Heterogeneous Fleets of Active and Passive Floating Sensors for River Studies

• •

Andrew Tinka

Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720
e-mail: tinka@berkeley.edu

Qingfang Wu

Civil and Environmental Engineering, University of California, Berkeley, Berkeley, CA 94720
e-mail: qingfangwu@berkeley.edu

Kevin Weekly

Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720
e-mail: kweekly@berkeley.edu

Carlos A. Oroza

Civil and Environmental Engineering, University of California, Berkeley, Berkeley, CA 94720
e-mail: coroza@berkeley.edu

Jonathan Beard

Mechanical Engineering, University of California, Berkeley, Berkeley, CA 94720
e-mail: beard.jonathan@gmail.com

Alexandre M. Bayen

Electrical Engineering and Computer Sciences, Civil and Environmental Engineering, University of California, Berkeley, Berkeley, CA 94720
e-mail: bayen@berkeley.edu

Lagrangian sensing for tracing hydrodynamic trajectories is an innovative approach for studying estuarial environments. Actuated Lagrangian sensors are capable of avoiding obstacles and navigating when active and retain a passive hydrodynamic profile that is suited for Lagrangian sensing when passive. A heterogeneous fleet of actuated and passive drifting sensors is presented. Data assimilation using a high-performance computing (HPC) cluster that runs the ensemble Kalman filter (EnKF) is an essential component of the estuarial state estimation system. The performance of the mixed capability fleet and the data assimilation backend is evaluated in the context of a landmark 96-unit river study in the Sacramento-San Joaquin Delta region of California. © 2015 Wiley Periodicals, Inc.

1. INTRODUCTION

Understanding the movement of water through river and estuarial environments is critical for many environmental management problems. Examples include predicting the outcome and impact of silt disturbed by dredging operations, maintaining the health of fish populations by understanding the factors that affect their migratory patterns, and assessing the vulnerability of freshwater resources to contaminant release or other unpredicted events.

On May 9, 2012, the Floating Sensor Network (FSN) team deployed 28 motorized, active drifters and 68 passive drifters in the Sacramento River near its junction with the Georgiana Slough, near the town of Walnut Grove, California. Figure 1 shows the fleet of drifters. The operation demonstrated the communication, obstacle avoidance, navigation, and data-gathering capabilities of the FSN fleet and gathered flow data for use in demonstrations of an online ensemble Kalman filter- (EnKF-) based assimilation using a HPC cluster. A smaller-scale pilot experiment was conducted earlier, on April 12, 2012 (28 days previously, with roughly similar tidal conditions). This article describes the experimental method, gives an overview of the server and assimilation infrastructure, and presents the results of fleet movement analysis and preliminary hydrodynamics assimilation results.

1.1. Drifting Lagrangian Sensors

In situ sensing refers to sensing techniques in which a device is in direct contact with the environmental phenomena it measures. (In contrast, *remote* sensing refers to techniques such as analysis of satellite imagery, in which measurements

Figure 1. Drifter fleet on the Walnut Grove Public Dock on May 9, 2012, prior to deployment. In the background, with yellow tops, are actuated drifters, described in Section 2.1; in the foreground, with orange tops, are passive drifters, described in Section 2.2. Photo credit: Jérôme Thai.

are taken from afar.) *In situ* sensing in fluid environments is classified into *Eulerian* and *Lagrangian* techniques, using the terminology for the different reference frames in hydrodynamics. Eulerian sensors are fixed to the external reference frame (e.g., the river bank) and take measurements from the water as it passes. Lagrangian sensors float freely in the fluid itself and gather measurements about the water as it moves along a trajectory.

Although some Lagrangian sensors may measure physical characteristics of the water in which they are immersed (e.g., dissolved constituents, temperature), the distinguishing characteristic of a Lagrangian sensor is that it moves freely in the water and measures its position over time. A well-designed Lagrangian sensor should act like an "ideal particle" in the water flow, and so the time series of its position should allow direct estimation of the velocity of the water in which it was immersed. This "trajectory" information is a useful observation of water movement in an environment. In the hydrodynamics literature, such sensor devices are called *drifters*. Drifter design has always been constrained by the positioning and communications technologies available. Modern oceanography began using drifters based on underwater acoustic communication in the 1950s (Swallow, 1955). Acoustic technology dominated until 1978, when the Argos satellite service gave oceanographic researchers a global location and data uplink system (Clark, 1989). Power, cost, and size constraints meant that Argos-based drifters (Davis, 1985; Bitterman & Hansen, 1986; Niiler, Davis, & White, 1987) were better suited to

oceanography than inland environments such as rivers and estuaries.

Positioning techniques relying on acoustics or the Argos system accurate has given way to GPS positioning, which is low cost and low power, driven by technological advances in consumer products. Therefore, inland studies with smaller, lower-cost units are made possible. Furthermore, short-range radio RF communication links are suitable technologies for inland studies because radio base stations can be set up with nearby infrastructure. GPS-carrying river drifters have been the focus of development by the FSN project at University of California, Berkeley (Tinka, Rafiee, & Bayen, 2013), and other groups (Perez, Bonner, Kelly, & Fuller, 2003; Austin & Atkinson, 2004). As well, studies in regions with well-developed civilian infrastructure, such as the continental United States, can take advantage of the mobile phone network for communications.

1.2. The Need for Actuated Drifters

Passive drifters have proved very successful for oceanographic applications, where fixed infrastructure can be impractical to install. A logical extension of the oceanographic research in Lagrangian sensing is to develop sensors for nearshore environments such as rivers and bays. Lagrangian sensors can be used in this context to better monitor the flow of freshwater and the transport of constituents therein. Specific examples include assessing vulnerabilities to contaminant spills or infrastructure failure in

critical water-resource regions, planning reservoir release and gate control policies to affect the intrusion of salt water, and monitoring the effect of heavy agriculture use on fresh water supplies.

For the river environment, however, fixed infrastructure is more viable to install, and drifters, without supervision, are likely to become stuck on obstacles, rendering them useless for flow tracking. Therefore, this vision requires long-term operation of the drifters, which in turn requires some ability to avoid obstacles such as debris or shallow regions. Thus, our design has added motors to the floating units, giving them the ability to avoid these hazards. The FSN project is the first to design motorized drifters that are both low cost and manufacturable, leading to the production of a fleet of 40 and a demonstration of their success in a field experiment.

Actuation, of course, is directly at odds with the primary goal of a Lagrangian sensor, namely, to match velocity with the surrounding water to measure its velocity and capture its trajectory. When the actuated drifter uses its propellers to move, its velocity differs significantly from the water. Actuation should therefore be used sparingly to increase the utility of the sensor; also, the estimation methods that use the drifter's sensed velocity must be flexible enough to handle the actuated drifter's trajectory. The control scheme for actuation and the assimilation method for actuated drifters are described in sections 4 and 3, respectively.

Although the active drifters present a solution to the obstacle and navigation challenge, their greater complexity leads to higher cost and difficulty of operation. The approach presented in the present article finds a practical solution via designing a heterogeneous fleet of active and passive drifters, combined with a careful experimental design to leverage the strengths of each over the experimental domain.

1.3. Data Assimilation and the Ensemble Kalman Filter

The main difficulty of using Lagrangian data is due to the well-known fact that Lagrangian motion is often affected by local flow perturbations, and Eulerian and Lagrangian behaviors are not simply related to each other. One way of looking at the data assimilation process is as a "translation" of the observed data from a Lagrangian framework to an Eulerian framework, as well as reconciling the observed data with other known information about the system.

Our objective is to develop a sensing-modeling system capable of predicting regional flows and transport in the Sacramento-San Joaquin Delta in real time without dependence on historical data. One major advantage of drifting sensors is that they can be rapidly deployed and retrieved in a wide variety of environments. To best take advantage of this flexibility, the data assimilation techniques used with the drifters should not require extensive model tuning or calibration (steps that typically involve gathering large amounts of historical data and repeatedly running the model until it performs well).

Data assimilation has been widely used in river hydraulics, hydrology, atmospheric sciences, and oceanography, in order to provide forecasts based on the solutions of underlying partial differential equations by assimilating measurement data into these models (Evensen, 2009; Kalnay, 2003; Nodet, 2006; Castaings et al., 2006; Romanowicz, Young, & Beven, 2006; Vrugt & Robinson, 2007; Lehikoinen, Huttunen, Finsterle, Kowalsky, & Kaipio, 2010; Honnorat, Monnier, & Le Dimet, 2009; Wang, Chau, Cheng, & Qiu, 2009).

In recent years, variants of the EnKF proposed by Evensen (Evensen, 1994) are becoming the main data assimilation techniques in atmospheric and oceanic sciences (Van Leeuwen & Evensen, 1996; Houtekamer & Mitchell, 1998, 2001; Hamill & Snyder, 2000; Hamill, Whitaker, & Snyder, 2001; Keppenne & Rienecker, 2002).

The main difference between the existing ensemble-based schemes lies in the generation of the analysis ensemble and handling the model/measurement noise. In our application, a systematic Bayesian approach proposed in (Kaipio & Somersalo, 2005; Huttunen & Kaipio, 2007a, 2007b; Nissinen, Kolehmainen, & Kaipio, 2011; Tossavainen, Percelay, Stacey, Kaipio, & Bayen, 2011a) is used to recover from the modeling errors with additive Gaussian noise processes.

The flow model we have chosen is the *River, Estuary, and Land Model* (REALM) flow model developed by the *Lawrence Berkeley National Lab* (LBNL) (Ateljevich et al., 2009). At each time step, the server receives new real data from the floating sensors, and the shallow water model generates a collection of states representing the evolution of the processed inputs. The EnKF server compares these real data to the gathered estimated output in order to provide the best estimation of the flow conditions in the field.

1.4. Contributions of this Article

This article presents a system-level description of a heterogeneous FSN, including actuated drifters, passive drifters, and communication relays; the description of a field operation to demonstrate system functionality and to validate core assumptions about drifter operations in estuarial environments; an analysis of the utility of the actuation capability of the active component of the fleet; the design and implementation of a back-end data assimilation server cluster; and an analysis of the assimilation system performance and output using the data gathered in these experiments. This experiment represents a significant step in the deployment of autonomous floating robotics systems in unstructured natural environments. To our knowledge, this experiment is the first of its kind in the field of estuarial hydrodynamics

Table I. Summary of Floating Sensor Network devices used in 2012 experiments.

Name	Generation 3 (Oroza et al., 2013)	Android (Beard et al., 2012)
Image		
Dimensions	13.0 cm diameter, 47.0 cm tall, 24.0 cm span across motor pods	13.0 cm diameter, 29.0 cm tall
Cost (parts)	$ 2 500	$ 500
Assembly time	10 h	0.25 h
Battery Life	24 h (11.1 V, 10.4 A h Li-ion)	48 h (7.2 V, 16 A h Li-ion)
Computation	Gumstix Overo: 720 MHz, 512 MB RAM, Linux 2.6	Motorola Defy: 800 MHz, 512 MB, Android 2.3
Communications	802.15.4	GSM
Propulsion	Twin parallel propellers	None

and the first that an in-depth analysis of this experiment has been presented.

