INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids **31**: 961–981 (1999)

COMPARISON OF EXTENDED AND ENSEMBLE KALMAN FILTERS FOR DATA ASSIMILATION IN COASTAL AREA MODELLING

H. MADSEN* AND R. CAÑIZARES

International Research Centre for Computational Hydrodynamics, Danish Hydraulic Institute, Agern Allé 5, DK-2970 Hørsholm, Denmark

SUMMARY

Data assimilation in a two-dimensional hydrodynamic model for bays, estuaries and coastal areas is considered. Two different methods based on the Kalman filter scheme are presented. These include (1) an extended Kalman filter in which the error covariance matrix is approximated by a matrix of reduced rank using a square root factorisation (RRSQRT KF), and (2) an ensemble Kalman filter (EnKF) based on a Monte Carlo simulation approach for propagation of errors. The filtering problem is formulated by utilising a general description of the model noise process related to errors in the model forcing, i.e. open boundary conditions and meteorological forcing. The performance of the two Kalman filters is evaluated using a twin experiment based on a hypothetical bay region. For both filters, the error covariance approximation sufficiently resolves the error propagation in the model at a computational load that is significantly smaller than required by the full Kalman filter algorithm. Furthermore, the Kalman filters are shown to be very robust with respect to defining the errors. Even in the case of a severely biased model error, the filters are able to efficiently correct the model. In general, the use of coloured model noise provides a numerically more efficient algorithm as well as a better performance of the filter. The error covariance approximation in the RRSQRT KF is shown to be more efficient than the error representation in the EnKF. For strongly non-linear dynamics, however, the EnKF is preferable.

KEY WORDS: Kalman filter; hydrodynamic modelling; coastal areas; data assimilation

1. INTRODUCTION

In general, the application of numerical models is hampered by the lack of knowledge of, or uncertainty related to, the physical parameters of the system, the initial and boundary conditions, and the external forcing. Furthermore, the model itself may be erroneous due to neglected or poorly described physical processes in the system equations and mathematical approximations (e.g. unresolved subgrid scale motions). On the other hand, measurements will always be sparse in both space and time, and hence will not be able to fully resolve the dynamics of the system at all spatial and temporal scales of interest. Data assimilation is a technique of combining any measurements of the state of the system with the model dynamics in order to improve the knowledge of the system.

* Correspondence to: Danish Hydraulic Institute, Agern Allé 5, DK-2970 Hørsholm, Denmark.

CCC 0271–2091/99/220961–21$17.50

Received July 1998
Revised January 1999

Existing data assimilation methods have primarily been developed for use in numerical weather prediction and open ocean models (see, e.g. review by Ghil and Malanotte-Rizzoli [1]). In this paper, data assimilation in a numerical hydrodynamic model for bays, estuaries and coastal areas is considered. Coupled numerical weather prediction and hydrodynamic models are now used operationally for storm surge forecasting [2]. In this respect, data assimilation techniques are adopted for real-time assimilation of water level measurements in order to improve the initial state of the system prior to a forecast.

At present, two different classes of data assimilation methods are prevalent: variational and sequential methods. Variational data assimilation consists of minimising a predefined cost function that measures the difference between model output and measurements over a certain time interval. A widely applied variational method is the adjoint method, which uses the model dynamics as strong constraints in the formulation of the cost function [3]. In coastal area modelling, the adjoint method has been applied for off-line estimation of model parameters [4]. A major drawback of the adjoint method, however, is the implicit assumption of a perfect model. To include model errors, a weak constraint formulation of the cost function should be adopted [5].

In sequential data assimilation, a recursive updating of the model solution is performed during a forward integration where model output and measurements are weighted according to the associated uncertainties. A classical sequential data assimilation method that is widely applied in operational numerical weather predictions is optimal interpolation [6,7]. This method has also been successfully applied in open ocean models [8]. However, since optimal interpolation is based on predefined model error statistics, which are assumed time invariant or are allowed to evolve according to a very simplified scheme, the updating procedure is not consistent with the dynamics of the system. In coastal area modelling, the dynamics are strongly influenced by the presence of land–sea boundaries and flooding and drying of tidal areas, and hence, the dynamical properties of the data assimilation scheme become extremely important.

The Kalman filter [9] is an efficient data assimilation method that explicitly accounts for the dynamic propagation of errors in the model. For linear models with known statistics of the system and measurement errors, the Kalman filter provides an optimal estimate of the state of the system, in terms of minimum estimation error covariance. In the case of non-linear model dynamics, an approximate Kalman filter algorithm (extended Kalman filter) in which the error propagation is based on a statistical linearisation of the model equation can be adopted. A major drawback of the Kalman filter, which makes it impracticable in high-dimensional systems, is the huge computational load and storage requirements associated with the propagation of the error covariance matrix. In recent years, several so-called suboptimal schemes [10] have been formulated which use different approximations of the error covariance modelling to reduce the computational burden.

Most suboptimal schemes are based on either a simplification of the model dynamics for propagation of the errors or approximations of the error covariance matrix. An example of the former approach can be found in Dee [11] who applied a simplified Kalman filter in an atmospheric flow model where the mass error propagation is obtained by simple advection, and the wind error propagation is subsequently evaluated by imposing geostrophic balance in the momentum equations. An alternative approach consists of propagating the error covariance matrix in a coarser grid than is used for propagation of the model itself [12]. Cohn and Todling [13] proposed a method where a reduced rank approximation of the singular value decomposition of the tangent linear operator is applied in the extended Kalman filter.

With respect to the approximation of the error covariance matrix, a significant reduction of the computational burden can be achieved by applying a steady state Kalman filter, where the Kalman gain can be calculated off-line. Examples of application of the steady state filter in coastal area modelling can be found in References [2,14,15]. The steady state filter prescribes time invariant error statistics and fixed measurement positions which may cause severe restrictions on the applicability of the method. For time varying Kalman filtering, an efficient scheme based on a reduced rank approximation of the error covariance matrix has been introduced. In this case, the error covariance matrix is approximated by a matrix of lower rank, including only the few, most significant eigenvectors of the matrix. Cohn and Todling [13] used this approach together with a Lanczos algorithm for the eigenvalue decomposition, whereas Verlaan and Heemink [16] used a square root factorisation. The reduced rank square root filter has been applied in coastal area modelling by Heemink *et al.* [15], Verlaan and Heemink [16], and Cañizares *et al.* [17,18].

The extended Kalman filter may provide poor results in the case of strongly non-linear dynamics. For instance, Evensen [19] found that the extended Kalman filter for a non-linear quasi-geostrophic ocean model resulted in an unbounded error covariance growth. Obviously, the extended Kalman filter can be improved by including higher-order moments in the error covariance approximation. However, in practice, for large systems this approach is not feasible since the computational load and storage requirements become overwhelming. To resolve the non-linearities in the error propagation, Evensen [20] introduced an ensemble Kalman filter approach based on Monte Carlo simulations. In this method, the error covariance matrix is represented by an ensemble of possible states that are propagated according to the full non-linear dynamics of the system. Evensen and van Leeuwen [21] applied the ensemble Kalman filter for assimilation of altimeter data in a quasi-geostrophic ocean model.

The objective of the present paper is to test and compare the reduced rank square root algorithm based on an extended Kalman filter formulation (RRSQRT KF) and the ensemble Kalman filter (EnKF) for assimilation of water level measurements in an existing two-dimensional hydrodynamic model (MIKE 21). In Section 2, the applied numerical model is briefly described. Based on a stochastic representation of the model and the measurement equations, the RRSQRT KF and EnKF algorithms are formulated in Section 3. In Section 4, results from an application of the two methods in a twin experiment based on a hypothetical bay region are presented. Finally, conclusions are given in Section 5.

2. HYDRODYNAMIC MODEL

The MIKE 21 hydrodynamic model is a general numerical modelling system for the simulation of unsteady two-dimensional flow, developed at the Danish Hydraulic Institute [22]. The basic partial differential equations (PDE) are the depth integrated continuity and momentum equations (shallow water equations)

$$\frac{\partial \xi}{\partial t}+\frac{\partial p}{\partial x}+\frac{\partial q}{\partial y}=0, \tag{1}$$

$$\frac{\partial p}{\partial t}+\frac{\partial}{\partial x}\left(\frac{p^2}{h}\right)+\frac{\partial}{\partial y}\left(\frac{pq}{h}\right)+gh\frac{\partial \xi}{\partial x}+\frac{gp\sqrt{p^2+q^2}}{C^2h^2}-\Omega q-fVV_x+h\frac{\partial}{\partial x}\left(\frac{p_a}{\rho_w}\right)=0, \tag{2}$$

$$\frac{\partial q}{\partial t}+\frac{\partial}{\partial y}\left(\frac{q^2}{h}\right)+\frac{\partial}{\partial x}\left(\frac{pq}{h}\right)+gh\frac{\partial \xi}{\partial y}+\frac{gp\sqrt{p^2+q^2}}{C^2h^2}-\Omega p-fVV_y+h\frac{\partial}{\partial y}\left(\frac{p_a}{\rho_w}\right)=0, \tag{3}$$

where x, y are the horizontal co-ordinates (m); t is the time (s); h is the water depth (m); ξ is the surface elevation (m); p, q are the flux densities in the x- and y-directions ($m^3\,s^{-1}\,m^{-1}$), $(p, q) = (v_x h, v_y h)$, where v_x and v_y are the depth averaged x and y velocities ($m\,s^{-1}$); g is the acceleration due to gravity ($m\,s^{-2}$); C is the Chezy bed resistance coefficient ($m^{1/2}\,s^{-1}$); Ω is the Coriolis parameter (s^{-1}); f is the wind friction factor; V, V_x, V_y are the wind speed and wind speed components in the x- and y-directions ($m\,s^{-1}$); p_a is the atmospheric pressure ($kg\,m^{-1}\,s^{-2}$); and ρ_w is the density of water ($kg\,m^{-3}$).

