Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Data Reconciliation of an Open Channel Flow Network using Modal Decomposition

Qingfang Wu*, Xavier Litrico† and Alexandre M. Bayen*
*Department of Civil and Environmental Engineering, UC Berkeley, CA, USA
{qingfangwu,bayen}@berkeley.edu
†Cemagref, UMR G-EAU, B.P. 5095, 34196 Montpellier Cedex 5, France
{xavier.litrico}@cemagref.fr

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 input; the flow properties at selected internal locations, as well as the other external boundary conditions, are defined as the output. Both inputs and outputs are affected by noise and we use the model to re-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 placed in the Delta as measurement points, and deploy 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.**

I. 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 [2] [3] [4] [5] which has been applied in several engineering fields [6] [7] [8]. The objective of the 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 [9]. However, in certain cases, and for given time intervals, dynamic effects can sometimes be neglected, leadin to simplified versions of the general approach, applicable to static models.

This article presents theoretical results applicable to data reconciliation for tidally forced network 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 [10] [11] [12]. With the assumption of a backwater curve model [13], a more realistic transfer matrix function has been introduced [14], which is used 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, e.g., 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.

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

II. Proposed Method

A. 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.

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 redundancy 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.

B. Linear Channel Network Model

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 [15], [16], [17], [18]. 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), $S_f(x,t)$ is the friction slope (m/m) modeled by Manning-Strickler's formula (6), with n is the Manning's roughness coefficient $(sm^{-1/3})$ [16].

$$S_f=\frac{Q^2n^2(T+2Y)^{4/3}}{(TY)^{10/3}} \tag{6}$$

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:

$$\frac{\partial}{\partial t}\begin{pmatrix}q(x,t)\\y(x,t)\end{pmatrix}=\left(\mathcal{A}(x)\frac{\partial}{\partial x}+\mathcal{B}(x)\right)\begin{pmatrix}q(x,t)\\y(x,t)\end{pmatrix} \tag{7}$$
$$(x,t)\in[0,X]\times[0,+\infty)$$

where,

$$\mathcal{A}(x)=\begin{pmatrix}-2V_0 & -\alpha_0\\ -\frac{1}{T_0} & 0\end{pmatrix}\qquad \mathcal{B}(x)=\begin{pmatrix}\beta_0 & \gamma_0\\ 0 & 0\end{pmatrix} \tag{8}$$

The application of Laplace transform to the linear PDE system (8) leads to the following *ordinary differential equations* (ODEs) in the variable x, with a complex parameter s.

$$\frac{d}{dx}\begin{pmatrix}q(x,s)\\y(x,s)\end{pmatrix}=\mathcal{A}^{-1}(x)\left[\mathcal{B}(x)-s\mathbb{I}_2\right]\begin{pmatrix}q(x,s)\\y(x,s)\end{pmatrix} \tag{9}$$

Following the method developed in [13], and further modified in [14], 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:

$$\begin{pmatrix}q(x,s)\\y(x,s)\end{pmatrix}=\mathcal{G}(x,X,s)\begin{pmatrix}q(0,s)\\y(X,s)\end{pmatrix} \tag{10}$$

$\mathcal{G}(x,X,s)$ 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) Transfer Matrix Model for Channel Networks: The model (10) 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 data assimilation work, in which we need estimations 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 (10) 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.

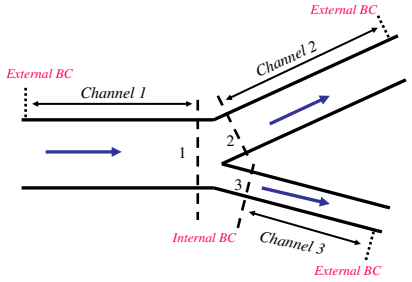

Fig. 1. *Representation the system of interest for estimating internal boundary conditions (Internal BC) using data reconciliation on external boundary conditions (External BC): data is given at the three external conditions of channel* 1, 2 *and* 3*; the state of the system is computed at three internal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