This article is built on previous contributions related to the FSN project. Previous work essential to the system description is reproduced in summary here. This work includes design studies of the individual drifter devices (Oroza, Tinka, Wright, & Bayen, 2013; Beard, Weekly, Oroza, Tinka, & Bayen, 2012), advances in data assimilation techniques (Wu & Bayen, 2015), and the Hamilton-Jacobi-Bellman-Isaacs partial differential equation- based navigation and obstacle avoidance control system for the actuated drifters (Weekly, Tinka, Anderson, & Bayen, 2014; Mitchell, Bayen, & Tomlin, 2005).

1.5. Overview of Article Structure

The rest of this article is organized as follows. Section 2 describes the field-deployed equipment used for this investigation: active sensor vehicles, passive sensor devices, communication relays, and the communications infrastructure. Section 3 describes the data assimilation methods used on the home servers to process the data gathered by the Lagrangian sensor devices into a state estimate of the river region under investigation. Section 4 documents the active sensor vehicle control for obstacle avoidance and navigation. Section 5 explains the field experiments performed in the Sacramento-San Joaquin Delta region that form the basis for the assimilation and fleet operation validation that are the core of this article. Section 6 evaluates the performance of the actuated and passive sensor fleet during the described experiment, with emphasis on the utility of the actuated vehicles and the results of the data assimilation process. Section 7 concludes with system-level evaluations

of the heterogeneous Lagrangian fleet concept as well as open questions for further investigation.

2. FLOATING SENSOR NETWORK SYSTEM DESCRIPTION

Table I provides a comparison of the two types of drifter used in this study; more details on their design are provided in Sections 2.1 and 2.2. The communications architecture, including installed communication relays necessary to improve field reliability, is described in Section 2.3. An essential part of the experimental system is the data assimilation back end, which is described in Section 3.

2.1. Active Sensor

The design of the active drifters was constrained by a number of functional requirements concerning actuation, form factor, and mission time. The design process is fully documented in (Oroza et al., 2013) and summarized here. Figure 2 shows an annotated front and side view of the active drifter.

The first consideration in designing the sensor was the "form factor," or overall size and shape of the sensor. This is an important constraint because it determines vehicle's effectiveness as a Lagrangian sensor. Previous studies have indicated that Lagrangian sensors must maximize their cross-sectional area to flow and should present a roughly symmetric drag profile (Tinka et al., 2013). To ensure that the vehicle is appropriate for Lagrangian sensing, its hull is a vertically oriented cylinder. It is designed to be hydrostatically stable (with a center of buoyancy (COB) located above its center of mass (COM) and to have a mass-to-volume ratio that keeps it low in the water (only the antennas extend above the waterline).
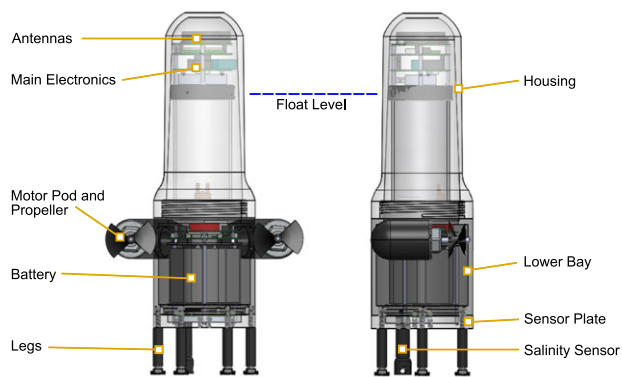
Figure 2. Front and side view of CAD drawing of active drifter; components labeled.

The functional requirements dictated by the form factor result in a large drag coefficient for the sensor, limiting the efficiency of the actuation system. Therefore, the system was designed to enable low-velocity, cross-stream movement and rotation but not upstream movement. The system consists of differential drive motors, located outboard of the main hull, and uses the onboard compass and a PID controller to stabilize the vehicle along a desired heading. The motors and propellers were sized to enable 0.3 m/s maximum forward velocity. The system is powered with an 11.1 V, 10.4 amp-hr lithium-ion (Li-ion) battery, enabling up to 24 hours of mission life with a 10 percent actuation duty cycle.

The upper housing is a repurposed water filter canister housing, while the lower bay and motor pods are constructed from CNC machined acetyl homopolymer (an engineering material also known as Delrin). This high-performance plastic maintains its dimensions when submerged and/or exposed to large variations in temperature, making the sensor suitable for a range of aquatic environments. Most of the actuation components, including the propellers, motors, and motor control boards, are hobbyist remote-control vehicle components. The DC motors are specialized low-speed, high-torque motors, which were chosen for high efficiency at low angular velocity. Parallel propeller shafts allow differential drive control for rotation and forward thrust, which is simpler and more reliable than a rudder or other steering mechanism.

We considered constructing a conical ducted shroud that mounts behind the removable engine pod to protect the propellers and mitigate entanglement problems with river vegetation or other material. In our experimental domain, however, we determined that river vegetation was not an issue as long as the sensors stayed away from the riverbanks. We decided that propeller protection was not worthwhile for this set of experiments.

The lower plate is designed to accept a water-facing sensor, which will be protected from impact by the six lower legs. This plate is modular and can be easily swapped out, allowing a variety of different sensors with varying seals and interfaces. In other experiments, we have used acoustic depth sensors, pressure transducers, salinity, and temperature sensors. In the experiments described in this article, the water-facing sensor was not used. The main sensing modality was the GPS positioning system.

The upper electronics consists of several off-the-shelf modules integrated with custom printed circuit boards (PCBs). A GPS module provides position data. Communications with the field team and the home server is provided by a short-range 802.15.4 radio and a GSM module; the communications scheme is explained in detail in Section 2.3. Orientation is determined with an electronic compass module. Low-level control of the motors and compass is performed with an Atmel XMEGA microcontroller, while mission control, data storage, and communications are handled by a Gumstix Overo embedded computer. The main computational load is the map-driven obstacle avoidance and navigation function, which is described in Section 4.

These actuated drifters are capable of autonomous navigation and movement; as such, they are mobile robots. We choose to refer to them as "active drifters" as opposed to "autonomous surface water vehicles" to reflect the functional priorities that motivated their design. The vertical cylinder form factor makes them excellent Lagrangian drifters but severely limits their performance as autonomous vehicles. In particular, their maximum speed is smaller than the water currents in which they are deployed. This is a particularly severe form of underactuation; upstream points, in general, are out of reach. In Section 4, we describe the control schemes for navigation and obstacle avoidance that are appropriate for these vehicles' capabilities.

Thus far, more than 40 active sensors have been produced and have been used in multiple experiments in the Sacramento-San Joaquin Delta region. Most of the custom components, including the PCBs and the acetyl homopolymer lower housing components, are fabricated by third-party companies; the final assembly process requires about 10 hr of Berkeley personnel time.

2.2. Passive Sensor

Although the autonomous navigation capabilities of the actuated drifters are important, they impose significant drawbacks: primarily high manufacturing and assembly cost, high maintenance workload, and high operational workload. A second sensor type, the passive drifter, was designed to complement the active drifter. A low-cost, easy-to-use sensor allows the fleet size to be scaled up economically, although the lack of actuation means they must be carefully deployed in river environments. The design and
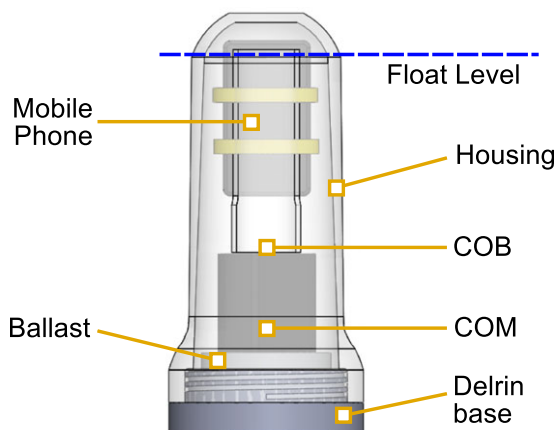
Figure 3. Passive drifter, with components annotated.

development of these devices is completely described in (Beard et al., 2012); an overview is provided here. Figure 3 shows the passive drifter with major components labeled.

The passive drifter design goals were to provide position tracking and communications capabilities similar to the active drifters, while reducing manufacturing and assembly cost, maintenance requirements, and operational workload, and increasing reliability and mission time. Figure 3 shows the passive drifter design, with annotations for the major components, COM, and COB.

Removing the need for actuation made several of these design goals possible. Motors consume far more power than the electronics required for positioning and communication; the operational time was easily extended to 48 hr using a similar battery to the active drifter model. While the active drifter has several static waterproof seals and two dynamic seals (the propeller motor shafts), the passive drifter has just one static waterproof seal; this increases reliability and reduces the maintenance workload.

The major reduction in manufacturing and assembly cost, as well as maintenance workload, came from a drastic simplification of the internal electronics. Although the active drifter integrated many individual functional modules (the GPS, the radio, the motor drivers, etc.) using several custom PCBs, all of the functions of the passive drifter could be accomplished with a mobile phone running specialized software. The mobile phone can gather positioning information using its integrated GPS and communicate this information to the home server using the GSM network. The mobile phone also provides greater reliability than the custom integrated components.

We selected the Motorola Defy smartphone for our passive drifters. In addition to meeting all of our functional requirements and being relatively inexpensive, the Defy is water resistant, which is convenient for field operations. (The passive drifter hull was designed to be watertight, but assembly and disassembly operations in the field always

have the potential for accidents.) At the time of manufacture, the Motorola Defy was the most economical water-resistant mobile phone that would support Android software development and provide the required sensor and communication modules.

Further reductions in costs were achieved by minimizing the number of machined parts; while the active drifter contains over 10 custom-machined components, the passive drifter has just two: the sealing base and the phone holster. A combination of battery requirements and the need to capture subsurface currents lead the mobile-phone-based drifter to retain its tall profile rather than shrink down to the size of the phone itself. By reducing component count and complexity, the final assembly time per passive drifter was reduced to just 15 min.

When deployed appropriately, the operational workload for the passive drifters is much lower than the active drifters. When the field team uses the active drifters, it must use a laptop computer and a radio to control the devices (e.g., to turn off the propellers for operator safety while the drifters are in and around the boat). Because turning off the propellers is safety critical, protocol requires radio confirmations of any commands sent to the drifters to ensure they were received and interpreted correctly. Although the passive drifter is not safety critical, it still needs some input from the operators; knowing when the drifter is in the boat or in the water makes data processing much easier. We leveraged the accelerometer on the mobile phone to provide a simple switching scheme; turning the drifter onto its side signals that it is not in the water and should not gather data. Although this is a very simple feature, it completely eliminates the need to open the drifter in the field or perform complicated over-the-radio communications, which makes it a much easier device to work with.

The FSN project has built 70 passive drifters for internal and external usage. Early experiments have shown that a vessel with a boat captain and two personnel equipped with poled nets (see Figure 12) can deploy, monitor, and retrieve 35–40 passive drifters. During typical experiments, passive drifters are deployed upstream at distance intervals, allowed to float, retrieved, and redeployed if they approach obstacles, and eventually are all retrieved downstream. With two boats, 70 passive drifters, and careful planning, the FSN group was capable of uniformly spanning a 5.5 km stretch of river with passive drifters (Beard et al., 2012).

