{y}^o(t_k) - \mathbf{H}(t_k)\boldsymbol{\xi}_i^f(t_k) + \mathbf{v}_i(t_k)], \tag{13}$$

where $\mathbf{v}_i(t_k)$ represents realizations of the noise process $\mathbf{v}(t_k)$.

2) RRSQRT FILTER

This approach, introduced by Cohn and Todling (1995, 1996) and Verlaan and Heemink (1995, 1997), is based on the idea of projecting the full covariance matrix for the state estimate onto a smaller subspace represented by the q leading eigenvectors $\mathbf{l}_i$, $i = 1, \ldots, q$ of $\mathbf{P}$. In matrix form $\mathbf{P} \simeq \mathbf{L}\mathbf{L}^{\mathrm{T}}$, where $\mathbf{L}$ is a matrix formed with the q leading eigenvectors $\mathbf{l}_i$.

The first step is initialization:

$$\begin{aligned} &\mathbf{x}^a(t_0) = \mathbf{x}_0, \\ &\mathbf{P}^a(t_0) = \mathbf{L}^a(t_0)[\mathbf{L}^a(t_0)]^{\mathrm{T}}, \quad \text{and} \\ &\mathbf{L}^a(t_0) = [\mathbf{l}_1^a(t_0), \ldots, \mathbf{l}_q^a(t_0)]. \end{aligned} \tag{14}$$

In the forecast step of the algorithm, the propagation of these modes ($\mathbf{l}_i$) using the original model is

$$\mathbf{x}^f(t_k) = \mathcal{M}[\mathbf{x}^a(t_{k-1})], \tag{15}$$

$$\mathbf{l}_i^f(t_k) = \frac{1}{\epsilon}\{\mathcal{M}[\mathbf{x}^a(t_{k-1}) + \epsilon\mathbf{l}_i^a(t_{k-1})] - \mathcal{M}[\mathbf{x}^a(t_{k-1})]\}, \tag{16}$$

where ϵ is a small perturbation, usually chosen close to 1.

To represent the growth of the covariance, the reduced matrix $\mathbf{L}$ is augmented with extra columns represented by the columns of error covariance of the system (matrix $\mathbf{Q}$)

$$\tilde{\mathbf{L}}^f(t_k) = [\mathbf{l}_1^f(t_k), \ldots, \mathbf{l}_q^f(t_k), \mathbf{Q}(t_{k-1})^{1/2}]. \tag{17}$$

A projection onto the q leading eigenvectors of the matrix $\tilde{\mathbf{L}}^f(t_k)\tilde{\mathbf{L}}^f(t_k)^{\mathrm{T}}$ is performed using a matrix projection $\mathbf{\Pi}^f$:

$$\mathbf{L}^f(t_k) = \mathbf{\Pi}^f(t_k)\tilde{\mathbf{L}}^f(t_k). \tag{18}$$

This operation can be performed efficiently because $\mathbf{L}$ has a small number of columns. The state error covariance matrix can be computed with this new matrix $\mathbf{L}^f(t_k)$:

$$\mathbf{P}^f(t_k) = \mathbf{L}^f(t_k)[\mathbf{L}^f(t_k)]^{\mathrm{T}}. \tag{19}$$

3) A comparison of EnKF and RRSQRT

Before emphasizing the differences, we will first consider the similarities between the EnKF and RRSQRT filters.

First, both algorithms can be interpreted as square root–type filters. The RRSQRT is by nature (definition, the way it is constructed) a square root type, being based on the factorization of the state error covariance matrix $\mathbf{P}$ [Eq. (19)]. The EnKF can also be interpreted as a square root filter if we take Eq. (12) into account, where the state error covariance is represented by

$$\mathbf{P}^f = \frac{1}{N}\mathbf{E}\mathbf{E}^{\mathrm{T}} = \left[\frac{1}{\sqrt{N}}\mathbf{E}^{f,\mathrm{EN}}(t_k)\right]\left[\frac{1}{\sqrt{N}}\mathbf{E}^{f,\mathrm{EN}}(t_k)\right]^{\mathrm{T}}. \tag{20}$$

Second, both algorithms can be interpreted as ensemble-type filters. The EnKF is by definition based on the representation of the probability density of the state by a finite number of ensembles randomly generated [as in Eqs. (9) and (10)]. The RRSQRT approach can also be formulated as an EnKF, where the q ensemble members are no longer chosen randomly (as in the original EnKF), but are taken in the direction of the q leading eigenvectors of the covariance matrix $\mathbf{P}$. These similarities will later simplify the implementation of a hybrid algorithm that will couple these two approaches.

Now consider the differences, first the type of the errors produced by these two algorithms are different. The EnKF shows statistical errors and the RRSQRT leads to truncation errors (Hanea et al. 2004). A good estimation is only obtained with a large number of ensemble members, the RRSQRT is more effective with a small number of modes. Second, for a linear model the RRSQRT algorithm is suitable, but less robust in case of a nonlinear model, where the EnKF is more suitable. Third is the difference in computational time. The computational time, in case of linear model, for the EnKF is $O(N)$ and the RRSQRT is $O(N^2)$. Based on these observations, the two filter approaches can be concluded to be complementary from the theoretical point of view.

The complementary character of the filters can also be shown in real applications. Hanea et al. (2004) concluded that both algorithms tend to converge to the same accuracy with an increasing number of modes, and both have some drawbacks when the number of ensemble members grows. Because of its inherent statistical properties, the EnKF can handle the nonlinearity of the model much better, but requires a large number of ensemble members. For a small number of modes the RRSQRT accuracy is better than the EnKF accuracy (see Fig. 5 in Hanea et al. 2004).

In the next section a hybrid algorithm will be proposed that aims to combine the advantages of RRSQRT and EnKF and cancel their disadvantages by combining both methods.

c. *Partially orthogonal ensemble Kalman filter (POEnKF) and COFFEE*

One problem of the RRSQRT algorithm is that repeated projection on the leading eigenvalues leads to a systematic bias in forecast errors. The part of the covariance matrix that is truncated does not contribute significantly to improvements of the state estimate.

