

Modeling and prediction of environmental data in space and time using Kalman filtering

A. W. Heemink, A. J. Segers

Abstract. The Kalman filter is used in this paper as a framework for space time data analysis. Using Kalman filtering it is possible to include physically based simulation models into the data analysis procedure. Attention is concentrated on the development of fast filter algorithms to make Kalman filtering feasible for high dimensional space time models. The ensemble Kalman filter and the reduced rank square root filter algorithm are briefly summarized. A new algorithm, the partially orthogonal ensemble Kalman filter is introduced too. We will illustrate the performance of the Kalman filter algorithms with a real life air pollution problem. Here ozone concentrations in a part of North West Europe are estimated and predicted.

Keywords: Distributed modeling, Environmental modeling, Kalman filtering

1

Introduction

Most environmental processes vary in space and time. Modelling these processes statistically from large data sets is far from trivial. The key problem is that the number of degrees of freedom in these identification problems is extremely large. In addition, physically based information of the process is required. Using traditional geostatistical methods, the variability of the process is often assumed to be spatially homogeneous (Cressie 1991). Using principal oscillation patterns (POPs), the process is considered to be a linear combination of a very limited set of spatial patterns (Hasselmann 1988). Here the patterns are all constant, while the weighing of the patterns varies in time. Other approaches to space–time modelling have been published among others by Bogaert and Christakos (1997) and by Christakos and Vyas (1998).

Wikle and Cressie (1999, 2000), introduced Kalman filtering as a framework for space time modelling of environmental processes. They developed a reduced dimension space time filter to analyse very large data sets. Their approach is based on a general statistical space time description of the process. Using Kalman filtering it is also possible to include physical knowledge into the data analysis procedure. For some environmental processes it may be possible to include the

A. W. Heemink (✉), A. J. Segers
Faculty of Information Technology and Systems
Delft University of Technology,
Mekelweg 4, NL-2628 CD Delft,
The Netherlands
e-mail: A.W.Heemink@math.tudelft.nl

The authors wish to thank Prof. Peter Bultjes, Dr. Maarten van Loon and Dr. Martin Verlaan for their contributions to the project.

concept of mass conservation or an advection-diffusion type model into the analysis. The basic idea of our paper is to use physically based numerical models together with the data. Integrating physically based model results with measurements is often called data assimilation.

In order to use a Kalman filter for assimilating data into a numerical model, a stochastic model or system is developed to model the inaccuracies of the underlying deterministic model. By using a Kalman filter, the information provided by the resulting stochastic dynamical model and the measurements are combined to obtain an optimal (minimal estimation variance) estimate of the state of the system.

In the last decade Kalman filtering has gained acceptance as a powerful framework for integrating space time data with physically-based models (Ghil and Malanotte-Rizzoli 1991). However, using the standard Kalman filter implementation would impose an unacceptable computational burden. In order to obtain a computationally efficient filter, approximations have to be introduced. The ensemble Kalman filter is based on a representation of the probability density of the state estimate by a finite number of randomly generated system states. A serious disadvantage of this approach is that the statistical error of the approach decreases very slowly with the sample size. Another approach for solving the large scale Kalman filtering problem is to approximate the full covariance of the state estimate by a matrix with reduced rank. To obtain a low rank approximation of the covariance matrix, a singular value decomposition is used to select the leading eigenvectors (EOF's) of the covariance matrix. The disadvantage of this algorithm is that by neglecting some of the eigenvectors the covariance matrix is under estimated. It is well-known that underestimating the covariance may introduce filter divergence. We propose using the partially orthogonal ensemble Kalman filter (POEnKF), where the reduced rank approximation is used as a variance reductor for the ensemble Kalman filter (EnKF) (Heemink et al. 2001). This algorithm combines the best properties of both previous approaches. It is less sensitive to divergence problems and computationally more efficient than the ensemble Kalman filter.

We first describe in Sect. 2 the original Kalman filtering approach. In Sect. 3 we introduce a number of fast algorithms for solving Kalman filtering problems with a very large number of variables. Finally in Sect. 4 an application of the Kalman filter to an ozone analysis and prediction problem is discussed in detail.

2

Kalman filtering

2.1

An environmental model for air pollution

The air pollution model used for this research is a condensed version of the Long Term Ozone Simulator (LOTOS) (Bultjes 1992). LOTOS is an Eulerian grid model used to study the controlling phenomena of ozone over Europe. The LOTOS model is based on the advection diffusion equation including source and sink terms for emissions, chemistry and deposition:

$$\frac{\partial c}{\partial t} = -\frac{\partial uc}{\partial x} - \frac{\partial vc}{\partial y} + \frac{\partial}{\partial x} \left(k_h \frac{\partial c}{\partial x} \right) + \frac{\partial}{\partial y} \left(k_h \frac{\partial c}{\partial y} \right) + E + R(c) - S(c) \quad (1)$$

where c is the concentration of one constituent, u, v are the wind velocities in respectively x - and y -direction, k_h is the dispersion coefficient, E represents the emissions, R is the reaction term and S is the sink term. The horizontal grid

spacing is 1.0° longitude by 0.5° latitude (about 60×60 km). The lowest 2–3 km of the troposphere are modeled using three layers with a depth depending on the height of the mixing layer (part of the meteorological input). The model includes a chemistry according to the CBM-IV mechanism (Gery et al. 1989). The chemical state in a grid cell is described in terms of concentrations (ppb) of 26 chemical components including ozone, nitrogenous oxides, carbon bounds, and radicals. Emissions of nitrogenous oxides, volatile organic compounds, and carbon monoxide are injected in the lowest layer following time dependent profiles. Components such as ozone and nitrogenous oxides are subject to dry deposition. Concentrations at ground level are therefore computed from a deposition profile and stored in a deterministic ground layer.

The version of the model used in this study was limited to a grid of 12×12 grid cells covering England and Wales (Fig. 1). This area was selected for its rather isolated position, which ensures that the bulk of the NO_x and VOC load arise from local emissions. Only long periods of eastern wind lead to a substantial inflow of pollutants from the continent. Besides, in this area a large number of ozone measurements from rural sites are available online (DENR 1997).

In discrete time form, the model follows the state space representation:

$$x_{k+1} = f(x_k) \quad (2)$$

where the vector x contains the concentrations in ppb of all 26 species of the CBM-IV mechanism, for each of the $12 \times 12 \times 4$ grid cells in the model (including the deterministic ground layer). The model provides hourly average concentration of all components in the states.

