Article

The International Journal of Robotics Research
1–18
© The Author(s) 2015
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0278364915587723
ijr.sagepub.com
SAGE

# Data-driven robotic sampling for marine ecosystem monitoring

**Jnaneshwar Das[1], Frédéric Py[2], Julio B.J. Harvey[2], John P. Ryan[2], Alyssa Gellene[3], Rishi Graham[2], David A. Caron[3], Kanna Rajan[2] and Gaurav S. Sukhatme[2]**

**Abstract**

*Robotic sampling is attractive in many field robotics applications that require persistent collection of physical samples for ex-situ analysis. Examples abound in the earth sciences in studies involving the collection of rock, soil, and water samples for laboratory analysis. In our test domain, marine ecosystem monitoring, detailed understanding of plankton ecology requires laboratory analysis of water samples, but predictions using physical and chemical properties measured in real-time by sensors aboard an autonomous underwater vehicle (AUV) can guide sample collection decisions. In this paper, we present a data-driven and opportunistic sampling strategy to minimize cumulative regret for batches of plankton samples acquired by an AUV over multiple surveys. Samples are labeled at the end of each survey, and used to update a probabilistic model that guides sampling during subsequent surveys. During a survey, the AUV makes irrevocable sample collection decisions online for a sequential stream of candidates, with no knowledge of the quality of future samples. In addition to extensive simulations using historical field data, we present results from a one-day field trial where beginning with a prior model learned from data collected and labeled in an earlier campaign, the AUV collected water samples with a high abundance of a pre-specified planktonic target. This is the first time such a field experiment has been carried out in its entirety in a data-driven fashion, in effect "closing the loop" on a significant and relevant ecosystem monitoring problem while allowing domain experts (marine ecologists) to specify the mission at a relatively high level.*

**Keywords**

Robotic sampling, marine robotics, field robots, learning and adaptive systems, machine learning

## 1. Introduction

Many studies require persistent collection of physical samples for ex-situ analysis. For example, in the earth and environmental sciences, air, water, and soil samples are routinely collected to study properties that can only be observed ex-situ through laboratory analysis. Robotic sampling is attractive in these domains, enabling collection of physical samples at spatial and temporal scales previously infeasible. In marine ecosystem monitoring, our test domain, detailed understanding of plankton diversity and abundance requires complex morphological and molecular analysis in a laboratory setting. To aid such studies, an autonomous underwater vehicle (AUV) has been equipped with a water sample collection system capable of retrieving 10, 1.8 L, samples for laboratory analysis (Figure 1). Additionally, onboard sensors measure physical and chemical properties of water in-situ, in real-time. Using this AUV as a testbed, we present an opportunistic, data-driven, and iterative robotic sampling strategy that uses statistical machine learning to maximize the total abundance of plankton in the water samples collected. Specifically, we maximize cumulative plankton abundance in water samples collected during a campaign consisting of multiple AUV surveys. Starting with no prior information on the factors that drive the population abundance of the specific plankton being studied, our approach creates a probabilistic

[1]Department of Computer Science, University of Southern California, USA
[2]Monterey Bay Aquarium Research Institute, Moss Landing, USA
[3]Department of Biological Sciences, University of Southern California, USA

**Corresponding author:**
Jnaneshwar Das, Department of Computer Science, University of Southern California, Ronald Tutor Hall, 3710 S McClintock Avenue, Los Angeles, California 90089, USA.
Email: jnaneshd@usc.edu

Downloaded from ijr.sagepub.com by guest on August 7, 2015

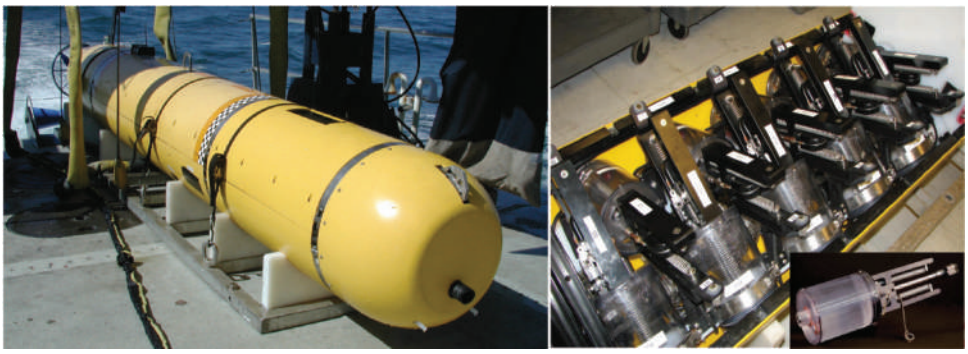


**Fig. 1.** The Dorado AUV with an onboard water sample collection system consisting of 10 1.8 L "gulpers" that can be triggered by the onboard computer. Real-time measurements by the AUVs sensor suite can guide physical sample collection decisions.

model for plankton abundance with environmental covariates sensed in-situ as its inputs. During a survey, the AUV collects water samples using predictions from this model, and the plankton abundance is labeled through laboratory analysis after the completion of the survey. The freshly labeled data is used to update the model, resulting in improved sampling accuracy during subsequent surveys. Extensive retrospective simulation studies (carried out by mining historical field data) consistently demonstrate the collection of samples with progressively higher abundance over successive surveys. We note that the previous state of the art in underwater vehicle deployment was to task the vehicle using a sequence of waypoints. Our work allows a domain expert to task the vehicle at a much higher level, e.g. "gather samples that are high abundance in a particular type of plankton", without needing to specify where to sample.

When labeled data from previous campaigns is available, our approach can be used to maximize the total plankton abundance from a single survey. During a six hour field trial in Monterey Bay in October 2013, the AUV collected water samples with an abundant amount of Pseudo-nitzschia (PN), a toxigenic genus of phytoplankton known to cause harmful algal blooms. To do so, the AUV used an onboard predictive model trained on 87 water samples collected during a campaign in 2010. The samples collected during the 2013 trial were used by marine biologists to study the physiology of PN.

## *1.1. Related work and contributions*

A body of literature exists for robotic monitoring of terrestrial and marine environments (Anthony et al., 2014; Batalin et al., 2004; Bryson et al., 2010; Hitz et al., 2012; Julian et al., 2012; Shkurti et al., 2012; Stealey et al., 2008; Tokekar et al., 2010); however, the focus has been on properties that are directly measurable by onboard sensors, i.e. observable in-situ. In this setting, approaches for computing probabilistic models for environmental processes has been explored (Garg et al., 2012; Kim et al., 2013; Singh et al., 2010). Strategies for sensor placement have been developed to generate optimal maps of properties such as temperature and pH (Krause and Guestrin, 2007; Krause et al., 2006). Extended to mobile sensors, informative-paths have been computed for robots to maximize information gain from in-situ measurements (Binney et al., 2013; Low et al., 2009; Singh et al., 2009; Zhang and Sukhatme, 2007). For terrestrial ecological monitoring, multi-modal and multi-scale data from unmanned aerial vehicles (UAVs) has been fused using Bayesian techniques to obtain maps of terrestrial vegetation (Reid et al., 2013). Efficient planning for fertilization has been demonstrated using in-situ measurements of nitrogen levels using UAVs and ground robots (Tokekar et al., 2013).

The problem of samples whose desired properties of interest cannot be analyzed in-situ but require ex-situ analysis (e.g. in a laboratory) has been investigated to a lesser extent. In the marine domain, the Dorado AUV (Figure 1) has been used to collect water samples abundant in the desired properties of interest, guided by onboard measurements of environmental covariates (Fox et al., 2007; Garcia-Olaya et al., 2012; Harvey et al., 2012a,b; Zhang et al., 2010); however, the key design parameters such as spacing between samples and thresholds have been specified by scientists. This inhibits scaling of such approaches to robot teams, as well as experiments over weeks or months that may demand frequent parameter tuning. Most importantly, many open problems in ecosystem monitoring are plagued with an insufficient amount of the prior information needed to specify such parameters precisely, resulting in sub-optimal performance. This poses a significant bottleneck to the role of robots in large-scale earth exploration. Our work addresses this issue by using statistical machine learning to train robots with past experiments, to guide future sampling decisions. The key to this is a probabilistic model that predicts organism abundance from environmental covariates measurable in-situ by the robot.

Examples can be found in the ecological modeling literature where hidden phenomena such as harmful algal blooms have been predicted from environmental parameters such as temperature, salinity, upwelling, and rainfall indices (Anderson et al., 2010). In terrestrial environmental

Downloaded from ijr.sagepub.com by guest on August 7, 2015

monitoring literature, the maximum entropy method has been used to predict geographic distribution of terrestrial species from environmental covariates of presence only samples (Elith et al., 2011); however, acquiring data for building such models is challenging. To this end, the Dorado AUV has been used for oceanographic studies of intermediate nepheloid layers (INLs) for prediction and sampling using a classifier trained on scientist labeled datasets. Rules were specified by scientists for water sample spacing, along with appropriate thresholds on predicted signals allowing the targeting of high probability INL samples for analysis in the lab (Fox et al., 2007; Garcia-Olaya et al., 2012). The Dorado AUV has also been used to acquire water samples precisely at chlorophyll peaks and thermal fronts using the properties of vertical yo-yo profiles (yo-yo profiles are vertical sawtooth patterns carried out by AUVs during a survey between specified depth envelopes) (Ryan et al., 2010; Zhang et al., 2010, 2012). Inspired by these efforts, we take a unified look at modeling and sampling with the goal of maximizing a desired property of interest, in this case the plankton abundance, in the acquired physical samples. We exploit the Bayesian optimization literature where sequential and batch-mode sampling has been studied for both active learning (i.e. improving model accuracy) and bandit-setting (maximizing the reward, or minimizing the cumulative regret) (Chen and Krause, 2013; Srinivas et al., 2009). We use the Gaussian process upper confidence bound (GP-UCB) algorithm, a sequential sampling policy with stochastic guarantees regarding cumulative regret after multiple trials (Srinivas et al., 2012). During each trial, the GP-UCB algorithm defines a utility function that computes the utility of each candidate sample in the input space, given the probabilistic estimates of the predicted sample value, i.e. the mean and the variance. Acquiring the sample with maximum GP-UCB utility in each trial results in progressively higher valued samples, hence reducing the cumulative regret over multiple trials. To address the selection of multiple samples during each trial, various approaches for Bayesian batch optimization have been developed. In this setting, instead of one sample being acquired and labeled in each trial, a batch of samples are collected and labeled. This results in improved time efficiency, i.e. comparable performance in a fewer number of trials. A hybrid Bayesian batch optimization algorithm uses the expected improvement on a Gaussian process (GP) model to switch between sequential sample acquisition, and acquisition of batches of samples (Azimi et al., 2012). The GP-UCB algorithm has been extended to the batch setting where given full knowledge of the input space, multiple samples are acquired during each trial for parallel evaluation (Desautels et al., 2014). In contrast to these approaches, our problem demands an *opportunistic* batch-mode Bayesian optimization strategy where no prior spatial map of target organism distributions is available; this is ideal in marine environments where organisms are constantly advected by ocean currents (Ryan et al., 2014a,b). Hence, the AUV needs to make decisions online as it carries out its survey. We provide an extension of the GP-UCB algorithm where a single sample is drawn in each trial, to fit our problem where $k$ samples are acquired during a survey, and labeled upon survey completion.

