

RESEARCH ARTICLE

10.1002/2014WR015270

Variational Lagrangian data assimilation in open channel networks

Qingfang Wu¹, Andrew Tinka², Kevin Weekly², Jonathan Beard³, and Alexandre M. Bayen^{1,2}

¹Department of Civil and Environmental Engineering, University of California, Berkeley, California, USA, ²Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California, USA, ³Department of Mechanical Engineering, University of California, Berkeley, California, USA

Key Points:

- Estimation of flow condition using Lagrangian measurement data
- The data assimilation method applicable to complex hydraulic networks
- The method has been validated with a large-scale field experiment

Correspondence to:

Q. Wu,
qingfangwu@berkeley.edu

Citation:

Wu, Q., A. Tinka, K. Weekly, J. Beard, and A. M. Bayen (2015), Variational Lagrangian data assimilation in open channel networks, *Water Resour. Res.*, 51, 1916–1938, doi:10.1002/2014WR015270.

Received 11 JAN 2014

Accepted 17 JAN 2015

Accepted article online 28 JAN 2015

Published online 3 APR 2015

Abstract This article presents a data assimilation method in a tidal system, where data from both Lagrangian drifters and Eulerian flow sensors were fused to estimate water velocity. The system is modeled by first-order, hyperbolic partial differential equations subject to periodic forcing. The estimation problem can then be formulated as the minimization of the difference between the observed variables and model outputs, and eventually provide the velocity and water stage of the hydrodynamic system. The governing equations are linearized and discretized using an implicit discretization scheme, resulting in linear equality constraints in the optimization program. Thus, the flow estimation can be formed as an optimization problem and efficiently solved. The effectiveness of the proposed method was substantiated by a large-scale field experiment in the Sacramento-San Joaquin River Delta in California. A fleet of 100 sensors developed at the University of California, Berkeley, were deployed in Walnut Grove, CA, to collect a set of Lagrangian data, a time series of positions as the sensors moved through the water. Measurements were also taken from Eulerian sensors in the region, provided by the United States Geological Survey. It is shown that the proposed method can effectively integrate Lagrangian and Eulerian measurement data, resulting in a suited estimation of the flow variables within the hydraulic system.

1. Introduction

The Sacramento-San Joaquin River Delta in California is experiencing a drastic decline in fresh water resources, while the water demand in California continues increasing [Müller-Solger *et al.*, 2002]. Large-scale numerical flow models, such as Delta Simulation Model II (DSM2), sponsored by the California Department of Water Resources, have been used as crucial water resources management tools, providing information about tidal forcing, and salinity transport in the bays and channels of the Delta. A number of factors affect the performance of these state-of-the-art models, such as parameter calibration, mesh generation, and choice of numerical solver. More importantly, the performance of the model largely relies on the determination of open boundary conditions and initial conditions.

Traditionally, the boundary and initial conditions required in the large-scale numerical flow models have been achieved via the Eulerian observations near the boundaries, for example, tidal gauge data, or through satellite data retrieval. Unfortunately, these measurements at large watershed have their intrinsic limitations, specifically small coverage and sparse sampling [Molcard *et al.*, 2006]. Furthermore, installed Eulerian sensors have often experienced many failures, such as broken gauges, sensor drifts, improper use of measuring devices, and other random sources [Albuquerque and Biegler, 1996; Wu *et al.*, 2009a].

In the last decades, techniques using surface and subsurface Lagrangian drifters have significantly advanced, and take measurements while the drifters move along a trajectory. Lagrangian data, in particular collected from surface drifters, provide instant information about the flow, which can be used to describe flow advection and eddy dispersion. Compared to the traditional Eulerian sensors, Lagrangian sensors have several advantages: the devices are inexpensive to build and maintain, and with the advancement in GPS technology, they are capable of providing more and more detail information about the dynamic flows. Because of this merit, Lagrangian data have been widely studied in the last two decades, in particular, in the field of atmospheric and oceanic sciences [Miller *et al.*, 1994; Fisher and Lary, 1995; Honnorat *et al.*, 2009]. The effectiveness of the measurement has been verified in many flow studies in numerous meteorological

and oceanic systems [Anthes, 1974; Poulain *et al.*, 1992; Kuznetsov *et al.*, 2003]. In recent years, one particular utilization of drifter measurements is to map currents and their variability in various flow regions, for example, on the North Atlantic [Garraffo *et al.*, 2001; Flatau *et al.*, 2003], tropical Atlantic [Xie and Carton, 2004], and so on.

Lagrangian measurements are especially valuable in channel networks where Eulerian sensors are sparse or not reliable. As an example, in the entire California Delta (approximately 1100 square miles), there are less than 20 fixed USGS stations installed to monitor the flow, whereas the total length of waterways in the area is more than 700 miles. In this case, traditional Eulerian measures usually result in rather coarse flow estimation in the network domain, while Lagrangian data, reflecting local flow information in essence, would help to refine the estimation in any specific channel. Also, as an alternative of current fixed Eulerian stations in channel networks, a drifter fleet, capable of carrying multiple physical, chemical, or environmental sensors as needed, can be promptly deployed in any specific flow domain of interest, or the areas where some unexpected event happens.

It is noted, however, that raw Lagrangian flow data are rather difficult to use and interpret, due to the fact that Lagrangian motion is often affected by local flow perturbations, which are caused by various physical processes, such as turbulence, surface wind, vertical mixing, etc. [Cowen and Monismith, 1997]. For example, one of the main obstacles of assimilating Lagrangian data into a hydrodynamic model in channel networks is the well-known tidal trapping phenomenon [Fischer *et al.*, 1979]. The trapping mechanism makes water elevation and velocity out of phase, inducing flow dispersion and eddy diffusion at the junctions of channels. The drifter trajectory at these junctions, due to the turbulent mixing processes, usually displays a stochastic “spaghetti-like” shape, which is indicative of slow currents. However, these small-scale features are not fully addressed by the numerical models, and furthermore cause oscillation in the data assimilation process. Lagrangian data acquired from the sensors therefore need to be preprocessed for “smoother” trajectories to exclude any small-scale perturbations, or measurement error due to hardware malfunctioning [Honnorat *et al.*, 2009]. Another difficulty in processing Lagrangian data are that most numerical models for geophysical systems are solved either on a fixed grid in space or as spectral models, which do not relate to the Lagrangian observations directly in terms of the model variables [Salman *et al.*, 2006].

Lagrangian flow data collected in the field are usually further integrated into underlying *shallow water equations* (SWEs) model in one or two dimensions. The technique, namely data assimilation, is an approach to obtain the “best” state estimation of a hydraulic system, given measurements, and a specific model. Here “best” usually refers to the minimization of an error norm. The estimation results are considered optimal in a sense that they minimize differences in the collected observations within certain spatial and temporal scales, and given dynamical or statistical relationships defined by the model [Gunson and Malanotte-Rizzoli, 1996; Nodet, 2006; Tinka *et al.*, 2009]. Through the data assimilation technique, the observational data will compensate for poorly specified model parameters and variables, such as boundary conditions, initial conditions, and other physical processes not incorporated into the model.

Data assimilation can be approached in many different ways. Most of the existing data assimilation methods are categorized into two groups: sequential assimilation methods [Ide *et al.*, 1997] and variational assimilation methods [Navon, 2009, 1986]. Sequential methods, including ensemble Kalman filtering [Salman *et al.*, 2006, 2008; Fan *et al.*, 2004; Tossavainen *et al.*, 2008; Evensen and Van Leeuwen, 2000], particle filtering [Van Leeuwen, 2009; Budhiraja *et al.*, 2007; Van Leeuwen and Melanie, 2013], optimal statistical interpolation [Molcard *et al.*, 2003], Newtonian relaxation (sometimes called nudging) method [Paniconi *et al.*, 2003], and so on, usually involve a series of state analysis and updates, where the observational data are incorporated into the state one step at a time. Consequently, the implementation of these methods requires significant computing resources. Variational assimilation methods, on the other hand, could reduce the overall computational resources, since only one single optimization computation is performed using all the observational data. In classical variational data assimilation, the adjoint method is used to efficiently compute the gradient of the criterion [Kamachi and O'Brien, 1995; Doucet *et al.*, 2001].

In this article, we develop a *quadratic programming* (QP)-based variational method to assimilate Lagrangian flow data acquired in channel networks, without consuming much of computational resources. The method poses the estimation of the flow state in a channel network as an optimization problem, by minimizing a quadratic cost function—the norm of the difference between the drifter observations and the model

velocity predictions—and expressing the constraints in terms of linearized equalities and inequalities. The problem can then be efficiently solved using any fast and robust algorithms available for these kind of posted problems.

Furthermore, we use a one-dimensional *linearized Saint-Venant model* to represent the flow state, not only because it is easy to implement and enables efficient computation, but also because, in many practical cases, if the boundary and initial conditions are properly quantified, the one-dimensional estimation results are adequate for decision making and water supply management to retrieve critical flow characteristics in the domain of interest.

To assess the performance of the proposed QP method, we investigated a distributed network of channels, subject to quasiperiodic tidal forcing, in the Sacramento-San Joaquin River Delta. A field operational test was carried out with a fleet of 100 surface drifters, deployed within approximately 0.55 km² of the river network. During the approximately 4 h experiment, 325,000 GPS readings were taken from the surface drifters and collected, in real time, onto a central server. It is the first experiment of this kind, conducted at such scale, in which high-density Lagrangian data have been collected in a river environment. We demonstrate that the proposed QP approach can successfully handle this drifter data to obtain accurate estimations, and, as a result, the channel network system is adequately simulated using one-dimensional linearized Saint-Venant model. The results also imply that the flow estimations in a channel network can be promptly obtained in a robust and accurate manner, despite of the uncertainties or inaccuracy inherent from model simplifications. Thus, flow estimations could be provided even as fast as real time and give water resource management a useful tool to understand the river system.

The QP-based variational method was first introduced to tidal channel studies to estimate the open boundary conditions using simulated Lagrangian data [Strub *et al.*, 2009], and its applicability to a one-dimensional channel network models was initially verified with specific channel data in Wu *et al.* [2009b]. An early version of the QP assimilation method on a two-dimensional linear hydrodynamic model over a short time domain was presented in Tinka *et al.* [2010]. In this article, we extend the work with a much more comprehensive set of experimental data, a realistic one-dimensional shallow water model, and a complete treatment of the experimental method, numerical schemes, and hardware platform.

The theoretical contributions of this article are the formulation of inverse QP schemes developed from linearized one-dimensional shallow water model with linear constraints, and the use of Lagrangian measurements to reconstruct the distributed flow state. The practical contributions lie in the fact that this is the first time the QP method is actually applied to Lagrangian flow data for assimilation in a tide-driven system.

The remainder of the article is organized as follows: section 2 introduces the mathematical flow models in open channels: a linear Saint-Venant model in a single river reach is derived after linearizing and discretizing the governing equations, and a linear channel network model is constructed enforcing the flow compatibility at the junctions of channels. Section 3 formulates the quadratic programming method and specifies the cost function. Section 4 describes our drifter experiment in the Sacramento-San Joaquin River Delta performed on 9 May 2012. Section 5 elaborates the postprocessing of the drifter data, starting from data filtering, and following with data assimilation. Section 6 presents the assimilation results and discusses the effectiveness of the method by correlating the model predictions with flow velocity data collected either at the Eulerian stations or from the Lagrangian drifters. Section 7 summarizes our drifter measurements and data assimilation.

2. Hydrodynamic Model in Tide-Driven Channel Network

2.1. Linearized Saint-Venant Model in Tide-Driven Hydrodynamic System

Linearized one-dimensional Saint-Venant equations (also called shallow water equations in its one-dimensional form) are commonly used in open-channel hydrodynamic systems which describe the dynamics of the shallow water flow [Litrice and Fromion, 2004a, 2006; Sanders, 2001; Sanders and Katopodes, 1999]. They usually consider the first-order flow perturbations around the steady state flow variables, assuming the quantity of the perturbation is much smaller than the steady state flow [Litrice and Fromion, 2009, 2004b; Wu *et al.*, 2007]. This assumption does not always hold in a tide-driven hydrodynamic system, where the perturbed velocity $v(x, t)$ is sometimes comparable to the steady state velocity $V_0(x)$. Here

$(x, t) \in (0, L) \times \mathbb{R}^+$. Water depths in the channel network, on the other hand, are generally large quantities and do not vary remarkably with respect to time. In such case, we introduce the mean state flow variables $V_T(x, t)$ and $Y_0(x)$ to quantify the flow properties.

