International Conference on Hydroinformatics

2014

Variational Data Assimilation With Telemac. Proof Of Concept For Model State Correction On The Berre Lagoon 3D-Model

Sophie Ricci

Andrea Piacentini

Anthony Weaver

Riadh Ata

Nicole Goutal

11th International Conference on Hydroinformatics
HIC 2014, New York City, USA

VARIATIONAL DATA ASSIMILATION WITH TELEMAC. PROOF OF CONCEPT FOR MODEL STATE CORRECTION ON THE BERRE LAGOON 3D-MODEL

SOPHIE RICCI (1), ANDREA PIACENTINI (1), ANTHONY WEAVER (1), RIADH ATA (2), NICOLE GOUTAL (2)

(1): Sciences de l'Univers au CERFACS, URA1875, Toulouse, France

(2): Laboratoire National d'Hydraulique et Environnement, Chatou, France

TELEMAC is a component of the open-source integrated suite of solvers TELEMAC-MASCARET for use in the field of free-surface flow that solves the Reynolds Averaged Navier-Stokes equations. Generally speaking, uncertainties in the model formulation itself due to simplified physics and also in the input fields to the model such as the boundary conditions, initial conditions and hydraulic parameters translate into errors in the simulated hydraulic variables. In spite of significant advances in numerical schemes, description of geographical data (topography, bathymetry) and environmental conditions (hydrologycal and meteorological fields), the representation of the true state of a system as well as its forecasted state remains imperfect and some of these limits can be overcome combining observations with simulation via data assimilation techniques. This paper presents the implementation of a 3D-Var FGAT variational data assimilation algorithm as a proof of concept for improving TELEMAC simulations and forecast. The demonstration is made on the Berre lagoon application with TELEMAC-3D: the salinity state is sequentially corrected assimilating in-situ salinity measurements.

I- INTRODUCTION

The TELEMAC-3D (T3D) software developed by Electricité de France (EDF R&D) is used to represent the hydrodynamics flow in the Berre lagoon as well as the stratified salinity and temperature fields. The Berre lagoon is a receptacle of 1000 Mm3 where salty water from the Mediterranean Sea, through the Canal de Caronte, meets fresh water discharged by the hydroelectric plant at Saint-Chamas and by natural tributaries. The proper representation of the stratified salinity and temperature fields as well as the 3D currents was identified as a valuable research objective with direct applications for both electricity production and ecological matters. Preliminary studies on the calibrated 11 vertical plans T3D model (Durand and Razafindrakoto, 2011[1]) were carried out to quantify the difference between a reference simulation and XBT (eXpendable BathyThermograph) sensors observations on a test period. A positive salinity drift was detected, probably resulting from the under-estimation of the fresh water inputs and the over estimation of evapotranspiration. While an artificial fresh water input at Caronte allows to partly correct the mean salinity, the temporal intra- and inter- annual estimation of this correction is difficult to estimate and the hydraulics state simulated by the model remains imperfect: the currents tend to be under-estimated and the difference between the simulated salinity and the observations can reach up to several g/l. The present work

illustrates how a variational data assimilation (DA) algorithm was implemented to sequentially correct the salinity state combining measurements with the numerical simulation output and thus improving predictions. While the merits of DA methods have been largely demonstrated in hydrology (Weerts and El Serafy, 2006 [2], Moradkhani et *al.*, 2005 [3]) and global and coastal ocean fields (Weaver et *al.*, 2003 [4], Moore et *al.*, 2011 [5]) they are yet to be fully taken advantage of in lakes and lagoon hydrodynamical modelling systems.
In this paper, the TELEMAC model is presented in Section 1. Then the 3D-Var FGAT algorithm is presented in Section 2. Preliminary results from the 3D-Var FGAT system are presented in Section 3 showing the merits of var DA for hydrodynamics as a proof of concept of the Berre case.

