

Adaptive Sampling of Transient Environmental Phenomena with Autonomous Mobile Platforms

by

Victoria Lynn Preston

B.S., Olin College of Engineering (2016)

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of

Master of Science in Aeronautics and Astronautics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND
WOODS HOLE OCEANOGRAPHIC INSTITUTION

September 2019

© 2019 Massachusetts Institute of Technology and
Woods Hole Oceanographic Institution. All rights reserved.

Author
Department of Aeronautics and Astronautics
August 22, 2019

Certified by
Anna Michel
Associate Scientist of Applied Ocean Physics and Engineering, WHOI
Thesis Supervisor

Certified by
Nicholas Roy
Professor of Aeronautics and Astronautics, MIT
Thesis Supervisor

Accepted by
Sertac Karaman
Associate Professor of Aeronautics and Astronautics
Chair, Graduate Program Committee

Accepted by
David Ralston
Associate Scientist with Tenure of Applied Ocean Physics and Engineering
Chair, Joint Committee on Applied Ocean Science and Engineering

Adaptive Sampling of Transient Environmental Phenomena with Autonomous Mobile Platforms

by

Victoria Lynn Preston

Submitted to the Department of Aeronautics and Astronautics
on August 22, 2019, in partial fulfillment of the
requirements for the degree of
Master of Science in Aeronautics and Astronautics

Abstract

In the environmental and earth sciences, hypotheses about transient phenomena have been universally investigated by collecting physical sample materials and performing *ex situ* analysis. Although the gold standard, logistical challenges limit the overall efficacy: the number of samples are limited to what can be stored and transported, human experts must be able to safely access or directly observe the target site, and time in the field and subsequently the laboratory, increases overall campaign expense. As a result, the temporal detail and spatial diversity in the samples may fail to capture insightful structure of the phenomenon of interest.

The development of *in situ* instrumentation allows for near real-time analysis of physical phenomenon through observational strategies (e.g., optical), and in combination with unmanned mobile platforms, has considerably impacted field operations in the sciences. In practice, mobile platforms are either remotely operated or perform guided, supervised autonomous missions specified as navigation between human-selected waypoints. Missions like these are useful for gaining insight about a particular target site, but can be sample-sparse in scientifically valuable regions, particularly in complex or transient distributions. A skilled human expert and pilot can dynamically adjust mission trajectories based on sensor information. Encoding their insight onto a vehicle to enable adaptive sampling behaviors can broadly increase the utility of mobile platforms in the sciences.

This thesis presents three field campaigns conducted with a human-piloted marine surface vehicle, the ChemYak, to study the greenhouse gases methane (CH_4) and carbon dioxide (CO_2) in estuaries, rivers, and the open ocean. These studies illustrate the utility of mobile surface platforms for environmental research, and highlight key challenges of studying transient phenomenon. This thesis then formalizes the *maximum seek-and-sample* (MSS) adaptive sampling problem, which requires a mobile vehicle to efficiently find and densely sample from the most scientifically valuable region in an *a priori* unknown, dynamic environment. The PLUMES algorithm — **P**lume **L**ocalization under **U**ncertainty using **M**aximum-**V**alu**E** information and **S**earch — is subsequently presented, which addresses the MSS problem and overcomes

key technical challenges with planning in natural environments. Theoretical performance guarantees are derived for PLUMES, and empirical performance is demonstrated against canonical uniform search and state-of-the-art baselines in simulation and field trials.

Ultimately, this thesis examines the challenges of autonomous informative sampling in the environmental and earth sciences. In order to create useful systems that perform diverse scientific objectives in natural environments, approaches from robotics planning, field design, Bayesian optimization, machine learning, and the sciences must be drawn together. PLUMES captures the breadth and depth required to solve a specific objective within adaptive sampling, and this work as a whole highlights the potential for mobile technologies to perform intelligent autonomous science in the future.

Thesis Supervisor: Anna Michel

Title: Associate Scientist of Applied Ocean Physics and Engineering, WHOI

Thesis Supervisor: Nicholas Roy

Title: Professor of Aeronautics and Astronautics, MIT

Acknowledgments

Thanks first goes to my family and friends for their support and willingness to entertain hours of conversation about robot boats and marine methane. I especially wish to acknowledge my parents for the freedom they provided me throughout my life to explore my interests, ultimately enabling me to become the independent (and quirky) person I am today. I also sincerely thank my partner, Bill, for his enthusiasm about my work, ability to engage in thoughtful conversations about robotics, and his enduring ability to make me laugh when the going gets tough.