The mean state velocity $V_T(x, t)$ is a function of both location and time, generally derived from the historical flow data, reflecting the historical tide information; $Y_0(x)$ is the steady state water depth, reflecting the geometry of the hydrodynamic system. For a tide-driven system, the first-order velocity (respectively water depth) perturbations are represented as $v(x, t) = V(x, t) - V_T(x, t)$ (respectively, $y(x, t) = Y(x, t) - Y_0(x)$). $V(x, t)$ is the average velocity (m/s) across cross section $A(x, t) = T(x) \cdot Y(x, t)$, where $Y(x, t)$ is the water depth (m) and $T(x)$ is the free surface width (m) for a rectangular cross section.

The linearized Saint-Venant equations can be expressed as:

$$y_t + Y_0(x)v_x + V_0(x)y_x + \frac{dY_0(x)}{dx}v - \alpha_0 y = 0 \tag{1}$$

$$(V_T + v)_t + V_T(x)v_x + gy_x + \beta_0(x)v - \gamma_0(x)y = 0 \tag{2}$$

with $\alpha_0(x)$, $\beta_0(x)$ and $\gamma_0(x)$ given by

$$\alpha_0 = \frac{V_T(x, t)}{Y_0(x)} \frac{dY_0(x)}{dx} \tag{3}$$

$$\beta_0 = \frac{g}{V_T(x, t)} \left[2S_b - (2 - F_0^2) \frac{dY_0(x)}{dx} \right] \tag{4}$$

$$\gamma_0 = \frac{4T_0g}{3Y_0(x)(T_0 + 2Y_0(x))} \left[S_b - (1 - F_0^2) \frac{dY_0(x)}{dx} \right] \tag{5}$$

S_b is the bed slope m/m, g is the gravitational acceleration (m/s²), and F_0 is the Froude number.

The chosen of $V_T(x, t)$ inherently reflect the historical tidal condition on the experimental date. In practice, we chose the tidal velocity from the previous day or several days before. An iteration over this velocity trajectory would likely result in a more accurate flow estimation; however, a simple and effective approach using the first guess of $V_T(x, t)$ has been proven to work very well.

During the model setup, the water stage, combined with geometry data measured by the Cross Section Development Program (CSDP) [Tom, 1998], was used to determine the total water depth, which was subsequently incorporated in the model. In the 1-D Linearized Shallow Water model, a uniform rectangular channel is assumed to represent the river geometry for certain branch. The bed level slope was identified by linearly interpolating between the bed levels of three major hydrometric stations.

2.2. Numerical Discretization Scheme

Equations (1) and (2) allow both implicit and explicit discretization schemes to be implemented. The implicit scheme is chosen here, as it is not restricted by the Courant–Friedrichs–Lewy (CFL) condition, and the inconvenient small time step can be avoided.

Applying the Preissman implicit finite difference scheme [Chau and Lee, 1991] to equations (1) and (2) leads to:

$$f(j\Delta x, k\Delta t) \approx \frac{\theta}{2}(f_{j+1}^{k+1} + f_j^{k+1}) + \frac{1-\theta}{2}(f_{j+1}^k + f_j^k) \tag{6}$$

$$\frac{\partial f}{\partial x} \approx \theta \frac{f_{j+1}^{k+1} - f_j^{k+1}}{\Delta x} + (1-\theta) \frac{f_{j+1}^k - f_j^k}{\Delta x} \tag{7}$$

$$\frac{\partial f}{\partial t} \approx \frac{f_{j+1}^{k+1} + f_j^{k+1} - f_{j+1}^k - f_j^k}{2\Delta t} \tag{8}$$

where $f(x, y)$ is the flow variables (either v or y in our case), $\theta \in (0, 1)$ is a time weighting coefficient, j denotes the space index and k is the time index. This scheme has the advantage of allowing nonuniform spatial grids and is unconditionally stable as long as $\theta > 0.5$. In practice, θ is usually set to 1 to achieve a

fully implicit scheme. This enables a more flexible schematization of the river, especially in the case of strongly varying cross sections. The time step is a function of the required accuracy only and can be chosen freely.

The linear form model for a single channel l can be represented as:

$$E_{k,l}X_{k+1,l} = A_{k,l}X_{k,l} + B_{k,l}U_{k,l} \tag{9}$$

where $X_{k,l}$ is the state variable for channel l at time t_k :

$$X_{k,l} = (v_{k,1,l}, y_{k,1,l}, \dots, v_{k,l,l}, y_{k,l,l})^T \tag{10}$$

The velocity and stage perturbation at location $j\Delta x$ and time t_k in channel l are denoted as $v_{k,j,l}$ and $y_{k,j,l}$, respectively. The downstream point of channel l is denoted as l_j and the index of the upstream point is 1. The boundary conditions at time t_k are represented by $U_{k,l}$, which includes the velocity at the upstream and water stage at the downstream end:

$$U_{k,l} = (u_{k,1,l}, y_{k,l,l})^T \tag{11}$$

The matrices $E_{k,l}$, $A_{k,l}$, and $B_{k,l}$ are constructed by assembling equations (1) and (2) discretized using above numerical method equations (6–8).

2.3. Network Model

The one-dimensional channel network model is constructed by decomposing the channel network into individual channel reaches, and applying the linear model equation (9) to each branch. Internal boundary conditions are imposed at every junction to ensure flow mass and energy conservation. For a river junction illustrated in Figure 1, the linear equations of hydraulic internal boundary conditions at the junction are assured by mass and energy conservation as follows.

Assuming no change in storage volume within the junction, the continuity equation can be expressed as:

$$v_{k,l_1,1} \cdot T_1 = v_{k,l_2,2} \cdot T_2 + v_{k,l_3,3} \cdot T_3 \tag{12}$$

where T_l is the average free surface width for channel l , $l = 1, 2, 3$.

Assuming the flow in all the branches meeting at the junction is subcritical, the equation for energy conservation can be approximated by a kinematic compatibility condition as follows:

$$y_{k,l_1,1} = y_{k,l_2,2} = y_{k,l_3,3} \tag{13}$$

To model the entire network, the conservation equations for each individual channel and interior junctions are assembled together. The flow variables inside the domain are therefore related by a linear equation:

$$E_k X_{k+1} = A_k X_k + B_k U_k \tag{14}$$

where X_k is the concatenated vector of $X_{k,l}$ and U_k denotes the external boundary conditions of the channel network system.

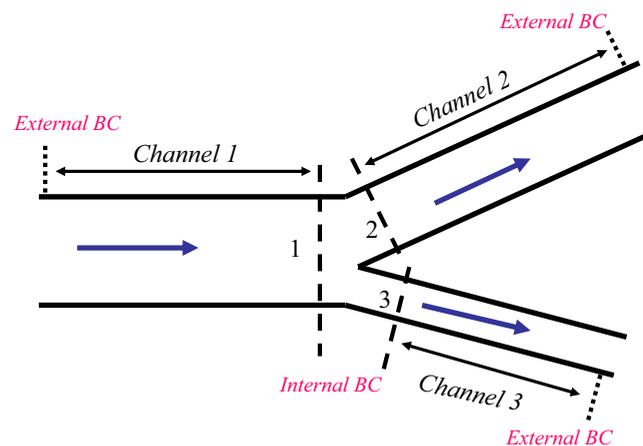


Figure 1. Flow compatibility of channel junctions: (1) no change in storage volume within the junction, equation (12); (2) kinematic compatibility condition, equation (13).

3. Formulation of the Data Assimilation Method

3.1. “Pseudo-Lagrangian” Data Assimilation

One major issue to assimilate Lagrangian data into hydrodynamic models is the quantification of the connection between Lagrangian measurements and Eulerian velocity. A simple and intuitive solution to this challenge is to approximate the Eulerian field by

dividing the position displacement δr with the sampling period δt , and the Eulerian velocity is thus estimated as $\delta r / \delta t$. Such a method is usually called “Pseudo-Lagrangian” Data Assimilation [Molcard *et al.*, 2005]. The method works well when the sampling period δt is much smaller than the Lagrangian correlation time scale T_l [Griffa, 1996; Veneziani *et al.*, 2004].

3.2. Cost Function

In an ideal case in which the flow measurement devices are active and the measurement data are reliable, we can construct a cost function to minimize the difference between the most valid observation data and model state variables. The standardized framework was specified in Ide *et al.* [1997] and notations are defined as follows:

1. X_k : Concatenated vector of state variables (u, h) for all mesh points at time t_k .
2. X_B : Background term vector to improve well-posedness and convergence of the problem. It is a “first guess” of the state of the system, usually derived from the historical data, and can be further refined in the data assimilation process.
3. U_k : Vector of boundary conditions at time t_k .
4. Y_k : Vector of observed variables at time t_k , namely the velocity components u and the water height h at some mesh points at a time instant t_k .
5. B : Covariance matrix of the background error.
6. R_k : Covariance matrix of the observation error at time t_k .
7. H_k : Observation operator, which projects the state vector X_k into the observation subspace containing Y_k .

Our data assimilation strategy is to search for the initial state X_0 and boundary conditions U_k that minimize the ℓ^2 -norm of the difference between the state and observation variables and the difference between the initial state X_0 and the background term X_B :

$$\mathcal{J}(X_0, U_k) = (X_0 - X_B)^T B^{-1} (X_0 - X_B) + \sum (Y_k - H_k[X_k])^T R_k^{-1} (Y_k - H_k[X_k]) \quad (15)$$

The objective function expressed in equation (15) is a function of the initial state and boundary condition of the system.

The observation operator, H_k , is nonlinear in general variational data assimilation schemes. However, in a “Pseudo-Lagrangian” data assimilation process, in which the observations and state variables both represent velocity, H_k will be a time-varying observation matrix. For simplicity, if we take the assimilation time step the same as the observation sampling time, H_k would be a $(0, 1)$ matrix, with element $i, j=1$ if the drifter associated with measurement i was in the cell associated with the state variable j at time k .

The covariance matrices B and R_k essentially specify the weights given to the error terms in equation (15). A reasonable choice of B and R_k is $B=b1$ and $R_k=r1$, where b is determined by the quality of the background term X_B , and r should reflect the accuracy of the observations.

3.3. Quadratic Program Formulation of Data Assimilation

The linearization of network constraints makes it possible to pose the data assimilation problem as an optimization problem, with the positive semi-definite quadratic cost function expressed in equation (15):

$$\begin{aligned} \text{minimize } \mathcal{J}(X_0) &= \frac{1}{2} X^T P X + q^T X \\ \text{subject to } G X &\leq h \\ AX &= b \end{aligned} \quad (16)$$

In the previous formulation, X is the vertical concatenation of all state vectors from time t_0 to the end of the data assimilation period t_{\max} , and P and q are formed by expanding all the terms in equation (15). Equation $AX = b$ represents the flow dynamics constrained by equations (1–13). In order to reduce the search time, G and h are used to keep the search in a realistic set.

