Perspectives for a data assimilation system in Tagus Estuary

Ângela Canas[a,*]

[a] *Instituto Superior Técnico, Departamento de Engenharia Mecânica, Av. Rovisco Pais, nº 1, 1049-001 Lisbon, Portugal*

Abstract

In this work a simplified extended Kalman filter correcting state along pre-specified error variability modes is applied, in a twin experiment, to a 2D hydrodynamic model of a large shallow Northeast Atlantic tidal estuary - the Tagus Estuary (Portugal) – and adjacent coast to correct a forecast error derived from mean sea level perturbation with tide gauge measurements. Forecast error departures from normal probability distribution are assessed to verify optimality of the filter correction. Considering a system state comprised of water level and zonal and meridional horizontal velocity components, the error variability modes are derived from the EOF analysis of historical model forecast estimates. The filter performance is assessed in several coastal locations in the context of different number of measurement locations and use of different error variability modes.

Important departures from normal probability distribution in the forecast error are detected in some of the locations, which are explained by bathymetry shallowness and circulation. The forecast state variability is very much dominated by astronomic tide, whose presence is evident in the first two EOFs which represent more than 70% of total variability. Error variability, which is not related to tide, appears to be captured only on lower eigenvalue EOFs and very affected by local processes. The consideration of the tide modes in the filter correction modes and 5 measurement locations causes degradation of forecasts in the velocity field, which is found to be partially compensated inside the estuary by the use of a larger number of measurement locations (10). With a basis of two-three correction modes chosen from lower eigenvalue EOFs and a scarce measurement network (5 locations) the filter achieves RMSE reductions in the range of 5%-25% both in water level and velocity fields, although filter correction remains inappropriate in velocity some coastal locations. The filter performance appears limited essentially by the capability of the correction modes to represent error variability. Based on previous research, some ways to progress in research are presented.

Keywords: Estuarine dynamics; Tide gauges; Modelling; Kalman filters; Twin experiment; MOHID Water; Portugal; Lisbon; Tagus Estuary; 38º44'N9º08'W

* Tel.: +351 218419435; fax: +351 218419423.
E-mail address: angela.maretec@ist.utl.pt.

1. Introduction

The Tagus Estuary (Fig.1) is one of the largest estuaries in Western Europe. It is an inundated valley with a submerged area of 370Km^2 in high tide maximum and 265 Km^2 in low tide minimum, with about one third of its surface consisting in intertidal areas.

Fig. 1

It is composed, from Tagus River to mouth, of three morphologically different areas: the upper, middle and lower estuary. The upper estuary has an inner delta and large marshland with small depth (5m maximum), directly influenced by the Tagus River flow. The middle estuary is a wide (maximum 14km) small depth (7m on average) area. Finally, the lower estuary has a small width (2km) and large depth (30m) channel with the cities of Lisbon (the Portuguese capital city) and Almada at its banks. Outside the estuary the continental shelf is narrow with an abrupt border slope of about 10% (Coelho, 2002). The dominant driver of hydrodynamics in the area is astronomic tide, with dominant semidiurnal period and maximum tide amplitude of 2m in spring tides.

According with the European Water Framework Directive (2000/60/EC) recommendations for the monitorization of the water quality in important surface water bodies using operational modelling, the Tagus Estuary is currently the object of an water quality pre-operational model (Fernandes, 2005) explored by Instituto Superior Técnico (Technical University of Lisbon).

These forecasts are affected by several error sources mainly associated with the boundary conditions, such as bathymetry, meteorological forcing and mean sea level evolution, and horizontal grid resolution (Leitão, 2002; Fernandes, 2005). This uncertainty and the need to avoid an expensive increase of grid resolution, makes a possible solution to improve the model solution the use data assimilation of actual system information gathered by measurements, as advocated by Mourre et al. (2006).

Data assimilation consists in a tool for optimal combination of real world information from measurements and dynamical processes information from models (e.g. Brusdal et al., 2003). Assimilation methods are commonly classified in two main groups (e.g. Malanotte-Rizzoli and Tziperman, 1996): sequential methods, based on statistical estimation theory, and variational methods, based on optimal control theory. The statistical estimation approach, based on the Kalman filter, is characterized by a lower initial investment in coding and the capability to incorporate gradual developments (Brasseur and Verron, 2006). The Kalman filter provides the best unbiased estimation of system state in the case of a linear model and perfectly known statistics for model and observations error. For application to non linear

models an Extended Kalman filter (EKF) is developed by Jazwinski (1970): this uses a linearization of the model for the time evolution of the system state forecast error covariance. A basic assumption of Kalman filters is that the error probability distribution function can be completely described by the first two statistical momentums, the average and the variance, meaning that the probability distribution function must be normal. In the presence of a non linear model this is necessarily not true, as evidenced by the work of Auclair and De Mey (2003), and the filter correction, which only takes into account these two moments, becomes non optimal.

One of the most promising EKF derived schemes is the Singular Evolutive Extended Kalman (SEEK) filter (Pham et al., 1998a). In this scheme the EKF computational cost and knowledge demands for filter initialization are reduced by considering the state forecast error covariance in a subspace determined by the dominant variability modes, which compose the correction basis. This subspace approximation has been shown to be physically justifiable (e.g. Brasseur and Verron, 2006) and that has been used in several coastal model applications in the Mediterranean Sea (e.g. Hoteit et al., 2004, 2005; Triantafyllou et al., 2005). SEEK is one of the preferred methods for small scale data assimilation applications and one of the most suitable methods for operational ocean forecasting systems (Brusdal et al., 2003).

A fixed correction basis version of this filter, the Singular Fixed Extended Kalman (SFEK) filter, is presented by Brasseur et al. (1999) and applied with success in coastal applications (for a recent example in the Mediterranean Sea see Barth et al., 2007). This filter reduces further the computational cost, which is of the same order of magnitude as a model simulation while the full SEEK cost is typically one order of magnitude more expensive. This makes it adequate for a first assessment of assimilation and measurement network capabilities while supporting a future expansion to more advance schemes as the full SEEK filter or variants (Brasseur and Verron, 2006).

The research on SEEK filter application is very scarce in terms applications not located in the Mediterranean region. This region is characterized at the basin scale by small astronomic tide amplitude which eliminates the potential problems that intertidal areas, through model non linearity generation pose to data assimilation. In the tide dominated region of the Northeast Atlantic recent data assimilation applications (Cañizares et al., 2001; Mourre et al., 2006; Torres et al., 2006) use the Ensemble Kalman filter (Evensen, 2003). This scheme has a better performance with highly non linear models (Madsen and Cañizares, 1999) but at the expense of a much larger computational cost (Brusdal et al., 2003). A variant of SEEK filter has been developed to improve the behaviour of the filter in non linear models, the SEIK

filter (Pham et al., 1998b), but this has also be mainly used in the Mediterranean region (Triantafyllou et al., 2005; Hoteit et al., 2005). The quantification of model non linearity is not generally explicitly addressed in research; one exception is the work of Auclair and De Mey (2003). This is a knowledge lack that hampers a proper choice of the most suitable data assimilation method to use in each specific data assimilation problem.

The Tagus Estuary and adjacent coastline have been the place of a network of operational tide gauges (Fig. 1). These gauges could potentially provide measurements for a future data assimilation system which could help improve forecasts of an operational model for the estuary.

In this paper is presented the implementation and validation of a first data assimilation application for a hydrodynamic model of the Tagus Estuary and adjacent coast using a small coast SFEK filter and tide gauge measurements to correct a mean sea level perturbation in model results. In a twin models experiment the normality of the forecast error probability distribution function is tested to assess the empirical non linearity of the model. This provides a quantification of empirical non linearity of the Tagus Estuary model which can be used in the future to contrast with other modelling applications. Then the data assimilation option is explored, paving the way for actual tide gauge assimilation tests at this estuarine site and increasing the research on SEEK filter applications in tide dominated systems.

2. Materials and methods

2.1. Model

The hydrodynamic model used in this work is an improved version of the pre-operational model of Tagus Estuary. It is accomplished with the MOHID Water Modelling System (Miranda et al., 2000; Martins et al., 2001; Leitão et al., 2005), a modular system, developed by Instituto Superior Técnico and Hidromod comprising a primitive equations hydrodynamic module with hydrostatic and Boussinesq approximations. MOHID Water has been used for several estuarine and coastal hydrodynamic and water quality studies (Cancino and Neves, 1998; Ruiz-Villarreal et al., 2002; Leitão et al., 2005; Vaz et al., 2007), including for the Tagus Estuary (Braunschweig et al., 2003).