- **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:

$$q_1(X_1,s) = q_2(0,s) + q_3(0,s) \tag{11}$$

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:

$$y_1(X_1,s) = y_2(0,s) = y_3(0,s) \tag{12}$$

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:

$$M(s)Z(s) = 0 \tag{13}$$

where $Z(s)$ is the concatenated vector of all $[q_i(0,s), q_i(X_i,s), y_i(0,s), y_i(X_i,s)]^T$, $(i = 1,\cdots,N)$; $Z(s)$ 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 (13). 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 (13). Model (13) now has the form:

$$Z_{otherBC} = R(s) Z_{givenBC} \tag{14}$$

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

$$\begin{pmatrix} q_i(x,s) \\ y_i(x,s) \end{pmatrix} = \begin{pmatrix} g_{i,11}(x,X_i,s) & g_{i,12}(x,X_i,s) \\ g_{i,21}(x,X_i,s) & g_{i,22}(x,X_i,s) \end{pmatrix} \begin{pmatrix} q_i(0,s) \\ y_i(X_i,s) \end{pmatrix} \tag{15}$$

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, $i = 1,\cdots,N$.

C. 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 (8) to a "static" model, in which the measurable variables are linked by an algebraic relationship in the frequency domain:

$$P(s)Y(s) = 0 \tag{16}$$

where $Y(s) = [Y_1, Y_2, Y_3, \cdots] \subseteq Z(s)$ is a vector of noise free measurements, $P(s)$ 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 ϵ :

$$Y_m = Y + \epsilon \tag{17}$$

This noise vector ϵ 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 σ_i is the noise standard deviation for each measurement.

The objective of data reconciliation is to obtain estimated values $\hat{Y}$ close to the measurements Y_m while satisfying the "static" linear model (16). 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 (16). The reconciliation problem in the spectral domain now becomes a least square problem with linear constraints. It reads:

$$\begin{aligned} \textbf{min.} \quad & f = (\hat{Y} - Y_m)^T W^{-1} (\hat{Y} - Y_m) \\ \textbf{s.t.} \quad & P\ \hat{Y} = 0 \end{aligned} \tag{18}$$

This quadratic program can be solved by a variety numerical tools, in particular CPLEX, MATLAB, CVX etc [19]. We hereby use the method suggested by [20] to solve the above data reconciliation problem. The constrained optimization problem is transformed into a corresponding unconstrained problem [19], using the Lagrange multiplier vector ν . The Lagrangian of the problem reads:

$$L(\hat{Y}, \nu) = (\hat{Y} - Y_m)^T W^{-1} (\hat{Y} - Y_m) + 2\nu^T P\hat{Y} \tag{19}$$

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

$$\begin{aligned} \frac{\partial L}{\partial \hat{Y}} &= 2(\hat{Y} - Y_m)^T W^{-1} + 2\nu^T P = 0 \\ \frac{\partial L}{\partial \nu} &= 2\ P\hat{Y} = 0 \end{aligned} \tag{20}$$

Rewrite the above equations as:

$$\begin{pmatrix} W^{-1} & P^T \\ P & \mathbf{0}_{\dim(\nu)\times\dim(\nu)} \end{pmatrix} \begin{pmatrix} \hat{Y} \\ \nu \end{pmatrix} = \begin{pmatrix} W^{-1} Y_m \\ \mathbf{0}_{\dim(\nu)\times 1} \end{pmatrix} \tag{21}$$

Thus,

$$\hat{Y} = \begin{pmatrix} \mathbf{I}_{\dim(Y_m)} \\ \mathbf{0}_{\dim(\nu)\times\dim(Y_m)} \end{pmatrix}^T \begin{pmatrix} W^{-1} & P^T \\ P & \mathbf{0}_{\dim(\nu)\times\dim(\nu)} \end{pmatrix}^{-1} \begin{pmatrix} W^{-1} Y_m \\ \mathbf{0}_{\dim(\nu)\times 1} \end{pmatrix} \tag{22}$$

where, matrices **I** and **0** are Identity Matrix and Zero Matrix of appropriate size.
The reconciliated measurements $\hat{Y}$, can then be used to obtain the desired internal boundary conditions using equation (13, 14, 15).