In certain environments, a passive drifter requires more support than an active drifter. A lack of actuation puts the drifter at risk of washing up on shore; they are usually operated close to personnel and in more obstacle-free waters. Deployment and retrieval must be carefully planned to ensure operational requirements, such as distance traveled, sampled area, sample variation, and time spent sampling, are met, while such variables as danger/safety of personnel are optimized. The field experiment described in
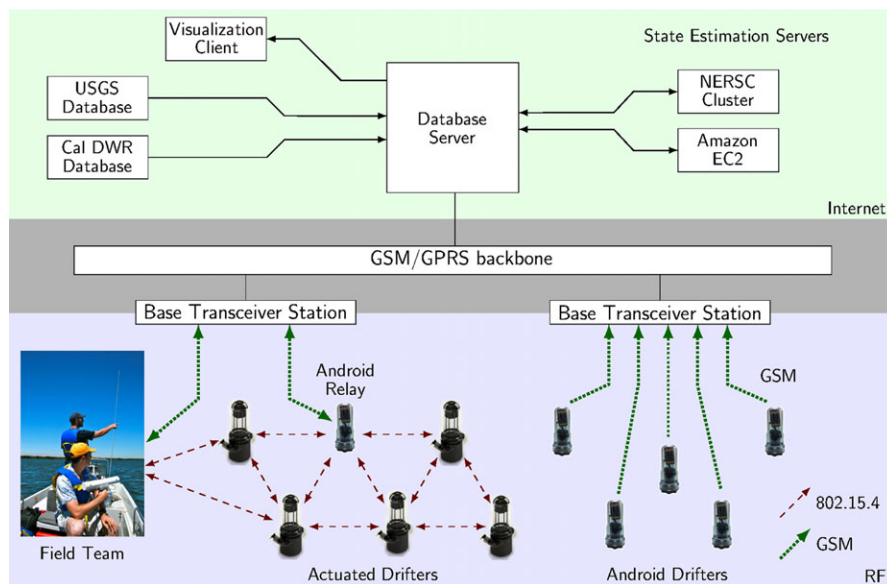
Figure 4. Communication architecture, showing the flow of data from drifters in the field to the database server and computation servers via the GSM service. GSM/GPRS backbone and base transceiver station (mobile phone tower) are provided by a contracted GSM provider.

Section 5 was designed so that the passive drifters would cover lower-risk regions of the experimental domain, while active drifters covered the junction and obstacle-laden region. Combining the fleet in this way allows greater coverage at lower cost than could be achieved by either sensor type on its own.

2.3. Communications Architecture

Figure 4 shows the communication links between various elements of the system. Data collected by the active and passive drifters is communicated back to the database server using the *general packet radio service* (GPRS) of GSM. The Android smartphone on board each passive sensor provides the necessary GPRS functionality. Our original design for the active drifters included two communication modules: a Motorola G24 OEM GSM module for direct communication with the server and a Digi XBee-PRO 802.15.4 module for short-range communication with other drifters and the field team. Reliability issues prevented us from using the G24 GSM module, however, and so the active drifters communicated solely through the XBee module.

The XBee-PRO module conforms to the IEEE 802.15.4–2006 draft standard for low-power mesh networking. Our experience in outdoor environments shows that point-to-point links of 100 m are reliable, and we have seen connectivity at distances of 1 km. To bridge between the 802.15.4 short-range networking and the database servers, we built 10 specialized Android drifters carrying a XBee-PRO module as well as an Android smartphone. These devices, called

"Relays," were put in static locations around the experimental environment. They did not gather data themselves but simply collected the data from the active drifters and transmitted them to the database server via GSM.

Field teams carried laptop computers with GSM modules and XBee-PRO modules. The range of the teams' XBee-PRO modules is extended using a 2.4-GHz Yagi-style directional antenna aimed by the radio operator. The active drifters, which have an internal omnidirectional 2.4-GHz wire antenna, can be sent commands for diagnosis and troubleshooting. Capabilities include enabling and disabling the motors, running or terminating processes on the Gumstix, and querying various values such as mission state or sensor readings. These commands can be sent directly over the XBee link or can be sent indirectly over GSM through the database server and the Android relays. During the April 12, 2012, and May 9, 2012, experiments, no commands were sent to the active drifters while they were under way; they operated autonomously.

The database server acts as a central repository for gathered data and assimilation results. In addition to the Lagrangian data collected by the drifters, relevant data from USGS and the California Department of Water Resources (DWR) sensor stations are collected and stored. The sensor data are sent to the computational cluster, which can be either a collection of Amazon Elastic Computing Cloud (EC2) processors or the NSERC computational cluster. Results from the assimilation process are stored on the same database server and queried by the visualization application, which can be accessed on the Web by any browser.
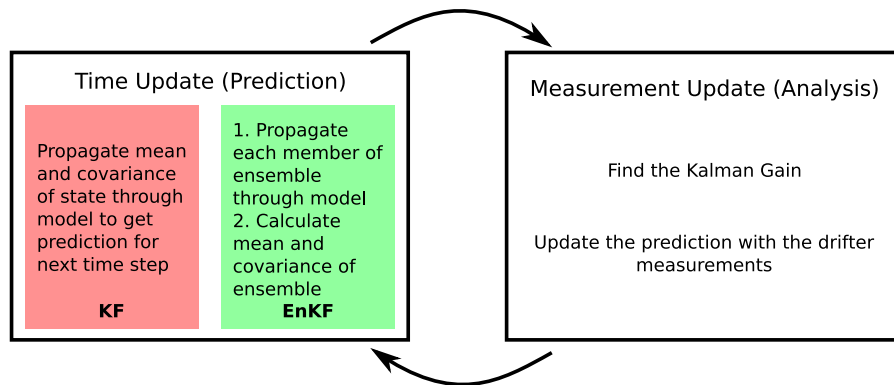
Figure 5. Overview of the Kalman filter and the ensemble Kalman filter. Each one cycles through a time update step and a measurement update step. The principal difference between the two schemes is how the state estimate is represented as a distribution and how this distribution is updated during the time update step. In the Kalman filter, the state estimate distribution is represented by a mean vector and covariance matrix, which can be directly updated by the linear model. In the EnKF, the distribution is represented by an ensemble of samples from the distribution, which are individually updated by the nonlinear model. The mean and covariance are aggregated from the ensemble.

3. ASSIMILATION METHOD AND BACK-END INFRASTRUCTURE

We are aiming to develop a water monitoring system for rivers and estuaries in order to observe the current speed, salinity, and depth of the water on a real-time map. These maps can be used for tracking the water current and salinity as well as other contaminants. These data are relevant for maintaining the hydrodynamic system, such as the Sacramento-San Joaquin Delta as well as in case of floods or water contamination. To provide a real-time representation of the Delta hydrodynamic state, data assimilation technique is applied at the back-end computational cluster. These methods rely on reconstructing Eulerian velocity information from consecutive measurements of drifter positions, which are then assimilated into Eulerian flow models using the EnKF.

The EnKF (Evensen, 1994, 2003) is a sequential data assimilation method related to the Kalman filter and extended Kalman filter. As a Monte Carlo approximation of the classical Kalman filter, EnKF replaces the covariance in the Kalman filter with the sample covariance computed from an ensemble of realizations. The estimated state of a system, and the covariance of that state estimate, is tracked as an *ensemble* of individual states. Each member of the ensemble is individually propagated through a forward model of the underlying system and corrected with new measurement data as it arrives. Because the EnKF does not need to maintain the state covariance matrix, it can be implemented efficiently for high-dimensional problems (Mandel, Cobb, & Beezley, 2011). Figure 5 compares the estimation cycle of the EnKF with the classical Kalman filter.

One of main challenges regarding the assimilation of Lagrangian information into hydrodynamic models is to quantify 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 evaluated as $\delta r / \delta t$. The estimated velocity will be subsequently assimilated into hydrodynamic models. Such a method is usually called pseudo-Lagrangian data assimilation. The method works well when the sampling period δt is much smaller than the Lagrangian correlation time scale.

A summary of the EnKF algorithm follows.

(1) Initialization: Generate an ensemble of N states, $\{\xi_{0|0}^{(i)}\}$, representing the initial estimate of the system. During our initialization process, an ensemble of N states is generated to characterize the uncertainty of the state. For each of the ensemble members, the forward simulation under perturbed boundary conditions has been run for a certain period of time until a stable state is reached. Such a state will be used as $\{\xi_{0|0}^{(i)}\}$, representing the $i - th$ initial estimate of the system.

(2) Time update:

$$\xi_{k|k-1}^{(i)} = F\left(\xi_{k-1|k-1}^{(i)}\right) + w_{k-1}^{(i)}$$

$$\theta_{k|k-1} = \frac{1}{N} \sum_{i=1}^{N} \xi_{k|k-1}^{(i)}$$

$$E_{k|k-1} = \left[\xi_{k|k-1}^{(1)} - \theta_{k|k-1} \quad \cdots \quad \xi_{k|k-1}^{(N)} - \theta_{k|k-1}\right],$$

where F is the forward system model, $\{w_{k-1}^{(i)}\}$ are samples from the model noise process, $\theta_{k|k-1}$ is the

premeasurement estimate at time step k, and $E_{k|k-1}$ is the matrix of deviations from the mean.

(3) Measurement update:

$$\Gamma_{k|k-1} = \frac{1}{N-1} E_{k|k-1} E_{k|k-1}^T$$

$$K_k = \Gamma_{k|k-1} H_k^T \left(H_k \Gamma_{k|k-1} H_k^T + R_k \right)^{-1}$$

$$\xi_{k|k}^{(i)} = \xi_{k|k-1}^{(i)} + K_k \left(y_k - H_k \xi_{k|k-1}^{(i)} + \epsilon_k^{(i)} \right),$$

where $\Gamma_{k|k-1}$ is the covariance of the premeasurement estimate ensemble, K_k is the Kalman gain for time step k, H_k is the observation operator (mapping each element of the observation vector to the state estimate), R_k is the covariance of the observation errors, y_k is the observation vector at time k, $\{\epsilon_{k-1}^{(i)}\}$ are samples from the observation noise process, and $\{\xi_{k|k}^{(i)}\}$ is the ensemble representing the postmeasurement estimate at time step k.

The drifters' observed water velocities at each time step form the observation vector, y_k. The locations of the drifters at that time are encoded in the observation operator, H_k. An important advantage of this formulation is that elements of the observation vector can be "disconnected" from the measurement update step by setting the relevant row of the H_k matrix to zero. This allows us to disregard individual drifters when necessary, specifically, when they are applying actuation and their movement does not match the water flow. When a drifter finishes moving and resumes unactuated drifting movement, it is reconnected to the measurement update via nonzero elements in the H_k matrix, and its measurements once again have influence on the state update. No additional steps to reinitialize or reset the assimilation procedure are necessary; drifters can be connected or disconnected from the measurement update step as needed throughout the assimilation process.