At closed boundaries the flow perpendicular to the boundary is set to zero. At open boundaries the surface elevation is prescribed. Owing to the non-linearity of the dynamics, a secondary boundary condition has to be imposed at open boundaries in order to close the solution. In MIKE 21 fluxes along the open boundary has to be specified, and in this respect three different options are available: (1) the flux is set to zero, (2) the flux is obtained by extrapolation in space, and (3) the flow direction is specified, implying that the flux can be computed internally by MIKE 21. With the defined boundary conditions and with prescribed initial values of surface elevations and flux densities, (1)–(3) form a well-posed boundary value problem.

MIKE 21 uses a finite difference approximation to solve the PDEs where the difference terms are expressed on a staggered grid. A time-centred alternating direction implicit (ADI) scheme is adopted. The equations are solved in one-dimensional sweeps, alternating between the x- and y-directions. In the x-sweep, the continuity equation and the momentum equation in the x-direction are solved with respect to ξ at time step $k+\frac{1}{2}$ and p at time step $k+1$ using the known variables ξ_k, p_k, $q_{k-1/2}$ and $q_{k+1/2}$. In the y-sweep, the continuity equation and the momentum equation in the y-direction are solved with respect to ξ_{k+1} and $q_{k+3/2}$ using $\xi_{k+1/2}$, $q_{k+1/2}$, p_k and p_{k+1}.

3. KALMAN FILTER

3.1. Kalman filter update

For the implementation of the Kalman filter in MIKE 21, the numerical model has to be formulated in a state–space form. The state variables to be considered are surface elevations and the depth averaged x and y velocities in every point of the horizontal grid. The Kalman filter algorithm is based on a recursive two-time step formulation. The numerical scheme in MIKE 21, however, involves the y velocity at three time steps. To express this scheme using only two time steps, the y velocity at time steps $k+\frac{1}{2}$ and $k-\frac{1}{2}$ are included in the state vector. The numerical scheme based on (1)–(3) can then be written as

$$x_k = \Phi(x_{k-1}, u_k), \tag{4}$$

where $x_k = (\xi_k, v_{x,k}, v_{y,k+1/2}, v_{y,k-1/2})$ is the state vector, and u_k is the forcing of the system in terms of the surface elevations at open boundaries, and the meteorological forcing components in the momentum equations (wind stress and pressure gradient).

For modelling the uncertainty of the system, it is assumed that model errors are mainly related to errors in the forcing terms. At open boundaries, the tidal component of the surface elevation can usually be obtained with a relatively high accuracy, whereas the variations due to the meteorological effects may contain large errors (e.g. due to generation of a surge outside the model domain). The errors in the meteorological forcing terms are partly caused by uncertainties of the meteorological observations, and partly related to the physical description of the wind stress component and determination of the wind friction factor.

The error processes are assumed to be less spatially variable than the water flow process [14], and the discrete error processes can thus be defined on a grid G2, which is coarser than the model grid G1. A stochastic representation of the system equation (4) can then be written

$$x_k = \Phi(x_{k-1}, u_k + \Lambda\varepsilon_k), \tag{5}$$

where ε_k contains the model error in every grid point of G2, and Λ is a matrix that represents the sequence of linear interpolations between G2 and G1. The model error process is assumed unbiased, and the error statistics (variance structure as well as spatial and temporal correlation structure) are assumed known.

Measurements z_k of the state of the system are assumed to be available at certain points in the model grid G1. Measurement errors are partly related to the measurement equipment, and partly related to the uncertainty caused by the use of point measurements to represent grid averages. The stochastic representation of the measurement equation reads

$$z_k = C_k x_k + \eta_k, \tag{6}$$

where C_k is a matrix that describes the relation between measurements and state variables, and η_k is a random measurement error with zero mean and known covariance matrix R_k.

Now, a one-step ahead forecast of the state of the system is denoted by x_k^{f}, according to the model operator $\Phi(\bullet)$, cf. Equation (4). The uncertainty of this forecast is described by the error covariance matrix P_k^{f}. If measurements are available, cf. Equation (6), the model forecast and the measurements can be combined to obtain an updated estimate of the state of the system. The Kalman filter update of the state vector and the error covariance matrix are given by

$$x_k^{\mathrm{a}} = x_k^{\mathrm{f}} + K_k(z_k - C_k x_k^{\mathrm{f}}), \tag{7}$$

$$P_k^{\mathrm{a}} = P_k^{\mathrm{f}} - K_k C_k P_k^{\mathrm{f}}, \tag{8}$$

where K_k is the Kalman gain matrix,

$$K_k = P_k^{\mathrm{f}} C_k^T [C_k P_k^{\mathrm{f}} C_k^T + R_k]^{-1}, \tag{9}$$

which serves as a weighting function of the model forecast and measurements and depends on the associated errors P_k^{f} and R_k. In Equations (7)–(9), superscripts f and a refer to forecast and analysis (or update) respectively.

The forecast error covariance matrix P_k^{f} is a combination of time propagation of P_{k-1}^{a} and forcing of the system by model errors. For large systems, the propagation of the errors is the main bottleneck, imposing an unacceptable computational burden. Let n denote the dimension of the state vector (in the order 10^3–10^5 in models of realistic complexity), the propagation of the error covariance matrix requires $2n$ as much computing effort as is required to advance the deterministic model. The Kalman filter algorithms described below represent two different approaches for approximating the error covariance propagation that significantly reduces the computational burden.

3.2. Reduced rank square root filter

In the extended Kalman filter, the propagation of the error covariance matrix is based on a statistical linearisation of the model dynamics. In the case of a white system noise process, the forecast step is given by

$$x_k^{\mathrm{f}} = \Phi(x_{k-1}^{\mathrm{a}}, u_k), \tag{10}$$

$$P_k^{\mathrm{f}} = F_k P_{k-1}^{\mathrm{a}} F_k^T + G_k(\Lambda Q_k \Lambda^T) G_k^T, \quad F_k = \left.\frac{\partial \Phi}{\partial x}\right|_{x = x_k^{\mathrm{f}}}, \quad G_k = \left.\frac{\partial \Phi}{\partial u}\right|_{x = x_k^{\mathrm{f}}}, \tag{11}$$

where Q_k is the covariance matrix of the system noise, defined on grid G2.

The RRSQRT approximation of the extended Kalman filter uses a square root algorithm as well as a lower rank approximation of the error covariance matrix. Denote by S_{k-1}^{a} the approximation of rank M of the square root of the error covariance matrix P_{k-1}^{a}. The propagation of the error covariance matrix is then given by

$$S_k^{\mathrm{f}} = [F_k S_{k-1}^{\mathrm{a}} | G_k \Lambda Q_k^{1/2}], \tag{12}$$

where $Q_k^{1/2}$ is the square root of Q_k. The matrix S_{k-1}^{a} has M columns, where M is chosen much smaller than the dimension n of the state vector. To calculate the derivatives in F_k and G_k, a finite difference approximation of $\Phi(\bullet)$ is adopted. Thus, the propagation of the error covariance matrix requires M plus q (the total number of noise points) model integrations, which is much smaller than the $2n$ integrations required in (11).

The propagation step in (12) increases the number of columns in the error covariance matrix from M to $M+q$. To reduce the number of columns, and hence keep the rank of the matrix constant throughout the simulation, a lower rank approximation of S_k^{f} in (12) is applied by keeping only the M leading eigenvectors of the error covariance matrix. The reduction can be achieved either by a singular value decomposition of S_k^{f} or by an eigenvalue decomposition of the matrix $(S_k^{\mathrm{f}})^T S_k^{\mathrm{f}}$ [16,18].

