Submitted, 48th IEEE Conference on Decision & Control
Shanghai, China, December 16-18, 2009

Inverse Modeling for Open Boundary Conditions in Channel Network

Qingfang Wu*, Mohammad Rafiee†, Andrew Tinka* and Alexandre M. Bayen*
*Department of Civil and Environmental Engineering, UC Berkeley, CA, USA
{qingfangwu,tinka,bayen}@berkeley.edu
†Department of Mechanical Engineering, University of California, Berkeley, CA,USA
{rafiee}@berkeley.edu

***Abstract*— An inverse modeling problem for systems governed by first-order, hyperbolic partial differential equations subject to periodic forcing is investigated. The problem is described as a PDE constrained optimization problem with the objective of minimizing the norm of the difference between the observed inputs and the model outputs. After linearizing and discretizing the governing equations using an implicit discretization scheme, linear constraints are constructed which leads to a quadratic programming formulation of the estimation problem. The utility of the proposed approach is illustrated by considering a channel network in the Sacramento San-Joaquin Delta in California, subjected to tidal forcing. The dynamics of the hydraulic system are modeled by the linearized Saint-Venant equations. The available data are the drifter positions as they circulat in the experiment domain. The inverse modeling problem is to estimate open boundary conditions by considering a finite number of dominant tidal modes. It is shown that the proposed method gives an accurate estimation of the flow variables at the boundaries and intermediate locations within the system.**

I. INTRODUCTION

The Sacramento San Joaquin Delta in California is experiencing drastic declines in fresh water resources, while the water demand

in California is increasing. Large-scale numerical flow models, such as DSM2 [3] and REALM [3], sponsored by the California Department of Water Resources, have been used as crucial water resources management tools, providing information about tidal forcing and salinity transport in the bays and channels of the Delta. A number of factors affect the performance of these state-of-the-art models, such as parameter calibration, mesh generation, and choice of numerical solver. More importantly, the performance of the model largely relies on the determination of open boundary conditions.

Traditionally, these open boundary conditions are obtained either via Eulerian observations near the boundaries, such as tidal gauge data, or through satellite data retrieval. Unfortunately, these measurements for large watershed have their intrinsic limitations, such as small coverage and sparse sampling [13]. Furthermore, installed Eulerian sensors are proven to have many failures, such as broken gauges, process leaks, sensor drifts, improper use of measuring devices, and other random sources [1].

In the last two decades, the techniques using surface and subsurface Lagrangian buoys have been significantly developed. The Lagrangian data, in particular those collected from surface drifters, provide information about the flow which can be used to describe flow advections and eddy dispersions. For this reason, Lagrangian data have been highly valued and extensively used in numerous meteorologic and oceanic models [2] [18]. The Lagrangian data assimilation problem can be approached in different ways, including variational method [15], ensemble Kalman filtering [21] [8], optimal statistical interpolation [12], and Newtonian relaxation [16]. In this article, we present a quadratic programming (QP) based method introduced by [22] [20] to determine the open boundary conditions in tidal channel networks by using Lagrangian measurements of the flow. More specifically, we derive the velocity field in a channel network solely from the position information collected by drifters. The proposed method is to minimize the norm of the difference between the drifter observations and model velocity predictions, subject to the constraints given by discretized linear equations. One of the major contributions of this article is to pose the problem of estimating the open boundary conditions of a channel network as a quadratic programming problem by minimizing a quadratic cost function and posing the constraints as a set of a linear form of equalities and inequalities. The proposed quadratic program can be solved using fast and robust algorithms, , and it is capable of providing reliable open boundary conditions for any flow simulations.

To verify the proposed QP method, we investigate a distributed network of channels, subject to quasi-periodic tidal forcing, in the Sacramento-San Joaquin Delta. The main obstacle of applying a linear model in the channel networks is the well-known tidal trapping phenomenon [9]. The trapping mechanism makes water elevation and velocity not in phase, which results in the flow dispersion and eddy diffusion at the junctions of channels. The drifter trajectory at these junctions, because of the turbulent mixing processes, usually display a stochastic spaghetti-like shape, which is indicative of slow currents. Another contribution of the article is to successfully assimilate this chaotic drifter data, and, as a result, the channel network system can be adequately simulated using one-dimensional Linear Saint Venant Model.

The article is organized as follows: Section II introduces the mathematical flow model in open channels. A linear Saint Venant model in a single river reach is derived after linearizing and discretizing the governing equations. A linear channel network model is constructed considering the flow compatibility condition at the junctions of channels. Section III formulates the quadratic programming method withby applying standard data assimilation techniques. Section IV describes the experiment protocol in the Sacramento San Joaquin Delta, and displays the drifter trajectory generated by nonlinear Shallow Water Model simulations. Section V shows the application of the inverse modeling procedure. The effectiveness of the method is substantiated by correlating the model estimations with field data at selected locations in the network. Section VI summarizes our studies and presents the scope of future work.