1) POEnKF

The idea behind the POEnKF (Heemink et al. 2001) is to use an extra ensemble to account for the information lost by truncation. The algorithm has two important components: the q leading eigenvectors $\mathbf{l}_i$ of the covariance matrix and the N randomly generated ensemble members $\boldsymbol{\xi}_i$. These two key components are propagated using the forecast step of the RRSQRT algorithm [Eqs. (14)–(17)] and the ensemble members $\boldsymbol{\xi}_i$ are propagated through the model by the forecast step from the EnKF algorithm [Eqs. (9)–(11)]. It is like having two almost independent algorithms. The point where the two approaches interact forms the analysis step of the algorithm.

In the POEnKF time-update step, the RRSQRT time update is used in the directions indicated by the leading eigenvectors of the forecast covariance matrix. A separate ensemble is used to represent the remaining space. A disadvantage of the POEnKF algorithm is that the covariance for this remaining space is initially obtained by generating the ensemble for the whole state; afterward, the members of ensemble that give information in the same direction as the leading eigenvectors are removed by an orthogonal projection. The contribution of the ensemble is therefore ignored in the subspace spanned by the leading eigenvectors.

2) COFFEE

A better way of combining the two methods (EnKF and RRSQRT) is to remove the contributions in the direction of the leading eigenvalues from the start with no need for orthogonal projection.

This can be done starting from the POEnKF algorithm. The new approach is called COFFEE (Heemink et al. 2001). In the initialization step of this sequential filter we define two sets: $\mathbf{l}_i^a$ are the leading eigenvectors of $\mathbf{P}_0$ and the N ensemble members $\boldsymbol{\xi}_i^a$, which are generated this time to represent only the part from the covariance matrix that will be truncated: $\mathbf{P}_0 - \mathbf{L}_0^a[\mathbf{L}_0^a]^{\mathrm{T}}$.

In the forecast step of COFFEE the $\mathbf{l}_i$ are propagated using the forecast step of the RRSQRT algorithm [Eqs. (14)–(17)] and the ensemble members ξ_i are propagated via the model by the forecast step from the EnKF algorithm [Eqs. (9)–(11)]. However, this time the random ensemble is generated with an extra covariance equal to the covariance of the truncated part of the RRSQRT part; furthermore, there is no system error for the ensemble part. This changes Eq. (11) into

$$\mathbf{E}^{\mathrm{COFFEE}}(t_k) = [\boldsymbol{\xi}_1^f(t_k) - \mathbf{x}^f(t_k) + \boldsymbol{\eta}_1(t_k), \ldots, \boldsymbol{\xi}_N^f(t_k) - \mathbf{x}^f(t_k) + \boldsymbol{\eta}_N(t_k)], \quad (21)$$

where

$$\mathbf{E}[\boldsymbol{\eta}_i(t_k)\boldsymbol{\eta}_i(t_k)^{\mathrm{T}}] = [\mathbf{I} - \boldsymbol{\Pi}^f(t_k)]\mathbf{L}^f(t_k)\mathbf{L}^f(t_k)^{\mathrm{T}}[\mathbf{I} - \boldsymbol{\Pi}^f(t_k)]^{\mathrm{T}}. \quad (22)$$

In the forecast step of COFFEE, the covariance matrix is truncated in the RRSQRT part, and the rest that is of no interest is represented by an ensemble and added to the EnKF part.

The analysis step is the same as in the POEnKF, where the only exception is that there is no longer a need for a projection:

$$\mathbf{P}^f(t_k) = \mathbf{L}^f(t_k)\mathbf{L}^f(t_k)^{\mathrm{T}} + \frac{1}{N-1}\mathbf{E}^{\mathrm{COFFEE}}(t_k)\mathbf{E}^{\mathrm{COFFEE}}(t_k)^{\mathrm{T}}. \quad (23)$$

The reduced-rank part acts as a variance reducer for the ensemble filter, thus reducing the statistical errors of the Monte Carlo approach. From another point of view, by coupling an EnKF in the reduced-rank filter, the covariance matrix is prevented from being underestimated any longer, eliminating the filter-divergence problem of the reduced-rank approach. For a better understanding of how the COFFEE algorithm was constructed, all the steps for EnKF, RRSQRT, and COFFEE algorithms are presented in Table 1.

3. EUROS model and ground-based observations

The algorithms presented in section 2 were tested in a data assimilation experiment using EROS, an Eulerian atmospheric chemistry transport model developed at the National Institute of Public Health and the Environment (Van Pul et al. 1996; Matthijsen et al. 2001). The model, applied to simulating emissions, transport processes, chemical transformation, and dry and wet deposition processes of various air-polluting components, can be used to examine the time and spatial behavior of SO_x, NO_x, O_3, and volatile organic compounds (VOCs) in the lower troposphere over Europe. The model area extends over a large part of Europe. The base grid consists of 52×55 grid cells with a $0.55° \times 0.55°$ latitude–longitude resolution (see Fig. 1).

The vertical stratification of EUROS consists of four layers, from the surface up to 3000 m. The height of these layers varies during the day with the thickness of the mixing layer. The chemical scheme used in EUROS is a simplified 15-component scheme with 17 reactions. The 15-component scheme, based on a model described by Stedman and Williams (1991), describes the essential characteristics of the life cycles and time scales involved in the formation of ozone in the boundary layer.

The measurements used in the data assimilation process are hourly ground-based measurements of O_3 at rural background sites below 500 m above sea level. Sites near industrial or densely populated areas were discarded because these are influenced too much by local emissions to be well represented on a coarse grid. The mountains sites are also discarded because the local topography has a large influence on the results (Segers 2002). The O_3 data were available via the AirBase database (available online at http://air-climate.eionet.europa.eu/databases/airbase/index_html) for the year 2000 for about 240 stations in Europe. When performing data assimilation, it is customary to split the observation stations into two sets: one for the assimilation and the other solely for validation of the assimilated concentrations. Hanea et al. (2004) showed assimilation with Kalman filter algorithms, the EUROS model, and ozone observations in Europe to work well.

According to Flemming et al. (2003), a standard deviation of the error for each measurement is assumed to be equal to 10 μg m^{-3} for O_3. These errors account for both the uncertainty in the actual concentrations measured at the specific station and for a representation error, reflecting the size of the grid cell.