The hydrodynamic model is composed of two barotropic 2D domains (Fig. 2). The larger domain 1 contains the Portuguese and Galician coast (from 36ºN to 45ºN and from 12ºW to 6º9'W), with grid resolution from 4km (at boundary) to 2km (at area of domain 2), providing tidal boundary conditions for a smaller domain 2. This domain (from 38ºN to 40ºN and from 10º30'W to 9ºW), with grid resolution from 2km (at boundary) to 300m (at Guia outfall location), contains the Tagus Estuary and adjacent

coastal interest area. For both domains is used an Arakawa C grid type. The number of grid cells is 218x324 for domain 1 and 162x162 for domain 2 (longitude x latitude).

Fig. 2.

The bathymetries (Fig. 2) are constructed from detailed bathymetric surveys data, available for Tagus Estuary and adjacent coast (informing most of the area comprised in domain 2) and Galician coastal areas (the Spanish data is produced by Instituto Hidrográfico de Cádiz and provided by METEOGALICIA), which are completed with ETOPO 2' resolution data.

The model boundary conditions consist in horizontal open boundary conditions, the vertical surface boundary condition and water inputs from rivers. The open boundary conditions are defined in domain 1 by imposing the tidal water level from the FES 95.2 tidal model (Le Provost et al., 1998) with constant mean sea level of 2.08m using the Blumberg and Kantha (1985) radiation boundary condition. In domain 2 it is considered the Flather (1976) radiation boundary condition considering as reference solution the domain 1 water level and water flow solutions at the boundary cells. The connection between the two domains solutions is performed one-way, with information passing from domain 1 to domain 2 only.

The choice of horizontal open boundary conditions and grid type is supported by the conclusions of the work of Leitão (2002) and Leitão et al. (2005) with MOHID Water in similar applications.

The vertical surface boundary condition is defined by forcing the model with the ECMWF ERA40 re-analysis daily fields (http://data.ecmwf.int/data/d/era40_daily/) for surface zonal and meridional wind velocities, which have 6 hours time resolution and 2º30'x2º30' spatial resolution.

The water inputs from rivers are considered for River Tagus, in the form of daily averaged flows measured in the Omnias hydrometric station (provided by INAG - Instituto Nacional da Água), and the other two main rivers draining to the estuary, Sorraia and Trancão (together representing less than a quarter of the Tagus River flow), in the form of a constant daily average flow. Other water sources have negligible effect on barotropic hydrodynamics of the estuary (Fernandes, 2005).

2.2. Data assimilation scheme

The SFEK filter used in this work follows the formulation of Pham et al. (1998a) and Brasseur et al. (1999) and it is described in this section following the notation of Ide et al. (1997).

The objective of the filter is to combine the model forecast state vector (x^f), which includes the model variables which are corrected directly by the assimilation, with information from the measurements vector

(y^o) to obtain a better estimation of the state, which is called the analysis (x^a). The algorithm is constituted by the initialization, the forecast step and the analysis step.

The initialization consists in obtaining a matrix representing the modelled system error covariance (error is assumed null averaged) with a low rank r. In this work this is done by assuming that the model state variability is representative of the model state error variability and performing Empirical Orthogonal Function (EOF) analysis of the covariance matrix obtained from historical model states and retaining only the dominant eigenvectors or EOFs (Hoteit et al., 2005). Latter it is found that this is not an adequate methodology and some dominant EOFs are discarded. The resulting covariance matrix for the initial state forecast, P^f, has thus the form

$$P^f(t_0) = LUL^T \qquad (1)$$

where U is the rxr diagonal matrix containing, in decreasing order, the r highest valued eigenvalues (λ) and L is the nxr matrix containing in columns the corresponding EOFs, which represent the initial filter correction directions basis, and T means the matrix transpose. The representation capability of each EOF (L^i) in reproducing state variability over time is expressed by its expansion coefficient α^i

$$\alpha^i(t) = \frac{P^f L^i}{\lambda^i} \qquad (2)$$

For simplification of the filter operations, P is instead taken as

$$P^f(t_0) = L_0 L_0{}^T \qquad (3)$$

$$L_0 = L[U]^{1/2} \qquad (4)$$

so that the U_0 corresponding to L_0 through equation 1 is provided by the identity matrix.

The filter receives at initial time the U_0 and L_0 and the initial system state.

After initialization the filter algorithm proceeds with the forecast-analysis cycle until the final simulation time is reached. The forecast step includes the following equation:

$$x^f(t_i) = M_{i-1}\left[x^a(t_{i-1})\right] \qquad (5)$$

M_{i-1} is the non linear model operator performing the evolution from time i-1 to time i.

The analysis step consists in the actualization of U_{i-1} to U_i and in obtaining the analysis:

$$U_i^{-1} = \rho U_{i-1}^{-1} + L_0{}^T H^T R^{-1} H L_0 \qquad (6)$$

$$x^a(t_i) = x^f(t_i) + L_0 U_0 L_0{}^T H^T R^{-1}\left[y_i^o - H\left[x^f(t_i)\right]\right] \qquad (7)$$

R is the covariance matrix of measurement error, considered uncorrelated between measurements. H is the observation operation, which transforms the model state variables in the measured variables. Both R and H are assumed time fixed for this work and H is considered a linear operator. ρ is a parameter valued between 0 and 1 called forgetting factor, which accounts, in a simple way, for several filter error sources such as the model error and the errors incurred by the filter assumptions (e.g. use of a low-rank covariance). This account is made through the amplification, increasing with decreasing value of forgetting factor, of the last state forecast error covariance. A value of 1 represents commonly the assumption of the perfect model, i.e., a model which does not introduces error in the state forecast.

The SFEK filter implementation in the MOHID Water framework is performed with the development of an assimilation pre-processor and a sequential assimilation module in MOHID Water. The pre-processor generates the initial state forecast error covariance in the form of L_0. The L_0 matrix thus produced and files containing the measurements are provided as input for the MOHID Water simulation. The SFEK filter is implemented in MOHID Water in a specific new module.

Intertidal areas are dealt in the assimilation scheme by considering for correction in every analysis step all the grid cells that are outside of the land mask, disregarding if they are actually covered with water in the model forecast for that specific time. All intertidal areas are considered in state definition and are included in EOF analysis. At the instants when these cells are not covered the model provides a reference value for state, which is not thought to affect the variability assessment except for the reduction of the actual number instants considered for covariance calculation. The transition from uncovered-covered state is always processed by the model based on each model time step results. This procedure is thought to assure that a smooth field of state variables exist when the intertidal areas cells become covered, reducing potential assimilation shock in these areas.

2.3. Twin experiment

The experiment is designed in order for the system error and the measurement error to be random, uncorrelated between them and unbiased, a requirement for Kalman filter methodology to be directly applicable (Jazwinski, 1970). Also the system error is intended to be realistic, so as to gather information on filter behaviour which could be useful in future real measurements data assimilation applications in the Tagus Estuary model.

The state properties considered in the twin experiment are the water level and the zonal (West to East) and meridional (South to North) components of horizontal velocity in every domain 2 grid computation

point, i.e. cell center for water level and cell faces for the velocity components; this causes that the dimension of the multivariate state vector is 47963. The option to consider velocity components instead of velocity modulus is chosen because in most tide gauge locations the water level correlates better with individual velocity components than with the velocity modulus (correlation coefficient in the range 0.04-0.61 instead of 0.01-0.21).

The measurements for assimilation are provided by the results of the Tagus Estuary model referred in section 2.1, which is called the true model in the twin experiment framework. The measurements consist in the water level taken at tide gauge locations with an added random normally distributed perturbation with null average and 0.02m standard deviation.

These are assimilated in a perturbed model derived by a random perturbation introduced at the mean sea level (water level without the tide signal) considered at the tide boundary condition of domain 1. This normally distributed perturbation has null average and 0.1m standard deviation, following the standard deviation of mean sea level error detected from contrast with tide gauge measurements available for the period from 01/11/1972 to 31/01/1973. Detiding is accomplished with the T_Tide harmonic analysis software (Pawlowicz et al., 2002).

True and perturbed model simulations, without assimilation, are made for the period from 01/10/1972 to 31/01/1973. The simulations start from an initial state composed by a 2.08m water level (the mean sea level calculated from a long time series of tide gauge 1 measurements) and null velocities and a gradual connection of the imposed tide water level (domain 1) and wind velocity forcing in the first day (both domains). The perturbed model simulation without assimilation is called the free run.

The perturbed model forecast error relative to the true model is analysed at tide gauge locations to assess the validity of the assumption of a normal probability distribution of the forecast error made in the Kalman filter methods and quantify the empirical non linearity of the model. The error is previously normalized and several descriptive statistics are calculated: average, skewness, kurtosis, percentage of values with absolute value smaller than 2. The values of these statistics are contrasted with the values characteristic of a normal distribution, i.e., null average and skewness, kurtosis valued 3 and 95% of values with absolute value smaller or equal to 2. This work is based on the methodology of Auclair and De Mey (2003). A qualitative classification of model non linearity is achieved by considering threshold levels for the statistics calculated.