The online nature of our sample collection problem is addressed using optimal stopping theory, which concerns the right time to carry out a particular action (e.g. acquire a physical water sample). Specifically, we seek a strategy to select a batch of $k$ water samples online in order to maximize the total utility of the sample set. Here the utility is generated online by the GP-UCB algorithm operating on probabilistic predictions from the organism abundance model as the AUV makes in-situ measurements of environmental covariates. Two algorithms are applicable to our problem, the multi-choice hiring algorithm (Girdhar and Dudek, 2009) and the submodular secretary algorithm (Bateni et al., 2010). Both are derived from the secretary (or hiring) problem that concerns an irrevocable hiring decision from a stream of $N$ rankable candidates arriving independent and identically distributed (IID) (Ferguson, 1989). The optimal solution to this problem has a probability of $1/e$ of selecting the top candidate. Extending the secretary problem to a case where multiple candidates need to be hired, Girdhar and Dudek (2009) present the multi-choice secretary algorithm with a fixed threshold to choose the $k$ top-ranked secretaries with a probability of $1/ke$. Where candidates can be rated instead of ranked, Bateni et al. (2010) present the *submodular secretary algorithm* that seeks to maximize a set function that defines efficiency of the selected secretarial group based on their overlapping skills. Of the two choices, we use the submodular secretary algorithm in our work for a couple of reasons. First, as opposed to the multi-choice hiring algorithm where the candidates are ranked and a probabilistic guarantee is presented for choosing the $k$ top-ranked candidates, the submodular secretary algorithm provides a mechanism to select the set of candidates with the highest cumulative rating. This is achieved by using a sum set function that reflects the total utility of the selected set of physical samples. Second, we found the submodular secretary algorithm to be robust regarding correlations among the samples observed by the AUV as it periodically passed through a thin layer of chlorophyll (phytoplankton). The multi-choice hiring algorithm, on the other hand, uses a fixed threshold to select $k$ candidates which results in the acquisition of spatially neighboring samples at the cost of future potentially high-value samples. We discuss this in more detail in Section 2.

Our work presents the first opportunistic, data-driven, and iterative robotic sampling strategy for physical sample collection, exploiting algorithms from probabilistic modeling, Bayesian optimization, and optimal stopping theory. The methodology is evaluated using different sampling strategies within a novel framework that mines previously collected AUV data to emulate campaigns with different initial conditions. Our approach based on cumulative regret minimization outperforms other approaches both for the offline scenario where we have full knowledge of a survey

Downloaded from ijr.sagepub.com by guest on August 7, 2015

in advance, and the online scenario where sampling decisions are made sequentially with no knowledge of the future. Additionally, water samples acquired during the one-day field trial showed a high abundance of PN, facilitating studies of its physiology.

Through the results, we demonstrate that our data-driven and iterative sampling strategy "closes the loop" on model learning and knowledge extraction for ecosystem monitoring. Input from domain experts is focused on tasking the vehicle at a high level, and labeling the physical samples collected. Although applied in the context of marine ecosystem monitoring, our approach is general and can be scaled to other domains where opportunistic collection of physical samples is necessary. By freeing scientists from manual specification of parameters that can instead be learned from past data, our approach has the potential to expand the role of robotic explorers beyond observational environmental monitoring, to being actors, persistently retrieving high-quality physical samples from hazardous and extreme environments inaccessible to humans.

### 1.2. Outline

The structure of the paper is as follows. In Section 2, we present the problem statement and describe the technical approach. Simulation studies to compare various strategies are presented in Section 3, followed by a discussion of the results from a field trial to target a harmful algae in Section 4. We conclude with a discussion of future directions in Section 5.

## 2. Technical approach

The physical sample collection problem for marine ecosystem monitoring is as follows. During a field campaign consisting of multiple AUV surveys, water samples containing a desired organism in high abundance must be collected. The AUV is equipped with water samplers, each of which can be triggered only once per survey. This constraint eliminates cross-contamination of collected water samples. The path for a survey is predetermined. For example, a scientist plans a lawnmower (or radiator) pattern over a region and the AUV is tasked with acquiring water samples along the survey track. Alternatively, surveys can be dynamic, carrying out a predetermined pattern relative to the frame of reference of an advecting patch of water tagged by a GPS-tracked drifter (Das et al., 2012). The goal is to develop a principled strategy for collecting water samples during such surveys, several of which typically comprise a campaign.

During a survey, the AUV makes irrevocable sample collection decisions *online* on a sequential stream of predicted organism abundance values, with no knowledge of the valuation of future samples. At the end of each survey, the batch of $k$ collected water samples are analyzed ex-situ and the labeled data is assimilated back into the training dataset to update a probabilistic predictive model. At the beginning of a study, when no data is available to guide sample collection decisions, random physical sampling is carried out followed by labeling to learn an initial model. We posit that as new labeled data is assimilated into the existing predictive model at each iteration (i.e. survey), the sample quality should improve in subsequent surveys. The problem can be posed as a minimization of cumulative regret for the physical sample values (in our problem the organism abundance) from all the surveys of a campaign. Cumulative regret is the total regret of all samples over multiple surveys of a campaign.

Our approach is opportunistic by design, where sampling decisions are carried out on an arbitrary survey without prior knowledge of the spatial distribution of the property of interest. A logical extension of this work is to compute informative paths instead of using predetermined surveys; however, we argue that the online opportunistic approach has a practical benefit in the marine domain. In the coastal ocean, the geographic distribution of properties of interest change rapidly due to environmental forcing such as water currents (Das et al., 2010). Hence, the computation of informative paths is subject to the availability of recent[1] data from a pilot survey, a scenario often infeasible due to operational constraints. Dynamics of the spatial distribution of organisms also affects the organism abundance model. Our analysis shows that spatial parameters, i.e. latitude, longitude, and depth, are not necessary as inputs to the predictive model of organism abundance, as long as relevant environmental covariates are observed and used as inputs. For example, temperature captures depth trends when predicting phytoplankton abundance. Additionally, the training dataset usually consists of samples analyzed across seasons, spanning multiple weeks, months, or years. This is acceptable for marine ecological studies since organisms are known to show similar trends during similar seasons. By ignoring geographic parameters, our modeling approach is able to exploit a small training dataset collected over a large spatial and temporal range to learn the "environmental niche" of the target organism.

### 2.1. Problem formulation

An AUV equipped with $k$ physical samplers is tasked to obtain water samples with high values of a property $b \in \mathbb{R}^+$ that can only be measured offline (i.e. ex-situ). The AUV measures environmental feature vectors $\{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$, $\mathbf{z} \in \mathbb{R}^D$ at corresponding spatial locations $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}, \mathbf{x} \in \mathbb{R}^3$. The set of locations $\mathbf{X} = \{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$ constitutes a geographic survey. $T$ such surveys are carried out, collectively forming a campaign, and $k$ water samples can be collected during each survey. Three practical considerations constrain this problem. Firstly, for every survey, $N >> k$. For the AUV used in this work, $k = 10$, and $N$ is upwards of 45,000 subject to survey duration. Hence, the sample collection algorithm has to be highly selective. Secondly, labeling of the acquired samples requires ex-situ laboratory analysis to measure the value of $b$ in each of the $k$ samples, and can only be performed offline when the

Downloaded from ijr.sagepub.com by guest on August 7, 2015

AUV is recovered at survey termination. Hence, measurements of $b$ are not available *during* the survey, and any computation that depends on $b$ can only be carried out at the completion of a survey after sample analysis. Thirdly, due to the opportunistic nature of our approach no pilot survey is carried out, and hence prior information on the spatial distribution of environmental feature vector $\mathbf{z}$ is not available. Related to this constraint, reiterative sampling of spatially and temporally variable physical features (e.g. phytoplankton thin-layers) in the marine environment is notoriously difficult in water masses that are constantly advecting due to ocean currents. To a smaller extent, precisely revisiting a physical feature is also difficult due to state-estimation errors state-estimation errors – absolute positioning through. Absolute positioning through GPS is not available underwater; hence state-estimation is carried out through dead-reckoning, with corrections through periodic surfacing, e.g. once every 30 minutes. The lack of prior information demands that the batch of $k$ water samples be collected *online*.

Under the above constraints, the autonomous water sample collection problem for a campaign of $T$ surveys is as follows. Let $B = \{b_1,\ldots,b_k\}$ denote a set of $k$ physical water samples that can be acquired by the AUV during each survey. We are given an initial set of water samples acquired randomly from $p$ pilot surveys, $B_{\text{pilot}} = \{B_1,\ldots,B_p\}$. The goal is to find a policy that adaptively acquires the sample set $B_{\text{adaptive}} = \{B_{p+1},\ldots,B_T\}$ from subsequent surveys using a model $g$ trained from pilot survey data, and updated after completion of each survey. Each sample in the set $B_j$ from the jth survey (where $p < j \leq T$) is acquired online using prediction $b = g(\mathbf{z})$, with no knowledge of the predicted value of future samples. The sample set $B_j$ is labeled after the termination of the $j$th survey when the AUV is retrieved and collected water samples analyzed in the lab. The goal is to minimize the cumulative regret of all the water samples collected during a campaign given by $B_{\text{campaign}} = \{B_{\text{pilot}}, B_{\text{adaptive}}\}$. Here, regret of a physical sample is the difference between the value of the best sample (i.e. the highest target organism abundance) that could have been acquired from the survey by an optimal strategy, and the value of the sample acquired by an online algorithm. For the set $B_j$, we compute the average regret of $k$ samples, and the goal is then to minimize the cumulative *average* regret of samples collected during the campaign.

Using the collected and laboratory-analyzed samples from each survey, we iteratively learn the predictive model $g : \mathbb{R}^D \mapsto \mathbb{R}$ that maps the in-situ measured environmental feature vector $\mathbf{z}$ to the hidden property of interest $b$. Specifically, we use GP regression (Rasmussen, 2006), a Bayesian function approximation technique to learn a *probabilistic* model. GP regression output is the predictive distribution for the hidden property of interest $b$, for a given deterministic candidate environmental feature vector $\mathbf{z}$. The probabilistic predictions of the GP regression model are converted to a real-valued sample *utility* in a Bayesian optimization setting to maximize long-term reward, i.e. minimize cumulative regret over multiple trials (or surveys of a campaign). Finally, since sample collection has to be performed online on a stream of candidate sampling locations with no knowledge of future locations, we employ optimal stopping theory to determine a theoretically sound best-choice algorithm to make this decision. Specifically, we use the submodular secretary algorithm to choose the $k$ highest utility (best water sampling locations) candidates from each survey *online*. These three subproblems are described in the following sections.