II- TELEMAC-3D MODEL

TELEMAC-3D is a component of the open-source integrated suite of solvers TELEMAC-MASCARET for use in the field of free-surface flow. TELEMAC-MASCARET is managed by a consortium of core organizations and used for dimensioning and impact studies. The numerical code TELEMAC-3D solves the Reynolds Averaged Navier-Stokes equations (RANS) in their hydrostatic and non-hydrostatic versions (Hervouet, 2007 [6]):

$$\nabla.\vec{U} = 0 \quad (1)$$

$$\frac{\partial U}{\partial t} + \vec{u}.\nabla U = -\frac{1}{\rho}\frac{\partial P}{\partial x} + \nabla.\left(\nu\vec{\nabla}U\right) + f_x \quad (2)$$

$$\frac{\partial V}{\partial t} + \vec{u}.\nabla V = -\frac{1}{\rho}\frac{\partial P}{\partial y} + \nabla.\left(\nu\vec{\nabla}V\right) + f_y \quad (3)$$

$$\frac{\partial W}{\partial t} + \vec{u}.\nabla W = -\frac{1}{\rho}\frac{\partial P}{\partial z} + \nabla.\left(\nu\vec{\nabla}W\right) + f_z \quad (4)$$

where $\vec{u} = (U,V,W)$ is the velocity vector, and *P* is the pressure, ν the diffusion coefficient that includes molecular and turbulent viscosities and $\vec{f} = (f_x, f_y, f_z)$ is the forcing vector. The numerical code uses algorithms based on finite-element or finite-volume methods where space is discretised in the form of an unstructured grid of prismatic elements obtained by extruding the 2D triangular mesh in the vertical direction. Boundary conditions are described by impermeable conditions at the bottom, free surface and various lateral boundaries such as walls, dykes, river banks or beaches. Eqs. 1-4 are coupled with an advection-diffusion equation for the tracer dynamics:

$$\frac{\partial C}{\partial t} + \vec{u}.\nabla C = S_C + \nabla.\left(\left(\kappa_C + \frac{\nu}{\sigma}\right)\vec{\nabla}C\right) \quad (5)$$

where C is the tracer field, κ_C is the tracer diffusion coefficient, σ = 0.72 is the Schmidt number (called Prandlt number if the tracer is the temperature), S_c t is the tracer source term. The turbulence is handled, for instance, through a k-ε model associated to Boussinesq hypothesis to account for buoyancy.
Errors in the description of the mesh (density, refinement, prism geometric features), the boundary conditions, the turbulence model, the numerical schemes translate into errors in TELEMAC-3D outputs. Once identified, these uncertainties should be quantified, classified and then reduced. In spite of constant advances in numerical methods, computational resources, mesh definition and input data acquisition (especially topography and bathymetry), a significant part of these uncertainties remain and DA appears as a complementary way of improving simulations.

III- VARIATIONAL DATA ASSIMILATION ALGORITHM

An incremental variational DA algorithm is implemented to correct the initial salinity state at the beginning of a time window over which several observations are available. The 3D-Var FGAT (3D variational method with First Guess at Appropriate Time) formulation is used here; it provides a correction to the initial state and the assumption is made that the dynamics of a perturbation to the salinity state can reasonably be approximated by a persistent model (this comes down to simplifying the classical Incremental 4D-Var to the 3D-Var FGAT and significantly reduces the cost of the DA algorithm). The DA algorithm formulates the difference between the numerical model outputs and the observations over the assimilation window $[t_0,t_T]$ as a function of the initial state of the system $\mathbf{x}=\mathbf{x}(t_0)$, also called the control vector. This cost function is regularized by a background term that penalizes the distance to the background state $\mathbf{x}^b$ which is the model estimate of this initial condition (simulated prior to the assimilation). The statistics of its errors are described by the background error covariance matrix **B**. The observation vector $\mathbf{y}^o$ is a vector of size N that gathers the observations available in space and time over the assimilation window. The statistics of its errors are described by the observations error covariance matrix **R** (assumed to be diagonal in the following). The inverse of the background and observation covariance error matrices define the weighting matrices of the quadratic terms in the cost function J

