Ocean Modelling 29 (2009) 85–93

Contents lists available at ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

Inverse estimation of open boundary conditions in tidal channels

Issam S. Strub *, Julie Percelay, Mark T. Stacey, Alexandre M. Bayen

Department of Civil and Environmental Engineering, University of California, 604 Davis Hall, Berkeley, CA 94720-1710, USA

ARTICLE INFO

Article history:
Received 7 October 2008
Received in revised form 23 January 2009
Accepted 4 March 2009
Available online 14 March 2009

Keywords:
Tidal channels
Boundary conditions estimation and prediction
Quadratic programming
Lagrangian data assimilation

ABSTRACT

This article presents a novel algorithm for the estimation of open boundary conditions in river systems where tidal forcing is present. This algorithm uses a linearisation of the model equations. With the help of a linear discretisation scheme, the article presents a quadratic programming formulation of the estimation problem in which the control variables are the coefficients of the dominant tidal modes. This method is implemented for a scenario in which only Lagrangian observations from drifters are available to measure flow quantities. The performance of the algorithm is evaluated using numerical experiments and comparing estimation results with boundary conditions from a river located in the Sacramento Delta. The sensitivity of the algorithm to the number of modes estimated and its predictive capabilities are also assessed.

1. Introduction

The study of tidal forcing in bays and estuaries is crucial to the monitoring of water quality issues in these areas (Fischer et al., 1979; Yi et al., 1989). Numerically, these flows can be simulated using a full three-dimensional approach, or the one- or two-dimensional shallow water equations (Chadwick et al., 2004). In any case, the numerical model needs to be thoroughly calibrated with parameters depending on geometry and flow features. In addition, a knowledge of open boundary conditions is required for the model to perform adequately. Such boundary conditions can be inferred from measurements made with fixed sensors placed at the boundaries of the domain of interest or as is developed in this article, using drifters that are circulating in the system under consideration; it is also possible to extend the computational domain beyond the boundaries of interest but knowledge of the flow properties at a boundary is still required. Data assimilation techniques can be used to incorporate these observations into the model; they originated several decades ago in meteorology and oceanography (Anthes, 1974; Le Dimet and Talagrand, 1986; Sasaki et al., 1955). A number of different methods have been introduced over the years, which include variational methods (Navon, 1998, 1985), ensemble Kalman filtering (Evensen, 2007; Kuznetsov et al., 2003), optimal statistical interpolation (Molcard et al., 2003), or the nudging method (Ishikawa et al., 1996; Paniconi et al., 2003). The topic of inverse estimation of boundary conditions in meteorology and oceanography has been studied over the past few decades starting with the pioneering work (Sasaki et al., 1955) and more recently (Navon, 1985; Shulman et al., 1998; Yang and Hamrick, 2005; Yi et al., 1989; Zhang et al., 2003). In this article, a novel quadratic programming based variational data assimilation algorithm will be applied to the estimation of open boundary conditions for tidal channel flows.

The need for boundary condition estimation is driven in our case by operational requirements: we are interested in deploying drifter fleets in specific areas of the Sacramento-San Joaquin Delta for which current sensing infrastructure and modelling capabilities are insufficient. The Sacramento-San Joaquin Delta is at risk of extensive levee failures, which could be caused by earthquakes, flood or human activities. In such an event, the water quality of the Delta will be negatively affected, due to sediment suspension, salinity intrusion, and potentially agricultural contaminants. It is critical that transport models be available for use following such an event, even in the case where the entire geometry of the system has been altered. Unfortunately, existing models of the Sacramento-San Joaquin Delta rely heavily on historical data sets for calibration. These models would be of limited utility if the system were radically altered, as would occur in the case of extensive levee failures. The use of rapidly deployed Lagrangian drifters and the estimation of boundary conditions is motivated by this need: we aim to develop a sensing-modelling system that is capable of predicting regional flows and transport in the Delta in real-time without dependence on historical data. The timescale of interest for this analysis is on the timescale of days: we would like to be able to project transport patterns forwards in time by a few days based on only a few days of data.

* Corresponding author. Tel.: +33 1415 694 3709.
E-mail address: strub@ce.berkeley.edu (I.S. Strub).

1463-5003/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ocemod.2009.03.002

The immediate objective is to estimate open boundary conditions for a model of flows in tidal channels of the Sacramento-San Joaquin Delta. To this end, we develop variational data assimilation techniques in which a cost function measuring the norm of the difference between observations gathered by drifters and model predictions is minimised, subject to constraints given by the discretised model equations; the control variables are the coefficients of the dominant tidal modes present in the upstream boundary condition. A similar variational approach for Lagrangian data assimilation in rivers applied to bottom topography estimation was presented in (Honnorat et al., 2009). The estimation of boundary conditions in the Sacramento-San Joaquin Delta using fixed sensors and data reconciliation techniques was presented in (Wu et al., 2009) which presents some similarities with our work, but specifically tackles the problem of Eulerian measurements. While a two-dimensional model could be chosen for the identification of the boundary conditions, its computational cost is high and can be avoided by using simpler models such as a one-dimensional model. The choice we make here is motivated by the desire to have a rapid, robust estimate of boundary conditions that can be used in real-time flow simulations. A somewhat similar approach was used in (Gejadze and Monnier, 2007) as a one-dimensional shallow water model was combined with a local two-dimensional model through optimal control methods. In (Gejadze and Copeland, 2006; Gejadze et al., 2006), a method for the estimation of open boundary conditions for the Navier–Stokes equations with a free surface from fixed depth and velocity measurements is developed. All the articles mentioned above use the adjoint method combined with a quasi Newton solver to solve a variational problem. The adjoint method has the main drawback of a high computational cost as 50–100 iterations are usually required which corresponds to 100– 200 numerical resolutions of the direct and adjoint equations. Additionally, the nonlinearity of the problem means that convergence to a global minimum is not guaranteed. In this article, we use a quadratic programming approach which eliminates both issues. Indeed, since the cost function is quadratic, this problem can be solved as a quadratic program provided the constraints are linear. Another novelty in this article is the use of drifters for the estimation of boundary conditions while the other articles previously mentioned rely on fixed measurement stations. Note also that unlike adjoint based optimisation, the quadratic programming technique used in the present article does not require the definition and resolution of an adjoint (backward) problem.

This article is organised as follows: we start by presenting the basics of tidal channels and we state the problem solved in the article; we then use standard data assimilation terminology to develop the estimation algorithm; numerical experiment settings are then described before the results for estimation and prediction of tidal flow are presented and analysed.

2. Model

2.1. Tidal channels

Tidal channels are bodies of water in which periodic changes in the water level and velocity field occur under the influence of tides, along the dominant tidal frequencies such as the K1 tide generated by the Sun, the M2 tide generated by the Moon or the MK3 shallow water tide created by the nonlinear interaction between the K1 and M2 tides resulting from the bottom friction. In channel networks, tidal trapping (Fischer et al., 1979) can occur due to phase effects, with the flow directions changing at different times in the deeper and shallower branches; for example, the flow may be ebbing in a shallow channel while a nearby channel is already flooding. The implications for transport and dispersion are profound and tidal trapping is likely the dominant dispersion observed in such systems.