The bias of the perturbed model state relative to true model state is verified to be small, bellow 5% of the centered RMSE (square root of the difference between the square RMSE and the square bias) in the vast majority of tide gauge locations for all state properties and always bellow 15%, so that the filter has an appropriate setting.

The domain 2 results of the free run after 01/11/1972 are considered for EOF analysis to obtain the state error covariance for use in assimilation simulations. The results are restricted to a set of 245 states sampled every 6 hours (which corresponds approximately to the Nyquist frequency of the semidiurnal tide harmonic components) due to available computational resources. This number of states seems adequate for estimating the covariance in this 2D model, as 200 states have been used for estimating the covariance for SFEK in 3D models (Barth et al., 2007), which potentially have more complex dynamics. Each state is normalized before the EOF analysis using as normalizing factor the inverse of the state property domain averaged standard deviation calculated from the free run results (similarly to Hoteit et al., 2005). In order to assess the processes being captured by EOFs, the EOF analysis is made also to the true model results and to perturbed model partial results obtained for the estuary and the out of estuary areas of domain 2.

Several perturbed model simulations with SFEK filter assimilation are made for the period from 01/11/1972 to 15/11/1972 (containing a spring-neap tide period) starting from the perturbed model state at 01/11/1972. In all simulations is considered a ρ value of 0.75, which is found to provide better filter in previous data assimilation tests (Canas et al., 2008). The filter performance is assessed in assimilating a different number of measurements, 5 or 10, and different sets of filter correction modes. The smaller set of measurements is taken at the tide gauge locations with the largest state free run RMSE and higher variability in the true and perturbed models as observed in the leading EOFs, in order for the filter to be sensible to measurement dynamics: the tide gauges chosen are 3, 5, 6, 8 and 12. These are tide gauges concentrated essentially in the Tagus Estuary exit channel (5, 6, 8), but also capturing the dynamics of the inner estuary (12) and the adjacent coast (3) (Fig. 1). For the 10 measurements simulation these are taken in all tide gauges locations except 1, 2, 7, 14 and 15 which has the smallest RMSE in the free run. The error of the perturbed model state relative to the true model in assimilation simulations is assessed after analysis step for all state properties through the calculation of the RMSE (Root Mean Square Error) obtained by equation (N is the number of variables):

$$RMSE=\sqrt{\frac{\sum\left(x^{t}-x^{p}\right)^{2}}{N}} \qquad (8)$$

3. Results and discussion

3.1. Forecast error study

The statistics for the normalized forecast error are presented in Table 1 for water level, Table 2 for zonal velocity and Table 3 for meridional velocity.

Table 1.

Table 2.

Table 3.

For the water level the model response is approximately linear except in the middle estuary (10) and the upper estuary (13 and 14). The model non linearities are more pronounced in the velocity fields. In zonal velocity these are distributed through the coast (2), lower estuary (7) and mainly in the upper estuary (13, 14 and 15). In the meridional velocity non linearity is appreciable in most locations, being present in the coast (2, 3) and in the three subdivisions of the estuary.

According to Table 4, model non linearities are associated to coastal bays outside the estuary and shallowness, river discharges and channels inside the estuary. The grid resolution, which is variable through the domain, is not apparently connected with non linearity.

Table 4.

3.2. EOF analysis

The EOF analysis of the perturbed model simulation results (without assimilation) reveals that the four EOFs corresponding to the leading eigenvalues represent about 85% of the total state variability: 55%, 17%, 10%, 4%, explained respectively by EOF 1, EOF 2, EOF 3 and EOF 4.

The dominant 2 EOFs present the signature of tide: a marked semidiurnal period in the evolution of the expansion coefficients. This is visible in the spatial patterns depicted in the leading EOF (Fig. 3 a)): an inverse correlation observed in the variability inside and outside the estuary in water level, derived from the propagation through the domain of the tidal wave, and the visibly high variability in the zonal velocity in the lower estuary channel. The larger variability in the meridional velocity is in the middle and upper estuary constrained areas. It is also interesting to note that, while the water level variability has comparable magnitude in most of the domain, for both components of velocity a large discrepancy exist between the different areas of the domain.

Fig. 3.

The first two EOFs are very similar in the true and perturbed models results and are robust to a change in domain extension, being present in results of EOF analysis considering only the estuary or only the area out of estuary. This is visible in the very high correlation coefficients observed between the expansion coefficients, as presented in Table 5.

Table 5.

EOFs 3 and 4 spatial patterns are not concordant and the expansion coefficients are not correlated in true and perturbed model results and are not so robust to change in domain extension (Table 5). The tide signature is not apparent in the expansion coefficients and hence it is considered that these modes of forecast variability are in fact representing modes of forecast error, which is by construction not related to tide. Since these EOFs are very different for the estuary and out of estuary there is evidence that these should be reflecting local hydrodynamic processes.

The contrast between the variability of original states and the variability of reconstructed states, using only the leading 4 EOFs (Fig. 3 b)), outlines that the magnitude of the difference is very dependent of location in the domain and also of the state property. In fact, in cells with larger depth and in the lower estuary channel the variability is generally well described. The same does not hold in lower depths, in the continental shelf and in the middle and upper estuary intertidal areas, reflecting the preponderance there of local hydrodynamic processes which cannot be captured by the full domain calculated EOFs.

A way to assess the general effect in the data assimilation of considering specific set of EOFs in a twin experiment is to compare the distribution through the domain of the standard deviation obtained with the states reconstructed with those EOFs and of the forecast error. This is a simple way to check if the correct covariances are being considered in the filter correction basis: a well conditioned filter correction basis should reflect in the relative magnitude of the standard deviation it conveys the relative magnitude of the forecast error standard deviation. In order to simplify this analysis only the state in the tide gauges locations are presented in Fig. 4.

Fig. 4.

It is possible to verify that considering the first 2 EOFs the standard deviation distribution of the reconstructed states and the standard deviation distribution of the forecast error are not at all concordant in the water level field. While the error has a marked pattern of decrease from the coast to the upper estuary, in the reconstructed states the standard deviation is larger in the lower and middle estuary. The

information conveyed by the following EOFs is more concordant with the relative magnitude of variability depicted by the error, specially the EOFs 3 and 4. In the velocity fields the discrepancy is not so severe, although the lower eigenvalue EOFs present a better representation of error standard deviation distribution.

The fact that the variability distribution is badly described in the water level field and but reasonable in the velocity fields evidences that the covariances between water level and velocity are not representing the error actual covariances. This is strong evidence that the data assimilation correction would not be adequate with the use of the two leading modes in the correction basis, causing forecast degradation instead of improvement. This degradation is clearly caused by the tide having strong influence on state variability and not a relevant effect in forecast error dynamics. These results suggest that the traditional assumption in SEEK filter research of the forecast error variability being represented by the forecast state variability is not applicable in the present case. However, it is expected that this degradation effect would be absent if the forecast error origin would be directly related with tide (e.g. incorrect definition of tide solution at open boundaries), and this traditional assumption would be adequate. Tide representation is not uncommonly imperfect in coastal models as the Tagus Estuary model and this has been linked to grid resolution (Fortunato et al., 1997; Fernandes, 2005) making the present work a valuable contribution for future attempts to perform tide correction through data assimilation in the Tagus Estuary.

Following these results, several model simulations with data assimilation are performed. In the section 3.3 are presented the most representative results, which concern the effect of the number of measurement locations and the influence of the choice of EOFs for the filter correction basis.

3.3. Data assimilation

3.3.1. Number of measurement locations

A first set of data assimilation simulations considers the leading 4 EOFs as filter correction basis and assess the sensibility of filter performance to different number of measurement locations. The analysis RMSE results are presented in Fig. 5 for tide gauge locations.

Fig. 5.

It can be seen that with 5 measurement locations the filter is able to correct the water level in most tide gauges with 5%-25% error reductions, more expressive at the coast and lower and middle estuary. In the velocity fields the locations with larger RMSE present also useful corrections, particularly in the zonal velocity. Globally the improvements concentrate essentially in the lower estuary. Worse results, with an

increase of the error relative to the free run, are however observed in the upper estuary locations, nearby Tagus and Sorraia river mouths. Outside the estuary the velocity fields are also badly corrected by the filter, with over 50% increase of RMSE. The determinant factor in the filter performance seems to be the magnitude of variability present in the filter correction basis, where the lower estuary variability is very well represented. Also it is clearly visible the effect of the distorted covariance conveyed by the two first EOFs referred in section 3.2 in the coastal and lower estuary results.

