# Data Reconciliation of an Open Channel Flow Network using Modal Decomposition

Qingfang Wu [a,*], Xavier Litrico [b], Alexandre M. Bayen [a]

[a]*Department of Civil and Environmental Engineering, UC Berkeley, CA, USA*

[b]*Cemagref, UMR G-EAU, B.P. 5095, 34196 Montpellier Cedex 5, France*

**Abstract**

This article presents a method to estimate flow variables for an open channel network governed by the linearized Saint-Venant equations and subject to periodic forcing. The discharge at the upstream end of the system and the stage at the downstream end of the system are defined as the model inputs; the flow properties at selected internal locations, as well as the other external boundary conditions, are defined as the outputs. Both inputs and outputs are affected by noise and we use the model to estimate this data. A spatially-dependent transfer matrix in the frequency domain is constructed to relate the model input and output using modal decomposition. A data reconciliation technique is used to incorporate the error in the measured data and results in a set of reconciliated external boundary conditions; subsequently, the flow properties at any location in the system can be accurately constructed from the input measurements. The applicability and effectiveness of the method is demonstrated with a case study of the river flow subject to tidal forcing in the Sacramento-San Joaquin Delta, in California. We used existing USGS sensors in place in the Delta as measurement points, and deployed our own sensors at selected locations to produce data used for the validation. The proposed method gives an accurate estimation of the flow properties at intermediate locations within the channel network.

*Key words:* linearized Saint-Venant equations, hyperbolic partial differential equations, distributed transfer function, modal decomposition, data reconciliation

* Corresponding author. Tel: 510-642-9569
 *Email addresses:* `qingfangwu@berkeley.edu` (Qingfang Wu),
`xavier.litrico@cemagref.fr` (Xavier Litrico), `bayen@berkeley.edu` (Alexandre M. Bayen).

# 1 Introduction

In hydraulic systems, numerous factors could lead to measurement errors, for example broken gauges, process leaks, sensor drifts, improper use of measuring devices, and other random sources [1]. Data reconciliation, is an effective method to tune-up the measurement data [21] [5] [10] [2], which has been applied in several engineering fields [11] [3] [20] [25]. The objective of data reconciliation is to use information redundancy to handle errors in real-time measurements.
In the field of process control, data reconciliation is a part of the general state estimation or reconstruction process for dynamical systems, using Kalman filtering [8]. However, in certain cases, and for given time intervals, dynamic effects can sometimes be neglected which leads to simplified versions of the general approach, applicable to static models.

This article presents theoretical results applicable to data reconciliation for tidally forced networks of open channels. Using the modal decomposition techniques, we are able to transform dynamic constraints into static constraints in the frequency domain, and subsequently obtain a static data reconciliation problem, which is easier to resolve and can lead to accurate results. Generally, this static data reconciliation problem is to minimize the measurement errors while satisfying the static constraints of the proposed model.

The proposed linear network model is constructed on the basis of analytical solutions to the *Linearized Saint-Venant equations* (LSVE) in the frequency domain [9] [12] [4]. With the assumption of a backwater curve model [22], a more realistic transfer matrix function has been introduced [16], which we use in the present article.

This article extends the general transfer matrix function approach to a channel network. A spatially-dependent transfer matrix is constructed, relating a selected set of model inputs to the output variables. The transfer matrix is a function of channel width, channel length, bed slope, mean discharge, mean stage and Manning coefficient. This set of parameters needs to be chosen carefully to characterize the geometry of the channels, as the uncertainty of the parameters would contribute to the errors in the model output.

With this linear model in the frequency domain, the static data reconciliation problem is shown to be equivalent to a quadratic problem. The objective function used in the present study is a weighted $L_2$-norm of the difference between the measured and reconciliated data. The linear network model constructed

serves as the constraints in the optimization problem. A closed-form optimal solution is obtained, resulting in a set of reconciliated boundary data consistent with both the linear network model and the statistical assumptions on measurement errors. Subsequently, we apply the reconstructed boundary conditions to the linear network model to obtain an accurate forward simulation of the flow within of the network domain.

This article is organized as follows: Section II introduces the general framework of linear models, i.e., LSVE in the frequency domain, the spatially-dependent transfer matrix. A channel network model featuring one-dimensional non-uniform flow is subsequently described, and the solution of the data reconciliation problem in the static case is addressed. Section III applies the linear model to a channel network in the Sacramento - San Joaquin Delta in California. Static data reconciliation is applied to handle the errors in the measurements. The effectiveness of the method is assessed by correlating the model estimations with field data at three intermediate locations in the network, which serve as validation points. Section IV summarizes the study and presents the scope of our future work.

## 2   Proposed Method

### 2.1   General Considerations

The general class of hydraulic system studied in the present article is a distributed network of channels subject to quasi-periodic tidal forcing. Sensing on this hydraulic system is done using fixed Eulerian *U.S. Geological Survey* (USGS) sensors, subject to measurement errors. The motivation of the work is to derive the "most likely" flow conditions from the measured data available in a given period of time (at least a few weeks to catch the tidal period), and forecast the future flow variables. The goal of the method is thus not real time estimation, as traditionally done in data assimilation [23], but forecast based on measured forcing.

The flow variables are related to each other by a mathematical model. Therefore, if the measurement data was error free, it would satisfy the model. Because the number of points at which the variables are measured is usually larger than needed to fully prescribe the model, there is "information redundancy" in the system. Once information redundancy exists, data reconciliation can be implemented to account for measurement error.

The ultimate goal of data reconciliation is to use such information redun-

dancy in a system to have the data self-corrected using the model. An effective data reconciliation method allows the detection of any inconsistent or biased measurements, and furthermore provides corrected values (namely estimated measurements).

It should be noted that any information redundancy is model-specific. We therefore need to first construct a "good" hydraulic model to characterize the flow system, as described in the following section.

## 2.2  Linear Channel Network Model

### 2.2.1  Transfer Matrix Representation of Saint-Venant Model

The *Linearized Saint-Venant Equations* (LSVE) have been widely used in the open-channel hydraulic systems literature [7], [15], [17], [14]. They describe the perturbed discharge $q(x,t)$ and stage $y(x,t)$ with two coupled *partial differential equations* (PDEs). For a rectangular cross-section, these equations are given by:

$$T_0 \frac{\partial y}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{1}$$

$$\frac{\partial q}{\partial t} + 2V_0(x)\frac{\partial q}{\partial x} - \beta_0(x)q + \alpha_0(x)\frac{\partial y}{\partial x} - \gamma_0(x)y = 0 \tag{2}$$

where $\alpha_0(x)$, $\beta_0(x)$ and $\gamma_0(x)$ are given by:

$$\alpha_0 = (C_0^2 - V_0^2)T_0 \tag{3}$$

$$\beta_0 = -\frac{2g}{V_0}\left(S_b - \frac{dY_0}{dx}\right) \tag{4}$$

$$\gamma_0 = gT_0\left[(1+\kappa_0)S_b - (1+\kappa_0 - (\kappa_0 - 2)F_0^2)\frac{dY_0}{dx}\right] \tag{5}$$

with $\kappa_0 = 7/3 - 8Y_0/(3(2Y_0 + T))$; $T_0$ is denoted as a uniform width at the free surface, $C_0 = \sqrt{gY_0}$ is the wave celerity, $F_0 = V_0/C_0$ is the Froude number, $V_0 = Q_0/A_0$ is the steady state velocity, $Q_0$ is the average discharge along the channel and $Y_0(X)$ is the average stage at the downstream point of the channel, $X$ is the river reach length $(m)$, $S_b$ is the bed slope $(m/m)$.

The upstream and downstream boundary conditions are the upstream discharge perturbation $q(0,t)$ and the downstream stage perturbation $y(X,t)$, respectively. The initial conditions are given by $y(x,0) = 0$ and $q(x,0) = 0$ for all $x \in [0, X]$.