2.2

Stochastic state space model

The deterministic model (2) is able to produce the main trend of pollutant propagation, and it provides a reasonable picture of the transport dynamics of air

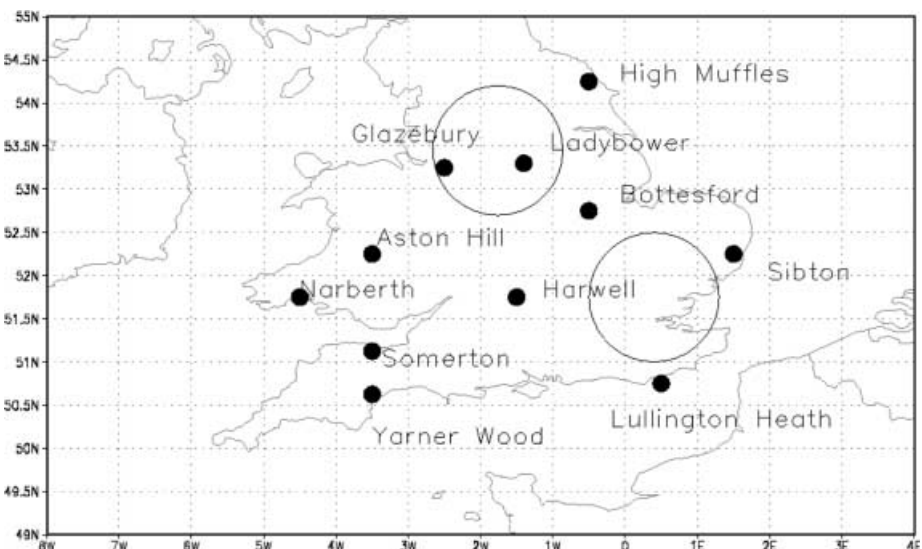


Fig. 1. Domain of the model area and the (rural) measurement sites. The circles denote the areas with the largest emissions of NO_x and VOC

pollution. However, since knowledge about the various physical phenomena is far from complete, the transport model is only an approximation and simplification of the real transport dynamics. The model results will never be perfectly consistent with the observations. By modelling the effects of variability and uncertainty in the natural system and model structure errors as system noise processes, a stochastic state space model is obtained:

$$x_{k+1} = f(x_k) + \Gamma_k w_k \quad (3)$$

where Γ_k is a noise input matrix and w_k is a Gaussian white noise process with covariance $\Sigma_{w,k}$. The initial condition x_0 is assumed to be Gaussian with mean \hat{x}_0 and covariance P_0 .

As mentioned before the model just described usually does not provide an accurate picture of the concentration. Therefore we want to use measurement information to improve our insight into the space time behaviour of the pollutant. Here we assume that the measurement y_k is related to the system state according to:

$$y_k = M_k x_k + v_k \quad (4)$$

where M_k is the measurement matrix and the measurement noise v_k is a Gaussian white noise processes with covariance $\Sigma_{v,k}$. The measurement noise is introduced to model measurement errors.

2.3

Conventional Kalman filtering for non linear systems

It is desired to combine the measurement modelled by relation (4) with the information provided by the system model (3) in order to obtain an optimal estimate of the system state x_k . If $\hat{x}_{k|l}$ is defined as the minimal variance estimator of x_k based on the measurements $y(1), \dots, y(l)$, and $P_{k|l}$ is the covariance matrix of the estimation error, recursive filter equations to obtain these equations are described below.

The optimal state estimate is propagated from measurement time k to measurement time $k + 1$ by the equations

$$\hat{x}_{k+1|k} = f(\hat{x}_{k|k}) \quad (5)$$

$$P_{k+1|k} = F_k P_{k|k} F_k^T + \Gamma_k \Sigma_{w,k} \Gamma_k^T \quad (6)$$

where

$$(F_k)_{ij} = \frac{\partial f(\hat{x}_{k|k})_i}{\partial (\hat{x}_{k|k})_j} \quad (7)$$

represents the tangent linear model.

At measurement time k the measurement y_k becomes available. The estimate is updated by the equations

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_k [y_k - M_k \hat{x}_{k+1|k}] \quad (8)$$

$$P_{k+1|k+1} = P_{k+1|k} - K_k M_k P_{k+1|k} \quad (9)$$

where

$$K_k = P_{k+1|k} M_k^T [M_k P_{k+1|k} M_k^T + \Sigma_{v,k}]^{-1} \quad (10)$$

is the Kalman gain. The initial condition for the recursion is given by $\hat{x}_{0|0} = \hat{x}_0$ and $P_{0|0} = P_0$.

By using the Kalman filter, both sources of information, model information and measurement information are integrated to obtain an optimal reconstruction of the concentration distribution. From the modelling point of view, measurements are used to reduce the model errors and to improve the model results. From the data point of view, model information in the form of a stochastic system representation, is used as a physically based interpolation scheme to fill in the gaps in the data set both in space as in time.

3

Kalman filter algorithms for high dimensional space time systems

3.1

Ensemble Kalman filter (EnKF)

The EnKF was introduced by (Evensen 1994) and has been used successfully in many applications (Evensen and Van Leeuwen 1996; Houtekamer and Mitchel 1998; Canizares 1999). This Monte Carlo approach is based on a representation of the probability density of the state estimate by a finite number N of randomly generated system states. The algorithm does not require a tangent linear model and is very easy to implement. The (EnKF) for the model (3)–(4) can be summarized as follows (Burgers et al. 1998):

Initialization:

An ensemble of N initial states $(\xi_i)_{0|0}$ are generated to represent the uncertainty in x_0 .

Time update:

$$(\xi_i)_{k+1|k} = f((\xi_i)_{k|k}) + \Gamma_k(w_i)_k \quad (11)$$

$$\hat{x}_{k+1|k} = \frac{1}{N} \sum_{i=1}^N (\xi_i)_{k+1|k} \quad (12)$$

$$E_{k+1|k} = [(\xi_1)_{k+1|k} - \hat{x}_{k+1|k}, \dots, (\xi_N)_{k+1|k} - \hat{x}_{k+1|k}] \quad (13)$$

Measurement update:

$$P_{k+1|k} = \frac{1}{N-1} E_{k+1|k} E_{k+1|k}^T \quad (14)$$

$$K_{k+1} = P_{k+1|k} M_{k+1}^T [M_{k+1} P_{k+1|k} M_{k+1}^T + \Sigma_{v,k+1}]^{-1} \quad (15)$$

$$(\xi_i)_{k+1|k+1} = (\xi_i)_{k+1|k} + K_{k+1}(y_{k+1} - M_{k+1}(\xi_i)_{k+1|k} + (v_i)_{k+1}) \quad (16)$$

here $(\xi_i)_{k|k}$ is an ensemble of state vectors generated with the realizations $(w_i)_k$ and $(v_i)_k$ of the noise process w_k and v_k respectively. Note that in the final implementation of the algorithm P need not actually be computed (Evensen 1994).