III. Application to the Sacramento River

A. 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 USGS stations, *SDC*, *DLC*, *GSS*, and *GES*. The method described in the previous section enables us to do that.

Four USGS stations, named *SDC*, *DLC*, *GSS*, and *GES*, 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. 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:

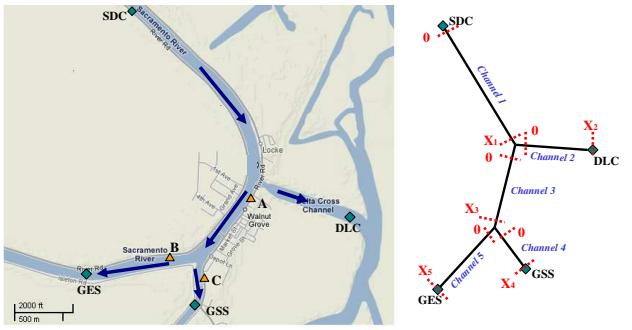

Fig. 2. *Test area in the Sacramento River and the Georgiana Slough.*

- 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).

B. 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:

$$\begin{aligned} Y_m = & [q_1^m(0,t), y_1^m(0,t), q_2^m(X_2,t), y_2^m(X_2,t), \\ & q_4^m(X_4,t), y_4^m(X_4,t), q_5^m(X_5,t), y_5^m(X_5,t)]^T \end{aligned} \tag{23}$$

where notation 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

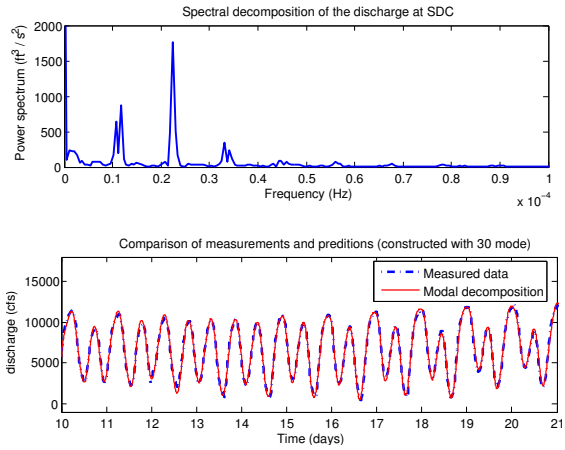

Fig. 3. *Spectral analysis of the discharge at the SDC Station.*

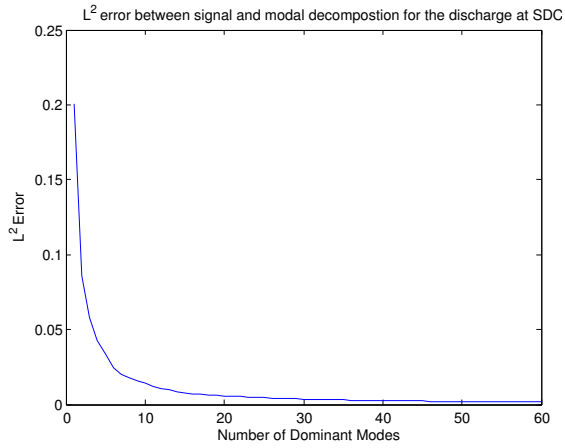

Fig. 4. *Relative percent error between measurement data and modal represtentions as a function of the number of modes chosen for the decomposition.*

tidal forcing. The measured variables are therefore expressed using modal decomposition:

$$Y_m = \sum_{k=0}^{N} \left[D_k e^{j\omega_k t} + \overline{D_k} e^{-j\omega_k t} \right] \tag{24}$$

where,

$$D_k = [d_k^{(1,1,0)}, d_k^{(1,2,0)}, d_k^{(2,1,X_2)}, d_k^{(2,2,X_2)}, d_k^{(4,1,X_4)}, d_k^{(4,2,X_4)}, d_k^{(5,1,X_5)}, d_k^{(5,2,X_5)}]^T \tag{25}$$

$D_k = [d_k^{(\alpha,\beta,\gamma)}]^T$ are the Fourier coefficients of the spectral decomposition of Y_m, where α, β, γ represent the channel number, discharge/stage variable, location of each channel , respectively. ω_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 indicate 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.