To facilitate the mathematical analysis, we rewrite the linearized Saint-Venant

equations as follows:

$$\frac{\partial}{\partial t}\begin{pmatrix} q(x,t) \\ y(x,t) \end{pmatrix} = \left(\mathcal{A}(x)\frac{\partial}{\partial x} + \mathcal{B}(x)\right)\begin{pmatrix} q(x,t) \\ y(x,t) \end{pmatrix} \qquad (x,t) \in [0,X] \times [0,+\infty)$$

$$(6)$$

Where,

$$\mathcal{A}(x) = \begin{pmatrix} -2V_0 & -\alpha_0 \\ -\frac{1}{T_0} & 0 \end{pmatrix} \qquad \mathcal{B}(x) = \begin{pmatrix} \beta_0 & \gamma_0 \\ 0 & 0 \end{pmatrix} \qquad\qquad (7)$$

The application of Laplace transform to the linear PDE system (6) leads to the following *ordinary differential equations* (ODEs) in the variable $x$, with a complex parameter $s$.

$$\frac{d}{dx}\begin{pmatrix} q(x,s) \\ y(x,s) \end{pmatrix} = \mathcal{A}^{-1}(x)\left[\mathcal{B}(x) - s\mathbb{I}_2\right]\begin{pmatrix} q(x,s) \\ y(x,s) \end{pmatrix} \qquad\qquad (8)$$

Following the method developed in [22], and further modified in [16], a transfer matrix $\mathcal{G}(x,X,s) = (g_{ij}(x,X,s))$ for the non-uniform channel relates the boundary conditions and intermediate flow variables, and is defined as:

$$\begin{pmatrix} q(x,s) \\ y(x,s) \end{pmatrix} = \mathcal{G}(x,X,s)\begin{pmatrix} q(0,s) \\ y(X,s) \end{pmatrix} \qquad\qquad (9)$$

$\mathcal{G}(x,X,s)$ is a function of channel length $X$, average discharge $Q_0$, average downstream depth $Y_X$, average width $T_0$, bed slope $S_b$, and Manning coefficient $n$ of the channel. The upstream and downstream boundary conditions are the upstream discharge perturbation $q(0,s)$ and the downstream stage perturbation $y(X,s)$, respectively. Because of the distributed nature of the system, this transfer function also depends on the coordinate $x$ in the channel, since it relates inputs $q(0,s)$ and $y(X,s)$ to the state of the system $q(x,s)$ and $y(x,s)$ at any $x$ in the channel. Please refer to the appendix for the details about the transfer matrix.

### 2.2.2 Transfer Matrix Model for Channel Networks

The model (9) can be readily applied to tidally driven channel networks. The problem of interest can be stated as follows, and is illustrated in Figure 1. Given a set of "external" boundary conditions of a network, at which we have measurements, reconstruct flow conditions at "internal" locations (also referred to as boundary conditions). This type of problem appears in our
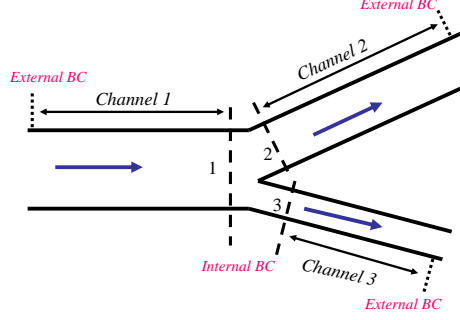
5

Fig. 1. *Representation the system of interest for estimating internal boundary conditions (Internal BC) using data reconciliation on external boundary conditions (External BC): data is given at the three external conditions of channel 1, 2 and 3; the state of the system is computed at three internal locations of the system labeled 1, 2 and 3.*

data assimilation work, in which we need estimates of boundary conditions at locations where fixed sensors are not available. The fundamental approach to build a network model is as follows:

- **Step 1**: Decompose the channel network into individual channel reaches, and apply the linear model (9) to each branch. For each of the river reach indexed by $i$, the flow variables $q_i(x, s)$, $y_i(x, s)$ denote the perturbed discharge and stage in the frequency domain respectively, $X_i$ denote the length of the channel. The junction of the river reach is defined as the node of the channel network.
- **Step 2**: Impose the internal boundary conditions at every junction to ensure flow compatibility. Considering a simple river junction illustrated in Figure 1: for each channel $i$, we will have the upstream boundary condition $y_i(0, s)$, $q_i(0, s)$, and downstream boundary condition $y_i(X_i, s)$, $q_i(X_i, s)$. If these boundary conditions are at the inside nodes of the channel network, they are called internal boundary conditions, otherwise they are labeled as external boundary conditions. The linear relationships 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:

$$q_1(X_1, s) = q_2(0, s) + q_3(0, s)$$

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:

$$y_1(X_1, s) = y_2(0, s) = y_3(0, s)$$

where $X_i$ is the downstream point of each channel $i$, and 0 is the upstream point of each channel $i$.

- **Step 3**: Assemble the equations for each individual channel and interior junctions together to model the entire network. The flow variables at the boundary of each channel are represented by a linear relationship:

$$M(s)Z(s) = 0 \qquad (10)$$

where $Z(s)$ is the concatenated vector of all $[q_i(0, s), q_i(X_i, s), y_i(0, s), y_i(X_i, s)]^T$, where $i = 1, \cdots, N$, $Z(s)$ is thus the vector comprising the discharge and stage variables at the upstream and downstream ends of all channels; $M(s)$ is a matrix of appropriate dimension encoding the previous constraints.
- **Step 4**: Evaluate the unmeasured flow variables inside the channel network.

  a) Specify the interior boundaries.

  In a channel network system, we define a subjective subset of boundary conditions ($Z_{givenBC} \subset Z$), which leads to a unique solution of model (10). This subset should satisfy: $dim(Z_{givenBC}) = dim(Z) - Rank(M)$. All the other unknown boundary variables (interior and external), denoted as $Z_{otherBC} = Z \setminus Z_{givenBC}$, are therefore estimated with model (10). Model (10) now has the form:

$$Z_{otherBC} = R(s)Z_{givenBC} \qquad (11)$$

where $R(s)$ is a matrix of appropriate size. Given $Z = \begin{pmatrix} Z_{givenBC} \\ Z_{otherBC} \end{pmatrix}$, $M(s) = \begin{bmatrix} R(s) \, | -\mathbb{I} \end{bmatrix}$.

  b) Estimation of the perturbed discharge and stage along the channel. It is achieved by a simple application of transfer function analysis:

$$\begin{pmatrix} q_i(x, s) \\ y_i(x, s) \end{pmatrix} = \begin{pmatrix} g_{i,11}(x, X_i, s) & g_{i,12}(x, X_i, s) \\ g_{i,21}(x, X_i, s) & g_{i,22}(x, X_i, s) \end{pmatrix} \begin{pmatrix} q_i(0, s) \\ y_i(X_i, s) \end{pmatrix} \qquad (12)$$
$$i = 1, \cdots, N$$

where, $\mathcal{G}_i(x, X_i, s) = (g_{i,jk}(x, X_i, s))$ is the distributed transfer matrix based on the information of channel $i$.

### 2.3  Data Reconciliation

In practice, the measured data called $Y_m$ is normally a superset of the data required to uniquely define the system. i.e., $Z_{givenBC} \subset Y_m \subseteq Z$. When this is

the case, we can use the information redundancy and apply data reconciliation to detect and handle the measurement errors. Data reconciliation requires a process model and statistical characteristics of the measurements.

Using modal decomposition, we are able to convert the dynamic model (6) to a "static" model, in which the measurable variables are linked by a algebraic relationship in the frequency domain:

$$P(s)Y(s) = 0 \tag{13}$$

where $Y(s) = [Y_1, Y_2, Y_3, \cdots] \subseteq Z(s)$ is a vector of noise free measurements, and where $P(s)$ is a sub-matrix of $M(s)$ with the appropriate dimension.

It is assumed that the measurements are independent and subject to an additive noise. The measured data $Y_m$ is composed of the "ideal" measurements vector $Y$ and a noise vector $\epsilon$ :

$$Y_m = Y + \epsilon \tag{14}$$

This noise vector $\epsilon$ is assumed to follow a Gaussian distribution with zero mean and **weight matrix** $W = \text{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2)$. Here $\sigma_i$ represents the noise standard deviation for each measurement. Note that the matrix $W$ is chosen (positive definite), and that its components are picked to have the error to be minimized, in the present case a weighted $L^2$ norm.