Based on the square root approximation of rank M, S_k^f, the error covariance matrix can be calculated as $P_k^{\mathrm{f}} = S_k^{\mathrm{f}}(S_k^{\mathrm{f}})^T$, and subsequently used for the Kalman filter update, cf. Equations (7)–(9). The expensive matrix multiplication, however, can be reduced significantly when the data assimilation is based on *in situ* measurements that are sparsely represented in space. In this case, the measurement matrix C_k has only a few non-zero elements, and only the columns in P_k^{f} that corresponds to the non-zero elements in C_k have to be calculated. Furthermore, if measurement errors are uncorrelated, a sequential updating algorithm that process one measurement at a time can be implemented [23], and hence avoiding the matrix inversion in (9). In this case, it is not necessary to calculate the forecast error covariance matrix, and the sequential updating can be performed using S_k^{f} directly (see details in Cañizares *et al.* [18]).

In the case of coloured (time correlated) system noise, the Kalman filter is defined by using an augmented state vector that includes the system noise components. Assuming that the noise process at time step k, ε_k, only depends on the previous time step $k-1$, the augmented stochastic model equation can be written as

$$\begin{pmatrix} x_k \\ \varepsilon_k \end{pmatrix} = \begin{pmatrix} \Phi(x_{k-1}, u_k + \Lambda\varepsilon_k) \\ A\varepsilon_{k-1} + \delta_k \end{pmatrix}, \tag{13}$$

where A represents the correlation model, and δ_k is a white noise process with covariance matrix Q_k. The propagation of the error covariance matrix corresponding to the augmented state vector is now given by

$$S_k^{\mathrm{f}*} = \begin{pmatrix} S_k^{\mathrm{f}} \\ \Sigma_k^{\mathrm{f}} \end{pmatrix} = \left[\begin{pmatrix} F_k & G_k\Lambda \\ 0 & A \end{pmatrix} \begin{pmatrix} S_{k-1}^{\mathrm{a}} \\ \Sigma_{k-1}^{\mathrm{a}} \end{pmatrix} \middle| \begin{pmatrix} 0 \\ Q^{1/2} \end{pmatrix} \right]. \tag{14}$$

To solve (14), a finite difference approximation of $\Phi(\bullet)$ is adopted that requires M model integrations [24]. Since the augmented state vector contains variables with different scales of magnitude, $S_k^{\mathrm{f}*}$ is normalised prior to the eigenvalue decomposition [24].

3.3. Ensemble Kalman filter

In the ensemble Kalman filter [20], the statistical properties of the state vector are represented by an ensemble of possible state vectors. Each of these vectors is propagated according to the dynamic system subjected to model errors, and the resulting ensemble then provides estimates of the forecast state vector and the error covariance matrix. In the measurement update, the Kalman gain matrix obtained from (9) is applied for each of the forecast state vectors. To account for measurement errors, the measurements are represented by an ensemble of possible measurements [25]. The resulting updated sample provides estimates of the updated state vector and the error covariance matrix. The ensemble Kalman filter can thus be summarised as follows:

1. Each member of the ensemble of M state vectors is propagated forward in time according to the dynamics of the system and the specified model error, i.e.

$$x^{\mathrm{f}}_{i,k} = \Phi(x^{\mathrm{a}}_{i,k-1}, u_k + \Lambda\varepsilon_{i,k}), \quad i = 1, 2, \ldots, M, \tag{15}$$

   where the model error $\varepsilon_{i,k}$ is randomly drawn from a predefined distribution with zero mean and covariance matrix Q_k.
2. In general, the forecast of the state vector can be calculated as a certain quartile of the ensemble forecast. Usually, the mean value is adopted, i.e.

$$\hat{x}^{\mathrm{f}}_k = \bar{x}^{\mathrm{f}}_k = \frac{1}{M}\sum_{i=1}^{M} x^{\mathrm{f}}_{i,k}. \tag{16}$$

   The error covariance matrix of the forecast is then estimated from the ensemble as

$$P^{\mathrm{f}}_k = S^{\mathrm{f}}_k (S^{\mathrm{f}}_k)^T, \qquad S^{\mathrm{f}}_{i,k} = \frac{1}{\sqrt{M-1}}(x^{\mathrm{f}}_{i,k} - \bar{x}^{\mathrm{f}}_k), \tag{17}$$

   where $S^{\mathrm{f}}_{i,k}$ is the ith column in S^{f}_k.
3. An ensemble of size M of possible measurements is generated

$$z_{i,k} = z_k + \eta_{i,k}, \quad i = 1, 2, \ldots, M, \tag{18}$$

   where z_k is the actual measurement vector and $\eta_{i,k}$ is the measurement error that is randomly generated from a predefined distribution with zero mean and covariance matrix R_k.
4. Each ensemble member is updated according to the updating scheme in (7), and based on the updated ensemble, the updated state vector and error covariance matrix are estimated, cf. Equations (16) and (17).

For uncorrelated measurement errors, a sequential updating scheme similar to the one described above can be applied using S^{f}_k directly [26], and hence avoiding the expensive calculation of P^{f}_k in (17) and the matrix inversion in (9).

For time coloured system noise the EnKF can be defined using an augmented state vector formulation. In this case, an ensemble of model errors are propagated according to the correlation model

$$\varepsilon^{\mathrm{f}}_{i,k} = A\varepsilon^{\mathrm{a}}_{i,k-1} + \delta_{i,k}, \quad i = 1, 2, \ldots, M, \tag{19}$$

where $\delta_{i,k}$ is randomly drawn from a predefined distribution with zero mean and covariance matrix Q_k.

The main features of the RRSQRT KF and EnKF algorithms are shown in Table I for comparison.

4. EXPERIMENTS

4.1. Test layout

To evaluate the performance of the RRSQRT KF and the EnKF data assimilation schemes, a twin experiment was carried out. A hypothetical bay with a grid of 21×20 points, and a grid size of 10×10 km^2 was used. The model bathymetry is shown in Figure 1. The Chezy bed resistance coefficient varies with the depth, representing values in the range 30–45 $m^{1/2}\,s^{-1}$ (largest values in the deepest areas). The flow is forced by a sinusoidal variation of the surface elevation at the open northern boundary with a period of 12 h and an amplitude range of 2 m. Meteorological forcing is included using wind and pressure fields from an artificially generated moving cyclone that moves in a west–east direction with a speed of 50 km per 6 h (see Figure 2). The main flow describes a Kelvin wave moving anti-clockwise in the bay region.

It is assumed that model errors are correlated in time according to a first-order auto-regressive process, i.e.

$$\varepsilon_k = \alpha \varepsilon_{k-1} + \delta_k, \qquad \varepsilon_0 = 0, \tag{20}$$

where α is the lag-one auto-correlation coefficient, and δ_k is a Gaussian white noise component with zero mean and prescribed covariance structure. It is assumed here that the model error covariance is constant in time and that the spatial correlation structure can be described by an exponential model, i.e. the covariance matrix Q has the elements

$$Q_{ij} = \rho^{d_{ij}} \sigma_{si} \sigma_{sj}, \tag{21}$$

where σ_{si}^2 is the variance of the ith element in δ_k, ρ is the correlation coefficient between two elements of δ_k separated by a unit distance, and d_{ij} is the distance between nodes i and j. The measurement errors are assumed to be independent of the model errors, mutually uncorrelated and homogeneous in time, i.e. the covariance matrix reads

$$R = \mathrm{diag}[\delta_{m1}^2, \ldots, \sigma_{mp}^2], \tag{22}$$

Table I. Summary of the EnKF and RRSQRT KF algorithms

	EnKF	RRSQRT KF
Model forecast	Ensemble estimate	Deterministic model forecast
Error propagation	Propagation of ensemble according to full non-linear model dynamics	Propagation of error covariance matrix using tangent linear model operator
Model error forcing	Part of ensemble propagation	Matrix algebra
Representation of error covariance matrix	Ensemble estimate	Reduced rank approximation of square root of covariance matrix
Storage requirements	$M \times (n+q)$	$(M+q) \times (n+q)$
Computational costs	M model integrations	M model integrations Eigenvalue decomposition

M is the ensemble size in the EnKF and the number of leading eigenvalues in the RRSQRT KF; n is the number of state variables; q is the number of noise points.