The goal of the data assimilation experiments in this research is to see how the hybrid filters perform with a large-scale application, and what one can gain from it. The nonlinearity measure is also a way of understand-

TABLE 1. The explicit equations for all the steps in the EnKF, RRSQRT, and COFFEE algorithms.

EnKF algorithm	RRSQRT algorithm	COFFEE algorithm
	1) Initialization step	
An ensemble of N states $\boldsymbol{\xi}_i^a(\mathbf{l}_0)$ are generated to represent the uncertainty in $\mathbf{x}_0$.	$\mathbf{x}^a(t_0) = \mathbf{x}_0$, and $\mathbf{L}^a(t_0) = [\mathbf{l}_1^a(t_0), \ldots, \mathbf{l}_q^a(t_0)]$ (T1)	$[\mathbf{L}^a(t_0)\mathbf{E}^a(t_0)] = [\mathbf{l}_1^a(t_0), \ldots, \mathbf{l}_q^a(t_0), \boldsymbol{\xi}_1^a(t_0), \ldots, \boldsymbol{\xi}_N^a(t_0)]$ (T2)
	2) Forecast step	
$\boldsymbol{\xi}_i^f(t_k) = \mathcal{M}[\boldsymbol{\xi}_i^a(t_{k-1})] + \mathbf{w}_i(t_{k-1})$ (T3)	$\mathbf{x}^f(t_k) = \mathcal{M}[\mathbf{x}^a(t_{k-1})]$ (T6)	$\mathbf{L}^a$ is updated using the same Eqs. (T6)–(T9) from the RRSQRT algorithm.
$\overline{\mathbf{x}}^f(t_k) = (1/N)\Sigma_{i=1}^N \boldsymbol{\xi}_i^f(t_k)$ (T4)	$\mathbf{l}_i^f(t_k) = (1/\epsilon)\{\mathcal{M}[\mathbf{x}^a(t_{k-1}) + \epsilon\mathbf{l}_i^a(t_{k-1})] - \mathcal{M}[\mathbf{x}^a(t_{k-1})]\}$ (T7)	The ensemble $\boldsymbol{\xi}_i^a$ is updated using the same Eqs. (T3)–(T4).
$\mathbf{E}^{f,\mathrm{EN}}(t_k) = [\boldsymbol{\xi}_1^f(t_k) - \overline{\mathbf{x}}^f(t_k), \ldots, \boldsymbol{\xi}_N^f(t_k) - \overline{\mathbf{x}}^f(t_k)]$ (T5)		Equation T5 becomes $\mathbf{E}^f(t_k) = [\boldsymbol{\xi}_1^f(t_k) - \mathbf{x}^f(t_k) + \boldsymbol{\eta}_1(t_k), \ldots, \boldsymbol{\xi}_N^f(t_k) - \mathbf{x}^f(t_k) + \boldsymbol{\eta}_N(t_k)]$. (T8)
	3) Truncation step	
—	$\tilde{\mathbf{L}}^f(t_k) = [\mathbf{l}_1^f(t_k), \ldots, \mathbf{l}_q^f(t_k), \mathbf{Q}(t_{k-1})^{1/2}]$ (T9) $\mathbf{L}^f(t_k) = \boldsymbol{\Pi}^f(t_k)\tilde{\mathbf{L}}^f(t_k)$ (T10)	Representation of the truncated energy (the part that is truncated in case of RRSQRT is taken into account here) $\mathbf{E}[\boldsymbol{\eta}_i(t_k)\boldsymbol{\eta}_i(t_k)^{\mathrm{T}}] = [\mathbf{I} - \boldsymbol{\Pi}^f(t_k)]\mathbf{L}^f(t_k)\mathbf{L}^f(t_k)^{\mathrm{T}}[\mathbf{I} - \boldsymbol{\Pi}^f(t_k)]^{\mathrm{T}}$ (T11)
	4) Analysis step	
$\mathbf{P}^f(t_k) = (1/N-1)\mathbf{E}^{f,\mathrm{EN}}(t_k)[\mathbf{E}^{f,\mathrm{EN}}(t_k)]^{\mathrm{T}}$ (T12)	$\mathbf{P}^f(t_k) = \mathbf{L}^f(t_k)[\mathbf{L}^{\mathrm{f}}(t_k)]^{\mathrm{T}}$ (T15)	$\mathbf{P}^f(t_k) = \mathbf{L}^f(t_k)\mathbf{L}^f(t_k)^{\mathrm{T}} + (1/N-1)\mathbf{E}^f(t_k)\mathbf{E}^f(t_k)^{\mathrm{T}}$ (T20)
$\mathbf{K}(t_k) = \mathbf{P}^f(t_k)\mathbf{H}(t_k)^{\mathrm{T}}[\mathbf{H}(t_k)\mathbf{P}^f(t_k)\mathbf{H}(t_k)^{\mathrm{T}} + \mathbf{R}(t_k)]^{-1}$ (T13)	$\mathbf{K}(t_k) = \mathbf{P}^f(t_k)\mathbf{H}(t_k)^{\mathrm{T}}[\mathbf{H}(t_k)\mathbf{P}^f(t_k)\mathbf{H}(t_k)^{\mathrm{T}} + \mathbf{R}(t_k)]^{-1}$ (T16)	$\mathbf{K}(t_k) = \mathbf{P}^f(t_k)\mathbf{H}(t_k)^{\mathrm{T}}[\mathbf{H}(t_k)\mathbf{P}^f(t_k)\mathbf{H}(t_k)^{\mathrm{T}} + \mathbf{R}(t_k)]^{-1}$ (T21)
$\boldsymbol{\xi}_i^a(t_k) = \boldsymbol{\xi}_i^f(t_k) + \mathbf{K}(t_k)[\mathbf{y}^o(t_k) - \mathbf{H}(t_k)\boldsymbol{\xi}_i^f(t_k) + \mathbf{v}_i(t_k)]$ (T14)	$\mathbf{x}^a(t_k) = \mathbf{x}^f(t_k) + \mathbf{K}[\mathbf{y}^o(t_k) - \mathbf{H}(t_k)\mathbf{x}^f(t_k)]$ (T17) $\tilde{\mathbf{L}}^a(t_k) = \{[\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k)]\mathbf{L}^f(t_k), \mathbf{K}(t_k)\mathbf{R}(t_k)^{1/2}\}$ (T18) $\mathbf{L}^a(t_k) = \boldsymbol{\Pi}^a(t_k)\tilde{\mathbf{L}}^a(t_k)$ (T19)	Followed by Eq. (T14) from the ENKF algorithm and Eqs. (T17)–(T19) from the RRSQRT algorithm.

ing how the hybrid algorithms deal with the nonlinear character of the assimilation.