The objective of data reconciliation is to obtain estimated values $\hat{Y}$ close to the measurements $Y_m$ while satisfying the "static" linear model (13). This can be formulated as an optimization problem with linear constraints. The cost function to minimize is the weighted quadratic error between the measurements $Y_m$ and the reconciliated data $\hat{Y}$. The constraints are given by the model (13). The reconciliation problem in the spectral domain now becomes a least square problem with linear constraints. It reads:

$$\begin{align} \textbf{min.} \quad & f = (\hat{Y} - Y_m)^T W^{-1}(\hat{Y} - Y_m) \\ \textbf{s.t.} \quad & P\hat{Y} = 0 \end{align} \tag{15}$$

We hereby use the method suggested by [13] to solve the above data reconciliation problem. The constrained optimization problem is transformed into a corresponding unconstrained problem [6], using the Lagrange multiplier vector $\nu$ . The Lagrangian of the problem reads:

$$L(\hat{Y}, \nu) = (\hat{Y} - Y_m)^T W^{-1}(\hat{Y} - Y_m) + 2\nu^T P\hat{Y} \tag{16}$$

In order to obtain the unknown variables, take partial derivatives and set them

to zero:

$$\frac{\partial L}{\partial \hat{Y}} = 2(\hat{Y} - Y_m)^T W^{-1} + 2\nu^T P = 0$$
$$\frac{\partial L}{\partial \nu} = 2\, P\hat{Y} = 0 \qquad (17)$$

Rewrite the above equations as:

$$\begin{pmatrix} W^{-1} & P^T \\ P & \mathbf{0}_{\dim(\nu) \times \dim(\nu)} \end{pmatrix} \begin{pmatrix} \hat{Y} \\ \nu \end{pmatrix} = \begin{pmatrix} W^{-1} Y_m \\ \mathbf{0}_{\dim(\nu) \times 1} \end{pmatrix} \qquad (18)$$

Thus,

$$\hat{Y} = \begin{pmatrix} \mathbf{I}_{\dim(Y_m)} & \mathbf{0}_{\dim(Y_m) \times \dim(\nu)} \end{pmatrix} \begin{pmatrix} W^{-1} & P^T \\ P & \mathbf{0}_{\dim(\nu) \times \dim(\nu)} \end{pmatrix}^{-1} \begin{pmatrix} W^{-1} Y_m \\ \mathbf{0}_{\dim(\nu) \times 1} \end{pmatrix} \qquad (19)$$

where, matrices $\mathbf{I}$ and $\mathbf{0}$ are Identity Matrix and Zero Matrix of appropriate size. Note that this program can be solved numerically easily, with standard optimization software such as CPLEX, MATLAB, CVX etc[6].

The reconciliated measurements $\hat{Y}$, can then be used to obtain the desired internal boundary conditions using equation (10, 11, 12).

**Remark 1** *The proposed approach thus consists in assuming high confidence in the model, and finding the "best" estimation, i.e. the estimation which minimizes measurement error. This is a standard procedure in data reconciliation.*

## 3  Application to the Sacramento River

### 3.1  Description of the system and assumption

The Sacramento-San Joaquin Delta in California is a valuable 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 usual problems of inaccuracy and measurement errors for interested sensing systems. The area of interest for our experiment is located around the junction of the Sacramento River and the Georgiana Slough, as shown in Figure 2. 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. For experimental purposes, we need the boundary conditions at the three locations labeled A, B and C, but only get the measurements at SDC,
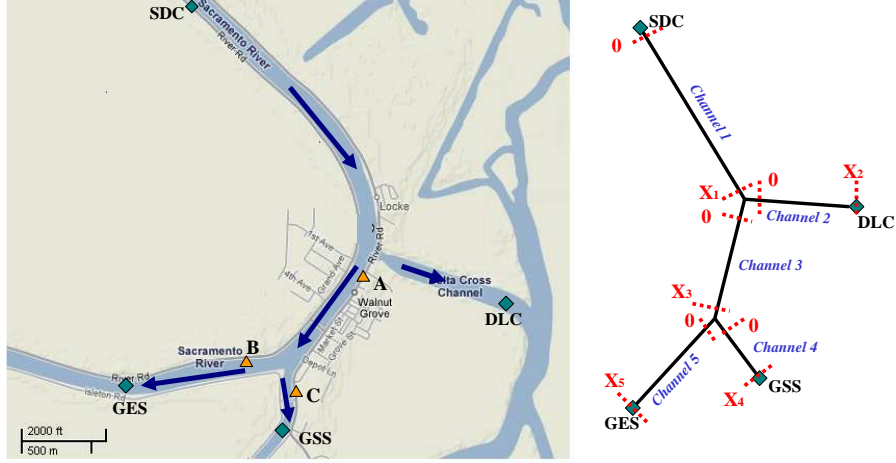
Fig. 2. *Test area in the Sacramento River and the Georgiana Slough.*

DLC, GSS and GES. The method described in the previous section enables us to do that.

Four USGS stations, named SDC, DLC, GES, and GSS, are located at the external boundaries of this deployment field. The stations are marked as squares in Figure 2. Both discharge and stage are collected every 900 seconds at these stations. Note that in the USGS measurement system, only the stage are measured directly, the discharge data is estimated by a rating curve, which is a relation between stream stage and streamflow. The relation of stream stage to streamflow is always changing, and need to be calibrated frequently. It will introduce errors if the rating curve has not been validated in time. More detailed information can be found at http://ga.water.usgs.gov/edu/measureflow.html. The field data was collected between 10/23/2007 and 11/13/2007. The raw field data is noisy, and the measurement errors are assumed to follow a normal Gaussian distribution. 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.

These assumptions have been verified in practice during experimental field deployments performed by our lab. The model parameters are the average free surface width $T_{0i}$, the average bottom slope $S_{bi}$, the average Manning's coefficient $n$, the average discharge $Q_i$, and the average downstream stage $Y_{Xi}$

for each channel $i$ ($i = 1, \cdots, 5$). These parameters are known to us experimentally. Based on measurements available to us at the *SDC, DLC, GSS*, and *GES*, the field data at three intermediate locations in the channel network are chosen to assess the accuracy of the method (at locations marked in triangles).

## 3.2   Modal Decomposition of the Measured Data

Since both the discharge and stage are measured at the four USGS stations (*SDC, DLC, GSS*, and *GES*), the measured flow variable vector $Y_m$ is:

$$Y_m = [q_1^m(0,t), y_1^m(0,t), q_2^m(X_2,t), y_2^m(X_2,t), q_4^m(X_4,t), y_4^m(X_4,t),$$
$$q_5^m(X_5,t), y_5^m(X_5,t)]^T \tag{20}$$

where $m$ stands for measured. The fundamental idea is to decompose the measured variables $Y_m$ into a finite sum of $N$ dominant oscillatory modes. In the case of a channel network influenced by the ocean at the downstream end, these modes are essentially the dominant modes produced by tidal forcing. The measured variables are therefore expressed using modal decomposition:

$$Y_m = \sum_{k=0}^{N} \left[ D_k e^{j\omega_k t} + \overline{D_k} e^{-j\omega_k t} \right] \tag{21}$$

where,

$$D_k = \left[ d_k^{(1,1,0)}, d_k^{(1,2,0)}, d_k^{(2,1,X_2)}, d_k^{(2,2,X_2)}, d_k^{(4,1,X_4)}, d_k^{(4,2,X_4)}, d_k^{(5,1,X_5)}, d_k^{(5,2,X_5)} \right]^T \tag{22}$$

$D_k = [d_k^{(\alpha,\beta,\gamma)}]^T$ are the Fourier coefficients of the spectral decomposition of $Y_m$, where $\alpha$, $\beta$, $\gamma$ represent the channel number, discharge/stage variable, location of each channel reach respectively. $\omega_k$'s are the set of frequencies used for modal decomposition.