II. Hydrodynamic Model in Tidal Channel Network

A. Saint-Venant Model

The Saint-Venant equations are non-linear hyperbolic PDEs that describe the dynamics of one-dimensional flow in open-channel hydraulic systems [5],[7]. For a rectangular cross-section, these equations are given by:

$$Y_t + (VY)_x = 0 \tag{1}$$
$$V_t + VV_x + gY_x + g\left(S_f - S_b\right) = 0 \tag{2}$$

for $(x,t) \in (0,L)\times\Re^+$, where L is the river reach m, $V(x,t)$ is the average velocity (m/s) across cross-section $A(x,t) = T(x)\cdot Y(x,t)$, $Y(x,t)$ is the water-depth (m), $T(x)$ is the free surface width (m) for rectangular cross-section, $S_f(x,t)$ is the friction slope (m/m), S_b is the bed slope m/m, g is the gravitational acceleration (m/s^2). The boundary conditions are $V(0,t) = V_0(t)$ and $Y(L,t) = Y_0(t)$. The initial conditions are given by $V(x,0)$ and $Y(x,0)$ for $x \in [0,L]$. The friction slope is empirically modeled by the Manning-Strickler's formula

$$S_f = \frac{V^2 n^2 (T+2Y)^{4/3}}{(TY)^{4/3}} \tag{3}$$

where n is the Manning's roughness coefficient $(sm^{-1/3})$. Under the proper constant boundary conditions, equations (1), (2) admit a *steady state solution*. Let the flow variables corresponding to the steady state condition be denoted by $V_0(x)$, $Y_0(x)$ etc. where $x \in [0,L]$. The steady state equations are given by

$$V_0(x)\frac{dY_0(x)}{dx} + Y_0(x)\frac{dV_0(x)}{dx} = 0 \tag{4}$$
$$\frac{dY_0(x)}{dx} = \frac{S_b - S_{f0}}{1 - F_0(x)^2} \tag{5}$$

where $C_0 = \sqrt{gY_0}$ is the wave celerity, $F_0 = V_0/C_0$ is the Froude number and V_0 is the steady state velocity. In this article, we assume the flow to be *sub-critical* ($F_0 < 1$) and *non-uniform*.

Remark 1 (Non-Uniform flow.): In case of natural channels, the shape, size, and slope may vary along the stream length x. In the case of non-uniform flow, the flow variables vary along the length of the channel: the velocity $V_0(x) \neq V_0 \neq V_X$ and the stage $Y_0(x) \neq Y_0 \neq Y_X$.
This non-uniform flow can be best approximated using a backwater profile model [10] [11].

B. Linearized Saint-Venant Model

Equation (2) of the Saint-Venant model is nonlinear in the flow variables V and Y. Each term $f(V,Y)$ in the Saint-Venant model can be expanded in Taylor series around the steady state flow variables $V_0(x)$ and $Y_0(x)$. Considering only the first order perturbations: $f(V,Y) \approx f(V_0,Y_0) + (f_V)_0\, v(x,t) + (f_Y)_0\, y(x,t)$ where the first order perturbations in velocity (resp. stage) is given by $v(x,t) = V(x,t) - V_0(x)$ (resp. $y(x,t) = Y(x,t) - Y_0(x)$). The linearized Saint-Venant model for the perturbed flow variables v and y is:

$$y_t + Y_0(x)v_x + V_0(x)y_x + \frac{dY_0(x)}{dx}v - \alpha_0(x)y = 0 \tag{6}$$
$$v_t + V_0(x)v_x + gy_x + \beta_0(x)v - \gamma_0(x)y = 0 \tag{7}$$

with $\alpha_0(x)$, $\beta_0(x)$ given by:

$$\alpha_0(x) = \frac{V_0(x)}{Y_0(x)}\frac{dY_0(x)}{dx} \tag{8}$$
$$\beta_0(x) = \frac{g}{V_0(x)}\left[2S_b - (2 - F_0^2)\frac{dY_0(x)}{dx}\right] \tag{9}$$
$$\gamma_0(x) = \frac{4T_0 g}{3Y_0(x)(T_0 + 2Y_0(x))}\left[S_b - (1 - F_0^2)\frac{dY_0(x)}{dx}\right] \tag{10}$$

In the above equations, to emphasize that the free surface width T is uniform, it is denoted as T_0 and the dependence on x is omitted for readability. The upstream and downstream boundary conditions are respectively given by the upstream velocity perturbation $v(0,t)$ and the downstream stage perturbation $y(X,t)$. The initial conditions are given by $y(x,0) = 0$ and $v(x,0) = 0$ for all $x \in [0,X]$.

C. One-dimensional numerical scheme

The Preissman implicit finite difference scheme [4] is applied to these equations (6), (7):

$$f(x,t) \approx \frac{\theta}{2}(f_{j+1}^{k+1} + f_j^{k+1}) + \frac{1-\theta}{2}(f_{j+1}^{k} + f_j^{k}) \tag{11}$$
$$\frac{\partial f}{\partial x} \approx \theta\frac{f_{j+1}^{k+1} - f_j^{k+1}}{\Delta x} + (1-\theta)\frac{f_{j+1}^{k} - f_j^{k}}{\Delta x} \tag{12}$$
$$\frac{\partial f}{\partial t} \approx \frac{f_{j+1}^{k+1} + f_j^{k+1} - f_{j+1}^{k} - f_j^{k}}{2\Delta t} \tag{13}$$

where $f(x,y)$ is the flow variables (either v or y in our case), $\theta \in (0,1)$ is a time weighting coefficient, denotes the space step and k the time step. This scheme has the advantage of allowing non-equidistant grids Δx and is unconditionally stable as long as $\theta > 0.5$. This enables a more flexible schematization of the river, especially in the case of strongly varying cross sections. The time step is a function of the required accuracy only and can be chosen freely.
The discretization form of equation (6), (7) can be written as:

$$\frac{y_{j+1}^{k+1}+y_j^{k+1}-y_{j+1}^k-y_j^k}{2\Delta t}=$$
$$-Y_0(x)\left[\theta\frac{v_{j+1}^{k+1}-v_j^{k+1}}{\Delta x}+(1-\theta)\frac{v_{j+1}^k-v_j^k}{\Delta x}\right]$$
$$-V_0(x)\left[\theta\frac{y_{j+1}^{k+1}-y_j^{k+1}}{\Delta x}+(1-\theta)\frac{y_{j+1}^k-y_j^k}{\Delta x}\right]$$
$$-\frac{dY_0(x)}{dx}\left[\frac{\theta}{2}(v_{j+1}^{k+1}+v_j^{k+1})+\frac{1-\theta}{2}(v_{j+1}^k+v_j^k)\right]$$
$$+\alpha_0(x)\left[\frac{\theta}{2}(y_{j+1}^{k+1}+y_j^{k+1})+\frac{1-\theta}{2}(y_{j+1}^k+y_j^k)\right] \quad (14)$$
$$\frac{v_{j+1}^{k+1}+v_j^{k+1}-v_{j+1}^k-v_j^k}{2\Delta t}=$$
$$-V_0(x)\left[\theta\frac{v_{j+1}^{k+1}-v_j^{k+1}}{\Delta x}+(1-\theta)\frac{v_{j+1}^k-v_j^k}{\Delta x}\right]$$
$$-g\left[\theta\frac{y_{j+1}^{k+1}-y_j^{k+1}}{\Delta x}+(1-\theta)\frac{y_{j+1}^k-y_j^k}{\Delta x}\right]$$
$$-\beta_0(x)\left[\frac{\theta}{2}(v_{j+1}^{k+1}+v_j^{k+1})+\frac{1-\theta}{2}(v_{j+1}^k+v_j^k)\right]$$
$$+\gamma_0(x)\left[\frac{\theta}{2}(y_{j+1}^{k+1}+y_j^{k+1})+\frac{1-\theta}{2}(y_{j+1}^k+y_j^k)\right] \quad (15)$$

Using the discretization form of the linearized Saint-Venant Equations (14) (15), the linear model for a single channel i can be represented as:

$$E_{k,i}X_{k+1,i}=A_{k,i}X_{k,i}+B_{k,i}U_{k,i} \quad (16)$$

where $X_{k,i}$ is the state variable

$$X_{k,i}=(v_{k,i,1},y_{k,i,1},\cdots,v_{k,i,l_i},y_{k,i,l_i})^T \quad (17)$$

$U_{k,i}$ is boundary conditions at time $k\Delta t$

$$U_{k,i}=(v_{k,i,1},y_{k,i,l_i})^T \quad (18)$$

where l_i denotes the downstream point of each channel i, and 1 is the upstream point of each channel i. $E_{k,i}$, $A_{k,i}$ and $B_{k,i}$ are matrices determined by numerical method above. $v_{k,i,j}$ and $y_{k,i,j}$ are the velocity and stage perturbation at location $j\Delta x$ at time $k\Delta t$ in channel i.

D. Linear Network Model

A linear channel network model can be constructed by decomposing the channel network into individual channel reaches, and apply the linear model (16) to each branch. The internal boundary conditions are also imposed at every junction to ensure flow compatibility. Considering a simple river junction illustrated in Figure 1, the linear equations of hydraulic internal boundary conditions at a junction are specified by equations of mass and energy conservation. Assuming no change in storage volume within the junction, the continuity equation can be expressed by:

$$v_{k,1,l_i}\cdot T_1=v_{k,2,1}\cdot T_2+v_{k,3,l}\cdot T_3$$

When the flows in all the branches meeting at a junction are subcritical, the equation for energy conservation can be approximated by a kinematic compatibility condition as:

$$y_{k,1,l_i}=y_{k,2,1}=y_{k,3,1}$$

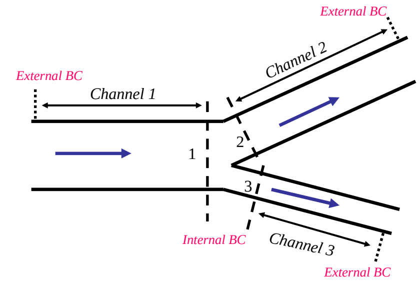


Fig. 1. *Flow compatibility of channel junctions*

The equations are assembledfor each individual channel and interior junctions together to model the entire network. The flow variables inside the domain are represented by a linear relationship:

$$E_kX_{k+1}=A_kX_k+B_kU_k \quad (19)$$

where X_k is the concatenated vector of $X_{k,i}$ and U_k is the boundary conditions of the channel network system.
The boundary conditions of (23) are given by

$$U_k=\left[u(k,i,j)|_{\partial\Omega_{upstream}},y(k,i,j)|_{\partial\Omega_{downstream}}\right] \quad (20)$$

and initial conditions given by

$$X_0=0 \quad (21)$$

The linear network model parameters are the average free surface width $T_{0,i}$, the average bottom slope $S_{b,i}$, the average Manning's coefficient n, the average velocity $V_{0,i}$, and the average downstream stage $Y_{l_i,i}$ for each channel i ($i=1,\cdots,13$). These parameters are known to us experimentally.

III. Variational Data Assimilation Using Quadratic Programming

A. General Considerations

In this section, open boundary condition estimation is formulated using the information of velocity and position measurements provided by a number of drifters which are released in a channel network. Following standard procedure in variational data assimilation, the cost function used by this article consists minimizing the difference between measured velocity at the location of the drifter and velocity estimated by the model. With the linear model constrain, the problem can be formulated as a QP and solved efficiently. Furthermore, under the assumption that tidal flow variables can be expressed by dominant oscillatory modes, the number of estimation variables is extremely reduced.

B. Notations

We employ the traditional notation of variational data assimilation in discrete time and space [17]:

- X_k : Vector of state variables (v,y) for each mesh point at time $k\Delta t$.
- Y_k : Vector of observed variables at time $k\Delta t$.
- R_k : Covariance matrix of the observation error at time $k\Delta t$.