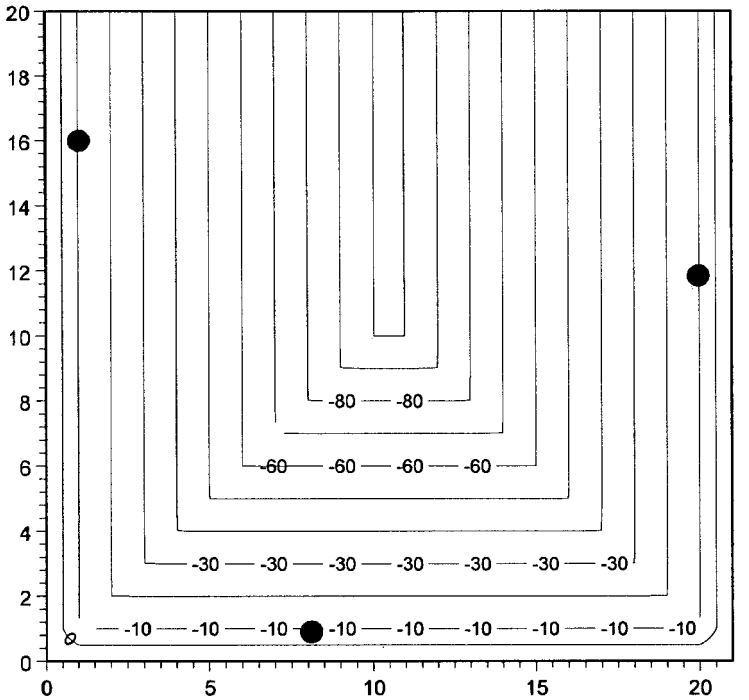

Figure 1. Bathymetry of hypothetical bay (depth in m). Dots indicate measurement locations.

where σ_{mi}^2, $i = 1, 2, \ldots, p$ is the variance of the error of the ith measurement.

The reference state (true state) of the system is simulated using the deterministic model with the sinusoidal boundary and meteorological forcing described above. A simulation time of 48 h with a time step of 15 min is used. Water level measurements are extracted from the true state simulation and disturbed by adding random errors generated from a Gaussian distribution with covariance matrix R_k, cf. Equation (22). Three measurement locations are considered, grid points (1, 16), (8, 1) and (20, 12) respectively (Figure 1). A wrong state of the system is simulated by adding errors to the forcing terms. The Kalman filter update is obtained by assimilating water level measurements from the three locations into the wrong model forecast in every time step.

The performance of the two Kalman filter updating schemes is evaluated by comparing the root mean square error (r.m.s.e.) between the true and updated surface elevations fields. To compare the error covariance estimates of the two filters, the spatial distribution of the average standard deviation (S.D.) of the surface elevations is calculated. To minimise the influence from the initial conditions, the r.m.s.e. and S.D.s are calculated using results only from the last 24 h of simulation. Global performance measures are calculated as spatial averages of the r.m.s.e. and the S.D.

4.2. Errors in open boundaries

First, errors in the open boundary are considered. Errors in the surface elevations at the open boundary are randomly generated from a first-order auto-regressive process, cf. Equation (20), with a lag-one auto-correlation coefficient $\alpha = 0.9$. The residuals are generated from a Gaussian distribution with covariance matrix given by (21) with a constant S.D. $\sigma_s = 0.1$ m,

and spatial correlation coefficient $\rho = 0.9$. The r.m.s.e. of the wrong model is shown in Figure 3. The random measurement errors are generated using $\sigma_{mi} = 0.05$ m at all locations.

In the first test, the two Kalman filters were applied using the true error statistics (i.e. the statistics used to generate the random errors in the model forcing). When the filters are provided with the true error statistics, the error covariance estimate should be close to the r.m.s.e., any differences being caused by the applied approximations in the Kalman filter schemes. The performance of the RRSQRT KF generally depends on the rank of the square root approximation of the error covariance matrix (number of leading eigenvalues used to define the matrix). Similarly, the performance of the EnKF depends on the ensemble size. The number of leading eigenvalues and the ensemble size (both denoted M in the following) are important parameters to control since they determine the computational load and storage requirements.

In Figure 4 the average r.m.s.e. and S.D. are shown for different numbers of leading eigenvalues in the RRSQRT KF and ensemble sizes in the EnKF. For small M, the updating schemes are not very efficient (large r.m.s.e.), and, in addition, the S.D. is underestimated. For both filters, the r.m.s.e. decreases and the S.D. increases for increasing M. For M smaller than about 100, the r.m.s.e. of the RRSQRT KF is smaller than that of the EnKF, whereas for larger M the r.m.s.e. of the two filters are virtually identical. The S.D. is slightly larger for the RRSQRT KF.

The main difference between the two filters is related to the convergence conditions. For the RRSQRT KF the r.m.s.e. stabilises around $M = 40$, whereas the EnKF requires an ensemble size of about $M = 100$ to obtain convergence. Thus, the reduced rank approximation of the error covariance matrix in the RRSQRT KF is more efficient than the ensemble representation

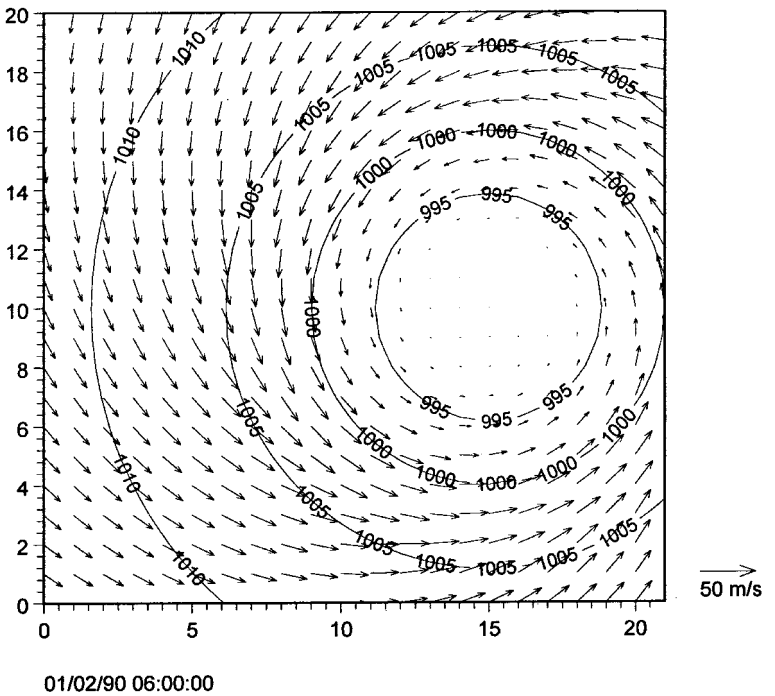

Figure 2. Generated cyclonic wind (m s^{-1}) and pressure fields (hPa) after 30 h.

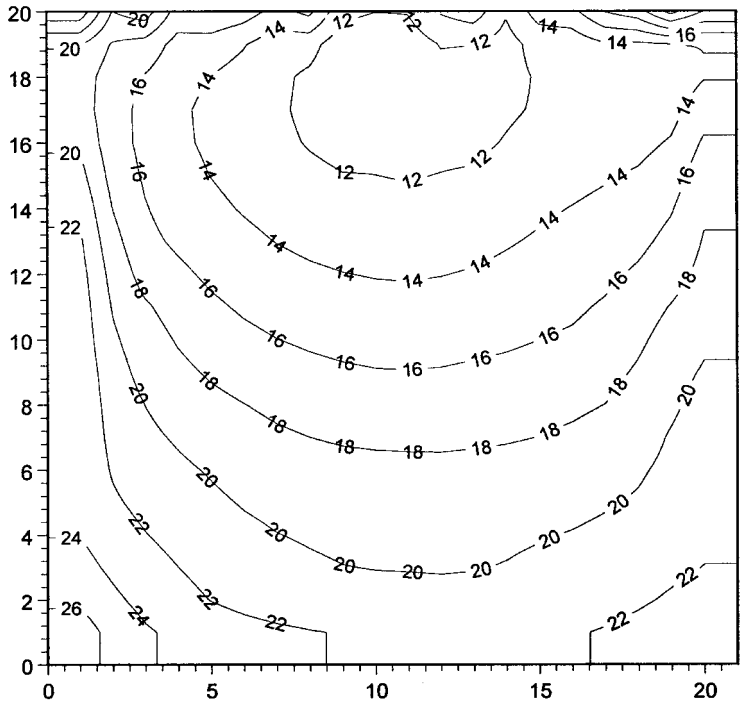

Figure 3. Spatial distribution of r.m.s.e. (cm) of the wrong model.

in the EnKF. This feature is due to the fact that the RRSQRT KF considers only the most significant eigenvectors of the error covariance matrix, i.e. noise in the error representation is filtered out in every time step. The performance of the EnKF, especially for small ensemble