The assimilation of 10 measurements leads generally to a better performance of the SFEK filter relative to the 5 measurements case (Fig. 5). In fact, in locations where the assimilation of 5 measurements introduced errors (e.g. upper estuary) the RMSE increase is avoided or controlled. The performance in the water level field is maintained in the range of 5%-25% but the velocity fields for most locations now present a lower RMSE than in the free run with reductions in the range 5%-45%, particularly for zonal velocity. This improvement does not completely avoid the severe inconsistencies introduced by an incorrect correction basis in the coastal velocity field but evidences the marked influence of local processes in forecast error dynamics. It is expected that the use of more measurement locations would benefit the data assimilation scheme performance.

3.3.2. Composition of the correction basis

The analysis RMSE of data assimilation simulations with changing composition of correction basis are presented in Fig. 6 for the tide gauge locations. These simulations are performed using the smaller number of measurement locations to put in more evidence the effect of correction basis choice.

Fig. 6.

It is observed that the consideration of the two leading EOFs in the correction basis causes generally worse filter corrections than considering only the lower eigenvalue EOFs, the expected filter response given the analysis of reconstructed state standard deviation presented in section 3.2. The best results are achieved with a correction basis composed of EOFs 3-4 or EOFs 3-5, with reductions of RMSE of 5%-25% in the water level field and similar reductions in the velocity fields, although the improvements remain more expressive in the water level field. The use of EOFs 1 and 2 in conjunction with lower eigenvalue EOFs seems to provide worse results globally than to opt just to use the first two EOFs, which suggests that the two sets of EOFs, 1-2 and 3-4, describe distinct hydrodynamic processes and inconsistencies are possible generated at correction steps in consequence of this.

Although the general results suggest that set EOFs 1-2 should be inadequate in the filter correction basis, it is verified that it provides a more adequate control of RMSE increase relative to the free run in coastal locations 1 and 3 than the lower eigenvalue EOFs. Since the covariance described by the leading EOFs only is distorted relative to the actual error covariance, as demonstrated in section 3.2, this should be the result of a lower magnitude of the covariance conveyed by the last EOFs.

The results obtained for the set of EOFs 3-6 are generally worse than using EOFs 3-4 or 3-5 and in some cases even worse than considering the first two EOFs. This is probably a consequence of the lower precision by which EOF 6 is calculated, due to the small magnitude of the variability represented, which introduces errors in the correction step. This imprecision in EOF analysis is thought to affect also, though in a lower degree, EOFs 3-5. Hence, the representation of forecast error in the covariance estimated in the EOF analysis could be improved if a way was devised to improve the precision with which the non tidal variability is analysed.

From Fig. 6 is possible also to verify that the use of only the lower eigenvalue EOFs 3-4 or 3-5 provides a very similar filter response in the middle and upper estuary as the one observed with the use of 10 measurement locations. This suggests that filter correction basis inadequacies can be partially be compensated by increasing the extension of the assimilated measurement network inside the estuary, adding extra evidence of this effect to the results presented in section 3.3.1.

Given these results the main limitation for a better filter performance seems to be the representation of the error variability in the filter correction basis. Local processes are important for the description of the forecast error in this twin experiment. Additional data assimilation simulations are made using a full SEEK filter and the set of EOFs 1-4 as initial correction basis. These provide inexpressive filter corrections in both water level and velocity fields and are not shown here. These SEEK filter results indicate that a badly conditioned correction basis at the filter initialization cannot be corrected by allowing the evolution with time of correction modes.

The use of a global EOF analysis for the generation of the correction basis is shown to have limitations in this Tagus Estuary case because the dominance of tide in state forecast variability causes that forecast error variability modes to be calculated with small precision. Ways to deal with this limitation can be investigated in the future. Previous studies provide some evidence on how to proceed in this research. Hoteit et al. (2001) provides a SEEK filter with a local correction basis. Brasseur et al. (1999) present a learning mechanism which can be used in a correction basis evolving SEEK filter to correct the basis

modes at correction steps, using the information from the contrast of model forecast and measurements. This last option would profit from the consideration of a large measurement network.

At this stage of research it is not clear if the use of SEIK filter would provide a considerable increase of data assimilation performance relative to a SEEK or a SFEK filter, since model non linearities are not perceived as the determinant factor limiting the data assimilation performance.

5. Conclusions

A simplified extended Kalman filter data assimilation scheme is implemented for a barotropic hydrodynamic model for the Tagus Estuary and adjacent coast and applied in the correction of a mean sea level error with tide gauge measurements in a twin experiment. The forecast error probability distribution function is found to have important departures from a normal distribution in some locations, particularly in the velocity field. These departures are explained by the shallowness of bathymetry and circulation characteristics of the estuary and adjacent coast and should cause the Kalman filter application to provide a non optimal correction of state.

Analysis of the model forecast by EOF analysis evidence a strong dominance of astronomic tide in state variability, concentrating in the first two modes. The forecast error considered in the twin experiment is confirmed not to be significantly affected by tide but instead by local hydrodynamic processes.

The results of data assimilation simulations, considering pre-specified correction modes derived from EOF analysis, indicate that the two tide dominated modes should not be considered for correction basis composition as they cause, through a distorted error covariance, important degradation of forecast in some locations, especially in the velocity field. This distortion is found to be partially compensated inside the estuary with the use of a large measurement number. A number of modes in the correction basis larger than 3 is found to be inadequate due to possible error introduction in the forecast through small precision estimated low eigenvalue EOFs.

Acknowledgments

This work was financed by a grant from Fundação para a Ciência e Tecnologia, for the author's Ph.D. studies, under contract SFRH / BD / 14185 / 2003, by the European Commission through the project INSEA (http://www.insea.info/) under contract SST4-CT-2005-012336 and by project EASY (http://www.project-easy.info/) of the INTERREG IIIB - ATLANTIC AREA program.

The author would like to thank Ibrahim Hoteit, George Triantafyllou and his colleagues from the Hellenic Center for Marine Research, and Nuno Vaz and Guillaume Riflet from MARETEC for useful discussion

on EOF analysis, SEEK filtering and research results. Acknowledge is also made to Aires dos Santos and Paulo Leitão for review of a previous version of this paper and support supplied as the author's Ph.D. supervisors.

References

Auclair, F., De Mey, P., 2003. Space-time structure and dynamics of the forecast error in a coastal circulation model of the Gulf of Lions. Dynamics of Atmosphere and Oceans, 36, 309-346.

Barth, A., Alvera-Azcárate, A., Beckers, J.-M., Rixen, M., Vandenbulcke, L., 2007. Multigrid state vector for data assimilation in a two-way nested model of the Ligurian Sea. Journal of Marine Systems, 65, 41-59.

Blumberg, A., Kantha, L., 1985. Open boundary condition for circulation models. Journal of Hydraulic Engineering, ASCE, 111, 237-255.

Brasseur, P., Ballabrera-Poy , J., Verron, J., 1999. Assimilation of altimetric data in the mid-latitude oceans using the Singular Evolutive Extended Kalman filter with an eddy-resolving, primitive equation model. Journal of Marine Systems, 22, 269-294.

Brasseur, P., Verron, J., 2006. The SEEK filter method for data assimilation in oceanography: a synthesis. Ocean Dynamics, 56, 560-661.

Braunschweig, F., Martins, F., Chambel, P., Neves, R., 2003. A methodology to estimate renewal time scales in estuaries: the Tagus Estuary case. Ocean Dynamics, 53, 137-145.

Brusdal, K., Brankart, J., Halberstadt, G., Evensen, G., Brasseur, P., van Leeuwen, P., Dombrowsky, Verron, J., 2003. A demonstration of ensemble-based assimilation methods with a layered OGCM from the perspective of operational ocean forecasting systems. Journal of Marine Systems, 40-41, 253-289.

Canas, A., Santos, A., Leitão, P., 2008. Implementation and validation of a SFEK data assimilation application for an hydrodynamic model of the Tagus Estuary. Proceedings of the XI International Symposium on Oceanography of the Bay of Biscay, Revista de Investigación Marina, 3, 159-160.

Cancino, L., Neves, R., 1998. Hydrodynamic and sediment suspension modelling in estuarine systems part II: Application to the Western Scheldt and Gironde estuaries. Journal of Marine Systems, 22 (2-3), 117-131.

Cañizares, R., Madsen, H., Jensen, H., Vested, H., 2001. Developments in Operational Shelf Sea Modelling in Danish Water. Estuarine, Coastal and Shelf Science, 53, 595-605.

Coelho, H., 2002. Modelação de Processos Físicos Relacionados com a Circulação Oceânica na Margem Continental Ibérica. Ph.D. Thesis, Universidade Técnica de Lisboa, Instituto Superior Técnico, Lisbon, Portugal, unpublished.

Evensen, G., 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dynamics, 53, 343-367.

Fernandes, R., 2005. Modelação Operacional no Estuário do Tejo. M. Sc.Thesis, Universidade Técnica de Lisboa, Instituto Superior Técnico, Lisbon, Portugal, unpublished.

