Learning uncertainty in ocean current predictions for safe and reliable navigation of underwater vehicles

Geoffrey A. Hollinger[*]
Robotics Program
School of Mechanical, Industrial & Manufacturing Engineering
Oregon State University
Corvallis, OR 97331
geoff.hollinger@oregonstate.edu

Arvind A. Pereira[†]
Clover Network Inc.
Mountain View, CA 94041
arvind@clover.com

Jonathan Binney[‡]
Savioke
Menlo Park, CA 94025
binney@savioke.com

Thane Somers
Robotics Program
School of Mechanical, Industrial & Manufacturing Engineering
Oregon State University
Corvallis, OR 97331
somersth@oregonstate.edu

Gaurav S. Sukhatme[§]
Department of Computer Science
Viterbi School of Engineering
University of Southern California
Los Angeles, CA 90007
gaurav@usc.edu

Abstract

Operating autonomous underwater vehicles (AUVs) near shore is challenging - heavy shipping traffic and other hazards threaten AUV safety at the surface, and strong ocean currents

[*]http://research.engr.oregonstate.edu/rdml/
[†]http://robotics.usc.edu/∼ampereir/
[‡]http://robotics.usc.edu/∼binney/
[§]http://robotics.usc.edu/∼gaurav/

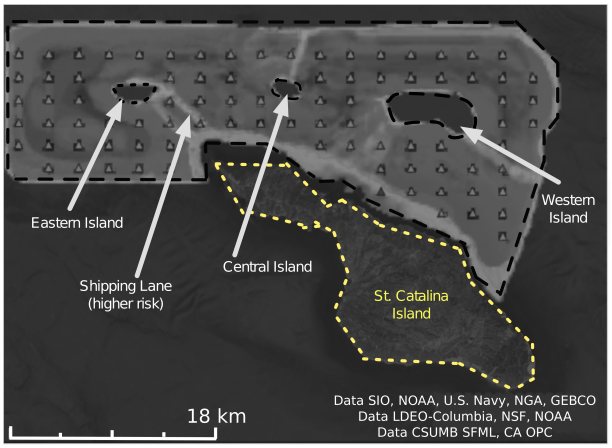
impede navigation when underwater. Predictive models of ocean currents have been shown to improve navigation accuracy, but these forecasts are typically noisy, making it challenging to use them effectively. Prior work has explored the use of probabilistic planners, such as Markov Decision Processes (MDPs), for planning in these scenarios, but prior methods have lacked a principled way of modeling the uncertainty in ocean model predictions, which limits applicability to cases where high fidelity models are available. To overcome this limitation, we propose using Gaussian processes (GPs) augmented with interpolation variance to provide confidence measures on predictions. This paper describes two novel planners that incorporate these confidence measures - (1) a *stationary risk-aware GPMDP* (for low-variability currents) and (2) a *non-stationary risk-aware NS-GPMDP* (for faster and high-variability currents). Extensive simulations indicate that the learned confidence measures allow for safe and reliable operation with uncertain ocean current models. Field tests of the planners on Slocum gliders over several weeks in the ocean demonstrate the practical efficacy of our approach.

1 Introduction

Buoyancy-driven Autonomous Underwater Vehicles (AUVs) like Slocum gliders (see Figure 1 (a)), with proven endurance of several weeks, are a natural choice for persistent monitoring of the ocean (Smith et al., 2010b). Such slow-moving AUVs operating in coastal regions are often at risk of colliding with navigational hazards, such as ships (Merckelbach, 2012), or of being swept away toward land by strong currents. Deterministic planning techniques ignoring ocean currents (Pereira et al., 2013) often perform poorly in practice due to significant errors in navigation while the glider dead-reckons underwater. Corrections for averaged current fields experienced by gliders over long dives rarely help improve navigational accuracy in future dives, due to significant variability in currents both spatially and temporally.

While it is possible to measure ocean currents directly on gliders, the required power restricts possible applications (Woithe et al., 2011). In the absence of measurements, predictive models of ocean currents can be used to compute paths for gliders through current fields (Thompson et al., 2010; Smith et al., 2010a). Predictions of ocean currents utilize forward simulations of prior data, and they incorporate large-scale measurements from satellite and HF-radar. We focus on the Regional Ocean Monitoring Systems (ROMS) predictions of ocean currents (Shchepetkin and McWilliams, 2005), which are now available for much of

(a) Slocum glider at surface near Catalina island



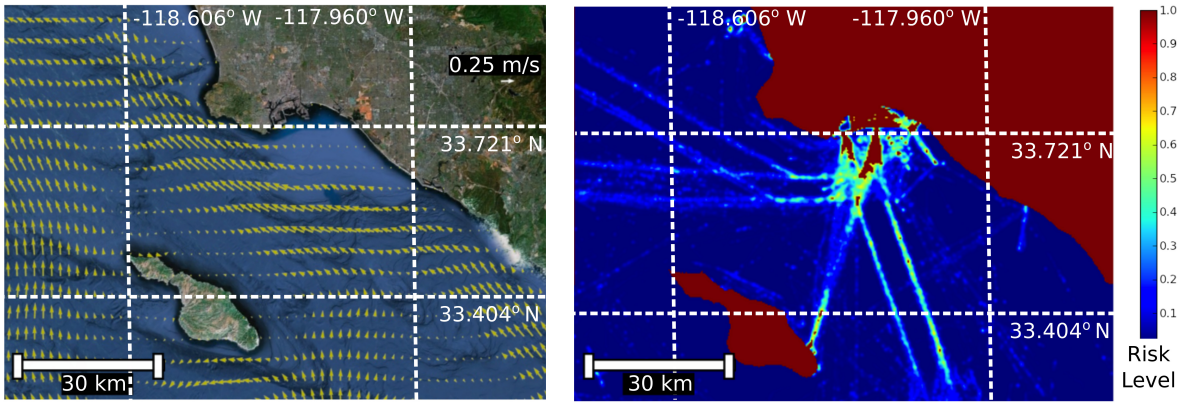
(b) Virtual map used to field test planners

Figure 1: (a) Slocum gliders are valuable assets in coastal observatories. When at the surface, an inflatable air bladder helps raise the antennae in the tail approximately 30 cm above the water surface allowing the robot to communicate via Iridium or radio modem and receive GPS updates. A glider at the surface is at risk of collision with boats and ships due to its small profile. (b) A virtual island map built for the Southern California coastal ocean that shows the region we use in this paper to evaluate our planners.

the U.S. West Coast, Pacific Ocean, Gulf of Mexico, and Atlantic Ocean (`https://www.myroms.org/`). An example is shown in Figure 2a for the Southern California Bight region. These predictions are provided daily and give forecasts up to 48 hours in advance. Such predictions are inherently uncertain (Lermusiaux, 2006), and this uncertainty affects the navigation accuracy of gliders in the ocean (Smith et al., 2013).

When gliders surface to acquire a GPS fix and transmit their data, they are vulnerable to damage by passing ship traffic. This danger motivates the development of risk-aware planning techniques that avoid areas with high shipping activity. Data from the automatic identification system (AIS) can be used to estimate shipping traffic in an area. In the example in Figure 2, the traffic going between the mainland and the island to the south consists mainly of diurnal tourist traffic to and from St. Catalina island. The shipping traffic along the primary shipping lanes near the mainland which enter/leave the ports of Los Angeles and Long beach do not show much variation between day and night.

In our prior work, we proposed algorithms capable of planning risk-aware paths by using predictive ocean current models (Pereira et al., 2013). These techniques required creating fixed transition models using stochastic simulations of a glider model in a field of predicted ocean currents. The goal of the planners was to minimize the expected risk of surfacing in regions where the glider might collide with ships and other navigational hazards given noisy predictions of the currents. A limitation of prior work was the lack of a principled way to estimate the uncertainty in these predictions (when not provided by the model), which



(a) Ocean currents predictions for Southern California

(b) Risk map for Southern California

Figure 2: (a) Ocean currents in the Southern California Bight region predicted by Regional Ocean Modeling Systems (ROMS) (Shchepetkin and McWilliams, 2005). (b) A collision risk map built using automatic identification system (AIS) data for shipping traffic (indicated by the warmer lines), and thresholding on bathymetry data for identifying land (red color) and navigation hazards indicated by warmer colors.

limited applicability to cases where high fidelity ocean current models were available.

As a result of the lack of uncertainty estimation, prior algorithms needed to act conservatively to avoid catastrophic failures. In the present work, we remove this limitation by utilizing learning methods to predict the uncertainty of the ocean currents, and we incorporate these estimates into probabilistic planners in a principally distinct way. We employ a data-driven technique where we use Gaussian processes augmented with interpolation variance to provide confidence measures about the time-series predictions from the Regional Ocean Modeling System (ROMS).

This data-driven approach allows for estimation of the uncertainty we expect in different parts of the operational region. The benefit of learning uncertainty in ocean current models for planning on AUVs has not been previously studied in prior work, and our work fills this gap. A preliminary version of the proposed approach was first described in our prior conference paper (Hollinger et al., 2013). The present paper includes the introduction of a non-stationary variant of the planning algorithm as well as more extensive simulations and field trials comparing the proposed techniques.

The key contributions of this paper are (1) the design and evaluation in simulation of two probabilistic planners that incorporate model confidence, which build upon the Markov Decision Process (MDP) framework – namely a stationary GPMDP and a non-stationary NS-GPMDP, (2) field trials and simulated comparisons of these planners using different amounts of ocean model knowledge. An additional contribution of this work is a software architecture that includes (1) a real-time uncertainty estimate of regional ocean current models

using a probabilistic data-driven method and (2) a Markov Decision Process (MDP) based path planner, which are both tested on physical deployments of two gliders to perform long-distance travel in the ocean.

The remainder of this paper is organized as follows. We first discuss related work in AUV path planning (Section 2) and highlight the need for probabilistic planning and learning. We then present risk-aware planning MDPs (Section 3) and learning approaches for predictive ocean current models (Section 4). We next present simulations (Section 5) and experiments (Section 6) demonstrating the effectiveness of our proposed approach. Finally, we draw conclusions and discuss avenues for future work (Section 7).

2 Related Work

Optimal path planning is the process of generating a sequence of waypoints from a start configuration to a desired goal configuration under constraints and a metric (e.g., avoiding obstacles and requiring minimal time/energy). Several well-known path and motion planning algorithms have been discussed in (Latombe, 1991; LaValle, 2006). Most classical planning techniques in the artificial intelligence literature are designed for use on a discretized world, due to ease of implementation on computers. Heuristic search techniques, such as A^*, have been very popular for deterministic planning in the artificial intelligence and robotics literature. These techniques use heuristics based on domain knowledge to guide the search, preventing unnecessary expansions whenever possible. Variants of these techniques, such as D^* and D^*-lite, are capable of performing efficient updates to the costs for nodes in the search graph when new knowledge of the world becomes available in dynamic scenarios(Koenig and Likhachev, 2002; Stentz, 1995).

While discrete path planning algorithms have been popular since they are easy to implement, sampling-based motion planning algorithms, such as rapidly-exploring randomized trees (RRTs) and probabilistic roadmaps (PRMs)b have gained popularity as they can solve planning problems in high-dimensional continuous state spaces (LaValle and Kuffner, 2001; Ferguson and Stentz, 2006; Kobilarov et al., 2012; Kavraki et al., 1996). These algorithms are well suited to problems requiring plans online as robots execute motion. These planners typically work well for robots with accurate sensors and high control authority. In contrast, Slocum gliders have low control authority, since ocean currents are often faster than their nominal speed, and they lack the ability to observe the ocean currents directly. Thus, we focus on planning techniques better suited for situations with uncertainty in observation and control of the state of the robot: Markov Decision Processes (MDPs) (Bellman, 1957; Burnetas and Katehakis, 1995).