The flow in such a channel can be modelled using two-dimensional depth-averaged shallow water equations such as (Vreugdenhil, 1994):

$$\frac{\partial h}{\partial t} + \vec{u} \cdot \nabla h + h \nabla \cdot \vec{u} = 0 \tag{1}$$

$$\frac{\partial u}{\partial t} + \vec{u} \cdot \nabla u = -g \frac{\partial h}{\partial x} - g \frac{\partial b}{\partial x} + f_x \tag{2}$$

$$\frac{\partial v}{\partial t} + \vec{u} \cdot \nabla v = -g \frac{\partial h}{\partial y} - g \frac{\partial b}{\partial y} + f_y \tag{3}$$

where (u, v, h) represent the velocity components and water depth, with $\vec{u} = (u, v)$, b the bottom elevation, g the acceleration of gravity and (f_x, f_y) the viscous forces term.

For the simulations presented in this article, we will use a specific realisation of (1)–(3) incorporating friction forces and a specific turbulence model:

$$\frac{\partial h}{\partial t} + \vec{u} \cdot \nabla h + h \nabla \cdot \vec{u} = 0 \tag{4}$$

$$\frac{\partial u}{\partial t} + \vec{u} \cdot \nabla u = -g \frac{\partial h}{\partial x} - g \frac{\partial b}{\partial x} + F_x + \frac{1}{h} \nabla \cdot (h \nu_t \nabla u) \tag{5}$$

$$\frac{\partial v}{\partial t} + \vec{u} \cdot \nabla v = -g \frac{\partial h}{\partial y} - g \frac{\partial b}{\partial y} + F_y + \frac{1}{h} \nabla \cdot (h \nu_t \nabla v) \tag{6}$$

The friction forces are given by the following Manning law:

$$F_x = -\frac{1}{\cos \alpha} \frac{g m^2}{h^{4/3}} u \sqrt{u^2 + v^2} \tag{7}$$

$$F_y = -\frac{1}{\cos \alpha} \frac{g m^2}{h^{4/3}} v \sqrt{u^2 + v^2} \tag{8}$$

where h is the total depth of water, $\vec{u} = (u, v)$ is the velocity field, g is the acceleration of gravity, b is the bottom elevation, ν_t is the coefficient of turbulence diffusion obeying the so-called k-epsilon model (see (Rastogi and Rodi, 1978) for more details), $\alpha = \alpha(x, y)$ is the slope of the bottom, and m is the Manning coefficient (usually denoted by n). In the present case, the Manning coefficient is chosen to be constant in time and space and equal to 0.02, corresponding to a highly frictional bottom. Finally, t is time and x, y are horizontal space coordinates.

2.2. Problem description

An accurate simulation of tidal trapping relies on a prediction of the phase differences which is usually estimated according to historical data sets. In this article, our goal is to estimate open boundary conditions, more precisely the amplitudes and phases corresponding to the main four to eight tidal modes using velocity and position measurements provided by a number of drifters which are released on a portion of the Sacramento River located in the Sacramento-San Joaquin Delta in California. While the model used for the data assimilation problem is two-dimensional, the estimation of the boundary conditions is performed on a one-dimensional model which yields substantial benefits with respect to the size of the problem solved and therefore the computational time. The use of a one-dimensional estimation of the open boundary conditions is justified by the rapid lateral adjustment time of the free surface (< 1 min); further, the velocity profile can be reconstructed using interpolation from the average velocity in the cross-section. For example, in the case of the software package TELEMAC used in the article, the velocity profile is reconstructed with the assumption that the velocity is proportional to the square root of the depth along the cross-section (Hervouet and Haren, 2002).

The algorithm proposed by this article consists in minimising the difference between measured velocity at the location of the drifter, and velocity estimated by the model. The constraints are the constitutive equations for the water, discretised in a linear manner. The decision variables, i.e. the variables with respect to which the optimisation is run are the upstream boundary conditions in the form of the coefficients of the dominant tidal modes. The problem thus consists in identifying the proper forcing at the upstream boundary of the domain, using Lagrangian sensing only.

In the next section, we present the cost function to be minimised using the standard formulation of variational data assimilation, the one-dimensional shallow water equations and linearised numerical scheme used as constraints.

3. Estimation algorithm

3.1. Notations

We start this section by defining the variables which will appear in the following. We employ the traditional notations of variational data assimilation, set forth in (Ide et al., 1997).

- X_i : vector of discretised state variables, namely the velocity component u and the water height h for each mesh point at a time instant t_i,
- Y_i : vector of observed variables, namely the velocity components u and (potentially) the water height h for some mesh points at a time instant t_i,
- $\mathbf{R}_i$: covariance matrix of the observation error (difference between the value of the state variables and observed variables at each mesh point), taken equal to the identity in this article,
- H_i : $H_i = h_i^o \circ h^I$ is the observation operator; the operator h_i^o projects the space into the observation subspace. The operator h^I is the interpolation function. In general H_i is nonlinear although we manage to use a linear operator in our case by using the a posteriori knowledge of the position of the measurements, therefore encoded as a time dependent observation matrix.

We minimise the following cost function:

$$\mathscr{J}^o(U_{\text{up}}) = \sum_{i=0}^{n}(Y_i^o - H_i[X_i])^T \mathbf{R}_i^{-1}(Y_i^o - H_i[X_i]) \tag{9}$$

with respect to the boundary conditions U_{up} upstream, using as constraints the one-dimensional linear discretised shallow water equations presented next.

3.2. One-dimensional numerical scheme

We choose to use a one-dimensional linearised shallow water model to estimate the upstream boundary condition, and then use this as input in a nonlinear two-dimensional solver which will generate a simulation of the flow taking into account the full geometry of the river. Compared to implementing a two-dimensional shallow water model, this method presents the added advantage of greatly reducing the size of the quadratic program (by a factor at least 20 in our case) therefore allowing us to run the algorithm over extended time horizons (50 h in the present case) for an acceptable computational cost.

The nonlinear one-dimensional shallow water equations can be written as:

$$\frac{\partial H}{\partial t} + \frac{\partial HU}{\partial x} = 0 \tag{10}$$

$$\frac{\partial U}{\partial t} + U\frac{\partial U}{\partial x} + g\frac{\partial H}{\partial x} + g\frac{\partial b}{\partial x} + g\frac{U|U|}{C^2 H} = 0 \tag{11}$$

where H is the water depth, b the one-dimensionally averaged bottom elevation, U the velocity, and C the Chézy friction coefficient related to the Manning coefficient (presented earlier for the two-dimensional model) by: $C = \frac{H^{\frac{2}{3}}}{m}$.