In this implementation, the forward model of the system is the REALM hydrodynamic model (Ateljevich et al., 2009) of the Sacramento-San Joaquin Delta. Simulating the forward evolution of each member of the ensemble represents a significant computational load. Fortunately, like many Monte Carlo computations, this phase of the algorithm is "embarrassingly parallel." Each forward simulation has no dependencies on the other members of the ensemble, and so each forward simulation can be delegated to a node of a computational cluster with very little data interchange or communication.

The REALM model is established on a Cartesian grid with embedded boundary discretization of the shallow-water equations. Users of the REALM model have different options for the grid size and spacing. We used a regular grid with fixed size, which turned out to work well. An adaptive grid with varying step sizes could also have been chosen and would possibly be more feasible if an even

larger hydrodynamic system were evaluated. Any hydrodynamic model applied to real-world, unstructured environments must be able to handle irregular boundaries. REALM can faithfully represent the shoreline via grid generation from digital elevation models. Other structures, such as gates, can be represented via implicit functions and constructive solid geometry. These techniques are known to be robust and efficient when applied to PDE finite volume models like REALM (Ligocki, Schwartz, Percelay, & Colella, 2008).

The main idea on the cluster side is to run multiple instances of the shallow water flow model (i.e., REALM) in parallel on a computer cluster. Each model instance communicates its output to the server, which executes data assimilation using the EnKF algorithm every 3 min. We employed parallel computation and adaptive mesh refinement for rapid computation.

A computer cluster used in HPC consists of a group of computers linked to each one another in order to behave as a single system. The computers are interconnected via a high-speed network, and each one contains homogeneous hardware and software. The main objective of HPC clusters is to use the processing power of multiple nodes in parallel. This parallelization requires communication between the nodes while processing the tasks if the tasks are dependent.

The *National Energy Research Scientific Computing Center* (NERSC), located at Lawrence Berkeley National Laboratory, maintains and operates Carver, a HPC cluster of 1,202 nodes, each consisting of 8 Intel Nehalem processors with 4 GB of memory (NERSC, 2012). An overview of the data flow from the drifters in the field to the Carver computational nodes is given in Figure 6. The drifters make TCP connections to a single FSN computer, which aggregates their incoming sensor data and sends it to a Postgres database within Carver. The database acts as a central repository for gathered data and assimilation results. In addition to the Lagrangian data collected by the drifters, relevant data from USGS and DWR sensor stations is collected and stored.

These computations are done using real-time data retrieved by the drifters in the water. The new real-time measurements provided by the drifters are then assimilated with the outputs of each flow model running on the client nodes. Once the data assimilation is done, each model's state is updated and sent back to its corresponding client to start a new cycle. The network communication is managed using a Java application. Figure 6 shows the data flow in the system. As mentioned earlier, the individual forward simulations are executed on nodes of the Carver cluster. Two simulations are executed on each of 250 nodes, for a total ensemble size of 500. The rest of the algorithm is then executed on a single "Server" node; every 3 min, the results of the individual simulations are aggregated and the "Measurement Update" step is performed on the server node. At this load, the system can operate at real-time speeds. Results from the assimilation process are finally stored on the same database
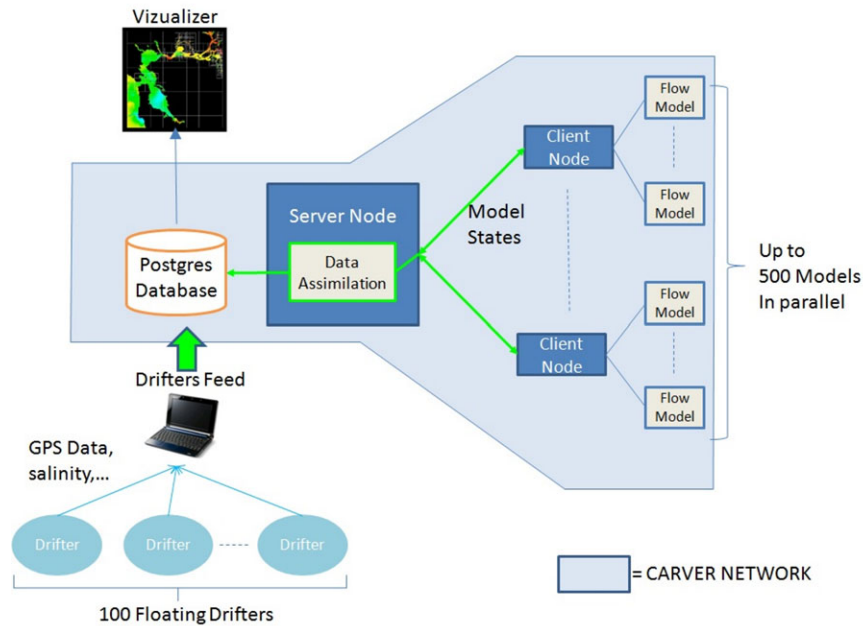
Figure 6. Data flow from drifters in the field to the Carver cluster at NERSC.

server and queried by the visualization application, which can be accessed on the Web by any browser.

4. ACTIVE VEHICLE CONTROL

The Generation 3 drifters were designed to mobilize proactively to avoid large static obstacles such as the shoreline or docks. The obstacle avoidance technique uses a pair of *minimum-time-to-reach* (MTTR) functions, numerically defined on a grid over the experimental domain. For any position in the domain, the first MTTR function gives the minimum time for an unpowered drifter to reach an obstacle, given a worst-case disturbance bounded by a maximum speed (chosen as 0.05 m/s). The disturbance represents known and unknown factors that could push the drifter toward undesirable locations, which, for robustness, must be assumed to *always* push the drifter undesirably. The disturbance is not the river current, which is incorporated into the dynamics of the drifter but can include *deviations* from the estimated river current. The second MTTR function gives a minimum time for an active drifter, using its maximum control authority (chosen as 0.2 m/s) in the optimal direction, to reach the "safest" area of the river, typically the center. The negative gradient of this function is also the optimal control input to reach the safe area.

The MTTR functions for a static flow scenario are derived from numerical solutions to *Hamilton-Jacobi-Bellman-Isaacs* (HJBI) equations, calculated offline prior to the experiment (Weekly et al., 2014). The results are saved into a "policy file" and loaded onto the drifters (See Figure 7 for a description of the binary file). The 3347 m × 3179 m domain shown in Figure 8, is discretized into a 1674 × 1590 grid of 2 m × 2 m cells and generates a 13.3 MB-sized policy file. Note that although the policy files are static by nature, logic on the drifters allows switching between different policy files based on time of day or on their numerical ID. This allows us the method to account for tidal cycles over the day and also to separate the fleet into groups by their IDs, which enables the results shown in Figure 14.

The input data to the offline policy generation algorithm (i.e., the algorithm that computes the HJBI solution) are three files. First is the flow-field estimate, typically provided by REALM, which gives the river velocity on a grid defined over the river areas in the experimental domain. Second is a "center-line" map, a binary image drawn manually and representing the safest area of the river, which is typically the geographical center, unless we have prior knowledge such as the topology of the river bed being asymmetrical. Finally, an obstacle map is provided, which is a binary image drawn manually over the geometry of the river and defines the regions that the drifter should never enter, such as landmasses and docks.

The obstacle avoidance technique also allows the drifter to choose one or more possible paths to travel down if there is a junction in the river. This is accomplished by adjusting the two binary images described above. First, we manually annotate the 'center-line map to only proceed down the desired path, then we incorporate a virtual obstacle into the obstacle map, blocking the undesirable path. Therefore, the algorithm will determine a control strategy
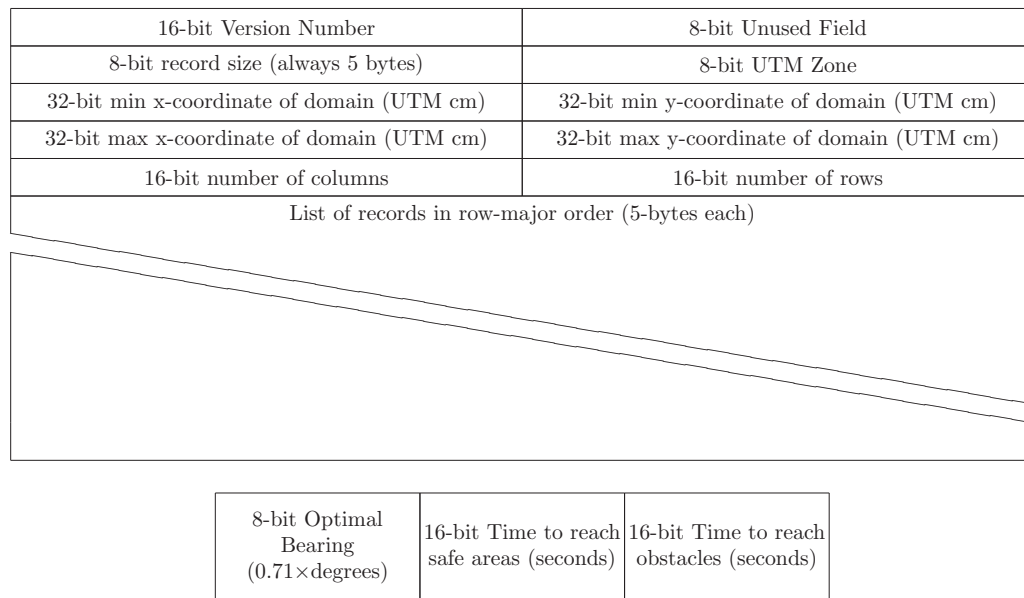
16-bit Version Number	8-bit Unused Field
8-bit record size (always 5 bytes)	8-bit UTM Zone
32-bit min x-coordinate of domain (UTM cm)	32-bit min y-coordinate of domain (UTM cm)
32-bit max x-coordinate of domain (UTM cm)	32-bit max y-coordinate of domain (UTM cm)
16-bit number of columns	16-bit number of rows
List of records in row-major order (5-bytes each)	

8-bit Optimal Bearing (0.71×degrees)	16-bit Time to reach safe areas (seconds)	16-bit Time to reach obstacles (seconds)

Figure 7. Binary format of policy files. Top: Field description of policy file. Bottom: Description of the 5-byte fields corresponding to cells of the discretized domain.

that avoids that area, the same as it would avoid the shoreline. For the May 9, 2012, field experiment, which included a junction in the experimental domain, we divided our active drifters into two groups, and they were given separate control strategies. For the Georgiana Slough, we adjusted the obstacle map to fill in the part of the Sacramento River west of the junction. Likewise, for the Sacramento River group, we filled in the Georgiana Slough areas of the obstacle map.

In Figure 8 are pictorial representations of the three data inputs used for the May 9, 2012, field experiment in Walnut Grove. The flow-field input is a grid of $2\,\text{m} \times 2\,\text{m}$ cells containing vector-valued velocities with a maximum magnitude of 2.33 m/s. Figure 8a shows this data, subsampled to a $60\,\text{m} \times 60\,\text{m}$ cell size. The data that were available for the region only specified the velocities in a small area around the split; therefore, we specified the rest of the needed data as no-flow regions. Our hypothesis, verified in simulation (Weekly et al., 2014), is that errors in the flow-field file are irrelevant for straight areas of the river for the purpose of obstacle avoidance. Figure 8b shows the two binary images for the Georgiana Slough path: The green region illustrates the obstacle data, and the black line down the center gives the target region toward which the drifter should navigate when it is in danger. Figure 8c illustrates the two binary images for the Sacramento River path.