For a standard MDP, the transition models are assumed to be stationary and not to change over time. Certain cyclic fluctuations in environmental conditions may not be captured by a short horizon length (e.g., tidal fluctuations (Keulegan, 1967)). In order to properly model periodic fluctuations, it is important to develop planners that utilize non-stationary transition models capable of utilizing the variations in the ocean currents to their advantage. Prior work has included the application of semi-Markov Decision Processes to allow for more sophisticated loitering behavior (Das et al., 1999) as well as switch-state MDPs that allow for the representation of multiple uncertainty models within the same planning framework (Brunskill et al., 2010). We propose the development of non-stationary transition models, which can be used to solve for non-stationary MDP policies appropriate to ocean current models.

This work focusses on risk avoidance, a well studied problem in robotics, spanning from unmanned ground vehicles (Soltani et al., 2002) to unmanned aerial vehicles (De Filippis et al., 2011). Most of this work uses deterministic strategies for risk minimization using heuristic search techniques. Our prior paper (Pereira et al., 2013) described a deterministic planning algorithm which used A^* to plan risk-aware paths (without considering ocean currents). A recent paper (Merckelbach, 2012), studied the risk of collisions between glider AUVs and ships. The resulting collision probability model indicates that the probability of a collision is proportional to shipping density. We construct risk maps based on a similar model for planning low-risk paths for AUVs, while also considering how noisy ocean predictions affect our surfacing outcome. In (Grasso et al., 2010) a decision support system using automatic identification system (AIS) data and ocean current predictions determines the safety of a set of waypoints chosen by human pilots. We extend this by providing an algorithm that chooses waypoints so as to minimize the expected risk of surfacing in hazardous locations or collisions with land bodies.

Fast marching and level sets are techniques based on wavefront expansion that provide efficient solvers for the motion of AUVs through time-varying current fields. Fast marching (Sethian, 2001; Sethian and Vladimirsky, 2003) is a technique which solves the Eikonal equations of a front whose speed function cannot change sign. Fast marching has been successfully applied in the context of planning paths for AUVs in (Petres et al., 2007). A related technique, level sets (Osher and Sethian, 1988; Sethian, 2001), relaxes the constraint by allowing the speed function of the wavefront to change signs. Recently, (Lolla et al., 2012) used level sets to plan paths for AUVs in flow fields. However, these techniques do not take into account the uncertainty of the ocean current field. In contrast, we focus on algorithms that also reason about the noise in predictive ocean current models such as Markov Decision Processes.

Most work in path planning for AUVs relies on deterministic techniques (e.g., (Carroll et al., 1992)) where

bathymetry, exclusion zones, and ocean current databases help to generate path corridors along great circle routes. Witt and Dinbabin describe an algorithm designed to find energy-optimal paths through search by using time-varying ocean current predictions (Witt and Dunbabin, 2008). Kruger et al., developed planning algorithms for AUVs in fast-flowing estuarine environments (Kruger et al., 2007). All of these algorithms are unsuitable for slow-moving vehicles like Slocum gliders, which lack good control authority. When it comes to deterministic planning for gliders, Fernandez-Perdomo et al., developed an algorithm called constant time surfacing A^*, where predictive ocean current models are used to plan time-optimal paths (Fernández-Perdomo et al., 2010). Similar graph-based planners for various cost-functions are described in (Eichhorn, 2010); further extensions using parallelization can also be found in (Eichhorn and Kremer, 2011). All these planners require low noise in the ocean current predictions to be practically useful. The first attempt at studying the effect of noisy predictive ocean current models in path planning is found in (Thompson et al., 2010), where the performance of wavefront expansion based planners under uncertainty in current predictions was studied. Results from their work indicate that noise in predictions has a significant effect on the planning performance. Our own prior work (Pereira et al., 2013), was the first to develop planners capable of reasoning probabilistically about noisy current predictions, but lacked a principled way to determine noise values for building transition models. The present paper describes planners using transition models based on data-driven GPs to learn the noise model.

Predictions of ocean currents have previously been used by several research groups to improve the navigation capabilities of autonomous vehicles and improve the safety of their operation (Fernández-Perdomo et al., 2010; Pereira et al., 2013). Lower bounds on navigation error have also been derived to estimate the path following performance of underwater vehicles (Szwaykowska and Zhang, 2011). While these prior works provide a basis for navigation of AUVs in the open ocean, they are lacking analysis of the uncertainty of predictions. Therefore, we develop the necessary tools to utilize these uncertainty predictions and improve path planning methods.

Prior work has examined the development of confidence measures for various ocean processes using straightforward statistical tools (Willmott et al., 1985) as well as more sophisticated Bayesian models (Kerman et al., 2008). In (Lermusiaux, 2006), a comprehensive method for the quantification, prediction, and estimation of uncertainties of ocean dynamics is described, along with the challenges involved in modeling ocean fields in a very large scale. A rigorous computational method called error subspace statistical estimation (ESSE) is described with several illustrative examples for its use. ESSE has components of time-varying basis functions, multi-scale initializations, and stochastic ensemble predictions, which are combined with data-assimilation.

However, ESSE does not provide approximate governing equations for the time-evolving error covariance bases, and it is best suited for data-assimilation of complex processes.

GPs have been used successfully to improve the accuracy of tasks such as large-scale terrain modeling (Vasudevan et al., 2009), scientific planetary surveying (Thompson and Wettergreen, 2008), and robotic grasping (Dragiev et al., 2011). Interpolation variance provides an alternative measure of uncertainty that also incorporates variability of the prediction into the variance. Interpolation variance was originally introduced for the estimation of geostatistical data (Yamamoto, 2000; Yamamoto and da Rocha, 2008) and was recently applied as a method for guiding adaptive sampling by mobile robots (Kim et al., 2012). To our knowledge, we are the first to apply interpolation variance to the prediction of ocean currents and the first to integrate it into the action model of a probabilistic planner in (Hollinger et al., 2013).

3 Risk-aware planning using MDPs and predictive ocean current models

Our work proposes using the Markov Decision Process (MDP) framework to design risk-aware trajectories for AUVs that incorporate confidence estimates of uncertainty in ocean currents. An MDP is described by a tuple (S, A, P, T, H, R), where S is the set of all possible states of the system $(s_1, s_2, ...)$, A is the set of all actions $(a_1, a_2, ...)$ a glider can take from each state (Mausam and Kolobov., 2012). The transition function $P : S \times A \times S \times D \rightarrow [0, 1]$, is a mapping specifying the probability $P(s_k|a_h, s_j, t_i)$ of arriving at state s_k given that the action a_h was executed at time t_i when the glider was in state s_j. T is a set of decision epochs or discrete time steps at which actions need to be taken $(0, 1, 2, ...T_{max})$. The reward function $R : S \times A \times S \times T$, provides a finite reward value $R(s_k|a_h, s_j, t_i)$ obtained when the system goes from state s_j to state s_k due to the execution of action a_h at time step t_i.

MDPs make the first-order Markov assumption, where mappings P and R depend only upon the state the system is currently in (and not its history). An MDP is well suited for a glider path planning problem because the state of the system is known (through GPS updates) when the glider is at the surface, but is highly uncertain when the glider is dead reckoning underwater. We provide rewards for the glider to get to the goal and penalize it proportional to the risk of the expected outcomes of performing actions from each state. States are assumed to have a linear additive utility function $U(R_t, R_{t+1}, ...) = \mathbf{E}[\sum_{t'=t}^{T_{max}} \gamma^{t'-t} R_{t'}]$, which is the expected sum of all the rewards (possibly discounted) that can be collected by the glider during

the epoch being considered.

A finite horizon MDP is one where the maximum number of time steps T_{max} is finite, and in such a case γ is typically assumed to be 1. An infinite horizon MDP is one where T_{max} is ∞, and in this case γ (the discount factor) is usually chosen to be some value between 0 and 1. A discount factor less than 1 makes a glider prefer rewards which can be obtained earlier. For these expected linear additive utilities, a policy is as good as the expected discounted rewards it will yield. The discount factor is also useful in ensuring convergence of value iteration. Setting $\gamma = 1$ makes the glider indifferent to when it will receive rewards. A policy π is a set of actions that the glider executes from being in each state. An optimal policy π^* is one whose value function $U^*(s)$ (called the optimal value function), dominates the value function for all other policies over all histories.

In (Pereira et al., 2013), we introduced the *risk-aware MDP* and described how it can be used to plan paths through ocean currents using predictive ocean current models and risk maps constructed using AIS data. Figure 1 (b) shows the planning environment which we use in this work both in evaluation of the planners (in simulation) as well as to conduct field experiments. This figure also shows how we have discretized the map using a regular grid (resolution approximately 2 km) to obtain states s used throughout this paper. The transition model P relies upon this discrete grid of states, with possible actions being the choice of moving from one state to any of its 8-connected neighbors. This transition function $P(s''|s, a(s, s'))$ describes the probability of ending up in state s'', given we choose to take action a (state s to state s').

In order to estimate the transition model P, a large number of simulations (> 30) using a standard kinematic model (see (Graver, 2005), (Bhatta, 2006), (Pereira et al., 2013)) are performed, between every pair of states. For the work reported in this paper we use depth-averaged currents for all simulations. The currents used for each simulation are perturbed with additive Gaussian noise as shown in Equations 1 and 2 to account for measurement noise. We note that if we have information about the noise in the ocean current predictions, we can develop a more accurate model of the perturbations. The distribution of surfacing locations for each set of trials gives us an estimate of the transition function for that action. This allows us to estimate the transition function $P(s''|s, a(s, s'))$ describing the probability of ending up in state s'', given we choose to perform action a (attempting to go from state s to state s').

When performing simulations for generating the transition model, we choose a representative noise for the currents either with a fixed standard deviation (as a baseline) or with a standard deviation learned using Gaussian process regression (see Section 4). The following equations describe how the simulated currents

are obtained when simulated noise is drawn from a Gaussian distribution $\mathcal{N}(0, \sigma^2)$. For easting (u_{pred}) and northing (v_{pred}) components of the ocean currents in the ROMS model, we have:

$$u_{sim}(x, y, t) = u_{pred}(x, y, t) + \sigma_u \tag{1}$$

$$v_{sim}(x, y, t) = v_{pred}(x, y, t) + \sigma_v \tag{2}$$

Now we turn our attention to generating the reward function R. The reward function is a mapping from an action a taken at a state s. The expected risk of performing this action, which is given by the average of the risk at the final surfacing location for each simulation. This risk is obtained from risk maps, such as the one shown in Figure 2. In Equation 3, $R(s_i, a_k)$ is the reward for performing action a_k from state s_i, f_j is the final surfacing location of the jth simulation involving glider movement from location s_i under action a_k. N is the total number of simulations. Risk is negative because it is a cost - higher risk corresponds to lower reward.

$$R(s_i, a_k) = -\frac{\sum_{j=1}^{N} risk(f_j|s_i, a_k)}{N} \tag{3}$$

There are several well known algorithms which can be used to solve MDPs (see (Mausam and Kolobov., 2012) for an extensive list). Goal locations carry a one-time positive reward and are terminal states. We employ the value iteration algorithm described in (Russell and Norvig, 2003) to solve our MDPs. This involves solving the Bellman backups corresponding to Equation 4. Here $V^*(s)$ is the maximum expected discounted utility which can be collected by the glider from state s with a discount factor γ and it is called a Bellman backup at state s.

$$V^*(s) = \max_{a \in A} \left[\sum_{s' \in S} P(s'|a, s)[R(s'|a, s) + \gamma V^*(s')] \right] \tag{4}$$

Value iteration consists of iteratively computing $V^*(s)$ for all states s, until $V^*(s')$ converges with the lefthand side equal to the righthand side as expressed in Bellman equation (Bellman, 1957). In this work we use a value of $\gamma = 1$, although in general one might choose a discounting factor < 1 to ensure that value

iteration converges. Once value iteration has converged, the optimal policy π^* for each state s is

$$\pi^*(s) = \underset{a \in A}{argmax} \left[\sum_{s' \in S} P(s'|a,s)[R(s'|a,s) + \gamma V^*(s')] \right]. \tag{5}$$

The policy $\pi^*(s)$ informs the glider of the action a it should take which will give it the maximum utility (which in most cases takes it to the goal so as to collect the one-time positive reward). Notice that it is also possible to index the transition models, rewards, optimal value function and optimal policy by the decision epoch. This creates a non-stationary model.