To my advisers, Anna Michel and Nicholas Roy, I extend sincere gratitude for their mentorship and insight throughout my work. I recognize the fortune I have to work with two incredible professionals in two unique fields, and the opportunities for field work and learning this has afforded. Additionally, working with their labs, WHOI DSL Michel Lab and MIT CSAIL RRG, has been a pleasure. Being surrounded by thoughtful and creative minds everyday is a privilege; and I'm especially thankful to count these individuals among my friends. I particularly thank my research collaborators Dr.Cara Manning, Dr.David Nicholson, Dr.Yogesh Girdhar, Genevieve Flaspohler, and Kevin Manganini for their insight, patience, support, and positive attitudes throughout our work together. They have taught me so much about how to be a good scientist, engineer, and communicator.

As I hit this milestone on my way towards earning a PhD, I look forward to what the future has in store. I would like to dedicate this thesis to my parents, Matthew and Elizabeth Preston, to whom I owe my intellectual curiosity, confidence, and perseverance.

Contents

1	Introduction	15
1.1	Overview of Field Standards in Environmental and Earth Sciences . . .	17
1.1.1	The “Gold Standard”	18
1.1.2	<i>In situ</i> Instrumentation and Unmanned Platforms	19
1.2	Maximum Seek-And-Sample	20
1.2.1	Representing Belief	21
1.2.2	Heuristic Reward	22
1.2.3	Decision-Making and Planning in Continuous Domains	23
1.3	Thesis Overview	25
2	Technical Background and Related Works	27
2.1	Measuring and Modeling Natural Phenomena	28
2.1.1	Bayesian Inference Techniques	29
2.1.2	Bayesian Representations	32
2.1.3	Gaussian Processes	33
2.2	Environmental Sensing as a Robotics Problem	35
2.2.1	Markov Decision Processes (MDPs)	35
2.2.2	Partially Observable MDPs (POMDPs)	36
2.3	Reward Specification	38
2.3.1	Overview of Information Measures	38
2.3.2	Rewards in Bayesian Optimization	40
2.4	Decision-Making under Uncertainty	43
2.4.1	Offline Planning	43

2.4.2	Online Planning	43
2.5	Robots in the Wild	46
3	Transience in Marine Science: Greenhouse Gas Emissions	49
3.1	Studying Methane and Carbon Dioxide	50
3.1.1	Carbon Dioxide	51
3.1.2	Methane	52
3.1.3	Impacts	52
3.2	The ChemYak Mobile Platform	54
3.2.1	Measurement Analysis	57
3.3	Wastewater Effluent in Tidal Estuaries	58
3.3.1	Overview of Field Work and Analysis	59
3.3.2	Results	60
3.3.3	Significance and Role of Transience	66
3.4	Methane Bubble Plumes in the Cascadia Margin	66
3.4.1	Overview of Field Work and Analysis	68
3.4.2	Results	69
3.4.3	Significance and Role of Transience	75
3.5	Spring Freshet River Inflow in Arctic Estuary	76
3.5.1	Overview of Field Work and Analysis	78
3.5.2	Results	80
3.5.3	Significance and Role of Transience	88
3.6	Sampling Transient Phenomenon and Motivating Adaptive Regimes	89
4	Adaptive Sampling for Transient Phenomenon	91
4.1	Maximum Seek-and-Sample POMDP Formalism	94
4.2	Representing Belief with Gaussian Processes	96
4.2.1	Kernels for Natural Phenomenon	96
4.3	Planning in Continuous State, Observation Spaces	100
4.3.1	Analysis of Continuous Observation MCTS	104
4.4	Maximum-value Information Heuristic Reward	106

4.5	Results in Static Environments	109
4.5.1	Evaluation	110
4.5.2	Bounded Convex Environments	111
4.5.3	Non-Convex Environments	115
4.6	Results in Dynamic Environments	119
4.6.1	Transience Rejection in Maximum-Value Cycling	119
4.6.2	Transience Incorporation	122
4.6.3	Target Tracking	125
4.7	Discussion	128
5	Conclusions	131
5.1	Thesis Contributions	132
5.1.1	Marine Sciences	132
5.1.2	Informative Path Planning and Adaptive Sampling	134
5.2	Future Work	135
5.2.1	Representing Scientific Phenomenon for Planning	135
5.2.2	Multi-Objective Missions	136
5.2.3	Longterm Monitoring	136
5.2.4	Multi-Agent Systems	137
5.3	Final Thoughts	137