5. EXPERIMENT DESCRIPTION

Experiments were conducted on April 12, 2012, and May 9, 2012, at the junction of the Sacramento River and Georgiana Slough near Walnut Grove, California. Figure 9 shows the spatial domain of the experiment. April 12 was a pilot deployment for the major May 9 field operational test. Figure 10 show the conditions as measured at a USGS field station in the region, demonstrating that the tidal conditions on the two days were similar (as is to be expected, because they were 28 days apart). The wind was 15 km/h SSE (gentle) on April 12 and 6 km/h S (light) on May 9. The water surface velocity was approximately 0.46 m/s (1.5 ft/s) in the outgoing (from northeast to southwest) direction. This is a noninverted tidal condition. The original plan was to deploy all the drifters from the Walnut Grove Public Dock (point A in Figure 9) (also see Figure 11), allow them to propagate through the junction, retrieve them at downstream points B and C (see Figure 12), and then redeploy them at points D and E while time permitted. Unfortunately, on May 9, there was a significant underwater construction operation happening at the junction (box F), requiring a midexperiment change of plans: drifters were initially released from point A and picked up around area F, then redeployed at points D and E, and then cycled from B–D and C–E. Table II shows the deploy and release schedule for the two experimental days.

The goal of the experiment was to determine the feasibility and utility of the heterogeneous fleet of drifters in an environment with obstacles and navigational challenges and to evaluate the back-end data assimilation techniques in the context of a real-time assimilation effort.

On April 12, we deployed 24 active drifters, and on May 9, we deployed 28 active drifters. On each day, 22
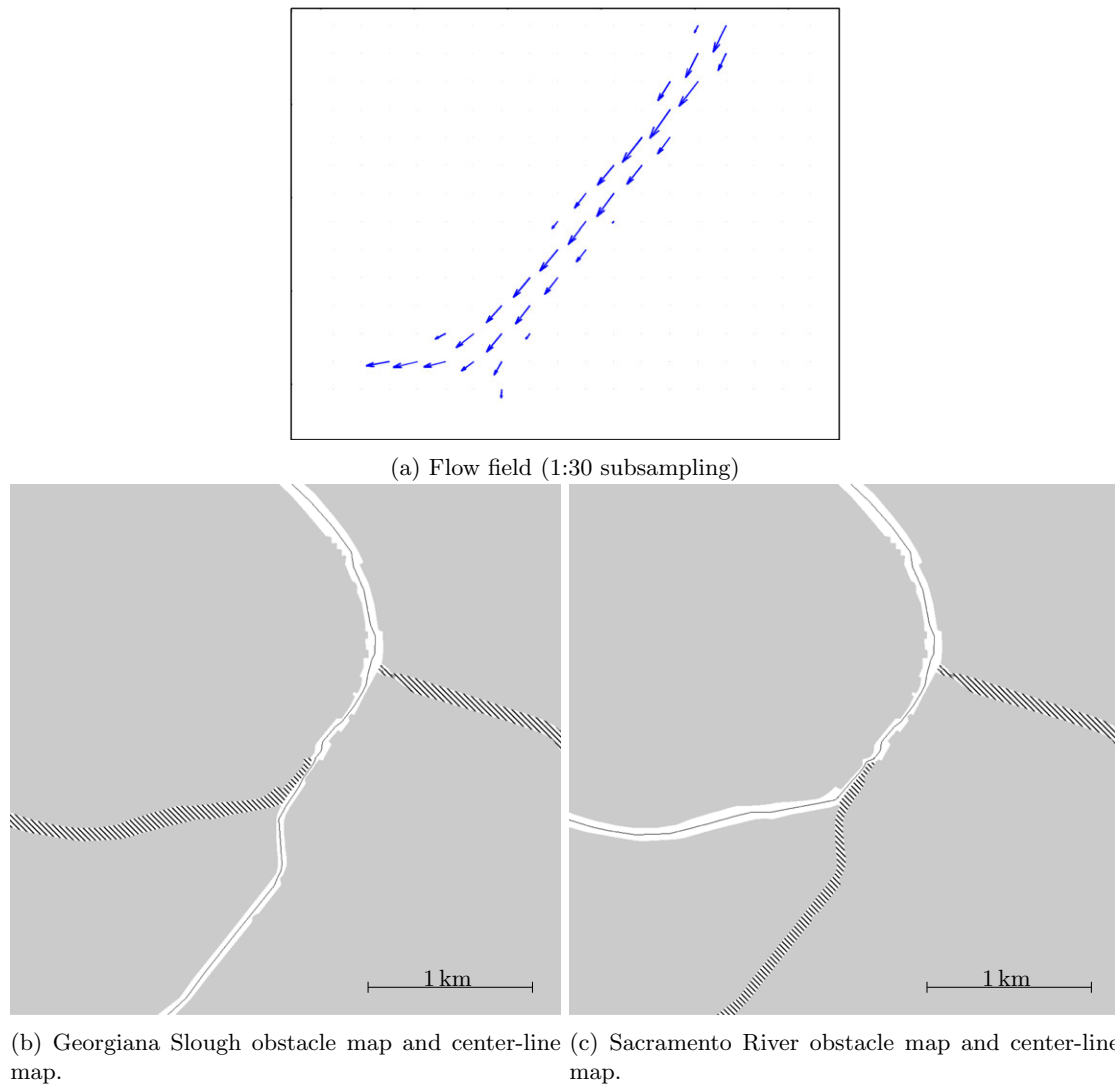
(a) Flow field (1:30 subsampling)

(b) Georgiana Slough obstacle map and center-line map.

(c) Sacramento River obstacle map and center-line map.

Figure 8. Input files to active control algorithm. Gray regions are actual geographical obstacles, whereas the hatched regions are virtual obstacles. Obstacle maps are composed of the gray and hatched regions and center-line maps are composed of the black center-line in each.

operated successfully. The remaining drifters (two on April 12, six on May 9) had internal electronics failures; when they failed to communicate, they were removed from the experiment. No manual intervention (aside from removing the failed drifters at the beginning) was required during the experiment; the operational drifters all successfully used actuation to avoid obstacles and select the appropriate channel of the Sacramento River/Georgiana Slough junction.

On April 12, we deployed 55 passive drifters, and on May 9, we deployed 68 passive drifters. All of the passive drifters worked successfully on April 12; one of the 68 passive drifters on May 9 failed to upload data. We did not filter out these failed passive drifters during the experiment; it

was deployed and retrieved like all the others, and the data-gathering failure was detected later. As described above, the passive drifter deployment was designed to eliminate the need for navigational moves and to reduce the risk of obstacles. None of the passive drifters needed to be removed from the experiment early.

6. PERFORMANCE EVALUATION

6.1. Fleet Control and Navigation

Figure 13 shows a detailed view of one active drifter moving through the Sacramento River downstream of the junction with the Georgiana Slough. Active propulsion segments
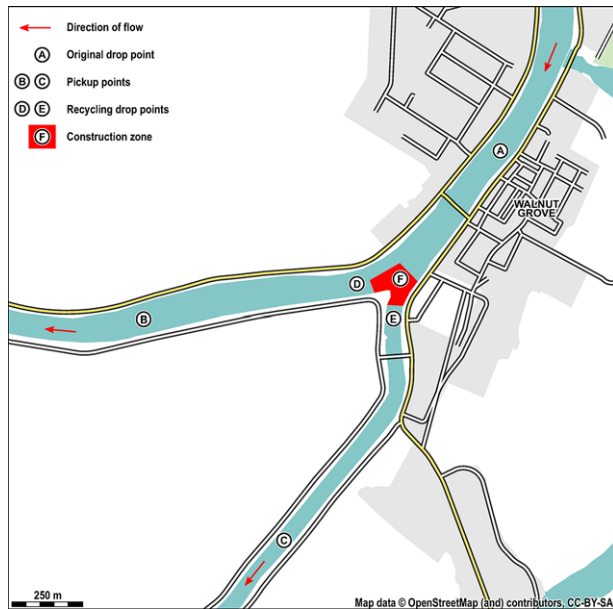
Figure 9. Map of experimental region near Walnut Grove, California. The Sacramento River flows from the upper right corner (north-east) toward the left side (west). The Georgiana Slough splits from the Sacramento River and flows to the bottom edge (south). Points A–F on the map indicate the deployment, pickup, and recycling locations for the experiments, as well as the construction zone that complicates the experiment at the junction.

Figure 11. Throwing an active drifter from the dock into the Sacramento River. Photo credit: Roy Kaltschmidt.

triggered by proximity to the shoreline can be seen in the red track segments. As the drifter floats downstream (from right to left on the map), it repeatedly approaches the south bank of the Sacramento River, triggering propulsion for obstacle avoidance. Note that, even though the drifter applied propulsion to return to the center of the river as quickly as

possible, the continued influence of the flow field caused it to move at a roughly 45° angle to the straight-line shortest distance. Figure 14 shows the same data as Figure 13, for the entire fleet, over each day of movement.

Figure 15 shows the fraction of the fleet involved in different kinds of moves over the course of the mission. Every propulsion action was classified as a "navigation" or "obstacle avoidance" move. Navigation moves were those actions
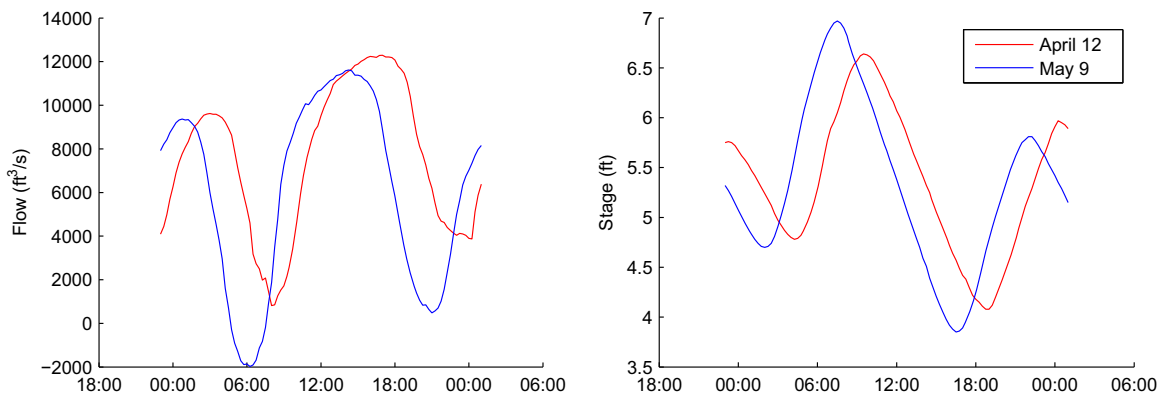

Figure 10. Water conditions at the GES field station on the experimental days. The flow (left) and stage (right) show the tidal effect on the movement and height of water in the experimental region. On May 9, flow reversed briefly around 6 a.m., but during the experimental periods, the water was always flowing in the normal direction.