The MDP is solved to obtain the optimal policy corresponding to the noise in the transitions. These policies are used during execution of the plan by the glider. A drawback of using a fixed noise value is that the currents have uniform noise throughout the planning graph. This is a limiting assumption, which we overcome by using the Gaussian process framework to estimate noise values using historical current data. We also provide extensions to non-stationary models within this framework.

Figure 3 depicts the general procedure used to produce policies from ocean current predictions and the risk map. We begin with historical ROMs data and predictions for the day we are interested in. We estimate the interpolation variance (see Section 4) associated with the ocean current predictions for the day indexed by location and time. These variance estimates are used to create transition models which can be stationary (applicable over 6, 12 or 24 hour periods), or non-stationary (models varying with time). The transition models are used to solve the MDP using classical methods like value iteration. The risk map is used to provide costs for the reward function of the MDP. Once solved, we obtain a policy which provides the action the glider must execute to get the maximum utility given the glider's current state.

4 Learning uncertainty in ocean current models

Rather than using a fixed noise value for the ocean currents, we can use machine learning methods to develop a principled estimate of uncertainty for predictions of large-scale ocean processes. We propose augmenting available predictions through the use of Gaussian process learning methods. In this paper, we focus on the prediction of ocean surface currents, due to the availability of accurate HF-Radar measurements for these values.

The ocean currents at a given latitude lat, longitude lon, and time t are represented by a vector $\mathbf{x}_i = (u, v)$, where u and v are the components of the currents along the latitude and longitude axes respectively. The

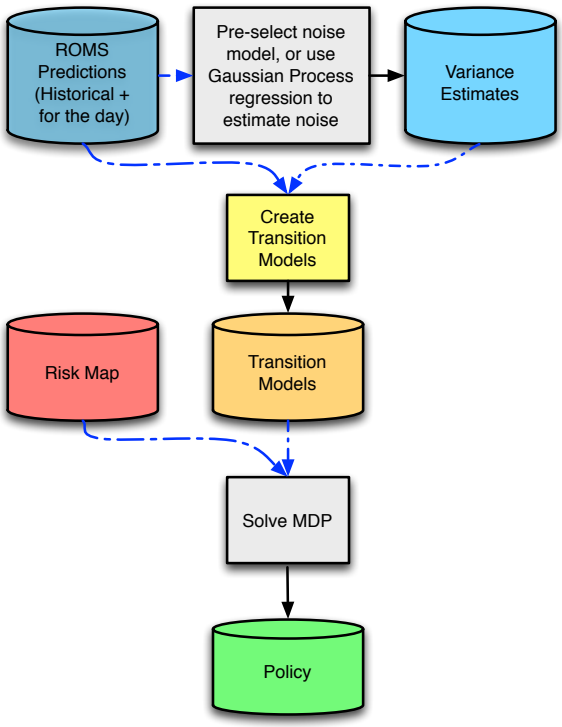


Figure 3: General procedure used to produce policies for autonomous underwater vehicles from ocean current predictions and a risk map of shipping activities.

positions are discretized based on the latitude and longitude, and time is discretized into hours. At a given time T, we are given historical data for times $t = \{T-1, T-2, \ldots\}$ back several months or years. We are also given predictions from the ROMS (Shchepetkin and McWilliams, 2005) data for $t = \{T+1, T+2, \ldots, T+48\}$ (two days in the future). Given this data, we want to provide better predictions for the points of time in the future as well as confidence bounds for these predictions.

4.1 Gaussian process regression

We propose modeling ocean currents using non-parametric Bayesian regression in the form of Gaussian processes (GPs) (Rasmussen and Williams, 2006). A GP models a noisy process $z_i = f(\mathbf{x}_i) + \varepsilon$, where $z_i \in \mathbb{R}$, $\mathbf{x}_i \in \mathbb{R}^d$, and ε is Gaussian noise. In the case of ocean current modeling, the function f represents a mapping from the latitude, longitude, and time values for a data point $\mathbf{x}_i$ to the u or v component of the ocean currents (i.e., the horizontal and vertical components). This data point may be a hindcast (past), nowcast (present), or forecast estimate. We note that the Gaussian measurement noise ε should not be confused with the uncertainty in the functional relationship f, which will later be estimated using the interpolation variance. In addition, the functional relationship f may be stationary (i.e., non-varying with

respect to space and time) or non-stationary. We examine both cases in our experiments and simulations.

We are given some data of the form $D = [(\mathbf{x}_1, z_1), (\mathbf{x}_2, z_2), \ldots, (\mathbf{x}_n, z_n)]$, where $\mathbf{x}_i$ represents a vector of latitude, longitude, and time values for a data point i, and z_i represents either the u or v component of the currents at that point and time. This formulation decouples the prediction of u and v, an assumption that could be relaxed using existing methods (e.g., using the techniques in (Álvarez and Lawrence, 2011)). We refer to the $d \times n$ matrix of $\mathbf{x}_i$ vectors as $\mathbf{X}$ and the vector of z_i values as $\mathbf{z}$.

To fully define a GP, we must choose a covariance function that relates the points in $\mathbf{X}$ to each other. We expect that data points close to each other in time and space will have high correlations, which provides a smoothing effect on the data. In addition, there is a periodic effect that arises due to tidal processes, which creates a correlation between points separated by 12 hours, 24 hours, etc. To capture both the spatial correlations and the periodic correlations, we apply the following kernel based on the squared exponential:

$$\begin{aligned} k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp[& - w_{lat}(\mathbf{x}_i^{lat} - \mathbf{x}_j^{lat})^2 \\ & - w_{lon}(\mathbf{x}_i^{lon} - \mathbf{x}_j^{lon})^2 \\ & - w_t(\mathbf{x}_i^t - \mathbf{x}_j^t)^2 \\ & - w_p \sin^2(\pi(\mathbf{x}_i^t - \mathbf{x}_j^t)/12)]. \end{aligned} \tag{6}$$

The hyperparameter σ_f^2 represents the process noise, and the hyperparameters w_{lat}, w_{lon}, w_t, and w_p represent weighting for the latitude, longitude, time, and periodic correlations. Having defined the kernel, combining the covariance values for all points into an $n \times n$ matrix $\mathbf{K}$ and adding a Gaussian observation noise hyperparameter σ_n^2 yields $\mathbf{K}_z = \mathbf{K} + \sigma_n^2\mathbf{I}$, where $\mathbf{I}$ is the $n \times n$ identity matrix.

The following equations predicts the mean function value (u or v value) $\mu(\mathbf{x}_*)$ and variance $\mathbb{V}_{gp}(\mathbf{x}_*)$ at a selected point $\mathbf{x}_*$ given the historical and prediction training data:

$$\mu(\mathbf{x}_*) = \mathbf{k}_*^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{z}, \tag{7}$$

$$\mathbb{V}_{gp}(\mathbf{x}_*) = k(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}_*^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{k}_*, \tag{8}$$

where $\mathbf{k}_*$ is the covariance vector between the selected point $\mathbf{x}_*$ and the training inputs $\mathbf{X}$. This model

gives a mean and variance for a particular latitude, longitude, and future time. An important aspect of this model is the derivation of a variance, which provides a measure of confidence of each prediction based on the correlations of the points used to predict it. The predicted mean value acts as a smoothing operator on the data. We note that in this work, we are not claiming that the GP mean provides a more accurate prediction of the ocean currents, since the focus is on uncertainty estimation.

The Gaussian process variance described above gives an estimate of the uncertainty of a prediction based on the data sparsity around that point and the estimated hyperparameters. However, this estimate does not take into account variability of the ocean currents around that point, which is an important predictor of measurement confidence. To improve on the GP variance uncertainty predictions, we propose using a method based on the interpolation variance (Yamamoto, 2000; Yamamoto and da Rocha, 2008; Kim et al., 2012). Given that a GP has already been learned, and letting $\boldsymbol{\mu}$ be the vector of all $\mu(\mathbf{x}_*)$ values, the interpolation variance can be estimated as:

$$\mathbb{V}_{iv}(\mathbf{x}_*) = \mathbf{k}_*^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}(\mathbf{z} - \boldsymbol{\mu})^T(\mathbf{z} - \boldsymbol{\mu}), \tag{9}$$

The interpolation variance incorporates correlations between the data learned using the GP framework as well as variability of correlated data to determine the predicted variance. This estimate of variance provides a richer representation that supports both data sparsity and data variability. The GP correlation matrix $\mathbf{K}$ is used as part of the computation of the interpolation variance. In order to calculate this matrix, it is necessary to learn the hyperparameters, which effectively provide the weighting between near points for calculating the interpolation variance. We use the GP framework to estimate these hyperparameters and to produce a principled correlation matrix. The interpolation variance then calculates how far the measured ROMS values are from the GP estimate, which provides a measure of uncertainty that we later show correlates to error in the ROMS predictions.

4.2 Local approximation and block learning

The modeling technique proposed above can be used for small datasets. However, ocean current models can correspond to long periods of time and a large portion of space. The computation time required by the GP scales $O(n^3)$ in the number of data points, which makes it infeasible for any large dataset more than approximately 2000 points. As an alternative, we propose estimating the predictions using a subset of the

data corresponding to the points we expect to be most correlated. To choose these points, we store the data in a KD-tree, where the relative weighting of space and time are adjusted to fit the values in the kernel (e.g., the periodic correlations are preserved). Such techniques have been applied successfully for terrain modeling (Vasudevan et al., 2009).