List of Figures

1-1	Simple illustration of myopic and nonmyopic planning	24
3-1	The ChemYak	55
3-2	Wareham River: field site	61
3-3	Wareham River: ChemYak transects	62
3-4	Wareham River: salinity, temperature, and gas relationships	63
3-5	Wareham River: spatio-temporal trends	65
3-6	Cascadia Margin: distribution of corrected measurements	69
3-7	Cascadia Margin: Stonewall Bank ChemYak tracks	71
3-8	Cascadia Margin: Yachats ChemYak tracks	72
3-9	Cascadia Margin: surface methane and oxygen	73
3-10	Cascadia Margin: methane depth profiles	74
3-11	Cambridge Bay: reference map	77
3-12	Cambridge Bay: field site and summary ChemYak tracks	79
3-13	Cambridge Bay: calibration results for gas analyzer	80
3-14	Cambridge Bay: depth profiles	83
3-15	Cambridge Bay: spatio-temporal surface trends	84
3-16	Cambridge Bay: salinity, temperature, and gas relationships	85
3-17	Cambridge Bay: depth profiles binned by estuary region	86
3-18	Cambridge Bay: spatio-temporal trends of the pycnocline	87
4-1	GP with RBF kernels	97
4-2	GP with time-varying RBF kernel	98
4-3	GP with Periodic kernels	98

4-4	GP with compositional Periodic-RBF kernel	99
4-5	GP with dynamics model embedded in kernel	100
4-6	Illustration of continuous-observation Monte-Carlo Tree Search	101
4-7	Convergence of MVI vs UCB heuristic reward	109
4-8	Overview of simulation environments	110
4-9	Distribution of accumulated MSS reward in static convex simulations	112
4-10	Coral head localization mission overview	113
4-11	Coral head map and ASV	115
4-12	Distribution of accumulated MSS reward in static non-convex simulations	116
4-13	Snapshot of unknown non-convex map scenario	118
4-14	Robust performance under unknown transience	121
4-15	Emergent monitoring behavior with modeled transience	123
4-16	Comparison of PLUMES and UCB-MCTS in modeled transience	124
4-17	Trivial dynamic tracking case	125
4-18	Tracking an unknown moving source	127

List of Tables

4.1	Summary of metrics over static, convex simulation and field trials of PLUMES and baselines	112
4.2	Summary of metrics over static, non-convex simulation and car trials of PLUMES and baselines	116

Chapter 1

Introduction

Environmental and earth sciences are an undertaking to observe and explain natural planetary processes and attributes. An interdisciplinary combination of physical, biological, chemical, geological, and information sciences, these academic fields have provided considerable insight about the places in which human life and activity are closely intertwined. In order to investigate hypotheses or characterize new discoveries, the vast majority of these disciplines require either physically realized samples (e.g., rock specimens, water samples) or direct measurements of physical attributes (e.g., temperature, gaseous concentration).

Collecting these samples or taking direct physical measurements of environmental phenomena is a considerable logistical and scientific challenge. Sample collection requires that a human expert or technician can physically access the region of interest and interact with the phenomenon. However, in many environments, this is infeasible either due to hazards or remoteness of the region. Unmanned mobile platforms have increased the number of feasible study sites by explicitly going in place of human-occupied vessels. Contemporaneous development and improvement of *in situ* instrumentation has further improved the resolution, accuracy, and scope of observations that can be gathered. Mobile “observatories” combine both unmanned vehicles and *in situ* sensors to gather data at various deployment scales, including: global continuous surveying (e.g., passive marine floats [1]), local monitoring (e.g., wire profilers [2]), and targeted sample retrieval (e.g., ROV [3]).

Canonically, these mobile agents navigate passively, are piloted by a human operator, or autonomously traverse fixed, predetermined paths. This thesis presents three case studies in which a human-piloted marine surface vehicle was used to study the distribution of greenhouse gases in the upper water column of coastal rivers, estuaries, and near-shore seas. These case studies illustrate the fidelity to which mobile observatories can resolve different chemical events, and highlight a major drawback with canonical methods: data sparsity in regions of interest.

Chemical expressions are *transient phenomena*, which display dynamic spatio-temporal behavior with temporal variation on the order of hours to days. When the robotic mission is relatively short compared to the temporal variation of the phenomenon, recovering the spatial distribution of the phenomenon dominates. In human-piloted missions, either uniform-coverage of the region is performed or reactive trajectories are executed. In both strategies, the number of samples of a target phenomenon (e.g., high concentrations of a gas species), may be relatively low if the volume of the expression is small relative to the target region, or time is spent following misleading chemical signals. When the robotic mission is on a similar timescale to the temporal variation of the phenomenon, then both spatial and temporal aspects must be recovered, further complicating these missions.