Figure 12. Two boat crews retrieving Android drifters near the Sacramento River/Georgiana Slough junction. Photo credit: Roy Kaltschmidt.

taken by the drifter to determine which lane of the river to follow, whereas obstacle avoidance moves were actions taken by the drifter to avoid colliding with the shoreline. Notice that both May 9 and April 12 have similar patterns of use over time, as the fleet moves from the initial deployment location at the Walnut Grove dock, to the lane split region downstream of the bridge, then to the junction of the Sacramento River and Georgiana Slough. In aggregate, on the April 12 operation, 5.7% of fleet time was spent on obstacle moves, while 4.9% of fleet time was spent on navigation moves. During the May 9 operation, 4.4% of fleet time was spent on obstacle moves, while 6.3% of fleet time was spent on navigation moves.

In Figure 16, the position of the fleet within the river channel is quantified by taking the mean of the MTTR values for the drifters' positions over time. The active drifters and passive drifters are aggregated separately to illustrate the difference in their movement through the channel. As described in Section 4, two different sets of MTTR feedback maps were used; active drifters were assigned either to the Georgiana Slough or to the Sacramento River branch of the junction through the assignment of policy file. Finding the MTTR values for the passive drifters was done in postprocessing; of the two possible policy files, the "best case" file was applied at each time step to each passive drifter.

The solid lines in Figure 16 show the average minimum-time-to-obstacle of the two drifter ensembles. This is taken as a proxy for how close the drifters come to the shoreline and other obstacles such as marinas and docks. When the active drifters perform navigational moves to select a channel,

Table II. Timeline of the experimental procedures.

	April 12	May 9
0800	Arrive Walnut Grove; begin setup	Arrive Walnut Grove; begin setup
0830		
0900		
0930		
1000	Start releasing all drifters	
1030	Finish releasing all drifters	
1100	Drifting downstream	Start releasing all drifters
1130	All drifters retrieved	Finish releasing all drifters
1200	Second active deployment	All retrieved; re-release passives at D/E
1230	All drifters retrieved	Cycling continues
1300	Start cycling passives	Cycling continues
1330	Cycling continues	Cycling continues
1400	Cycling continues	Cycling continues
1430	Retrieve all drifters	Retrieve all drifters
1500	All boats depart	All boats depart
1530	Leave Walnut Grove	Leave Walnut Grove

they actually get closer to the shore than they would have if they had remained passive; this can be seen in the early stage of each experiment, where the time-to-shore value of the active fleet (the solid red line) drops more rapidly
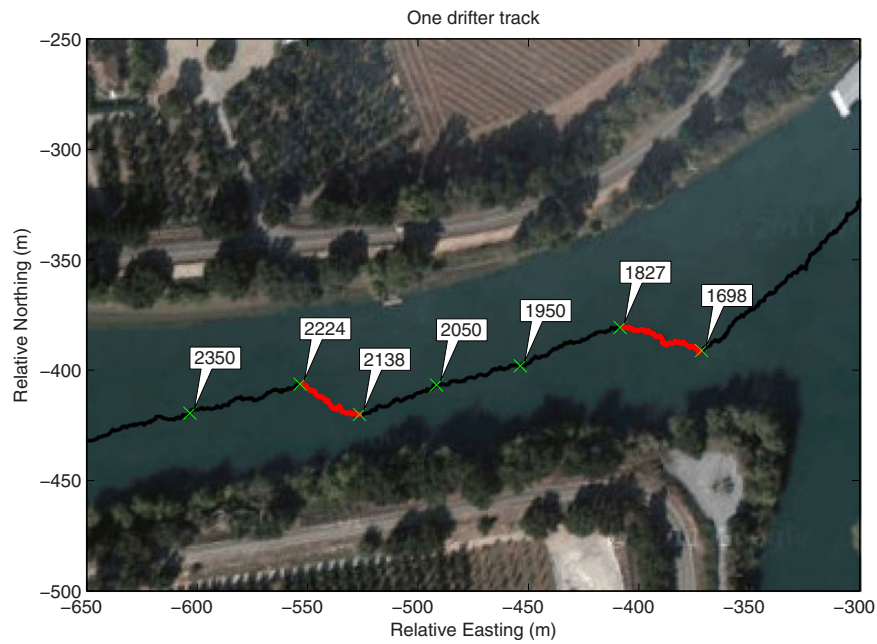
Figure 13. Detail view of one active drifter track through experimental region. Red track indicates active propulsion. Numbers indicate mission time in seconds.
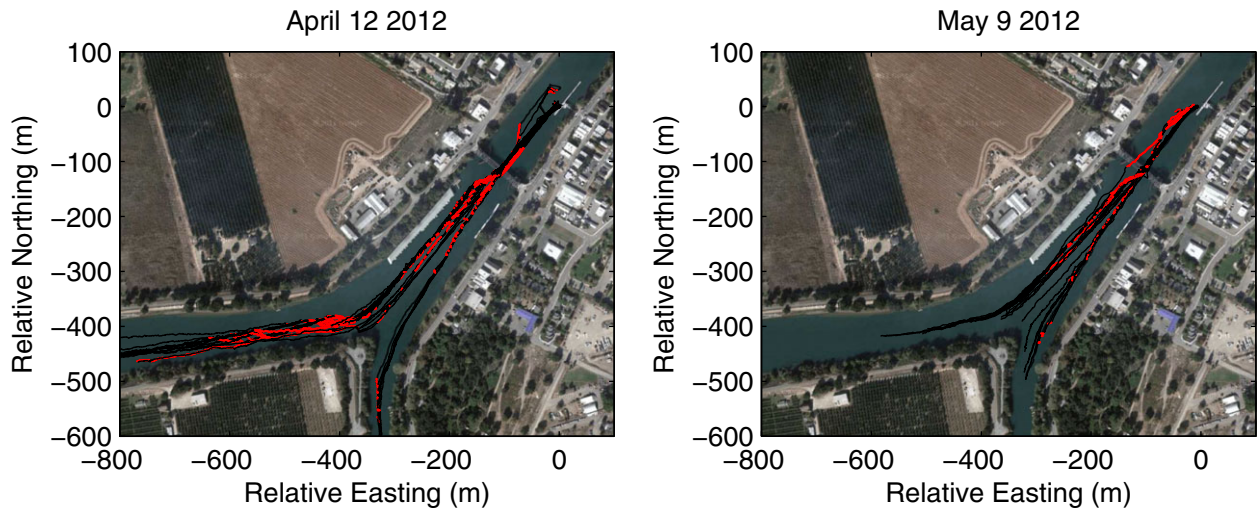

Figure 14. Overview of active drifter fleet movement in experimental region on both experimental days (22 active drifters on each day). Propelled movement is highlighted in red.

than the passive fleet (the solid blue line). Notice that the aggregate time-to-shore value for the passive fleet is below 300 s for a substantial portion of the experiment; this indicates that significant portions of the fleet were close enough to the shore that, if they had been active drifters, their propulsion would have been triggered for an obstacle-avoidance move. The time-to-shore value for the active fleet stays mostly above 300 s, for the same reason: Any active

drifter that comes closer to the shore uses its propulsion to move away.

The dotted lines in Figure 16 show the average minimum-time-to-center of the two drifter ensembles. Unlike the minimum-time-to-obstacle time series, there is not a consistent, substantive difference in the aggregate values for the two ensembles. This may be due to the fact that the minimum possible value for the MTTR function is zero;
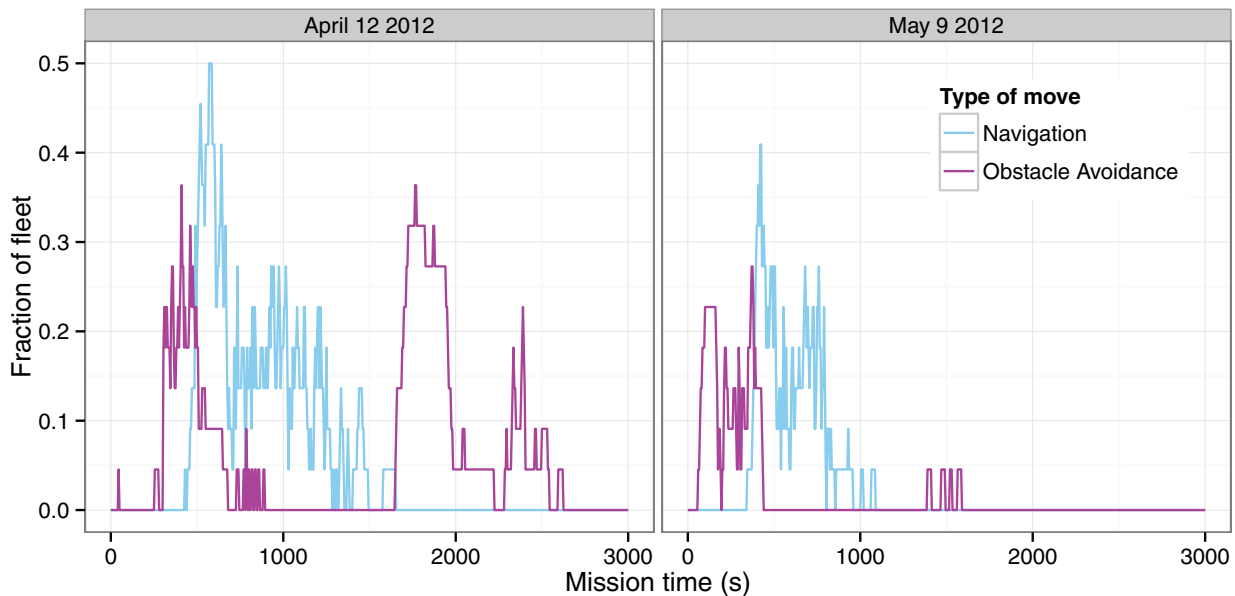
Figure 15. Fraction of fleet (22 active drifters) involved in different types of propulsion during the mission.
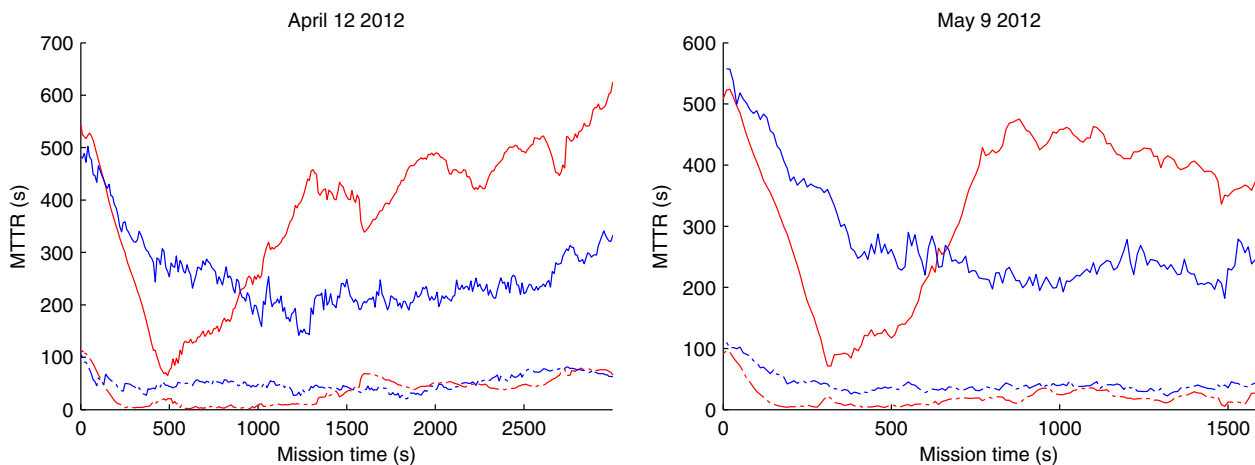

Figure 16. Mean MTTR values for active and passive component of fleet. Red lines are active drifters, and blue lines are passive drifters. Solid lines are shore MTTR values, and dotted lines are center MTTR values.

the low variance in the minimum-time-to-center time series may be caused by a "saturation at zero" effect.