For most practical problems Eq. (11) is computationally dominant. As a result the computational effort required for the EnKF is approximately N model

simulations. The errors in the state estimate are of statistical nature and decrease very slowly with the sample size ($\approx \frac{1}{N}$). This is one of the very few drawbacks of this Monte Carlo approach.

3.2

Reduced rank square root filter (RRSQRT)

Another approach for solving large scale Kalman filtering problems is to approximate the full covariance matrix of the state estimate by a matrix with reduced rank. This approach was introduced by Cohn and Todling (1995, 1996), and Verlaan and Heemink (1995, 1997). The latter used a more robust square root formulation for the filter implementation. Algorithms based on similar ideas have been proposed and applied by Lermusiaux (1997) and Pham et al. (1998).

The reduced rank approaches can also be formulated as an EnKF where the q ensemble members are not chosen randomly, but in the directions of the q leading eigenvectors of the covariance matrix (Verlaan and Heemink 1997). As a result these algorithms do not require a tangent linear model.

The RRSQRT filter algorithm (Verlaan and Heemink 1997) is based on an approximation of the covariance matrix of the state estimate according to $P = LL^T$, where L is a matrix with the q leading eigenvectors (eigenvectors with largest eigenvalues) l_i , $i = 1, \dots, q$ of P as columns. The algorithm can be summarized as follows:

Initialization:

$$\mathbf{x}_{0|0} = \mathbf{x}_0$$

$$L_{0|0} = [(l_1)_{0|0}, \dots, (l_q)_{0|0}]$$

Time update:

$$\hat{\mathbf{x}}_{k+1|k} = f(\hat{\mathbf{x}}_{k|k}) \quad (17)$$

$$l_{i,k+1|k} = \frac{1}{\varepsilon} (f(\hat{\mathbf{x}}_{k|k} + \varepsilon l_{i,k|k}) - f(\hat{\mathbf{x}}_{k|k})) \quad (18)$$

$$\tilde{L}_{k+1|k} = [l_{1,k+1|k}, \dots, l_{q,k+1|k}, \Gamma_k \Sigma_{w,k}^{1/2}] \quad (19)$$

$$L_{k+1|k} = \Pi_{k+1|k} \tilde{L}_{k+1|k} \quad (20)$$

where $\Pi_{k+1|k}$ is a projection onto the q leading eigenvectors of the matrix $\tilde{L}_{k+1|k} \tilde{L}_{k+1|k}^T$ and where ε is chosen close to 1.

Measurement update:

$$P_{k+1|k} = L_{k+1|k} L_{k+1|k}^T \quad (21)$$

$$K_{k+1} = P_{k+1|k} M_{k+1}^T [M_{k+1} P_{k+1|k} M_{k+1}^T + \Sigma_{v,k+1}]^{-1} \quad (22)$$

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + K_{k+1} (\mathbf{y}_{k+1} - M_{k+1} \hat{\mathbf{x}}_{k+1|k}) \quad (23)$$

$$\tilde{L}_{k+1|k+1} = [(I - K_{k+1} M_{k+1}) L_{k+1|k}, K_{k+1} \Sigma_{v,k+1}^{1/2}] \quad (24)$$

$$L_{k+1|k+1} = \Pi_{k+1|k+1} \tilde{L}_{k+1|k+1} \quad (25)$$

where $\Pi_{k+1|k+1}$ is a projection onto the q leading eigenvectors of the matrix $\tilde{L}_{k+1|k+1} \tilde{L}_{k+1|k+1}^T$. As in the EnKF the full covariance matrix need not to be computed (Verlaan and Heemink 1997). This measurement update is not the most efficient procedure. Equations (19)–(20) are, however, more general than measurement update proposed by Verlaan and Heemink (1997). Equations (24)–(25) holds for arbitrary filter gains K_k and not only for gain matrices satisfying Eq. (22). This becomes important if the RRSQRT algorithm is used as part of the POEnKF described in the next subsection.

For smaller values of q Eq. (18) is computationally dominant resulting in a computational effort of $q + 1$ model simulations. The projection (20) requires $O(q^3)$ computations. As a result for very large q this part of the algorithm becomes time consuming too (see Canizares 1999). The errors of the algorithm are caused by the linearized dynamics in Eq. (18) and the representation of the covariance matrix by only the q leading eigenvectors. Because a number of eigenvectors are neglected the covariance matrix is underestimated. As a result the algorithm is sensitive to filter divergence problems. This problem can be avoided by choosing q relatively large, but this obviously reduces the computational efficiency.

3.3

Partially orthogonal ensemble Kalman filter (POEnKF)

The EnKF and the RRSQRT filter are both of the more robust square root type since both algorithms are formulated in terms of the square root of the covariance matrix (respectively $E_{k+1|k}$ and $L_{k+1|k}$). Also, both algorithms are of the ensemble type. The ensemble filter is based on N randomly chosen ensemble members, while in the reduced rank filter the ensemble members are chosen deterministically in the direction of the q leading eigenvectors. Because the two algorithms both have a very similar algorithmic structure, they can be integrated relatively easy (Heemink et al. 2001).

The ensemble of the POEnKF consists of two parts. The q leading eigenvectors l_i of the covariance matrix plus N randomly chosen ensemble members ξ_i . The ensemble members l_i are updated using a RRSQRT algorithm and ξ_i by using the EnKF algorithm. The two algorithms interact with each other a the measurement update. Here only the information of the random ensemble members orthogonal to the q leading eigenvectors is used.

The POEnKF algorithm can be summarized as follows:

Initialization:

$$[L_{0|0}, E_{0|0}] = [(l_1)_{0|0}, \dots, (l_q)_{0|0}, (\xi_{q+1})_{0|0}, \dots, (\xi_{q+N})_{0|0}]$$

where l_i are the leading eigenvector of P_0 and ξ_i are generated randomly to represent the uncertainty in x_0 .