The estimation of the kernel hyperparameters $\theta = (\sigma_f, \sigma_n, w_{lat}, w_{lon}, w_t, w_p)$ in large-scale Gaussian processes is also a challenging task. The standard method is maximizing the likelihood of the measurements given the data and the hyperparameters (Rasmussen and Williams, 2006):

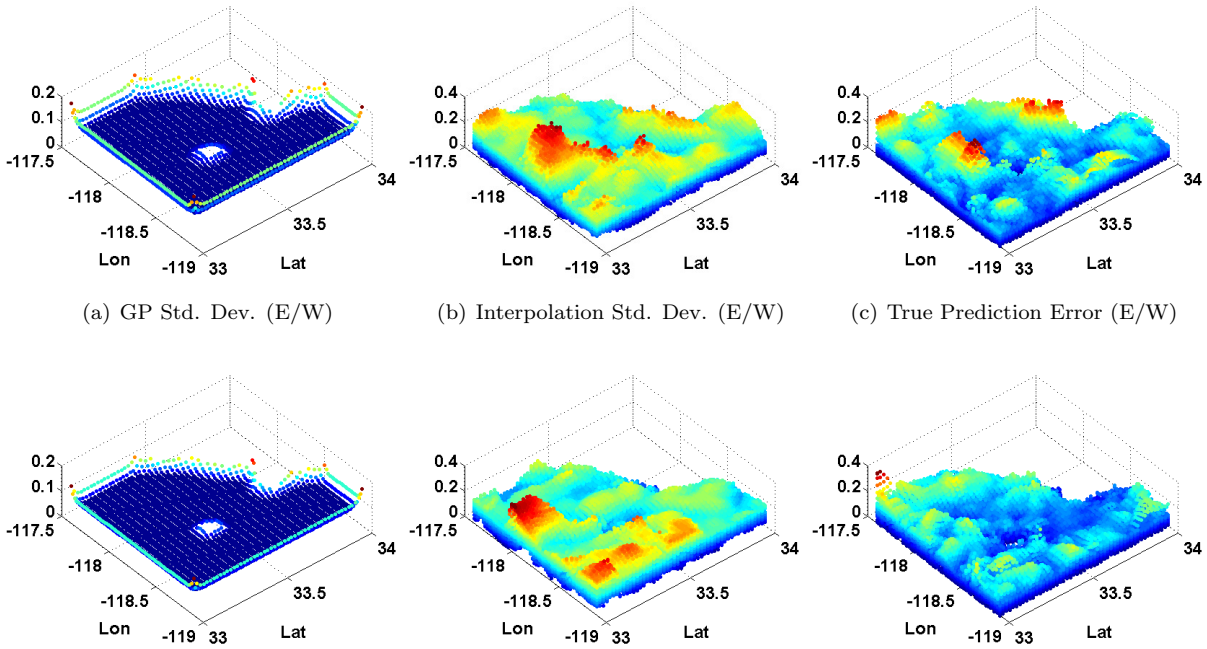
$$\log p(\mathbf{z}|\mathbf{X}, \theta) = -\frac{1}{2}\mathbf{z}^T\mathbf{K}_z^{-1}\mathbf{z} - \frac{1}{2}\log|\mathbf{K}_z| - \frac{n}{2}\log 2\pi, \tag{10}$$

This likelihood can be maximized using conjugate gradient optimization for small and moderately sized datasets. For the large ROMS dataset, we use a block learning approach that learns hyperparameters for subsets of the data. The block learning approach divides the data into equally sized segments and independently learns the hyperparameters for each segment. The final hyperparameter values are then found by averaging these local hyperparameter values. We use this method to learn separate hyperparameters for each day.

4.3 Comparison of uncertainty predictions

The data processing was performed on a single desktop with a 3.2 GHz Intel i7 processor with 9 GB of RAM. The kernel hyperparameters were learned independently for each day by dividing the training data from each day into 500 measurement blocks and then estimating the hyperparameters for each block using conjugate gradient optimization. The final hyperparameter values for a day were then found by averaging over all blocks. Using this block learning approximation and the KD-tree inference approximation, the GP took approximately 6 minutes to compute all predictions for a single day. These results use 100 points in the space/time KD-tree for the local GP prediction at each point. There are a total of 2560 locations estimated over a 24 hour period for predictions up to two days ahead (total of 123,880 point predictions per day). The data points are available at 1 hour intervals with 2 km resolution. The proposed method is highly parallelizable, since the hyperparameter estimation of each block is independent, and the prediction for each point can be performed independently once the hyperparameters are learned and the KD-tree is built.

Our methods provide confidence measures based on the underlying variance within the GP. In the formulation



(a) GP Std. Dev. (E/W) (b) Interpolation Std. Dev. (E/W) (c) True Prediction Error (E/W)

(d) GP Std. Dev. (N/S) (e) Interpolation Std. Dev. (N/S) (f) True Prediction Error (N/S)

Figure 4: Comparison of the Gaussian process variance and the interpolation variance measures of uncertainty for the prediction of ocean currents in the Southern California Bight on August 7, 2012. The interpolation variance accounts for variability in the ocean currents when making its prediction of uncertainty, and it shows a stronger correlation with the true prediction error.

above, the Gaussian process variance prediction provides an estimate of error due to data sparsity and measurement noise. Since the ROMS data are available at a uniform resolution across time (hourly) and space (2 km resolution), this value provides a limited measure of confidence that is fairly homogeneous. With a constant data resolution, the GP predicts low uncertainty throughout the space with high uncertainty on the edges, where the data density sharply decreases. This is clearly not a useful measure of uncertainty for predicting the error in the estimates of the ocean currents.

The alternative method based on the interpolation variance incorporates variability in the surrounding currents as a component of the uncertainty. Figure 4 shows example variance maps for a day in August, 2012 as well as the true prediction error (compared to the next day's nowcast, which would not have been available during planning). The results demonstrate that the interpolation variance provides a much richer measure of uncertainty that qualitatively matches with the true prediction error.

We also compare the correlation coefficients (R-values) between the interpolation variance and the Gaussian process variance correlated with the true prediction error. If the variance is an accurate representation of uncertainty, we would expect to see a positive correlation between the variance and the true error. Ta-

ble 1 shows the R-values for three months in 2012. The interpolation variance shows a positive correlation with the true prediction error, while the Gaussian process variance shows essentially no correlation.[1] We note that these correlations values are still fairly low, but we show in the following section that even this weak correlation is sufficient to provide improved performance when planning the trajectories of underwater vehicles.

Table 1: Correlation coefficients (R-values) for Gaussian process variance and interpolation variance relative to true prediction error (shown separately for N/S component and E/W components of the ocean current vectors)

Month (2012)	*June*	*July*	*August*
GP variance R-value (E/W)	-0.0519	-0.0178	-0.0277
Interp. variance R-value (E/W)	0.1383	0.1260	0.1655
GP variance R-value (N/S)	-0.0652	-0.0245	-0.0271
Interp. variance R-value (N/S)	0.1745	0.1400	-0.1323

4.4 Risk-aware Gaussian process MDP using stationary transition models

The methods described above can be incorporated into path planners for AUVs to improve the safety and reliability of operation. Path planners that incorporate data from ocean currents clearly stand to benefit from improvements in the accuracy of ocean current predictions. In addition, confidence estimates are useful both for planners that reason probabilistically as well as for planners that reason about worst-case instances. We now examine how the proposed GP prediction methods could be useful for improving the paths generated by such planners.

We first propose the risk-aware Gaussian process MDP (GPMDP) planner, which uses the maximum interpolation variance among the states s and s' of each transition edge, $e(s, s')$, over every time interval $t \in [t_1, t_2]$ to generate the transition model for the MDP. To compute the maximum variance for the generation of the transition model for a particular transition edge, we use Equations 11 and 12. Each transition model is then computed by conducting (> 30) simulations of transitions between each pair of states where the u and v components of the ocean currents used for the simulation are drawn from Graussians $\mathcal{N}(u, \sigma_u)$ and $\mathcal{N}(v, \sigma_v)$. The standard deviations σ_u and σ_v for the easting and northing current components are found using Equations 1 and 2 respectively:

$$\sigma_u^2 = \max_{t \in [t_1, t_2]} \left\{ \sigma_u^2(s, t), \sigma_u^2(s', t) \right\}, \tag{11}$$

[1] The slight negative correlation for the GP variance is likely due to the high predicted variances in the bordering edges of the dataset, which do not correspond to increased errors in the prediction.

$$\sigma_v^2 = \underset{t \in [t_1, t_2]}{max} \left\{ \sigma_v^2(s,t), \sigma_v^2(s',t) \right\}. \tag{12}$$

With improved confidence measures describing the amount of noise in the ocean current predictions, the transition models used by the GPMDP are more representative of the true errors than the models using a constant prediction noise value over the entire planning graph. Thus, we would expect the plans generated with the GPMDP to provide improvements in the execution speed, reliability, and safety versus methods using a fixed uncertainty.

4.5 Risk-aware Gaussian process MDP using non-stationary transition models

Stationary models are useful when there is low variability in the ocean currents. When planning paths through regions with high variability in currents, a natural option is to use non-stationary transition models during planning. To do this, we make use of the interpolation variance estimates learned using Gaussian process regression on the predictions, and we use these models in a risk-aware non-stationary GPMDP (NS-GPMDP). The transition model P in this planner uses a discrete grid of both states and time, with possible actions defined as moving from one state to any of its 8-connected neighbors at some time. We perform a number of simulations for every action performed for each state $s \in S$ at time t, where S is the set of all states in our planning region. This allows us to estimate the transition function $P(s''|s, a(s,s'), t)$ describing probability of surfacing in state s'' when performing that action $a(s,t)$.

The reward function is still computed using Equation 3. The optimal value function is given by

$$V^*(s,t) = \underset{a \in A}{max} \left[\sum_{s' \in S} P(s,a,s',t)[R(s,a,s',t) + \gamma V^*(s', t+1)] \right], \tag{13}$$

where $V^*(s,t)$ is the maximum expected utility the glider can collect at state s and time t. The optimal policy π^* corresponding to this optimal value function is deterministic Markovian and is given by

$$\pi^*(s,t) = \underset{a \in A}{argmax} \left[\sum_{s' \in S} P(s,a,s',t)[R(s,a,s',t) + \gamma V^*(s', t+1)] \right]. \tag{14}$$

The NS-GPMDP uses transition models up until the maximum lookahead time from the predictions of ocean currents. Due to this, when the start and goal are sufficiently separated, the NS-GPMDP can be susceptible to computing a policy which makes no progress toward the goal, but in fact chooses to hover in a low-risk

region in the vicinity of this node. This is reasonable because the glider prefers to minimize the risk of colliding with obstacles it may encounter while making toward the goal if it is aware that it cannot make it to the goal (where it may get a high goal reward). We observe in Section 5.2 (both in simulation as well as field tests), that this results in policies which timeout instead of taking the glider to the goal. However, these policies can be considered to "safe" policies in that they avoid risky regions and way for conditions to improve rather than risking collision.

One limitation of this approach is that the Markov assumption restricts modeling dependence between neighboring locations. In many cases, ocean currents in neighboring locations are strongly correlated, which is not captured by the Markov property. However, we note that correlations in uncertainties of neighboring locations are still captured through the interpolation variance. Since this variance is used to find the transition probability within the MDP, coupling is captured in that way.

5 Planner simulations

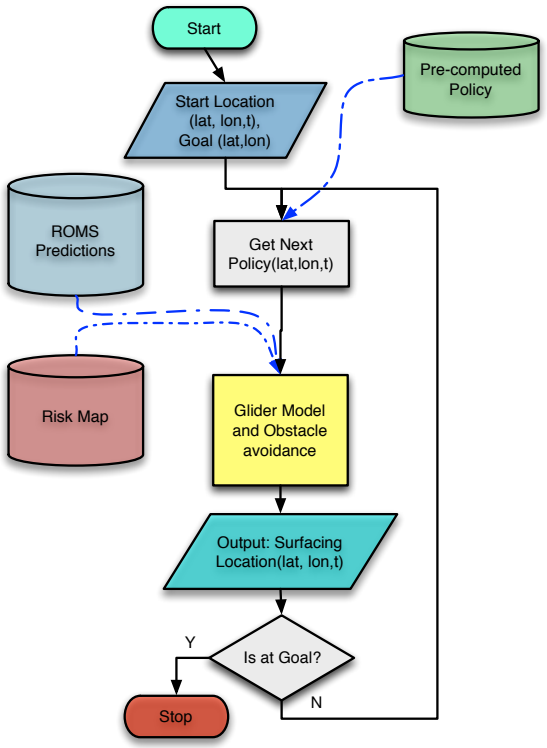


Figure 5: Basic flowchart explaining how simulations of the glider are performed for both the stationary and non-stationary planners.

We evaluate our planners in simulation using the virtual map shown in Figure 1 (b). A disadvantage of using virtual land bodies (e.g., the three islands) is that ocean currents may not behave exactly as modeled[2], especially when the true currents flow through an island (a situation which is unlikely with a real island).