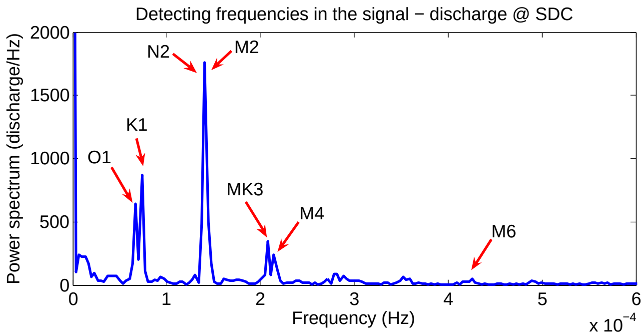


Fig. 2. *Spectral analysis of the discharge at the SDC Station.*

- H_k : Observation operator, which projects the state vector X_k into the observation subspace containing Y_k.

We deploy D passive drifters in the network to collect Lagrangian measurements of the velocity in the system, and try to estimate boundary conditions by minimizing the ℓ^2-norm of the error between the observed data and the the corresponding model predictions:

$$\mathcal{J} = \sum_k (Y_k - H_k[X_k])^T R_k^{-1} (Y_k - H_k[X_k]) \qquad (22)$$

This positive semi-definite quadratic cost function is constrained by:

$$E_k X_{k+1} = A_k X_k + B_k U_k \qquad (23)$$

In this way, the variational data assimilation problem can be posed as a Quadratic Programming problem:

$$\begin{aligned} \min \quad & \frac{1}{2} X^T P X + q^T X \\ \text{s.t.} \quad & GX \leq h \\ & FX = b \end{aligned} \qquad (24)$$

where X is the concatenated vector of X_k from time 0 to the final time step; P is a symmetric matrix and q is vector containing the information of Y_k, H_k and R_k; F and b are the block diagonal matrix of A_k and B_k. Normally G and h are 0, and the QP can be solved by a linear system. In our case, we may impose heuristic inequality constraints to reduce the search space.

C. Decision Variables

The decision variables of the QP problem (24) are the flow variables at the open boundaries. If it is expressed in the time domain, the number of decision variables would be equal to the number of boundaries times the number of time steps. Using spectrum analysis, it is obvious that flow variables in a tidal system can be modeled by seven dominant tidal modes, as seen in Figure 2,

These dominant tidal modes are listed in Table I. Thus, any flow variables at the boundaries can be specified as:

$$u(k\Delta t) \approx \sum_{l=0}^{7} \left[d_l e^{j\omega_l k\Delta t} + \overline{d_l} e^{-j\omega_l k\Delta t} \right] \qquad (25)$$

TABLE I

THE DOMINANT TIDAL MODES IN SACRAMENTO DELTA

Tide	Tide Period(hours) T_l	Tide Frequency (s^{-1}) $\omega_l = \frac{2\pi}{T_l}$
K1	23.9345	$7.2921 \cdot 10^{-5}$
M2	12.4206	$1.4082 \cdot 10^{-4}$
MK3	8.1771	$2.1344 \cdot 10^{-4}$
M4	6.2103	$2.8104 \cdot 10^{-4}$
M6	4.1202	$4.2360 \cdot 10^{-4}$
O1	25.8193	$6.7598 \cdot 10^{-5}$
N2	12.6584	$1.3788 \cdot 10^{-4}$

where $\omega_l = \frac{2\pi}{T_l}$is the frequency associated with one of the seven dominant tidal periods. The decision variables of this inverse modeling problem become the unknown coefficients d_l corresponding to specified tidal frequencies for each boundary to be estimated. In this way, the number of decision variables is reduced, which speeds up the convergence of QP process.

IV. EXPERIMENT PREPARATION

A. Experiment Protocol

The following is a description of an experiment to test the proposed method. The intuitive way to test it is to assimilate field Lagrangian data into our linear model and compare estimated the boundary conditions with Eulerian measurements at the boundaries. Since the field instrument development and data collection is still under process, the Lagrangian drifter data in the article is generated by using Telemac 2D [19], a fully nonlinear Shallow Water Equation (SWE) solver, with an unstructured triangular grid mesh and finite element method. The virtually simulated drifter data will be replaced by hardware experiment field measurements in future studies.

A set of fixed Eulerian U.S. Geological Survey (USGS) sensors (see Figure 3(a)) on this hydraulic system is employed as the boundary conditions for the model simulation, and a finite number of passive drifters are released virtually in the experiment period. During the inverse modeling process, only these simulated drifter data are used to re-construct open boundary conditions, which are then compared with the initial boundary setting. Another set of USGS Eulerian sensors along with the deployed fixed Acoustic Doppler Current Profiling (ADCP) instrumentation and Water Pressure Sensors (see Figure 3(b)) are used to validate the flow characteristics inside the experiment domain. The flow chart of the experiment process is shown in Figure 4.

B. Two-dimensional Shallow Water Equations and Numerical Forward Simulation

In this subsection, we will set the forward simulation and introduce the Lagrangian measurements.

1) Two-dimensional Shallow Water Equations: The governing hydrodynamic equations for forward simulation are

(a) USGS Sensor station (b) Deployable ADCP sensor

Fig. 3. ***Left:*** *USGS Sensor station at* GSS*, used as a measurement sensor.* ***Right:*** *ADCP sensor deployed in the Sacramento River, used in Section 5.3 for gathering the validation data*

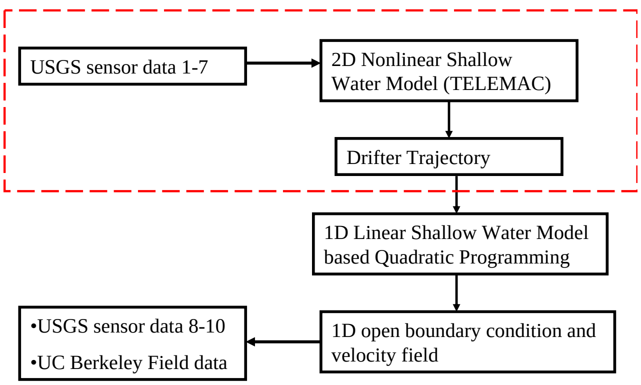


Fig. 4. *Data Assimilation Flow Diagram. The red box denotes the part which in the future will be replaced by hardware experiment filed measurements.*