$$\mathrm{J}(\mathbf{x})=\frac{1}{2}(\mathbf{x}-\mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b)+\frac{1}{2}(G(\mathbf{x})-\mathbf{y}^o)^T\mathbf{R}^{-1}(G(\mathbf{x})-\mathbf{y}^o). \quad (6)$$

where G is the generalized observation operator that allows to compare the model trajectory (computed from the background initial state) to the observations over the assimilation window. G is the composition of the T3D dynamics model $M_{t0,tT}$, then mapped to the observation space using the observation operator H (extracting the simulated information at the observation time and space).The initial state that minimizes the cost function is called the analysis $\mathbf{x}^a$. It can be integrated forward in time to produce a forecast beyond the assimilation time window.

The minimization of the non-quadratic cost function J is usually achieved as a sequence of minimizations of approximated quadratic functions where a local linearization of the generalized observation operator is used. This is the incremental formulation that aims at identifying a correction $\delta\mathbf{x}$ to the background state such that $\mathbf{x}^a = \mathbf{x}^b + \delta\mathbf{x}^a$. The generalized observation operator is linearized, noted **G**, around a reference state usually chosen as the background, that requires the formulation of the tangent-linear $\mathbf{M}_{t0,tT}$, of the nonlinear model $M_{t0,tT}$ approximated by the identity operator in the 3D-Var FGAT, and **H** of the observation operator H. The incremental cost function J_{inc} reads:

$$\mathrm{J}_{inc}(\delta\mathbf{x})=\frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x}+\frac{1}{2}(\mathbf{G}\delta\mathbf{x}-\mathbf{d})^T\mathbf{R}^{-1}(\mathbf{G}\delta\mathbf{x}-\mathbf{d}) \quad (7)$$

where **d** is the innovation vector that denotes the difference between the observation vector and the background trajectory integrated from the background state with $M_{t0,tT}$. The optimal increment is obtained by setting the gradient of J_{inc} to zero but since the size of the problem does not allow for an algebric formulation of the solution, an approximate solution is found by iteratively solving a linear system, in the dual space (Nocedal and Wright, 2006 [7]) spanned by vectors of the size of the observation vector (much smaller than the control space). The details and the implementation of this algorithm called RBCG (Restricted B-preconditioned Conjugate Gradient) are given in Gurol et *al.* (2013) [8].

IV- RESULTS FOR A SINGLE OBSERVATION EXPERIMENT

In order to validate the 3D-Var FGAT algorithm, a single observation is assimilated over an assimilation window with a diagonal **B** matrix (meaning that the spatial correlation of the model state errors are not represented so far). Here, the DA procedure comes down to computing a weighted average between the simulated and the observed salinity at the observation point, where the background and observation weights are given by the background and observation error variances in **B**. When both variances are set arbitrarily to 0.25 psu^2, and given that the observed salinity is equal to 26.434100 psu while the simulated salinity is 26.6386 psu, the analysis increment given by the RBCG minimization is $\delta\mathbf{x}$ =-0.1022501 psu, which is, as expected half of the difference between the simulation and the observation. It should be noted that the RBCG converges in one single iteration and it was also verified that when the variances are modified, the analysis changes accordingly: it remains close to the background when the observation error variance increases and gets closer to the observation when the background error variance increases. This validation test only allows for a local correction of the salinity state while it is expected that each observation should translate into a spatially coherent correction to the salinity state. To do so, the extra-diagonal terms in **B** (that describe the spatial correlations of the model state errors) are represented using an implicitly formulated 3D diffusion equation. This method and its implementation with an implicit scheme are presented in Mirouze and Weaver (2010) [9]. In the present framework, the correlation functions are described applying the diffusion operator with different diffusion coefficients in the vertical and horizontal directions that relate to the vertical and horizontal correlation length scales denoted L_v and L_h.