It should be noted that the computational cost of solving the quadratic program above is very low, as an implicit discretization algorithm is adopted, allowing for a large time step discretion. Furthermore, the

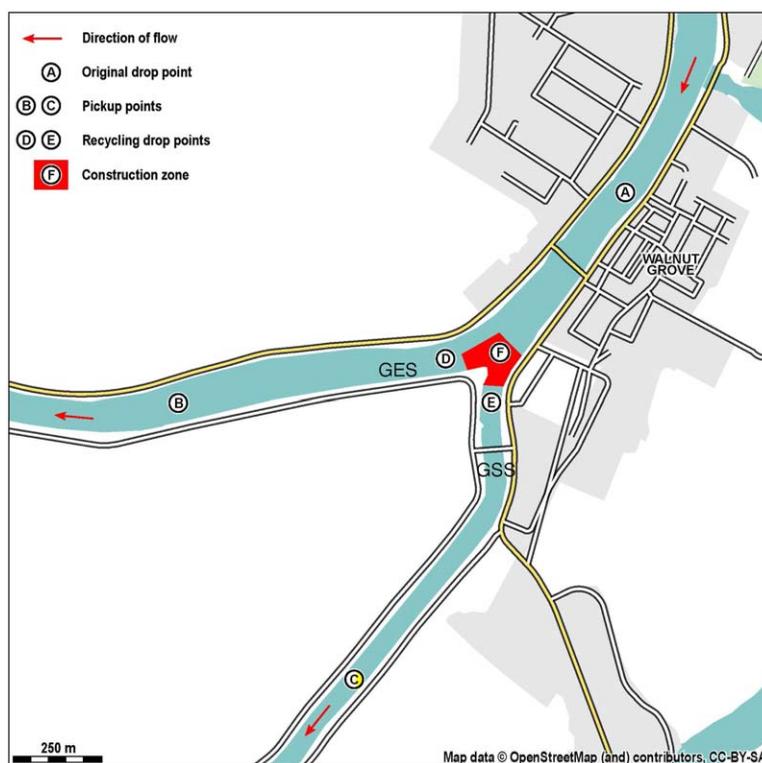


Figure 2. Annotated map of the Walnut Grove experimental area for the 9 May field operational test.

uniqueness of an optimum X is not guaranteed, due to the rank-deficiency of the observation matrix H_{ki} ; however, we can still exclude all the linear subspaces from the set of optimal solutions by employing the background term X_B [Tinka et al., 2010].

4. A Field Operational Test in the Sacramento River and Georgiana Slough

4.1. General Description of the Field Operational Test

On 9 May 2012, an experimental deployment of 70 drifters was conducted at the junction of the Sacramento River and Georgiana Slough near Walnut Grove, California. Figure 2 shows the spatial deployment area.

The Sacramento River is the larger channel including labels A and B, while the Georgiana Slough splits off from the Sacramento River at F and continues south toward label C.

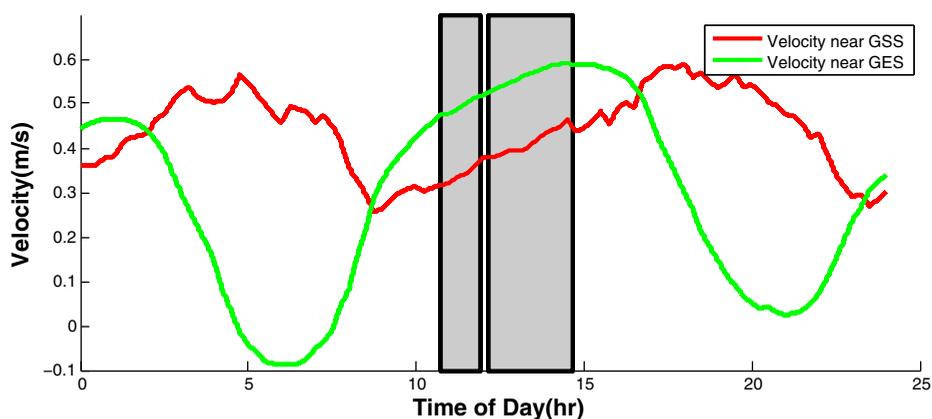


Figure 3. Overview of the experimental timeframe. The two gray boxes denote the time domain during which the drifter data were collected: deployment, sensing, retrieval, and redeployment.



Figure 4. Android drifter usage during the 9 May experiment. Photo credit: Berkeley Lab-Roy Kaltschmidt.

The flow velocity at surface along the Sacramento River was approximately 0.46 m/s (1.5 ft/s) in the outgoing (from northeast to southwest) direction. This is the noninverted tidal condition. Figure 3 shows the water velocity at two U.S. Geological Survey (USGS): GES, GSS over time during the experiment.

The original plan was to deploy all of the drifters from the Walnut Grove Public Dock (label A in Figure 2), allow them to propagate through the junction, retrieve them at downstream points B and C, then recycle them at point D and E for the rest of the experimental run.

Unfortunately, on 9 May there was a significant underwater construction operation happening at the junction (box F, in Figure 2), requiring a midexperiment change of plans: drifters were initially released from A



Figure 5. Android drifter set for the study with complete parts list.

Trajectory of Drifter 37 on 05/09/2012

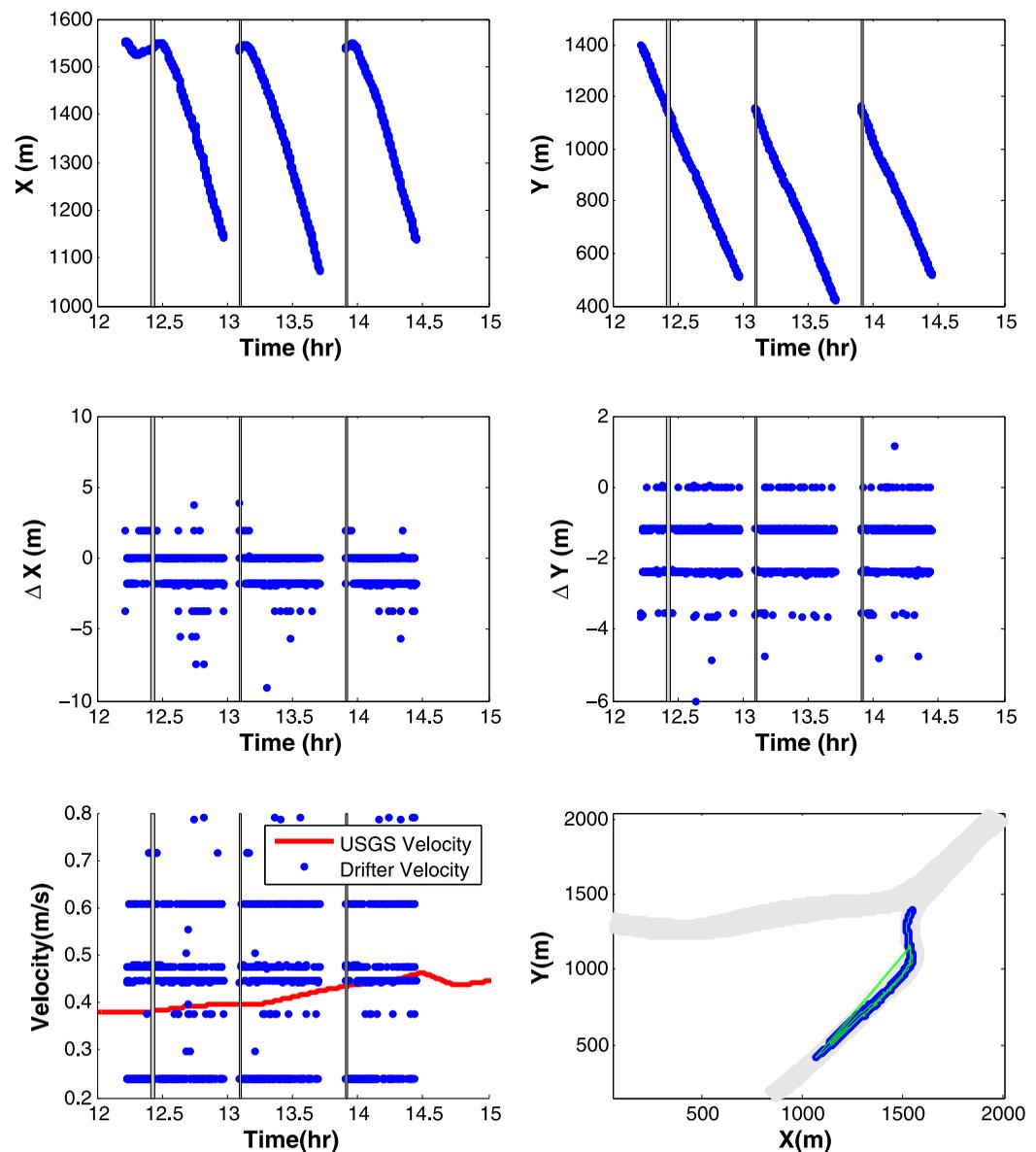


Figure 6. (top) x and y coordinates of the GPS measurements as a function of time, of Drifter 37 during the field operational test. (middle) Corresponding deviations of the x and y coordinates between successive measurements. These exhibit quantization due to the GPS-capturing system of the Android phones used. (bottom left) Corresponding velocities obtained by finite differences of successive measurements, compared against USGS measurements. (bottom right) GPS measurements captured during the experiment, plotted on map of the domain, with USGS sensor station marked. The gray boxes indicate the time when the drifters were passing by USGS stations.

and picked up around F, then redeployed at D and E, then cycled from B–D and C–E. Drifter retrievals and redeployments were performed by two boat teams of three workers (one pilot and two drifter retrievers), as seen in Figure 4.

4.2. Android Drifters

The fleet of drifters used in the 9 May experiment consisted of UC Berkeley’s recently developed “Android drifters” [Beard et al., 2012]. Each Android drifter is an inexpensive assembly consisting of a mobile phone running the Android operating system, a lithium-ion battery to extend the phone life, a waterproof enclosure based around a water filter canister, and supporting mechanical parts (Figure 5). Drifter sensor-gathering operations are executed by a custom application written for the Android operating system environment.

Drifter 37: Passing by GSS 3 times on 05/09/2012

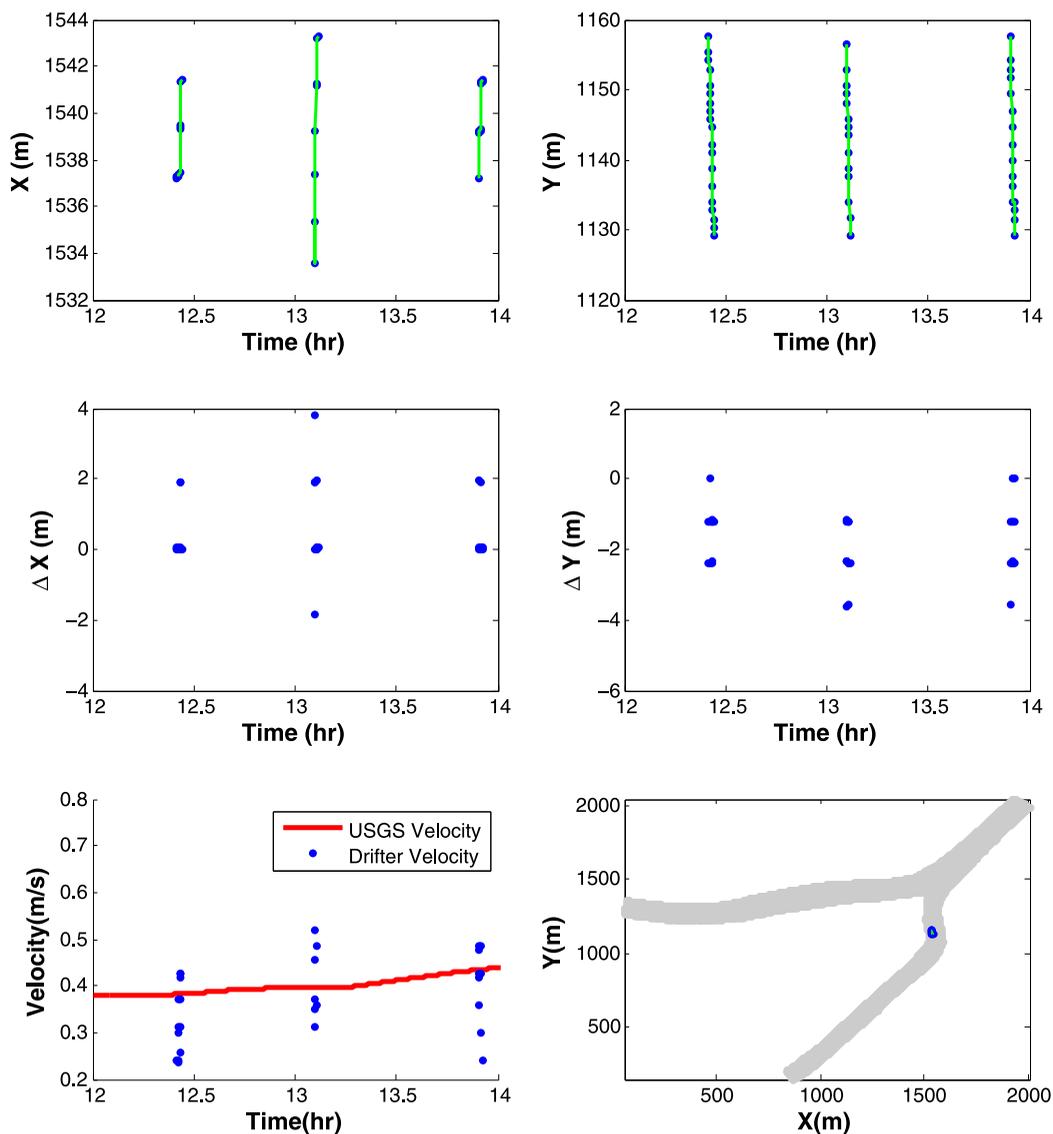


Figure 7. The data of Figure 6, filtered to include only points in which Drifter 37 was near the USGS station. (top) x and y tracks of the GPS measurements as a function of time. (middle) Corresponding increments in x and y between the successive GPS tracks. (bottom left) Corresponding velocities obtained by finite differencing the position (compared to USGS measurements). (bottom right) Measurement locations of the above data.