Time update:

Time update equations (11)–(13) of the EnKF algorithm

Time update equations (17)–(20) of the RRSQRT algorithm

Measurement update:

$$E_{k+1|k}^\perp = \Pi_{k+1|k}^\perp E_{k+1|k} \quad . \quad (26)$$

where $\Pi_{k+1|k}^\perp$ is a projection of the random ensemble members orthogonal to the first q ensemble members l_i .

$$P_{k+1|k} = L_{k+1|k} L_{k+1|k}^T + \frac{1}{N-1} E_{k+1|k}^\perp E_{k+1|k}^\perp \quad (27)$$

$$K_{k+1} = P_{k+1|k} M_{k+1}^T [M_{k+1} P_{k+1|k} M_{k+1}^T + \Sigma_{v,k+1}]^{-1} \quad (28)$$

Measurement update equation (16) of the EnKF algorithm for the ensemble $E_{k+1|k}$
Measurement update equations (23)–(25) of the RRSQRT algorithm

232 For small values of q the time propagation equations for the ensemble is computationally dominating. As a result for most practical problems the computational effort for the POEnKF is approximately $N + q$ times the effort required for one model simulation.

By integrating the EnKF and the RRSQRT filter the best of both are combined. The reduced rank part acts as a variance reductor for the ensemble filter reducing the statistical errors of this Monte Carlo approach significantly (Hammersley and Handscomb 1964). At the other hand by embedding the reduced rank filter in an EnKF the covariance is not underestimated, eliminating the filter divergence problems of the reduced rank approach (also for very small numbers of q). As a result q can be chosen on the basis of efficiency arguments and not for stabilizing the filter algorithm.

4

Application to ozone analysis and prediction problems

In this section we will describe the setup and the results of filter experiments with the LOTOS model. First, a description of the stochastic model will be given, Here the emission input of the model is modelled as being stochastic. Second the test period and the available measurements are described. Finally, the results of the analysis and prediction experiments are presented, including a comparison of the performance of the various filter algorithms.

4.1

Stochastic model

Important model errors are caused by erroneous emissions. The emission model is based on time profiles, which assign fractions of yearly averaged emissions to individual hours. Since the time profiles are based on average statistics, the actual emissions might be quite different from the modelled one. The yearly averaged emissions might also contain errors.

The following stochastic state space model is used for LOTOS with stochastic emissions (Zhang et al. 1999; Segers et al. 2000):

$$\begin{pmatrix} x_{k+1} \\ \delta e_{k+1} \end{pmatrix} = \begin{pmatrix} f(x_k, e_k[1 + \delta e_k]) \\ \alpha \delta e_k \end{pmatrix} + \begin{pmatrix} 0 \\ \sqrt{1-\alpha^2} \Gamma \end{pmatrix} w_k \quad (29)$$

where x_k denotes the concentration array of the deterministic model f as described in Sect. 2.1. The model state x_k has been augmented with the total emission in mol/min for each specific hour, equal to the default emission e_k multiplied with a correction factor $(1 + \delta e)$. In this research, the emissions of NOx, VOC and CO are considered to be uncertain, varying independent from each other with standard deviations of 30, 50, and 30% respectively. The emissions are disturbed in all grid

cells covering England and Wales equally. The factor δe is modelled as a colored noise process (Jazwinski 1970), forced by white noise input w_k . A time correlation parameter $\alpha = 0.95$ ensures that samples of δe are rather smoothed in time.

The specification of the model error statistics has been chosen on the basis of expert opinions, since there is hardly any information on model errors available. The various noise parameters of the Kalman filter have been validated by monitoring the innovations of the filter and by checking whether their theoretical statistics based on the model assumptions are in agreement with the measured statistics of the innovations (Jazawinski 1970; Segers et al. 2000).

4.2

Measurements and observation error

A 6 day period from August 5 till August 10, 1997 was selected as a test period. Ozone measurements for 11 rural sites are available during this period (Fig. 1). During the first days of the test period, the ozone concentrations are rather low, especially in the southern part of the domain. Investigation of the meteorological data shows this could be explained from a rather high cloud cover. Later on, the cloud cover decreases, and measured ozone levels start to show high peaks during the day, with maximum values of 100 ppb. The model simulates the lower concentrations in the beginning and the high peaks later on correct, but is not able to reproduce the height of the peaks accurately. Some of the peaks are underestimated with more than 30 ppb. Part of this misfit might be explained from the coarse resolution of the model, which tends to spread local high concentrations over a larger area. However, the height of the peaks are not reproduced for almost all sites at the same time, suggesting a more systematic underestimation of the ozone production.

The variability in the measurements is large: the measured values sometimes differ with more than 10 ppb from 1 h to another, while a long term mean over several hours is almost constant. One explanation is the occurrence of measurement errors, which are estimated to be 5–10% of the measured value (Tilmes and Zimmermann 1998). Local weather conditions are also expected to have a large impact on the variability, since a single hour of sunshine in a period with clouded sky might lead to a higher ozone level.

The ozone measurements from Harwell, Bottesford, and Glazebury were used for assimilation. The ozone levels at these sites are under direct influence of the largest emission sources.

4.3

Experiment: 6 days assimilation

In a first experiment, the selected ozone measurements have been assimilated in the model to see the impact on the ozone concentrations at other locations in the grid and on other components in the state. An ensemble filter with 40 ensemble members was chosen to obtain a benchmark for future experiments, since it produced an accurate answer to the nonlinear filter problem. Repeating this experiment did not change the results of the filter experiment significantly.