In order to densely sample interesting, dynamic phenomena in a natural environment, intelligent, adaptive strategies are necessary to respond to stimulus. *Adaptive sampling* is an autonomy technique which uses a history of observations to inform navigation goals and mobile agent behavior to optimize a high-level scientific objective. For principled sampling in spatio-temporal systems, adaptive sampling regimes can perform inference over a robot's belief of a phenomenon based on historical observations, and strategically navigate to intercept interesting events. Key questions with respect to the design and performance of an adaptive algorithm for the study of transient phenomena are:

- How can scientific objectives be encoded as an adaptive sampling problem for robotic systems?

- How can transient phenomena at multiple temporal scales be modeled in order to inform robotic decisions?
- How should scientific value be quantified in an unknown environment?
- In the face of uncountably infinite possible states of an unknown environment, what performance guarantees can an autonomous framework provide with respect to valuable sample collection?

This thesis presents the PLUMES algorithm, which addresses these questions and overcomes core technical challenges to allow a mobile agent to intelligently seek and densely sample from the most scientifically valuable region in an unknown natural environment. The rest of this chapter is structured as follows: Section 1.1 discusses the practical challenges of sampling and collecting observations in natural environments, further motivating autonomous unmanned vehicles and introducing key vocabulary. Section 1.2 presents the motivating adaptive sampling problem of this thesis, the *maximum seek-and-sample* (MSS) problem, and its core technical challenges. Section 1.3 summarizes the contributions of this thesis and previews the remaining chapters.

1.1 Overview of Field Standards in Environmental and Earth Sciences

Within the environmental and earth sciences, sample collection and physical measurements are largely taken on *field campaigns*, which are dedicated trips to a target region. Specialized equipment to observe, measure, or capture a phenomenon of interest is carried on a campaign. Alternatives to field campaigns include stationary observatories, passive drifters, or remote technologies (i.e., satellites), most of which are calibrated using reference curves generated from physical samples collected and processed by human experts. Human-staffed field campaigns can target specific geographic locations which may not have long-term infrastructure present, or may not be accessible by other means (e.g., satellite optical instrumentation cannot penetrate

water). Even in regions which can be observed by other means, data collected on field campaigns serve to re-calibrate instrumentation, supplement a dataset with specialized measurements, or dynamically target a particular phenomenon.

1.1.1 The “Gold Standard”

Physically collecting (i.e., removing) singular, discrete samples of a target environmental phenomenon is widely considered the “gold standard” in many disciplines [4]. In order to physically collect samples of an environment, it requires that a human expert or mobile system is able to safely navigate to and from the target region and be able to appropriately store the specimen for *ex situ* analysis in a laboratory. In marine campaigns, sample collection could involve pumping water samples from depth to a ship, or sending a Remotely Operated Vehicle (ROV) to the seafloor to retrieve sediment cores. In biological or chemical sciences, sample storage may involve using a poisoning chemical to terminate biological processes or adding preservatives for long-term storage. Physical specimen collection and subsequent laboratory analysis provides insurance that the target phenomenon of interest is directly observed and measured. Generally, replicates are made (either by collecting multiple samples at the same location or dividing a large specimen into separate samples) and independently processed to constrain procedural error and noise.

Although the gold standard, there are notable drawbacks to relying on physical sample collection alone. The first is logistical cost: due to the monetary expense and time associated with finding, collecting, transporting, storing, and processing specimen, the total number of samples that may be realized for a region may be relatively few compared to the geographic extent, spatial volume, or transience of the phenomenon. This is known as “data sparsity,” and makes characterizing distributions of a phenomenon in space or time challenging. In the worst case, these samples may not be representative of the true underlying distribution of a phenomenon, leading to incorrect interpretations or misleading conclusions. The second drawback is related to data analysis comparability across independent studies of similar phenomena. Based upon recent inter-laboratory studies in environmental chemistry, it is

unclear whether laboratories can draw consistent [5] or inconsistent [6] conclusions from equivalent samples. In general, these contradictory results inspire caution with respect to comparing published datasets or relying too closely on physical samples alone to draw conclusions.

1.1.2 *In situ* Instrumentation and Unmanned Platforms

Given the drawbacks of solely using collected samples, it is necessary to supplement samples with other direct observational data. *In situ* instrumentation records measurements of observations in near-real time of physical attributes that either could be confirmed by subsequent laboratory analysis, or simply may be difficult to extract in a laboratory environment at all (i.e., local temperature). Generally, this equipment is powered electrically, and the actual measurement that is observed is expressed as a voltage, which can be converted to a physical quantity by applying a calibration curve. *In situ* instrumentation is generally calibrated with a set of accepted laboratory measurements, so raw observations by this equipment can be directly compared with simultaneously collected physical samples that are processed *ex situ*.