We replace the friction term from Eq. (11) by an empirical drag coefficient $r = g\frac{|U_0|}{C^2 H_0}$ and linearise Eqs. (10) and (11) in the case of small perturbations around a nominal flow (U_0, H_0), with $U = u + U_0$, $H = h + H_0$. We then neglect the derivatives of the nominal variables (see, for example, Wesseling, 2001) to obtain:

$$\frac{\partial h}{\partial t} + H_0\frac{\partial u}{\partial x} + U_0\frac{\partial h}{\partial x} = 0 \tag{12}$$

$$\frac{\partial u}{\partial t} + U_0\frac{\partial u}{\partial x} + g\frac{\partial h}{\partial x} + g\frac{\partial b}{\partial x} + ru = 0 \tag{13}$$

The Preissmann scheme (see Preissmann, 1961; Venutelli, 2002) is applied to these equations yielding:

$$\begin{aligned}
h_{j+1}^{i+1} - h_{j+1}^{i} + h_j^{i+1} - h_j^i &= -H_0\frac{\Delta t}{\Delta x}(u_{j+1}^{i+1} - u_j^{i+1} + u_{j+1}^i - u_j^i) \\
&\quad - U_0\frac{\Delta t}{\Delta x}(h_{j+1}^{i+1} - h_j^{i+1} + h_{j+1}^i - h_j^i) \\
u_{j+1}^{i+1} - u_{j+1}^{i} + u_j^{i+1} - u_j^i &= -U_0\frac{\Delta t}{\Delta x}(u_{j+1}^{i+1} - u_j^{i+1} + u_{j+1}^i - u_j^i) \\
&\quad - g\frac{\Delta t}{\Delta x}(h_{j+1}^{i+1} - h_j^{i+1} + h_{j+1}^i - h_j^i) \\
&\quad - ru_j^i - g\frac{2\Delta t}{\Delta x}(b_{j+1} - b_j)
\end{aligned}$$

where j denotes the space step and i the time step. This is an implicit scheme which allows us to choose fewer time steps than in the explicit case. This scheme was used as constraints in a quadratic program minimising the L^2 norm (defined by $||x||_2 = (\sum_{i=1}^N x_i^2)^{\frac{1}{2}}$ where $x = (x_1, \ldots, x_N)$ is a vector with N components) of the difference between the observed velocity and water height and the variables obtained through the Preissmann scheme. This algorithm will be used in the following to estimate boundary conditions for tidal flows, running a simulation for 50 h and attempting to obtain up to eight modes.

3.3. Modelling assumptions

We can now pose the estimation problem as a quadratic optimisation problem. Using simulated drifter trajectories, we estimate the amplitude and phase associated with four to eight dominant frequencies and the signal thus obtained is compared to the original flow. The coefficients representing the amplitudes of the main modes are estimated using a data assimilation algorithm which consists in minimising the cost function (9) with respect to the constraints given by the Preissmann scheme detailed earlier. The boundary condition is specified as:

$$u_{\text{up}}(t) = a_0 + \sum_{k=1}^{N_{\text{modes}}} a_k\cos(\omega_k t) + b_k\sin(\omega_k t) \tag{14}$$

where a_0 is the mean flow and $\omega_k = \frac{2\pi}{T_k}$ is the frequency associated to one of the periods identified in the spectral analysis of the data. The coefficients $(a_k), 0 \leqslant k \leqslant N_{\text{modes}}$ and $(b_k), 1 \leqslant k \leqslant N_{\text{modes}}$, fully characterise the boundary conditions and the data assimilation problem is posed in a way which makes them decision variables.

3.4. Optimisation program

As mentioned above, the control variables are the coefficients $a_0, a_1, \ldots, a_{N_{\text{modes}}}$ and $b_1, \ldots, b_{N_{\text{modes}}}$ in Eq. (14). The discretised dynamics of the flow are encoded in the form of linear constraints through the Preissmann scheme developed above. While the numerical scheme is implicit, i.e. it can be written in the form $EZ_{i+1} = AZ_i + Bu_i$ (where Bu_i comes from the bottom elevation term in the one-dimensional shallow water equations), the quadratic program can incorporate them at no further cost or complication. Note that this is a major advantage of the quadratic programming approach over techniques such as Kalman filtering, which would require an explicit inversion of the Preissmann scheme, so it can be expressed as a linear time invariant system $Z_{i+1} = AZ_i + Bu_i$. Implicit constraints of the present form can be incorporated in quadratic programs at no further cost. In addition, no initial guess is required, which is an advantage of our method over some other method (in particular specific implementations of adjoint based formulation of data assimilation problems).

We concatenate the vectors Z_i for all i into a single vector called X, and abbreviate the dynamics constraints by $CX = b$, where this equation encodes the discretised flow equations. The search space $\{(a_k, b_k), 1 \leqslant k \leqslant N_{\text{modes}}\}$ from which all other quantities depend is allowed to evolve in a set of feasible initial conditions dictated by flow physics. Because the (u_j^i, h_j^i) cannot take arbitrary values, the space in which X evolves can also be restricted to increase the speed of convergence of the method. These two constraints are encoded in the form of an inequality, $GX \leqslant h$. Finally, using the previous variable definitions, Eq. (9) can be written as:

$$\begin{array}{ll} \text{minimise} & J(X_{\text{up}}) = \frac{1}{2}X^T PX + q^T X + r \\ \text{subject to} & GX \leqslant h \\ & AX = b \end{array}$$

The constraints are thus as follows:

- (u_j^i, h_j^i) must verify the Preissmann scheme at all space and time steps.
- State variables have to satisfy boundary and initial conditions (X_{up} represents the boundary conditions upstream and is related to the state of the system X through the linearised Preissmann scheme), encoded by the linear equality constraint.
- Inequality constraints need to keep (u_j^i, h_j^i) in a realistic set (in our case positivity of the water height and bound on the absolute value of the velocity $|u| < 10$ m/s).

Note that the computational cost of solving the quadratic program above is very reasonable, because the vector X is a concatenated vector of u_j^i and h_j^i resulting from the discretisation of Eqs. (12) and (13), which is a contribution of this article. The approach consisting in linearising Eqs. (4)–(6) and performing the identification with these linearised discrete constraints can also be encoded as a quadratic program, but by nature of the grid, would not necessarily result in a tractable approach.

4. Numerical experiment setting

A standard benchmark for the validation of data assimilation algorithms is a numerical experiment which consist in comparing the state of the system with and without data assimilation. For this, a *forward* simulation is run from time t_0 to time T using a two-dimensional nonlinear shallow water model and, yielding the so-called *true state* at every time instant between t_0 and T. At a chosen time t_1, drifters are released from the upstream end and their trajectories are simulated using a Runge–Kutta method and the vector field provided by the nonlinear shallow water forward simulation; then, a data assimilation process is started and the estimated boundary conditions are computed and compared with the true boundary conditions. In the present case, the true state of the river and the position of the drifters are generated using the two-dimensional nonlinear shallow water solver TELEMAC (Hervouet and Haren, 2002) and we realise several experiments with and without tidal flow reversal in order to test the influence of the number of modes estimated on the accuracy of the algorithm. The boundary conditions for the two-dimensional forward simulation are computed using the Delta Simulation Model II (DSM2) (Anderson and Mierzwa, 2002) (Fig. 3), a one-dimensional model of the Sacramento-San Joaquin Delta that provides discharge and surface elevation at various locations every hour.