C. Hydraulic Model of Sacramento River and Georgiana Slough

The open-channel network system in this study 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:

$$\begin{pmatrix} q_i(X_i,s) \\ y_i(0,s) \end{pmatrix} = \begin{pmatrix} g^n_{i,11}(X_i,X_i,s) & g^n_{i,12}(X_i,X_i,s) \\ g^n_{i,21}(0,X_i,s) & g^n_{i,22}(0,X_i,s) \end{pmatrix} \begin{pmatrix} q_i(0,s) \\ y_i(X_i,s) \end{pmatrix} \quad i = 1,\cdots,5. \tag{26}$$

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

$$\begin{aligned} &y_1(X_1,s) = y_2(0,s); \quad y_2(0,s) = y_3(0,s); \\ &q_1(X_1,s) = q_2(0,s) + q_3(0,s); \\ &y_3(X_3,s) = y_4(0,s); \quad y_4(0,s) = y_5(0,s); \\ &q_3(X_3,s) = q_4(0,s) + q_5(0,s) \end{aligned} \tag{27}$$

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 (26) and (27). These flow variables are related in a linear model $M(s)Z(s) = 0$ (Equation (13)), with

$$\begin{aligned} Z(s) =&[q_1(0,s), y_1(0,s), q_1(X_1,s), y_1(X_1,s), q_2(0,s), y_2(0,s), \\ &q_2(X_2,s), y_2(X_2,s), q_3(0,s), y_3(0,s), q_3(X_3,s), y_3(X_3,s), \\ &q_4(0,s), y_4(0,s), q_4(X_4,s), y_4(X_4,s), q_5(0,s), y_5(0,s), \\ &q_5(X_5,s), y_5(X_5,s)]^T \end{aligned}$$

Here, $M(s)$ is a 16 by 20 matrix, which encodes the 16 equations comprised of (26) (five channels), and (27) (internal boundary conditions).

Since 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 (26) (27).

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 (14). More specifically,

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

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

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

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

The parameters of the model are listed in Table I. The mean discharge (Q_{0i}) of the channels 1, 2, 4, 5 are 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 the measurement error, the mean discharge at channel 3 is set to be: $Q_{03} = [(Q_{01} - Q_{02}) + (Q_{04} + Q_{05})]/2$.

TABLE I

PARAMETERS FOR THE SACRAMENTO RIVER AND GEORGIANA SLOUGH

Channel	Q_{0i}	Y_{Xi}	T_{0i}	S_{bi}	n	X_i
$i=1$	186.73m^{3}s^{-1}	5.61m	115m	-0.00004	0.0323m$^{-1/3}$s	2800m
$i=2$	83.89m^{3}s^{-1}	4.04m	110m	-0.00009	0.0323m$^{-1/3}$s	2000m
$i=3$	113.08m^{3}s^{-1}	7.74m	110m	-0.00004	0.0323m$^{-1/3}$s	1300m
$i=4$	58.07m^{3}s^{-1}	4.02m	56m	-0.00019	0.0323m$^{-1/3}$s	600m
$i=5$	65.24m^{3}s^{-1}	5.27m	89m	-0.00004	0.0323m$^{-1/3}$s	1600m

D. 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 (14), the measurable variables are linked by a static relationship of the following form:

$$P(s)Y(s) = 0 \tag{28}$$

where,

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

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 (22) with the static model (28), reconciliated measurements $\hat{Y}$ can be calculated. Assume that $\hat{Y}$ is in the form:

$$\hat{Y} = \sum_{k=0}^{N} \left[B_k e^{j\omega_k t} + \overline{B_k} e^{-j\omega_k t} \right] \tag{29}$$