[19]:

$$\frac{\partial h}{\partial t} + \vec{u} \cdot \nabla h + h \nabla \cdot \vec{u} = 0 \quad (26)$$

$$\frac{\partial u}{\partial t} + \vec{u} \cdot \nabla u = -g \frac{\partial \eta}{\partial x} + F_x + \frac{1}{h} \nabla \cdot (h \nu_t \nabla u) \quad (27)$$

$$\frac{\partial v}{\partial t} + \vec{u} \cdot \nabla v = -g \frac{\partial \eta}{\partial y} + F_y + \frac{1}{h} \nabla \cdot (h \nu_t \nabla v) \quad (28)$$

The friction forces are given by the following Manning law:

$$F_x = -\frac{1}{\cos \alpha} \frac{g n^2}{h^{4/3}} u \sqrt{u^2 + v^2} \quad (29)$$

$$F_y = -\frac{1}{\cos \alpha} \frac{g n^2}{h^{4/3}} v \sqrt{u^2 + v^2} \quad (30)$$

$$(31)$$

where h is the total depth of water, $\vec{u} = (u, v)$ is the velocity in the domain, g is the gravity acceleration, η is the free surface elevation, ν_t is the coefficient of turbulence diffusion, α is the bed slope of river bottom, and n is the Manning coefficient. The boundary condition and initial condition are given by:

$$u(x, y, t)|_{\partial \Omega_{land}} = 0, v(x, y, t)|_{\partial \Omega_{land}} = 0 \quad (32)$$

$$(u(x, y, t), v(x, y, t))|_{\partial \Omega_{upstream}} = f(x, y, t) \quad (33)$$

$$\eta(x, y, t)|_{\partial \Omega_{downstream}} = g(x, y, t) \quad (34)$$

$$u(x, y, 0) = u_0, v(x, y, 0) = v_0, h(x, y, 0) = h_0, \quad (35)$$

where $\partial \Omega$ represents the boundaries of our computational domain and f, g are known functions.

2) Lagrangian drifters: The deployed drifters is modeled as passive Lagrangian tracers. In this framework, the drifters move with the local flow streamline, obeying the following equations:

$$\frac{d x_D(t)}{dt} = u[x_D(t), y_D(t), t] \quad (36)$$

$$\frac{d y_D(t)}{dt} = v[x_D(t), y_D(t), t] \quad (37)$$

with the drifter initial conditions

$$x_D(t) = x_{D,0}, y_D(t) = y_{D,0} \quad (38)$$

3) Numerical solution: The numerical solution of the 2D shallow-water equations and drifter positions, is computed using a commercial hydrodynamic software Telemac 2D [19]. Telemac 2D uses a streamline upwind Petrov-Galerkin based finite element solver for hydrodynamic equations. The turbulence and mixing process at the estuaries are considered. To generate the drifter data, a forward simulation is run from time t_0 to time t_1 with given boundary conditions to stabilize the flow. At t_1, drifters are released randomly inside the domain and their trajectories are simulated using a Runge-Kutta method and the velocity field provided by the nonlinear shallow water forward simulation. The data assimilation process estimated the boundary conditions, which are compared with the previously given boundary conditions, along with the flow variables at intermediate locations within the watershed.

V. Case Study: The Sacramento Delta

A. General introduction the the Sacramento Delta

The Sacramento-San Joaquin Delta in California is a valuable fresh water resource and an integral part of California's water system. This complex network covers 738,000 acres interlaced with over 1,150 km of tidally-influenced channels and sloughs. This network is monitored by a static sensor infrastructure subject to the usual problems of inaccuracy and measurement errors for sensing systems. The area of interest for our experiment covers Sacramento River, Cache Slough, Steamboat Slough, Sutter Slough, Minor Slough, Delta Cross Channel and Georgiana Slough as shown in Figure 5. Most of the time, the direction of mean river flow is from north to south, as indicated with arrows. During the tidal inversion, the water flows in the opposite way.

Ten USGS stations, named HWB, RYI, SRV, HWV, SUT, SSS, SDC, DLC, GES, and GSS, are scatterly located in this experiment field. The stations are marked as diamonds in

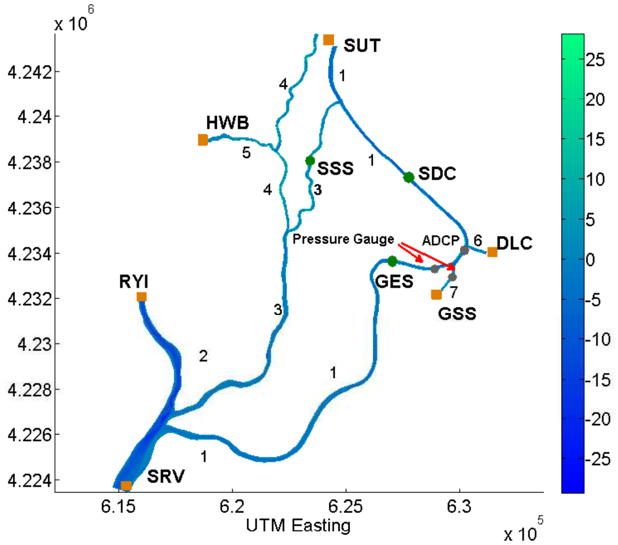


Fig. 5. *Experiment area in the Sacramento River (1), Cache Slough (2), Steamboat Slough (3), Sutter Slough (4), Minor Slough (5), Delta Cross Channel (6) and Georgiana Slough (7).*

Figure 5. Both velocity and stage are collected every 900 seconds at these stations. Note that in the USGS measurement system, only the stage are measured directly. The velocity data is estimated by a rating curve, which is a relation between stream stage and stream flow. The relation of stream stage to stream flow is always changing, and need to be calibrated frequently. It will introduce errors if the rating curve has not been validated in time.