4.1. Description of the experimental protocol

While the algorithm described in this article can be applied to any river, access to data such as an accurate bathymetry and boundary conditions is required. This led to the implementation of these algorithms on a body of water located in the Sacramento-San Joaquin Delta (Fig. 1) near the town of Walnut Grove; the geographical proximity of this area is also helpful in the perspective of conducting field experiments such as releasing drifters or installing sensors to measure boundary conditions. The Delta is a web of channels and reclaimed islands formed by the confluence of the Sacramento River, itself fed by the northern Sierra Nevada runoff, and the San Joaquin River, which mingle with smaller tributaries to create a 700 mile labyrinth of sloughs and waterways surrounding 57 reclaimed islands (see, for example, Lund et al., 2007; McClurg, 2000, for a detailed presentation of the Delta). The Delta is of great importance to the state of California for a number of reasons; in particular, it represents a source of drinking water for more than 20 million Californians and most of California's farmland relies on water tributary to the Delta for irrigation. Water quality, including concentrations of salt, suspended sediment and contaminants, is a major concern as it affects directly the potability of drinking water supplies, the productivity of farmland and the viability of organisms in the aquatic ecosystems. While there are Delta-scale transport models that have been calibrated with historical data, these models are unlikely to be valid if the Delta is radically altered by an extreme event (earthquake or flood). Instead, we seek a modelling solution that relies only on real-time data collected from rapidly deployable drifters integrated into a regional model of flows and transport that can predict conditions over the coming days. Performing this analysis on a subregion in the Delta allows us to focus on regions of interest for Delta management, but creates a technical challenge through the presence of tidal open boundaries. Our goal here is to establish a fast, robust method for estimating open boundary conditions in a tidal channel network.

In the map in Fig. 1, the area influenced by tidal forcing goes beyond the deployment area in Walnut Grove and affects channels within several miles of the portion of the Sacramento River where the study is conducted. In this article, we will consider a section of the Sacramento River, located between the Delta Cross Channel and the Georgiana Slough. The bathymetry is provided by United States Geological Survey (USGS) as shown in Fig. 2.

4.2. Tidal flow in the Sacramento River

Eight years of hourly data coming from a DSM2 simulation of the Delta were available for the study. The spectral analysis of the DSM2 data for the month of October 2006 shows eight dominant frequencies (see Fig. 4). A similar work is done with DSM2 data from March 2006, a situation in which there is no tidal reversal (see bottom Fig. 3). The spectral analysis of the flow during the month of March 2006 shows a highly dominant nominal flow with

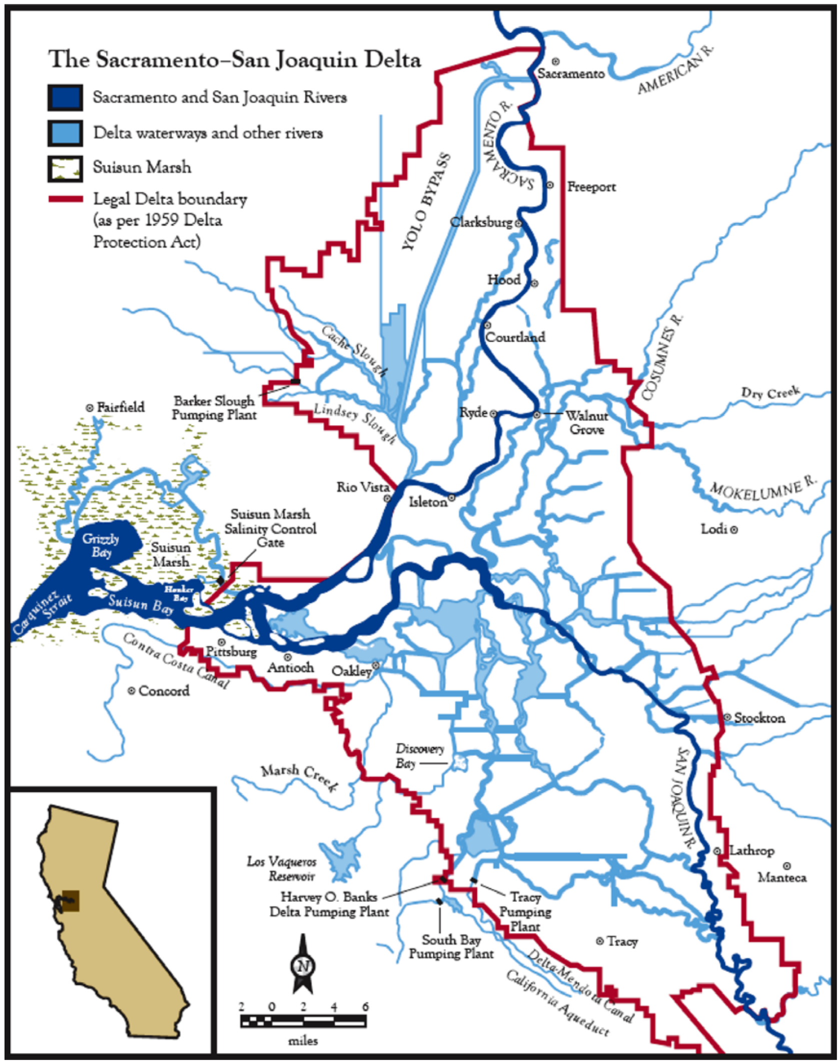

Fig. 1. The Sacramento-San Joaquin Delta (from Lund et al., 2007). The portion of the Sacramento River considered for the experiment is located near the town of Walnut Grove, in the centre of the figure.

minimal amplitudes for the other modes. These results are summarised in Table 1.

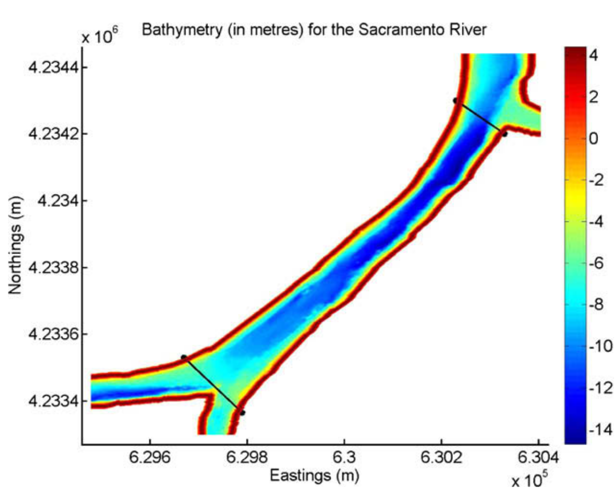

Fig. 2. Bathymetry in the Sacramento River (m) in NGVD (*National Geodetic Vertical Datum*) datum used for the present study. The bathymetry on this 930 m section of the Sacramento River goes from −14 m in the deepest part to +4 m on the river banks. The boundaries of the domain are delimited by solid black lines.

4.3. Initial and boundary conditions

The boundary and initial conditions of Eqs. (4)–(6) are given by

$$u(x,y,t)|_{\partial\Omega_{\text{land}}} = 0, \quad v(x,y,t)|_{\partial\Omega_{\text{land}}} = 0 \tag{15}$$

$$(u(x,y,t),v(x,y,t))|_{\partial\Omega_{\text{upstream}}} = f(x,y,t) \tag{16}$$

$$\eta(x,y,t)|_{\partial\Omega_{\text{downstream}}} = g(x,y,t) \tag{17}$$

$$u(x,y,0) = u_0, \quad v(x,y,0) = v_0, \quad h(x,y,0) = h_0 \tag{18}$$

where $\partial\Omega$ represents the boundaries of our computational domain and f, g are known functions (in the present case obtained through a DSM2 simulation as explained earlier).