## 2.2. Probabilistic model

The goal of the probabilistic model is to learn the mapping from $\mathbf{z} = [\text{temperature, salinity}, \ldots] \in \mathbb{R}^D$, measured at spatial locations $\mathbf{x} = [\text{latitude, longitude, depth}] \in \mathbb{R}^3$, to a desired parameter of interest, in our case plankton abundance $b \in \mathbb{R}^+$. We use GP regression to learn this mapping, with the assumption that the joint distribution of the observed plankton abundance is Gaussian. In this setting, the underlying process can be defined completely by a mean function (assumed to be zero without loss of generality), and a covariance or kernel function that captures the correlation between data points in the input space. The kernel function enforces smoothness constraints on the trained function, where the observed values for closer input samples are more correlated than the ones farther apart. This is a desirable feature when modeling the "environmental niche" of a target organism where the organism abundances are correlated in the input space of environmental covariates. The training data $\mathbf{T} = \langle \mathbf{z}_1, b_1 \rangle, \langle \mathbf{z}_2, b_2 \rangle, \ldots, \langle \mathbf{z}_N, b_N \rangle$ is drawn from the GP

$$b = g(\mathbf{z}) + \epsilon \tag{1}$$

where $\epsilon$ is a Gaussian noise term. Using the training data, we can compute the posterior mean and covariance for an unobserved test data point using the following equations

$$\mu(\mathbf{z}^*) = \mathbf{k}(K + \sigma_n^2 I)^{-1}\mathbf{b} \tag{2}$$

$$\sigma^2(\mathbf{z}^*) = k(\mathbf{z}^*, \mathbf{z}^*) + \mathbf{k}(K + \sigma_n^2 I)^{-1}\mathbf{k} \tag{3}$$

Here, $\sigma_n^2$ is the measurement noise, $K$ is the covariance or Gram matrix, generated using a kernel function $k$ that captures correlations between data points. For our work, we use a squared-exponential kernel function given by

$$k(\mathbf{z}_p, \mathbf{z}_q) = e^{1/(2\lambda^2)\|\mathbf{z}_p - \mathbf{z}_q\|^2} \tag{4}$$

where $\lambda$ is the decorrelation length scale. $\mathbf{k}$ is a vector of correlations between the test data point and the training data points computed using the kernel function $k$. We used the GPML toolbox (Rasmussen and Nickisch, 2013) for MATLAB to learn the dominant input variables and hyperparameters from training data.

For a measured environmental feature vector, $\mathbf{z}^*$, the model predicts plankton abundance through the posterior distribution $p(b^*|\mathbf{T}, \mathbf{z}^*) = \mathcal{N}(b^*|\mu(\mathbf{z}^*), \sigma^2(\mathbf{z}^*))$, where $\mu(\mathbf{z}^*)$

Downloaded from ijr.sagepub.com by guest on August 7, 2015

and $\sigma^2(\mathbf{z}^*)$ are the mean function and covariance function, representing the learned GP model. We train the GP regression model on shore, and use it on board the AUV for real-time predictions of the mean plankton abundance, and the associated prediction variance.

## 2.3. Bayesian optimization

The GP regression model enables probabilistic prediction of plankton abundance given a measured environmental feature vector $\mathbf{z}$. The goal is to design a sampling policy that minimizes cumulative regret for samples post laboratory analysis. Bayesian optimization addresses the computation of the utility of candidate samples to minimize cumulative regret over multiple trials. The balance between exploiting a known model to maximize reward (high-abundant samples), and improving the model accuracy to ensure the global optimum is not missed by exploring points with high variance is discussed in the machine learning literature in the context of the multi-armed bandit problem (Audibert et al., 2009). The goal is to develop a sampling policy that chooses the next best sample, taking into account the balance between exploration and exploitation, i.e. using the mean *and* the variance of the abundance predictions. Given a prediction of mean and variance for a candidate input data point, the sampling policy concerns maximization of a utility function $h : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ where utility $u = h(\mu(\mathbf{z}), \sigma^2(\mathbf{z}))$. The goal is to maximize the utility function over the input space to obtain the best sample $\mathbf{z}_t$ for the $t$th trial

$$\mathbf{z}_t = \arg\max_{\mathbf{z}\in\mathbf{Z}^*} h(\mu_{t-1}(\mathbf{z}), \sigma^2_{t-1}(\mathbf{z})) \tag{5}$$

Hence, the utility function $h$ computes the *sample utility* using the mean and variance estimates from the previous trial, i.e. $\mu_{t-1}(\mathbf{z})$ and $\sigma^2_{t-1}(\mathbf{z})$ respectively for a candidate sample $\mathbf{z}$. For an efficient utility function to guide sampling, we use GP-UCB, a sequential stochastic optimization strategy that uses a GP regression model to minimize cumulative regret over $T$ trials, with regret bounds (Srinivas et al., 2009, 2012). Using the learned GP regression model, instead of seeking the maximum of the mean or the variance independently, the GP-UCB algorithm prescribes the utility function for a candidate sample $\mathbf{z}$

$$h(\mu(\mathbf{z}), \sigma(\mathbf{z})) = \mu(\mathbf{z}) + \beta_t^{1/2}\sigma(\mathbf{z}) \tag{6}$$

The best candidate $\mathbf{z}_t$ for timestep $t$ is given by

$$\mathbf{z}_t = \arg\max_{\mathbf{z}\in D} \mu_{t-1}(\mathbf{z}) + \beta_t^{1/2}\sigma_{t-1}(\mathbf{z}) \tag{7}$$

where $\mathbf{z}_t$ is the sampling candidate for the $t$th trial, and $\beta_t$ is a constant that grows logarithmically with each trial, given by

$$\beta_t = 2\log\left(\frac{|D|\, t^2\pi^2}{6\delta}\right) \tag{8}$$

Here, $|D|$ is the number of input dimensions, $t$ is the trial number (or iteration), and $\delta$ is a parameter that defines the probability of the regret bound being satisfied. By targeting points that maximize a combined function of mean and variance, the GP-UCB algorithm strikes a balance between exploration and exploitation, with the goal of keeping the average regret within bound after $T$ trials.

The GP-UCB algorithm expects update of the model, i.e. functions $\mu(\mathbf{z})$ and $\sigma^2(\mathbf{z})$, after every sample has been acquired and labeled. In our problem, however, the GP model can only be updated once a *batch* of $k$ samples have been retrieved and labeled after a survey. The GP-UCB algorithm has been investigated for the batch setting (GP-BUCB) for parallelizing exploration–exploitation trade-offs, but the sampling decisions are made with complete knowledge of the predicted utility for the input space (Desautels et al., 2014). Based on this map, the GP-BUCB algorithm selects the $k$ candidate samples to be labeled for each trial. However, in our approach for opportunistic sampling the AUV can predict the utility of the sample only *during* the survey as it acquires real-time measurements of environmental covariates, without knowledge of future sample utilities. In this setting, the GP-BUCB algorithm cannot be applied, and instead we simply use the GP-UCB algorithm by selecting samples with the top $k$ utilities *online* for every trial, instead of planning for a batch of $k$ samples as prescribed by the GP-BUCB algorithm. Empirical results presented in Section 3 demonstrate the efficacy of our approach. The pseudo-code for our batch-update GP-UCB algorithm is described in Algorithm 1.

We define the average regret $R_t$ of a sample set from a trial $B_t$ as the average difference between the sum of abundance of acquired samples, and the best sum possible if values of all candidate samples are known. Note that the values of all candidate samples are not available to the batch-update GP-UCB algorithm during an actual trial (i.e. deployment); however, for analyzing the performance (Section 3), we have deployment datasets for which the values of all samples are known. For a trial $t$, $R_t = \Sigma_{i=1}^{k}(b_i^* - b_i)/k$, where $b_i^*$ is the $i$th top ranked sample in hindsight, and $b_i$ is the $i$th acquired sample. To evaluate the performance of a sampling algorithm, we use cumulative regret over $T$ surveys, defined as $R = \Sigma_{t=1}^{T} R_t$. A lower cumulative regret shows better long term reward collection by the algorithm.

## 2.4. Optimal stopping theory

Taking into account the lack of a priori information about the utility of candidate samples, the batch-update GP-UCB algorithm demands online collection of $k$ top-utility samples (i.e. Algorithm 1, line 3). An additional constraint is that sample collection decisions are irrevocable by design to avoid cross-contamination across samples. To maximize the total utility of the $k$ irrevocable samples, the online algorithm needs to have the following three properties. Firstly, the algorithm should not be too greedy, thereby

Downloaded from ijr.sagepub.com by guest on August 7, 2015

**Algorithm 1.** Batch-update GP-UCB algorithm.

**Data:** Input dimension $D$, GP Prior $\mu = 0$, $\sigma_0$, $k$
**for** $t \leftarrow 1$ **to** $T$ **do**
  $\beta_t = 2\log\left(\frac{|D|t^2\pi^2}{6\delta}\right)$;
  Choose top $k$ arguments $\mathbf{Z}_t = \mathbf{z}_1 ... \mathbf{z}_k$ corresponding to top $k$ peaks of $\mu_{t-1}(\mathbf{z}) + \beta_t^{1/2}\sigma_{t-1}(\mathbf{z})$;
  Sample set $B_t = g(\mathbf{Z}_t) + \epsilon_t$;
  Perform Bayesian update to obtain $\mu_t$ and $\sigma_t$;
**end**

missing future samples with higher utility than those already acquired. Secondly, the sampling algorithm should not be too conservative, resulting in an opportunity cost of fewer than $k$ samples collected during the survey. Optimal sampling would be exactly $k$ samples collected during each survey. Finally, it is undesirable to manually specify parameters for making sampling decisions. Typically, in a campaign with multiple surveys spanning several days, a scientist will have to manually specify thresholds for environmental parameters. While this may be feasible over short durations (i.e. days) or for a single robot, it becomes unrealistic as the temporal scale and/or the number of robots are increased. We believe scenarios where multiple AUVs are employed for autonomous environmental data and water sample collection, over extended durations, are not far off.