The field data was collected between 11/12/2007 0:00am to 8:30am. In addition, the following simplifications for the flow model have been made in this study:

- The flow can be represented by a one dimensional model.
- The channel geometry is fixed, as the effects of sediment deposition and scour are negligible during the experiment period.
- The channel geometry can be modeled by a rectangular cross-section.
- The lateral and vertical accelerations are negligible.
- The pressure distribution is hydrostatic.
- There is no significant jump along the bathymetry of the channel, and the bed slope is smooth and small.
- The water surface across any cross-section is horizontal.

B. Drifter Data Generation

Telemac 2D is used to perform a non linear flow simulation using velocity data measured by USGS station SRV, RYI, GSS, and stage data measured by DLC and SUT. The geometry of the area is complex; thus we use an unstructured finite element mesh (41375 nodes, 74983 triangular elements). The bottom friction is modeled using Manning's law. The Manning coefficient is chosen to be constant in time and space, and equal to 0.02, corresponding to a straight gravel bottom [6]. The turbulence process is included such that the flow streamline at the estuaries are similar to the real world. The simulation runs for two and a half hours before the release of the drifters so that a stable state is reached. The drifters are released from 2:30AM to 6:30PM on November 11th 2007. This time period was chosen to capture the highly variable flow in Sacramento Delta. We release a total of 39 drifters during the experiment (6 hours). The first thirteen drifters are released at 2:30AM on the centerline of selected sub-channels. Then at 4:30AM, 6:30am we release another two sets of thirteen drifters respectively. Drifter positions are recorded every 60 seconds until the end of the experiment at 8:30AM. Figure 6 shows the drifter trajectories and the snapshots of the drifter positions corresponding to the three releases of the drifters.

C. Implementation of the algorithm

Following the method described in III, we assimilate the drifter data generated by TELEMAC (as described in V-B) to reconstruct the boundary conditions at SRV, RYI, GSS, DLC and SUT. The reconstructed boundary condition data is shown and compared to measured data in Figures 7. From the figures, the estimated data is very close to the measurements. The QP problem was expressed by the optimization modeling language AMPL and solved with CPLEX. With 13 sub-channels, the assimilation process takes approximately 65 minutes to compute with a 2.33 GHz Pentium dual core processor.

Without loss of generality, the flow variables measured by USGS sensor GES, SDC, SSS, HWB and the velocity and stage data recorded by selective deployable UC Berkeley sensors (marked as grey circles Figure 5) are used to achieve the model validation. The simulation results are shown in Figure 8.

The difference between the modeled data and measurements is further evaluated in Table II. Three primary evaluation measures are analyzed here:

- The maximum value is the maximum difference between the estimated and measured data at the same time steps.
- The coefficient of efficiency E is defined as [14]:

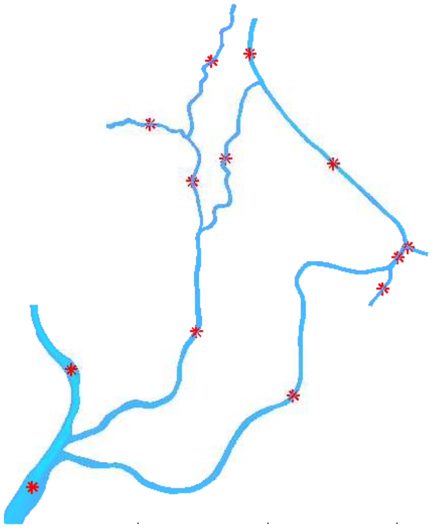
$$E = 1 - \left[\frac{\sum_{i=1}^{N}(\hat{u}_i - u_i)^2}{\sum_{i=1}^{N}(u_i - \overline{u_i})^2}\right] \tag{39}$$

  where u_i is the flow variable of interest (for example v_i or y_i in this study), $\hat{u}_i$ is the modeled flow variable, $\overline{u_i}$ is the mean of u_i, for $i = 1$ to N measurement events. If the measured data is perfect, $E = 1$. If $E < 0$, the corresponding measurement is not reasonable and must be excluded from the modeling procedure.
- The correlation coefficient ρ is given by:

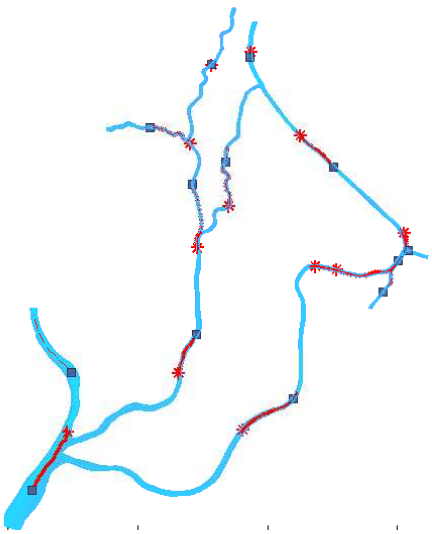
$$\rho = \frac{\sum_1^N (u_i - \overline{u_i})(\hat{u}_i - \overline{\hat{u}_i})}{\sqrt{\sum_1^N (u_i - \overline{u_i})^2 \sum_1^N (\hat{u}_i - \overline{\hat{u}_i})^2}} \tag{40}$$

  where $\overline{\hat{u}_i}$ represents the mean of model estimated flow for $i = 1$ *to* N measurement events. If the measured data is perfect, $rho = 1$