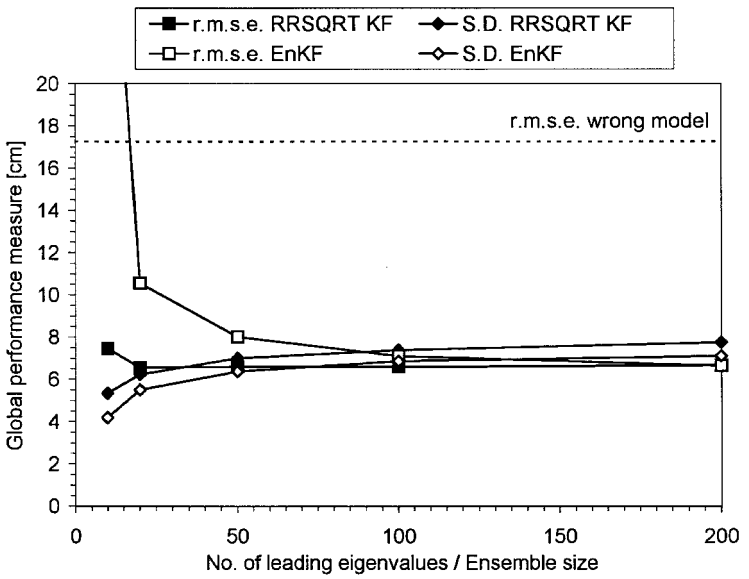

Figure 4. Spatial average r.m.s.e. and S.D. as a function of the number of leading eigenvalues in the RRSQRT KF and the ensemble size in the EnKF.

sizes, can be improved by limiting the sampling space. In this case, the random sampling of the residual model error δ_k is performed within an hyperellipsoid that encloses the value of δ_k with probability p, i.e.

$$\delta_{i,k}^T Q_k^{-1} \delta_{i,k} \leq \chi_q^2(p), \tag{23}$$

where $\chi_q^2(p)$ is a χ^2 variable with q degrees of freedom (dimension of the noise vector) and probability p. By reducing the random sampling space, the noise in the error representation, which is especially pronounced for small ensemble sizes, can be reduced. However, reduction of the sampling space also introduces an underestimation of the true error covariance matrix, and hence, one should not apply too severe limitations.

The computational load of the two filters for different values of M is shown in Figure 5. For the EnKF, the computational load corresponds roughly to the load of M model integrations. For the RRSQRT KF the computational load is highly dependent on the eigenvalue decomposition, which becomes relatively more expensive for increasing M. In the present case, the computational load for the RRSQRT KF with $M = 50$ is virtually identical to the load for the EnKF with an ensemble size of $M = 100$. Thus, in this case the same performance is achieved for almost the same computational costs.

In Figure 6, the r.m.s.e. and S.D. are shown for the RRSQRT KF with $M = 40$ and for the EnKF with $M = 100$. Assimilation of surface elevations in the three measurement positions reduces significantly the RMSE in most parts of the area, except for points close to the open boundary (compare with the r.m.s.e. of the wrong model in Figure 3). The spatial distribution of the r.m.s.e. and the S.D. are very similar for the two filters. Moreover, both filters provide estimates of the S.D. that are very close to the observed r.m.s.e., indicating that the error propagation in the filter is sufficiently resolved.

The general performance and convergence of the filters depend on the definition of the coloured noise. Different tests were performed using the same layout as above but with

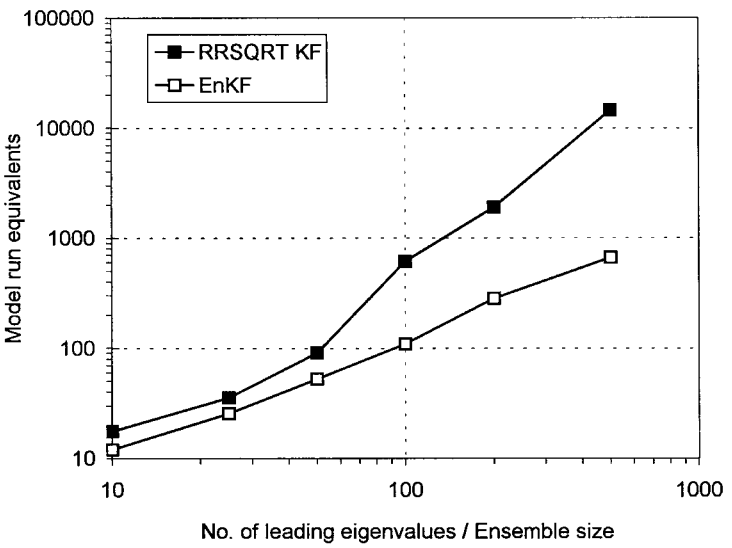

Figure 5. Computational load of the RRSQRT KF and the EnKF, in terms of model run equivalents, as a function of the number of leading eigenvalues and ensemble size respectively.

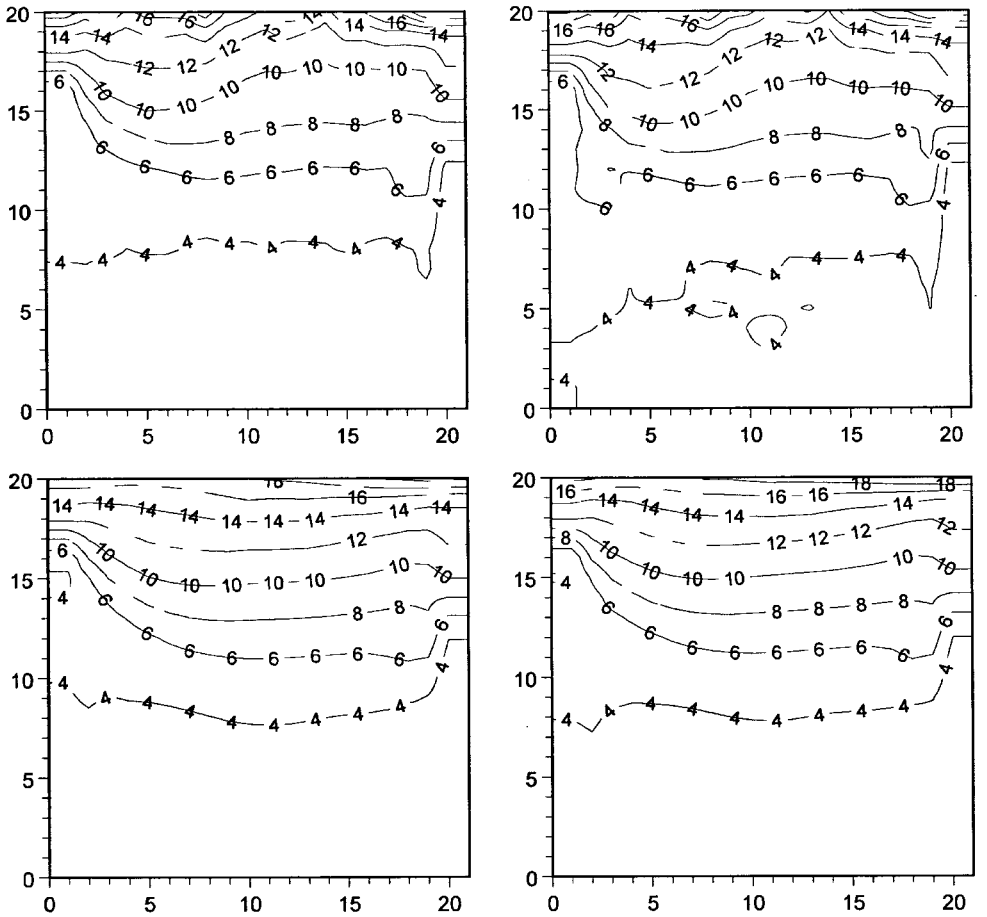

Figure 6. Left: r.m.s.e. (cm) (top) and S.D. (cm) (bottom) for the RRSQRT KF. Right: r.m.s.e. (top) and S.D. (bottom) for the EnKF.

different values of the lag-one auto-correlation coefficient. For smaller correlations, the r.m.s.e. converges for larger values of M. In addition, the filter becomes less efficient in the sense that the reduction of the global r.m.s.e. is smaller. Thus, the use of coloured noise implies a numerically more efficient algorithm as well as a better performance of the filter.

In practice, the model error structure and error statistics are unknown, or only partly known, and hence an important property of the Kalman filter is its robustness with respect to specification of the model error. A number of sensitivity tests have been carried out, applying the filters on the wrong model as defined above but using different error statistics (i.e. varying α, σ_s, ρ, σ_{mi}). Generally, when the error statistics are not correctly specified, the performance of the filters deteriorates. However, the filters are in most cases very robust, implying a reduction of the r.m.s.e. in the same order as the reduction obtained by the filters based on the true statistics. The estimated S.D. is more sensitive to the definition of the error statistics, implying a severe under or overestimation when the error statistics are grossly misspecified. With respect to definition of the error statistics the stability of the filter is another important issue. When the uncertainty of the measurements is very small compared with the model

uncertainty, the RRSQRT KF becomes unstable. The EnKF does not have any stability problems in this case, although it produced poor results.