Optimal stopping theory addresses the problem of when to take a particular action (e.g. acquire a sample). In our problem the desired action is to trigger a water sampler $k$ times during each survey. Considering the case of a single decision, the hiring (or secretary) problem (Ferguson, 1989) addresses the optimal selection of the top candidate from $n$ applicants interviewed sequentially for a job. In this setting candidates are ranked after each interview and the recruiter has to make an irrevocable hire or reject decision immediately after the interview. Assuming candidates arrive for the interview IID, the optimal strategy then is to observe the first $n/e$ candidates without hiring (training window), and select the next best candidate. If no better candidate is found, then the last candidate is hired. This strategy results in the best candidate being selected $1/e$, or 37% of the time. Various approaches have been discussed to extend the hiring problem to the case where the top $k$ candidates need to be chosen instead of a single candidate. Using a modified training window size of $n/(ke^{1/k})$, Girdhar and Dudek (2009) presents an algorithm to choose the $k$ top-ranked secretaries with a probability of $1/ke$. The secretary algorithm has also been extended to the case where the candidates can be rated instead of being ranked, with the goal of maximizing the expectation of a submodular set function that defines efficiency of the selected secretarial group based on their overlapping skills. Called the submodular secretary problem (Bateni et al., 2010), this is the ideal formulation for our problem for a couple of reasons. First, as opposed to the case where the candidates are ranked and a probabilistic guarantee is presented for choosing the $k$ top-ranked candidates (Girdhar and Dudek, 2009), the submodular secretary algorithm provides a mechanism to select the set of candidates with the top sum of individual rating of candidates. For a ground set $V$ of possible candidates, the submodular secretary algorithm is an approximation algorithm for finding the set $S$ with the maximum expected skill measured using a submodular set function $F : 2^V \rightarrow \mathbb{R}$. Specifically, this *skill function* is chosen to be a set-sum function $F(S) = \Sigma_{v \in S} v$, where the skill of set $S$ is the sum of the values of the set elements. The set-sum function $F$ is an additive measure with the property $F(A) + F(B) = F(A \cap B) + F(A \cup B) \forall A, B \subseteq V$, and hence monotone submodular (Bach, 2011).[2] Exploiting the submodularity of the set-sum function $F$, the submodular secretary algorithm seeks to find $\text{argmax}_S F(S)$ such that $S \subset V$, $|S| = k$ by splitting the stream of candidates into $k$ equal windows and applying the secretary algorithm on each section. This strategy results in a competitive ratio of $(1 - 1/e)/11$. We use the submodular secretary algorithm to make online choices for water samples to be collected, where the skill function $F$ captures the sum of observed organism abundance of the selected set of water samples $S$.

The second reason for using the submodular secretary algorithm has to do with the determination of thresholds. The multi-choice hiring algorithm uses a single threshold to select $k$ candidates, which in our application can result in the acquisition of spatially neighboring samples at the cost of future potentially higher value samples. This happens when the AUV carries out vertical yo-yo profiles through three-dimensional geographic space and iteratively samples a vertically narrow (meters thick), chlorophyll-rich feature. Hence, neighboring samples can be correlated when the AUV traverses through a layer with high target organism abundance. The submodular secretary algorithm demonstrates a better worst case performance than the multi-choice secretary algorithm under these circumstances since samples are split uniformly into $k$ equal windows prior to independent application of the secretary algorithm to each window, resulting in exactly one sample per window. The size of each window is determined from the planned deployment time, and typically consists of multiple windows. Pseudo-code for the submodular secretary algorithm is described in Algorithm 2.

## 3. Simulation studies

We evaluated our approach by mining previously collected AUV data from an eight-day campaign in 2005, consisting

Downloaded from ijr.sagepub.com by guest on August 7, 2015

**Algorithm 2.** Submodular secretary algorithm to maximize utility of $k$ online water samples.

```
Data: Gulp set B = ∅, number of gulpers= k, total survey samples expected= N, stopping parameter= r, current trial= t
Result: Gulp set B
Window duration N_w = ⌊N/k⌋;
Observation duration N_o = ⌊N_w/r⌋;
Start survey: n = 0;
Set window count w = 1;
Set best utility u* = 0;
while survey time n ≤ N do
  while not at end of current window w do
    read current environmental feature vector z_n;
    use GP model to compute utility of candidate sample u(z_n) = μ(z_n) + β^{1/2}σ(z_n);
    if within observation window then
      if u_n > u* then                                  /* update utility threshold */
        Set u* = u_n;
      end
    else                                                /* sampling window */
      if no candidate added to B from current window w then
        if u_n > u* then
          add candidate at n to B;
        end
      end
    end
  end
  if no candidate added to B from current window w then   /* sample at end of window */
    add current candidate to B;
  end
  Increment window count w while w ≤ k;
  Set best utility u* = 0;
end
```

of at least two surveys per day and 17 surveys in total. Physical sample collection was emulated by taking AUV survey environmental data, and hiding one of the measured parameters, chlorophyll fluorescence (a proxy for phytoplankton abundance; the property of interest), from the sampling algorithm during each survey. Our framework used batches of $k$ emulated "gulps" (or samples) with high chlorophyll fluorescence for each survey, with the goal of minimizing cumulative regret of samples from the whole campaign. Chlorophyll fluorescence, although measured in-situ by the AUV, was subsequently converted into a hidden ground-truth by our framework, and only revealed at the end of each survey. The workflow for our simulation framework is shown in Figure 2. Being a proxy for algal biomass, chlorophyll fluorescence was also an ideal candidate for our analysis since it evaluates the prediction of biological features from environmental covariates only, ignoring geographic parameters such as latitude, longitude, and depth.

Chlorophyll fluorescence was predicted by a probabilistic model that uses environmental covariates measured by the AUV as inputs. These predictions were used to collect $k$ samples during every simulated survey and samples were labeled at the end of each survey when our framework revealed the measured chlorophyll fluorescence to the sampling algorithm. The newly labeled data was assimilated into the training dataset to update the predictive model, and used to guide sampling decisions in subsequent surveys.

To compute a predictive model for chlorophyll fluorescence, we used $p$ pilot surveys to collect $pk$ random samples. The GPML toolbox (Rasmussen and Nickisch, 2013) for MATLAB (2013) was used to learn a GP regression model from the $pk$ records. In subsequent surveys this model was used to guide sampling with the goal of minimizing cumulative regret of the whole campaign.

The $k$ samples acquired during each subsequent survey formed the set $B = \{b_1, b_2, \ldots, b_k\}$. An optimal set $B^+$ was computed offline on each survey resulting in $k$ samples with maximum abundance (global chlorophyll fluorescence peaks). To serve as a baseline, we also computed a set, $B_r$, that consisted of $k$ random samples without replacement from each survey. To determine how well the sampling policy performs in an exploitation-exploration setting over multiple surveys, we used the average regret of acquiring the set $B$ instead of the optimal set $B^+$. To evaluate the accuracy of the learned model for different choices of environmental parameters as inputs, we computed the correlation coefficient for predicted and observed values of chlorophyll fluorescence for all samples collected during each survey.

Based on the approach chosen for sample labeling and model updates, we investigated three cases. First, we stopped labeling after $p$ pilot surveys, and used the resulting model for all subsequent sampling decisions. This was indexed as the initial or INI strategy. Next, by continually

Downloaded from ijr.sagepub.com by guest on August 7, 2015

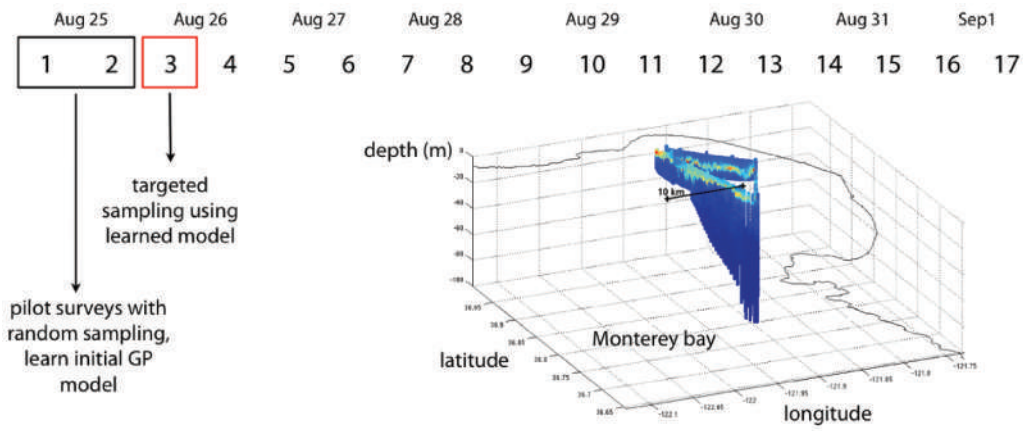



**Fig. 2.** A campaign of 17 AUV surveys was carried out over eight days in August 2005. We used chlorophyll fluorescence, one of the in-situ measurements, as the property of interest, and emulated collection of samples with high values for chlorophyll fluorescence. *pk* samples are randomly sampled from the first *p* surveys ($p = 2$ for day 1), followed by training and update of a GP model to guide sample collection in subsequent surveys. The inlay at right shows a synoptic view of chlorophyll fluorescence from one of the surveys, with warmer colors depicting higher values. The same triangular survey pattern was repeated 17 times over eight days.

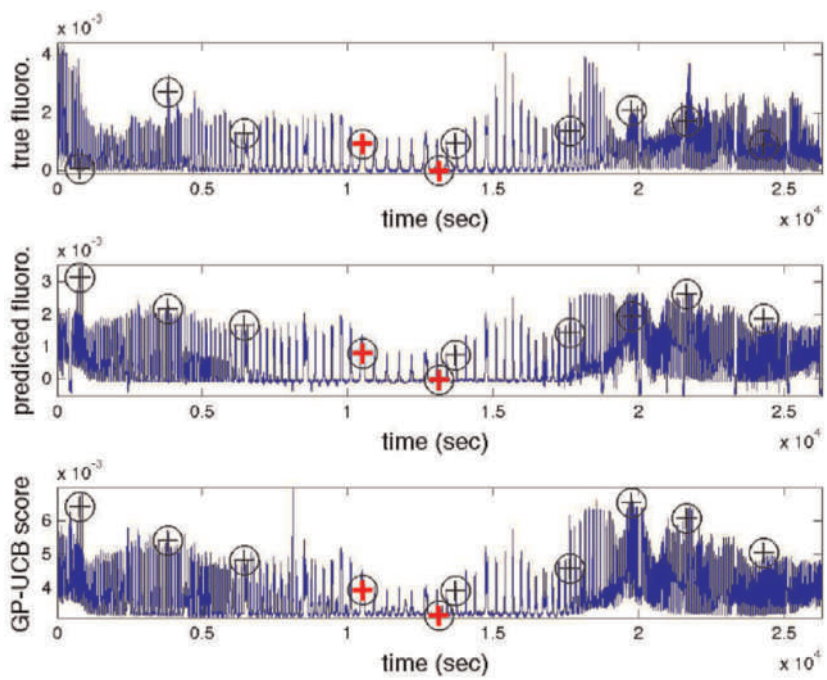



**Fig. 3.** Plots showing the measured true fluorescence signal (top), predicted fluorescence mean signal (center), and the GP-UCB score during survey 11/17 of one of the simulated campaigns (bottom). Black crosshairs show the gulps taken within sampling windows using the submodular secretary algorithm, and red crosshairs show the gulps taken at window ends due to lack of better candidates within the sampling windows in question.

labeling after each survey, we updated the model using only data from the previous *p* surveys (with $p = 2$, the window or WIN strategy). Finally, when we used all the data from previous surveys to keep the GP model updated, we generated the case for all data. For each of the three update methods, we used three utility functions, i.e. mean, variance, and GP-UCB. Along with random sampling, this resulted in 10 methods in total. We emulated 100 campaigns for each method, by starting with a new set of *k* samples drawn randomly from the first *p* pilot surveys. Figure 3 and Figure 4 show the signals (true, predicted mean, and GP-UCB utility) and the transect plots (true and predicted mean) for survey 11/17 of one of the simulated campaigns. In this example, the submodular secretary algorithm has been used on the GP-UCB utility, predicted using a model that was learned using all data until survey 10/17.