Previous generations of drifters developed by the Floating Sensor Network team included custom developed electronics and additional features, notably autonomous propulsion [Oroza et al., 2013]. The design priorities for this generation of drifter were to increase the amount of collected data through higher fleet numbers (requiring lower production cost per unit), greater reliability, and improved manufacturability. Avoiding custom electronics and taking advantage of the positioning and communication features of modern mobile phones allowed us to meet these design requirements. Operating passive drifters in river environments, however, leads to greater personnel involvement for deployment, retrieval, and protection.

The core functionality of the Android app is to transmit time-stamped GPS positions to a remote server over the cellular network. The app user interface features the ability to start and stop the service of transmitting and logging, a display of current GPS data and orientation, and a menu to set the drifter ID, measurement frequency and server address. During an experiment, drifters are subject to numerous

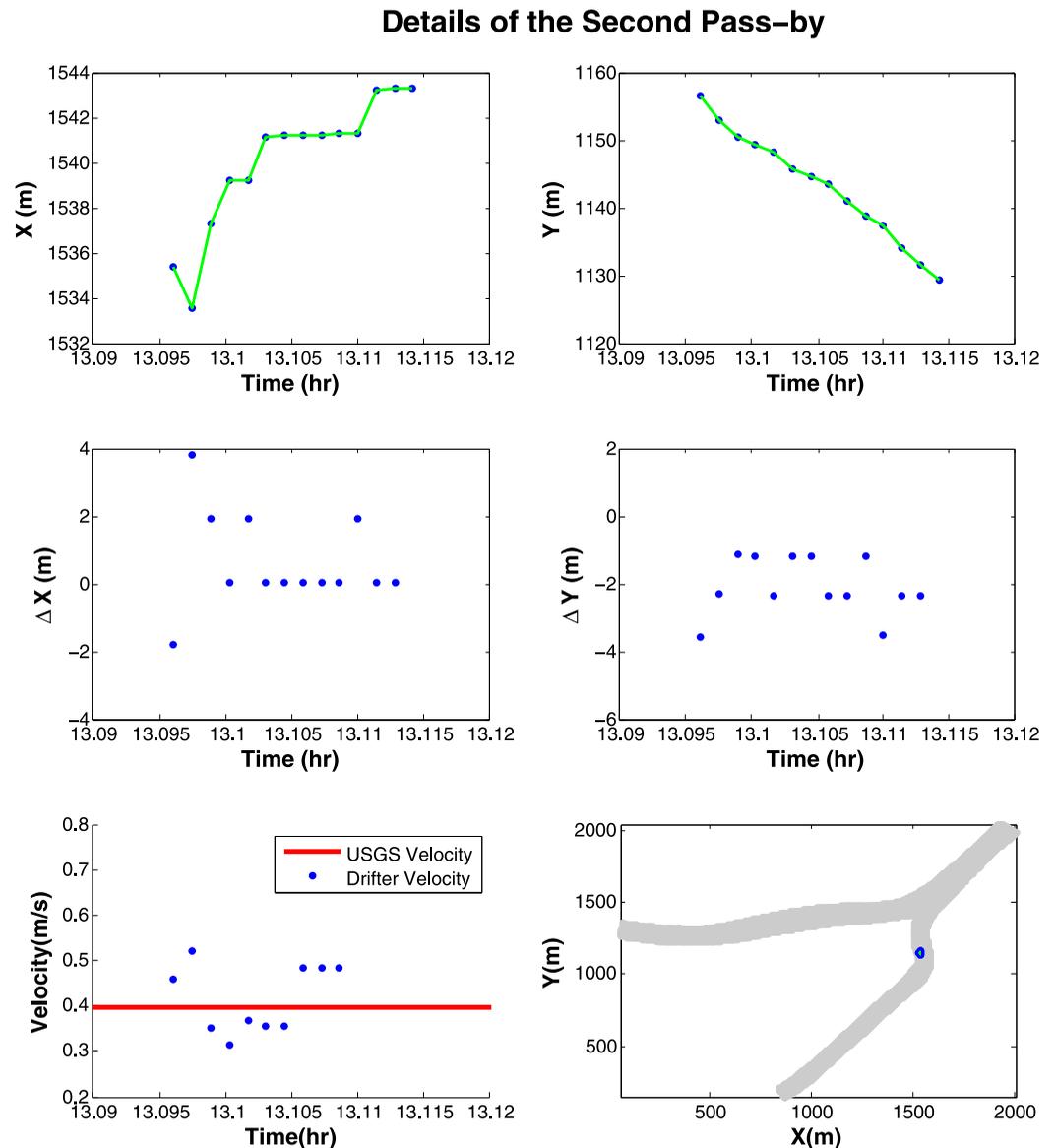


Figure 8. The data of Figure 7, filtered to include only the second pass-by, around time 13:00, of the experiment. (top) x and y tracks of the GPS measurements as a function of time. (middle) Corresponding increments in x and y between the successive GPS tracks. (bottom left) Corresponding velocities obtained by finite differencing the position (compared to USGS measurements). (bottom right) Measurement locations of the above data.

environments, most notably: storage in containers, floating in water, and transportation in boats between deployments.

5. Implementation of Data Assimilation

5.1. Overview of the Drifter Data From 9 May Experiment

The GPS drifter data were collected via the *LocationManager* service of the Android operating system [Inc., 2013]. Under normal conditions (i.e., clear sky and the GPS is locked onto more than four satellites), the operating system will provide the user application with a new position estimate every 1 s. The GPS coordinates are provided by the operating system as a pair of double-digit (64-bit floating point) values representing global latitude and longitude in decimal degrees. Immediately upon receiving these coordinates, the location is transformed into UTM coordinates. Our customized software transmits the latest GPS coordinates every 5 s, rounds them to the nearest centimeter, and stores them on an onboard memory.

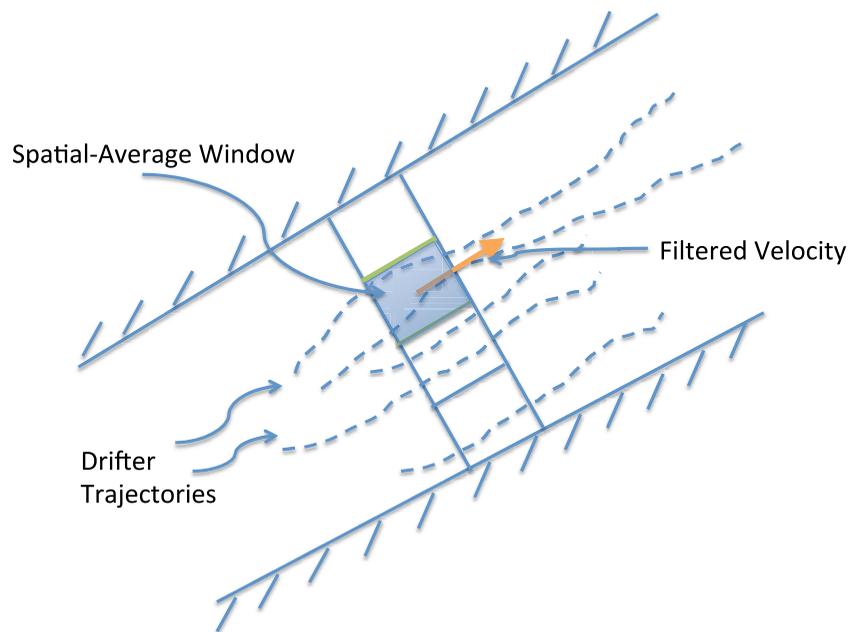


Figure 9. Rectangle spatial-time window to remove perturbations from Lagrangian drifters. The shadowed box is the space-time window W . Dash lines denote the drifter trajectories, and orange arrow denotes the filtered velocity.

Figure 6 shows the GPS measurements collected by Drifter 37 during the field operational test. Figure 7 shows a different view of the data where only points that are near the GSS station are represented. Furthermore, Figure 8 plots the points during the second travel that Drifter 37 passes by the GSS station. The drifter data taken close to the GSS station are of interest because it can be directly correlated with the velocity measurements taken at that station. Measurements taken further away could be at points where the local river velocity does not match the velocity of the GSS station.

As shown in the bottom left plots in Figures 7 and 8, the flow velocity obtained by finite differencing the drifter positions are correlated with that estimated by USGS gauging stations. During our 9 May experiment,

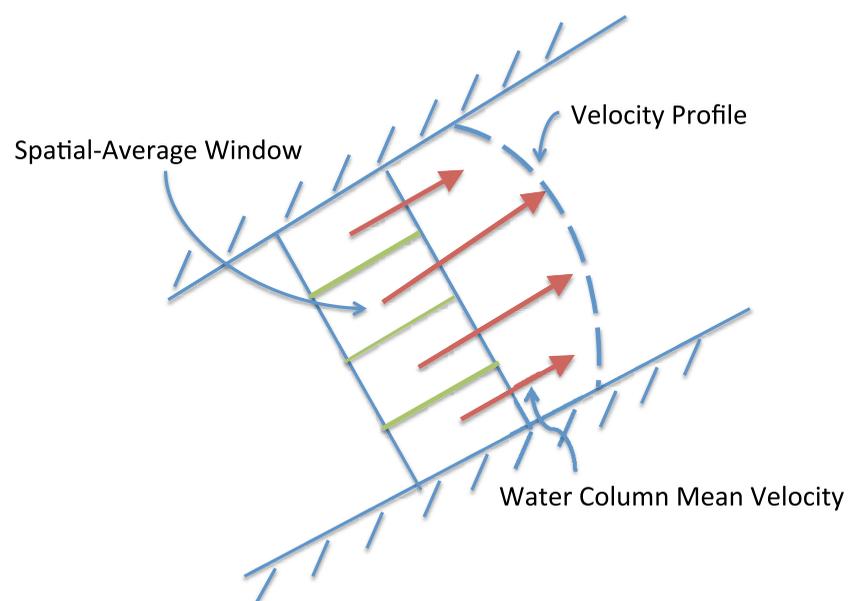


Figure 10. Velocity profile in a cross section. The average velocity across the river cross section $\bar{u}(X, t)$ is calculated with the water column mean velocity $u_{\text{column}}(X, t)$ at each cell equation (19).