With the ability to take a hundred-fold more observations than traditional sample collection, *surveying* strategies have emerged in which a mobile platform carrying an instrument or suite of instruments is used to collect observations across a specified target region. Generally, it is desirable to perform a *one-shot mission*, or a single deployment of a mobile platform to collect data in a target region. Cost, time limit, and vehicle safety all play a role in this preference. Among the most common strategies in the environmental and earth sciences used to perform one-shot missions is *uniform coverage* which is also known as *lawnmowing* or *boustrophedonic* search* [7], e.g., [8–11]. In uniform coverage, a target region is traversed by making long sweeping motions across one dimension of the region while incrementally moving in the other dimension; much like how one may methodically mow a field. Although this method enjoys simplicity in execution and guarantees with regards to spatial resolution, it

*from boustrophedon, which is a form of writing such that the lines alternate directions (right to left followed by left to right)

can suffer from data sparsity in interesting regions due to its non-adaptivity.

“Autonomy” in environmental and earth sciences has generally been in the service of automating simple, predefined navigation tasks. However, with smaller and more powerful computers, improved performance of *in situ* sensors, and continued development in robust control and planning algorithms, intelligent autonomy is fully realizable in practice. With intelligent autonomy comes the ability to specify more complicated adaptive missions, to coordinate multi-vehicle fleets, and to create theoretically sound guarantees about the overall quality of samples that can be collected.

1.2 Maximum Seek-And-Sample

In many environmental and earth science applications, experts want to collect scientifically valuable samples of a maximum (e.g., an oil spill source). This maximum could correspond with collecting the most *pristine sample* in an environment, the most productive source in an environment, the highest concentration area, or similar attributes. The *maximum seek-and-sample* (MSS) problem is pervasive in many disciplines, however, natural environments are difficult to examine because they are typically unknown *a priori*, are *continuously distributed*, *partially observable*, and display *transience*. This thesis presents the PLUMES adaptive sampling framework to overcome these challenges to efficiently find and sample from the global maximum of a natural environment.

Informative path planning (IPP) is a broad field which develops adaptive sampling frameworks. To encode the challenging aspects of the MSS problem, a partially-observable Markov decision process (POMDP) [12] is defined. A POMDP is a general model for decision-making under uncertainty, and is composed of a state space, observation space, action set, transition function, and reward function. In the MSS POMDP, the state and observation space are continuous functions, where the state represents the partially-observable underlying phenomenon. The action set defines the behaviors that an agent can select from, and the transition function defines how the robot’s internal model of the world, known as *belief*, will update according to

observations collected during an action. The reward function encodes the scientific objective of the MSS problem, giving value only to samples sufficiently close to the global maximum. Solving a POMDP exactly is generally intractable [13], and the MSS POMDP is additionally complicated by both continuous state and observation spaces, and the sparse MSS reward function. This presents the three core challenges that PLUMES addresses: compactly representing the robot’s belief of the state of the transient phenomenon, overcoming reward function sparsity, and performing online search in a belief-space over continuous functions.

1.2.1 Representing Belief

Every observation that a mobile agent collects provides partial information about the underlying state of the world. In order to plan actions, the robot must consider what it *believes* the complete, true state of the environment is based upon these observations. A belief representation functionally stores historical observations, incorporates new observations, and can be queried while the vehicle plans what the next best action is to take. Generally, this requires that the belief representation take the form of a probability distribution over possible states, where a sample drawn from the distribution is one possible realization of the environmental phenomenon. This allows the robot to query the belief representation for a possible world model, evaluate actions with respect to the phenomenon distribution within that model, and consider the uncertainty associated with that model. In the MSS problem, the underlying distribution is a continuous function, and observations which may update the belief state are also continuous, thereby requiring that the belief representation be a probability distribution over an uncountably infinite number of possible environments.

In order to compactly represent and plan with continuous states and observations, it is common in literature to discretize the state space [14, 15]. In the MSS problem, discretization over space and time is complicated by the reward function, which requires precision about the location of the global maximum. Poor choices for discretization parameters will yield poor estimates of the maximum, and subsequently poor performance for any planner. Thus, to represent the robot’s belief over continu-

ous states conditioned on a history of observations without discretization, PLUMES uses Gaussian Processes (GPs) [16]. The GP is parameterized by a mean function and a covariance kernel function. Since the environment is unknown *a priori*, the mean function is assumed to be uniform (e.g., uninformative). The kernel function can be used to encode both spatial and temporal relationships in an environment based upon human knowledge; for example the dispersion characteristic of a gas species in the atmosphere, or the dynamics of an oil-leaking ship. This thesis provides insight about designing kernel functions for transient phenomena, and properties of kernel functions which allow for efficient inference about natural environments.