Flather, R., 1976. A Tidal Model of the North-West European Continental shelf. Memoires Societe Royale des Sciences de Liege, 6e serie, tome X, 141-164.

Fortunato, A., Baptista, A., Luettich, R., 1997. A Three-Dimensional Model of Tidal Currents in the Mouth of Tagus Estuary. Continental Shelf Research, 17(4), 1689-1714.

Hoteit, I., Pham, D.-T., Blum, J., 2001. A Semi-Evolutive Partially Local Filter for Data Assimilation. Marine Pollution Bulletin, 43 (7-12), 164-174.

Hoteit, I., Korres, G., Triantafyllou, G., 2005. Comparison of extended and ensemble based Kalman filters with low and high resolution primitive equation ocean models. Nonlinear Processes in Geophysics, 12, 755-765.

Ide, K., Courtier, P., Ghil, M., Lorenc, A., 1997. Unified Notation for Data Assimilation: Operational, Sequential and Variational. Journal of Meteorological Society of Japan, 75(1B), 181-189.

Jazwinsky, A., 1970. Stochastic Processes and Filtering Theory. Academic Press, New York.

Leitão, P., 2002. Integração de Escalas e de Processos na Modelação do Ambiente Marinho. Ph.D. Thesis, Universidade Técnica de Lisboa, Instituto Superior Técnico, Lisbon, Portugal, unpublished.

Leitão, P., Coelho, H., Santos, A., Neves, R., 2005. Modelling the main features of the Algarve coastal circulation during July 2004: A downscaling approach. Journal of Atmospheric & Ocean Science, 10 (4), 421-462.

Le Provost, C., Lyard, F., Molines, J., Genco, M., Rabilloud, F., 1998, A hydrodynamic ocean tide model improved by assimilating a satellite altimeter derived data set. Journal of Geophysical Research – Oceans, 103, 5513-5529.

Madsen, H., Cañizares, R., 1999. Comparison of extended and ensemble Kalman filters for data assimilation in coastal area modelling. Int J. Numer. Meth. Fluids, 31, 961-981.

Malanotte-Rizzoli, P., Tziperman, E., 1996. The Oceanographic Data Assimilation Problem: Overview, Motivation and Purposes. In: Malanotte-Rizzoli, P. (Ed.), Modern Approaches to Data Assimilation in Ocean Modeling. Elsevier Oceanography Series, 61, Elsevier Science B.V., Amsterdam, pp. 3-17.

Martins, F., Leitão, P., Silva, A., Neves, R., 2001. 3D modelling in the Sado estuary using a new generic vertical discretization approach. Oceanologica Acta, 24 (1), 551–562.

Miranda, R., Braunschweig, F., Leitão, F., Martins, F., Santos, A., 2000. MOHID 2000 – a coastal integrated object oriented model. Hydraulic Engineering Software, VIII.

Mourre, B., Crosnier, L., Le Provost, C., 2006. Real-time sea-level gauge observations and operational oceanography. Philosophical Transactions of the Royal Society, 364, 867-884.

Pawlowicz, R., Beardsley, B., Lentz, S., 2002. Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. Computer & Geosciences, 28, 929-937.

Pham, D., Verron, J., Roubaud, M., 1998 (a). A singular evolutive extended Kalman filter for data assimilation in oceanography. Journal of Marine Systems, 16, 323-340.

Pham, D., Verron, J., Gourdeau, L., 1998 (b). Singular evolutive Kalman filters for data assimilation in oceanography. C. R. Acad. Sci., 326, 255-260.

Ruiz-Villarreal, M., Montero, P., Taboada, J., Prego, R., Leitão, P., Pérez-Villar, V., 2002. Hydrodynamic Model Study of the Ria de Pontevedra Under Estuarine Conditions. Estuarine, Coastal and Shelf Science, 54, 101-113.

Torres, R., Allen, J., Figueiras, F., 2006. Sequential data assimilation in an upwelling influenced estuary. Journal of Marine Systems, 60, 317-329.

Triantafyllou, G., Hoteit, I., Petihakis, G., 2005. Ecosystem Modelling and Data Assimilation of Physical Biochemical Processes in Shelf and regional Areas of the Mediterranean Sea. Appl. Num. Anal. Comp. Math., 2, 262-280.

Vaz, N., Dias, J., Leitão, P., Nolasco, R., 2007. Application of the Mohid-2D model to a mesotidal temperate coastal lagoon. Computers & Geosciences, 33, 1204-1209.

Verlaan, M., 1998. Efficient Kalman Filtering Algorithms for Hydrodynamic Models. Ph.D. Thesis, Technical University of Delft, Delft, The Netherlands, unpublished.

Figure captions

Fig. 1. The Tagus Estuary. The cities of Lisbon and Almada and the two most important rivers that drain to the estuary, Tagus and Sorraia, together with the Guia wastewater submarine outfall are presented. The tide gauges network is indicated by the numbered locations.

Fig. 2. The Tagus Estuary model two domains. Domain 1 contains the Portuguese and Galician coast and domain 2 includes the Tagus Estuary and adjacent coast. Bathymetry is shown.

Fig. 3. Perturbed model state EOF 1 for state properties (a) and standard deviation difference between originally sampled and reconstructed states (b). The fields for water level, zonal velocity and meridional velocity (Mer. velocity) are presented in absolute values. Analysis is performed using model results from 01/11/1972 to 01/01/1973 sampled every 6 hours. The reconstructed fields for water level, zonal velocity (Zonal vel.) and meridional velocity (Mer. vel.) are obtained for the same time instants of the original sampled states using the leading 4 EOFs.

Fig. 4. Standard deviation of reconstructed perturbed model forecast states (rec. state) using a different set of EOFs and standard deviation of forecast error of states used in EOF analysis for tide gauge locations. Magnitude of values for reconstructed states using EOFs 1-2 and 1-4 is provided by left axis and magnitude for the remaining values is provided by right axis.

Fig. 5. SFEK filter analysis RMSE using different number of measurement locations. RMSE results for water level (a), zonal velocity (b) and meridional velocity (c) obtained with the leading 4 EOFs and 5 and 10 measurement locations are presented as a fraction of the free run simulation RMSE (left axis); values not shown present larger RMSE. In broken line are presented the RMSE of the free run (right axis). RMSE is calculated with the values for the period from 01/11/1972 to 15/11/1972 sampled every 6 hours.

Fig. 6. SFEK filter analysis RMSE using different EOFs as filter correction basis. RMSE results for water level (a), zonal velocity (b) and meridional velocity (c) obtained with the leading EOFs 1-2, 1-4, 3-4, 3-5 and 3-6 are presented as a fraction of the free run simulation RMSE (left axis); values not shown present larger RMSE. In broken line are presented the RMSE of the free run (right axis). RMSE is calculated with the values for the period from 01/11/1972 to 15/11/1972 sampled every 6 hours.

Tables

Table 1. Descriptive statistics for water level state normalized error for tide gauge locations from the perturbed model simulation. Normalized error is obtained by the quotient between error and error standard deviation. Statistics are calculated from hourly Domain 2 results for the period 01/11/1972 0h - 31/01/1973 17h. Tide gauges indicated from open ocean (Peniche) to Tagus River (Vila Franca de Xira).

Tide gauge	Average	Skewness	Kurtosis	% Absolute values ≤ 2	Important departure from normal distribution[a]
1 - Peniche	-0.03	-0.07	2.89	95.91	No
2 - Sesimbra	0.01	0.01	2.91	95.82	No
3 - Cascais	-0.01	0.00	2.81	96.00	No
4 - Paço de Arcos	-0.01	0.13	2.97	95.91	No
5 - Trafaria	-0.03	0.07	2.95	95.59	No
6 - Cacilhas	-0.05	-0.10	3.28	94.60	No
7 - Alfeite	-0.05	-0.01	3.19	95.41	No
8 - Lisboa	-0.04	-0.06	3.20	94.96	No
9 - Cabo Ruivo	0.01	0.07	2.95	95.55	No
10 - Seixal	0.14	-0.01	3.22	94.69	Yes
11 - Montijo	0.00	0.11	3.27	95.10	No
12 - Alcochete	0.00	-0.08	3.29	94.91	No
13 - Ponta de Erva	0.06	-0.08	4.89	93.64	Yes
14 - Póvoa de Santa Iria	0.01	-0.16	4.96	94.28	Yes
15 - Vila Franca de Xira	-0.03	-0.07	2.89	95.91	No

Note: [a] Yes if any of the following threshold levels is exceeded: 0.1 for absolute average, 0.5 for absolute skewness, 0.5 for absolute Kurtosis difference from 3, 1% for absolute value number; No if otherwise.

Table 2. Descriptive statistics for zonal velocity state normalized error for tide gauge locations from the perturbed model simulation. Normalized error is obtained by the quotient between error and error standard deviation. Statistics are calculated from hourly Domain 2 results for the period 01/11/1972 0h - 31/01/1973 17h. Tide gauges indicated from open ocean (Peniche) to Tagus River (Vila Franca de Xira).