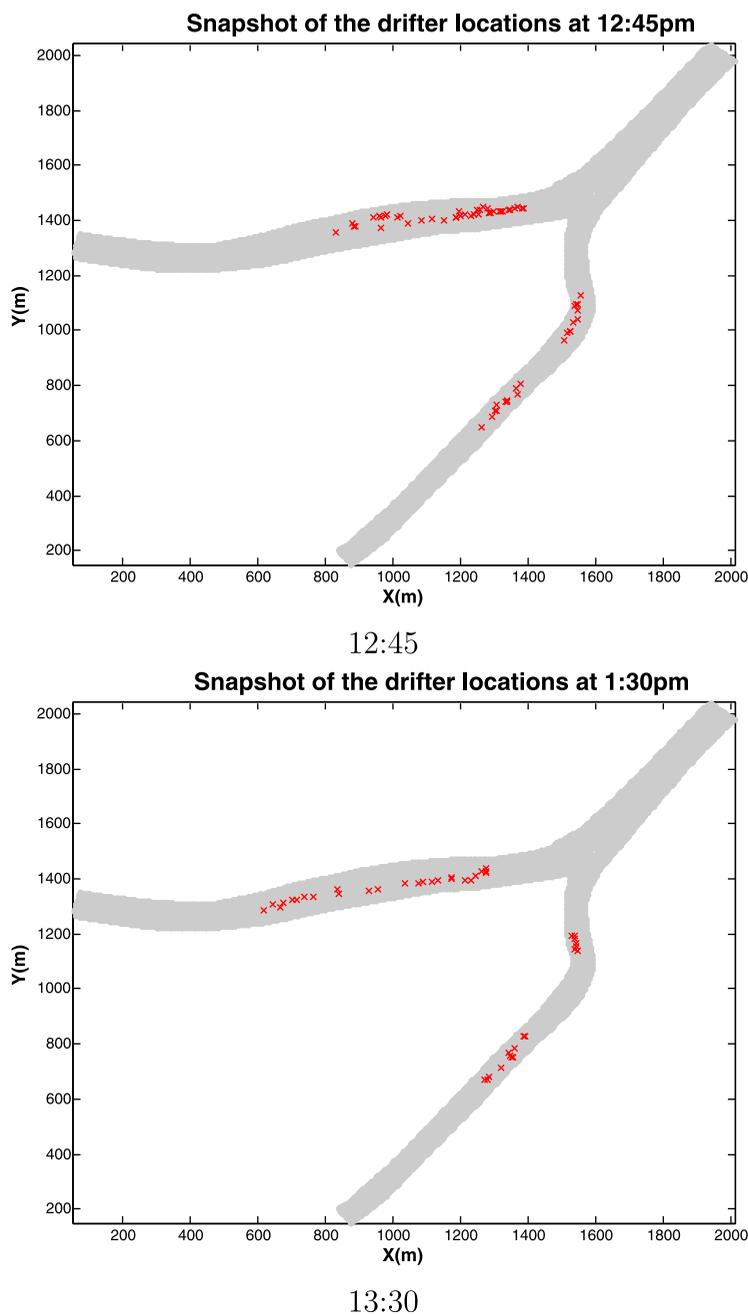


Figure 11. Snapshot of the Android drifter fleet at two given times during the 9 May 2012 field operational test.

one specific drifter (Drifter 37) passed by the gauging stations three times, and velocities calculated from the drifter data were consistent with the gauging station measurements. Using the USGS measurements as a reference, the drifter data fluctuated in a range of 25–30%.

The GPS coordinates provided by the Android operating system appear to be quantized to approximately 1.88 m in the x direction and 1.1 m in the y direction. Evidence of this is seen by the deviations of the x and y coordinates, which seem to correspond to discrete values, i.e., “quantized.” A possible explanation for the y direction is that the fifth decimal place in a decimal representation of latitude corresponds to 1.11 m for this region. Unfortunately, we were unable to find the same relationship for the precision in the x direction (the fifth decimal place for longitude is 0.88 m). This quantization problem is most likely due to the architecture of the GPS-capturing system of the Android phone. The discretization error can be mitigated by a filter

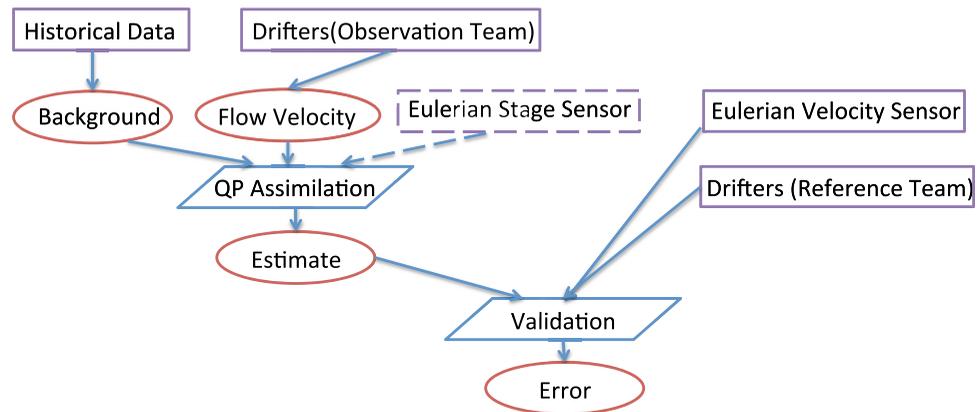


Figure 12. Flow chart for the data assimilation. Drifter data are divided into two parts for observation (used for the data assimilation) and validation purpose.

with low pass characteristics [Stewart and Pfann, 1998]. Therefore, the filtering step discussed in the following section is imperative to smooth out any discretization error.

The drifters transmit their data via a socket interface over the cell phone’s GSM connection to our Internet server. The coordinates are both stored and transmitted as ASCII-encoded decimal numbers, along with the drifter’s identifier, and a “valid” flag indicating whether the drifter is floating or in storage. The Internet server parses and stores the received data in a SQLite database for later retrieval. The server software also forwards the data, via a socket connection, to a computing cluster where data analysis can be performed.

5.2. Trajectory Filtering

In order to properly use the data, we collected from the 9 May experiment, we need to preprocess the “Pseudo-Lagrangian” data as described before. More specifically, when the time interval between successive position measurements is smaller than the Lagrangian integral time scale T_L , the flow velocity u can be approximated as finite difference of successive positions. However, in the case of experimental flows, such as the ones we are dealing with, the estimated velocity cannot be used directly in the data assimilation system, as the velocity of the drifters is perturbed by many physical processes that cannot be precisely simulated in the shallow water models. In this case, one should recognize that individual trajectories are largely unpredictable, and a statistical description is preferable. Ideally, when a large number of trajectory observations are available, a space-time filter should be applied to the original data set to remove any small-scale perturbations and measurement errors.

In this section, we define a local space-time averaging filter, and discuss its related features. The filtered drifter velocity $u_{\text{filtered}}(X, t)$ is defined as the mean velocity observed at time t and location X in a space-time window $\mathcal{W} = \mathcal{W}_t \times \mathcal{W}_s$, where \mathcal{W}_t denotes a temporal window around t and \mathcal{W}_s denotes the spatial neighborhood of $X(t)$.

$$u_{\text{filtered}}(X, t) = \frac{1}{\mathcal{T}} \sum_{i=1}^{N_{\text{obs}}} \int_{\mathcal{W}_t} \frac{d}{dt} X_i^{\text{obs}} \mathbb{1}_{X_i^{\text{obs}} \in \mathcal{W}_s} dt, \tag{17}$$

where

$$\mathcal{T} = \sum_{i=1}^{N_{\text{obs}}} \int_{\mathcal{W}_t} \mathbb{1}_{X_i^{\text{obs}} \in \mathcal{W}_s} dt \tag{18}$$

Table 1. ℓ^2 -Value for Assimilated Flow Velocity and Eulerian Sensors

Location	Sacramento River		Georgiana Slough	
	Full Fleet	Half Fleet	Full Fleet	Half Fleet
Filtering Window				
30 s window (drifter only)	0.06	0.09	0.07	0.08
150 s window (drifter only)	0.05	0.06	0.06	0.07
150 s window (drifter + stage)	0.02	0.03	0.03	0.03

It is important to select a suitable bin size for the filter. Choosing \mathcal{W}_t and \mathcal{W}_s too large yields an overly smooth mean while using too small bins subsumes eddy-like features. To determine a suitable window size, two factors are compromised against each other. On the one hand, the window

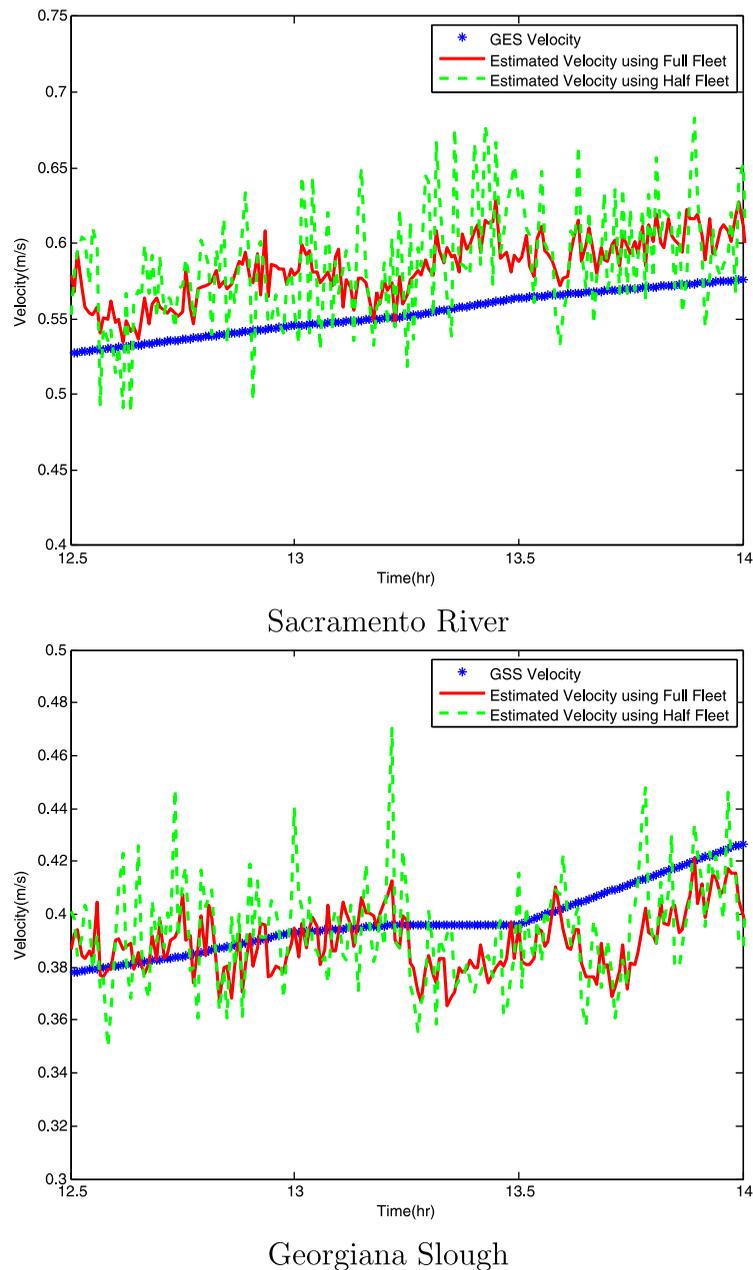


Figure 13. Assimilated flow velocity near USGS stations: (a) GES and (b) GSS using filtered drifter data with 30 s space-time window.

should be sufficiently large to ensure any small-scale perturbations being removed. On the other hand, however, it should not be too big to have the most important characteristics of the flow taken out. These space-time windows may be overlapped along the flow, resulting in a smooth flow representation.

In Figure 9, a rectangular spatial window is drawn showing Lagrangian drifters located within it from time t^0 to time t^f . A number of drifters travel through this spatial window during the time window \mathcal{W}_t . The observed drifter trajectories are represented with dotted lines. Essentially we “average out” the velocities of all the drifters passing through the specific spatial window during a given time interval, and denote the filtered velocity of these drifters with a solid arrow. The filtered velocity is thus considered to represent the local flow velocity at the specific time, and utilized in the subsequent flow computations.