Specific disadvantages or drawbacks to the system as implemented in this experiment were

1. The low reliability of the active drifter's GSM system drove us to install communication relays over the environment to ensure connectivity. This would not be a feasible solution for larger domains.
2. Despite the small size of the drifters, it would have been unwise to let them drift through the active underwater construction site at the junction of the Georgiana Slough

and the Sacramento River. We were less concerned with the possibility of causing harm to the construction operation; our primary concern was the difficulty and danger of retrieving a drifter were it to get caught in the construction equipment. We are very grateful to the DWR personnel, who accommodated our experiment with a brief shutdown of operations while we moved through their area.

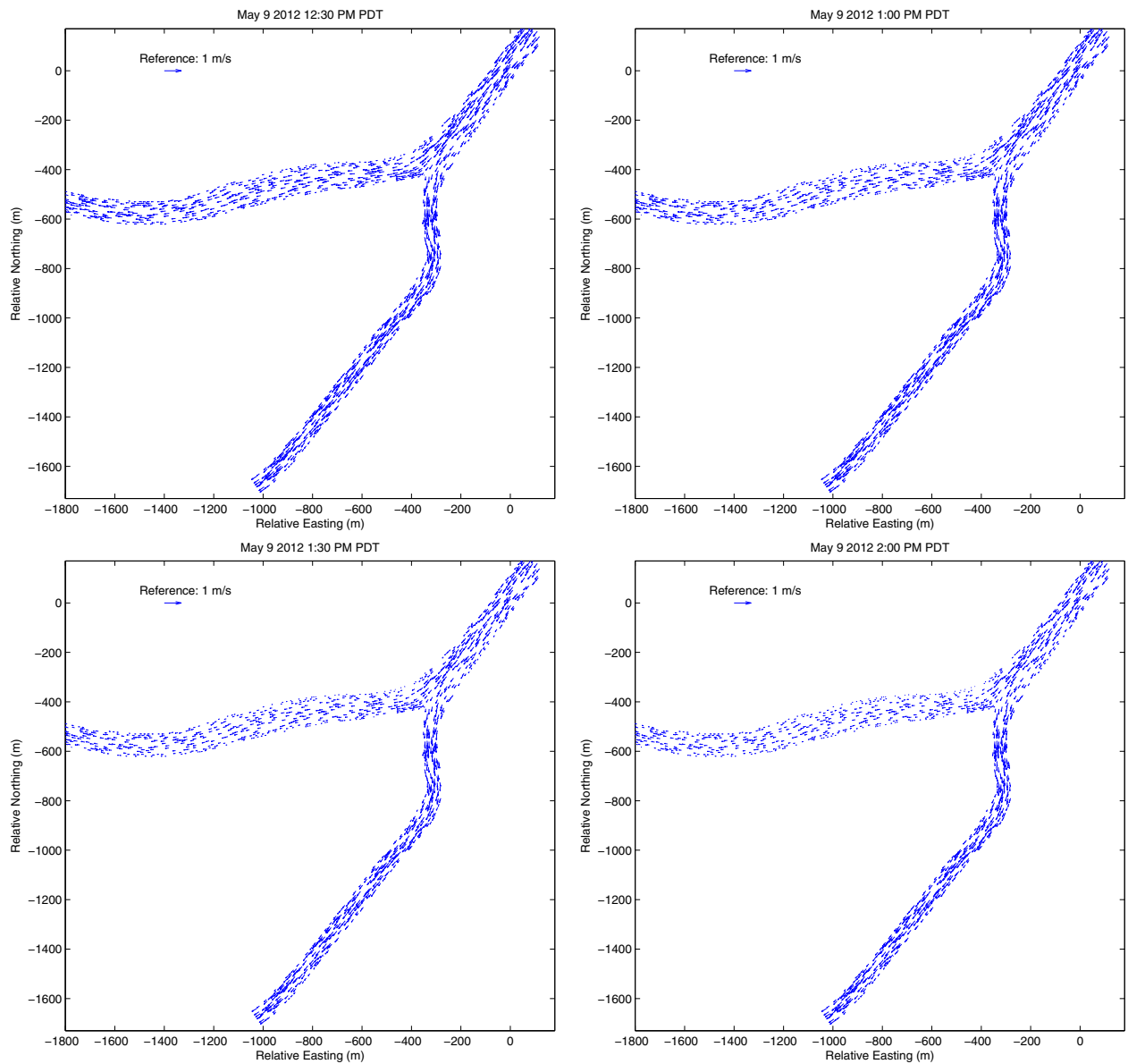In Section 7, we discuss how future work will address these shortcomings.

Figure 17. Quiver plots of assimilation results for flow fields during May 9 experiment. Each arrow represents the estimated water velocity at that location; each of the four facets represents a snapshot in time of the estimated flow.

6.2. Assimilation Results

Assimilation experiments were conducted for 2 hr, using the flow data collected during the May 9, 2012, experiment. The EnKF was implemented with a different number of ensemble members and inflation factors. (Inflation factors are methods for adaptively adjusting the EnKF model's standard deviation prescription in order to avoid model divergence.) We tested the system under different data configuration and model setups. It should be noted that there were no USGS Eulerian measurements available that can

be used as boundary conditions in the experimental domain. Therefore, we use the data reconciliation method developed in (Wu, Litrico, & Bayen, 2009) to get an estimation of discharges, and pose it to be the first guess of the data assimilation process in this section.

Figure 17 shows the flow field at each node from the model grid points. The grid size was selected as 20 by 20 m. It is determined for this specific experiment, based on our computational capacities. The arrow size represents the magnitude of the flow velocity, and the flow field is

demonstrated at 12:30, 1:00, 1:00, and 2:00 p.m. on May 9th, 2012, respectively. At the moment when drifter data were valid, the assimilated flow field was presented with drifter data incorporated; if the drifter data were unavailable, the model forward simulation was applied to ensure the continuity of the results.

Special attention has been paid to handling modeling errors that arise from the use of a simplified evolution model for the river flow and thus for the dynamics of the Lagrangian drifters. The modeling error in this study is modeled as a random variable with a Gaussian distribution. As the computation of the modeling error statistics is an intensive task, it should be performed prior to the actual estimation process.

The sensitivity of the assimilation system was studied by running the system with various inflation factors and ensemble sizes. The assimilation results are further compared with the forward model running from MIKE3. MIKE3 (Mike, 2008) is a commercial simulation package developed by the DHI Group. It features powerful three-dimensional (3D) modeling tools for free surface flows and associated sediment or water quality processes.

These sensitivity runs suggest that inflation plays an important role in the performance of the assimilation system. The inflation factor was first introduced in (Anderson & Anderson, 1999) to counteract the uncertainties that could not be specified in the filter covariance during the data assimilation process. The ensemble of forward model estimation is replaced by using the formula as below:

$$\xi_i = \rho(\xi_i - \bar{\xi}) + \bar{\xi}$$

The value of ρ can be determined either by trial-and-error or estimated adaptively as a parameter in the EnKF analysis process. In our experiment, we found that inflation indeed plays an important role in the data assimilation, once the model error is well estimated. Optimal inflation factor in our experiment was identified in the range of 1.1 to 1.2. The value varies, depending on different flow conditions. Increasing inflation from 1.1 to 1.2 improves the accuracy of the estimates, as inflation affects the spread of the ensemble and accounts for uncertainties that could not be specified in the filter covariance (Hoteit, Pham, & Blum, 2002). However, increasing the inflation factor to 1.3 caused some of the ensemble members to diverge during their forward integration with the model.

One of the most critical characteristics of EnKF is the finite size of the ensemble. The choice of the ensemble size must be large enough to account for the mean and spread of the prior distribution and allow an accurate representation of the covariance between the observation and the prior state. Generally, a larger ensemble size would lead to a smaller root mean square (RMS) error between model predictions and *in situ* observations, at the cost of increased computational load. There is, however, no rule of thumb of determining the ensemble size, as it can be affected by many different factors, such as model noise, measurement noise, and experiment settings. In our experiment, we found an ensemble size of 500 successfully provided a stable assimilation process and reasonable flow estimation. For simple cases, one might find 100-200 ensemble members adequate to provide good assimilation results.

7. CONCLUSIONS

The experiments conducted in Walnut Grove on April 12, 2012, and May 9, 2012, were the first ones to successfully use a fleet of 96 drifting Lagrangian sensors for estuarial hydrodynamic state estimation. A crucial component of this experiment was the ability of our robotic sensors to autonomusly correct their trajectory for navigational or obstacle avoidance purposes; without this autonomy, sensor fleets will not be able to scale above the number of human operators.

These experiments allowed us to demonstrate the following:

1. Small Lagrangian floating sensors are an effective way of gathering water flow information for an EnKF-driven assimilation process.
2. Actuated Lagrangian drifters can accomplish navigational tasks in unstructured environments and avoid natural obstacles such as shorelines; for the propulsion capabilities of the active drifter, a duty cycle of 10% is a reasonable first estimate for provisioning movement.
3. The trajectories of an actuated fleet will be significantly different than that of a passive fleet, in particular with respect to proximity to the shoreline and other obstacles. In supervised environments, the passive drifters are a good fit.

Future experiments will expand the utility of the FSN system by expanding the spatial and temporal domain of experiments and working toward real-time processing and assimilation of the incoming data. Solving the communication challenges and real-time assimilation objective is within our capability. Resolving the issue of interaction with other marine activities is an open problem; for the near future, we will have to mitigate this problem through careful experimental design.

The policy maps for actuation were computed offline, based on plausible estimates of the water currents. An improved system could generate these policy maps in real time, using the state estimates of the water currents produced by the assimilation process, and updating these maps on the drifters themselves via wireless communication as the experiment progressed. A future demonstration of this capability would be an important step toward the vision of a flexible fleet that could be deployed in unfamiliar environments without extensive preparation.