The time series for Glazebury in Fig. 2 show that the filter is able to decrease the difference between model and measurements significantly. The results of the filter are presented in the form of 2σ error bounds. The true ozone concentrations are expected to be within the bounds with a probability 95%. The high ozone levels during daytime that were not reproduced by the model are now estimated within the assumed measurement error. Only during the night, the very low concentrations are not reproduced by the filter. The difference with the model is

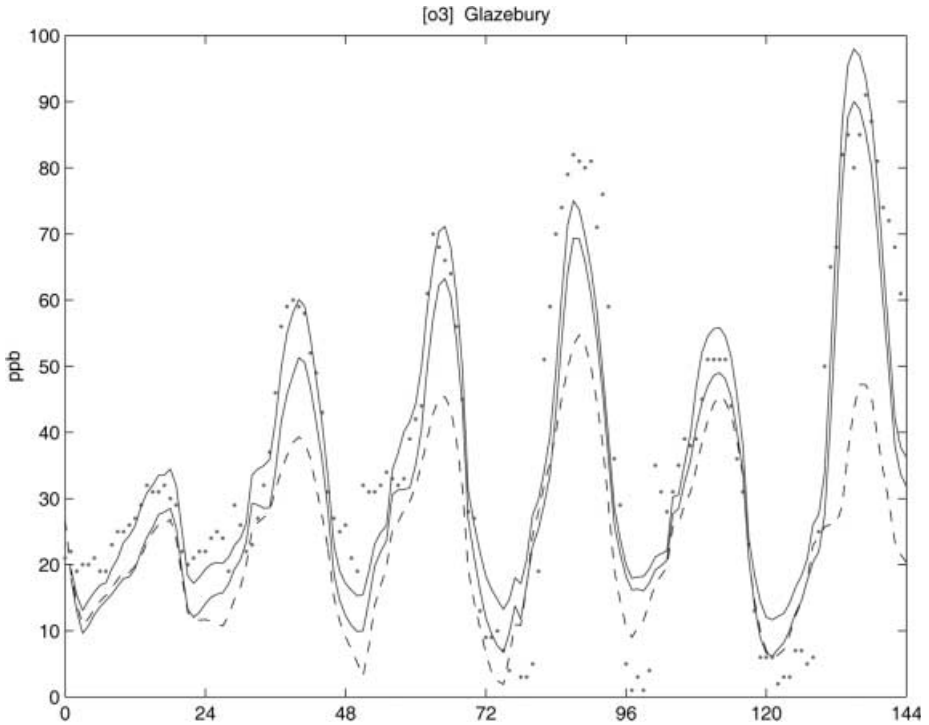


Fig. 2. Assimilation results for site Glazebury (assimilated): dots: measurements; dashed: model; solid: assimilated mean $\pm 2\sigma$

probably caused by other errors than those in the emissions, for example by errors in the deposition rates. Similar results were obtained for the other assimilated stations.

The spatial distribution of the adjustment at hour 87 (day 4, 15:00; Fig. 3) shows that the underestimation of ozone is corrected by a plume arising from the largest emission sources, flowing with the wind direction (north during day 4). The filter corrects the difference between model and measurements using decreased NO_x emissions and increased VOC emissions.

The filter tends decrease NO_x levels during the complete period. Comparison with NO_x measurements from Ladybower (Fig. 4) shows that this is during daytime in agreement with reality. The nighttime values are however decreased to far, which is another indication that the NO_x emissions are not the only error source in the model during the night.

Outside the plume, the adjustments of the ozone concentrations are minor. The ozone concentrations in Aston Hill for example not significant influenced by the changing emissions until day 6, when eastern wind has directed the emission plume towards the site. Clear sky conditions lead to increased ozone levels, underestimated by the model. Although the measurements are not assimilated, the filter produced an improved estimation of the ozone level due to the assimilation upwind from the site (Harwell). Improved results were also obtained for High Muffles, which is downwind from the assimilated sites Bottesford and Glazebury during episodes of southern wind. The results show that the filter technique is able to improve ozone simulations significantly, based on the assumption that the emissions are uncertain.

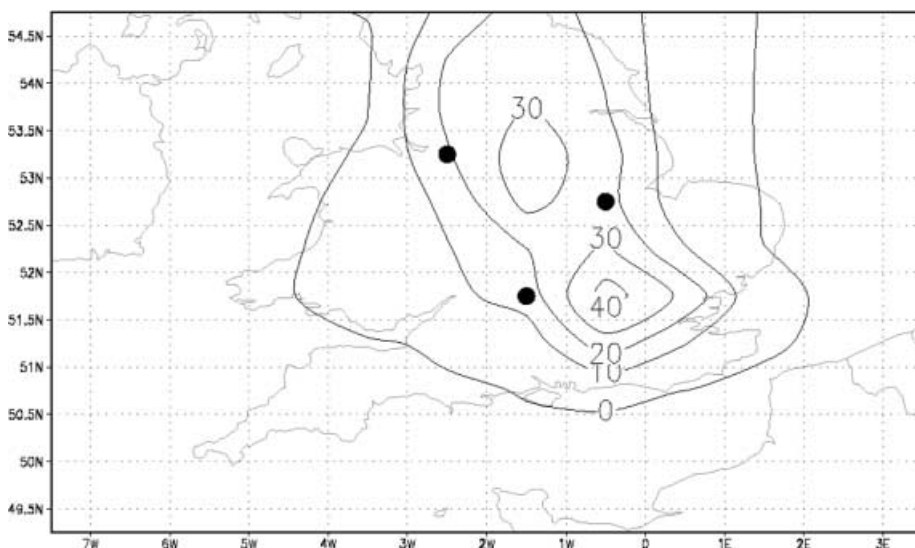


Fig. 3. Adjustment (assimilated mean minus model) of ozone concentration on day 4, 15:00. The adjustment of other components in the state show a similar spatial distribution. The largest adjustments are made around the most important emission sources, with values up to 60 ppbC for VOC and -8 ppbC for NO_x

4.4 Reconstruction of emissions

From the emission correction factors δe of the state, it is possible to obtain some insight in the actual values of the emissions estimated by the filter. Figure 5 shows the total amount of NO_x and VOC emitted in the area England/Wales by the model, and the 2σ error bounds during the 6-days assimilation experiment. Due to the modelling with time profiles, the default emissions show a periodic and block shaped pattern with high or low emission rates depending on different human activities during day or night, and week or weekend (day 5/6). The emissions of CO were found to be of minor impact for the ozone concentrations in comparison with the impact of NO_x and VOC, and are therefore omitted in the discussion. The adjustments to the NO_x emissions are rather small. During nighttime, the emissions are estimated to be smaller since the model tends to overestimate the ozone levels during the night. During daytime, the emissions start at a lower level following the trend from the night before, and increases to a peak emission just before sunset. The total amount of released NO_x has however hardly changed, from default 7.5 to 6.0 – 8.6×10^4 ppb N/min. Note the behavior in the morning of day 6: The night time emissions were estimated to be slightly increased, which cause a large additional emission when the emission profile changes to daytime rate. This higher rate is however soon regarded as a mistake, and the emissions fall back to a lower level. If the peak is omitted, the emissions follow a rather smoothed profile like the profiles on day 2, 3, and 5.