All the tests considered above have been based on unbiased model errors in accordance with the basic assumption of the Kalman filter. To evaluate the performance of the filters in the case of biased model errors, different tests were carried out where the introduced error at the open boundary is a combination of a random error and a systematic error, i.e. an error in the phase or in the amplitude. As an example, the r.m.s.e. of the wrong model, including a phase error of 1 h is shown in Figure 7. Generated model errors are very large in this case with an average r.m.s.e. of 90.7 cm. The two filters were applied using the parameters $\alpha = 0.9$, $\rho = 0.9$, $\sigma_s = 0.1$ m, $\sigma_{mi} = 0.05$ m, $M = 40$ for the RRSQRT KF, and $M = 100$ for the EnKF. Both filters provide very large reductions of the r.m.s.e., reducing the average to respectively 9.3 cm for the RRSQRT KF and 9.7 cm for the EnKF. The performance of the filters in this case is highly dependent on the definition of coloured noise. The spatial distribution of the r.m.s.e. of the EnKF using coloured ($\alpha = 0.9$) and white noise respectively is shown in Figure 8. In the case of white noise, larger r.m.s.e. are observed, with an average equal to 22.6 cm. When coloured noise is applied, the Kalman filter provides estimates of the model error (augmented part of the updated state vector, cf. Equation (13)), and hence the surface elevation at the open boundary is updated along with the state variables. The tests have shown that the Kalman filter with coloured noise is very efficient in tracking phase and amplitude errors at the open boundary.

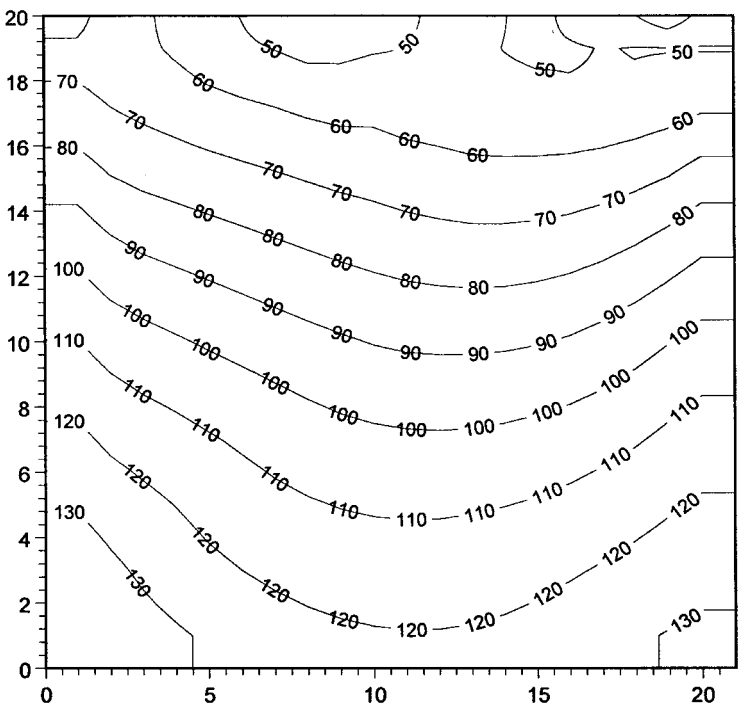

Figure 7. Spatial distribution of r.m.s.e. (cm) of the wrong model, including a phase error of 1 h at the open boundary.

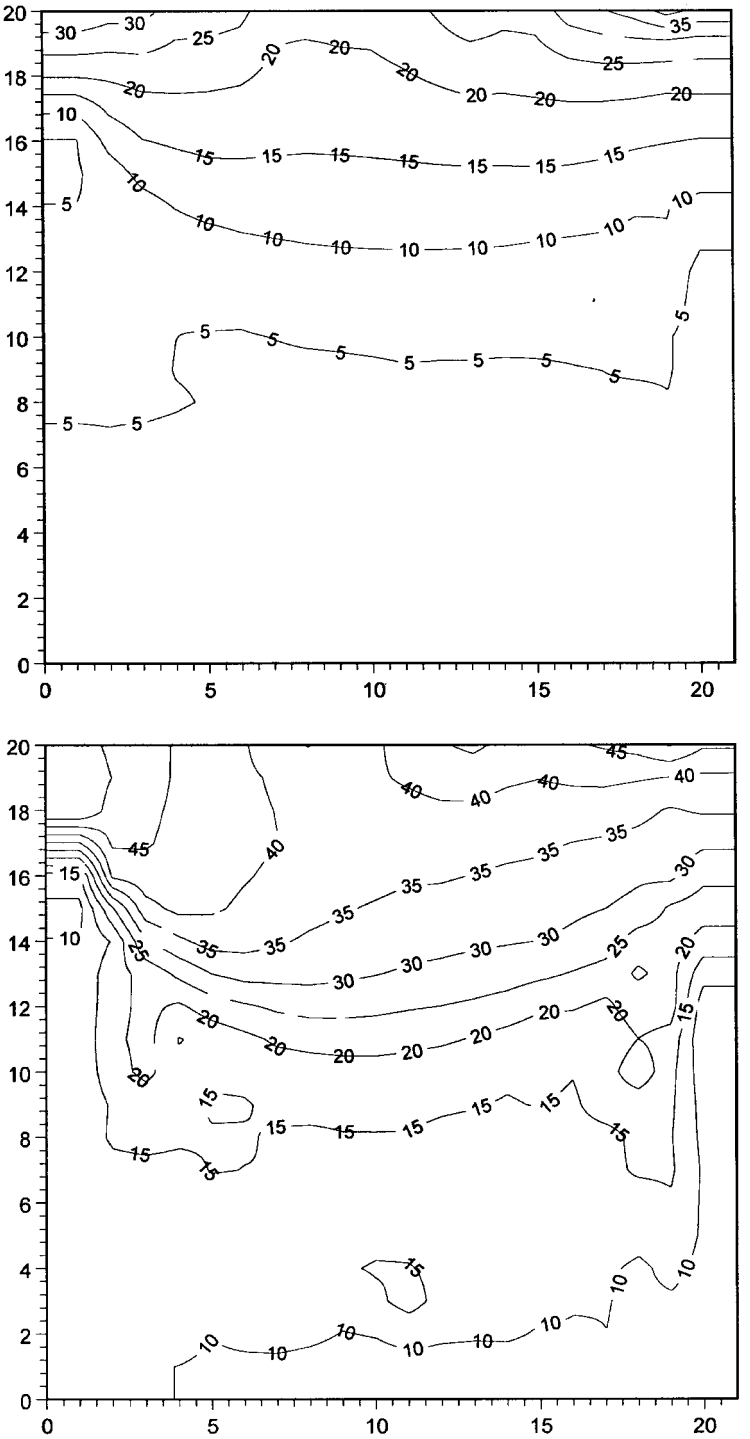

Figure 8. Spatial distribution of r.m.s.e. (cm) of the EnKF using coloured (top) and white noise (bottom) respectively.

4.3. Errors in meteorological forcing

The performance of the filters in the case of errors in the meteorological forcing terms has been tested using the same true model and measurements as defined above. A wrong model was simulated by adding randomly generated errors to the meteorological forcing terms using the following noise statistics: $\alpha = 0.97$, $\rho = 0.98$ and $\sigma_s = 0.00025$ m^2 s^{-2} (correlation parameters are taken from [14]). The r.m.s.e. of the wrong model in this case is shown in Figure 9. Compared with the wrong model based on disturbed boundary conditions, the error growth is more severe in the SW and SE parts of the region. Since the open boundary is assumed perfect, only very small errors are present close to the boundary.

Different simulations were performed to test the convergence and sensitivity of the two filters. Compared with the case with errors in the open boundaries, generally a larger number of leading eigenvalues in the RRSQRT KF and a larger ensemble size in the EnKF are necessary to obtain convergence. When the filters are provided with the noise statistics used to generate the wrong model, the RRSQRT KF converges for M about 70, whereas the EnKF requires an ensemble size of $M = 200$ to obtain convergence. In this case, the two algorithms have virtually the same computational load. In Figure 10 the r.m.s.e. and S.D. are shown for the RRSQRT KF with $M = 70$ and for the EnKF with $M = 200$. Large reductions in the r.m.s.e. are achieved; the average is reduced from 17.1 cm of the wrong model to 5.0 cm for both filters. Also, in this case the spatial distribution of the r.m.s.e. and the S.D. are very similar for the two filters, and both filters provide estimates of the S.D. that are very close to the observed r.m.s.e.