Figure 3 shows the spectral analysis for the discharge data at station $SDC$: There are three dominant tidal frequencies in the system: $\omega_1 = 2.31 \times 10^{-5}\ s^{-1}$ (or period 12.4 *hrs* tide, corresponding to the M2 tide generated by the moon), $\omega_2 = 1.16 \times 10^{-5}\ s^{-1}$ (or period 24 *hrs* tide, corresponding to the K1 tide generated by the sun) and a $\omega_3 = 1.11 \times 10^{-5}\ s^{-1}$ (or period 25 *hrs* tide). The power spectrum is cut-off at $70 ft^3/s^2$ to determine the 30 dominant frequencies. The second plot in Figure 3 and Figure 4 indicates that 30 modes are sufficient to capture the signal. The amplitude at 0 Hz is essentially the nominal stage. Similar arguments hold for the other measurements.
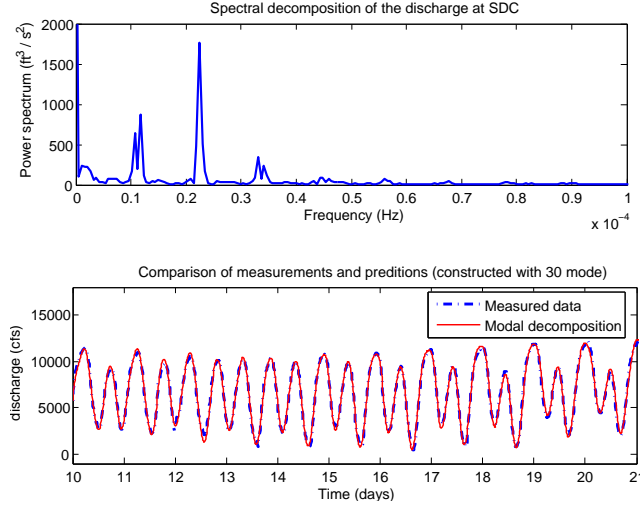
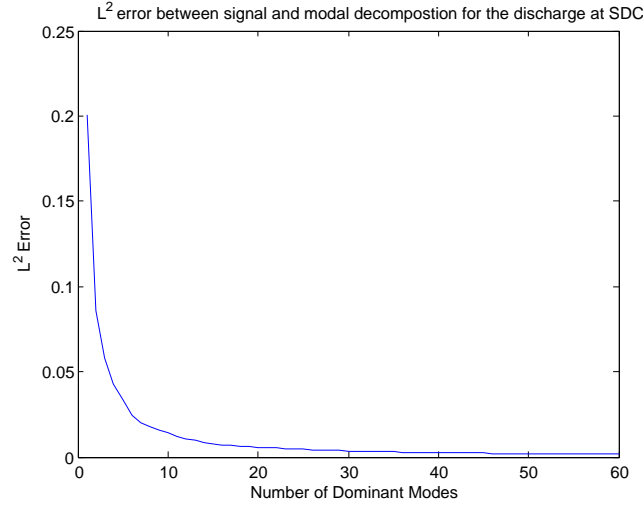Fig. 3. *Spectral analysis of the discharge at the SDC Station.*



Fig. 4. *Relative percent error between measurement data and modal representations as a function of the number of modes chosen for the decomposition.*

*3.3  Hydraulic Model of Sacramento River and Georgiana Slough*

The open-channel network system in this study here consists of five individual channels, as shown in Figure 2. For each channel, the discharge and stage at upstream and downstream are related by a non-uniform transfer matrix:

$$\begin{pmatrix} q_i(X_i, s) \\ y_i(0, s) \end{pmatrix} = \begin{pmatrix} g_{i,11}^n(X_i, X_i, s) & g_{i,12}^n(X_i, X_i, s) \\ g_{i,21}^n(0, X_i, s) & g_{i,22}^n(0, X_i, s) \end{pmatrix} \begin{pmatrix} q_i(0, s) \\ y_i(X_i, s) \end{pmatrix}$$

$$i = 1, \cdots, 5. \tag{23}$$

12

The linear relationships between internal boundary conditions at the two junctions are:

$$y_1(X_1, s) = y_2(0, s)$$
$$y_2(0, s) = y_3(0, s)$$
$$q_1(X_1, s) = q_2(0, s) + q_3(0, s)$$
$$y_3(X_3, s) = y_4(0, s)$$
$$y_4(0, s) = y_5(0, s)$$
$$q_3(X_3, s) = q_4(0, s) + q_5(0, s) \tag{24}$$

A total of twenty flow variables $q_i(x, s)$, $y_i(x, s)$ (for $x = 0$ or $X_i$, $i = 1, 2, \cdots, 5$) are included in the system (23) and (24). These flow variables are related in a linear model $M(s)Z(s) = 0$ equation (10), with

$$Z(s) = [q_1(0, s), y_1(0, s), q_1(X_1, s), y_1(X_1, s), q_2(0, s), y_2(0, s), q_2(X_2, s), y_2(X_2, s),$$
$$q_3(0, s), y_3(0, s), q_3(X_3, s), y_3(X_3, s), q_4(0, s), y_4(0, s), q_4(X_4, s), y_4(X_4, s),$$
$$q_5(0, s), y_5(0, s), q_5(X_5, s), y_5(X_5, s)]^T.$$

Here, $M(s)$ is a 16 by 20 matrix, which encodes the 16 equations comprised of (23) (five channels), and (24) (internal boundary conditions).
Since $\text{rank}(M(s)) = 16$, given four boundary flow variables $Z_{givenBC} \subset Z$, all the other sixteen boundary flow variables $Z_{otherBC} = Z \setminus Z_{givenBC}$ can be uniquely determined by the sixteen equations set (23) (24).
Let us assume that the four known external boundary conditions of the network are: the discharge at $SDC$: $q_1(0, s)$, the stage at $DLC$: $y_2(X_2, s)$, the stage at $GSS$: $y_4(X_4, s)$ and the stage at $GES$: $y_5(X_5, s)$. All the other boundary flow variables can be solved by equation (11). More specifically,

$$Z_{givenBC} = [q_1(0, s), y_2(X_2, s), y_4(X_4, s), y_5(X_5, s)]^T$$
$$Z_{otherBC} = [y_1(0, s), q_1(X_1, s), y_1(X_1, s), q_2(0, s), y_2(0, s), q_2(X_2, s), q_3(0, s), y_3(0, s),$$
$$q_3(X_3, s), y_3(X_3, s), q_4(0, s), y_4(0, s), q_4(X_4, s), q_5(0, s), y_5(0, s), q_5(X_5, s)]^T$$

$$R(s) = R_1(s)^{-1} R_2(s)^T$$

$$R_1(s) = \begin{pmatrix}
0 & 1 & -g_{1,12}(s) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & -g_{1,22}(s) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -g_{2,11}(s) & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -g_{2,21}(s) & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & g_{3,11}(s) & 0 & -1 & g_{3,12}(s) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & g_{3,21}(s) & 0 & -1 & g_{3,22}(s) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -g_{4,11}(s) & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -g_{4,21}(s) & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -g_{5,11}(s) & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -g_{5,21}(s) & 1 & 0 \\
0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & 0 & 0
\end{pmatrix}$$

$$R_2(s) = \begin{pmatrix}
g_{1,11}(s) & g_{1,21}(s) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & g_{2,12}(s) & g_{2,22}(s) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & g_{4,12}(s) & g_{4,22}(s) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g_{5,12}(s) & g_{5,22}(s) & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

Table 1
Parameters for the Sacramento River and the Georgiana Slough

| Channel | $Q_{0i}(\mathrm{m^3 s^{-1}})$ | $Y_{Xi}(m)$ | $T_{0i}(m)$ | $S_{bi}(\mathrm{m/km})$ | $n(\mathrm{m^{-1/3}s})$ | $X_i(m)$ | $C_0(m/s)$ | $L_0/C_0(\mathrm{s})$ |
|---|---|---|---|---|---|---|---|---|
| $i=1$ | 186.7 | 5.6 | 115 | -0.04 | 0.0323 | 2800 | 7.42 | 377.4 |
| $i=2$ | 83.9 | 4.1 | 110 | -0.09 | 0.0323 | 2000 | 6.30 | 317.5 |
| $i=3$ | 113.1 | 7.7 | 110 | -0.04 | 0.0323 | 1300 | 8.71 | 149.3 |
| $i=4$ | 58.1 | 4.0 | 56 | -0.19 | 0.0323 | 600 | 3.40 | 176.5 |
| $i=5$ | 65.2 | 5.3 | 89 | -0.04 | 0.0323 | 1600 | 7.19 | 222.5 |

The parameters of the model are listed in Table 1. The mean discharge $(Q_{0i})$ of the channels 1, 2, 4, 5 are computed using the measured discharge at *SDC, DLC, GSS, GES*, respectively. It is clear that the measurement

data are inconsistent, since $Q_{01} \neq Q_{02} + Q_{04} + Q_{05}$. To partially compensate for the measurement error, the mean discharge at channel 3 is set to be: $Q_{03} = [(Q_{01} - Q_{02}) + (Q_{04} + Q_{05})]/2$.