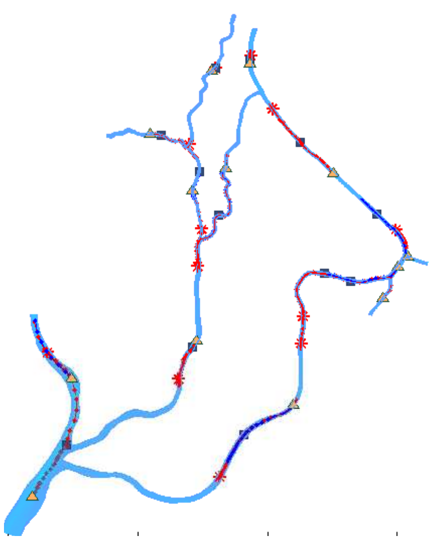
Figure 8 and Table II thus indicate that the proposed approach provides a good accuracy in the flow estimation.



(a) Time Step 0

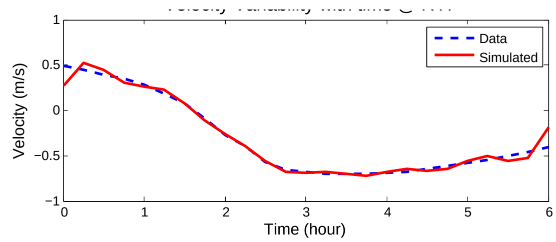


(b) Time Step 120

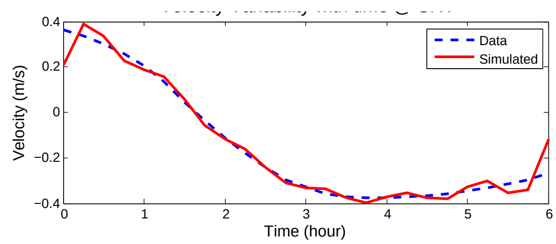


(c) Time Step 240

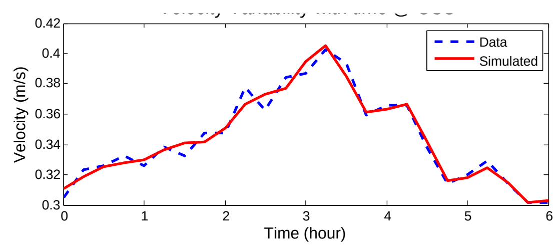
Fig. 6. *Drifter trajectories and their release positions. 13 drifters are released time step 0 ($*$), 13 drifters at time step 120 ($\square$) and 13 drifters at time step 240 (Δ) . The drifter positions are recorded every 60 second until the end of the experiment.*



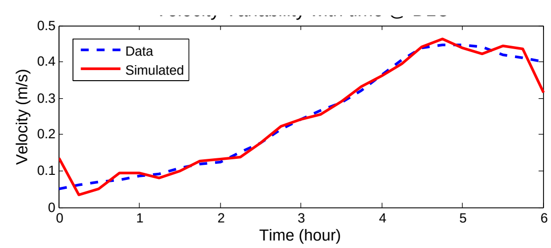
(a) Valocity Variability with Time at USGS station: RYI



(b) Valocity Variability with Time at USGS station: SRV



(c) Valocity Variability with Time at USGS station: GSS



(d) Valocity Variability with Time at USGS station: DLC

Fig. 7. *Comparison of the estimated boundary condition with USGS measurements at the boundary of the domain.*

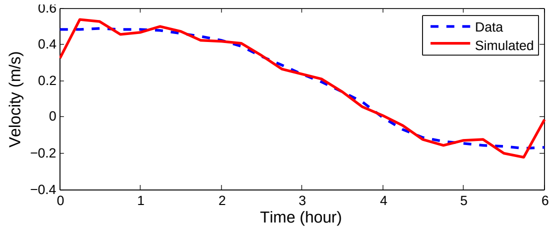
TABLE II

MAX-VALUE, ρ-VALUE AND E-VALUE FOR MODELED DATA AND MEASURED DATA

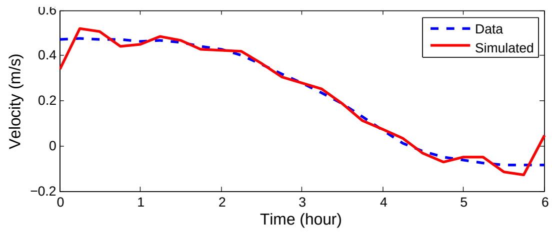
Variable	USGS Station	Max-value	E-value	ρ-value
Velocity	GES	0.05 m/s	0.9930	0.9975
	SDC	0.04 m/s	0.9368	0.9883
	SSS	0.055 m/s	0.9968	0.9985
	ADCP	0.07 m/s	0.9435	0.8923
Stage	GES	0.05 m	0.9889	0.9947
	SDC	0.12 m	0.9504	0.9759
	SSS	0.07 m	0.9847	0.9935
	Pressure Sensor I	0.06 m	0.8479	0.8743
	Pressure Sensor II	0.06 m	0.9345	0.8734

VI. CONCLUSIONS

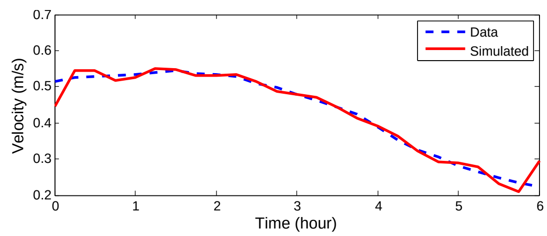
In this article we have presented a boundary condition estimation method for complex channel networks using Lagrangian measurement data. The solution is formulated as a