4. Results of the Kalman filter algorithms

In evaluating the filter algorithms we used, as error criterion, the root-mean-square (RMS) differences averaged over all sites with observations not used in the assimilation (so-called validation stations) and all hours between the observed ozone concentrations and those obtained using a particular filter. Experiments with the hybrid algorithms, POEnKF and COFFEE, were described in Heemink et al. (2001) using a twin experiment for a 2D advection diffusion equation. Both algorithms improve the convergence for a small ensemble size, with the COFFEE variant being more efficient. This is why we chose to use only the COFFEE filter for applications to large-scale nonlinear data assimilation problems.

a. System errors

The stochastic environment is based on both a specification of the errors of the deterministic model EUROS and on the observations. All sorts of model or input parameters can be defined as uncertain and optimized in the assimilation. In this study we applied noise to the emissions of NO_x and VOC, because they are major sources of uncertainties in atmospheric modeling, and have been shown to be good candidates for optimization in an assimilation (Hanea et al. 2004). In Hanea et al. (2004) one of the goals was to gain inside information in the uncertain parameters of the model and the impact of these parameters on the assimilation. We accomplished that by treating the parameters as model input, defining them as stochastic parameters, and making these parameters part of the model state using an augmented state vector. The details of the implementation of the noise are given in Hanea et al. (2004) and Velders et al. (2006, manuscript submitted to *J. Geophys. Res.*). The NO_x and VOC emissions are defined for seven different regions: Belgium–Netherlands–Luxembourg, Germany–Denmark, France, United Kingdom–Ireland, Spain–Portugal, Switzerland–Austria–Italy–Czech Republic–Slovakia–Hungary–Poland, and Norway–Sweden–Finland–

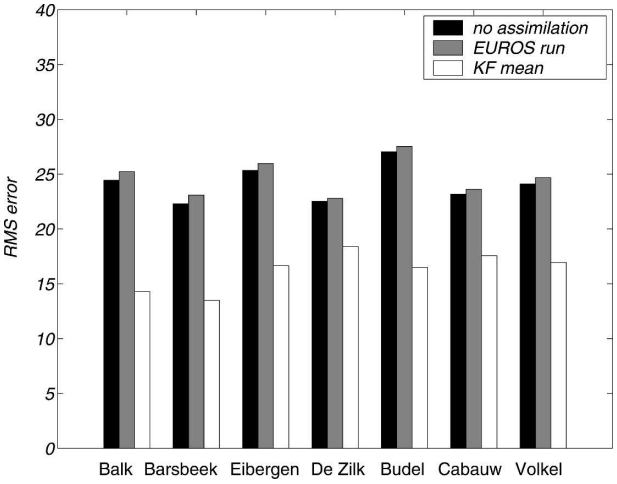

FIG. 1. RMS errors (μg m^{-3}) for an assimilation process (EnKF; 50 ensemble members) for all validation stations in the Netherlands for the EUROS model calculation, the Kalman filter mean, and the mean without assimilation. Simulations are for June 2000.

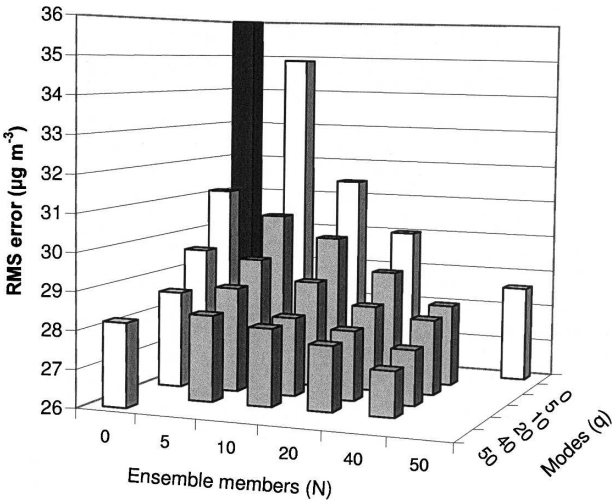

FIG. 2. RMS errors (μg m^{-3}) averaged over all validation stations in Europe, calculated with the EnKF, RRSQRT, and COFFEE filters for various numbers of modes q and ensemble members N. The columns with $q = 0$ contain the EnKF results and the columns with $N = 0$ contain the RRSQRT results. The RMS error of the EUROS model calculation $N = 0$ and $q = 0$ is 42.4 μg m^{-3} and it is truncated in the figure.

Estonia–Latvia–Lithuania. For the first four regions there are a large number of observations of O_3 at rural sites. Observations for the last three regions are included to reduce effects on the boundaries of the first four regions. There is not much information available in the literature on uncertainties in NO_x and VOC emissions, so for this study we set the uncertainties for the emissions for all countries at 10% for the first four regions and 20% for the last three, and applied a time correlation parameter of 24 h (Velders et al. 2006, manuscript submitted to *J. Geophys. Res.*).