Figure 1 presents the salinity increment on the horizontal plane when L_h=200m (Fig. 1-a) and vertical plane when L_v=0.5m (Fig. 1-b) for an observation assimilated at -5m at observation point SA1. It should be noted that a negative salinity correction is applied and spread from the observation point to the surrounding points delimited by the correlation length scales. When observations at SA1, at -5m, are assimilated every 15 minutes with an assimilation window of 1 hour, a correction to the initial salinity state is computed from the 3D-Var FGAT algorithm using 4 observations in the minimization process. This analysis is cycled over 24 hours and the results are presented in Fig. 2- for L_h=200 m and L_v=2 m. For each panel, the salinity is represented at SA1, at different vertical depths that correspond to the 5 XBT sensors positions, as a function of time over 24 h, at observation times only. The observations are represented in red, the T3D Free Run (no assimilation) is plotted in black, the background (for the current cycle) is plotted is green and the analysis is plotted in blue. First, it should be noted that at 5 m deep (where the observations are assimilated), the salinity is significantly improved and brought closer to the observations (the background and observation errors variances are set to 4 psu^2). The difference between the analysis and the observation is systematically reduced at the beginning of the assimilation window when the correction is applied, then the model is integrated over 1 hour and deviates from the observations. The analysis salinity value at the end of the assimilation window is the background initial salinity state for the following cycle. The 1-hour integration of the background state can thus be considered as a 1-hour forecast following the 1-hour assimilation window. It should then be noted that the 3D-Var FGAT algorithm improves the salinity over the assimilation period as well as over a forecast period of 1 hour. It should be noted that the salinity is corrected over the entire water column. Whether this correction improves or not the salinity depends on the coherence between the spatial correlation of the errors in the simulated salinity field and the correlation length-scales prescribed in **B**. A major result here is that the impact of the IC correction is limited in time and that a sequential assimilation of frequent enough observations is necessary, especially under the 3D-Var FGAT assumption.

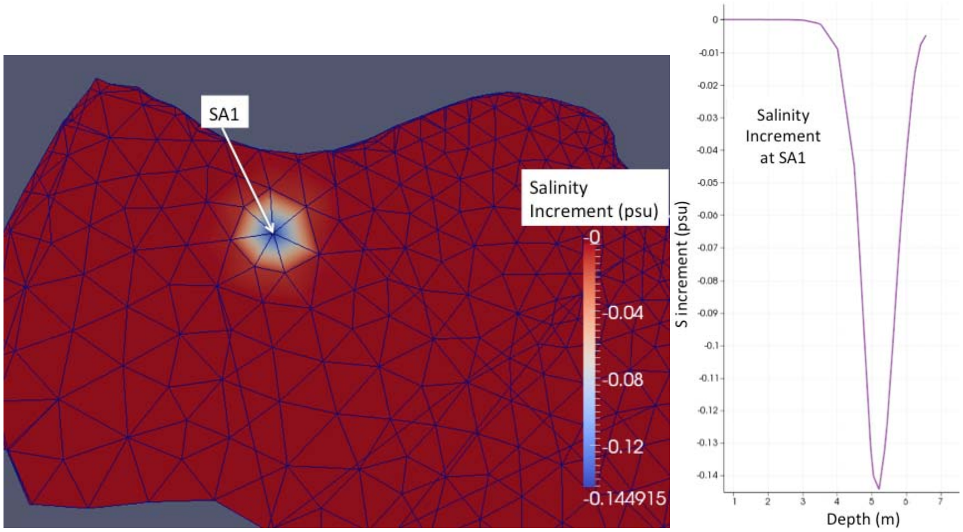

Figure 1: Salinity increment (in psu) for single observation assimilation in a- horizontal plane and b- vertical plane for a single observation assimilation at SA1 at -5m.