Simulations are performed by computing policies for one of each of these 6 goal locations: (33.4498°, −118.3593°), (33.5430°, −118.3316°), (33.4692°, −118.4181°), (33.5238°, −118.5881°), (33.4882°, −118.7014°) and (33.5223°, −118.7572°). When performing simulations with currents, we first compute the average direction of the currents during the simulation window (typically a week). Then for each pair of locations we compare the direction from the start to the goal with the average direction of the currents. If these directions are found to be within ±45° of each other, we consider moving between this ordered pair of locations to be "with the currents". If on the contrary, the directions are found to have a difference greater than ±135°, we consider movement to be "against the currents". By varying the locations of the gliders in simulation, we provide insight into the applicability of our approach in different ocean current configurations.

In simulation a glider model and ROMS ocean current predictions are used to simulate the action of the gliders motion at sea. Policies (stationary or non-stationary) are pre-computed and indexed by date and time ranges, allowing for easy reference when a glider is at the surface in the field. During simulation, the tuple of surfacing latitude, longitude, and time of surfacing are used to look up the best action to be taken from this state given the policy corresponding to that time. The policy provides a new latitude and longitude for the glider to aim for, which is then executed using the glider model until the glider is at the goal or aborts the mission due to a collision. Surfacing risk is accumulated at each surfacing location (with the assumption that time at surface is equal at every waypoint). A simulation run ends either in success (glider reaches goal), abort (glider nears land) or in a time-out (when the simulation has not ended in more than 250 hours).

Figure 5 shows the procedure followed when performing simulations of the gliders as well as during field tests of the planners with gliders at sea. The pre-computed policy for the non-stationary MDP is sensitive to the time, while the stationary MDP policy chooses a finite-horizon policy corresponding to the 6, 12 or 24 hour period for which the stationary policy has been created. In the simulations reported in this paper, we use policies valid for 12 hour periods. When the glider is at sea during field experiments, the same procedure holds; we use the latest pre-computed policy and look up the best action based on the state consisting of the latitude, longitude (and in the case of non-stationary planners, time). This action is translated to a glider

[2]A higher number of aborts due to potential collisions with land happen than we would expect with more realistic ocean currents going around these virtual islands.

Table 2: Fisher p-values for simulated comparisons between risk-aware planning methods on AUVs in fast ocean currents.

GPMDP vs. SAMDP	
With currents	0.7448 (not significant)
Against currents	0.01771 (significant)
NS-GPMDP vs. GPMDP	
With currents	0.8901 (not significant)
Against currents	$6.5 * 10^{-8}$ (significant)
Omniscient MDP vs. NS-GPMDP	
With currents	0.4776 (not significant)
Against currents	1.0 (significant, alternate hypothesis)

mission file, which is uploaded to and executed by the glider.

Simulation results are presented where we compare the GPMDP with transition models guided by noise estimates learned by Gaussian processes (see Section 5.1) to a fixed noise state-action MDP (SAMDP) policy that uses a single noise estimate across the entire map. The SAMDP is similar to the algorithms proposed in our prior work that do not account for uncertainty in the ocean current models (Pereira et al., 2013). Further simulations, presented in Section 5.2, are performed to compare the performance of the stationary GPMDP with a non-stationary NS-GPMDP (which is designed for use when currents have higher variability).

Finally, Section 5.3 presents simulation results of an omniscient non-stationary MDP planner that has access to all ocean predictions for a full week of data in the future at the start of planning. Such information is not available from existing models and is only possible when running simulations from historical data. Examining the omniscient MDP gives us an idea of how well the NS-GPMDP could do if it had access to longer-term estimates of the ocean currents. For all simulated comparisons, we perform Fisher's exact test to generate a p-value and determine if the reduction of collisions is statistically significant (p-value less than 0.1). All simulation results are shown in Figures 6 and 7. Tables 2 and 3 give summaries of all p-values for the simulations.

5.1 Comparison of stationary risk-aware MDP (SAMDP) with the stationary risk-aware GPMDP

In the state-action risk-aware MDP (SAMDP) we make the assumption that noise in the ocean currents is a constant fixed value. In contrast, the GPMDP uses historical data to derive a stationary transition model which encodes the uncertainty in the ocean predictions within it. Figures 6 and 7 show comparisons of these

Table 3: Fisher p-values for simulated comparisons between risk-aware planning methods on AUVs in slow ocean currents.

GPMDP vs. SAMDP	
With currents	0.7045 (not significant)
Against currents	0.3733 (not significant)
NS-GPMDP vs. GPMDP	
With currents	1.0 (significant, alternate hypothesis)
Against currents	0.985 (significant, alternate hypothesis)
Omniscient MDP vs. NS-GPMDP	
With currents	$1.35 * 10^{-11}$ (significant)
Against currents	0.1541 (not significant)

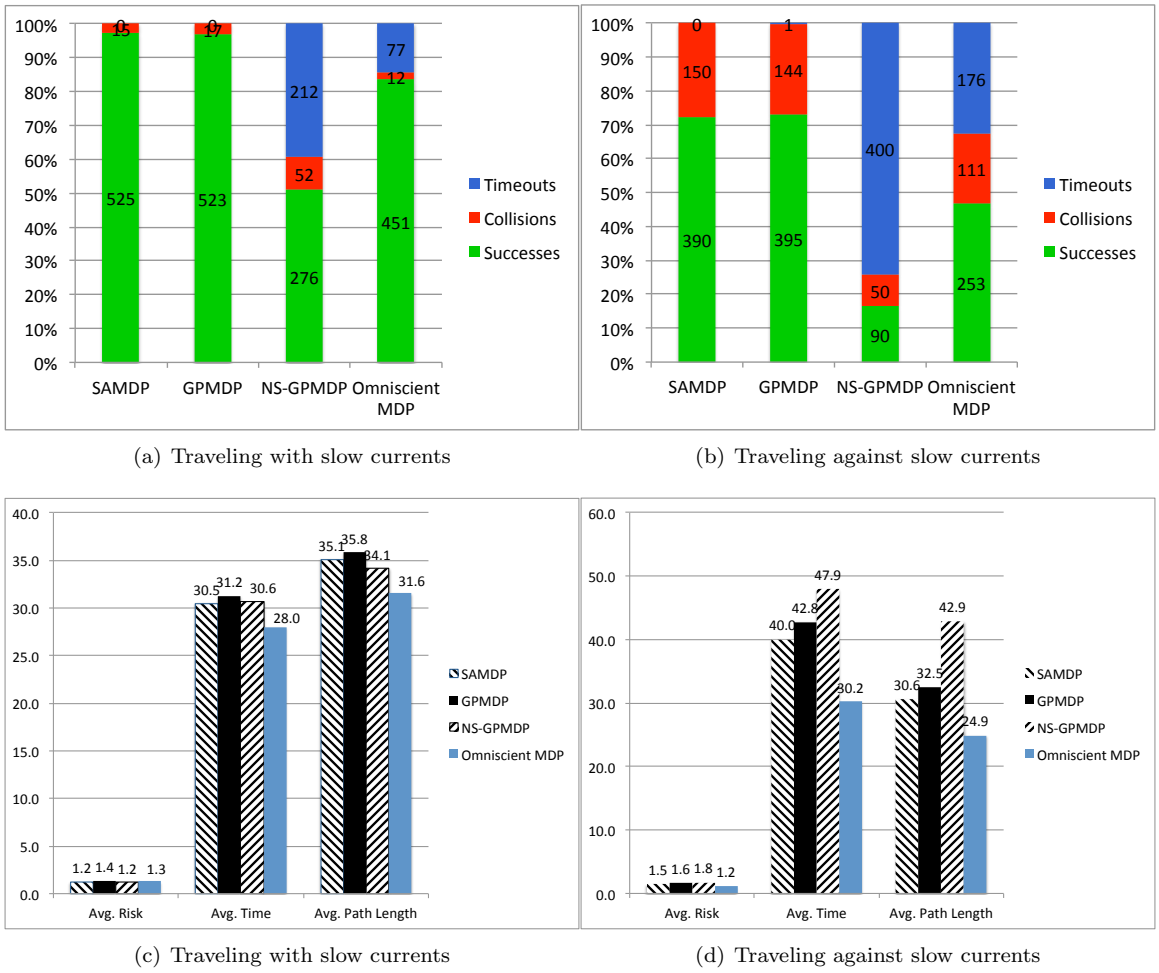
techniques for cases where the currents are slower than speed of glider (Aug 4 to Aug 11, 2012) and when currents are faster than speed of glider (Aug 17 to Aug 24, 2012). The fast currents exceeded the nominal speed of the glider (0.3 m/s) in a number of instances.

In the case where the glider is traveling with the currents we do not observe a significant change in collision rate between the SAMDP and GPMDP, with (p-value=.70) in slower currents and (p-value=0.74) in faster currents. We also do not observe any improvement through the use of the GPMDP when going against slow currents (p-value=0.37), but there is a significant improvement in reducing collisions when going against fast currents (p-value=0.018).

These results from simulation indicate that the confidence estimates in the GPMDP are likely to improve the performance of probabilistic planners when traveling against the ocean currents. When considering the results of simulations against fast currents, we notice that both of the stationary planners have a large number of collisions with land, which is a highly undesirable behavior. The interpolation variance estimated by the GPMDP uses a non-stationary kernel which provides estimates of the uncertainty in the current noise over time. The NS-GPMDP, which was described in Section 4.5, provides a means of reducing the collisions by preventing the planners from being overly goal-directed when conditions are risky.

5.2 Comparison of the non-stationary risk-aware GPMDP and the stationary risk-aware MDP

We now compare the non-stationary risk-aware GPMDP to the stationary risk-aware MDP. The results in Figures 6 and 7 show simulations using the stationary GPMDP and the non-stationary NS-GPMDP. When going with currents, we find that the NS-GPMDP does not significantly reduce the number of collisions as compared to the number of successfully executed paths. The GPMDP has success rates in excess of 86%,



(a) Traveling with slow currents

(b) Traveling against slow currents

(c) Traveling with slow currents

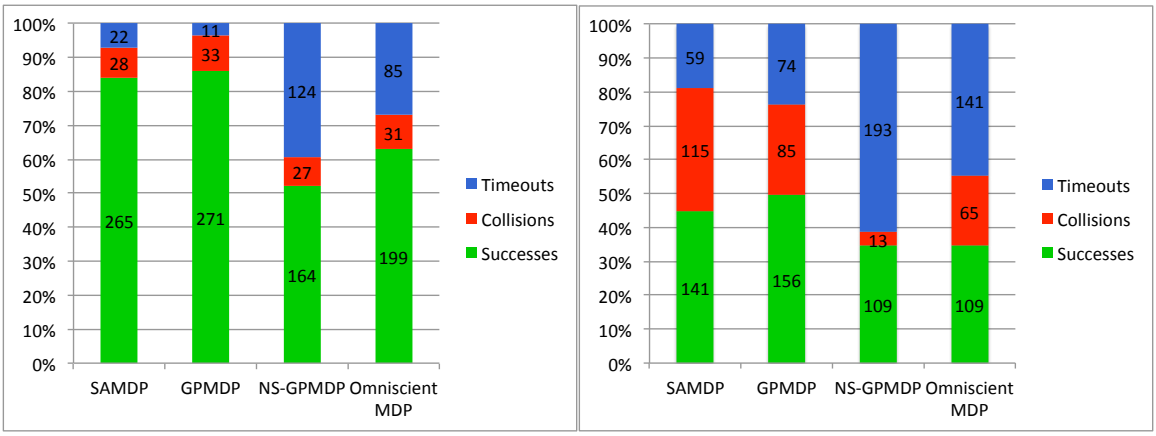
(d) Traveling against slow currents

Figure 6: Simulations comparing a fixe noise state-action MDP (SAMDP), a learned noise stationary MDP (GPMDP), a learned noise non-stationary MDP (NS-GPMDP), and learned noise omniscient non-stationary MDP with a week of future predictions (Omniscient MDP) for an autonomous underwater glider traveling both with and against slow ocean currents.

while the NS-GPMDP is successful only 50% of the time, choosing to time out instead of reaching the goal.