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

$$\begin{aligned} B_k = &[b_k^{(1,1,0)}, b_k^{(1,2,0)}, b_k^{(2,1,X_2)}, b_k^{(2,2,X_2)}, \\ &b_k^{(4,1,X_4)}, b_k^{(4,2,X_4)}, b_k^{(5,1,X_5)}, b_k^{(5,2,X_5)}]^T \end{aligned} \tag{30}$$

For specific dominant ω_k, $\quad k = 1, \cdots, N$, the coefficient vector B_k in the equation (29) is calculated by equation (22):

$$B_k = \begin{pmatrix} \mathbf{I}_{8,8} & \mathbf{0}_{8,4} \end{pmatrix} \begin{pmatrix} W^{-1} & P(s)^T \\ P(s) & \mathbf{0}_{4,4} \end{pmatrix}^{-1} \begin{pmatrix} W^{-1} D_k \\ \mathbf{0}_{4,1} \end{pmatrix} \tag{31}$$

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 betewen the reconciliated data and measurements is further evaluated in the Table II. Three primary evaluation measure are analyzed here:

- The maximum value is the maximum difference between the reconciliated and measured data at the same time points.
- The coefficient of efficiency E [21] is defined as:

$$E = 1 - \left[\frac{\sum_{i=1}^{N} (\hat{u}_i - u_i)^2}{\sum_{i=1}^{N} (u_i - \overline{u_i})^2} \right] \tag{32}$$

  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 $E < 0$, this measurement is not reasonable and must be excluded from the modeling procedures.
- The last statistic evaluation of the analysis is the correlation coefficient (ρ), given by:

$$\rho = \frac{\sum_1^N (u_i - \overline{u_i})(\hat{u}_i - \overline{\hat{u}_i})}{\sqrt{\sum_1^N (u_i - \overline{u_i})^2 \sum_1^N (\hat{u}_i - \overline{\hat{u}_i})^2}} \tag{33}$$

  where, $\overline{u_i}$ and $\overline{\hat{u}_i}$ represent the mean of measured and reconciliated flow for $i = 1$ *to* N measurement times, respectively.

TABLE II

Max-VALUE, ρ-VALUE AND E-VALUE FOR MODEL VALIDATION

Variable	$USGS$ Station	Max-value	E-value	ρ-value
Discharge	SDC	23.6599	0.9930	0.9975
	DLC	28.2284	0.9368	0.9883
	GES	13.0004	0.9968	0.9985
	GSS	18.4125	0.9368	0.8369
Stage	SDC	0.0539	0.9889	0.9947
	DLC	0.1180	0.9504	0.9759
	GES	0.0703	0.9847	0.9935
	GSS	0.0455	0.9938	0.9989

E. 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 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, along with the stage data at three locations, are used to test the model. Following the steps described in the Section II, the flow variables at the boundaries of each branch $Z_{otherBC}$ are calculated with reconciliated boundary conditions. Furthermore,