A run of the measurements with an EnKF with no assimilation was performed to see if the stochastic environment built around the EUROS model (Hanea et al. 2004) could capture the underlying physics of the atmospheric processes. In Fig. 1 the first bar represents the residual obtained when there is no assimilation process present. In comparison with the second bar, which represents the residual of the EUROS model, the first residuals are observed to be always smaller than those given by the deterministic model. One can calculate the following:

$$E[f(\xi)] = f[E(\xi)] + \frac{1}{2}\frac{\partial^2 f}{\partial \xi^2} E\{[\xi - E(\xi)]^2\} + \ldots, \tag{24}$$

where the "..." represents the higher-order terms. The second term in the right-hand side of the equality disappears only in case of a linear function. In the presence of nonlinearities it is nonzero, and therefore we have differences between the runs of the EnKF with out assimilating data and the model run. That means that the ensemble forecast is more accurate than EUROS and, in conclusion, 1) nonlinearities are important in our application and 2) the bias has an impact on both the model forecast and the assimilation process. Using the measurements in the data assimilation process gives us an idea about the impact of bias on the process. Having achieved an improvement in the estimation of the true state of the atmosphere given by the stochastic model, the assimilation process was turned on. Simulations with RRSQRT, EnKF, and COFFEE were performed in which the information from measurements was assimilated.

b. Convergence of the algorithms

The following filter setups were examined to gain insight into the efficiency during the assimilation process: EnKF and RRSQRT filter with $N = 5$, 10, 20, 50, 100, and 200 ensemble members/modes and a COFFEE with $q = 5$, 10, 20, and 40 modes and $N = 5$, 10, 20, and 40 ensemble members. The number of model evaluations required comes to N for the EnKF, $1 + q$ for the RRSQRT filter and $1 + q + N$ for the COFFEE filter. The assimilation period was the month of July 2000.

Considering first the RRSQRT and EnKF filter simulations in Figs. 2 and 3 allows us to see that the RMS error of the RRSQRT simulations decreases faster with an increasing number of modes than the error of the EnKF simulations, but the value for 100 and 200 modes/ensemble members is about the same

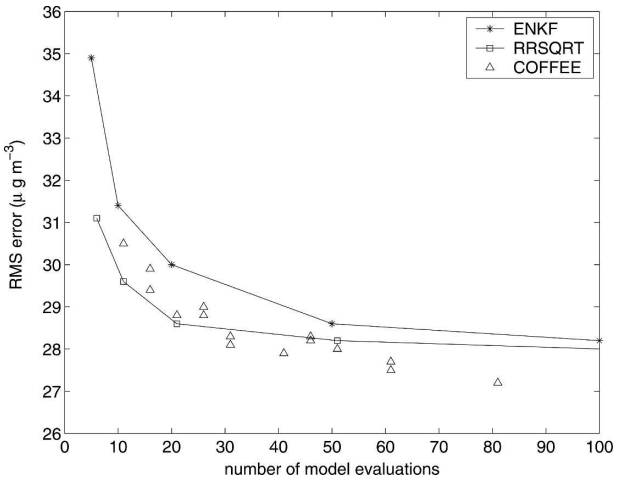

FIG. 3. RMS errors (μg m^{-3}) averaged over all validation stations in Europe, calculated with the EnKF, RRSQRT, and COFFEE filters vs the number of model evaluations. The RMS error of the EUROS model calculation (without assimilation) is 42.4 μg m^{-3}. The RMS errors for the runs with EnKF and RRSQRT with 200 ensemble members/modes are 28.1 and 27.9 μg m^{-3}, respectively.

for both. The RMS error of the COFFEE simulations is lower than the error of the RRSQRT and EnKF simulations for equal number of modes or ensemble members. This is clearly the result of the combination of using both modes and ensembles members, and an increased number of model evaluations. There does not seem to be a clear preference in the COFFEE results for either modes (q) or ensembles members (N) with respect to lower RMS errors. The RMS errors initially drop somewhat faster with fixed q and increasing N than with fixed N and increasing q, but similar values are obtained for $q = 20$, $N = 5$ and $q = 5$, $N = 20$, for example. The COFFEE filter produces better results than EnKF and RRSQRT filters for large number of modes/ensemble members. The COFFEE results for $N = 40$, $q = 10$ and $q = 40$, $N = 10$ are below the RRSQRT and EnKF results with 100 and 200 modes or ensemble members. The performances of the different filters are compared with respect to the RMS errors and the number of model evaluations required in the simulations. The number of model evaluations is not the same as the total central processing unit (CPU) time of a simulation, but a good indication for it. Considering the currently employed 15 chemical species, the EUROS model calculations use about 60% of the total CPU time in an EnKF simulation, about 55% in an RRQSRT simulation, and about 40% in a COFFEE simulation. The rest of the time is consumed by solving the Kalman filter equations and the handling of the observations. The fraction of the total CPU time used by the EUROS model will increase to 70%–90% of the total time when an extended version of the EUROS model with 32 components is used. Figure 3 shows the RMS errors obtained with the three filters versus the number of model evaluations.

The COFFEE results fall somewhere between the EnKF and RRSQRT results for less than about 30 model evaluations. For more than about 30 model evaluations, the COFFEE results are, in all cases, better than, or in a single case, equal to, the EnKF and RRSQRT results. The results of the COFFEE simulations with more than about 60 model evaluations turn out significantly better than all the EnKF and RRSQRT simulations (even better than those with 100 and 200 modes or ensemble members). From these simulations we can conclude that in our application the RRSQRT filter gives the best assimilation (i.e., the lowest RMS errors) when using less than 30 model evaluations. In cases where a better assimilation is required, the COFFEE filter algorithm shows the lowest RMS errors and also requires fewer model evaluations.

The model error covariance was set relatively high to compensate for the characteristic underestimation of forecast error variance by the RRSQRT. The impact on the performance of RRSQRT, EnKF, and COFFEE algorithms is not very significant. A choice of a very high $\mathbf{Q}$ tends to make the differences less.

A disadvantage of COFFEE is that it requires a state vector for each observation. A simulation might therefore require a large, or too large, amount of computer memory for applications with a large state vector (as in atmospheric applications), and a large number of observations (e.g., using satellite data).

c. *Nonlinearity*

The previously described filters vary in their treatment of the nonlinearity of an application (i.e., the EUROS model). When there are several ways at our disposal to apply data assimilation algorithms with different performance with respect to the nonlinearity of the model for a certain problem, it is important to know the degree of nonlinearity of this problem. Knowing how nonlinear the processes are, one can choose a suitable algorithm from the point of view of computational time and efficiency. This explains the introduction of a method for quantifying the nonlinearity of the data assimilation problem. This will help in choosing the right filters for an application.