The performance of the filters is strongly dependent on the definition of the spatial correlation structure. For smaller spatial correlation coefficients, the filters converge for larger

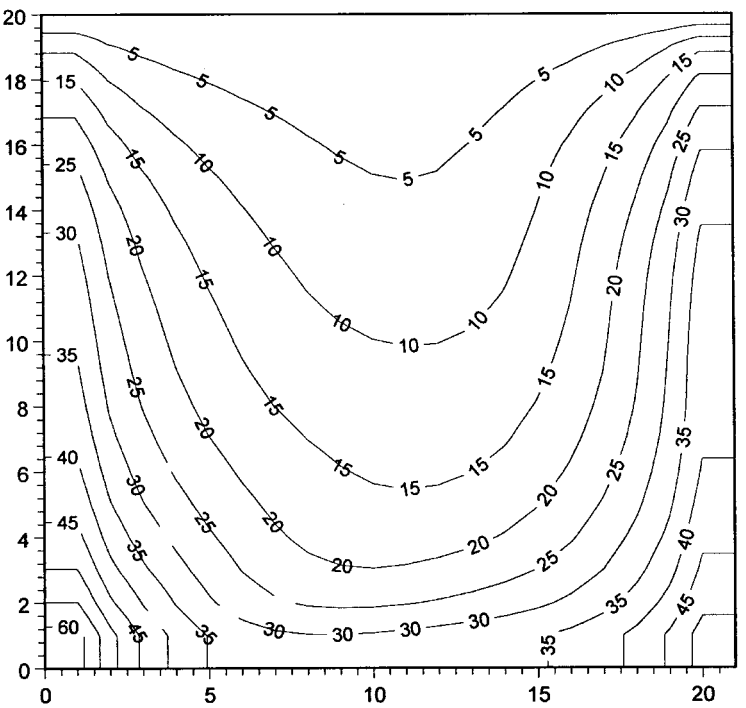

Figure 9. Spatial distribution of r.m.s.e. (cm) of the wrong model.

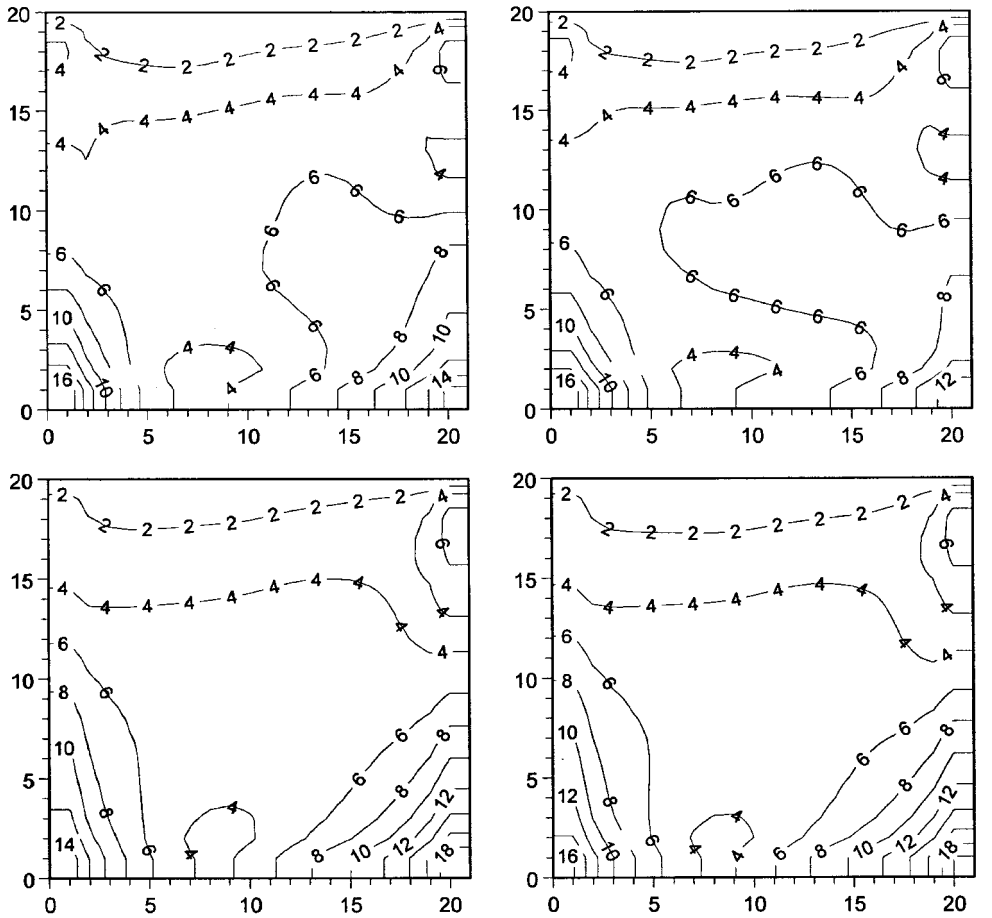

Figure 10. Left: r.m.s.e. (cm) (top) and S.D. (cm) (bottom) for the RRSQRT KF. Right: r.m.s.e. (top) and S.D. (bottom) for the EnKF.

values of M. Moreover, the reduction of the r.m.s.e. is less pronounced, and for the RRSQRT KF in some cases even an increase is observed. Thus, in general, a high spatial correlation should be used to obtain numerically more efficient updating schemes as well as a better model performance.

To evaluate the performance of the filters in the case of bias in the meteorological forcing, a test was performed where the wrong model is generated assuming that no wind is present. This test represents a very extreme case where the Kalman filter is adopted for the model to generate its own forcing. Figure 11 shows the r.m.s.e. of the wrong model and the updated model based on the EnKF with $\alpha = 0.97$, $\rho = 0.98$, $\sigma_s = 0.00025\ \mathrm{m^2\,s^{-2}}$, and $M = 200$ (the RRSQRT KF provides virtually similar results). A significant reduction of the r.m.s.e. is achieved; the average is reduced from 19.3 cm of the wrong model to 5.9 cm of the updated model. Also in this case, the use of coloured noise is very important for the efficiency of the updating scheme.

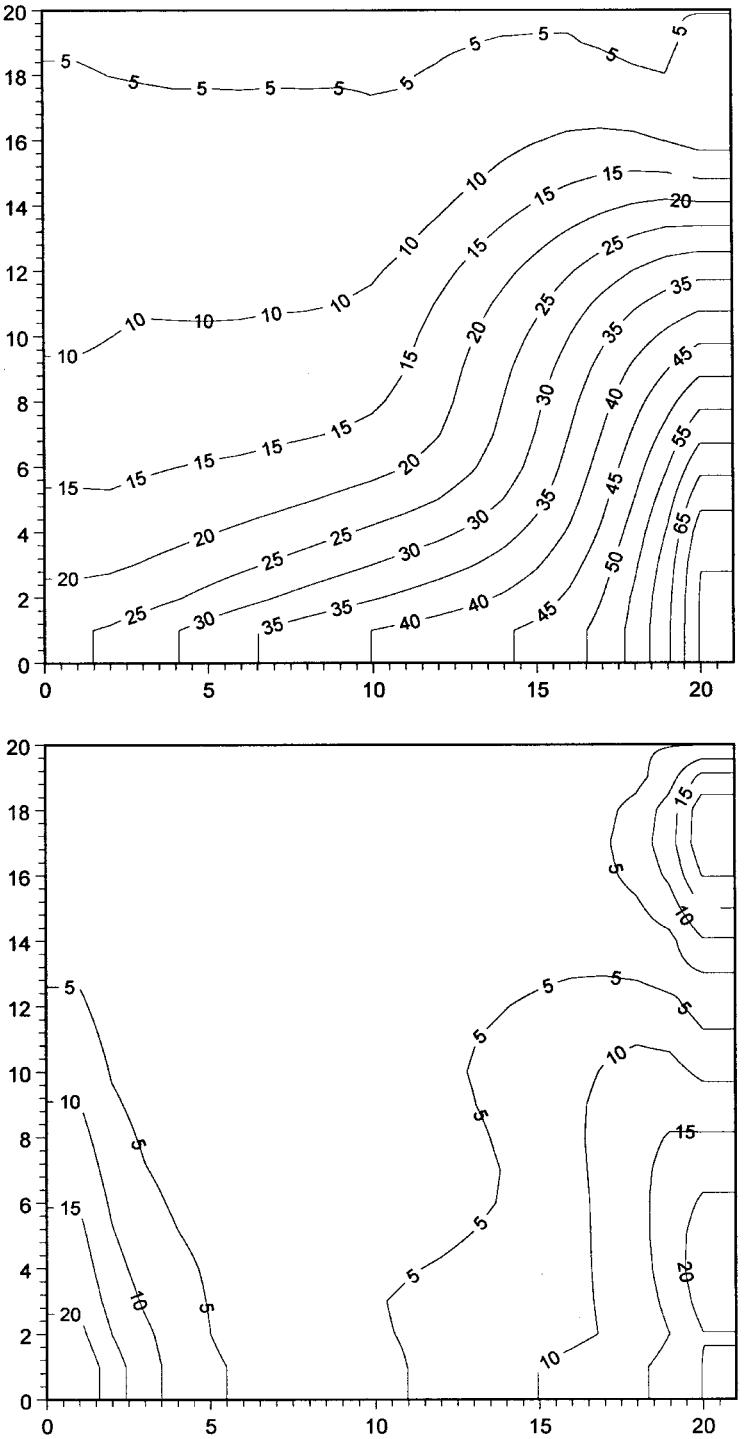

Figure 11. Spatial distribution of r.m.s.e. (cm) of the wrong model (top) and the updated model based on the EnKF (bottom).