Downloaded from ijr.sagepub.com by guest on August 7, 2015

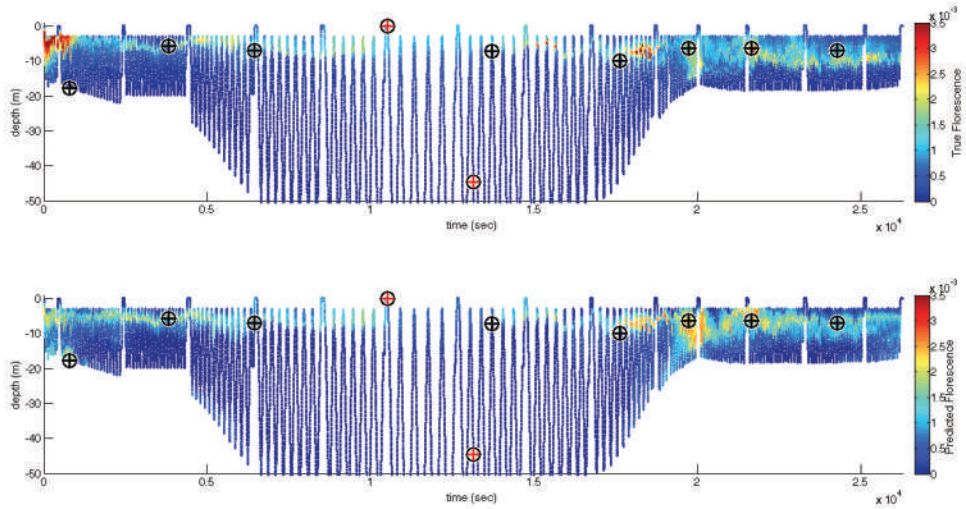



**Fig. 4.** The true (top) and predicted (bottom) chlorophyll fluorescence from survey 11/17 of one of the simulated AUV campaigns. It shows the AUV carrying out vertical profiles between the surface and a depth of 100 m (shown between the surface and 40 m depth for clarity). The submodular secretary algorithm is used on the utility computed from probabilistic abundance predictions, and black crosshairs show gulps taken within sampling windows. Red crosshairs show gulps taken at the window ends due to lack of better candidates within the sampling windows in question.

Figure 5 shows the regret and cumulative regret of each survey, averaged over 100 campaigns. We observe that the lowest cumulative regret is obtained when the GP-UCB algorithm is used with a model updated with all data up to previous survey (ALL-GPUCB). The worst performance is when random sampling is used during each survey. The variance driven utility function (VAR) is analogous to using a greedy strategy to maximize mutual information in the sequential setting (Krause and Guestrin, 2007). Although considered in the batch setting, the VAR utility function serves as a comparison of the GP-UCB utility with the scenario where global model accuracy is improved after every iteration. In contrast, the mean-driven utility function targets samples with highest predicted mean (reward), demonstrating only exploitation and no exploration. When using all data, the GP-UCB approach outperforms the other model update methods, demonstrating the advantage of the trade-off between exploration and exploitation. Figure 6 highlights the result of the mean-driven strategy where the model misses the global optimum in the high-backscatter and high-temperature region of the input space. Survey 12 shows high regret with random sampling, suggesting the presence of a stronger hotspot compared to other surveys. ALL-GPUCB shows a trend of lower regret with each survey, validating the claim for this work. Figure 7 shows the summary of campaign results for the 10 methods.

Our results also demonstrate how the GP-UCB algorithm learns the distribution of chlorophyll fluorescence in the backscatter–temperature space in a data-driven manner. Specifically, it successfully learns a known relationship between chlorophyll fluorescence, backscatter, and temperature. Backscatter is a measure of particle concentration and shows strong correlation with algal biomass; however, due to the presence of suspended sediment particles of comparable or greater concentration than algae, especially in deeper waters, additional information is necessary for the accurate prediction of chlorophyll fluorescence.

The GP-UCB algorithm learns the trend between backscatter and depth when predicting fluorescence by defining peak fluorescence in the high-temperature and high-backscatter region of the temperature–backscatter input space. This is illustrated in Figure 6, and results in the precise acquisition of chlorophyll fluorescence samples through targeting of high-backscatter samples that are also warmer, i.e. near the surface.

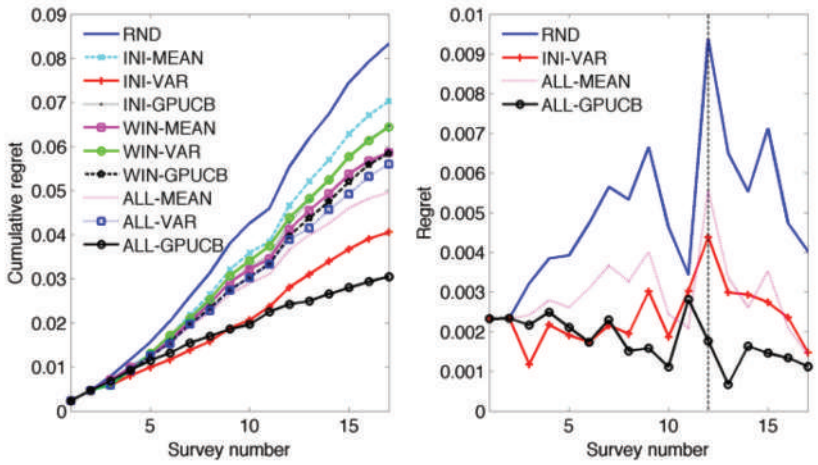
## 4. Field experiment

Our field experiment was carried out in October 2013, targeting PN a genus of phytoplankton known to cause potentially toxic blooms. The goal was to acquire AUV water samples rich in PN, over a course of a six hour survey carried out in northern Monterey Bay (Figure 8). The survey consisted of three 1 km $\times$ 1 km Lagrangian box patterns (Das et al., 2012), with the Dorado AUV conducting vertical yo-yos to a depth of 30 m while tracking a virtual drifter. The experimental design and a summary of results indicating high PN abundance in AUV water samples are discussed below.

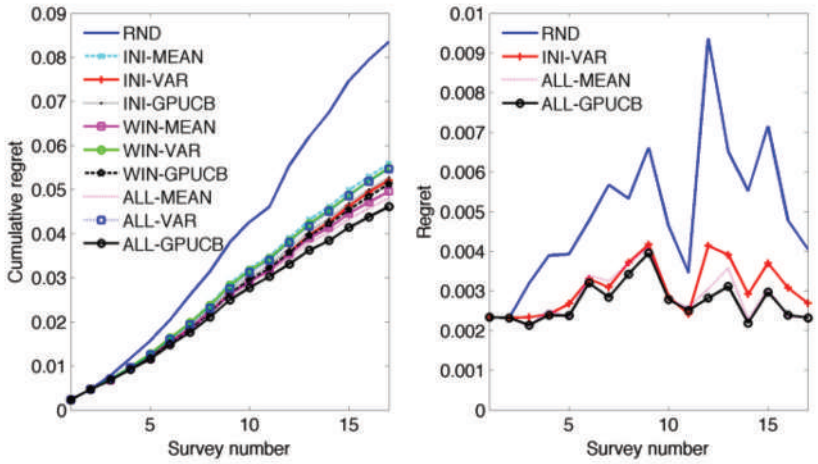
### *4.1. Experiment design*

A training dataset was composed from molecular analysis results of 87 water samples collected by the Dorado AUV during a field campaign in October 2010 (the same season) using the chlorophyll peak-capture algorithm (Zhang et al.,

Downloaded from ijr.sagepub.com by guest on August 7, 2015




(a) Offline sampling with full knowledge of the survey. The GP-UCB algorithm (black solid line and circles) has the lowest cumulative regret (left-hand plot) at the end of the campaign, and also has progressively lower regret as the campaign proceeds (right-hand plot). The variance-driven sampling policy (INI-VAR; red line, + symbol) trained on data from pilot surveys resulted in superior performance compared with the other sampling strategies. Additionally, for survey 12 (vertical dashed line in the right-hand plot), a strong phytoplankton hotspot resulted in high cumulative regret for the random sampling strategy. However, the GP-UCB policy demonstrated a reduction of regret in this survey.




(b) Online sampling using submodular secretary algorithm. The cumulative regret of the GP-UCB algorithm is lowest at the end of the campaign, although in the online case the performance is poorer than in the case with full knowledge.

**Fig. 5.** Cumulative regret measured over the course of the 17-survey campaign, and averaged over 100 simulated campaigns.