The assimilation results demonstrate the ability to produce a coherent, quantitatively sound estimate of the hydrodynamic state from Lagrangian data, in situations in which the experimental data are noisy and the model is imperfect. Our implementation of the EnKF system on the Carver cluster was made possible by the parallelizable nature of EnKF computations. However, several issues need to be addressed to ensure the stability of the data assimilation process, including uncertainties of the observations, differences between the forecast model and the real-world model, and the imperfect forward observation operator.

It is critical for a successful implementation of the EnKF to account for model error in an appropriate manner (Daley, 1992; Dee, 1995). Neglecting model error in the EnKF would lead to ensemble members spread too small (Houtekamer, Mitchell, & Deng, 2009; Isaksen, Fisher, & Berner, 2007). Extending the work in (Tossavainen, Percelay, Stacey, Kaipio, & Bayen, 2011b), will explore the approach for characterizing the model error in the river flow in channel networks on the basis of the Bayesian approximation error theory. The versatility of the theory enables us to model the error caused by numerous uncertainty sources simultaneously.

With a good estimation of model error, we can assume a suited physical model available in assimilation. In practice, however, an ensemble of vectors not globally orthogonalized has its tendency to collapse toward a small subspace. As a result, even for a perfect model, the background error covariance tends to be underestimated. These effects tend to underestimate the forecast error and, therefore, give too little weight to the observations, which can subsequently lead to the divergence of the filter (Anderson, 2003). To avoid this problem, multiplicative inflations and additive random perturbations (Constantinescu, Sandu, Chai, & Carmichael, 2007; Li, Kalnay, Miyoshi, & Danforth, 2009) may be appropriate tools to separate and enlarge the ensemble perturbations.

Further development of the Lagrangian sensor concept for estuarial studies would benefit from validation of the assimilation results through comparison with ground truth data. Hydrodynamic ground truth in field experiments is very difficult data to collect; indeed, this is the main motivation for the development of this sensor concept. Experiments for testing Lagrangian assimilation against ground truth might only be feasible in controlled environments.

Actuated mobile sensing will be a valuable tool for environmental studies in estuarial regions. Our work has established that a system of GPS-located, lightly actuated drifting sensors feeding an EnKF-based assimilation back end is a viable sensing technology. Miniaturization, cost reduction, and more complete recovery contingencies have the potential to make these sensing techniques a standard part of a hydrology researcher's toolbox.

REFERENCES

Anderson, J. L. (2003). A local least squares framework for ensemble filtering. Monthly Weather Review, 131(4), 634–642.

Anderson, J. L., & Anderson, S. L. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Monthly Weather Review, 127(12), 2741–2758.

Ateljevich, E., Colella, P., Graves, D. T., Ligocki, T. J., Percelay, J., Schwartz, P. O., & Shu, Q. (2009). CFD modeling in the San Francisco Bay and Delta. In Proceedings of the Fourth SIAM Conference on Mathematics for Industry, pages 99–107, San Francisco, California.

Austin, J., & Atkinson, S. (2004). The design and testing of small, low-cost GPS-tracked surface drifters. Estuaries, 27(6), 1026–1029.

Beard, J., Weekly, K., Oroza, C., Tinka, A., & Bayen, A. M. (2012). Mobile phone based drifting Lagrangian flow sensors. In IEEE 3rd International Conference on Networked Embedded Systems for Every Application (NESEA), pages 1–7. IEEE.

Bitterman, D. S., & Hansen, D. V. (1986). The design of a low cost tropical drifter buoy. In MDS '86: Marine Data Systems International Symposium, pages 575–581.

Castaings, W., Dartus, D., Honnorat, M., Le Dimet, F.-X., Loukili, Y., & Monnier, J. (2006). Automatic differentiation: A

tool for variational data assimilation and adjoint sensitivity analysis for flood modeling. In Automatic Differentiation: Applications, Theory, and Implementations, pages 249–262. Springer.

Clark, D. D. (1989). Overview of the Argos system. In OCEANS'89. Proceedings, volume 3, pages 934–939.

Constantinescu, E. M., Sandu, A., Chai, T., & Carmichael, G. R. (2007). Ensemble-based chemical data assimilation. I: General approach. Quarterly Journal of the Royal Meteorological Society, 133(626), 1229–1243.

Daley, R. (1992). Estimating model-error covariances for application to atmospheric data assimilation. Monthly Weather Review, 120(8), 1735–1746.

Davis, R. E. (1985). Drifter observations of coastal surface currents during CODE: The method and descriptive view. Journal of Geophysical Research, 90(C3), 4741–4755.

Dee, D. P. (1995). On-line estimation of error covariance parameters for atmospheric data assimilation. Monthly Weather Review, 123(4), 1128–1145.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research, 99(C5), 10143–10162.

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean Dynamics, 53, 343–367.

Evensen, G. (2009). Data assimilation: The ensemble Kalman filter. Springer.

Hamill, T. M., & Snyder, C. (2000). A hybrid ensemble Kalman filter–3d variational analysis scheme. Monthly Weather Review, 128(8).

Hamill, T. M., Whitaker, J. S., & Snyder, C. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. Monthly Weather Review, 129(11), 2776–2790.

Honnorat, M., Monnier, J., & Le Dimet, F.-X. (2009). Lagrangian data assimilation for river hydraulics simulations. Computing and Visualization in Science, 12(5), 235–246.

Hoteit, I., Pham, D.-T., & Blum, J. (2002). A simplified reduced order Kalman filtering and application to altimetric data assimilation in Tropical pacific acific. Journal of Marine Systems, 36(1), 101–127.

Houtekamer, P., Mitchell, H. L., & Deng, X. (2009). Model error representation in an operational ensemble Kalman filter. Monthly Weather Review, 137(7), 2126–2143.

Houtekamer, P. L., & Mitchell, H. L. (1998). Data assimilation using an ensemble Kalman filter technique. Monthly Weather Review, 126(3), 796–811.

Houtekamer, P. L., & Mitchell, H. L. (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. Monthly Weather Review, 129(1), 123–137.

Huttunen, J. M., & Kaipio, J. P. (2007a). Approximation error analysis in nonlinear state estimation with an application to state-space identification. Inverse Problems, 23(5), 2141.

Huttunen, J. M., & Kaipio, J. P. (2007b). Approximation errors in nonstationary inverse problems. Inverse Problems and Imaging, 1(1), 77.

Isaksen, L., Fisher, M., & Berner, J. (2007). Use of analysis ensembles in estimating flow-dependent background error variance. In Proceedings of the ECMWF Workshop on flow-dependent aspects of data assimilation, pages 11–13.

Kaipio, J. P., & Somersalo, E. (2005). Statistical and computational inverse problems. Springer.

Kalnay, E. (2003). Atmospheric modeling, data assimilation, and predictability. Cambridge University Press.

Keppenne, C. L., & Rienecker, M. M. (2002). Initial testing of a massively parallel ensemble kalman filter with the Poseidon isopycnal ocean general circulation model. Monthly Weather Review, 130(12).

Lehikoinen, A., Huttunen, J. M. J., Finsterle, S., Kowalsky, M. B., & Kaipio, J. P. (2010). Dynamic inversion for hydrological process monitoring with electrical resistance tomography under model uncertainties. Water Resources Research, 46(4).

Li, H., Kalnay, E., Miyoshi, T., & Danforth, C. M. (2009). Accounting for model errors in ensemble data assimilation. Monthly Weather Review, 137(10), 3407–3419.

Ligocki, T. J., Schwartz, P. O., Percelay, J., & Colella, P. (2008). Embedded boundary grid generation using the divergence theorem, implicit functions, and constructive solid geometry. In Journal of Physics: Conference Series, 125, 12080.

Mandel, J., Cobb, L., & Beezley, J. D. (2011). On the convergence of the ensemble Kalman filter. Applications of Mathematics, 56(6), 533–541.

Mike, D. (2008). Mike 3 flow model fm. Hydrodynamic and Transport Module, Scientific documentation.

Mitchell, I., Bayen, A., & Tomlin, C. (2005). A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games. IEEE Transactions on Automatic Control, 50(7), 947–957.

NERSC (2012). NERSC's Carver iDataPlex Cluster. Online: http://www.nersc.gov/systems/carver-ibm-idataplex/.

Niiler, P. P., Davis, R. E., & White, H. J. (1987). Water-following characteristics of a mixed layer drifter. Deep Sea Research Part A. Oceanographic Research Papers, 34(11), 1867–1881.

Nissinen, A., Kolehmainen, V. P., & Kaipio, J. P. (2011). Compensation of modelling errors due to unknown domain boundary in electrical impedance tomography. Medical Imaging, IEEE Transactions on, 30(2), 231–242.

Nodet, M. (2006). Variational assimilation of Lagrangian data in oceanography. Inverse Problems, 22(1), 245.

Oroza, C., Tinka, A., Wright, P. K., & Bayen, A. M. (2013). Design of a network of robotic Lagrangian sensors for shallow water environments with case studies for multiple applications. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 227(11), 2531–2548.

Perez, J. C., Bonner, J., Kelly, F. J., & Fuller, C. (2003). Development of a cheap, GPS-based, radio-tracked, surface drifter for closed shallow-water bays. In Proceedings of the IEEE/OES Seventh Working Conference on Current Measurement Technology, pages 66–69.

Romanowicz, R. J., Young, P. C., & Beven, K. J. (2006). Data assimilation and adaptive forecasting of water levels in

the river Severn catchment, United Kingdom. Water Resources Research, 42(6).

Swallow, J. C. (1955). A neutral-buoyancy float for measuring deep currents. Deep Sea Research (1953), 3(1), 74–81.

Tinka, A., Rafiee, M., & Bayen, A. M. (2013). Floating sensor networks for river studies. IEEE Systems Journal, 7(1), 36–49.

Tossavainen, O.-P., Percelay, J., Stacey, M., Kaipio, J. P., & Bayen, A. (2011a). State estimation and modeling error approach for 2d shallow water equations and Lagrangian measurements. Water Resources Research, 47(10).

Tossavainen, O.-P., Percelay, J., Stacey, M., Kaipio, J. P., & Bayen, A. (2011b). State estimation and modeling error approach for 2D shallow water equations and Lagrangian measurements. Water Resources Research, 47(10).

Van Leeuwen, P. J., & Evensen, G. (1996). Data assimilation and inverse methods in terms of a probabilistic formulation. Monthly Weather Review, 124(12), 2898–2913.

Vrugt, J. A., & Robinson, B. A. (2007). Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging. Water Resources Research, 43(1).

Wang, W.-C., Chau, K.-W., Cheng, C.-T., & Qiu, L. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. Journal of Hydrology, 374(3), 294–306.

Weekly, K., Tinka, A., Anderson, L., & Bayen, A. M. (2014). Autonomous river navigation using the Hamilton-Jacobi framework for underactuated vehicles. IEEE Transactions on Robotics, 30(5), 1250–1255.

Wu, Q., & Bayen, A. M. (2015). Variational Lagrangian data assimilation in open channel networks. Water Resources Research, 51(4), 1916–1938.

Wu, Q., Litrico, X., & Bayen, A. M. (2009). Data reconciliation of an open channel flow network using modal decompositions. Advances in Water Resources, 32(2), 193–204.