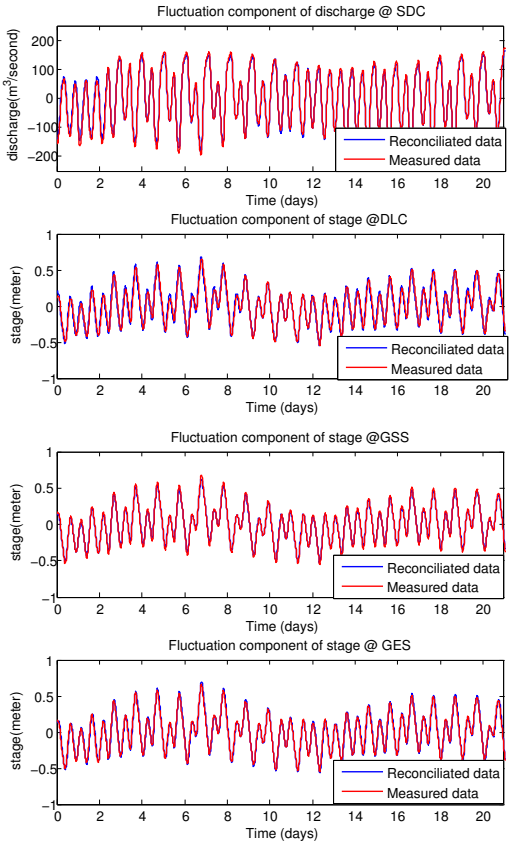

Fig. 5. *Reconciliated boundary condition data v.s. measured data.*

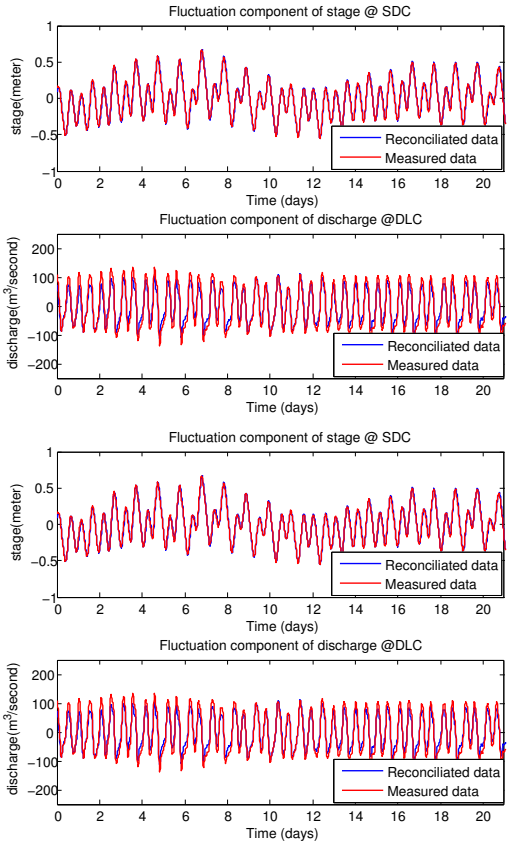

Fig. 6. *Reconciliated boundary condition data v.s. measured data..*

the flow variables along each branch are estimated using the non-uniform transfer matrix. The simulation results are shown in Figure 8.

Model calibration and validation are further evaluated using E-value and ρ-value. $\hat{u}$ here is the modal estimated flow variables.

If a model predicts observed variables perfectly, $E = 1$. If $E < 0$, the model's predictive power is worse than simply taking the average of observed values.

Table III summarizes the values of ρ and E in the validation sets of our channel flow model. Both ρ-values and E-values

TABLE III

ρ-VALUE AND E-VALUE FOR MODEL VALIDATION

Location	A	A	B	C
Variable	discharge	stage	stage	stage
E	0.9775	0.9643	0.9768	0.9612
ρ	0.9895	0.9876	0.9897	0.9875

are close to unity. The results in Table III and Figure 8 indicate that our model accurately simulates the channel flow.

IV. 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

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).*

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 an 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