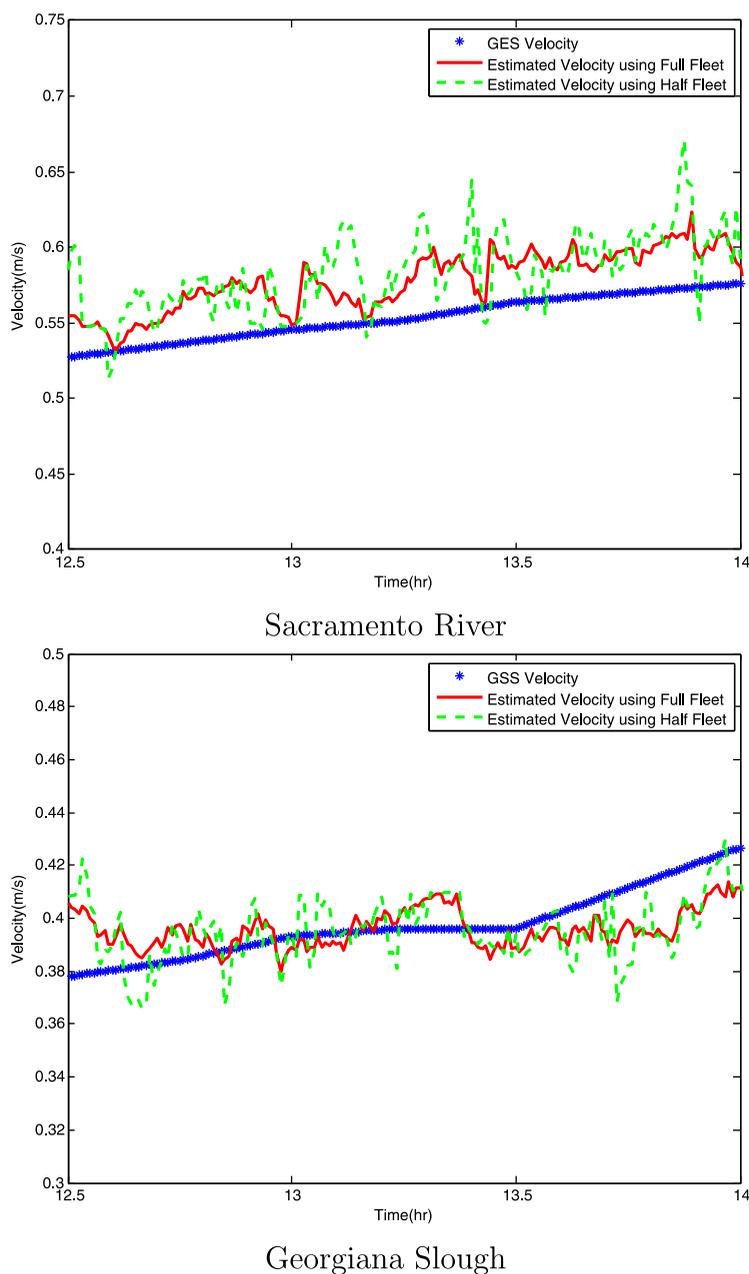


Figure 14. Assimilated flow velocity near USGS stations: (a) GES and (b) GSS using filtered drifter data with 150 s space-time window.

5.3. Average Velocity Across the Channel

Since our model relies on one-dimensional shallow water equations, the velocity in the model system is defined as the average flow velocity across the channel, and thus the data we obtained from the previous section $u_{\text{filtered}}(X, t)$ need to be further refined.

Generally, for flows in open channels and natural rivers, actual velocities in a cross section varies from the highest value near the channel center to the lowest value near overbanks or river bottom. In a river discharge measurement protocol recommended by the USGS, the mean water column velocity $u_{\text{column}}(X, t)$ in a shallow water system is determined by the average of velocities measured at the vertical locations 60% of the water depth. If the velocity profile follows the log-law-of-the-wall, the theoretical ratio of water column mean velocity over surface velocity would be 0.85. In practice, this ratio can be verified by an acoustic instrument, usually placed at fixed locations to measure the velocity profiles across the river. In the previous experiments, the velocity profiles were measured with an *Acoustic*

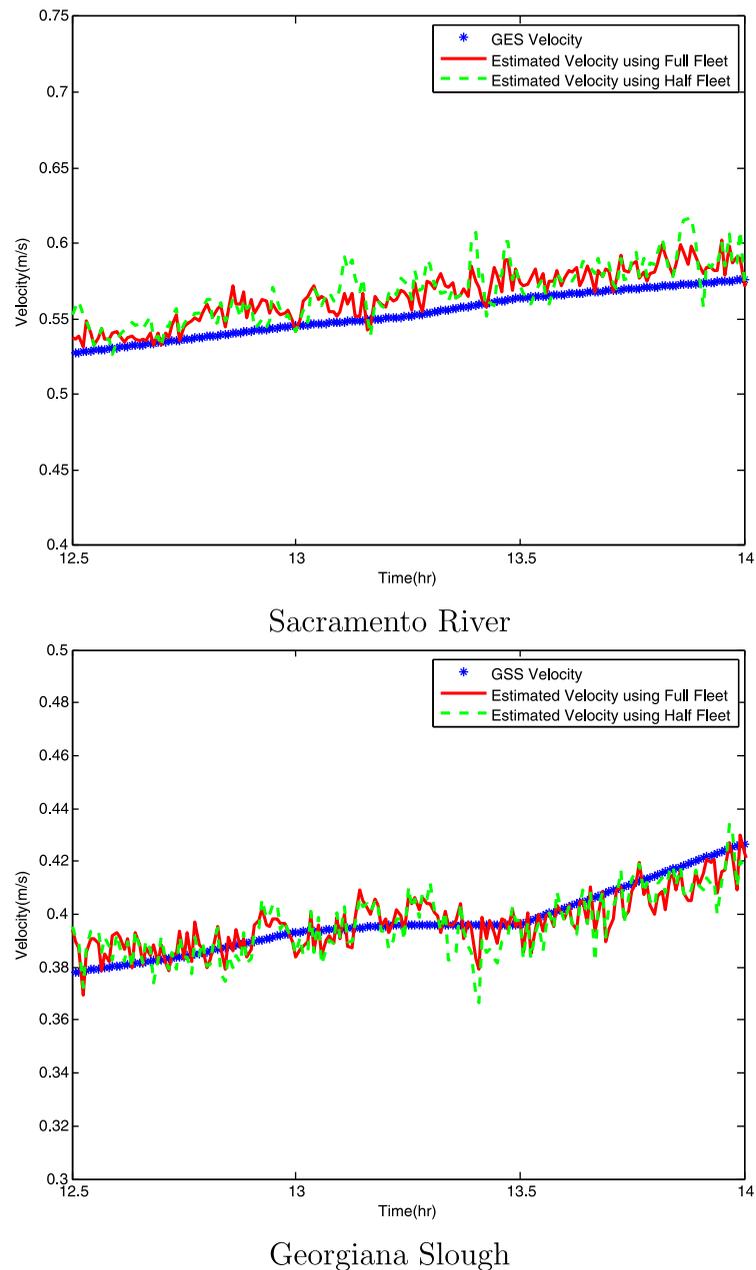


Figure 15. Assimilated flow velocity near around USGS stations: (a) GES and (b) GSS using both filtered drifter data with 150 s space-time window and Eulerian depth data.

Doppler Current Profiler (ADCP), and the computed mean velocity to surface velocity ratio is in the range of 0.80–0.93, with a mean value of 0.86 [De Serres et al., 1999].

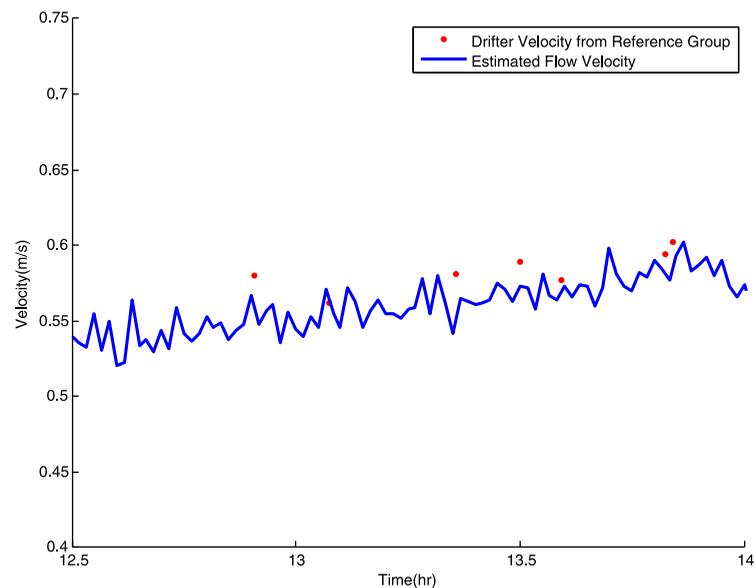
Once the water column mean velocity $u_{\text{column}}(X, t)$ is determined, the average velocity across the river cross section $\bar{u}(X, t)$ is readily evaluated (Figure 10):

$$\bar{u}(X, t) = \frac{1}{\mathcal{A}} \int u_{\text{column}}(x, t) h(x, t) dx \quad (19)$$

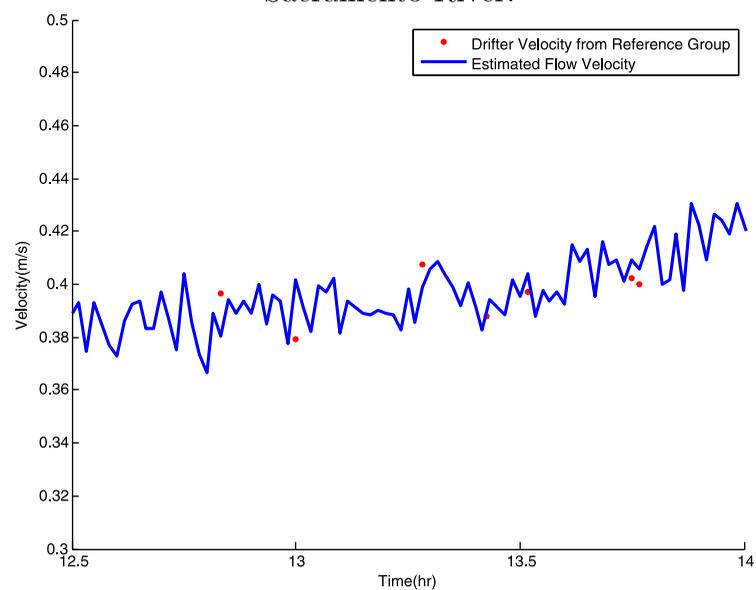
where

$$\mathcal{A} = \int h(x, t) dx \quad (20)$$

where $h(x, t)$ is the total water depth at the specific location x at time t .



Sacramento River.



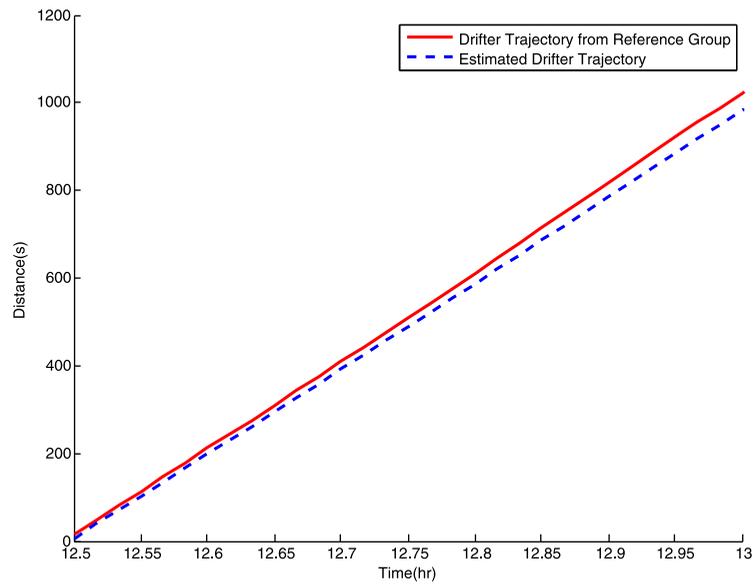
Georgiana Slough.

Figure 16. Comparison of assimilated flow velocity at certain location with drifter data from the Reference Group.