Concluding, the filter approach is able to produce more accurate ozone concentrations by adapting the emissions of NO_x and VOC. Including the correction factors for the emissions in the state provides useful insight in how the filters corrects discrepancies between model and measurements, although one should be careful with the interpretation. A conclusion from our experiments is that at least

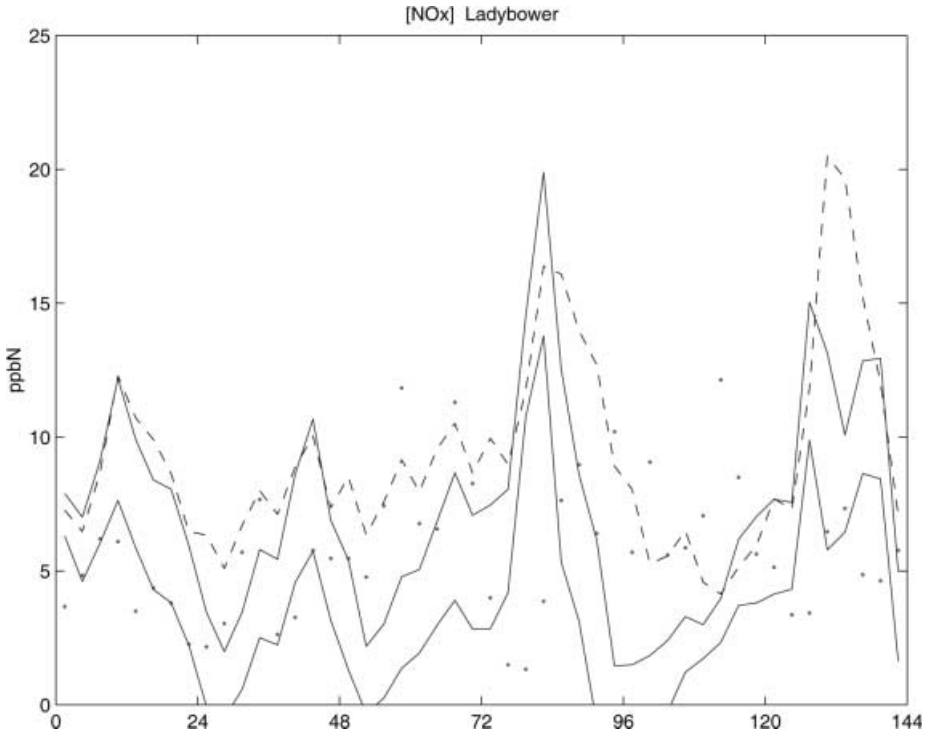


Fig. 4. Three hour averages of NOx concentrations in Ladybower according to measurements, the model, and the filter (same legend as Fig. 2). Model and filter results are a combination of values extracted from the mixing and reservoir layer, depending on the time

the time profiles used for the NOx emissions are subject to uncertainties, and that VOC emissions are under estimated.

4.5

Experiment: forecast of ozone concentrations

In a second experiment, the filter technique was tested for its capabilities to ozone forecasting. The mean state computed by the filter is in general a more accurate approximation of the true state than what is computed by a deterministic model run. If a model run is started with an assimilated mean as initial state, it is therefore expected to be more in agreement with measurements. This property could be used for an online forecasting system: The filter provides an accurate initial condition for a deterministic forecast run. Apart from the initial state, the quality of an ozone forecast also depends on the quality of the meteorological input. Since the forecast skill of numerical weather prediction is limited to about 5 days, one cannot expect an ozone forecast to be accurate over more than a few days.

The forecast skill of the filter has been tested starting from day 4, 15:00 (see Fig. 3 for the ozone distribution). During a first forecast run, model (29) was propagated to the end of the 6 day period forced by a noise input w_k equal to zero. The value of the emission adjustment δe_k decayed from its mean value at 15:00 to its first guess value zero with a rate of α^k . The results in Fig. 6 show that a forecast with zero noise input rapidly converges to a normal model run. After 24 h, the forecast is still in good agreement with the measurements (as the normal model run was). The high ozone level after 48 h is however not reproduced accurately. The forecast is still

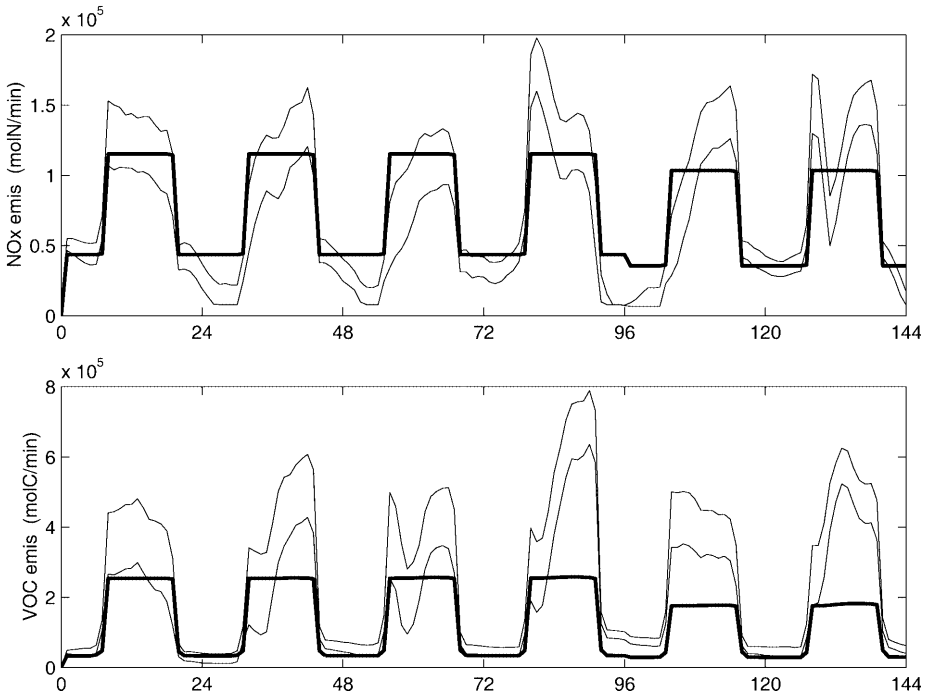


Fig. 5. Estimates of total emission in England/Wales; thick lines: model; thin lines: 2σ during assimilation

better than the model simulation, partly because of the better initial condition, and partly because still 10% of the emission adjustment is present in the state.