1.2.2 Heuristic Reward

In the MSS problem, the reward function is *sparse*; the single non-zero reward is placed at the maximum of the world. This presents challenges for selecting useful actions, especially at the beginning of a mission when few observations are available to build a useful belief model of the target phenomenon and the location of the maximum. Planning with sparse rewards requires long-horizon information gathering and is an open problem in robotics [17]. To alleviate this difficulty, less sparse heuristic reward functions are often used in place of the true reward function, to elicit an *explore-exploit* behavior in the planner. In the explore phase, actions that drive down uncertainty about the world are valued, and in the exploit phase the gathered knowledge is used to target actions which more directly satisfy the scientific objective.

Selecting an appropriate heuristic reward for the MSS problem requires designing an information-theoretic measure which elicits explore-exploit behavior, and efficiently converges to good estimates of the global maximum. In the field of Bayesian optimization (BO), black-box functions are typically used to optimize over an unknown function or distribution. The most commonly used heuristic reward in IPP is the Upper confidence Bound (UCB) reward [18, 19], which sums the predicted mean and variance of a proposed sample to assign a reward value. This thesis shows that for the MSS problem, UCB can lead to suboptimal behavior, and draws on work from the BO community to propose the *maximum value information reward* (MVI) [20]

for robotic systems. In contrast with UCB, MVI uses samples from the belief representation of the robot to distribute reward according to the possibility of finding the maximum at a particular location. As the robot examines promising regions and takes more observations, the belief representation and the location of the inferred maximum converges, and MVI places reward solely at this point, entering into the exploit phase. For the MSS problem, this reward convergence is attractive as it encourages efficient identification of and dense sample collection at the maximum.

1.2.3 Decision-Making and Planning in Continuous Domains

A planner searches over possible actions and environmental states (drawn from the belief representation) in order to find the most rewarding action or action sequence for the vehicle to take. Generally taking the form of a search tree, planners can be classified as either *myopic* or *nonmyopic*. As the name implies, myopic agents are short-sighted: they greedily take the best action based on local information. In contrast, a nonmyopic agent will choose the best action based on global information.

The difference between these planners is highlighted in a demonstrative scenario shown in Fig. 1-1, in which there are two modes in the distribution of the phenomenon of interest; mode A is greater than mode B in value. The robot is given the objective to collect as many high-valued samples as possible, knows the full distribution, is initialized in a random location, and can move 1 unit in any cardinal direction. If using a myopic planner, the robot will perform gradient ascent to the nearest mode and remain there (as any other action will locally collect less reward). In this bimodal world, if the robot is initialized near mode B, it will converge to this lower mode and fail to collect the globally optimal amount of reward, although the location of mode A is known. However, if using a nonmyopic planner, the robot may move away from mode B and towards mode A as it has the “foresight” to go to the globally optimal location, even if it requires locally suboptimal choices. Although in some environments a myopic planner may be suitable, PLUMES implements a nonmyopic planner to optimize the MVI heuristic for the attractive convergence and global planning properties.

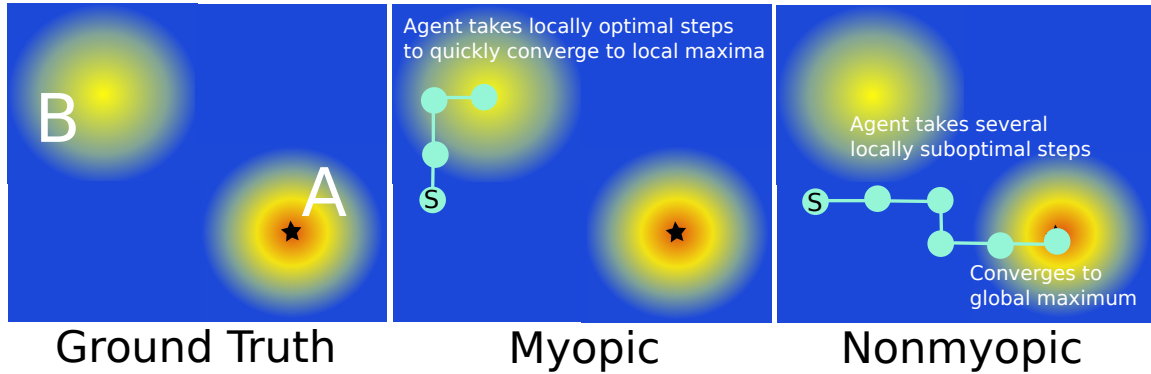


Figure 1-1: **Simple illustration of myopic and nonmyopic planning** A robot tasked with collecting high-valued samples is placed in a world with two modes, A and B. Mode A is greater (more valuable) than Mode B; the global maximum is marked with a star. The robot has full knowledge of the world. When the robot is initialized close to Mode B (at the marker labeled “S”), a myopic planner (center) will take a series of locally optimal decisions and converge on Mode B. A nonmyopic planner (right) will initially take several suboptimal actions in order to converge on Mode A because it optimizes over the globally optimal action sequence.