Tide gauge	Average	Skewness	Kurtosis	% Absolute values ≤ 2	Important departure from normal distribution[a]
1 - Peniche	0.01	0.01	2.99	95.37	No
2 - Sesimbra	-0.06	0.13	4.46	94.50	Yes
3 - Cascais	0.01	-0.05	2.93	95.28	No
4 - Paço de Arcos	-0.03	0.04	3.20	95.05	No
5 - Trafaria	0.00	0.11	3.41	95.32	No
6 - Cacilhas	-0.07	-0.07	3.10	95.37	No
7 - Alfeite	0.10	0.56	4.93	94.55	Yes
8 - Lisboa	-0.02	-0.08	3.09	95.50	No
9 - Cabo Ruivo	0.05	0.06	3.44	94.87	No
10 - Seixal	0.03	5.65	126.25	97.68	Yes
11 - Montijo	-0.02	-0.07	3.40	94.10	No
12 - Alcochete	0.03	0.12	3.37	94.14	No
13 - Ponta de Erva	-0.01	0.21	5.64	92.69	Yes
14 - Póvoa de Santa Iria	-0.01	0.97	8.27	93.78	Yes
15 - Vila Franca de Xira	-0.05	0.00	7.39	93.69	Yes

Note: [a] Yes if any of the following threshold levels is exceeded: 0.1 for absolute average, 0.5 for absolute skewness, 0.5 for absolute Kurtosis difference from 3, 1% for absolute value number; No if otherwise.

Table 3. Descriptive statistics for meridional velocity state normalized error for tide gauge locations from the perturbed model simulation. Normalized error is obtained by the quotient between error and error standard deviation. Statistics are calculated from hourly Domain 2 results for the period 01/11/1972 0h - 31/01/1973 17h. Tide gauges indicated from open ocean (Peniche) to Tagus River (Vila Franca de Xira).

Tide gauge	Average	Skewness	Kurtosis	% Absolute values ≤ 2	Important departure from normal distribution[a]
1 - Peniche	0.07	0.13	3.13	95.37	No
2 - Sesimbra	-0.07	-0.09	4.01	94.19	Yes
3 - Cascais	-0.03	-0.21	3.55	95.05	Yes
4 - Paço de Arcos	-0.02	0.05	3.02	95.46	No
5 - Trafaria	-0.10	0.01	3.34	95.14	No
6 - Cacilhas	0.08	0.73	5.14	93.55	Yes
7 - Alfeite	-0.06	-0.32	3.98	94.82	Yes
8 - Lisboa	0.00	-0.07	3.30	95.14	No
9 - Cabo Ruivo	0.04	0.01	3.40	95.19	No
10 - Seixal	0.01	-0.08	3.82	94.91	Yes
11 - Montijo	0.05	0.03	3.94	94.50	Yes
12 - Alcochete	0.02	0.12	3.25	94.87	No
13 - Ponta de Erva	-0.01	0.23	7.48	93.87	Yes
14 - Póvoa de Santa Iria	-0.02	0.79	6.76	93.78	Yes
15 - Vila Franca de Xira	-0.01	-0.14	8.41	93.69	Yes

Note: [a] Yes if any of the following threshold levels is exceeded: 0.1 for absolute average, 0.5 for absolute skewness, 0.5 for absolute Kurtosis difference from 3, 1% for absolute value number; No if otherwise.

Table 4. Departure from normal distribution of perturbed model error and bathymetric features of tide gauge locations. Tide gauges indicated from open ocean (Peniche) to Tagus River (Vila Franca de Xira).

Tide gauge	Important departure from normal distribution of Free-run error[a]	Grid resolution (mxm)	Bathymetry depth (m)	Type of location
Peniche	No	300x2000	56	Coast
Sesimbra	Yes (U, V)	600x2000	12	Coastal bay (southward)
Cascais	Yes (V)	500x300	7	Coastal bay (southward)
Paço de Arcos	No	400x300	5	Coast/estuary exit channel
Trafaria	No	300x300	11	Estuary exit channel
Cacilhas	Yes (V)	500x300	5	Estuary exit channel
Alfeite	Yes (U, V)	500x300	4	Intertidal
Lisboa	No	600x300	7	Estuary exit channel
Cabo Ruivo	No	700x500	3	Estuary
Seixal	Yes (WL, U, V)	700x300	-3	Narrow estuarine channel/intertidal
Montijo	Yes (V)	700x300	0	Estuarine channel/intertidal
Alcochete	No	700x500	1	Estuary
Ponta de Erva	Yes (WL, U, V)	700x500	0.5	Intertidal /riverside
Póvoa de Santa Iria	Yes (WL, U, V)	700x500	1	Narrow estuarine channel
Vila Franca de Xira	Yes (U, V)	700x1000	1	Estuarine channel/riverside

Note: [a] Yes if any of the following threshold levels is exceeded: 0.1 for absolute average, 0.5 for absolute skewness, 0.5 for absolute Kurtosis difference from 3, 1% for absolute value number; No if otherwise. WL = water level; U = zonal velocity; V = meridional velocity.

Table 5. Correlation coefficient between expansion coefficients of the leading EOFs as derived from true model, perturbed model global EOF analysis and perturbed model partial EOF analysis for the estuary and out estuary.

EOF	True model - perturbed model	Perturbed model global – perturbed model estuary	Perturbed model global – perturbed model out of estuary
1	0.99[a]	0.96[a]	0.97[a]
2	0.99[a]	0.97[a]	0.97[a]
3	0.14	-0.40[a]	0.88[a]
4	-0.16	-0.05	-0.94[a]

Note: [a] significant relative to the 1% level of significance.

Figures

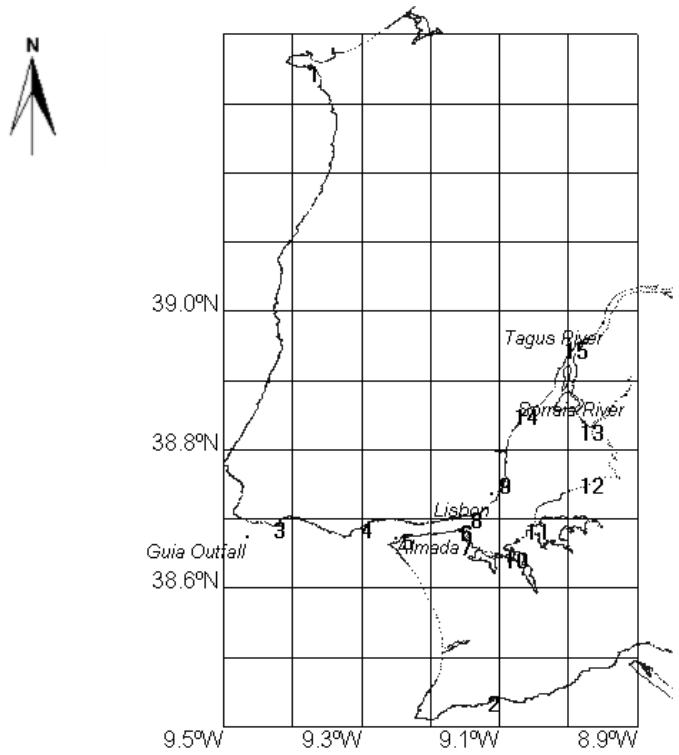

Fig. 1.

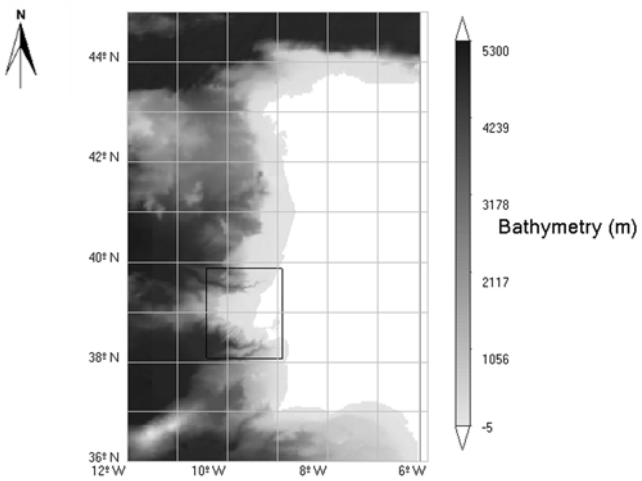

Fig. 2.