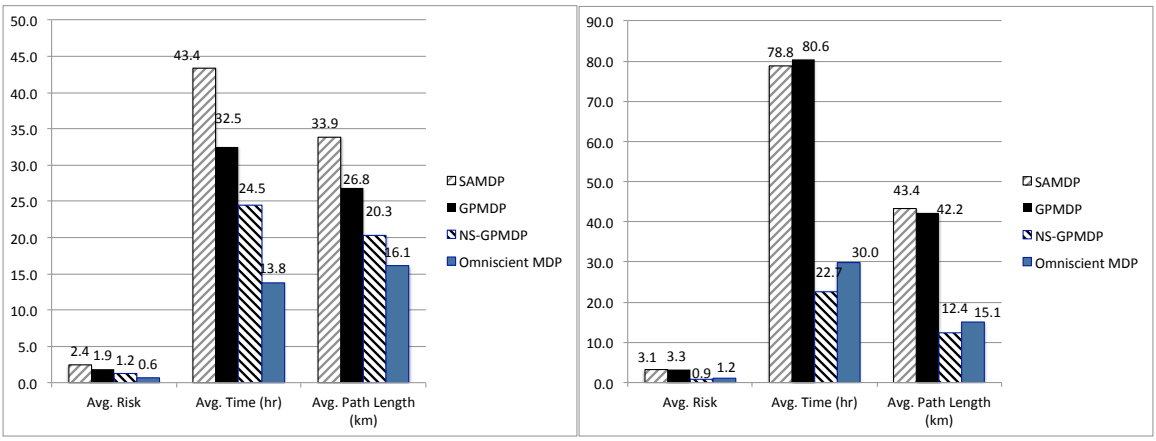
For the case where the currents go against the direction the glider, we notice that the NS-GPMDP has a much larger number of timeouts. In addition, the NS-GPMDP has significantly lower collisions than the stationary GPMDP for fast currents (p-value=$6.5 * 10^{-8}$) but higher collisions for slow currents (p-value=0.98). When traveling with the currents, we do not see a significant reduction in the number of collisions for fast currents (p-value=0.89), and we see a significant increase in the number of collisions for slow currents (p-value=1.0).

Here, the NS-GPMDP prefers not to make progress toward the goal in most cases because the short horizon for planning (approximately 48 hours) might be too short a time for the glider to make it successfully to the goal. This is quite evident in Figure 10(b) where the most of the states in the western portion of the planning



(a) Traveling with fast currents

(b) Traveling against fast currents



(c) Traveling with fast currents

(d) Traveling against fast currents

Figure 7: Simulations comparing a fixe noise state-action MDP (SAMDP), a learned noise stationary MDP (GPMDP), a learned noise non-stationary MDP (NS-GPMDP), and a learned noise non-stationary MDP with a week of future predictions (Omniscient MDP) for an autonomous underwater glider traveling both with and against fast ocean currents.

region do not have actions to take the glider toward the goal. States closer to the goal have actions going to the goal in the case where the glider goes against the currents. The limited duration of ocean current predictions prevents the NS-GPMDP from having better performance, except in the case where the goal is easily reachable from the state where the glider surfaces.

5.3 Gain of using an Omniscient MDP over a practical finite-horizon non-stationary MDP

In Section 5.2, we found that the NS-GPMDP often computes policies that make little progress toward the goal due to incomplete knowledge of the ocean currents. Although the knowledge about currents in the future is limited in real-world scenarios, we can develop a non-stationary MDP planner for simulated

comparison which concatenates together a week of ocean current data. We call this planner the Omniscient MDP because it has substantially more knowledge about the currents during the time of planning than is practically possible when operating in the real world without knowledge of the future.

Figure 6 shows results from a total of 4320 simulations conducted in relatively slow currents from August 12 to August 22, 2013. Here we notice that the performance of the stationary planners is preferable to that of the non-stationary planners, indicating that if the currents are always slower than the speed of the glider, it is better to use a stationary planner such as the state-action MDP or the stationary GPMDP. In this case, the conservative behavior of the non-stationary model does not provide benefit because the glider is able to overpower the slow moving ocean currents even with poor estimates.

When we compare the performance of the NS-GPMDP and the omniscient MDP, we notice that in slower currents the Omniscient MDP has fewer timeouts as compared to the NS-GPMDP. This is not surprising considering that the glider is likely to have a policy that can take it from the start to the goal within the week-long planning horizon. The Omniscient MDP significantly reduces the number of collisions versus the NS-GPMDP (p-value=$1.4 * 10^{-11}$) for the case when going with slow currents. However, when going against slow currents, the Omnsicient MDP performs somewhat worse than the NS-GPMDP (p-value=0.15).

Now, consider Figure 7, where the currents are relatively fast in comparison with the speed of the glider. In such cases, we have seen (as discussed in Section 5.2) that the NS-GPMDP can significantly reduces the number of collisions as compared to the stationary planners. This figure shows that providing additional knowledge about future currents does not yield a significant improvement over the performance of the NS-GPMDP in terms of reduction of crashes (p-value=0.48) when going with fast currents, though there is a slight reduction in the number of timeouts. When going against the fast currents, there was no reduction in collisions as compared to the NS-GPMDP (p-value=1). The number of timeouts decreased while the number of collisions increased.

These results indicate that for challenging regions like those shown in Figure 1 (b), strong currents often call for policies in which the glider should not attempt to make progress toward the goal. In such cases, the best policy appears to be to wait the faster currents out in a safe area, and attempt to move to the goal only when current speeds reduce. In these simulations, the currents stayed strong throughout the runs, and in such cases the NS-GPMDP's policies were similar to those of the Omniscient MDP in trying to avoid risky areas and remaining in safer low-risk regions.

Figures 6 and 7 (c) and (d) show statistics of the planners' average risk, average time, and average path

length during the simulations. These statistics are the averages calculated for the successfully executed paths. The stationary MDP planners have higher risk associated with them even though they have more successfully executed paths. This is because even though the NS-GPMDP and the Omniscient MDP timeout, they do so by oscillating between low-risk regions, accruing very little risk. We notice that the paths successfully executed by the non-stationary planner are, on average, shorter than those successfully executed by the stationary GPMDP, particularly when going against the currents. When going with the currents, the omniscient MDP further reduces the path length over the NS-GPMDP planner.

In some some cases the plan will only succeed about 50% of the time, even with the most reliable method. We note, however, operating autonomous gliders is inherently a dangerous practice (Merckelbach, 2012). The willingness of scientists to deploy gliders in such risky situations demonstrates the importance of collecting scientific information in the ocean. Our simulated results show that the proposed techniques improve safety, and these results are statistically significant in some cases. Our techniques also provide scientists with tools to determine "how risky" it is to deploy a glider on a given day. Ultimately the determination of whether the gliders should be deployed depends on the scientist's risk/reward requirements. Our tools provide them with more information and an algorithm for improving safety over the current state of practice.

6 Field trials

We report results from field trials of the stationary version of the GPMDP as well as the non-stationary NS-GPMDP. The experiments from the stationary GPMDP planner were performed in August, 2012, while the NS-GPMDP experiments were performed in August, 2013. Even though the GPMDP and NS-GPMDP trials were performed on different dates, sea state conditions in August, 2012 and August, 2013 had prevailing currents moving in similar directions with similar magnitudes. In both cases, we obtain ROMS data at 06:30 GMT and compute transition models based upon the GP's estimates of interpolation variance. Next, we solve the MDP using the desired goal to pre-compute the MDP policy for each planner. When the AUV surfaces, we use its reported position to look up the best action to execute from the indexed policy corresponding to this surfacing location. We create a mission file for the glider with this state as the desired waypoint and then instruct the glider to execute it. At each subsequent waypoint we repeat this procedure until the glider reaches the goal.

Slocum gliders manufactured by Teledyne Webb Research communicate back to shore using a proprietary server called the glider *Dockserver*. Much of the process for field trials with the planners is automated

using a Python framework for glider path planning, which includes the ability to communicate with a glider Dockserver, as well as to transfer mission files to the glider through the Dockserver. The framework, which we call the Glider Path Planning Library (GPPLib), also includes methods to download the latest ROMs update, create transition models, archive and reload them as well as algorithms to solve MDPs and shortest paths in planning graphs.

In the field tests of the stationary GPMDP conducted in August, 2012, we observed that the GPMDP was able to reach the goal faster than the SAMDP method tested in prior work (Pereira et al., 2013) (see Table 4). This seems to indicate that the GPMDP was better informed and took less time to arrive at the goal as a consequence of planning a better path. Figure 8 depicts the execution of the GPMDP, which began approximately at 11:00 GMT on August 18, 2012. The GPMDP's confidence measures produce a less conservative policy which moved the glider more quickly toward the goal. The glider oscillated a few times before crossing the western shipping lanes, before making its way to the goal at approximately 10:00 GMT on August 20, 2012. This was nearly 12 hours faster than the glider running the SAMDP reported in prior work in the same operational area with similar prevailing ocean current conditions (Pereira et al., 2013).

Figure 8 shows the policy from the NS-GPMDP being executed by the glider. The NS-GPMDP policy did not involve oscillation (since the goal was reachable within 48 hours when going with the currents). The planner had to cross one set of shipping lanes, and it chose to cross the shipping lane to the east that required a longer path around the eastern and central virtual islands. Statistics for each field run, including duration and path length, are given in Table 4. The field experiments serve primarily as one-shot demonstrations of the stationary and non-stationary planning approaches running in an integrated system. This integration shows the proposed ocean current-based MDP planning working in the ocean with the predictive uncertainty estimates. Additional field trials would be necessary to show a rigorous comparison of the proposed methods, but such extensive trials are outside the scope of this paper, given the cost and resources necessary.

An example snapshot of the non-stationary policy used during the execution on August 15, 2013 is shown in Figure 9(a), and several simulated trajectories used to test whether the mission should be sent to the glider during the field tests is shown in Figure 9(b). During field trials, if these simulations end up causing collisions, the mission is not loaded for execution on the glider and instead requires a human pilot's supervision. This is a useful feature which can help glider pilots with a decision support framework during real glider deployments.