4.4. Lagrangian drifters

In addition to the Eqs. (4)–(6) describing the flow dynamics, we model the deployed drifters as passive Lagrangian tracers. Hence, the drifters move with the local fluid velocity, obeying the following equations:

$$\frac{dx_D(t)}{dt} = u[x_D(t), y_D(t), t], \tag{19}$$

$$\frac{dy_D(t)}{dt} = v[x_D(t), y_D(t), t], \tag{20}$$

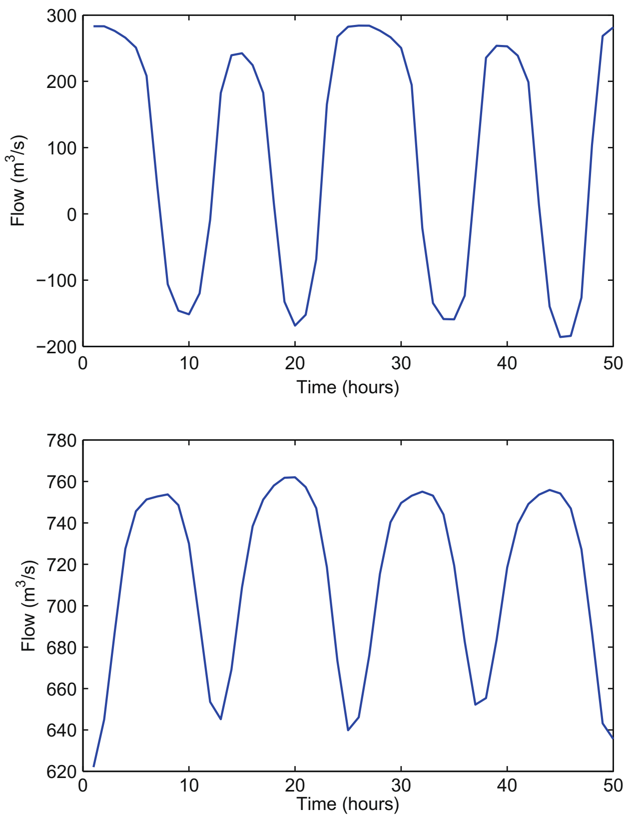

Fig. 3. Hourly evolution of the flow over a 50 h period in October 2006 (top) and March 2006 (bottom) used as boundary condition in the data assimilation algorithm.

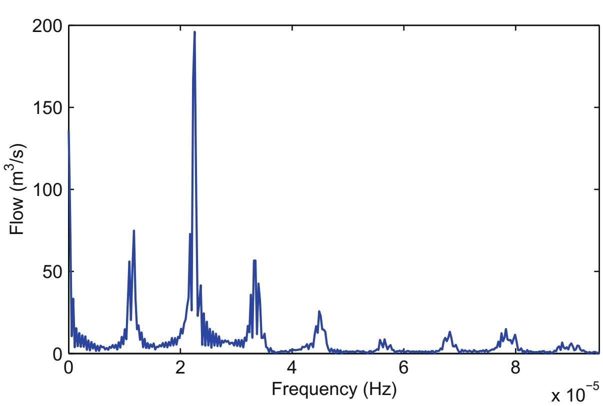

Fig. 4. Spectral decomposition of the flow in the Sacramento River using DSM2 data for the month of October 2006.

Table 1
Amplitudes of the main tidal modes in October 2006 (tidal inversion) and March 2006 (no inversion).

Tide	Amplitude (m^3/s) (October 2006)	Amplitude (m^3/s) (March 2006)
Mean flow	68	714
O1	56.2	23.2
K1	74.9	171
N2	72.9	21.2
M2	196	88.1
MK3	56.7	70.1
M4	25.7	42.7
M6	13.3	3.26
M8	14.9	2.46

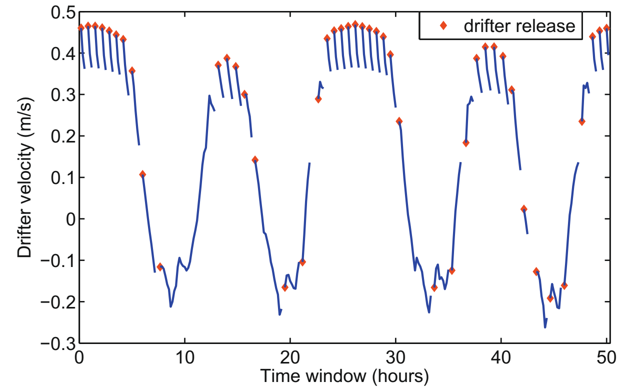

Fig. 5. Velocity of the drifter during the simulation in the case of tidal inversion (October 2006). The red dots indicate the release of a new drifter.

with the initial conditions

$$x_D(0) = x_{D,0}, \quad y_D(0) = y_{D,0}. \tag{21}$$

Eqs. (19) and (20) have to be understood for each x_D and y_D corresponding to a specific drifter.

The simulation runs for two and a half hours before the drifters' release so that a stable state is reached. The experiment is then run during 50 h with drifters being released.

A simulation is run using the two-dimensional nonlinear shallow water solver TELEMAC during a 50 h period in October 2006 during which tidal flow reversal occurs and during another 50 h period in March 2006 in which flow is unidirectional (due to freshwater flow). The release of drifters is simulated during this period (see Fig. 5) with a drifter being released from one boundary or the other depending on the direction of the flow since a tidal reversal occurs eight times during the simulation (drifter releases are indicated by red dots on the graph). The nearly vertical lines at the top of each sinusoid correspond to the velocity of the drifter along the river, the velocity decreasing as the drifter floats down the channel. The positions and velocities of the drifters are recorded every 10 min and used in the data assimilation algorithm. Note that for these numerical experiments, the only observations from the channel flow are those given by the drifters. A drifter is released from the upstream boundary whenever the previous drifter exits the domain so that there is one drifter in the model domain at all time during the simulation. The run time for a 50 h numerical experiment is approximately 1h:30.