## 3.4 Data Reconciliation

Let us assume that the measured variables $Y_m$ are independent and subject to a Gaussian distributed noise. Based on the static model (11), the measurable variables are linked by a static relationship of the following form:

$$P(s)Y(s) = 0 \tag{25}$$

where

$$P(s) = \begin{bmatrix} R(s)_{1,1} & R(s)_{1,2} & R(s)_{1,3} & R(s)_{1,4} & -1 & 0 & 0 & 0 \\ R(s)_{6,1} & R(s)_{6,2} & R(s)_{6,3} & R(s)_{6,4} & 0 & -1 & 0 & 0 \\ R(s)_{13,1} & R(s)_{13,2} & R(s)_{13,3} & R(s)_{13,4} & 0 & 0 & -1 & 0 \\ R(s)_{16,1} & R(s)_{16,2} & R(s)_{16,3} & R(s)_{16,4} & 0 & 0 & 0 & -1 \end{bmatrix}$$

given $Y(s) \setminus Z_{givenBC}$ is the first, sixth, thirteenth and sixteenth element of $Z_{otherBC}$. Now, combining the solution of the data reconciliation problem (19) with the static model (25), reconciliated measurements $\hat{Y}$ can be calculated. Assume that $\hat{Y}$ is in the form:

$$\hat{Y} = \sum_{k=0}^{N} \left[ B_k e^{j\omega_k t} + \overline{B_k} e^{-j\omega_k t} \right] \tag{26}$$

where $B_k = [b_k^{(\alpha,\beta,\gamma)}]^T$ is the Fourier coefficients vector of the spectral decomposition of $\hat{Y}$, and $\alpha$, $\beta$, $\gamma$ represent the channel number, discharge/stage variable, location of each channel reach respectively:

$$B_k = \left[ b_k^{(1,1,0)}, b_k^{(1,2,0)}, b_k^{(2,1,X_2)}, b_k^{(2,2,X_2)}, b_k^{(4,1,X_4)}, b_k^{(4,2,X_4)}, b_k^{(5,1,X_5)}, b_k^{(5,2,X_5)} \right]^T \tag{27}$$

For specific dominant $\omega_k$, $\quad k = 1, \cdots, N$, the coefficient vector $B_k$ in the equation (26) is calculated by equation (19):

$$B_k = \left( \mathbf{I}_{8,8} \; \mathbf{0}_{8,4} \right) \begin{pmatrix} W^{-1} & P(s)^T \\ P(s) & \mathbf{0}_{4,4} \end{pmatrix}^{-1} \begin{pmatrix} W^{-1} D_k \\ \mathbf{0}_{4,1} \end{pmatrix} \tag{28}$$

The reconciliated boundary condition data is shown and compared to measured data in Figures 5 and 6. For clarity, the mean flow has been subtracted from the plots in the interest of magnifying the display scale. From the figures,

the reconciliated data is very close to the measurements. The difference between the reconciliated data and measurements is further evaluated in Table 2. Three primary evaluation measures are analyzed here:

- The maximum value is the maximum difference between the reconciliated and measured data at the same time steps.
- The coefficient of efficiency $E$ is defined as [18]:

$$E = 1 - \left[ \frac{\sum_{i=1}^{N}(\hat{u}_i - u_i)^2}{\sum_{i=1}^{N}(u_i - \overline{u_i})^2} \right] \tag{29}$$

where $u_i$ is the flow variable of interest (for example $q_i$ or $y_i$ in this study), $\hat{u}_i$ is the reconciliated/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 last statistic evaluation of the analysis is the correlation coefficient $\rho$, given by:

$$\rho = \frac{\sum_{1}^{N}(u_i - \overline{u_i})(\hat{u}_i - \overline{\hat{u}_i})}{\sqrt{\sum_{1}^{N}(u_i - \overline{u_i})^2 \sum_{1}^{N}(\hat{u}_i - \overline{\hat{u}_i})^2}} \tag{30}$$

where $\overline{\hat{u}_i}$ represents the mean of reconciliated flow for $i = 1$ $to$ $N$ measurement events.

Table 2
Max-value, $\rho$-value and $E$-value for reconciliated data and measured data

| Variable | $USGS$ Station | Max-value | $E$-value | $\rho$-value |
|---|---|---|---|---|
| Discharge | SDC | 23.66 $m^3/s$ | 0.9930 | 0.9975 |
| | DLC | 28.23 $m^3/s$ | 0.9368 | 0.9883 |
| | GES | 13.00 $m^3/s$ | 0.9968 | 0.9985 |
| | GSS | 18.41 $m^3/s$ | 0.9368 | 0.8369 |
| Stage | SDC | 0.05$m$ | 0.9889 | 0.9947 |
| | DLC | 0.12$m$ | 0.9504 | 0.9759 |
| | GES | 0.07 $m$ | 0.9847 | 0.9935 |
| | GSS | 0.05 $m$ | 0.9938 | 0.9989 |

*3.5   Method Validation*

We used existing USGS sensors in place in the Delta as measurement points, and deploy our own sensors at selected locations to produce data used for the validation. We validate the method by using existing and deployable monitoring infrastructure: USGS fixed sensor stations (see Figure 7) are used as
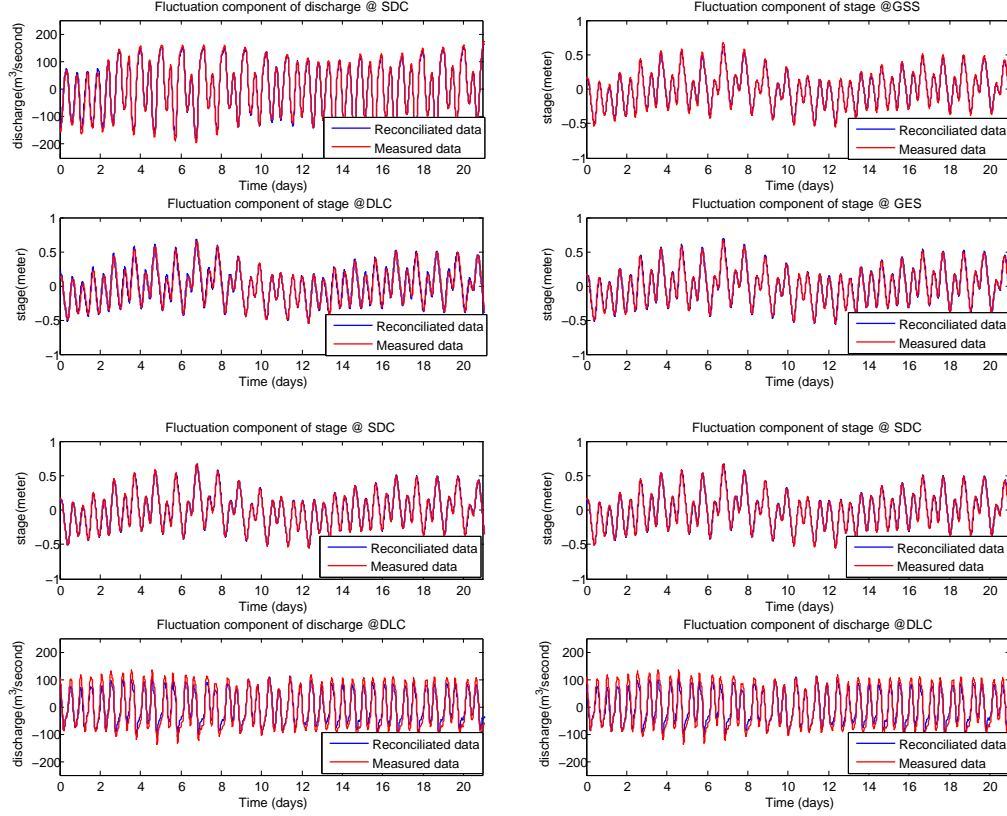
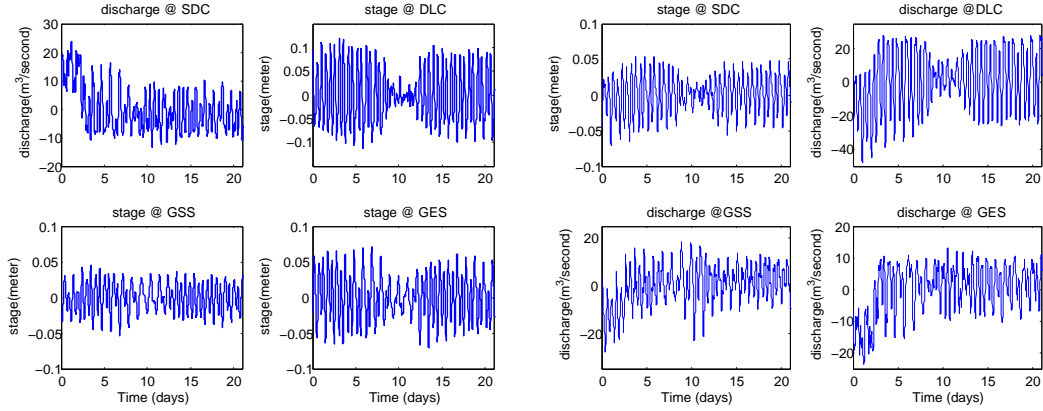Fig. 5. *Reconciliated boundary condition data vs. measured data.*