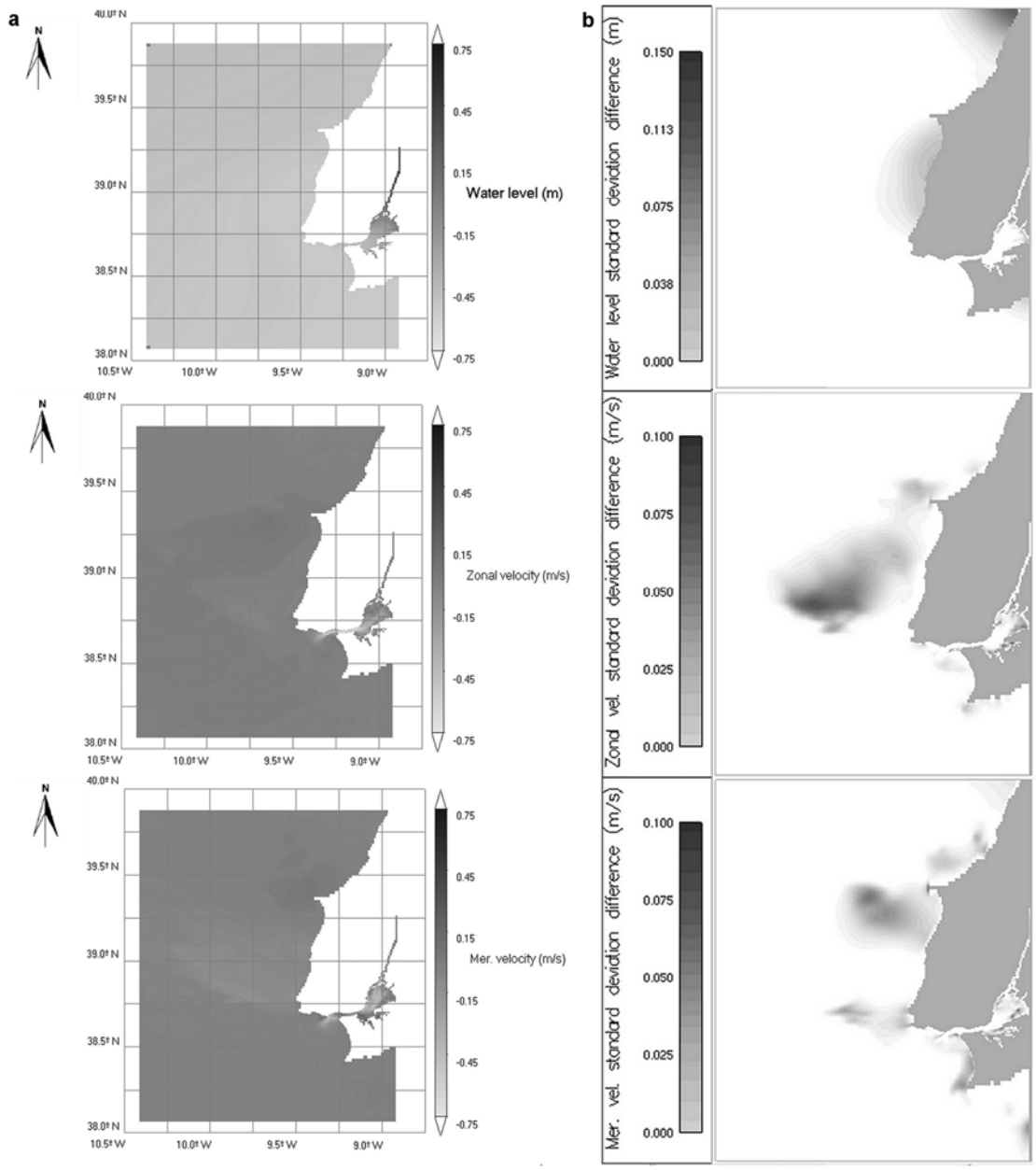
a
b
N
40.0° N
39.5° N
39.0° N
38.5° N
38.0° N
10.5° W
10.0° W
9.5° W
9.0° W
0.75
0.45
0.15
-0.15
-0.45
-0.75
Water level (m)
Zonal velocity (m/s)
Mer. velocity (m/s)
Wder level standard deviation difference (m)
Zond vel. standard deviation difference (m/s)
Mer. vel. standard deviation difference (m/s)
0.150
0.113
0.075
0.038
0.000
0.100
0.050
0.025

Fig. 3.

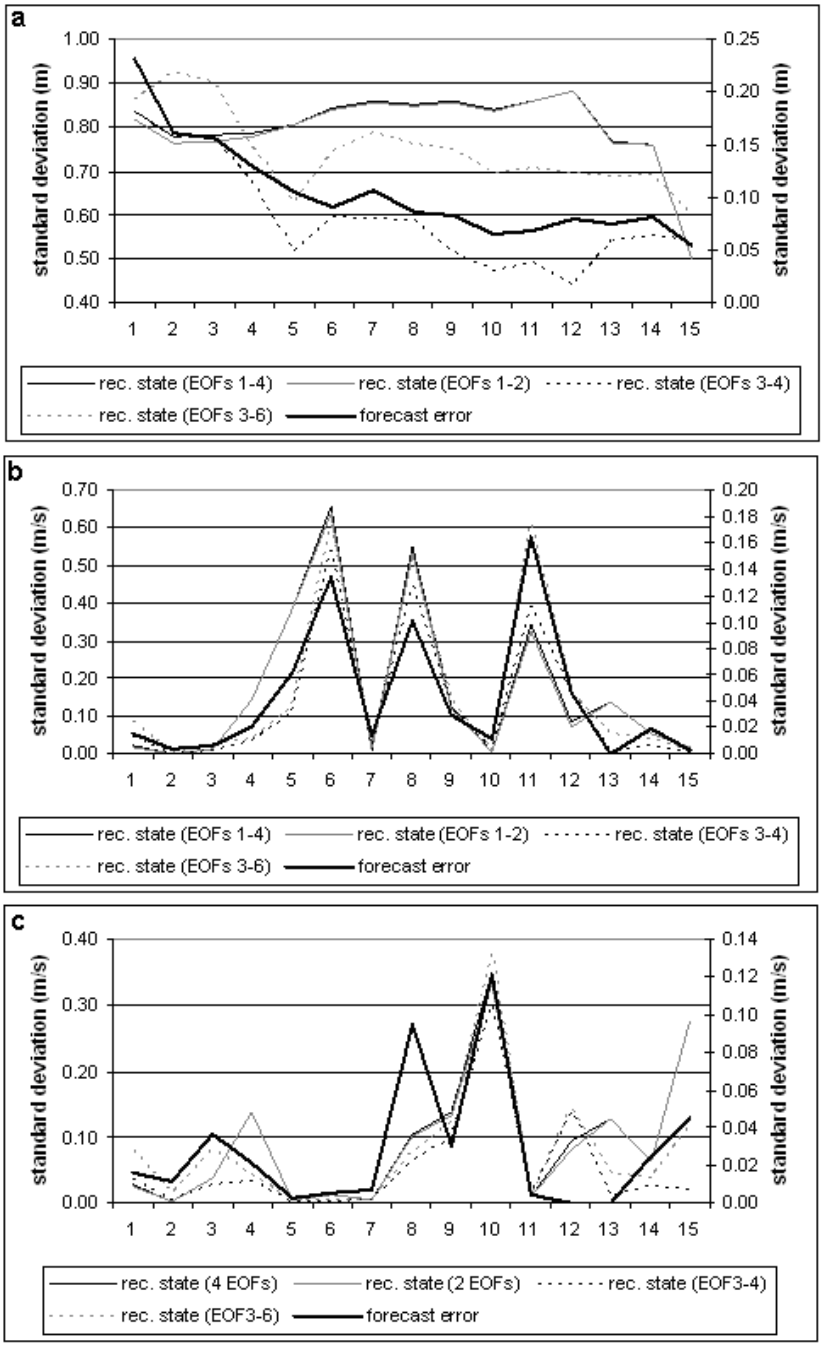
a
standard deviation (m)
standard deviation (m)
rec. state (EOFs 1-4)
rec. state (EOFs 1-2)
rec. state (EOFs 3-4)
rec. state (EOFs 3-6)
forecast error
b
standard deviation (m/s)
standard deviation (m/s)
rec. state (EOFs 1-4)
rec. state (EOFs 1-2)
rec. state (EOFs 3-4)
rec. state (EOFs 3-6)
forecast error
c
standard deviation (m/s)
standard deviation (m/s)
rec. state (4 EOFs)
rec. state (2 EOFs)
rec. state (EOF3-4)
rec. state (EOF3-6)
forecast error

Fig. 4.