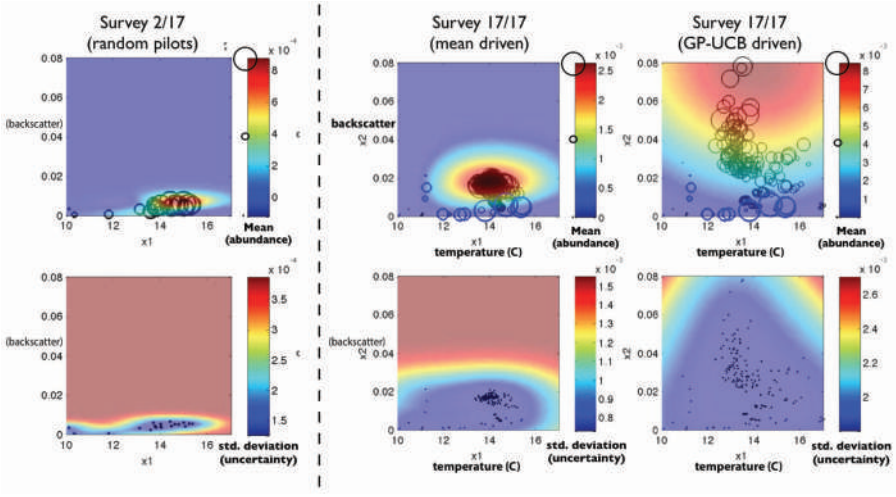
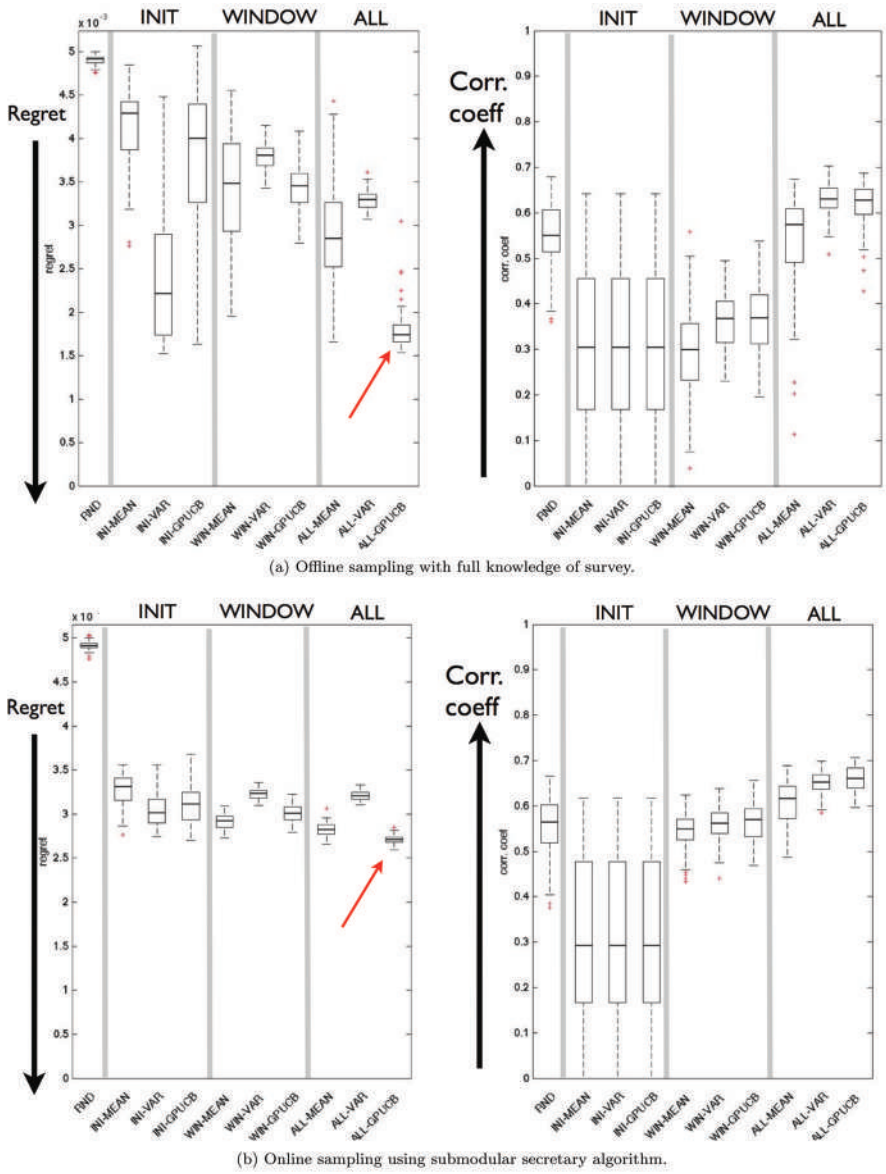



**Fig. 6.** Survey 2/17 and survey 17/17 during one of the emulated campaigns. The top plots and bottom plots show prediction means and variances, respectively, for the input space. Circles on the top plots and dots on the bottom plots show ex-situ sample locations. Circle sizes are proportional to the true value of chlorophyll fluorescence, after ex-situ labeling.

Downloaded from jfr.sagepub.com by guest on August 7, 2015




(a) Offline sampling with full knowledge of survey.

(b) Online sampling using submodular secretary algorithm.

**Fig. 7.** Summary statistics of regret and correlation coefficients for 100 simulated campaigns. Lower regret and a higher correlation coefficient are desirable. The results are divided into four parts based on how the samples are selected: random sampling used as a baseline (RND), sampling using a model learned from samples acquired during an initial $p$-survey pilot and never updated (INI), sampling using a model learned on data from a window of $p$ previous surveys (WIN), and sampling using a model learned on all available data so far. Regret statistics for the ALL-GPUCB algorithm demonstrated the lowest regret for both the offline and online scenarios (red arrows).

2010). Along with measured PN abundance, the in-situ measurements of temperature, salinity, chlorophyll fluorescence, dissolved oxygen, backscatter, and nitrate concentration from the AUVs onboard sensor suite were recorded in the training dataset. The hyperparameters for the kernel function, and the input variables were chosen using four-fold cross validation, and chlorophyll fluorescence and temperature were observed to be the dominant input parameters for prediction of PN abundance. Figure 9 shows the predicted mean and variance of PN abundance over a range of chlorophyll fluorescence and temperature values. Using the trained PN abundance model, the goal was to acquire water samples with PN in high abundance. Due to operational constraints, sampling was restricted to a single survey. Given that we had a model trained on a large number of analyzed water samples (equivalent to about nine surveys), we use an exploitation only sampling policy ($\beta = 0$) to maximize the expected PN abundance from the samples.

We ran the GP model for PN abundance on board the Dorado AUV during a deployment lasting six hours in north Monterey Bay, resulting in a value prediction for PN abundance $b^*$ for every measured environmental feature vector $\mathbf{z}^*$. The sample variance was not used during this trial because we used a mean-driven sampling strategy. We

Downloaded from ijr.sagepub.com by guest on August 7, 2015

included AUV surface time (necessary for communication with the vehicle) in our estimation of total survey duration. Since the AUV spent longer than expected on the surface due to variability in surface data transmission time, our algorithm collected eight gulps out of the tasked nine. The ninth gulp was not triggered because the last segment was not completed within the preset duration.

The total duration of the survey was set to be approximately 2.18 hours, with the total expected in-situ data points $N = 15{,}725$. The submodular secretary algorithm splits this window into nine segments, resulting in the segment size $N_w = 1747$, approximately 14 minutes. We used stopping parameter $r = 3$, resulting in the observation window length for each segment being approximately 5 minutes. A moving-average filter is used by the submodular secretary algorithm to filter out spikes in the sensor data while carrying out threshold updates, with a tolerance of 0.2 optical density (a proxy for abundance, measured by molecular analysis (Harvey, 2014)). This tolerance is also used when making firing decisions to avoid erroneous gulps resulting from data spikes.

## 4.2. Results

Figure 10 shows the AUV transect and the predicted PN mean abundance. Black crosses show the locations where gulps were taken. We observe that five out of eight gulps were taken in PN hotspots (the region in red showing high predicted abundance). Out of the remaining three samples, sample five appears to be at the boundary of the PN hotspot at a depth of approximately 14 m, and samples six and eight in deeper waters with low predicted abundance. The corresponding distributions of the gulp locations in the environmental parameter space (temperature, fluorescence) are shown in Figure 11. This figure reflects the distribution of gulps with respect to the PN abundance prediction model used. Gulp numbers one, two, three, four, and seven were taken close to the PN abundance prediction peak, demonstrating the performance of the submodular secretary algorithm in targeting prediction peaks.

The submodular secretary algorithm's sample collection decisions are shown in Figure 12, highlighting the gulp locations with respect to predicted PN abundance and adaptive threshold updates. The periodic variations in predicted PN abundance corresponded to the movement of the AUV through the predicted PN layer (high abundance between depths of 4 m to 10 m). The solid line shows how the thresholds were picked, with updates happening in, approximately, the first six minutes (observation window) of each segment with total duration of approximately 18 minutes, and the sampling happening in the remaining 12 minutes (sampling window). When a better candidate was not found in the sampling window, samples were taken at the end of the window. This explains the low PN abundance for gulps five, six, and eight. Data spikes were filtered during threshold update, and hence the sampling algorithm ignored extreme values.

**Fig. 8.** The trial site for the October 2013 Dorado AUV field experiment in northern Monterey Bay, marked with a white dot.

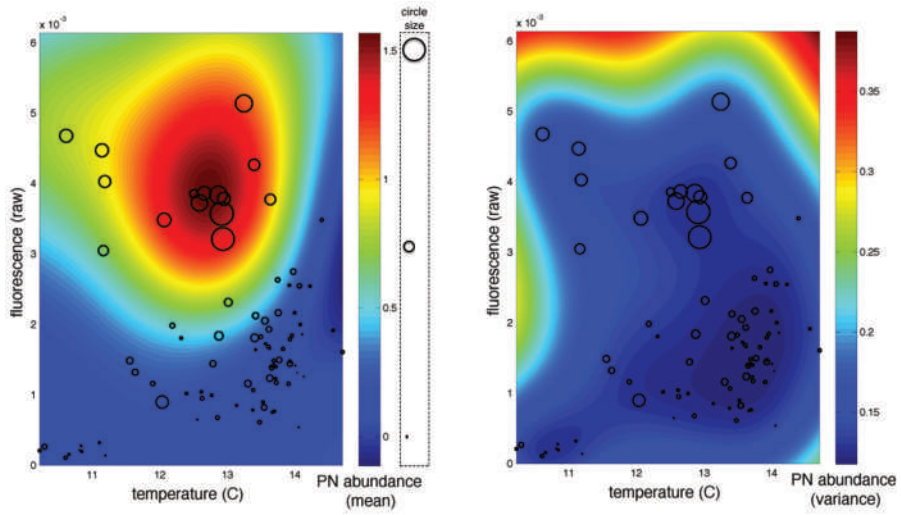
Figure 13 shows the distribution of gulps as vertical dashed lines overlaid on histograms for chlorophyll fluorescence (Figure 13(a)), temperature (Figure 13(b)), and the predicted PN abundance (Figure 13(c)). The first two figures show the distribution of gulps across the two input parameters, whereas the last figure shows the distribution of gulps across predicted PN abundance. A key observation from this figure is that the randomly fired gulps at the window ends essentially resulted in control samples. Because PN hotspots occupy only a small portion of the water column, the mode of the predicted PN distribution consists of negligible abundance values. Statistically, samples at window ends are likely to be at the mode, resulting in low value samples that are useful as controls.

Also, in the histograms of fluorescence and temperature, the distribution of samples acquired within the sampling window was close to the predicted PN abundance peaks (Figure 9). This is also evident in Figure 11 that shows the location of AUV measurements, and gulps are shown in environmental parameter space.

## 4.3. Ex-situ sample analysis

Following the experiment, the eight AUV water samples were morphologically analyzed (microscopy) in a marine microbiology laboratory to directly enumerate PN abundance. The comparative results of algorithm predicted and microscopy counts of PN abundances are shown in Figure 14. Consistent trends are evident between the predicted and directly counted PN abundances for the acquired samples. Note that the units for the predicted and enumerated

Downloaded from ijr.sagepub.com by guest on August 7, 2015




(a) PN model mean. (b) PN model variance.

**Fig. 9.** Trained PN model used for the October 17 trial. The size of the circles is proportional to the measured PN abundance through molecular analysis (highlighted next to the color scale for the PN model mean).

