Implementation and validation of a SFEK data assimilation application for an hydrodynamic model of the Tagus Estuary

Ângela Canas*[a], Aires dos Santos[a] and Paulo Leitão[b]

Introduction

Estuaries are important areas both in ecological and economic terms, namely because of their high biological productivity and as their banks are preferred areas for human populations and economic activity settings. Because of economic and social pressure, estuaries are currently in a worldwide scale being affected by water quality problems.

The European Water Framework Directive (2000/60/EC) stresses the estuarine vulnerability to human pressure and, in the context of surface water, advocates the monitorization of the water quality, using not only measurements analysis but also operational modelling.

The Tagus Estuary (38º44'N, 9º08'W), one of the largest in west coastal Europe, is located near the portuguese capital Lisbon in the most populated area of Portugal and particularly sensible to water pollution. It is the object of an water quality pre-operational model (Riflet *et al.*, 2007) explored by Instituto Superior Técnico (Technical University of Lisbon) producing useful results for several water and waste water management entities in the Lisbon area.

This model is presently under improvement. One of the major difficulties experienced is the water level prediction, very degraded in the Tagus river channel areas where grid resolution is unsufficient to resolve the bathymetry, affecting negativelly local hydrodynamics and water quality results. Data assimilation of water level *in situ* measurements provides an interesting solution for this problem (Mourre *et al.*, 2006) while avoiding the expensive increase of grid resolution.

In this work is presented the implementation and validation of a first data assimilation application for an hydrodynamic model of the Tagus Estuary, paving the way for the use data assimilation tests of real measurements at this estuarine site.

Methods

Hydrodynamic model

The hydrodynamic model used in this work is an improved version of the 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 with a primitive equations hydrodynamic module with hydrostatic and Boussinesq aproximations. It consists in a two barotrophic 2D domains: a larger domain containing the Portuguese and Galician coast providing tidal boundary conditions for a smaller domain containing the Tagus Estuary and adjacent coastal area. The model is forced with tidal solution, spatially variable wind fields and average flows for the Tagus River, Sorraia and Trancão rivers.

Data assimilation

Given the characteristics of the hydrodynamic model and the study area, the data assimilation method chosen is the SFEK filter, the fixed correction base version of the Singular Evolutive Extended Kalman (SEEK) filter (Pham *et al.*, 1998; Brasseur *et al.*, 1999), a suboptimal scheme of the Kalman filter which considers a reduced dimension of the model solution error covariance matrix, informed by the dominant Empirical Orthogonal Functions (EOF) of model estimate error covariance.

The filter is implemented in Mohid Water Modelling System in a specific new module, allowing the future expansion to other data assimilation schemes. For practical reasons the generation of initial covariance structure is in a specially designed preprocessing tool, allowing its use also in general EOF analysis studies and the reconstruction of analysed data with the calculated EOFs.

Validation

The validation of the application consists in the evaluation of the performance of the data assimilation scheme in the framework of a *twin test* (e.g. Pham *et al.*, 1998). As the *true model* is considered the hydrodynamic model application described before and the *wrong model* is obtained by the perturbation of the mean sea level imposed at the larger domain boundary with a 0.1m standard deviation white noise, producing a non-biased model error.

A set of assimilation trials are performed considering as state variables the water level and velocity meridional and zonal components of velocity in every grid cell for a two months period, where the *true model* results for water level at 5 tide gauges locations inside and outside the estuary are assimilated every 6 hours in the *wrong model* nested domain state using a SFEK filter, with the initial covariance structure based on *wrong model* historical error or on *wrong model* historical state. These trials consist in a sensibility analysis of the forgetting factor value and the use of objective analysis at the beginning of simulation. It is considered a EOF correction basis composed of the 4 dominant EOFs.

A first part of the work consists in the validation of EOF analysis made for *twin model*s nested domain without assimilation results, sampled every 6 hours for 2 months, contrasting the original data used for EOF analysis and the respective reconstructed data in terms of variability. The

[a] *Instituto Superior Técnico, Av. Rovisco Pais, nº 1, 1049-001 Lisbon, Lisbon, Portugal. Fax: +351 218419423; Tel:+351 218419435 ; E-mail: angela.maretec@ist.utl.pt*
[b] *Hidromod, Av. Manuel da Maia, nº 36, 3º Esq., 1000-201 Lisbon, Lisbon, Portugal. Fax: +351 214211272; Tel: +351 214211373; E-mail: paulo.chambel@hidromod.com*

second part of the work consists in the assessement of results of the assimilation trials, contrasting the *true model*, *wrong model* without assimilation and the *wrong model* with assimilation and assessing filter performance in measurement locations and other 6 tide gauges locations.