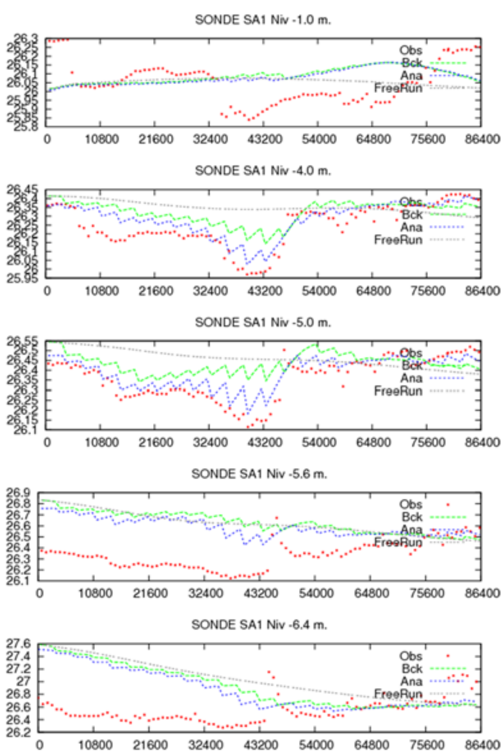

Figure 2: Salinity (y-axis in psu) at the 5 observation points at SA1 along the vertical for L_h=200m, L_v=2m over 24 assimilation cycles of 1 hour (x-axis in seconds). Only the observations at 5m deep are assimilated. Observations are plotted in red, model run without assimilation is plotted in black, DA analysis is plotted in blue and DA forecast in plotted in green. The DA time window is 1h.

V- CONCLUSIONS

A 3D-Var FGAT algorithm was implemented to improve the salinity field description assimilating in-situ salinity observations with a T3D model for the Berre lagoon. A major perspective for this work is to increase the forecast lead-time. To do so, the DA algorithm should also correct hydrological and meteorological forcing fields. This cannot be achieved under the 3D-Var FGAT assumption and requires the use of the tangent linear code of TELEMAC with respect to boundary conditions. Using the tangent linear and adjoint codes of TELEMAC with respect to the model state will also allow for the implementation of a classical 4D-Var algorithm that takes into account the dynamics of a correction to the IC and provides an increment that is in good agreement with the flow dynamics. Ongoing works with University of Aachen aim at automatically differentiating the TELEMAC numerical code using the NAG-compiler.

This proof of concept for DA with TELEMAC opens the way to a wide range of applications for EDF in hydraulics, for instance in the field of water quality, water resources management, and flood forecasting.

REFERENCES

[1] N. Durand, E. Razafindrakoto. "Synthèses et conclusions des études sur la modélisation numérique TELEMAC-3D de l'évolution des courants, de la salinité et de la température dans l'étang de Berre." , Note Interne EDF-LNHE, H-P74-2011-02289-FR, 2011.

[2] Weerts, A.H., and El Serafy, G.Y.H.: Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, Water Resour. Res., 42, W09403, doi:10.1029/2005WR004093., 2006.

[3] Moradkhani, H., Sorooshian, S, Gupta, H.V., and Houser, P.-R.: Dual state-parameter estimation of hydrological models using ensemble Kalman filter, Advances in Water Resources, 28, 135--147, 2005.

[4] Weaver A.T., Vialard J., Anderson D.L.T.: Three- and four- dimensional variational assimilation with a general circulation model of the tropical Pacific ocean. Part 1: formulation, internal diagnostics and consistency checks. Mon. Weather Rev. 131: 1360–-1378, 2003.

[5] Moore, A.M., Arango H.G., Broquet G., Powell B.S., Zavala-Garay J., Weaver A.T.: The Regional Ocean modelling System (ROMS) 4-dimensional variational data assimilation systems. Part I: System overview and formulation. Prog. Oceanogr. 91: 34–49, 2011.

[6] Hervouet JM, Hydrodynamics of free Surface flows. Wiley 2007.

[7] Nocedal, J., Wright, S.J.: Numerical Optimization. Series in Operations Research, Springer Verlag: Heidelberg, Berlin, New York, 2006

[8] Gurol, S., A. T. Weaver, A. M. Moore, A. Piacentini, H. G. Arango and S. Gratton. " B-Preconditioned Minimization Algorithms for Variational Data Assimilation with the Dual Formulation", Q. J. R. Meteorol. Soc. , in print.

[9] Mirouze, I., Weaver, A.T.: Representation of correlation functions in variational assimilation using an implicit diffusion operator. Q. J. R. Meteorol. Soc., 136, 421--1443, 2010.