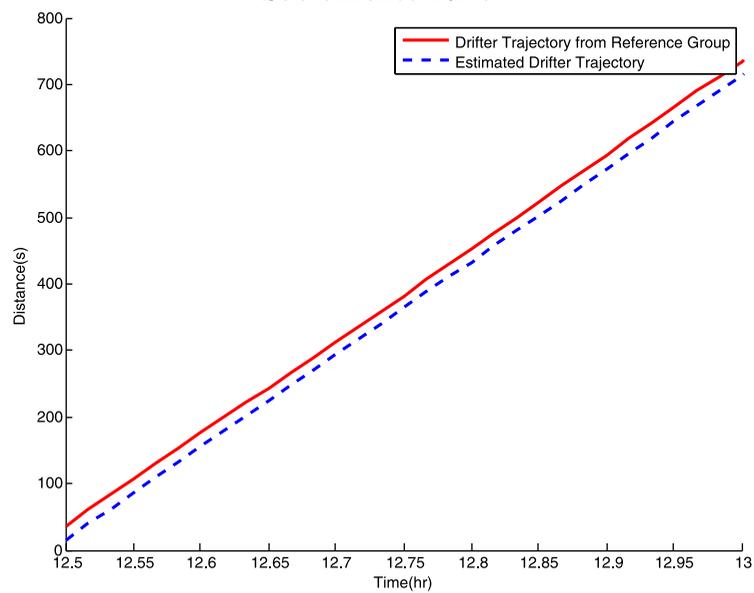
5.4. Assimilation Technique

Many data assimilation techniques, for example, adjoint and ensemble methods, are capable of processing Lagrangian data. The adjoint method has been successfully applied to many types of data assimilation problems, including weather applications, ground water flow studies, oceanography applications, and in shallow water flow estimations. However, the integration of the adjoint of a large-scale model backward in time is equivalent to several forward model simulations and thus remains computationally expensive. Several other Bayesian methods, e.g., ensemble Kalman filtering and particle filtering, also demand enormous computational resources.

The QP-based data assimilation in this article is implemented with the optimization modeling language AMPL and solved with IBM ILOG CPLEX [CPLEX, ILOG, 2007]. We chose CPLEX as the optimization solver, not only because it is very efficient and robust in large-scale optimization, but also because our cost function



Sacramento River.



Georgiana Slough

Figure 17. Comparison of assimilated drifter trajectory with the drifter trajectory from the Reference Group.

happens to be in a quadratic form, which is the designated utilization of the software. Moreover, this package also includes a distributed parallel algorithm for mixed integer programming to leverage multiple computers to solve difficult problems with millions of constraints and variables.

6. Data Assimilation Results

6.1. Overview

In this section, we present the data assimilation results of our 9 May 2012 field operational test. Two sets of data are tracked in our test: The first one is the drifter trajectories provided by the Lagrangian drifters. The second set of data consists of water depth and velocity at certain locations in the deployment domain, acquired with classical USGS Eulerian sensors.

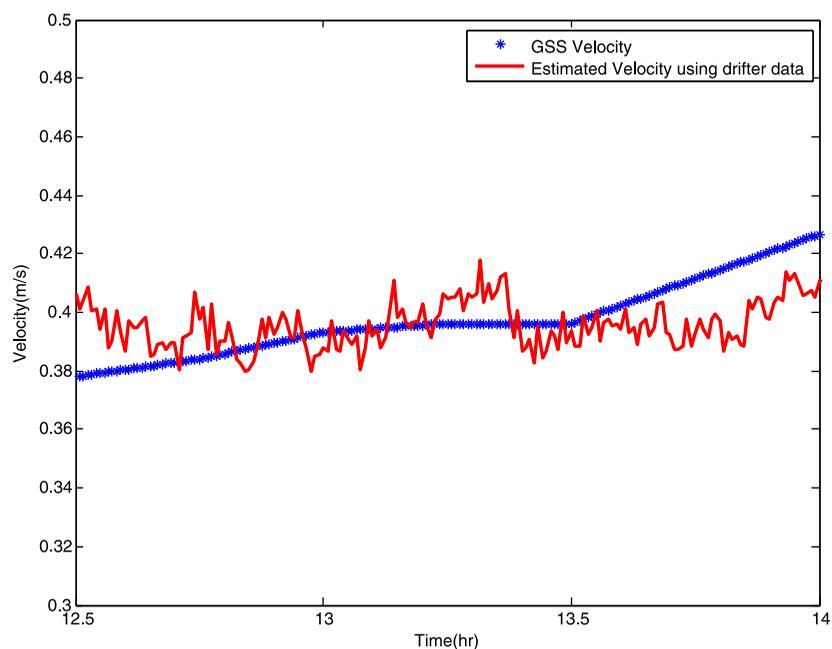


Figure 18. Comparison of the velocity data from the gauging station at GSS with estimated velocity using drifter data. This set of drifter data contains 5% random Gaussian noise.

Figure 11 gives two snapshots of the Android drifters during the field operational test to illustrate the coverage of the experiment.

Figure 12 shows the flow chart of the data assimilation. Historical data from the USGS are used to generate the background term for the QP process. The estimate of the state of the system is obtained by assimilating the drifter data from the observational team, with or without the Eulerian stage information. Either the Eulerian velocity data provided by USGS stations or the drifter data from a Reference Group (as defined later) are used for validation.

We have conducted 12 numerical experiments with different settings as summarized in Table 1 to validate the proposed Lagrangian data assimilation method. We have also investigated how the number of drifters and the filtering window could affect the results. In addition, water depth measurement from the Eulerian stations is specified to justify these results.

6.2. Assimilation With Lagrangian Data Only

Figures 13 and 14 present the flow velocity near the USGS stations predicted in the flow field recovered from the assimilation results, using 30 and 150 s filtering windows, respectively. We compare them with the velocity measured at two USGS stations for validation.

As described in the work of *Domeneghetti et al.* [2013], streamflow hydrograph relative to a specific gauging station and flood event is calculated by converting measured water level into flow rate by means of an existing stage-discharge relation, or rating-curve. The curve is generally calibrated over a series of $h(t)$ to $Q(t)$ pairs, where $h(t)$ is the water level measured at time t and $Q(t)$ is the concurrent river discharge, which, in turn, is often estimated through the velocity-area method [*Fenton and Keller, 2001*]. Note that the discharge-velocity relationship can result in rather rough flow estimation. This is exactly one of the advantages of using Lagrangian data in a river system, as the flow velocity is indeed an explicit measurement.

6.3. Assimilation With Both Lagrangian and Eulerian Data

We also have the local water depth measured at USGS stations included in the data assimilation, along with the filtered Lagrangian drifter trajectories.

Figure 15 demonstrates a remarkable improvement of the flow field estimation, over the estimation results which only include drifter trajectories. The assimilated results are significantly closer to the Eulerian flow velocities at two USGS stations.

6.4. Validation With Reference Drifters

In the two previous cases, the assimilated flow velocity is correlated with the Eulerian velocity at USGS stations. It is possible, however, to validate the assimilation method with another set of Lagrangian drifter data as well, which provides an additional validation procedure, given that no data provide “ground truth.”

We divide the Lagrangian drifters into two groups: one is for our regular data collection, from where the flow state is essentially derived; the other group, namely the Reference Group, is for validation purposes only, i.e., the drifter velocity acquired from these devices is correlated with the assimilated flow results.

Figure 16 shows the assimilation results with the Lagrangian data inputs from half of the drifter fleet, as well as the drifter data collected from the other half fleet (i.e., the Reference Group). A good agreement is demonstrated between the two data sets.

In addition, Figure 17 shows that the assimilated drifter trajectory is very close to the measured data from the Reference Group, with a small fluctuation.

6.5. Error Analysis

The difference between the assimilated data and measurements is further evaluated by computing the relative error ℓ^2 norm:

$$\ell^2 = \left(\frac{\sum_{i=1}^N (\hat{u}_i - u_i)^2}{\sum_{i=1}^N (u_i)^2} \right)^{\frac{1}{2}} \tag{21}$$

where u_i is the measured flow variable of interest (e.g., the flow velocity from Eulerian sensors in this study), \hat{u}_i is the estimated flow variable, for $i=1$ to N measurement events.

Table 1 lists the ℓ^2 norm values of 12 numerical experiments. It shows that the proposed approach possesses good flow estimation accuracy, especially when the Eulerian data are also included in the assimilation. Also, more drifters essentially carry more flow information into the assimilation, and consequently improve the quality of flow estimation. Another factor contributes to the estimation accuracy is the filtering time. The larger the time window (meaning the flow is observed in a coarser timeframe), the flow state exhibits more steady, and thus, not surprisingly, the flow estimation becomes more accurate.

The observation errors can be specified in two categories: random error and systematic error. GPS signals, for example, are treated as observations with random error. This inaccuracy is inherited from the GPS clock, as signals can be distorted in atmospheric disturbances, or reflected from buildings and other large solid objects, before they reach a receiver [Ward, 1997]. The random noise is usually assumed to follow the Gaussian distribution. Wind forcing and waves driven by passing boat, on the other hand, are considered as factors with systematic error.

Two more numerical experiments were conducted to quantify the sensitivity of the proposed method to the observation errors. Figure 18 shows that when the data contain random errors, the proposed method is able to remove the noise and recover the flow field; however, when the data contained systematic errors, the proposed method constructs a flow field with many deficits with oscillations. In some cases, the systematic error is so big that it might even cause the assimilation processes converge to some unreasonable values. In the case of 5% systematic errors, the proposed method leads to an analysis error around 4% (Table 2).

We also noticed that the mean values of the estimated velocity of Sacramento River in Figures 13–15 are a bit off from the USGS measurements. This is likely due to the system error from GES measurements, or some consistent wind effect in the Sacramento River.

6.6. Discussions on Assimilation Results

The assimilation results validate not only the suitability of the proposed flow estimation method using Lagrangian data, but also the effectiveness of different data filtering windows. It is noted that the estimation quality of the flow state in the hydrodynamic system may be

Table 2. Relative RMS Error ℓ^2 -Value for Assimilated Flow Velocity in the Presence of Drifter Data Errors

Error Type	5% Random Noise	5% Systematic Noise
ℓ^2	1.3%	3.86%

further improved if Eulerian information, when available, is also included in the data assimilation. And generally, more drifters would result in a better flow estimation.

The choice of the space-time window affects the results of the filtered data significantly. As shown in Figures 13 and 14, the spatial dimension of the windows is set to be 15 m, and two different time windows of 30 and 150 s are specifically selected in our field test. The plots indicate that assimilated flow velocity estimates become less “noisy” with a larger filtering window. The larger filtering window (meaning more drifter information is “averaged”) results in less fluctuations in the filtered Lagrangian data, and consequently a smoother flow state. This trend is consistent with that of the drifter data. Moreover, the signal quality is improved, and the signal magnitude is at the correct scale. We speculate that noise is either inherited from the Lagrangian data controlled by unknown physical processes, or affected by any erroneous drifter measurements, which can make the data assimilation results converge to unreasonable values.

Another interesting observation in Figures 13 and 14 is the effect of different drifter numbers involved in the data assimilation, i.e., half versus full fleet. In general, the more drifters included in the assimilation, the less noisy and the more plausible the flow estimation is. This is not surprising, since more drifters can essentially carry more flow information into the assimilation process.

7. Summary

In this article, we presented the flow estimation for complex channel networks using Lagrangian measurement data. This work is, to our knowledge, among the first successful applications of variational Lagrangian data assimilation in tidal-driven environment. The solution is formulated as an optimization problem based on minimizing the difference between measured Lagrangian data and modeled drifter trajectories, constrained by a one-dimensional implicit linear channel network model. The major advantage of the proposed formulation is that it requires low-computational cost, making the proposed data assimilation method applicable to many vast and complex hydraulic networks.

The effectiveness of the method has been validated with a field experiment in the Sacramento-San Joaquin Delta, in which the Lagrangian drifter data were collected from GPS equipped drifters and processed with proper filtering and assimilation method. Due to the accessibility restrictions, our drifters could only be deployed in one specific branch of the Delta channels. In addition, we did not have enough manpower to monitor and track the entire drifter fleet over a full tidal cycle. As a result, the experiment was constrained both temporally and spatially. However, this prototype approach is extendable to a full tidal cycle analysis, which will be composed of several repeats of drifter deployment and measurement.

In the future, we plan to extend the data assimilation techniques in a two-dimensional shallow water model, and apply statistical filtering to achieve a more accurate representation of the flow state in a tidal channel network.

Acknowledgments

The authors would thank many members of the UC Berkeley community who have contributed to the Georgiana Slough field experiment described in this article, in particular Mohammad Rafiee, Samitha Samaranayake, Jerome Thai, Agathe Benoit, Axel Parmentier, and Paul Borokhov. We also thank David Bell and his colleagues at the Romberg Tiburon Research Center of San Francisco State University, who provided boating support for the field experiment. Special thanks to three anonymous reviewers for their invaluable comments on this manuscript.