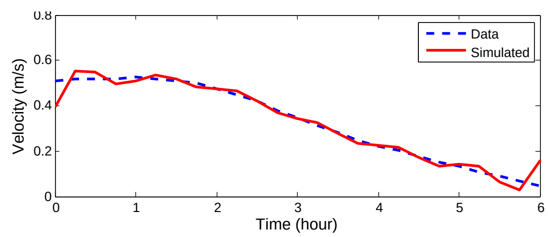
(a) Valocity Variability with Time at USGS station: GES



(b) Valocity Variability with Time at USGS station: SSS



(c) Valocity Variability with Time at USGS station: SDC



(d) Valocity Variability with Time at USGS station: ADCP

Fig. 8. *Validation of the model output with USGS and ADCP measurements inside of the domain.*

QP problem based on minimizing the difference between measured Lagrangian data and modeled drifter trajectory, constrained by a 1D implicit linear channel network model. A major advantage of the 1D QP formulation is that requires low computational cost, which makes the method applicable to vast and complex networks of open channels. Modal decomposition allows the estimated output be expressed in terms of dominant tidal frequency. This reduces the number of decision variables, which substantially lower computation complexity. The performance of the method has been demonstrated using a experiment setting in which the drifter data were generated by a 2D nonlinear shallow water model. From the results, the performance of the method has been validated.

Future works using the method include the use of real data collected from the GPS equipped drifters deployed in the Sacramento San Joaquin Delta. The effects of different deployment strategy and number of drifter on the performance of the method is also of interested.

VII. Acknowledgment

Xavier Litrico from CEMAGREF is gratefully acknowledged for fruitful discussions during his stay at Berkeley, which shaped the framework used for modeling shallow water systems, which follow from his work. Real-time flow data at USGS stations are downloaded from the California Data Exchange Center (CDEC) of Dept. of Water Resources (DWR) (http://cdec.water.ca.gov). The flow data used for validation is obtained from deployable sensors provided to us by Mark Stacey. We are grateful to Julie Percelay, Maureen Downing-Kunz and Mark Stacey for their help with the deployment of these sensors.

References

[1] J. S. Albuquerque and L. T. Biegler. Data reconciliation and gross-error detection for dynamic systems. *American Institute of Chemical Engineering Journal*, 42(10):2481–2586, 1996.
[2] R. A. Anthes. Data assimilation and initialization of hurricane prediction models. *Journal of Atmospheric Science.*, 31:702–719, 1974.
[3] California Department of Water Resources. *Enhanced Calibration and Validation of DSM2 HYDRO and QUAL.*
[4] K.W. Chau. Application of the preissmann scheme on flood propagation in river systems in difficult terrain. *Hydrology in Mountainous Regions. I - Hydrological Measurements; the Water Cycle*, (193), 1990.
[5] V. Chow. *Open-channel Hydraulics*. McGraw-Hill Book Company, New York, 1988.
[6] C. Crowe, D. Elger, and J. Roberson. *Engineering Fluid Mechanics*. US John Wiley and Sons, Inc, 2001.
[7] J.A Cunge, F.M. Holly, and A. Verwey. *Practical aspects of computational river hydraulics*. Pitman, 1980.
[8] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer-Verlag, 2007.
[9] H. B. Fischer. *Mixing in Inland and Coastal Waters: in inland and coastal waters*. Academic Press, 1979.
[10] X. Litrico and V. Fromion. Simplified modeling of irrigation canals for controller design. *Journal of Irrigation and Drainage Engineering*, 130(5):373–383, 2004.
[11] X. Litrico and V. Fromion. Boundary control of linearized saint-venant equations oscillating modes. *Automatica*, 42(6):967–972, 2006.
[12] A. Molcard, L.I.Piterbarg, A.Griffa, T.Ozgokmen, and A. Mariano. Assimilation of drifter observations for the reconstruction of the eulerian circulation field. *Journal of Geophysical Research*, 108(C3):3056, 2003.
[13] A. Molcard, A.C. Poje, T.M. Ozgokmen, and J. Sau. 06: Directed drifter launch strategies for lagrangian data assimilation using hyperbolic trajectories. *Ocean Modelling*, 12:268–289, 2006.
[14] J.E. Nash and J.V. Sutcliffe. River flow forecasting through conceptual models part i - a discussion of principles. *Journal of Hydrology*, 10:282–290, 1970.
[15] I.M. Navon. Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dynamics of Atmospheres and Oceans*, 27(C3):55–79, 1997.
[16] C. Paniconi, M. Marrocu, M. Putti, and M. Berbunt. Newtonian nudging for a richards equation-based distributed hydrological model. *Advances in Water Resources*, 26(2):161–178, 2003.
[17] S. Polavarapu, M. Tanguay, and L. Fillion. Four-dimensional variational data assimilation with digital filter initialization. *Monthly Weather Review*, 128(7):2491–2510, 1999.
[18] P.M. Poulain, D.S.Luther, and W.C.Patzert. Deriving inertial wave characteristics from surface drifter velocities - frequency variability in the tropical pacific. *Journal of Geophysical Research*, 97(C11):17947–59, 1992.
[19] Report EDF. *TELEMAC 2D. Version 5.2 – Principle note*, 2002.
[20] I.S Strub, J. Percelay, M.T.Stacey, and A.M.Bayen. Inverse estimation of open boundary conditions in tidal channels. *Ocean Modelling*, 2009.
[21] O.-P. Tossavainen, J. Percelay, A. Tinka, Q. Wu, and A. Bayen. Ensemble kalman filter based state estimation in 2d shallow water equations using lagrangian sensing and state augmentation. In *Proceedings of the 46th Conference of Decision and Control*, 2008.
[22] Q. Wu, X. Litrico, and A. Bayen. Data reconciliation of an open channel flow network using modal decomposition. *Advances in Water Resources*, 32:193–204, 2009.