One way of quantifying the relative importance of nonlinear system dynamics in data assimilation applications is based on the combination of a first-order extended and a truncated second-order Kalman filter, as introduced by Verlaan and Heemink (2001). For near-

linear dynamics the results of the two methods should be the same, while for nonlinear cases, the difference between them will increase with the degree of nonlinearity. The only difference between the extended Kalman filter and the truncated second-order filter is a bias term in the time propagation of the truncated second-order filter. If the error is defined by

$$\mathbf{e}^f(t_k) = \mathbf{x}^t(t_k) - \mathcal{M}[\mathbf{x}^a(t_{k-1})], \tag{25}$$

where $\mathbf{x}^t(t_k)$ is the true value of the state vector at time t_k, the bias can be defined as

$$\mathbf{b}^a(t_k) = E[\mathbf{e}^f(t_k)]. \tag{26}$$

Using these notations, we can derive the time propagation and the measurement propagation for the bias of the extended Kalman filter as follows:

$$\mathbf{b}^f(t_k) = \left[\frac{\partial \mathcal{M}}{\partial x}\right]_{x^a(t_{k-1})} \mathbf{b}^a(t_{k-1}) + \left[\frac{\partial^2 \mathcal{M}}{\partial x^2}\right]_{x^a(t_{k-1})} \mathbf{P}^a(t_{k-1}), \tag{27}$$

where

$$\mathbf{b}^a(t_k) = [\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k)]\mathbf{b}^f(t_k). \tag{28}$$

Bias, an average difference between model and observations over some time, can occur because of a number of reasons. First, the model may miss some of the relevant physics or the parameters of this model are not known accurately. Another cause is a bias in the observations (e.g., because the instrument is not calibrated well or drifts with age). But, there are also mathematical causes of biases (e.g., if a variable that is used in a nonlinear way contains some uncertainty, then the estimates may become biased).

In our approach we use the "mathematical" difference between the mean of the ensemble forecasts and the forecast of the ensemble mean as an estimate of nonlinearity. This is also known as the bias between ensemble mean and central forecast and can easily be shown to be zero for any linear model, even if the model itself is biased compared with the observations. Therefore, the bias is assumed to be related to nonlinear dynamics only. The length of the bias is a suitable measure for the error in the filter due to nonlinearities. The definition used in Verlaan and Heemink (2001) computes the length of the bias relative to the covariance, which is a measure of the total error

$$V = \|\boldsymbol{\beta}\|, \tag{29}$$

where

$$\boldsymbol{\beta} = (\mathbf{L}\mathbf{L}^{\mathrm{T}})^{-1}\mathbf{L}^{\mathrm{T}}\mathbf{b}. \tag{30}$$

The m-vector $\boldsymbol{\beta}$ contains the coefficients of the projection of $\mathbf{b}$ on the columns of $\mathbf{L}$. This number is not affected by scaling the state variables or by any other linear state transformation. The V value was shown to be sensitive to all aspects that define whether a nonlinear model complicates the filter problem or not. Experiments with the Lorentz model (Heemink et al. 2001) showed that for the different parameter settings controlling these aspects, the error made by the filter is related to the value of V. For the largest values of V, the more complex forecast methods provided more accurate results, while performance for small values was indifferent. A large value of V means a highly nonlinear system, while a small value of V indicates a near-linear problem. The disadvantage of this measure is that it depends on the number of modes used in the low-ranking algorithms. Therefore, a growing number of modes will automatically lead to a larger value of V. Nevertheless, a measure of the linear or nonlinear regime of the application is given by $V/\sqrt{n}$. If $V/\sqrt{n} \ll 1$ the bias is insignificant; if the value of $V/\sqrt{n} \gg 1$, the bias is significant (Heemink et al. 2001). To avoid this problem, Segers (2002) introduced two other measures based on the one presented above, but independent of the number of modes. The measures compare the bias vector with the subspace in which the filter assimilates measurements. For a low-ranking filter, the error between mean and true state is supposed to be a sample of $\mathbf{L}\mathbf{w}$ for some $\mathbf{w} \sim N(0, \mathbf{I}_m)$. The bias should be small compared with $\mathbf{L}\mathbf{w}$, otherwise the analysis is not able to account for the nonlinearity error. The measure used in this research computes the angle between the bias $\mathbf{b}$ and its projection $\mathbf{L}\boldsymbol{\beta}$ on the subspace spanned by $\mathbf{L}$:

$$\cos(\theta) = \frac{\|\mathbf{L}\boldsymbol{\beta}\|}{\|\mathbf{b}\|}. \tag{31}$$

If a large part of the bias is not in the subspace spanned by the columns of $\mathbf{L}$, this will be visible in the angle θ.

The nonlinearity values were investigated for a better understanding of the performance of the two classical algorithms, EnKF and RRSQRT. The value of V increases with the numbers of modes used in the different Kalman filters algorithms, as expected from theory. In Fig. 4 the values for V are plotted as time series.

The average value of V for the specified simulations increases from 2.8, when 5 modes are used, to 4.8 when 50 modes are used, and to 9.4 when 100 modes are used. At the same time, increasing the number of ensemble members causes the variability of V to drop; the patterns in the assimilation process can be identified when runs with 50 and 100 modes are considered. Figure 5

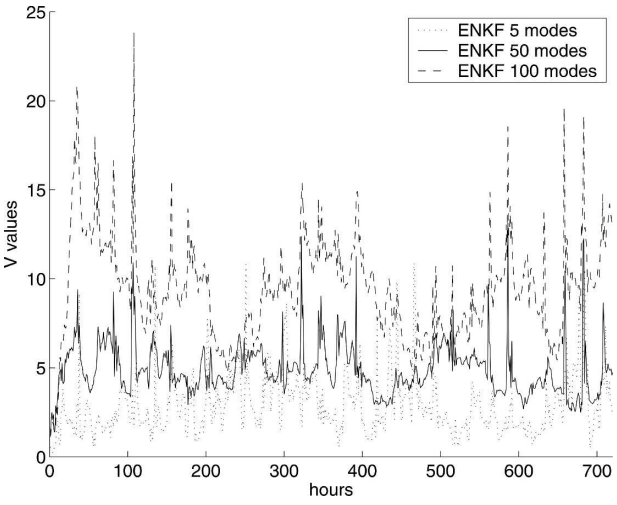

FIG. 4. The nonlinearity measure V calculated for the whole assimilation (hourly) period for EnKF simulations with different number of ensemble members.