Fig. 6. *The difference between Reconciliated data and measured data.*

measurement points (see exact location in Figure 2); deployable UC Berkeley sensors (see Figure 7) are placed at locations A, B, C on the map of Figure 2. The measurements were collected between 11/01/2007 and 11/12/2007, and serve as a validation data set for this method.

Location A is downstream of the junction of Sacramento river and Delta Cross channel; Location B is downstream of GSS branch; Location C is downstream of Sacramento Branch. Without loss of generality, the discharge at Location A,

17

Fig. 7. **Left:** *USGS Sensor station at* GSS, *used as a measurement sensor.* **Right:** *Deployable ADCP sensor, used in Section 3.5 for gathering the validation data (three of them were deployed between 11/01/2007 and 11/12/2007 in order to gather the data for this study).*
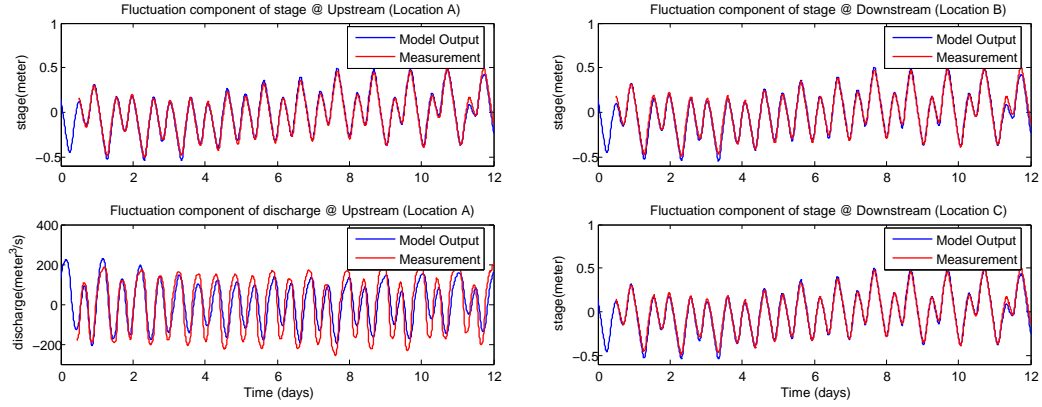


Fig. 8. *Validation of the model output with measurement using USGS measurements.*

along with the stage data at three locations, are used to test the model.

### 3.5.1 *Model validation with measured boundary conditions*

Following the steps described in Section II, the flow variables at the boundaries of each branch $Z_{otherBC}$ are calculated using Equation (11). Here, the measured discharge at SDC, stage at DLC, GES, GSS are used as $Z_{givenBC}$. The flow variables along each branch are estimated using the non-uniform transfer matrix (Equation (12)). The simulation results are shown in Figure 8.

Model calibration and validation are further evaluated using $E$-value and $\rho$-value. Table 3 summarizes the values of $\rho$ and $E$ in the validation sets of our channel flow model.

From Figure 8 and Table 3, it is clear that Location A is the location where the discharge is estimated withe least precision, as the characteristics of either

18

Table 3

$\rho$-value and $E$-value for model validation without reconciliation (Section 3.5.1)

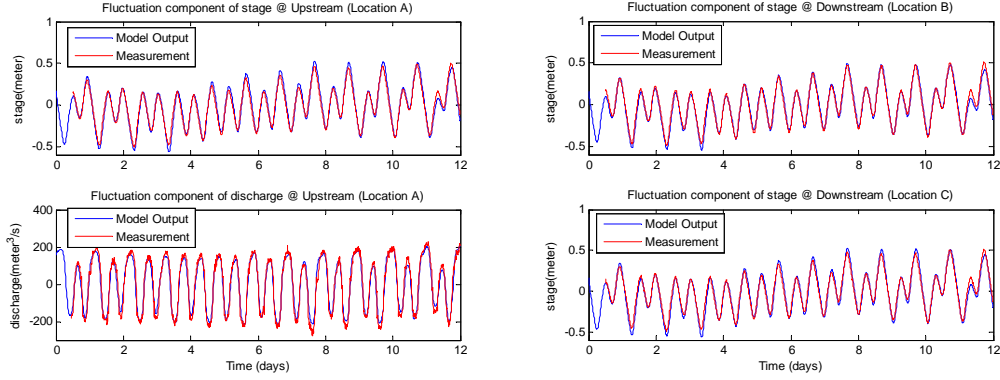| Location | A | A | B | C |
|----------|-----------|--------|--------|--------|
| Variables | discharge | stage | stage | stage |
| $E$ | 0.7219 | 0.9820 | 0.9796 | 0.9807 |
| $\rho$ | 0.8555 | 0.9922 | 0.9916 | 0.9927 |



Fig. 9. *Validation of the model output with measurement using reconciliated BC.*

the phase or the amplitude suffer a significant offset. Although we have a precise stage estimation at Locations A, B, C, the model does not provide enough information to characterize the flow in the experiment area.

### 3.5.2   Modal validation with reconciliated boundary conditions based on all the measurements

We will use reconciliated data shown in Figure 5 as $Z_{givenBC}$. The flow variables $Z_{otherBC}$, $q_i(s, x)$, $y_i(x, s)$ are calculated using Equations (11) and (12). The simulation results are shown in Figure 9.

The values of $\rho$ and $E$ are listed in Table 4. Both $\rho$-values and $E$-values are

Table 4

$\rho$-value and $E$-value for model validation after reconciliation (Section 3.5.2)

| Location | A | A | B | C |
|----------|-----------|--------|--------|--------|
| Variables | discharge | stage | stage | stage |
| $E$ | 0.9775 | 0.9643 | 0.9768 | 0.9612 |
| $\rho$ | 0.9895 | 0.9876 | 0.9897 | 0.9875 |

close to unity. Table 4 and Figure 9 thus indicate that the proposed model reconciliation approach provides a higher accuracy in the flow estimation.
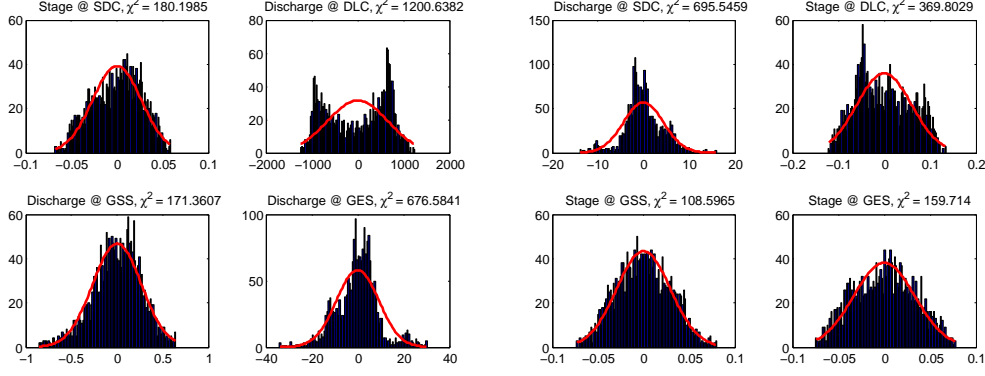
Fig. 10. *Error distribution between the reconciliated data and USGS measurements (perturbation discharge at DLC is added on purpose in the data set).*

### 3.5.3 Exclusion of erroneous sensors

The data reconciliation method enable us to detect and exclude sensors with erroneous measurements and further estimate the boundary conditions using the remaining properly working sensors. Based on known results in data reconciliation, when the reconciliated data is close to the "true" data, the difference between the reconciliated and measured data must follow a Gaussian distribution with a zero mean. If this condition is not satisfied, the sensor is deemed to be malfunctioning and not suitable for measurements.