This work was supported by NSF awards CNS-0845076, CNS-0931348, CNS-0615299, and California Sea Grant funding.

References

- Albuquerque, J. S., and L. T. Biegler (1996), Data reconciliation and gross-error detection for dynamic systems, *AIChE J.*, 42(10), 2841–2856.
- Anthes, R. A. (1974), Data assimilation and initialization of hurricane prediction models, *J. Atmos. Sci.*, 31, 702–719.
- Beard, J., K. Weekly, C. Oroza, A. Tinka, and A. M. Bayen (2012), Mobile phone based drifting lagrangian flow sensors, in *2012 IEEE 3rd International Conference on Networked Embedded Systems for Every Application (NESEA)*, pp. 1–7, IEEE Press, Liverpool, U. K.
- Budhiraja, A., L. Chen, and C. Lee (2007), A survey of numerical methods for nonlinear filtering problems, *Physica D*, 230(1), 27–36.
- Chau, K. W., and J. H. Lee (1991), Mathematical modelling of Shing Mun river network, *Adv. Water Resour.*, 14(3), 106–112.
- Cowen, E., and S. Monismith (1997), A hybrid digital particle tracking velocimetry technique, *Exp. Fluids*, 22(3), 199–211.
- CPLX, ILOG (2007), *11.0 Users Manual ILOG SA*, Gentilly, France.
- De Serres, B., A. G. Roy, P. M. Biron, and J. L. Best (1999), Three-dimensional structure of flow at a confluence of river channels with discordant beds, *Geomorphology*, 26(4), 313–335.
- Domeneghetti, A., S. Vorogushyn, A. Castellarin, B. Merz, and A. Brath (2013), Probabilistic flood hazard mapping: Effects of uncertain boundary conditions, *Hydrol. Earth Syst. Sci.*, 17(8), 3127–3140.
- Doucet, A., N. Freitas, and N. Gordon (Eds.) (2001), *Sequential Monte Carlo Methods in Practice*, Springer, Sci.+Business Media, N. Y.
- Evensen, G., and P. J. Van Leeuwen (2000), An ensemble Kalman smoother for nonlinear dynamics, *Mon. Weather Rev.*, 128(6), 1852–1867.
- Fan, S., L.-Y. Oey, and P. Hamilton (2004), Assimilation of drifter and satellite data in a model of the northeastern gulf of Mexico, *Cont. Shelf Res.*, 24(9), 1001–1013.
- Fenton, J. D., and R. J. Keller (2001), The calculation of streamflow from measurements of stage, *Tech. Rep. 01/6*, Coop. Res. Cent. for Catchment Hydrol., Melbourne, Australia.
- Fischer, H. B., E. J. List, R. C. Y. Koh, J. Imberger, and N. H. Brooks (1979), *Mixing in rivers, Mixing in Inland and Coastal Waters*, pp. 104–147, Elsevier BV, Amsterdam, Netherlands.
- Fisher, M., and D. Lary (1995), Lagrangian four-dimensional variational data assimilation of chemical species, *Q. J. R. Meteorol. Soc.*, 121(527), 1681–1704.

- Flatau, M. K., L. Talley, and P. P. Niiler (2003), The North Atlantic oscillation, surface current velocities, and SST changes in the subpolar North Atlantic, *J. Clim.*, *16*(14), 2355–2369.
- Garraffo, Z. D., A. J. Mariano, A. Griffa, C. Veneziani, and E. P. Chassignet (2001), Lagrangian data in a high-resolution numerical simulation of the North Atlantic: I. Comparison with in situ drifter data, *J. Mar. Syst.*, *29*(1), 157–176.
- Griffa, A. (1996), Applications of stochastic particle models to oceanographic problems, in *Stochastic Modelling in Physical Oceanography*, edited by R. J. Adler, P. Müller, and B. L. Rozovskii, pp. 113–140, Birkhäuser, Boston, Mass.
- Gunson, J. R., and P. Malanotte-Rizzoli (1996), Assimilation studies of open-ocean flows: Error measures with strongly nonlinear dynamics, *J. Geophys. Res.*, *101*(C12), 28,473–28,488.
- Honnorat, M., J. Monnier, and F.-X. Le Dimet (2009), Lagrangian data assimilation for river hydraulics simulations, *Comput. Visual Sci.*, *12*(5), 235–246.
- Ide, K., P. Courtier, M. Ghil, and A. Lorenc (1997), Unified notation for data assimilation: Operational, sequential and variational, *Practice*, *75*(1B), 181–189.
- Inc., G. (2013), *Using the Location Manager*. [Available at <http://developer.android.com/training/basics/location/locationmanager.html>.]
- Kamachi, M., and J. O'Brien (1995), Continuous data assimilation of drifting buoy trajectory into an equatorial Pacific Ocean model, *J. Mar. Syst.*, *6*(1), 159–178.
- Kuznetsov, L., K. Ide, and C. Jones (2003), A method for assimilation of Lagrangian data, *Mon. Weather Rev.*, *131*(10), 2247–2260.
- Litrico, X., and V. Fromion (2004a), Simplified modeling of irrigation canals for controller design, *J. Irrig. Drain. Eng.*, *130*(5), 373–383.
- Litrico, X., and V. Fromion (2004b), Frequency modeling of open-channel flow, *J. Hydraul. Eng.*, *130*(8), 806–815.
- Litrico, X., and V. Fromion (2006), Boundary control of linearized Saint-Venant equations oscillating modes, *Automatica*, *42*(6), 967–972.
- Litrico, X., and V. Fromion (2009), *Modeling and Control of Hydrosystems*, Springer, Sci.+ Business Media, London, U. K.
- Miller, R. N., M. Ghil, and F. Gauthiez (1994), Advanced data assimilation in strongly nonlinear dynamical systems, *J. Atmos. Sci.*, *51*(8), 1037–1056.
- Molcard, A., L. I. Piterberg, A. Griffa, T. M. Özgökmen, and A. J. Mariano (2003), Assimilation of drifter observations for the reconstruction of the Eulerian circulation field, *J. Geophys. Res.*, *108*(C3), 3056, doi:10.1029/2001JC001240.
- Molcard, A., A. Griffa, and T. M. Özgökmen (2005), Lagrangian data assimilation in multilayer primitive equation ocean models, *J. Atmos. Oceanic Technol.*, *22*(1), 70–83.
- Molcard, A., A. C. Poje, and T. M. Özgökmen (2006), Directed drifter launch strategies for Lagrangian data assimilation using hyperbolic trajectories, *Ocean Modell.*, *12*(3), 268–289.
- Müller-Solger, A. B., A. D. Jassby, and D. C. Müller-Navarra (2002), Nutritional quality of food resources for zooplankton (daphnia) in a tidal freshwater system (Sacramento-San Joaquin river delta), *Limnol. Oceanogr.*, *47*(5), 1468–1476.
- Navon, I. M. (1986), A review of variational and optimization methods in meteorology, in *Festive Volume of the International Symposium on Variational Methods in Geosciences*, vol. 5, pp. 29–34, Elsevier Sci.
- Navon, I. M. (2009), Data assimilation for numerical weather prediction: A review, in *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, edited by S. K. Park and L. Xu, pp. 21–65, Springer, Berlin.
- Nodet, M. (2006), Variational assimilation of Lagrangian data in oceanography, *Inverse Problems*, *22*(1), 245.
- Oroza, C., A. Tinka, P. K. Wright, and A. M. Bayen (2013), Design of a network of robotic Lagrangian sensors for shallow water environments with case studies for multiple applications, *Proc. Inst. Mech. Eng., Part C*, 0954406213475947 pp., SAGE Publications.
- Paniconi, C., M. Marrocu, M. Putti, and M. Verbunt (2003), Newtonian nudging for a Richards equation-based distributed hydrological model, *Adv. Water Resour.*, *26*(2), 161–178.
- Poulain, P.-M., D. S. Luther, and W. C. Patzert (1992), Deriving inertial wave characteristics from surface drifter velocities: Frequency variability in the tropical pacific, *J. Geophys. Res.*, *97*(C11), 17,947–17,959.
- Salman, H., L. Kuznetsov, C. Jones, and K. Ide (2006), A method for assimilating Lagrangian data into a shallow-water-equation ocean model, *Mon. Weather Rev.*, *134*(4), 1081–1101.
- Salman, H., K. Ide, and C. K. Jones (2008), Using flow geometry for drifter deployment in Lagrangian data assimilation, *Tellus, Ser. A*, *60*(2), 321–335.
- Sanders, B. F. (2001), High-resolution and non-oscillatory solution of the St. Venant equations in non-rectangular and non-prismatic channels, *J. Hydraul. Res.*, *39*(3), 321–330.
- Sanders, B. F., and N. D. Katopodes (1999), Control of canal flow by adjoint sensitivity method, *J. Irrig. Drain. Eng.*, *125*(5), 287–297.
- Stewart, R., and E. Pfann (1998), Oversampling and sigma-delta strategies for data conversion, *Electron. Commun. Eng. J.*, *10*(1), 37–47.
- Strub, I. S., J. Percelay, M. T. Stacey, and A. M. Bayen (2009), Inverse estimation of open boundary conditions in tidal channels, *Ocean Modell.*, *29*(1), 85–93.
- Tinka, A., I. Strub, Q. Wu, and A. M. Bayen (2009), Quadratic programming based data assimilation with passive drifting sensors for shallow water flows, in *Proceedings of the 48th IEEE Conference on Decision and Control, 2009 Held Jointly With the 2009 28th Chinese Control Conference (CDC/CCC 2009)*, pp. 7614–7620, IEEE Press, Piscataway, N. J.
- Tinka, A., I. Strub, Q. Wu, and A. M. Bayen (2010), Quadratic programming based data assimilation with passive drifting sensors for shallow water flows, *Int. J. Control*, *83*(8), 1686–1700.
- Tom, B. (1998), Chapter 6: Cross-Section Development Program, in *19th Annual Progress Report on Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Marsh*, State Water Resour. Control Board, Calif. Dep. of Water Resour., Off. of State Water Proj. Plann., Sacramento, Calif.
- Tossavainen, O.-P., J. Percelay, A. Tinka, Q. Wu, and A. M. Bayen (2008), Ensemble Kalman filter based state estimation in 2D shallow water equations using Lagrangian sensing and state augmentation, in *47th IEEE Conference on Decision and Control, 2008. CDC 2008*, pp. 1783–1790, IEEE Press, Piscataway, N. J.
- Van Leeuwen, P. J. (2009), Particle filtering in geophysical systems, *Mon. Weather Rev.*, *137*(12), 4089–4114.
- Van Leeuwen, P. J., and A. Melanie (2013), Efficient fully nonlinear data assimilation for geophysical fluid dynamics, *Comput. Geosci.*, *55*, 16–27.
- Veneziani, M., A. Griffa, A. M. Reynolds, and A. J. Mariano (2004), Oceanic turbulence and stochastic models from subsurface Lagrangian data for the northwest Atlantic ocean, *J. Phys. Oceanogr.*, *34*(8), 1884–1906.
- Ward, N. (1997), *Understanding GPS—Principles and applications*. Elliott D. Kaplan (Editor). £75. ISBN: 0-89006-793-7. Artech House Publishers, Boston & London. 1996, *J. Navig.*, *50*(01), 151, doi:10.1017/s0373463300023730.
- Wu, Q., S. Amin, S. Munier, A. M. Bayen, X. Litrico, and G. Belaud (2007), Parameter identification for the shallow water equation using modal decomposition, in *2007 46th IEEE Conference Decision and Control*, pp. 1584–1590, IEEE Press, Piscataway, N. J.
- Wu, Q., X. Litrico, and A. M. Bayen (2009a), Data reconciliation of an open channel flow network using modal decomposition, *Adv. Water Resour.*, *32*(2), 193–204.
- Wu, Q., M. Rafiee, A. Tinka, and A. M. Bayen (2009b), Inverse modeling for open boundary conditions in channel network, in *Proceedings of the 48th IEEE Conference on Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference (CDC/CCC 2009)*, pp. 8258–8265, IEEE Press, Piscataway, N. J.
- Xie, S.-P., and J. A. Carton (2004), Tropical Atlantic variability: Patterns, mechanisms, and impacts, *Earths Clim.*, 121–142.