After the glider arrived at the goal, we chose a goal location in the east of the map (to cause the glider to travel against the currents). This experiment began at 01:20 GMT on August 17, 2013 and is shown in

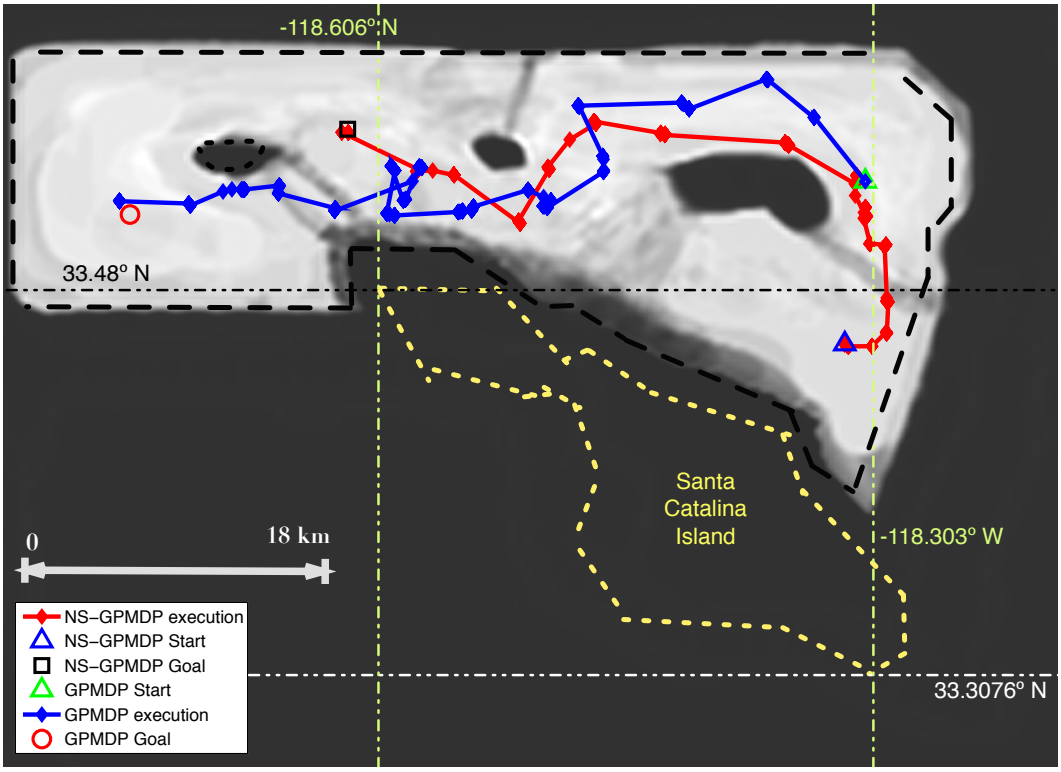
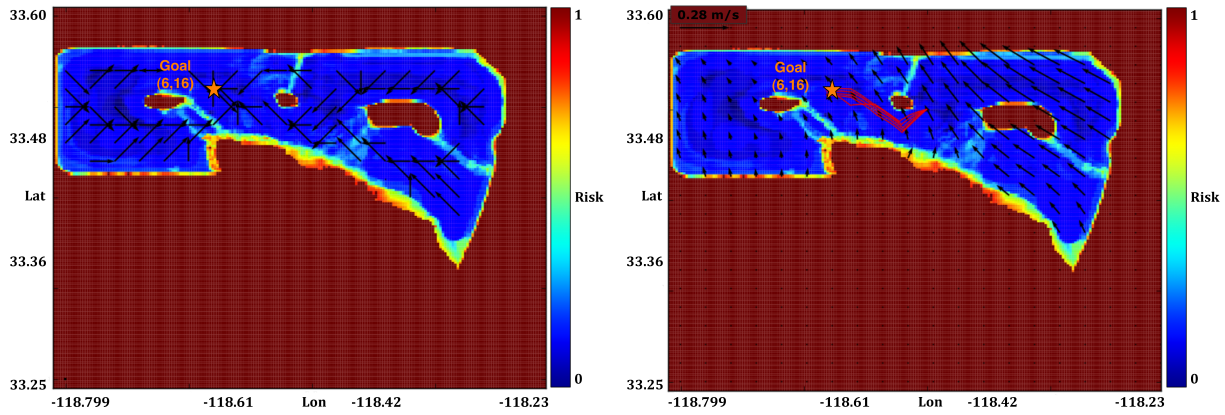


Figure 8: Field trials of stationary GPMDP and the non-stationary GPMDP overlaid for comparison. The field trial of the stationary GPMDP began at 10:59 GMT on August 18, 2012, at the location denoted by the blue triangle. The glider arrived at the goal denoted by the black square at 10:18 GMT on August 20, 2012. The path being executed by the GPMDP was generally with the direction of the currents. The field trial of the non-stationary GPMDP began at 07:33 GMT on August 14, 2013 and ended at 23:39 GMT on August 16, 2013.

Figure 10(a), where the glider started at the blue triangle. The goal location is depicted by the black square was too far for the glider to be able to realistically reach while traveling against the currents. We found that the glider made no progress toward the goal over the next 12 hours. The actions took the glider to low-risk regions of the map and caused it to oscillate there to ensure safety. Such conservative policies resulted in the NS-GPMDP timing out often in practice due to a lack of information about the ocean currents in the future. This is a practical limitation imposed upon the planner due to the limited time-horizon for which ocean current predictions are available.

Table 4: Statistics for the field trials performed with each algorithm.

Field Test	Start Time (GMT)	End Time (GMT)	Duration (HH:MM)	Path Length (km)
SAMDP (Pereira et al., 2013)	28-Jul-2012 06:12	30-Jul-2012 17:04	58:51	55.25
GPMDP	18-Aug-2012 10:59	20-Aug-2012 10:18	47:18	51.93
NS-GPMDP	14-Aug-2013 07:33	16-Aug-2013 23:39	64:05	43.82



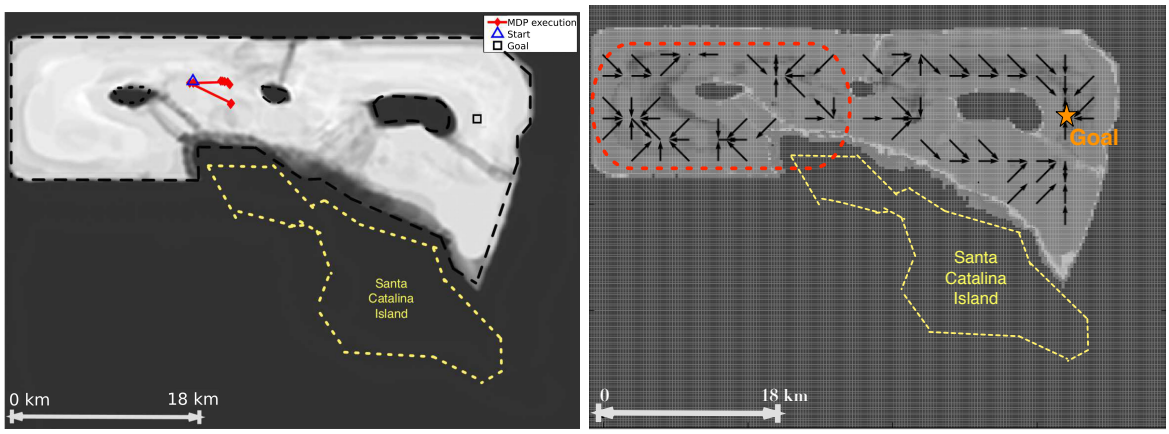
(a) Non-stationary policy on Aug 15, 2013

(b) Simulated trajectories to goal at same time

Figure 9: Screenshots from the Glider path planning library during the field trial on August 15, 2013. Red areas represent high risk of collision with land and passing ship traffic, while blue areas represent lower risk of collision. (a) NS-GPMDP policy computed when the glider surfaced at the location just south-east of the central island. (b) Forward simulations with varying amounts of noise in the ocean currents indicate that the glider is likely to successfully execute the computed policy from its current surfacing location.

Closer inspection of the NS-GPMDP policy revealed that the states shown in Figure 10(b) did not make progress toward the goal, while states outside this region made progress toward the goal. The policy shown in Figure 10(b) is a snapshot of the policy at 05:20 GMT on August 17, 2013. Notice how the actions are all risk-averse and take the glider to low-risk regions away from shipping lanes and land, where they oscillate back and forth between low-risk states. This is a natural solution to MDPs when the goal cannot be reached - they try to be risk-averse and find solutions which minimize the expected risk. Less informed planning strategies would likely choose to move through the uncertain ocean currents, risking damage or loss of the glider.

Successful field tests indicate that the proposed algorithms (along with optimizations, such as pre-computing and indexing policies) allow for probabilistic planners to be implemented and used on robots in the real world. Such automation in plan execution and decision making is essential to enable more complex data-driven planning for AUVs. Our system compensates for the lack of computational resources on-board the vehicle (as well as limitations, such as the large size of ROMs data and the transition models) by performing most of the computationally expensive work off-board and only uploading the final mission file to the vehicle for execution. Field trials of our planners resulted in execution of paths which were similar to those observed in simulation with currents from the same time period.



(a) Execution over 10 hours against currents

(b) Policy for NS-GPMDP against currents

Figure 10: (a) Execution of the NS-GPMDP on the Slocum glider. The experiment began at 01:20 GMT on August 17, 2013 at the location denoted by the blue triangle. The glider needed to travel against the currents, and chose to execute a conservative policy that did not make progress toward the goal. (b) Closer inspection of the policies generated revealed that none of the locations in the red region had goal-directed actions.

7 Conclusions

This paper addressed the problem of underwater vehicle navigation using noisy predictive ocean current models by developing planners based on the Markov Decision Process framework. The planners described in this paper utilize a data-driven approach for generating and using confidence measures in the uncertainty of noisy predictions during planning. The results demonstrate that learning confidence measures of the underlying ocean current models with a stationary Gaussian process model augmented with interpolation variance (GPMDP) allows for improved safety and reliability in slower ocean currents, and using a non-stationary model (NS-GPMDP) provides improvements with faster ocean currents. These planners provide a distinct way of dealing with noise in ocean current models versus the fixed noise state-action MDP (SAMDP) developed in prior work.

Extensive simulation results indicated a number of statistically significant gains over fixed noise MDPs by learning uncertainties and utilizing them during planning. We have observed improvements in simulation in the performance of the stationary GPMDP over the fixed-noise SAMDP, particularly when the speed of the currents are relatively slow in comparison with the nominal speed of the glider. We also observed that when the speed of the currents is substantially faster than the speed of the glider, the best solution is to utilize the NS-GPMDP. Table 5 summarizes the qualitative performance of planners when traveling with and against slow or fast currents based upon the results obtained from simulations. As ocean current models improve and provide predictions over longer horizons, the performance of our NS-GPMDP should resemble that of

the Omniscient MDP, which has access to information currently unavailable in existing predictions.

Table 5: Summary of planner performance. (✓ ⇒ good, × ⇒ poor). The stationary GPMDP consistently outperforms the other planners in slower currents, while the non-stationary NS-GPMDP outperforms the stationary planners in faster currents. If longer prediction horizons were available, the performance of the NS-GPMDP would be closer to that of the Omniscient MDP, which provided the most consistent performance across all types of currents.

		SAMDP	GPMDP	NSGPMDP	Omniscient MDP
Slow currents	With	✓✓	✓✓	×	✓
	Against	✓	✓✓	×	✓
Fast currents	With	✓	✓	✓	✓
	Against	××	×	✓✓	✓

The behaviors of the non-stationary planners in field tests were observed to be similar to those observed during simulation. We have tested our algorithms on gliders over several multi-day runs at sea and developed a semi-autonomous planning framework, which provides relative ease for automating risk-aware planning for underwater vehicles. The stationary GPMDP, which was tested in August 2012, outperformed the SAMDP in the field by reaching the goal safely approximately 10 hours earlier. The NS-GPMDP was also tested at sea in August 2013 and performed satisfactorily and in agreement with the simulated predictions. Our simulations and experimental results demonstrate that this approach (1) allows for more reliable trajectories, and (2) provides safe operation in environmental conditions in which it was previously unsafe to operate.

Future extensions of this work involve augmenting the reward functions of the GPMDP to allow the planners to automatically select between several goals using some multi-criteria model. Another avenue to investigate is the effect of changing the terminal rewards for the different planners. Our initial experience was that terminal rewards did not affect the trends in behavior. However, a more comprehensive evaluation in different domains would be necessary to make a strong claim. Another possibility for future extension would be to augment this work using a method based on Partially Observable Markov Decision Processes (POMDPs). The POMDP would address the effect of partial observation of information, such as ocean currents and shipping traffic, and provide actions accordingly. The key issue to address in this extension would be keeping the method scalable to large environments.

We would also like to investigate other data-driven approaches to modeling and estimating the uncertainty in general ocean models. Estimation of quantities like temperature, salinity, and chlorophyll content may benefit from the proposed data-driven approaches. Finally, we would like to test these techniques with more than two gliders to examine the scalability of our approach to multiple vehicles. Overall this paper

described two planners, each of which is useful in planning paths for gliders under uncertainty. Through the development and testing of these planners both in simulation and on real-world systems, our work has provided a suite of tools that allow operation of AUVs in ocean currents that would previously have yielded considerable risk of losing the vehicle.