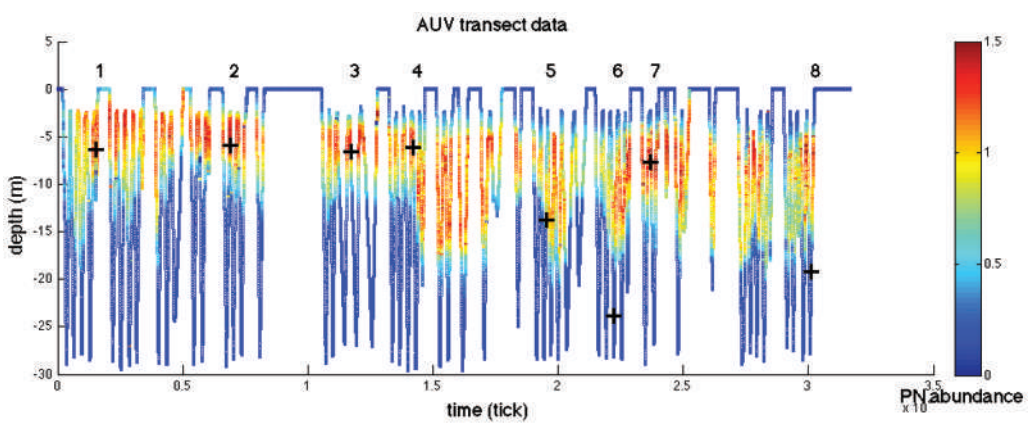



**Fig. 10.** Transect plot of the October 2013 AUV survey showing time on the *x*-axis and depth on the *y*-axis. The color shows predicted PN abundance, with red corresponding to high values. Black " + " symbols mark the locations where gulps were taken using our approach.

abundances follow from molecular and morphological analysis respectively, but both assess an abundance measure for the targeted organism (PN). The sample analysis results show that we successfully trained a model, ran it onboard the AUV to predict organism abundance in real time, and acquired samples that were rich in the targeted organism of interest. The fact that trends in data-driven predictions matched measurements obtained by taxonomic experts suggests that the predictive capabilities of our algorithm are accurate. Accomplishments of the ex-situ comparative AUV water sample analysis are as follows.

1. We designed and conducted an experiment in an entirely data-driven fashion. Predictive models were trained from laboratory analyzed water sample molecular detection data from a previous season, October 2010. The model succeeded in October 2013 owing to the model's predictive accuracy under similar environmental conditions associated with the Fall season. The PN abundance prediction algorithm successfully learned the "environmental niche" associated with the target organism and exploited this fact during the October 2013 field trial.
2. Scientists were able to observe the characteristics of predictive models and the behavior of the sampling algorithms through expressive plots. The use of probabilistic models is especially useful since scientists have a sense of the uncertainty inherent in predictive models, facilitating introspection.
3. By assimilating the observed PN abundance from the data back into the training dataset, the potential for even higher abundance samples in future field trials

Downloaded from ijr.sagepub.com by guest on August 7, 2015

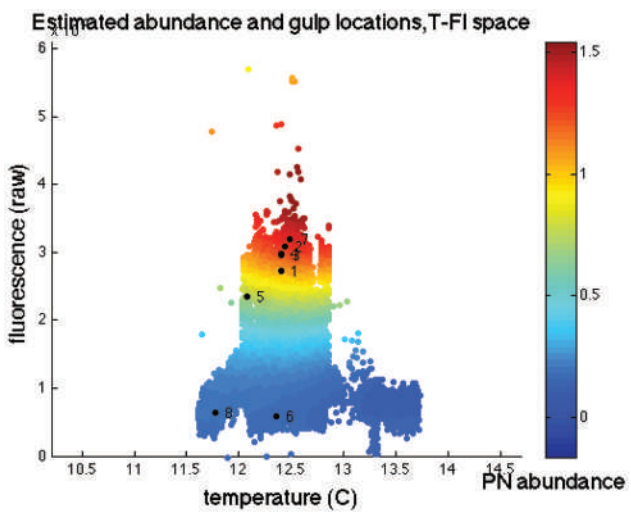



**Fig. 11.** Data points corresponding to in-situ measurements of temperature (*x*-axis) and fluorescence (*y*-axis) taken by the AUV during the survey, along with the predicted PN abundance (color), and the AUV water sample locations in the temperature–fluorescence space (numbered dots).

can be realized. This in effect "closes the loop" on marine ecosystem monitoring, with potential applications in a variety of other applications.

## 5. Conclusions

In this paper, we have presented a principled approach for data-driven opportunistic robotic sampling to collect batches of physical (water) samples iteratively. Our approach has wide application in the earth sciences, where physical samples can often only be labeled ex-situ after the completion of each robot survey. Using marine ecosystem monitoring as a test domain, we developed a sampling strategy to minimize cumulative regret of water samples collected during an AUV field campaign consisting of multiple surveys. We used previously collected data from pilot surveys to train a GP regression model for probabilistic prediction of target phytoplankton abundances. These predictions were used in a Bayesian optimization setting to maximize the utility of batches of physical samples collected during each subsequent survey, and labeled after completion of the survey. Since prior information on the distribution of predicted utility is not available during opportunistic surveys, optimal stopping theory considerations were used to maximize the utility of the sample batch online. We evaluated our work extensively using a simulation framework that emulated sample collection campaigns by mining previously collected AUV data. The results from 100 campaigns with different initial conditions, and for different sampling strategies, demonstrated the lowest cumulative regret for samples collected adaptively using

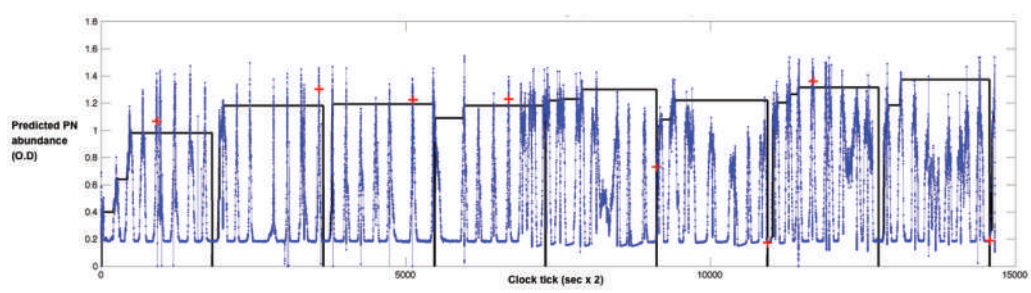



**Fig. 12.** The time series of predicted PN abundance from the deployment, with the threshold choices for each segment as determined by the submodular secretary algorithm.

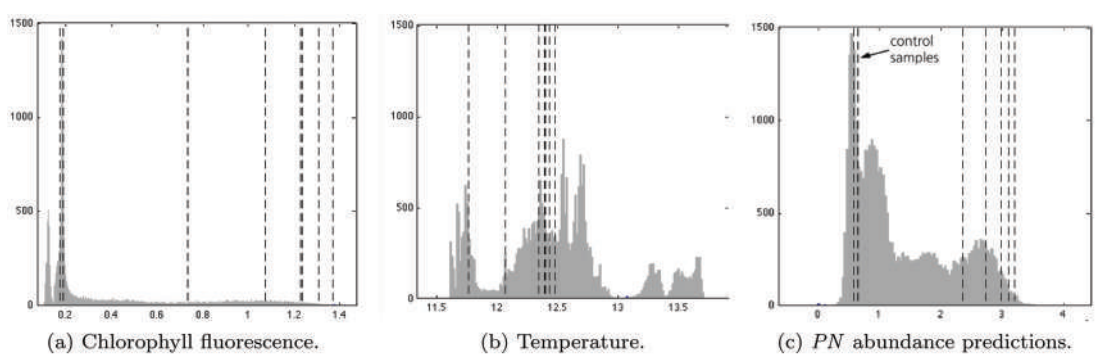
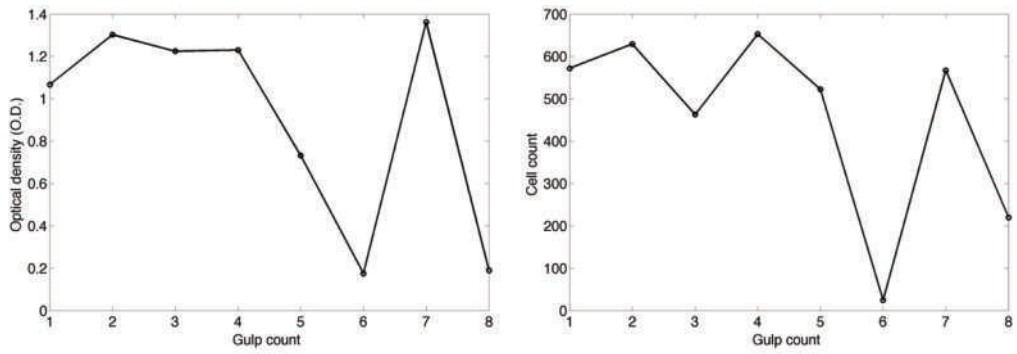



**Fig. 13.** Histograms showing the distribution of input measurements and predicted output (PN abundance) from the October 17 AUV survey. Vertical dashed lines show locations of the eight gulps acquired during the survey. The predicted PN abundance (c) demonstrates how gulps acquired at the end of sampling windows are by design random, and likely to happen at the mode of the predicted signal, which in the mean-driven strategy is PN abundance. The mode in this case corresponds to low-abundance samples that also serve as controls.

Downloaded from ijr.sagepub.com by guest on August 7, 2015




(a) Predicted PN abundance, measured in optical density (O.D.), a unit corresponding to molecular methods. (b) Measured PN abundance, estimated as cell count through microscopy (morphological method).

**Fig. 14.** A comparison of onboard PN abundance predictions, and the PN abundance measurements after laboratory analysis for the eight water samples collected during the field trial. A consistent trend was observed between the predicted and measured PN abundances, validating our hypothesis that data-driven physical sample collection can be carried out using a model trained on previously acquired data, without using geographical parameters as inputs.

the GP-UCB sampling policy, updated on available data from all previous surveys. Performance online, using the submodular secretary algorithm, showed the same trends. A one-day field trial was carried out using a GP regression model trained on 87 previously collected samples by an AUV. Using a mean-driven strategy, the AUV collected eight water samples, highly abundant in the harmful algal bloom producing diatom PN. This is the first time such a field experiment has been carried out in its entirety in a data-driven fashion, in effect "closing the loop" on model learning since newly collected data is always assimilated into the training dataset to improve model accuracy.

Our work illustrates a path to increased autonomy wherein a scientist can task the vehicle at a higher level instead of providing waypoints or other positional information. By iteratively learning the "ecological niche" of target organisms, our robotic sampling methodology will allow predictions to be made through improved ecological modeling. Finally, although tested in the context of marine ecosystem monitoring, our approach can be applied to a variety of other environmental monitoring problems that demand persistent collection of physical samples for ex-situ analysis. Examples include precision agriculture, forestry, surficial geology, aerobiological sampling, and air quality monitoring.

## Acknowledgements

We thank the David and Lucile Packard Foundation for supporting our work at Monterey Bay Aquarium Research Institute, and the crew of the R/V Zephyr and R/V Rachel Carson for help with deployments.