5. Results and analysis

5.1. Estimation of flow

Either four or eight modes are estimated using the data assimilation algorithm from which the flow is reconstructed and compared to the original DSM2 data. The relative *Root Mean Square error* (RMS) given by

$$E = \left(\frac{\sum_{i=1}^{50} |u^t_{\mathrm{up}}(i) - u^a_{\mathrm{up}}(i)|^2}{\sum_{i=1}^{50} |u^t_{\mathrm{up}}(i)|^2} \right)^{\frac{1}{2}}$$

where $u^t_{\mathrm{up}}(i)$ and $u^a_{\mathrm{up}}(i)$ represent, respectively, the true one-dimensional boundary conditions (DSM2 data) and the assimilated one-dimensional boundary conditions at time step i. The quadratic program is coded via the modelling language AMPL (Fourer et al., 2003) and solved with CPLEX (Ilog, 2008), a large scale quadratic programming solver, using eight iterations of a barrier method in less than 5 s on a desktop computer (Fig. 6). A one-dimensional grid

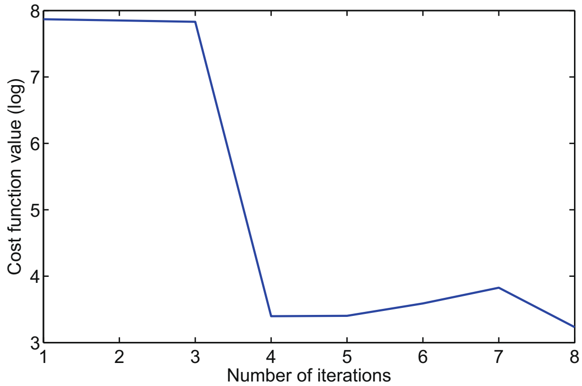

Fig. 6. Evolution of the cost function with respect to the number of iterations of the barrier method.

with 37 cells is used for the implementation of the one-dimensional shallow water equations as constraints and a 10 min time step is chosen. Depending on whether four or eight modes are used, the relative RMS error of the flow obtained is, respectively, 15.6% and 7.2%. The comparative plots of the reconstructed flow and the true flow are given in Fig. 7.

A numerical experiment similar to the one presented above realised in March 2006 yields a relative RMS error of 7.2% and 4.6% for four and eight modes, respectively (see Fig. 8). The errors are smaller than in the October 2006 case which may be explained by the fact that the linear model used in the data assimilation algorithm is a better fit when there is no flow reversal. Indeed, during the flow reversal around slack water, nonlinear effects are likely to be more pronounced given that the overall flow is very low.

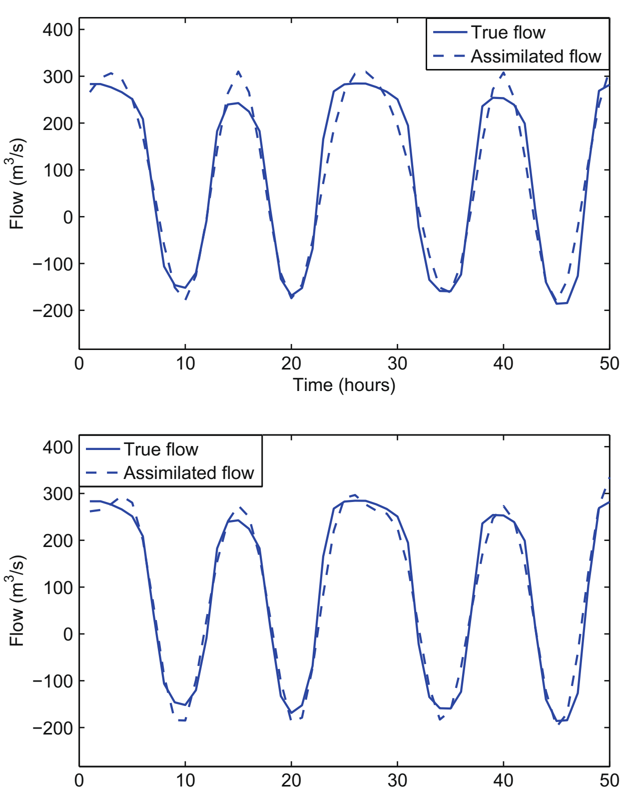

Fig. 7. Hourly evolution of the true and assimilated flow using four modes (top) or eight modes (bottom) over a 50 h period with tidal inversion (October 2006).

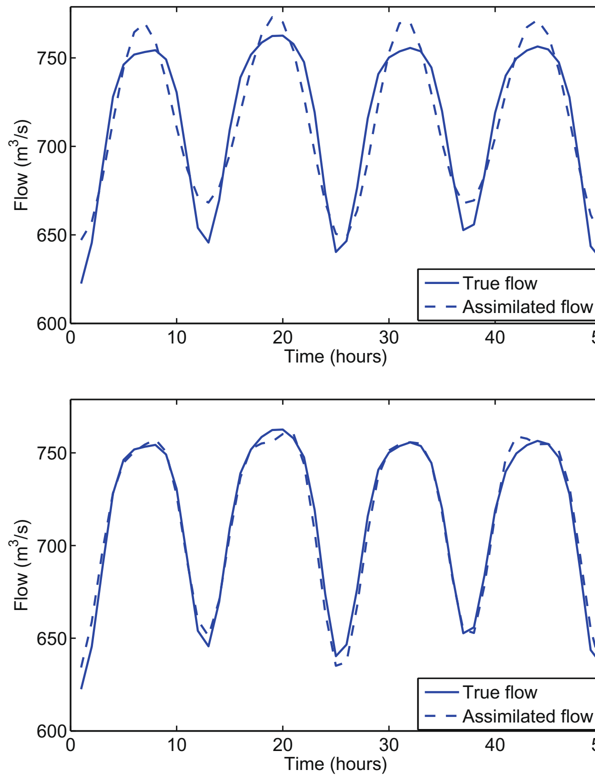

Fig. 8. Hourly evolution of the true and assimilated flow using four modes (top) or eight modes (bottom) over a 50 h period with no tidal inversion (March 2006).

5.2. Prediction of flow evolution

The next experiment aims to evaluate the predictive capabilities of this data assimilation algorithm. Our goal is to be able to make projections forwards in time based on limited drifter data. Because of lower frequency variability, most notably the spring-neap cycle and variations in freshwater flow here, the projections will necessarily deteriorate with time. Here, we aim to evaluate the window of time over which realistic projections can be made. Specifically, we will use 50 h of drifter data to estimate the amplitudes corresponding to the dominant frequencies. Then, we try to reconstruct the signal for a 1 week period and compare it with the DSM2 data (Table 2).

For the first data set from October 2006 which presents a flow reversal, the relative RMS error is 39.7% and 37.3% for four and eight modes, respectively (see Fig. 9). The increase of the flow in the last few days of the week considered is not predicted by either reconstructed flow hence the similar relative RMS error.

In the case from March 2006, the relative RMS error is 30.6% and 28.4% for four and eight modes, respectively (see Fig. 10). Once again, the decrease of the flow in the second half of the period considered is not predicted by either of the reconstructed flows which explains the almost identical error for four and eight modes signals.

Table 2
Relative RMS error of the estimated flow for four and eight modes in October 2006 (tidal inversion) and March 2006 (no inversion).