In order to test the performance of our method, we first intentionally add a large (10 times in magnitude) perturbation to the discharge data measured at DLC, and conduct the standard data reconcliliation procedure described in preceding sections. The probability density function of the difference between the reconciliated data and measured data $(\hat{Y} - Y_m)$ is calculated and compared in Figure 10. A Pearson's chi-square $(\chi^2)$ test is further applied to assess whether this probability distribution differs from a theoretical Gaussian distribution [19] [3]. The Pearson's $\chi^2$ statistic is calculated as:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \tag{31}$$

where,

$\chi^2 =$ the test statistic that asymptotically approaches a $\chi^2$ distribution
$O_i =$ an observed frequency
$E_i =$ an theoretical (Gaussian) frequency
$n =$ the number of possible outcomes of each event

From the data (see in particular subfigure in Figure 10), it is visually obvious that the difference between the measured and reconciliated discharge at station DLC does not follow the Gaussian distribution, implying that the sensor
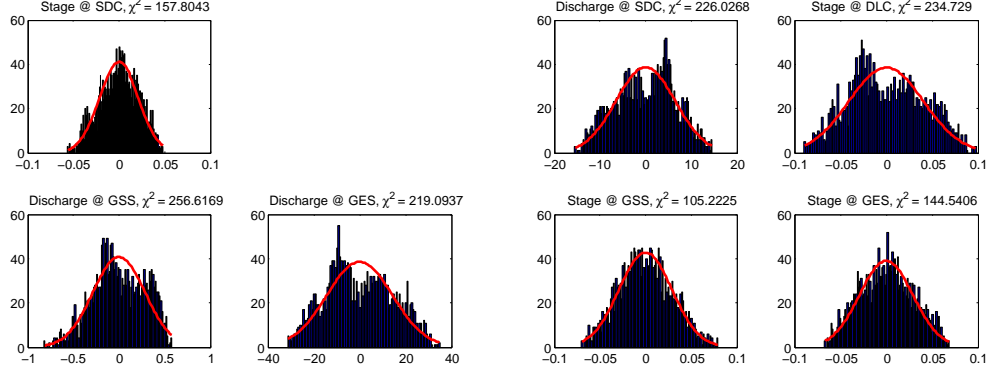
Fig. 11. *Error distribution between reconciliated and measured data (Discharge at DLC is removed from the data reconciliation process).*

is malfunctioning and should be excluded from the data sets.

We then remove the discharge at DLC from the measurement data, and repeat the data reconciliation procedure. The probability density functions are showed in Figure 11. The difference between the reconciliated data and USGS measurements is then more likely to follow a standard Gaussian distribution. In summary, the $\chi^2$ values are evaluated and assembled in Table 5. Note that a small $\chi^2$-value indicates that the observation data distribution is likely to follow a normal distribution.

Table 5
$\chi^2$ value for different cases.

| Variables | with wrong DLC-discharge data | without DLC discharge data |
|:---:|:---:|:---:|
| discharge at SDC | 695.55 | 157.80 |
| stage at SDC | 180.20 | 226.03 |
| discharge at DLC | 1200.64 | N/A |
| stage at DLC | 369.80 | 234.73 |
| discharge at GSS | 171.36 | 256.62 |
| stage at GSS | 106.60 | 105.22 |
| discharge at GES | 675.58 | 219.09 |
| stage at GES | 159.71 | 144.54 |

Furthermore, we use the same three locations (A, B, C) in the experiment domain to validate the reconciliated data. Table 6 compares the $\rho$-value and $E$-value for the two cases: with the erroneous sensor and without the erroneous sensor.

It is rather interesting to note that modal output is not affected by the intentionally perturbed sensor data, meaning that the data reconciliation method

Table 6

$\rho$-value and $E$-value for modal validation (with and without discharge @ DLC).

| Location | A | A | B | C |
|---|---|---|---|---|
| Variables | discharge | stage | stage | stage |
| $E$ (with wrong sensor) | 0.9777 | 0.9599 | 0.9762 | 0.9567 |
| $E$ (without wrong sensor) | 0.9676 | 0.9651 | 0.9788 | 0.9611 |
| $\rho$ (with wrong sensor) | 0.9889 | 0.9867 | 0.9894 | 0.9867 |
| $\rho$ (without wrong sensor) | 0.9893 | 0.9892 | 0.9908 | 0.9891 |

is robust enough to provide satisfactory boundary conditions even if one of the sensors is malfunctioning.

## 4 Conclusions

This article proposes a new method to estimate the flow variables in a channel network system subject to periodic forcing. A spatially-dependent channel network model is constructed in the frequency domain using LSWE transfer matrix for the non-uniform steady state case. Modal decomposition allows the output response to be expressed in terms of the spectral coefficients of the input variables and the transfer matrix coefficients evaluated at appropriate locations. Data reconciliation in this case is reduced to a static least-square minimization problem in the frequency domain, and enables a efficient reconstruction of noisy boundary measurements. Subsequently, the flow properties at any location in the system can be readily predicted. The approach proposed in this study has been applied to a channel network in the Sacramento-San Joaquin Delta, using four USGS fixed sensors as measurement points, the flow prediction being successfully validated at three intermediate locations of the channel system, using deployed sensors from UC Berkeley.

This method is now used for short term forecast of internal condition in the Georgiana Slough and Sacramento River, which we use for our experimental drifter and submarine deployments. This information is particularly useful for our ongoing data assimilation and inverse modeling studies currently underway, using Lagrangian sensors.

## Acknowledgments

**APPENDIX**

**Transfer Matrix Representation of Linearized Saint-Venant Model**
The application of Laplace transform to the linear PDE system (6) leads to the *ordinary differential equations* (ODEs) in the variable $x$, with a complex parameter $s$.

$$\begin{pmatrix} q(x,s) \\ y(x,s) \end{pmatrix}_x = \mathcal{G}_s(x) \begin{pmatrix} q(x,s) \\ y(x,s) \end{pmatrix} \tag{A-1}$$

In the case of designing and analyzing small-scale flood management systems, such as storm sewers and highway drainage, *Uniform Flow* is a good simplified model which assume the channel invert slope and the energy grade line slope are equal ($S_b = S_f$). However, in real-world situation, this assumption is rarely represented, and backwater approximation is introduced to simulate the *Non-uniform Flow*.