Generally, nonmyopic search is most successful when state representations, which are nodes in the search tree, can be revisited during multiple action sequence simulations. By revisiting a state multiple times, the estimate of the reward that may be accumulated by reaching or traveling through that state can be better constrained. Due to the continuous observation space in the MSS problem, an uncountable set of possible belief states arises (since each observation will be unique with probability 1), creating a degenerate tree. Practically, tree degeneracy leads to inefficient decisions and theoretically prevents any formal guarantees regarding planner performance. Given the importance of establishing confidence for an autonomy system in the environmental and earth sciences, the key challenge to overcome is related to addressing tree degeneracy in continuous, partially-observable domains and formulating performance guarantees of the framework. This thesis extends planners for fully-observable continuous-domains [21, 22] to partially-observable domains, and demonstrates that theoretical performance guarantees are preserved.

1.3 Thesis Overview

The contributions of this thesis are:

1. Presentation of three field campaigns which analyze transient chemical phenomena in near-shore seas and coastal estuaries and rivers:
 - (a) The effect of wastewater effluent on the Wareham River, MA, USA
 - (b) Surface expression of methane from bubble plumes in the Cascadia Margin
 - (c) Arctic outgassing driven by spring-time river inflow in Cambridge Bay, Nunavut, Canada
2. Formulation of the MSS problem as a POMDP
3. Presentation of the PLUMES adaptive sampling framework:
 - (a) Analysis of GP kernel functions to model spatio-temporal phenomena
 - (b) Identification, adaptation, and implementation of a BO black-box information-theoretic reward function for the MSS problem
 - (c) Novel analysis of a nonmyopic planner demonstrating guarantees in continuous, partially-observable environments
 - (d) Demonstration of improved performance over baseline techniques in simulated and field trials

The remainder of this thesis is organized as follows: Chapter 2 presents technical background and related work on the foundations of adaptive sampling. Chapter 3 presents the three field campaigns and scientific discoveries, and discusses the role of transience in those environments. In Chapter 4, the MSS problem is formalized and the PLUMES algorithm is described in detail, including core sections on the belief representation of transient phenomena, heuristic reward function, and nonmyopic planner. Comparisons with baseline planners from literature are also presented. A summary of results and brief discussion concludes in Chapter 5.

Chapter 2

Technical Background and Related Works

This thesis builds upon foundational work in probabilistic modeling, planning under uncertainty, and information measures in order to examine the development and use of intelligent robotic systems for environmental field campaigns. Central to this work is the concept of *adaptive sampling*, which is a technique that informs future behaviors of an agent or system by incorporating a history of observations. Within the fields of robotics and sensing, adaptive sampling techniques have been broadly applied in applications including:

1. Direct observation and study of environmental domains (e.g., [23–27])
2. Within localization frameworks to improve state estimates (e.g., [28–31])
3. Image processing (e.g., [32, 33])
4. Distributed sensor networks (e.g., [34–36])

This chapter will primarily frame the technical discussion of core adaptive sampling concepts within the context of environmental sampling with mobile robotic agents. Problem- and content-specific related work will be provided in each of Chapters 3 and 4 in technical discussions of the thesis contributions.

2.1 Measuring and Modeling Natural Phenomena

The intent of using robotic platforms in environmental and earth science campaigns is to collect “useful data” about target phenomena, which allow scientists to gain insight about an underlying process and address hypotheses. This insight is generally driven by comparing data to *scientific models*, which encode well-characterized physical relationships between phenomenon (e.g., relationship between hypoxic zones and high methane content in deep ocean waters). Numerical models based on first principles dominate much of the sciences [37–40]. These models are incredibly expressive and comprehensive. Collected data are generally used to tune free parameters or define new relationships to be incorporated into these models in order to match input conditions with realized physical observations [41].

Although the standard, numerical models are generally brittle to noisy or unexpected measurements and do not generalize well to new datasets, largely because they rely on an assumption of determinism. This brittleness has motivated the use of statistical or *probabilistic* models, which can capture uncertainty in measurements and physical relationships (e.g., [42–44]). Innovation in environmental probabilistic modeling complements innovations made in robotics; kriging is perhaps the most clear example of a probabilistic model that was developed originally for characterizing geostatic phenomenon [44] and has since been widely adopted in the form of Gaussian process regression [16] in robotics and machine learning communities.