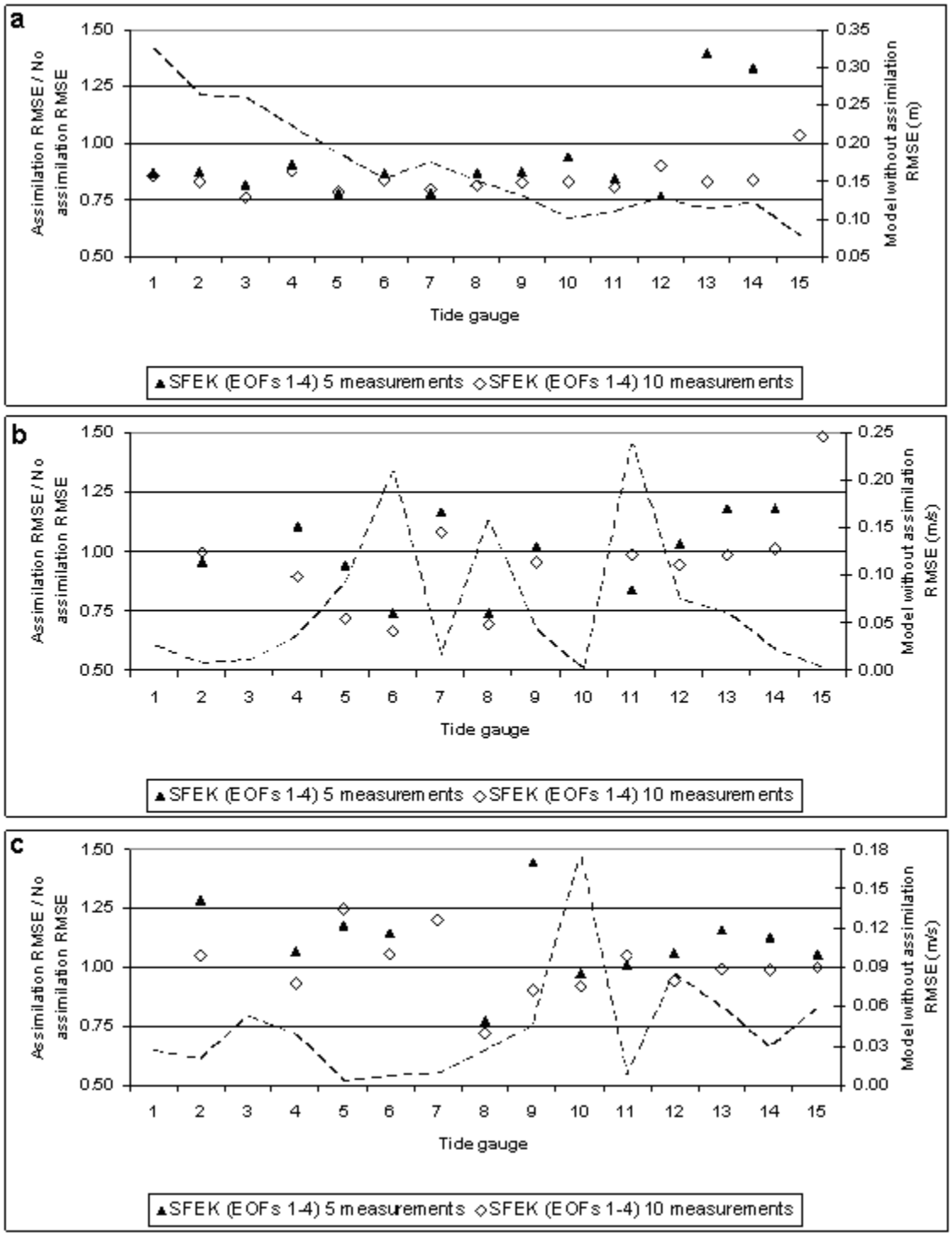
a
Assimilation RMSE / No assimilation RMSE
Model without assimilation RMSE (m)
Tide gauge
▲SFEK (EOFs 1-4) 5 measurements ◇SFEK (EOFs 1-4) 10 measurements
b
Assimilation RMSE / No assimilation RMSE
Model without assimilation RMSE (m/s)
Tide gauge
▲SFEK (EOFs 1-4) 5 measurements ◇SFEK (EOFs 1-4) 10 measurements
c
Assimilation RMSE / No assimilation RMSE
Model without assimilation RMSE (m/s)
Tide gauge
▲SFEK (EOFs 1-4) 5 measurements ◇SFEK (EOFs 1-4) 10 measurements

Fig. 5.

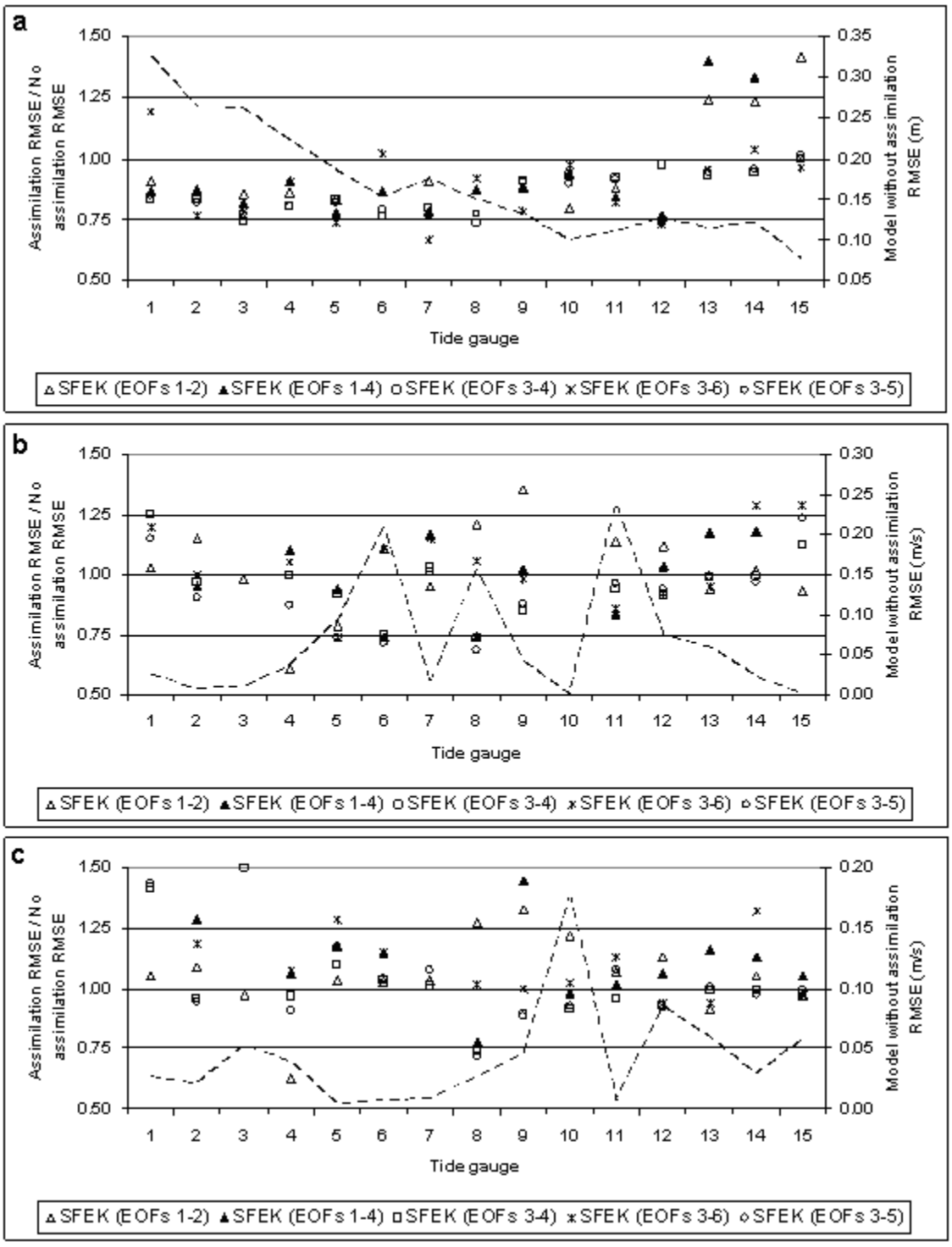
a
Assimilation RMSE / No assimilation RMSE
Model without assimilation RMSE (m)
Tide gauge
SFEK (EOFs 1-2) SFEK (EOFs 1-4) SFEK (EOFs 3-4) SFEK (EOFs 3-6) SFEK (EOFs 3-5)
b
Assimilation RMSE / No assimilation RMSE
Model without assimilation RMSE (m/s)
Tide gauge
SFEK (EOFs 1-2) SFEK (EOFs 1-4) SFEK (EOFs 3-4) SFEK (EOFs 3-6) SFEK (EOFs 3-5)
c
Assimilation RMSE / No assimilation RMSE
Model without assimilation RMSE (m/s)
Tide gauge
SFEK (EOFs 1-2) SFEK (EOFs 1-4) SFEK (EOFs 3-4) SFEK (EOFs 3-6) SFEK (EOFs 3-5)

Fig. 6.