In a second run, the value of δe_k was fixed to the value at day 4, 15:00, by setting $\alpha = 0$. This setup reflects the idea that errors in the emission profiles are rather persistent in time as proved by Fig. 5. Figure 6 shows that fixed emissions improve the forecast skill significantly. The forecast of the maximum ozone level is close to the assimilated value.

4.6

Computational aspects

For each of the filter types described in Sect. 3, the efficiency has been judged based on a large number of experiments. To get insight in the efficiency during more complicated assimilations, the nonlinear character of the stochastic model was slightly increased by introducing uncertain emissions of NOx and VOC in two disjunct area instead of one, each covering one are a with large emissions. The assimilation period was limited to the last three days of the previous described period. The following filter setups have been examined: an ensemble filter with $N = 5, 10, 15,$ or 20 members, a RRSQRT filter with $q = 1-5, 8, 12,$ or 20 modes, and a POEnKF with $q = 1, 2, 4, 8,$ or 12 modes and $N = 5$ random ensembles. The results obtained for a particular filter setup have been compared with the results obtained with an ensemble filter with 80 members, serving as a reference solution of the filter problem. We used as error criterion the root mean square over all grid cells and all hours between the ozone concentrations of the ‘true’ solution and those obtained with a particular filter. To see the impact of the random number used in the EnKF and POEnKF algorithm, each of the

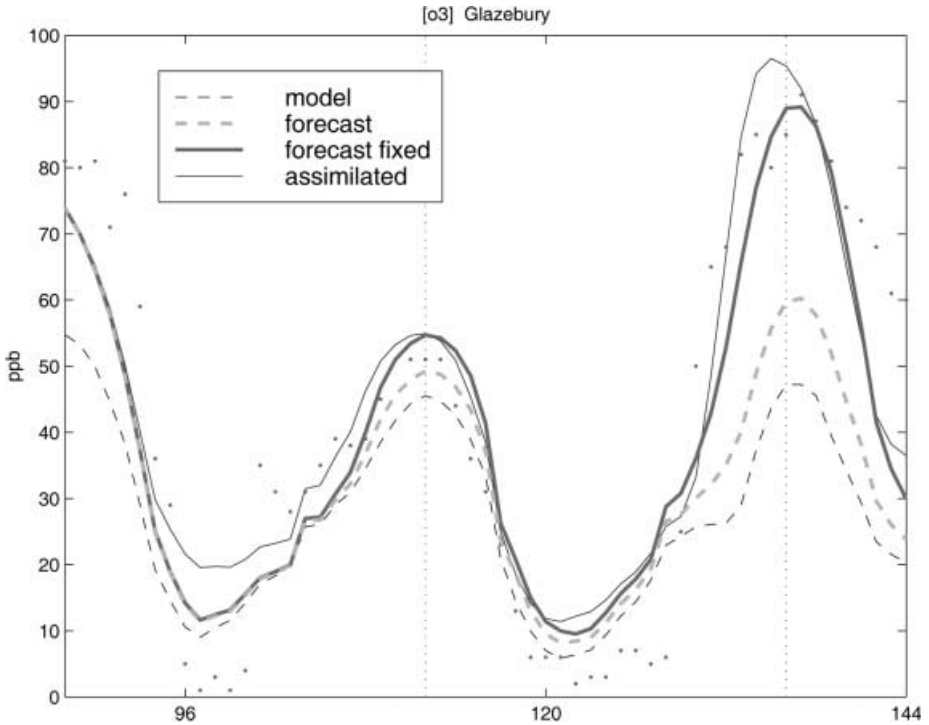


Fig. 6. Forecast of ozone concentration in site Glazebury, from 15:00 August 8 1997 to the end of August 10, as well as the assimilated mean value and a default model run. One and two day forecasts are marked by the dotted lines

experiments using one of these algorithms was repeated four times. The results are plotted in Fig. 7 versus the number of required model evaluations (N for EnKF, $1 + q$ for RRSQRT, and $1 + q + N$ for POEnKF). The rms error of the deterministic model was very large; the model underestimated the high ozone levels during daytime with 20–30 ppb, leading to an rms of 12.4. The results show that with only a few additional model evaluations, the filters are able to reduce the rms error with a factor 2. The slow convergence of the EnKF filter is shown by the large spread in the corresponding rms errors, even for large ensemble sizes. The results show that the number of modes used in the RRSQRT algorithm should exceed a critical level, before the filter converges. At least four modes are required for properly expression of the covariance. The best possible solution is obtained if the covariance is expressed in 12 modes or more. Comparing the results of the RRSQRT filter with the POEnK variant shows that the introduction of random ensembles only improves the results for small a number of modes significantly. Comparison of the POEnKF results with the EnKF results for five random ensembles shows that introduction of a few modes reduces the statistical error. Where the rms errors for the EnKF range from 3 to 6, the range is decreased to about 3–4 if two modes are ‘included’ in the algorithm.

5

Conclusions

In this paper the Kalman filter is introduced as a framework for analyzing space time data. Our approach is based on the use of physically based simulation models.

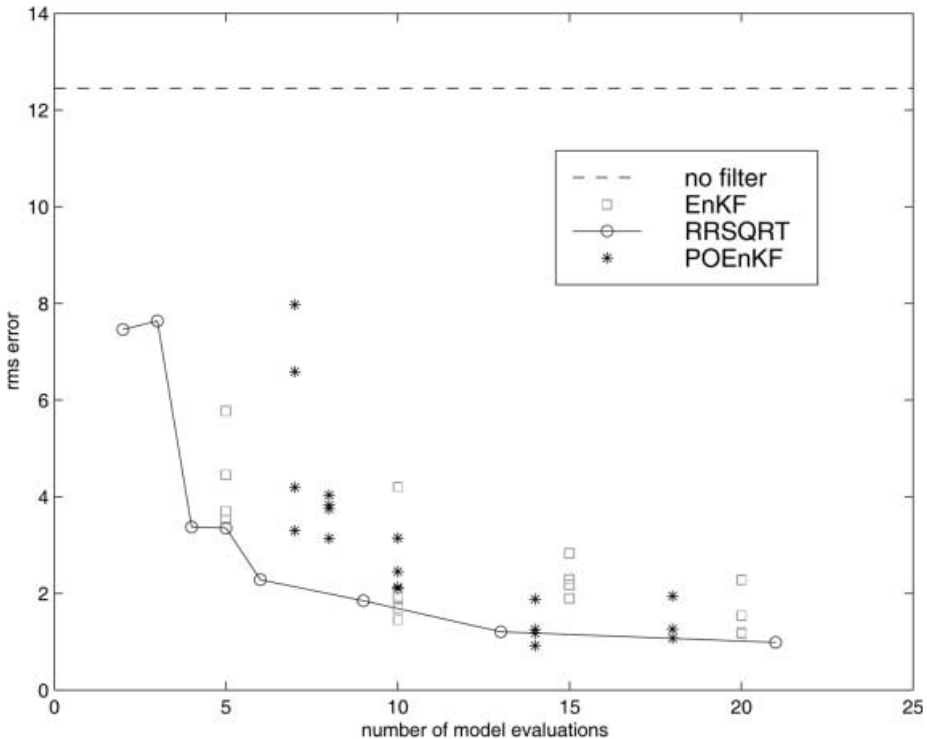


Fig. 7. Root mean square error for various filter setups versus the number of required model evaluation