Acknowledgments

The authors are grateful for the help from Matthew Ragan, Carl Oberg, Supreeth Subbaraya, Joerg Mueller, Christian Potthast, Bridget Seegers, Xiao Liu, Elizabeth Teel, Trevor Oudin, Gerry Smith, Gordon Boivin and everyone at the Wrigley Institute for Marine Sciences, for their assistance with field experiments. They helped at various times during deployments and glider recoveries. We thank Svetlana Vaz for assistance during the multi-day field experiments. The authors gratefully acknowledge Paul Reuter for providing us with the AIS data and Yi Chao, John Farrara and the rest of the team from JPL for providing us with ROMs data. We thank Sven Koenig, Burton Jones, David Caron and Ryan Smith for valuable discussions and insights during the course of this work. This research has been funded in part by the following grants: ONR N00014-14-1-0509, ONR N00014-09-1-0700, ONR N00014-07-1-00738, ONR N00014-06-10043, NSF IIS-1317815, NSF 0831728, NSF CCR-0120778, and NSF CNS-1035866.

References

Álvarez, M. A. and Lawrence, N. D. (2011). Computationally efficient convolved multiple output gaussian processes. *J. Machine Learning Research*, 12:1459–1500.

Bellman, R. E. (1957). *Dynamic Programming*. Princeton University Press, Princeton, NJ.

Bhatta, P. (2006). *Nonlinear Stability and Control of Gliding Vehicles*. PhD thesis, Princeton University, Dept. of Mechanical and Aerospace Engineering.

Brunskill, E., Kaelbling, L. P., Lozano-Perez, T., and Roy, N. (2010). Planning in partially observable switching-mode continuous domains. *Annals of Mathematics and Artificial Intelligence*, 58(3–4).

Burnetas, A. N. and Katehakis, M. N. (1995). Optimal adaptive policies for Markov Decision Processes. *Mathematics of Operations Research*, 22(1):222–255.

Carroll, K. P., McClaran, S. R., Nelson, E. L., Barnett, D. M., Friesen, D. K., and Williams, G. N. (1992). AUV path planning: An A* approach to path planning with consideration of variable vehicle speeds

and multiple, overlapping, time-dependent exclusion zones. In *Proc. Symp. Autonomous Underwater Vehicle Technology*, pages 79–84.

Das, T. K., Gosavi, A., Mahadevan, S., and Marchalleck, N. (1999). Solving semi-markov decision problems using average reward reinforcement learning. *Management Science*, 45(4):560–574.

De Filippis, L., Guglieri, G., and Quagliotti, F. (2011). A minimum risk approach for path planning of UAVs. *J. Intelligent Robot Systems*, 61:203–219.

Dragiev, S., Toussaint, M., and Gienger, M. (2011). Gaussian process implicit surfaces for shape estimation and grasping. In *Proc. IEEE Int. Conf. Robotics and Automation*, pages 2645–2850.

Eichhorn, M. (2010). Solutions for practice-oriented requirements for optimal path planning for the AUV SLOCUM glider. In *Proc. IEEE Oceans*, Seattle, WA.

Eichhorn, M. and Kremer, U. (2011). Opportunities to parallelize path planning algorithms for autonomous underwater vehicles. In *Proc. IEEE Oceans*, Kona, HI.

Ferguson, D. and Stentz, A. (2006). Anytime RRTs. In *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, pages 5369–5375.

Fernández-Perdomo, C., Cabrera-Gámez, J., Hernández-Sosa, D., Isern-González, J., Domínguez-Brito, A. C., Redondo, A., Coca, J., Ramos, A. G., Fanjul, E. A., and García, M. (2010). Path planning for gliders using Regional Ocean Models: Application of pinzón path planner with the ESEOAT model and the RU27 trans-Atlantic flight data. In *Proc. IEEE Oceans*, Seattle, WA.

Grasso, R., Cecchi, D., Cococcioni, M., Trees, C., Rixen, M., Alvarez, A., and Strode, C. (2010). Model based decision support for underwater glider operation monitoring. In *Proc. IEEE/MTS Oceans*, Seattle, WA.

Graver, J. G. (2005). *Underwater Gliders: Dynamics, Control and Design*. PhD thesis, Princeton University, Dept. of Mechanical and Aerospace Engineering.

Hollinger, G., Pereira, A., and Sukhatme, G. (2013). Learning uncertainty models for reliable operation of autonomous underwater vehicles. In *Proc. IEEE Int. Conf. Robotics and Automation*, pages 5573–5579.

Kavraki, L., Svestka, P., Latombe, J. C., and Overmars, M. (1996). Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Trans. Robotics and Automation*, 12(4):566–580.

Kerman, M. C., Jiang, W., Blumberg, A. F., and Buttrey, S. E. (2008). A comparison of robust metamodels for the uncertainty quantification (UQ) of New York Harbor oceanographic data. *J. Operational Oceanography*, 1(2):3–13.

Keulegan, G. H. (1967). Tidal flow in entrances; water-level fluctuations of basins in communication with seas. Technical Report 14, Committee on Tidal Hydraulics, Corps of Engineers, U.S. Army.

Kim, Y.-H., Shell, D., Ho, C., and Saripalli, S. (2012). Spatial interpolation for robotic sampling: Uncertainty with two models of variance. In *Proc. Int. Symp. Experimental Robotics*, Quebec City, Canada.

Kobilarov, M., Marsden, J. E., and Sukhatme, G. (2012). Global estimation in constrained environments. *Int. J. Robotics Research*, 31(1):24–41.

Koenig, S. and Likhachev, M. (2002). D* lite. In *Proc. AAAI*, pages 476–483.

Kruger, D., Stolkin, R., Blum, A., and Briganti, J. (2007). Optimal AUV path planning for extended missions in complex fast-flowing estuarine environments. In *Proc. IEEE Int. Conf. Robotics and Automation*, pages 4265–4270.

Latombe, J. (1991). *Robot Motion Planning*. Kluwer Academic Publishers, Boston, MA.

LaValle, S. M. (2006). *Planning Algorithms*. Cambridge University Press, Cambridge, U.K. Also available at http://planning.cs.uiuc.edu/.

LaValle, S. M. and Kuffner, J. J. (2001). *Algorithmic and Computational Robotics: New Directions*, chapter Rapidly-exploring random trees: Progress and prospects, pages 293–308. A K Peters, Wellesley, MA.

Lermusiaux, P. F. (2006). Uncertainty estimation and prediction for interdisciplinary ocean dynamics. *J. Computational Physics*, 217:176–199.

Lolla, T., Ueckermann, M. P., Yigit, K., Haley, P. J., and Lermusiaux, P. F. J. (2012). Path planning in time dependent flow fields using level set methods. In *Proc. IEEE Int. Conf. Robotics and Automation*, pages 166–173.

Mausam and Kolobov., A. (2012). *Planning with Markov Decision Processes: An AI Perspective*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan and Claypool Publishers.

Merckelbach, L. (2012). On the probability of underwater glider loss due to collision with a ship. *J. Marine Science and Technology*, 18:75–86.

Osher, S. and Sethian, J. A. (1988). Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Computational Physics*, 79:12–49.

Pereira, A., Binney, J., Hollinger, G., and Sukhatme, G. S. (2013). Risk-aware path planning for autonomous underwater vehicles using predictive ocean models. *J. Field Robotics*, 30(5):741–762.

Petres, C., Pilhas, Y., Patron, P., Petillot, Y., Evans, J., and Lane, D. (2007). Path Planning for Autonomous Underwater Vehicles. *IEEE Trans. Robotics*, 23(2):331–340.

Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. The MIT Press.

Russell, S. and Norvig, P. (2003). *Artificial Intelligence - A Modern Approach*. Prentice Hall.

Sethian, J. A. (2001). Evolution, implementation, and application of level set and fast marching methods for advancing fronts. *J. Computational Physics*, 169:503–555.

Sethian, J. A. and Vladimirsky, A. (2003). Ordered upwind methods for static Hamilton-Jacobi equations: Theory and algorithms. *SIAM J. Numerical Analysis*, 41(1):325–363.

Shchepetkin, A. F. and McWilliams, J. C. (2005). The Regional Oceanic Modeling System (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, 9(4):347–404.

Smith, R. N., Chao, Y., Li, P. P., Caron, D. A., Jones, B. H., and Sukhatme, G. S. (2010a). Planning and implementing trajectories for autonomous underwater vehicles to track evolving ocean processes based on predictions from a regional ocean model. *Int. J. Robotics Research*, 29(12):1475–1497.

Smith, R. N., Das, J., Heidarsson, H. K., Pereira, A. A., Arrichiello, F., Cetinic, I., Darjany, L., Garneau, M.-E., Howard, M. D., Oberg, C., Ragan, M., Seubert, E., Smith, E. C., Stauffer, B., Schnetzer, A., Toro-Farmer, G., Caron, D. A., Jones, B. H., and Sukhatme, G. S. (2010b). USC CINAPS builds bridges: Observing and monitoring the Southern California Bight. *IEEE Robotics and Automation Magazine*, 17(1):20–30.

Smith, R. N., Kelly, J., Nazarzadeh, K., and Sukhatme, G. S. (2013). An investigation on the accuracy of regional ocean models through field trials. In *Proc. IEEE Int. Conf. Robotics and Automation*, pages 3436–3442.

Soltani, A. R., Tawfik, H., Goulermas, J. Y., and Fernando, T. (2002). Path planning in construction sites: performance evaluation of the Dijkstra, A*, and GA search algorithms. *Advanced Engineering Informatics*, 16(4):291–303.

Stentz, A. (1995). The focussed D* algorithm for real-time re-planning. In *Proc. Int. Joint Conf. Artificial Intelligence*, pages 1652–1659.

Szwaykowska, K. and Zhang, F. (2011). A lower bound on navigation error for marine robots guided by ocean circulation models. In *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, pages 3584–3588.

Thompson, D. and Wettergreen, D. (2008). Intelligent maps for autonomous kilometer-scale science survey. In *Proc. Int. Symp. Artificial Intelligence, Robotics and Automation in Space*, Los Angeles, CA.

Thompson, D. R., Chien, S., Yi-Chao, Li, P. P., Cahill, B., Levin, J., Schofield, O., Balasuriya, A., Petillo, S., Arrott, M., and Meisinger, M. (2010). Spatiotemporal path planning in strong, dynamic, uncertain currents. In *Proc. IEEE Int. Conf. Robotics and Automation*, pages 4778–4783.

Vasudevan, S., Ramos, F. T., Nettleton, E. W., and Durrant-Whyte, H. F. (2009). Gaussian process modeling of large scale terrain. *J. Field Robotics*, 26(10):812–840.

Willmott, C. J., Davis, R. E., Feddema, J. J., Klink, K. M., Legates, D. R., Rowe, C. M., Ackleson, S. G., and O'Donnell, J. (1985). Statistics for the evaluation and comparison of models. *J. Geophysical Research*, 90(C12):8995–9005.

Witt, J. and Dunbabin, M. (2008). Go with the flow: Optimal AUV path planning in coastal environments. In *Proc. Australian Conf. Robotics and Automation*, pages 1–9.

Woithe, H. C., Boehm, D., and Kremer, U. (2011). Improving Slocum Glider Dead Reckoning Using a Doppler Velocity Log. In *Proc. IEEE Oceans*, Kona, HI.

Yamamoto, J. (2000). An alternative measure of the reliability of ordinary kriging estimates. *Mathematical Geology*, 32(4):489–509.

Yamamoto, J. and da Rocha, M. M. (2008). Properties and applications of the interpolation variance associated with ordinary kriging estimates. In *Proc. Int. Symp. Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, pages 70–77.