shows the plotted values of $V/\sqrt{n}$, obtained using an EnKF with a different number of ensemble members ($N = 5$, 10, 20, 50, and 100). The convergence of the average values with the number of ensemble members and a value at the limit of approximately 0.7 leads to the conclusion that the bias is not significant in the assimilation process. In the case of a highly nonlinear system this value should have been much larger than 1. The temporal behavior is not very consistent; many members are needed for convergence. Therefore, in this case, the assimilation process is not a highly nonlinear one and the hybrid algorithm, COFFEE, will result in an improvement in a linear regime. Another measure for the amount of information that can be gained by increasing the number of modes is the angle θ. In both the EnKF and RRSQRT cases, a larger part of the subspace is spanned with an increasing ensemble size, reducing the angle. Convergence of the angles in Fig. 6 slows down with the number of ensemble members/modes because part of the bias is in the direction with small random errors and also because of the slow convergence of sampling errors. The angle measure seems more related to ensemble size than to nonlinearity. The assimilation process can be concluded to be nonlinear, because an increasing number of modes is needed to catch the nonlinearities; however, at the same time, the nonlinearity regime is not highly nonlinear, because, at some point, the estimation for certain values of the number of modes cannot be improved.

5. Conclusions

In this research a hybrid filter algorithm, COFFEE, was implemented and tested for a large-scale atmospheric chemistry transport model.

Depending on the nonlinearity of the data assimilation process one or another algorithm is best suited. For the hybrid algorithm the nonlinearities play a key role

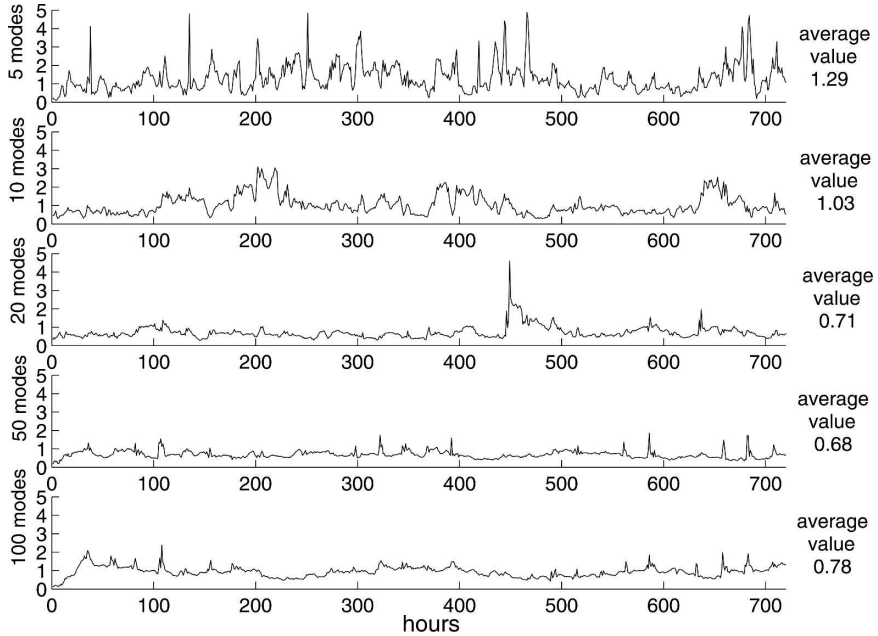

FIG. 5. The nonlinearity measure $V/\sqrt{n}$ for different numbers of ensemble members in EnKF simulations.

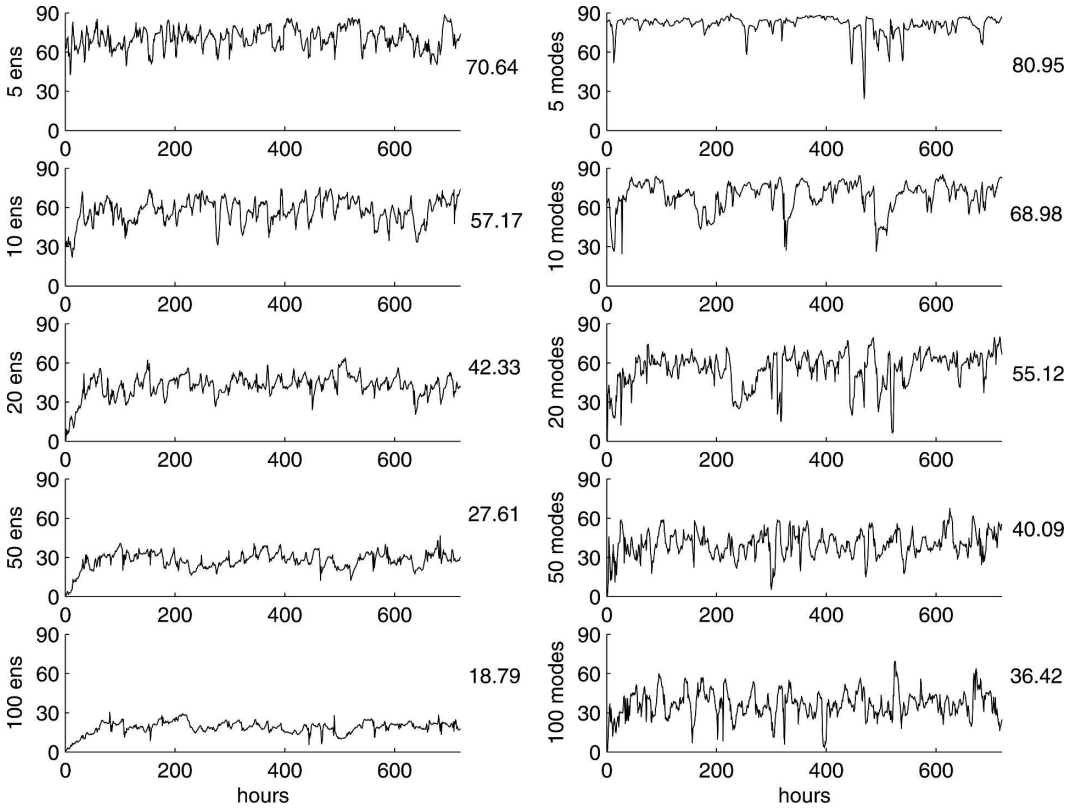

FIG. 6. The nonlinearity angle θ for EnKF and RRSQRT simulations with different numbers of ensemble members/modes.

in choosing the ensemble sizes q and N. Therefore, a measure of nonlinearity was used here to analyze the atmospheric chemistry data assimilation problem and to find out the impact of the nonlinearities on the data assimilation process and how much of these nonlinearities are captured by the filter. Based on this diagnostic, if the model is nearly linear, one should use fewer ensembles in the EnKF part of the COFFEE algorithm. On the other hand if the nonlinearity measure shows a nonlinear model the amount of ensembles used should be increased in order to compensate for the effects of nonlinearity.