Results and Discussion

The EOF analysis of the twin models historical results reveals that the dominant 5 EOFs represent about 95% and 85% of the total variance of model state, respectively for the *true model* and *wrong model*. The 2 dominant EOFs seem to be linked to tide, as semidiurnal periodic oscillation is visible in expansion coefficients. The high representability of a small set of EOFs highlight the simple dynamics of the 2D model, while the discrepancy in explained variance in both models should the result of an increase of complexity in dynamics due to the perturbation imposed on mean sea level. Contrast between the original and reconstructed states points that the explained variance is non uniform in state variables, suggesting the presence of noise accumulation/generation areas in model domain.

As for the *wrong model* historical error fields, the EOF analysis finds a similar explained variance for the 5 dominant EOFs to the one of *wrong model* historical fields but the tidal influence is no longer visible and velocity fields seem to be better described relatively to water level field, possibly the result of normalization in EOF analysis.

The assimilation trials performed using the *wrong model* error covariance structure for the forgetting factor values of 0.5, 0.75 and 1.0 (perfect model assumption) for the first 15 days of simulation resulted in improvements in the range 2% - 45% of centered RMSE in the water level values both at measurements or non measurement locations, which are more significative with smaller forgetting factor value. Contrastingly, the velocity fields are generally degraded by assimilation in all trials except the one with forgetting factor 1. A detail inspection of model results reveals that assimilation in most locations improves velocity results in the first 5 days of simulation but after begins to produce bad results. This is though to result from highly variable representativity of the fixed correction basis in SFEK scheme, error accumulation in domain and also from decrease covariance of model prediction error since the beginning of simulation, as the situation is dependent on forgetting factor value. Objective analysis has, comparatively to forgetting factor, a smaller effect in model predition error reduction.

Considering the initial covariance structure derived from the historical states of *wrong model* instead of the historical *wrong model* error provides similar results in qualitative terms, although quantitatively the water level improvements and velocity degradations tend to be more expressive. This behaviour seems to be the result of the higher initial covariance obtained from historical states, increasing the filter sensibility to innovation change, and the worse capability of the correction EOF basis in describing error behaviour, penalizing the velocity field.

Conclusions

The SFEK data assimilation scheme was implemented in Mohid Water Modelling System code sucessfully. The preprocessing tool for EOF analysis was validated and was found to be very useful in assessing patterns of model variability for the 2D hydrodynamic model for the Tagus Estuary in a simple *twin test* case.

A SFEK filter using a correcting basis of 4 EOFs, derived from model error covariance or model state covariance, and assimilating water level measurements has been capable to provide useful correction of water level predictions of the Tagus Estuary 2D hydrodynamic model in most of the tide gauge locations studied in a period of 15 days simulation. However, velocity prediction is generally degraded, possibly due to weak representability of the EOF basis and the accumulation of errors in the model domain over the simulation period.

These results suggest that a SEEK filter will have to be developed if an operative filter is to be used to improve predictions of the Tagus Estuary hydrodynamic model.

Acknowledgements

Special thanks to Ibrahim Hoteit, which provided EOF analysis code that was used as the basis for this work, and to George Triantafyllou and his colleagues from the HCMR for useful discussion on research results and state of the art on EOF analysis and SEEK filtering.

This work was financed by a grant from Fundação para a Ciência e Tecnologia, for the Ph.D. studies of Ângela Canas, under contract SFRH / BD / 14185 / 2003 and by the European Commission through the project INSEA (http://www.insea.info/) under contract SST4-CT-2005-012336.

References

Brasseur, P., J. Ballabrera-Poy, J. Verron, 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.

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

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

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

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

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

Riflet, G., P. Leitão, R. Fernandes, R. Neves, 2007. Assessing the quality of a pre-operational model for the portuguese coast. *Geophysical Research Abstrac*ts, 9, EGU2007-A-09979.