## Funding

This work was supported by National Science Foundation (grant numbers CCF-0120778 and IIS-1125015) and the Office of Naval Research (grant number 000140911031). The AUV data used for simulations was collected as part of the Office of Naval Research funded layered organization in the coastal ocean department research initiative (grant number N000140410311).

## Notes

1. Due to ocean dynamics, the geographical distribution of plankton can change dramatically within a couple of hours.
2. A set function is submodular if and only if $\forall A, B \subseteq V : F(A) + F(B) \geq F(A \cap B) + F(A \cup B)$. Alternatively, the marginal profit of each item should be non-increasing, i.e. $\forall B \subseteq A \subseteq V : F(A \cup a) - F(A) \leq F(B \cup b) - F(B)$. The set function $F : 2^V \rightarrow \mathbb{R}$ is monotone if $F(A) \leq F(B)$ for $A \subseteq B \subseteq V$.

## References

Anderson CR, Sapiano MRP, Prasad, et al. (2010) Predicting potentially toxigenic Pseudo-nitzschia blooms in the Chesapeake Bay. *Journal of Marine Systems* 83(3): 12–140.

Anthony D, Elbaum S, Lorenz A, et al. (2014) On crop height estimation with UAVs. In: *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, Chicago, USA, 14–18 September, pp. 4805–4812. IEEE.

Audibert JY, Munos R and Szepesvári C (2009) Exploration–exploitation tradeoff using variance estimates in multi-armed bandits. *Theoretical Computer Science* 410(19): 1876–1902.

Azimi J, Jalali A and Zhang-Fern X (2012) Hybrid batch Bayesian optimization. In: *Proceedings of the 29th international conference on machine learning (ICML)*, Edinburgh, UK, 26 June–1 July, pp. 1215–1222. Madison: Omnipress.

Bach F (2011) Learning with submodular functions: A convex optimization perspective. Report, arXiv: abs/1111.6453.

Batalin MA, Rahimi M, Yu Y, et al. (2004) Call and response: Experiments in sampling the environment. In: *Proceedings of the 2nd international conference on embedded networked sensor systems*, Baltimore, USA, 3–5 November, pp. 25–38. New York: ACM.

Bateni M, Hajiaghayi M and Zadimoghaddam M (2010) Submodular secretary problem and extensions. In: *Proceedings of the*

Downloaded from ijr.sagepub.com by guest on August 7, 2015

*approximation, randomization, and combinatorial optimization. Algorithms and techniques: 13th international workshop, APPROX 2010, and 14th international workshop, RANDOM 2010*, Barcelona, Spain, 1–3 September 2010, vol. 6302, pp. 39–52. Berlin: Springer Science & Business Media.

Binney J, Krause A and Sukhatme GS (2013) Optimizing waypoints for monitoring spatiotemporal phenomena. *International Journal of Robotics Research* 32(8): 873–888.

Bryson M, Reid AS, Ramos FT, et al. (2010) Airborne vision-based mapping and classification of large farmland environments. *Journal of Field Robotics* 27(5): 632–655.

Chen Y and Krause A (2013) Near-optimal batch mode active learning and adaptive submodular optimization. In: *International conference on machine learning (ICML)*, vol. 28, Atlanta, USA, 16–21 June, pp. 160–168.

Das J, Py F, Maughan T, et al. (2012) Coordinated sampling of dynamic oceanographic features with underwater vehicles and drifters. *International Journal of Robotics Research* 31(5): 626–646.

Das J, Rajan K, Frolov S, et al. (2010) Towards marine bloom trajectory prediction for AUV mission planning. In: *IEEE International Conference on Robotics and Automation*, Anchorage, USA, 3–7 May, pp. 4784–4790. IEEE.

Desautels T, Krause A and Burdick J (2014) Parallelizing exploration–exploitation tradeoffs with gaussian process bandit optimization. *The Journal of Machine Learning Research* 15(1): 3873–3923.

Elith J, Phillips SJ, Hastie T, et al. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17(1): 43–57.

Ferguson T (1989) Who solved the secretary problem? *Statistical Science* 4: 282–289.

Fox M, Long D, Py F, et al. (2007) In situ analysis for intelligent control. In: *Proceedings of IEEE/OES OCEANS conference*, Aberdeen, UK, 18–21 June, pp. 1–6. IEEE.

Garcia-Olaya A, Py F, Das J, et al. (2012) An on-line utility based multi-criteria approach for sampling dynamic ocean fields. *IEEE Journal of Oceanic Engineering* 37(2): 185–203.

Garg S, Singh A and Ramos F (2012) Learning non-stationary space-time models for environmental monitoring. In: *26th conference on artificial intelligence (AAAI), special track on computational sustainability and AI*, Ontario, Canada, 22–26 July. New York: ACM.

Girdhar Y and Dudek G (2009) Optimal online data sampling or how to hire the best secretaries. In: *Canadian conference on computer and robot vision*, Kelowna, Canada, 25–27 May, pp. 292–298. IEEE.

Harvey JBJ (2014) A 96-well plate format for detection of marine zooplankton with the sandwich hybridization assay. In: S Stricker and D Carroll (eds) *Methods in Molecular Biology: Developmental Biology of the Sea Urchin and Other Marine Invertebrates*. New York: Humana Press Incorporation.

Harvey JB, Ryan JP, Marin R, et al. (2012a) Robotic sampling, in situ monitoring and molecular detection of marine zooplankton. *Journal of Experimental Marine Biology and Ecology* 413: 60–70.

Harvey JBJ, Zhang Y and Ryan JP (2012b) AUVs for ecological studies of marine plankton communities: Intelligent algorithms on Dorado and Tethys AUVs enable precise water sampling for plankton research. *Sea Technology*, September.

Hitz G, Pomerleau F, Garneau ME, et al. (2012) Autonomous inland water monitoring: design and application of a surface vessel. *Robotics & Automation Magazine, IEEE* 19(1): 62–72.

Julian BJ, Angermann M, Schwager M, et al. (2012) Distributed robotic sensor networks: An information theoretic approach. *International Journal of Robotics Research* 31(10): 1134–1154.

Kim Y, Shell DA, Ho C, et al. (2013) Spatial interpolation for robotic sampling: Uncertainty with two models of variance. In: Desai JP, Dudek G, Khatib O, et al. (eds) *Experimental Robotics: Springer Tracts in Advanced Robotics*, Volume 88. Springer, pp. 759–774.

Krause A and Guestrin C (2007) Nonmyopic active learning of Gaussian processes: An exploration–exploitation approach. In: *Proceedings of the 24th international conference on machine learning (ICML)*, Corvalis, USA, 20–24 June, pp. 449–456. New York: ACM.

Krause A, Guestrin C, Gupta A, et al. (2006) Near-optimal sensor placements: Maximizing information while minimizing communication cost. In: *Proceedings of the 5th international conference on information processing in sensor networks*, Nashville, USA, 19–21 April, pp. 2–10. New York: ACM.

Low KH, Dolan J and Khosla P (2009) Information-theoretic approach to efficient adaptive path planning for mobile robotic environmental sensing. In: *Proceedings of the 19th international conference on automated planning and scheduling (ICAPS-09)*, Thessaloniki, Greece, 19–23 September. Menlo Park: AAAI.

MATLAB (2013) *version 8.1.0.604, R2013a*. Natick, MA: The MathWorks Incorporation.

Rasmussen CE (2006) *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press.

Rasmussen CE and Nickisch H (2013) GPML Gaussian Processes for Machine Learning Toolbox. http://mloss.org/software/view/263/.

Reid A, Ramos F and Sukkarieh S (2013) Bayesian Fusion for Multi-Modal Aerial Images. In: *Proceedings of robotics: science and systems*, Berlin, Germany, 24–28 June, pp. 1–8. Cambridge, MA: MIT Press.

Ryan J, Harvey J, Zhang Y, et al. (2014a) Distributions of invertebrate larvae and phytoplankton in a coastal upwelling system retention zone and peripheral front. *Journal of Experimental Marine Biology and Ecology* 459(0): 51–60.

Ryan J, McManus M, Kudela R, et al. (2014b) Boundary influences on HAB phytoplankton ecology in a stratification-enhanced upwelling shadow. *Deep Sea Research Part II: Topical Studies in Oceanography* 101(0): 63–79.

Ryan JP, Johnson S, Sherman A, et al. (2010) Intermediate nepheloid layers as conduits of larval transport. *Limnology & Oceanograhy Methods* 8: 394–402.

Shkurti F, Xu A, Meghjani M, et al. (2012) Multi-domain monitoring of marine environments using a heterogeneous robot team. In: *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 1747–1753.

Singh A, Krause A, Guestrin C, et al. (2009) Efficient informative sensing using multiple robots. *Journal of Artificial Intelligence Research* 34: 707–755.

Singh A and Ramos F, Durrant-Whyte H, et al. (2010) Modeling and decision making in spatio-temporal processes for environmental surveillance. In: *IEEE international conference on robotics and automation (ICRA)*, Anchorage, USA, 3–7 May, pp. 5490–5497. IEEE.

Downloaded from ijr.sagepub.com by guest on August 7, 2015

Srinivas N, Krause A, Kakade S, et al. (2012) Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory* 58(5): 3250–3265.

Srinivas N, Krause A, Kakade SM, et al. (2009) Gaussian process optimization in the bandit setting: no regret and experimental design. In: *Proceedings of the 27th international conference on machine learning (ICML)*, Haifa, Israel, 21–24 June, pp. 1015–1022. Madison: Omnipress.

Stealey MJ, Singh A, Batalin MA, et al. (2008) NIMS-AQ: A novel system for autonomous sensing of aquatic environments. In: *IEEE international conference on robotics and automation (ICRA)*, pp. 621–628.

Tokekar P, Bhadauria D, Studenski A, et al. (2010) A robotic system for monitoring carp in Minnesota lakes. *Journal of Field Robotics* 27(6): 779–789.

Tokekar P, Vander Hook J, Mulla D, et al. (2013) Sensor planning for a symbiotic UAV and UGV system for precision agriculture. In: *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, Tokyo, Japan, 3–7 November, pp. 5321–5326. IEEE.

Zhang B and Sukhatme GS (2007) Adaptive sampling for estimating a scalar field using a robotic boat and a sensor network. In: *IEEE international conference on robotics and automation (ICRA)*, Rome, Italy, 10–14 April. IEEE.

Zhang Y, McEwen RS, Ryan JP and Bellingham JG (2010) Design and tests of an adaptive triggering method for capturing peak samples in a thin phytoplankton layer by an autonomous underwater vehicle. *IEEE Journal of Oceanic Engineering* 35: 785–796.

Zhang Y, Ryan JP, Bellingham JG, et al. (2012) Autonomous detection and sampling of water types and fronts in a coastal upwelling system by an autonomous underwater vehicle. *Limnology Oceanography: Methods* 10: 934–951.

Downloaded from ijr.sagepub.com by guest on August 7, 2015