5. DISCUSSION AND CONCLUSIONS

Two different data assimilation schemes based on the Kalman filter algorithm have been implemented into an existing two-dimensional hydrodynamic model. The RRSQRT KF is based on the extended Kalman filter scheme in which the propagation of the error covariance matrix is performed using a statistical linearisation of the model dynamics. The error covariance matrix is approximated by a matrix of lower rank using a square root factorisation. The EnKF is based on a Monte Carlo simulation approach for propagation of errors, according to the full non-linear model dynamics. The error covariance matrix is approximated by the ensemble estimate.

The filtering problem has been formulated by utilising a general description of the model noise process related to errors in the model forcing, i.e. open boundary conditions and meteorological forcing. An augmented state vector formulation is adopted for a general implementation of time coloured model noise. In this case, the Kalman filter provides corrections of the forcing terms along with corrections of the state of the system during assimilation.

The performance of the two Kalman filters has been evaluated using a twin experiment based on a hypothetical bay region. The following conclusions were obtained:

1. The performance of the filters depends on the number of leading eigenvalues for the RRSQRT KF and the ensemble size for the EnKF. In this case the reduced rank approximation of the error covariance matrix in the RRSQRT KF is more efficient than the ensemble representation in the EnKF, implying a smaller number of leading eigenvalues than ensemble members to obtain the same model performance. The eigenvalue decomposition in the RRSQRT KF, however, implies that the computational load becomes more expensive than determined by the rank of the error covariance matrix, and hence to a certain extent balances the additional costs of the EnKF caused by the larger ensemble size.
2. When the filters are provided with the true noise statistics, the estimated error covariance closely resembles the observed r.m.s.e. by using only relatively few leading eigenvalues in the RRSQRT KF and ensemble members in the EnKF as compared with the dimension of the state vector. Thus, for both filters, the error covariance approximation sufficiently resolves the error propagation in the model at a computational load that is significantly smaller than required by the full Kalman filter algorithm.
3. The efficiency of the Kalman filter generally depends on the prescribed error statistics. In practical applications error statistics are never known with certainty, and hence the robustness of the Kalman filter with respect to the specification of the error statistics is important. The tests revealed that both filters are very robust. Even in the case of a severe biased model error, the filters are able to efficiently correct the wrong state. The estimation of the error covariance was seen to be more sensitive with respect to the definition of the error statistics.
4. In general, the use of coloured noise provides a numerically more efficient algorithm as well as a better performance of the filter. The use of coloured noise is especially important for tracking and correcting biased model errors. In the case of errors in the meteorological forcing, a high spatial correlation generally leads to more efficient updating schemes.

In the test examples presented herein both filters have shown to provide very efficient updating schemes. However, since the RRSQRT KF is based on an extended Kalman filter formulation, it is expected to work well only for weakly non-linear dynamics. For strongly non-linear dynamics, the RRSQRT KF may fail. The EnKF does not have this limitation and it is

expected to be particularly useful when large non-linearities and discontinuities are present, e.g. in the case of flooding and drying of tidal areas. Thus, for general applications, the EnKF seems to be the most appropriate method. On the other hand, the error covariance approximation in the RRSQRT KF has shown to be very efficient since it focuses only on the most important eigenvalues of the error covariance matrix, and hence is able to filter out noise in the error representation. Application of a hybrid Kalman filter scheme that combines the reduced rank approximation of the error covariance matrix in the RRSQRT KF and the Monte Carlo based error propagation in the EnKF is presently being investigated.

ACKNOWLEDGMENTS

This work was funded by the Danish National Research Foundation through the International Research Centre for Computational Hydrodynamics.

REFERENCES

1. M. Ghil and P. Malanotte-Rizzoli, 'Data assimilation in meteorology and oceanography', *Adv. Geophys.*, **33**, 141–266 (1991).
2. H.J. Vested, J.W. Nielsen, H.R. Jensen and K.B. Kristensen, 'Skill assessment of an operational hydrodynamic forecast system for the North Sea and Danish Belts', in D.R. Lynch and A.M. Davies (eds.), Quantitative Skill Assessment for Coastal Ocean Models, *Coast. Estuar. Studies*, **47**, 373–396 (1995).
3. W. Thacker and R.B. Long, 'Fitting dynamics to data', *J. Geophys. Res.*, **93**, 1227–1240 (1988).
4. R.W. Lardner, A.H. Al-Rabeh and N. Gunay, 'Optimal estimation of parameters for a two-dimensional hydrodynamical model of the Arabian Gulf', *J. Geophys. Res.*, **98**, 18229–18242 (1993).
5. M. Eknes and G. Evensen, 'Parameter estimation solving a weak constraint variational formulation for an Ekman model', *J. Geophys. Res.*, **102**, 12479–12492 (1997).
6. A.C. Lorenc, 'A global three-dimensional multivariate statistical interpolation scheme', *Mon. Weather Rev.*, **109**, 701–721 (1981).
7. R. Daley, *Atmospheric Data Analysis*, Cambridge University Press, Cambridge, UK, 1991.
8. M.M. Rienecker and R.N. Miller, 'Ocean data assimilation using optimal interpolation with a quasi-geostrophic model', *J. Geophys. Res.*, **96**, 15093–15103 (1991).
9. R.E. Kalman, 'A new approach to linear filter and prediction theory', *J. Basic Eng.*, **82D**, 35–45 (1960).
10. R. Todling and S.E. Cohn, 'Suboptimal schemes for atmospheric data assimilation based on the Kalman filter', *Mon. Weather Rev.*, **122**, 2530–2557 (1994).
11. D.P. Dee, 'Simplification of the Kalman filter for meteorological data assimilation', *Q. J. R. Meteorol. Soc.*, **117**, 365–384 (1991).
12. I. Fukumori and P. Malanotte-Rizzoli, 'An approximate Kalman filter for ocean data assimilation: an example with an idealized Gulf Stream model', *J. Geophys. Res.*, **100**, 6777–6793 (1995).
13. S.E. Cohn and R. Todling, 'Approximate data assimilation schemes for stable and unstable dynamics', *J. Meteorol. Soc. Jpn.*, **74**, 63–75 (1996).
14. A.W. Heemink, 'Identification of wind stress on shallow water surfaces by optimal smoothing', *Stochastic Hydrol. Hydraul.*, **4**, 105–119 (1990).
15. A.W. Heemink, K. Bolding and M. Verlaan, 'Storm surge forecasting using Kalman filtering', *J. Meteorol. Soc. Jpn.*, **75**, 305–318 (1997).
16. M. Verlaan and A.W. Heemink, 'Tidal flow forecasting using reduced rank square root filters', *Stochastic Hydrol. Hydraul.*, **11**, 349–368 (1997).
17. R. Cañizares, A.W. Heemink and H.J. Vested, 'Sequential data assimilation in a fully non-linear hydrodynamic model', in A. Müller (ed.), *Hydroinformatics '96*, Balkema, Rotterdam, 1996, pp. 463–470.
18. R. Cañizares, A.W. Heemink and H.J. Vested, 'Application of advanced data assimilation methods for the initialisation of storm surge models', *J. Hydraul. Res.*, **36**(4), 655–674 (1998).
19. G. Evensen, 'Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model', *J. Geophys. Res.*, **97**, 17905–17924 (1992).
20. G. Evensen, 'Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics', *J. Geophys. Res.*, **99**, 10143–10162 (1994).
21. G. Evensen and P.J. van Leeuwen, 'Assimilation of Geosat altimeter data for the Agulhas current using the ensemble Kalman filter with a quasi-geostrophic model', *Mon. Weather Rev.*, **124**, 85–96 (1996).
22. DHI, MIKE 21 HD (Hydrodynamic Module) Release 2.5, *User Guide and Reference Manual*, Danish Hydraulic Institute, Denmark, 1995.

23. C.K. Chui and G. Chen, *Kalman Filtering with Real-Time Applications*, Springer Series in Information Sciences No. 17, 2nd edn., T.S. Huang (ed.), Springer, Berlin, 1991.
24. R. Cañizares, 'On the application of data assimilation in regional coastal models', *Ph.D. Thesis*, IHE Delft/Technical University Delft, Delft, Netherlands, 1999.
25. G. Burgers, P.J. van Leeuwen and Evensen, G., 'On the analysis scheme in the ensemble Kalman filter', *Mon. Weather Rev.*, **126**, 1719–1724 (1998).
26. H. Madsen, 'On the use of Monte Carlo simulation methods for data assimilation in MIKE 21', *Internal Report*, International Research Centre for Computational Hydrodynamics, Danish Hydraulic Institute, Denmark, 1997.