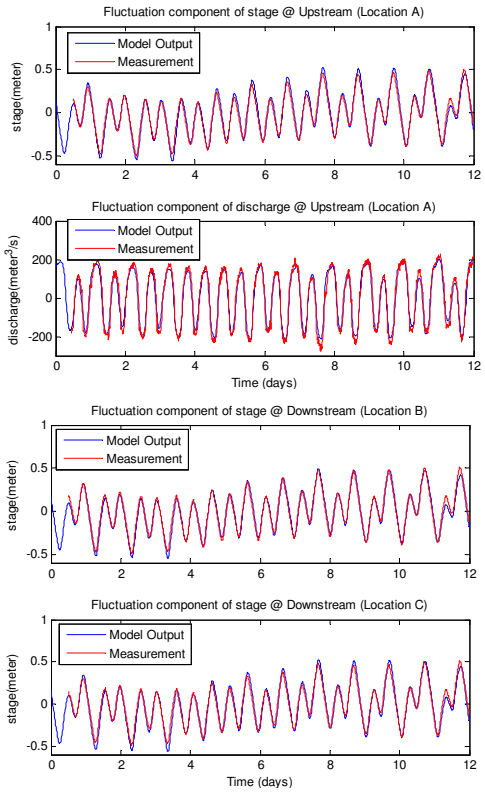

Fig. 8. *Validation of the model output with measurement.*

channel network in the Sacramento-San Joaquin Delta, using four USGS fixed sensors as measurement points. The flow predictions are 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.

Acknowledgements

This article was written when Xavier Litrico was visiting scholar at UC Berkeley. Financial support of Cemagref and France Berkeley Fund are gratefully acknowledged. 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 Professor Mark Stacey. We are grateful to Julie Percelay, Andrew Tinka, Maureen Downing-Kunz and Professor 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*, vol. 42, no. 10, pp. 2481–2586, 1996.

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

[3] N. Bedjaoui, X. Litrico, A. Lourosa, and J. R. Bruno, "Application of data reconciliation on an irrigation canal." in *7th conference on Hydroinformatics*, vol. II, 2006, pp. 1503–1510.

[4] C. M. Crowe, "Data reconciliation–progress and challenge," *Journal of Process Control*, vol. 6, no. 2, pp. 89–98, 1996.

[5] M. Alhaj-Dibo, D. Maquin, and J. Ragot, "Data reconciliation: A robust approach using a contaminated distribution," *Control Engineering Practice*, vol. 16, pp. 159–170, 2008.

[6] 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*, vol. 131, no. 3, pp. 291–297, 2005.

[7] M. J. Bagajewicz and E. Cabrera, "Data reconciliation in gas pipeline systems." *Industrial and Engineering Chemistry Research*, vol. 42, no. 22, pp. 5596–5606, 2003.

[8] J. Ragot and D. Maquin, "Reformulation of data reconciliation problem with unknown-but-bounded errors." *Industrial and Engineering Chemistry Research*, vol. 43, no. 6, pp. 1530–1536, 2004.

[9] C. K. Chui and G. Chen, *Kalman filtering with real time applications.* Berlin: Springer-Verlag, 1998.

[10] R. Corriga, F. F. Patta, S. Sanna, and G. Usai, "A mathematical model for open channel networks." *Applied Mathematical Modelling*, vol. 23, pp. 51–54, 1979.

[11] Y. Ermolin, "Study of open-channel dynamics as controlled process." *Journal of Hydraulic Engineering*, vol. 119, no. 1, pp. 59–71, 1992.

[12] J. Baume and J. Sau, "Study of irrigation cannal dynamics for control purposes." in *International Workshop in Regulation of Irrigation Canals*, Marrakech, Morroco, 1997.

[13] 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*, vol. 125, no. 6, pp. 338–344, 1999.

[14] X. Litrico and V. Fromion, "Simplified modeling of irrigation canals for controller design," *Journal of Irrigation and Drainage Engineering*, vol. 130, no. 5, pp. 373–383, 2004.

[15] V. Chow, *Open-channel Hydraulics*. New York: McGraw-Hill Book Company, 1988.

[16] X. Litrico and V. Fromion, "Frequency modeling of open channel flow," *Journal of Hydraulic Engineering*, vol. 130, no. 8, pp. 806–815, 2004.

[17] ——, "Boundary control of linearized saint-venant equations oscillating modes." *Automatica*, vol. 42, no. 6, pp. 967–972, 2006.

[18] 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*, vol. 12, no. 5, pp. 1029–1036, 1976.

[19] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.

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

[21] J. Nash and J. Sutcliffe, "River flow forecasting through conceptual models part i - a discussion of principles," *Journal of Hydrology*, vol. 10, pp. 282–290, 1970.