As a result of the nonlinearity, the ensemble forecast run without any assimilation is more accurate than EUROS estimations. Assimilating data into the model makes the assimilation process less nonlinear. This can be observed by analyzing the two measures of nonlinearity. The convergence of the average values for $V/\sqrt{n}$, with the number of ensemble members to a value of 0.7 leads to the conclusion that the bias is not significant in the assimilation process. Convergence of the angles slows down with the number of ensemble members/modes because part of the bias is in the direction where small random errors occur and also because of the slow convergence of sampling errors. The angle measure seems more related to ensemble size than to nonlinearity. The hybrid filter algorithm, COFFEE, is a suitable filter for assimilating data in our atmospheric chemistry model.

EnKF, compared with the RRSQRT filter, suffers from statistical errors and, therefore, converges more slowly for increasing ensemble size. On the other hand, EnKF is more suitable for strongly nonlinear systems and, therefore, the final estimate of EnKF will be more accurate.

The actual convergence behavior of both algorithms is highly related to the specific properties of the application and especially with the nonlinearity of the resulting data assimilation problem.

The COFFEE filter combines the best of both and includes both algorithms. Because of presence of the nonlinearities in our data assimilation problem, it is more accurate than the RRSQRT, and because COFFEE suffers less from statistical errors than EnKF, it converges faster (Heemink et al. 2001). Again, it depends on the specific application how the ensemble size N and the number of modes for RRSQRT part q have to be chosen to get the best performance. With less than

about 30 model evaluations, COFFEE results fall somewhere between EnKF and RRSQRT results. With more than about 30 model evaluations, the COFFEE results are almost always better than the EnKF and RRSQRT results. The results of the COFFEE simulations with more than about 60 model evaluations are significantly better than all the EnKF and RRSQRT simulations (even better than those with 100 and 200 modes or ensemble members).

REFERENCES

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: On the analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.,* **126,** 1719–1724.

Canizares, R., 1999: On the application of data assimilation in regional coastal models. Ph.D. thesis, Delft University of Technology, Delft, Netherlands, 133 pp.

Cohn, S. E., and R. Todling, 1995: Approximate Kalman filters unstable dynamics. *Second Int. Symp. on Assimilation of Observations in Meteorology and Oceanography,* Tokyo, Japan, WMO, 241–246.

——, and ——, 1996: Approximate data assimilation schemes for stable and unstable dynamics. *J. Meteor. Soc. Japan,* **74,** 63–75.

Elbern, H., H. Schmidt, and A. Ebel, 1997: Variational data assimilation for tropospheric chemistry modeling. *J. Geophys. Res.,* **102,** 15 967–15 985.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99,** 10 143–10 162.

——, and P. J. van Leeuwen, 1996: Assimilation of Geosat altimeter data for the Agulhas current using the ensemble Kalman filter with a quasigeostrophic model. *Mon. Wea. Rev.,* **124,** 85–96.

Flemming, J., M. van Loon, and R. Stern, 2003: Data assimilation for CTM based on optimum interpolation and Kalman filter. *26th NATO/CCMS Int. Tech. Meeting on Air Pollution Modeling and Its Application,* Instanbul, Turkey, NATO/CCMS.

Hanea, R. G., G. J. M. Velders, and A. W. Heemink, 2004: Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport model. *J. Geophys. Res.,* **109,** D10302, doi:10.1029/2003JD004283.

Heemink, A. W., M. Verlaan, and A. J. Segers, 2001: Variance reduced ensemble Kalman filtering. *Mon. Wea. Rev.,* **129,** 1718–1728.

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

Ide, K., P. Courtier, and A. C. Lorenc, 1997: Unified notation for data assimilation: Operational sequential and variational. *J. Meteor. Soc. Japan,* **75,** 181–189.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory.* Vol. 64, *Mathematics in Science and Engineering,* Academic Press, 376 pp.

Matthijsen, J., L. Delobbe, F. Sauter, and L. de Waal, 2001: Changes of surface ozone over Europe upon the Gothenburg protocol abatement of 1990 reference emissions. *Proceedings of EUROTRAC Symposium 2000,* P. M. Midgley, M. Reuther, and M. Williams, Eds., Springer-Verlag, 1384–1388.

Segers, A., 2002: Data assimilation in atmospheric chemistry models using Kalman filtering. Ph.D. dissertation, Delft University of Technology, Delft, Netherlands.

Stedman, J. R., and M. L. Williams, 1991: A trajectory model of the relationship between ozone and precursors emissions. *Atmos. Environ.,* **26A,** 1271–1281.

Van Pul, W. A. J., J. A. van Jaarsveld, and C. M. J. Jacobs, 1996: Deposition of persistent organic pollutants over Europe. *Air Pollution Modeling and Its Application XI,* S. E. Gryning and F. A. Schiermeier, Eds., Plenum, 203–212.

Verlaan, M., and A. W. Heemink, 1995: Data assimilation schemes for non-linear shallow water flow models. *Proc. Second Int. Symp. on Assimilation of Observations,* Tokio, Japan, WMO, 245–252.

——, and ——, 1997: Tidal flow forecasting using reduced-rank square root filters. *Stochastic Hydrol. Hydraul.,* **11,** 349–368.

——, and ——, 2001: Nonlinearity in data assimilation applications: A practical method for analysis. *Mon. Wea. Rev.,* **129,** 1578–1589.