- *Transfer Matrix for Uniform Flow*
  **Remark 2 (Uniform flow)** *In the case of uniform flow, the flow variables are constant along the length of the channel, i.e., the discharge $Q_0(x) = Q_0 = Q_X$ and the stage $Y_0(x) = Y_n$ (normal depth).*
  A closed-form solution of the linearized Saint-venant equations in the uniform flow case can be obtained, relating the flow variables at any point $x$ of the river reach $q(x,s)$, $y(x,s)$ to the boundary conditions $q(0,s)$ and $y(X,s)$ (referring to [15] for details).

$$\begin{pmatrix} q(x,s) \\ y(x,s) \end{pmatrix} = \begin{pmatrix} g_{11}^u(x,X,s) & g_{12}^u(x,X,s) \\ g_{21}^u(x,X,s) & g_{22}^u(x,X,s) \end{pmatrix} \begin{pmatrix} q(0,s) \\ y(X,s) \end{pmatrix} \tag{A-2}$$
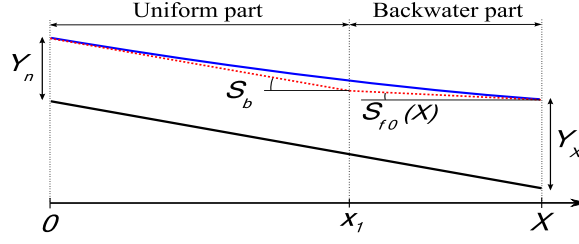
Fig. 12. *Backwater curve approximation*

where,

$$g_{11}^u(x, X, s) = \frac{\lambda_2 e^{\lambda_1 x + \lambda_2 X} - \lambda_1 e^{\lambda_2 x + \lambda_1 X}}{\lambda_2 e^{\lambda_2 X} - \lambda_1 e^{\lambda_1 X}}$$

$$g_{12}^u(x, X, s) = T_0 s \frac{e^{\lambda_1 x} - e^{\lambda_2 x}}{\lambda_2 e^{\lambda_2 X} - \lambda_1 e^{\lambda_1 X}}$$

$$g_{21}^u(x, X, s) = \frac{\lambda_1 \lambda_2}{T_0 s} \frac{e^{\lambda_2 x + \lambda_1 X} - e^{\lambda_1 x + \lambda_2 X}}{\lambda_2 e^{\lambda_2 X} - \lambda_1 e^{\lambda_1 X}}$$

$$g_{22}^u(x, X, s) = \frac{\lambda_2 e^{\lambda_2 x} - \lambda_1 e^{\lambda_1 x}}{\lambda_2 e^{\lambda_2 X} - \lambda_1 e^{\lambda_1 X}}$$

Here, $\lambda_1$ and $\lambda_2$ are the eigenvalues of the ODE system (A-1), and are given by:

$$\lambda_i(s) = \frac{2T_0 V_0 s + \gamma_0}{2\alpha_0} + (-1)^i \frac{\sqrt{4C_0^2 T_0^2 s^2 + 4T_0(V_0\gamma_0 - \alpha_0\beta_0)s + \gamma_0^2}}{2\alpha_0} \qquad \text{(A-3)}$$

In the following sections, the transfer matrix for the uniform case is denoted as $\mathcal{G}^u(x, X, s) = (g_{ij}^u(x, X, s))$.

- *Transfer Matrix for Non-uniform Flow*

To solve a hydraulic problem in realistic cases, the backwater approximation is assumed to study flow regimes, in which the water elevation is not constant along the reach [22]. Following the method in [22], and further modified in [16], the backwater curve defined by equation (A-5) is approximated by two straight lines, as shown in Figure (12).

$$\frac{dQ_0(x)}{dx} = 0 \qquad \text{(A-4)}$$

$$\frac{dY_0(x)}{dx} = \frac{S_b - S_{f0}}{1 - F_0(x)^2} \qquad \text{(A-5)}$$

The river reach is then decomposed into two parts: a uniform part and a backwater part. The intersection of the two parts is denoted by $x_1$. Let $x_u$ denote the location in the uniform part, $x_b = x - x_1$ denote the location in the backwater part, $X_u = x_1$ denote the length of the uniform part, and $X_b = X - x_1$ denote the length of the backwater part. Let $\mathcal{G}^u(x_u, X_u, s)$ and $\mathcal{G}^b(x_b, X_b, s)$ denote the transfer matrices for the uniform and backwater parts respectively; while $\mathcal{G}^b(x_b, X_b, s)$ has the same form as the trans-

24

fer matrix $\mathcal{G}^u(x_u, X_u, s)$. The transfer matrix for the non-uniform channel $\mathcal{G}^n(x, X, s) = (g_{ij}^n(x, X, s))$ compromises $\mathcal{G}^u(x_u, X_u, s)$ and $\mathcal{G}^b(x_b, X_b, s)$ with:

$$\begin{pmatrix} q(x, s) \\ y(x, s) \end{pmatrix} = \begin{pmatrix} g_{11}^n(x, X, s) & g_{12}^n(x, X, s) \\ g_{21}^n(x, X, s) & g_{22}^n(x, X, s) \end{pmatrix} \begin{pmatrix} q(0, s) \\ y(X, s) \end{pmatrix} \tag{A-6}$$

The detail of entries of the transfer matrix for the non-uniform case $\mathcal{G}^n(x, X, s) = (g_{ij}^n(x, X, s))$ is listed in [24].

# 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] M. ALHAJ-DIBO, D. MAQUIN, and J. RAGOT. Data reconciliation: A robust approach using a contaminated distribution. *Control Engineering Practice*, 16:159–170, 2008.

[3] M. J. BAGAJEWICZ and E. CABRERA. Data reconciliation in gas pipeline systems. *Industrial and Engineering Chemistry Research*, 42(22):5596–5606, 2003.

[4] J. P. BAUME and J. SAU. Study of irrigation cannal dynamics for control purposes. In *International Workshop in Regulation of Irrigation Canals*, Marrakech, Morroco, 1997.

[5] N. BEDJAOUI, X. LITRICO, D. KOENIG, J. RIBOT-BRUNO, and P.-O. MALATERRE. Static and dynamic data reconciliation for an irrigation canal. *Journal of Irrigation and Drainage Engineering*, in press, 2008.

[6] S. BOYD and L. VANDENBERGHE. *Convex Optimization*. Cambridge University Press, New York, 2004.

[7] V. CHOW. *Open-channel Hydraulics*. McGraw-Hill Book Company, New York, 1988.

[8] C. K. CHUI and G. CHEN. *Kalman filtering with real time applications*. Springer-Verlag, Berlin, 1998.

[9] R. CORRIGA, F. F. PATTA, S. SANNA, and G. USAI. A mathematical model for open channel networks. *Applied Mathematical Modelling*, 23:51–54, 1979.

[10] C. M. CROWE. Data reconciliation–progress and challenge. *Journal of Process Control*, 6(2):89–98, 1996.

[11] J. DELTOUR, E. CANIVET, F. SANFILIPPO, and J. SAU. Data reconciliation on the complex hydraulic system of canal de provence. *Journal of Irrigation and Drainage Engineering*, 131(3):291–297, 2005.

[12] Y. Ermolin. Study of open-channel dynamics as controlled process. *Journal of Hydraulic Engineering*, 119(1):59–71, 1992.

[13] G. Heyen, M. N. Dumont, and B. Kalitventzeff. Computer-aided design of redundant sensor networks. *Computers and Chemical Engineering*, 10:685–690, 2002.

[14] M. T. Lee and J. W. Delleur. A variable source area model of the rainfall-runoff process based on the watershed stream network. *Water Resources Research*, 12(5):1029–1036, 1976.

[15] X. Litrico and V. Fromion. Frequency modeling of open channel flow. *Journal of Hydraulic Engineering*, 130(8):806–815, 2004.

[16] X. Litrico and V. Fromion. Simplified modeling of irrigation canals for controller design. *Journal of Irrigation and Drainage Engineering*, 130(5):373–383, 2004.

[17] X. Litrico and V. Fromion. Boundary control of linearized saint-venant equations oscillating modes. *Automatica*, 42(6):967–972, 2006.

[18] 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.

[19] R.L. Plackett. Karl pearson and the chi-squared test. *International Statistical Review*, 51(1):59–72, 1983.

[20] J. Ragot and D. Maquin. Reformulation of data reconciliation problem with unknown-but-bounded errors. *Industrial and Engineering Chemistry Research*, 43(6):1530–1536, 2004.

[21] C. Reimers, J. Werther, and G. Gruhn. Flowsheet simulation of solids processes data reconciliation and adjustment of model parameters. *Chemical Engineering and Processing*, 47:138–158, 2008.

[22] J. Schuurmans, A. J. Clemmens, S. Dijkstra, A. Hof, and R. Brouwer. Modeling of irrigation and drainage canals for controller design. *Journal of Irrigation and Drainage Engineering*, 125(6):338–344, 1999.

[23] 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.

[24] Q. Wu, S. Amin, S. Munier, A. Bayen, X. Litrico, and G. Belaud. Parameter identification for the shallow water equations using modal decomposition. In *Proceedings of the 46th Conference of Decision and Control*, 2007.

[25] Q. Wu, X. Litrico, and A. Bayen. Open channel flow estimation and data reconciliation using modal decomposition. In *Proceedings of the 46th Conference of Decision and Control*, 2009.