Modes	Error (%) (October 2006)	Error (%) (March 2006)
Mean, K1, M2, MK3, M4	15.6	7.8
Mean, K1, M2, MK3, M4, M6, M8, O1, N2	7.2	4.6

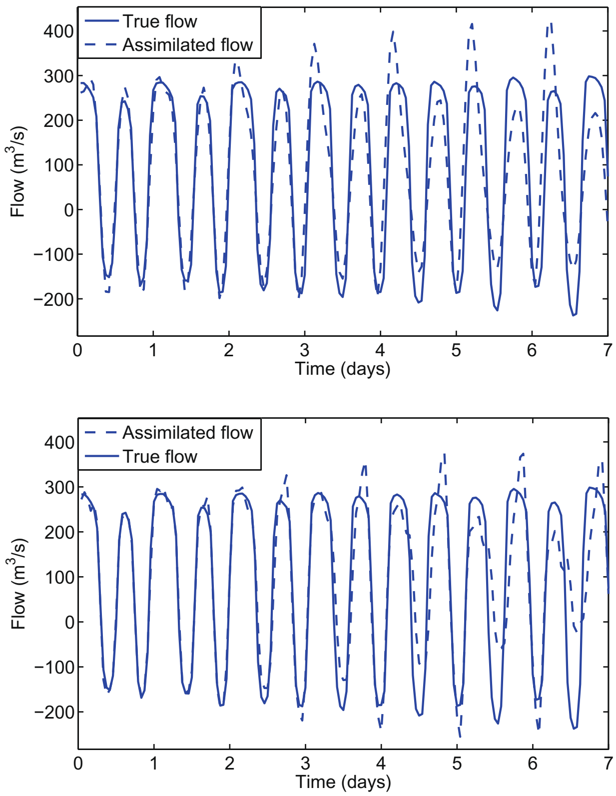

Fig. 9. Hourly evolution of the true and assimilated flow using four modes (top) or eight modes (bottom) over a 1 week period with tidal inversion (October 2006).

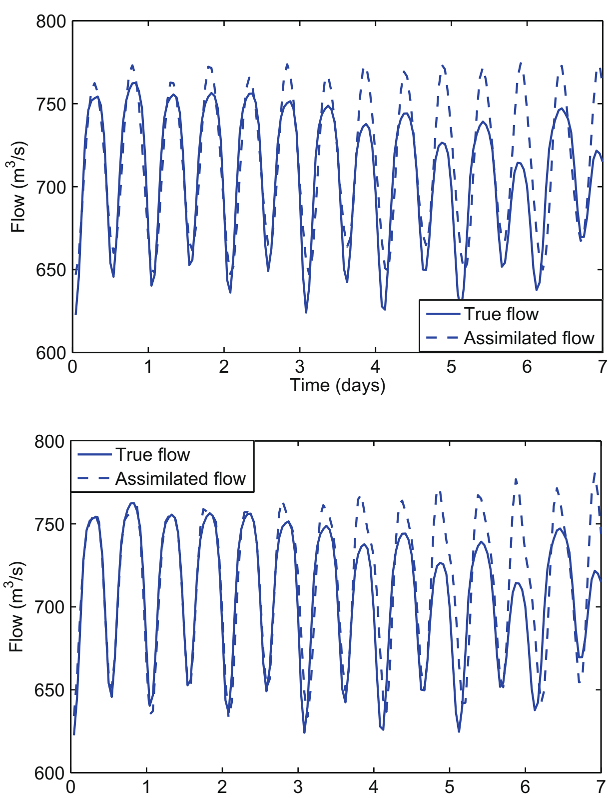

Fig. 10. Hourly evolution of the true and assimilated flow using four modes (top) or eight modes (bottom) over a 1 week hour period with no tidal inversion (March 2006).

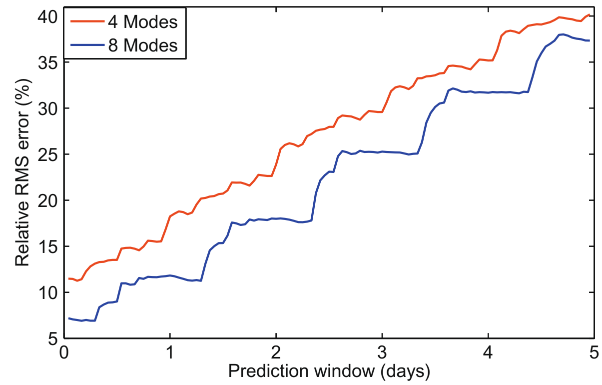

Fig. 11. Evolution of the relative RMS error after of the observation period with four and eight modes considered.

When the eight modes are considered, the results presented in Fig. 11 show that adding the larger lunar elliptic semidiurnal tide N2 does however reduce the error at the beginning of the prediction period (Fig. 11) but the prediction error remains above 30% over a 1 week period. It would appear that events other than tidal forcing (such as reservoir release or local precipitation) influence the flow in the Sacramento Delta; this is confirmed by the comparison between the spectra of the flow weeks before and 3 weeks after the experiment period (Fig. 12) which shows significant differences in the amplitude of the mean flow (238 m^3/s before and 160 m^3/s after the experiment) and the M2 tide (176 m^3/s before and 215 m^3/s after) thereby limiting the predictive capability of an algorithm based on identifying tidal components over a short period of time. In the case without tidal reversal, the amplitudes of the O1, M6, M8, and N2 tides are too small compared to the mean flow and other dominant tides for their inclusion to have a significant effect on the estimation error. The preceding results are summarised in Table 3.

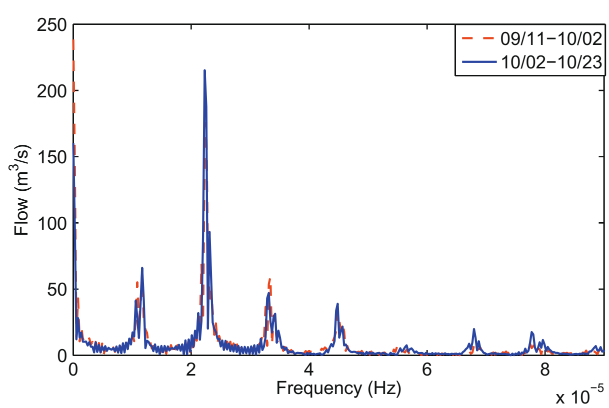

Fig. 12. Comparison of the spectra of the flow in the Sacramento River 3 weeks before the October 2006 experiment and 3 weeks after. Note the significant difference of amplitude in the mean flow and M2 tide.

Table 3
Relative RMS error of the estimated flow for four and eight modes in October 2006 (tidal inversion) and March 2006 (no inversion) after 50 h, 4 days, and 1 week.

Modes	50 h		4 days		1 week	
	(03/06)	(10/06)	(03/06)	(10/06)	(03/06)	
Four modes	15.6%	7.8%	22.6%	15.1%	39.7%	30.6%
Eight modes	7.2%	4.6%	17.9%	11.2%	37.3%	28.4%

6. Conclusion

In this article, a novel data assimilation algorithm was applied to estimate open boundary conditions for a tidal channel in the Sacramento-San Joaquin Delta. Its performance was assessed using numerical experiments. Given that tidal forcing is dominant, the focus was placed on identifying the main frequencies and then estimating the corresponding amplitudes. Using 50 h of DSM2 data, the flow in the Sacramento River was simulated using two-dimensional nonlinear shallow water solver TELEMAC and drifter trajectories were simulated both in a reversing tidal flow and when the flow does not reverse. The velocities of the drifters along the trajectories were used to estimate the amplitude of the main four to eight modes and enabled us to generate a reconstructed flow which is compared to the original DSM2 data. The relative root mean square error varies from 15.6% to 4.6% depending on whether a flow reversal occurs and on the number of modes estimated. This shows that during the period of time that observations are available, the inverse model is able to effectively estimate the harmonic constants and reproduce the flows. It is important to note that we have not constrained the tidal amplitudes or phases in any way: the only tidal information added are the frequencies of the dominant modes. For them, the case we consider here is a stringent one on which there is only one drifter available; most Lagrangian measurements include a number of drifters which would improve the ability to estimate boundary conditions.