By using Kalman filtering the high amount of information about the environmental process contained in these models can be integrated in the data analysis procedure. Because additional, physically based information is used, the final results of the analysis procedure is less sensitive for fluctuation in the data. A number of efficient filter algorithms for solving high dimensional filtering problems are described and applied to a real life ozone analysis and prediction problem. The results show that solving real life Kalman filtering problems is feasible. However computational burden is still large, at least an order of magnitude larger then required for the underlying deterministic transport model. The Kalman filter to improves the analysis and prediction results of the ozone concentration significantly by reconstructing the emissions. The main problem with the Kalman filtering approach is the specification of the model error. In our approach the emissions were chosen as the only error terms. There is, however, no consensus yet about the type model errors in air pollution models. This is still in the stage of expert opinions. We believe that in data model integration, getting quantitative information about the model error is one of the major challenges in the near future.

References

- Bogaert P, Christakos G** (1997) Stochastic analysis of spatiotemporal solute content measurements using a regression model. *Stochas. Hydrol. Hydraul.* 11(4): 267–295
- Builtjes P** (1992) The lotos – long term ozone simulation-project; summary report, TNO-report TNO-MW-R 92/240, TNO, Delft, The Netherlands
- Burgers G, Van Leeuwen PJ, Evensen G** (1998) Analysis scheme in the ensemble Kalman filter. *Geophys. Res. Lett.* 26: 1027–1030

- Canizares R** (1999) On the Application of Data Assimilation in Regional Coastal Models. PhD thesis, Delft University of Technology. A.A. Balkema, Rotterdam, 133 pp
- Christakos G, Vyas VM** (1998) A composite space/time approach to studying ozone distribution over Eastern United States. *Atmospheric Environ.* 32(16): 2845–2857
- Cohn SE, Todling R** (1996) Approximate data assimilation schemes for stable and unstable dynamics. *J. Meteorol. Soc. Japan* 74(1): 63–75
- Cohn SE, Todling R** (1995) Approximate Kalman filters for unstable dynamics. In: *Second Int. Symp. Assimilation of Observations in Meteorology and Oceanography*, Tokyo, pp. 241–246. WMO
- Cressie N** (1991) *Statistics for Spatial Data*. Wiley Inc., New York
- Evensen E** (1994) Sequential data assimilation with a nonlinear QG model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* 99: 10143–10162
- Evensen G, Van Leeuwen PJ** (1996) Assimilation of geosat altimeter data for the agulhas current using the Ensemble Kalman filter with a quasigeostrophic model. *Monthly Weather Rev.* 124(1): 85–96
- Gery M, Whitten G, Killus J, Dodge M** (1989) A photochemical kinetics mechanism for urban and regional scale computer modelling. *J. Geophys. Res.* 94(D10): 12,925–12,956
- Ghil M, Malanotte-Rizzoli P** (1991) Data assimilation in meteorology and oceanography. *Adv. Geophys.* 33: 141–266
- Hammersley J, Handscomb D** (1964) *Monte Carlo Methods*. Wiley and Sons Inc., London
- Hasselmann K** (1988) PIP's and POP's: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.* 93: 11,015–11,021
- Heemink AW, Verlaan M, Segers AJ** (2001) Variance reduced ensemble Kalman filtering. *Monthly Weather Rev.* 129: 1718–1728
- Houtekamer PL, Mitchel AL** (1998) Data assimilation using ensemble Kalman filter technique. *Monthly Weather Rev.* 126: 796–811
- Jazwinski AH** (1970) *Stochastic Processes and Filtering Theory*, Vol. 64 of *Mathematics in Science and Engineering*, Academic Press, New York
- Lermusiaux P** (1997) Error subspace data assimilation methods for ocean field estimation: theory, validation and applications. PhD thesis, Harvard University, Cambridge, MA
- Pham D, Verron J, Rouband M** (1998) A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Marine Sys.* 16: 323–340
- Segers AJ, Heemink AW, Verlaan M, Van Loon M** (2000), Non linear Kalman filter for atmospheric Chemistry Models, *Inverse Methods in global biogeochemical cyclus*. American Geophysical Union, pp. 139–146
- The UK National Qir Quality Information Archive** (1997) Department of the Environment, Transport and the Regions, <http://www.aeat.co.uk/netcen/airqual/>
- Tilmes S, Zimmermann J** (1998) Investigation on the spatial scales of the variability in measured near-ground ozone mixing ratios. *Geophys. Res. Lett.* 25(20): 3827–3830
- Verlaan M, Heemink AW** (1995) Data assimilation schemes for non-linear shallow water flow models. *Proc. Second Int. Symp. Assimilation of Observations*, Tokyo, Japan, pp. 247–252. WMO
- Verlaan M, Heemink AW** (1997) Tidal flow forecasting using reduced rank square root filters. *Stochastic Hydrol. Hydraul.* 11: 349–368
- Wikle CK, Cressie N** (1999), A dimension-reduced approach to space-time Kalman filtering. *Biometrika* 86: 815–829
- Wikle CK, Cressie N** (2000) Space-time statistical modeling of environmental data, In: *Todd Mowrer H, Congalton RG (eds), Quantifying Spatial Uncertainty in Natural Resources: Theory and Applications for GIS and Remote Sensing*. Ann Arbor Press, Chelsea Michigan, pp. 213–235
- Zhang X-F, Heemink A, Janssen L, Janssen P, Sauter F** (1999) A computationally efficient Kalman smoother for the evaluation of the ch4 budget in Europe. *Appl. Mathe. Model.* 23(2): 109–129