Considering projections in time beyond the observation period, the ability of the harmonics to track the true values is somewhat limited. An attempt was made to evaluate the predictive capabilities of the algorithm by comparing the reconstructed flow and DSM2 data over a period of 1 week. The prediction error was found to be between 22.6% and 11.2% after 4 days and 39.7% and 28.4% after 1 week. The relatively high value of the RMS error after 1 week appears to be due to the fact that tidal harmonics are limited in their ability to predict flows in this reach of the Sacramento-San Joaquin Delta, where, in addition to low frequency tidal components, unsteady freshwater flows influence the evolution. Indeed, a comparison of the tidal spectra 3 weeks before and after the experiment reveals significant differences in the amplitudes of the main tidal modes which limit the predictive capability of an algorithm based on identifying tidal components over a short period of time (50 h). Future research directions include studying the optimal placement of the drifters and the optimal release points in order to maximise the accuracy of the estimation.

References

Anderson, J., Mierzwa, M., 2002. An introduction to the Delta Simulation Model II (DSM2) for simulation of hydrodynamics and water quality of the Sacramento-San Joaquin Delta. Delta Modeling Section, Office of State Water Project Planning, California Department of Water Resources, Sacramento, CA.

Anthes, R.A., 1974. Data assimilation and initialization of hurricane prediction models. J. Atmos. Sci. 31, 702–719.

Chadwick, A., Morfett, J., Borthwick, M., 2004. Hydraulics in Civil and Environmental Engineering, fourth ed. Spon Press, London.

Evensen, G., 2007. Data Assimilation: The Ensemble Kalman Filter. Springer-Verlag, Berlin.

Fischer, H.B., List, E.J., Koch, R.C.Y., Imberger, J., Brooks, N.H., 1979. Mixing in Inland and Coastal Waters. Academic Press, New York.

Fourer, R., Gay, D.M., Kernighan, B.W., 2003. AMPL: A Modeling Language for Mathematical Programming, second ed. Thomson Brooks/Cole.

Gejadze, I.Y., Copeland, G.J.M., 2006. Open boundary control problem for Navier–Stokes equations including a free surface: adjoint sensitivity analysis. Comput. Math. Appl. 52 (8–9), 1243–1268.

Gejadze, I.Y., Copeland, G.J.M., Navon, I.M., 2006. Open boundary control problem for Navier–Stokes equations including a free surface: data assimilation. Comput. Math. Appl. 52 (8–9), 1269–1288.

Gejadze, I.Y., Monnier, J., 2007. On a 2D 'zoom' for the 1D shallow water model: coupling and data assimilation. Comput. Methods Appl. Mech. Eng. 196 (45–48), 4628–4643.

Hervouet, J.M., Van Haren, L., 2002. TELEMAC 2D. Version 5.2 – Principle note.

Honnorat, M., Monnier, J., Le Dimet, F.X., 2009. Lagrangian data assimilation for river hydraulics simulations. Comput. Visual Sci.. doi:10.1007/s00791-008-0089-x.

Ide, K., Courtier, P., Ghil, M., Lorenc, A., 1997. Unified notation for data assimilation: operational, sequential and variational. J. Met. Soc. Jpn. 75 (1B), 181–189.

Ilog, 2008. ILOG AMPL CPLEX System Version 11.0 User Guide.

Ishikawa, Y., Awaji, T., Akitomo, K., Qiu, B., 1996. Successive correction of the mean sea surface height by the simultaneous assimilation of drifting buoy and altimetric data. J. Phys. Oceanogr. 26, 2381–2397.

Kuznetsov, L., Ide, K., Jones, C.K.R.T., 2003. A method for assimilation of Lagrangian data. Mon. Weather Rev. 131 (10), 2247–2260.

Le Dimet, F.X., Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus 38A, 97–110.

Lund, J., Hanak, E., Fleenor, W., Howitt, R., Mount, J., Moyle, P., 2007. Envisioning futures for the Sacramento-San Joaquin Delta. Public Policy Institute of California, San Francisco, CA.

McClurg, S., 2000. Layperson's Guide to the Delta. Water Education Foundation, Sacramento, CA.

Molcard, A., Piterbarg, L.I., Griffa, A., Özgökmen, T., Mariano, A., 2003. Assimilation of drifter observations for the reconstruction of the Eulerian circulation field. J. Geophys. Res. 108 (C3), 3056.

Navon, I.M., October 15–17 1985. A review of variational and optimization methods in meteorology. In: Sasaki, Y.K. (Ed.), Proceedings of the International Symposium on Variational Methods in Geosciences. vol. 5 of Developments in Geomathematics, University of Oklahoma, Norman, OK, pp. 29–34.

Navon, I.M., 1998. Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. Dyn. Atmos. Oceans 27, 55–79.

Paniconi, C., Marrocu, M., Putti, M., Verbunt, M., 2003. Newtonian nudging for a Richards equation-based distributed hydrological model. Adv. Water Resour. 26 (2), 161–178.

Preissmann, A., 1961. Propagation des intumescences dans les canaux et rivières. In: 1st Congress of the French Association for Computation. Grenoble.

Rastogi, A.K., Rodi, W., 1978. Predictions of heat and mass transfer in open channels. ASCE J. Hydr. Div. 104 (3), 397–419.

Sasaki, Y., Gu, P., Yan, L., 1955. A fundamental study of the numerical prediction based on the variational principle. J. Meteor. Soc. Jpn. 33, 262–275.

Shulman, I., Lewis, J.K., Blumberg, A.F., Kim, B.N., 1998. Optimized boundary conditions and data assimilation with application to the M2 tide in the Yellow Sea. J. Atmos. Ocean. Tech. 15, 1066–1071.

Venutelli, M., 2002. Stability and accuracy of weighted four-point implicit finite difference schemes for open channel flow. J. Hydraul. Eng. 128 (3), 281–288.

Vreugdenhil, C.B., 1994. Numerical Methods for Shallow Water Flow. Kluwer Academic Publishers, Dodrecht.

Wesseling, P., 2001. Principles of Computational Fluid Dynamics. Springer, Berlin, New York.

Wu, Q., Litrico, X., Bayen, A., 2009. Data reconciliation of an open channel flow network using modal decomposition, Adv. Water Resour. 32, 193–204.

Yang, Z., Hamrick, J.M., 2005. Optimal control of salinity boundary condition in a tidal model using a variational inverse method. Estuar. Coast. Shelf Sci. 62, 13–24.

Yi, J., Gu, P., Yan, L., 1989. Physical and mathematical models of pollution dispersion in the Yangtze estuary. China Ocean Eng. 3 (1), 95–106.

Zhang, A., Wei, E., Parker, B.B., 2003. Optimal estimation of tidal open boundary conditions using predicted tides and adjoint data assimilation technique. Cont. Shelf Res. 23 (11), 1055–1070.