Probabilistic models are particularly attractive for robotic applications because they allow for decision-making to be informed by measures of certainty (or uncertainty). For example, one of the most ubiquitous probabilistic models in robotic navigation is an occupancy grid [45] which is used to model occupied, free, and unknown physical space based on observations of obstacles by an agent. The state of a grid cell is a probabilistic measure based upon the history of noisy observations collected by a robot platform and a model of the noise characteristic of the sensor. When space is uncertain, then the robot has the ability to collect more measurements to increase certainty, or choose alternate paths which may be believed to be

safer. However, unlike metric obstacles, environmental phenomenon are generally assumed to be complex, multimodal, dynamic, and stochastic. To suitably model such phenomena, this thesis primarily makes use of Gaussian processes (GPs) [16], which are a Bayesian nonparametric model from machine learning. Section 2.1.1 provides background on Bayesian inference, Section 2.1.2 describes Bayesian model representations, and Section 2.1.3 specifically presents GPs and discusses work in environmental sampling which has leveraged GPs for studying phenomena.

2.1.1 Bayesian Inference Techniques

Bayesian models, as the name implies, leverage Bayes Theorem as a basis for performing inference over a set of unknown parameters:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \quad (2.1)$$

where Bayes Theorem states that the *posterior distribution* of a set of parameters θ given a dataset y is proportional to the *likelihood* of the data given the set of parameters, and a *prior* distribution on the parameters. In environmental domains, the prior distribution is a way to encode knowledge about scientific principles or relationships. It has been observed that Bayesian techniques have become increasingly used in the environmental sciences [46] due to the flexibility of various Bayesian frameworks, uncertainty characterization, and robustness to variability and nondeterministic factors.

A Bayesian inference problem takes the following form: let $\mathcal{X} = \{X_0, \dots, X_{N-1}\}$ be a set of N random variables with a relationship described by joint distribution $\Pr(X_0, \dots, X_{N-1})$. Some subset of these random variables are observed, \mathcal{X}_{obs} . Bayesian inference allows for these observations to inform what values the unobserved, latent variables $\mathcal{X}_{lat} = \mathcal{X} \setminus \mathcal{X}_{obs}$ are likely to take. A direct application of Bayes Theorem and probability theory yields:

$$\Pr(\mathcal{X}_{lat}|\mathcal{X}_{obs}) = \frac{\Pr(\mathcal{X}_{lat}, \mathcal{X}_{obs})}{\Pr(\mathcal{X}_{obs})} = \frac{\Pr(\mathcal{X}_{lat}, \mathcal{X}_{obs})}{\int_{\mathcal{X}_{lat}} \Pr(\mathcal{X}_{lat}, \mathcal{X}_{obs})} \quad (2.2)$$

where the denominator is the marginal likelihood of the observations. In practice,

performing exact inference with Eq. 2.2 is computationally expensive to the point of being intractable: calculating the marginal likelihood is exponential in the number of latent variables. To overcome this challenge, approximation techniques are employed.

Variational Bayesian Inference One common approximate inference technique is variational Bayesian inference [47, 48], which approximates a posterior distribution with a well-behaved function class, $q^*(\theta) \approx p(\theta|y)$. In order to identify $q^*(\theta)$ from class Q , an optimization procedure over some distance measure f is performed:

$$q^*(\theta) = \arg \min_{q \in Q} f(q(\cdot), p(\cdot|y)) \quad (2.3)$$

The Kullback-Leibler (KL) divergence, $\text{KL}(\cdot||\cdot)$ is a common choice for the distance metric because of good empirical performance and adaptability to fast, streaming, and distributed systems [48]. By using KL divergence, a convenient simplification for the optimization problem arises:

$$\text{KL}(q||p(\cdot|y)) = \log p(y) - \int_{\Theta} q(\theta) \log \frac{p(\theta)p(y|\theta)}{q(\theta)} d\theta \quad (2.4)$$

$$q^*(\theta) = \arg \max_{q \in Q} \int_{\Theta} q(\theta) \log \frac{p(\theta)p(y|\theta)}{q(\theta)} d\theta \quad (2.5)$$

where the *ELBO* (evidence lower bound) of the KL divergence is the second term in Eq. 2.4 and the optimization can primarily occur with respect the ELBO as it only contains well-defined aspects of the model.

This only leaves selecting distributions Q that are well-behaved. A common choice is to apply the mean-field approximation:

$$Q = \{q : q(\theta) = \prod_{i=1}^n q_i(\theta_i)\} \quad (2.6)$$

which provides that the set of distributions over the latent parameters factorizes and allows low-dimensional representations to be considered. With this family of distributions, Eq. 2.5 can be solved using simple coordinate ascent approaches [47].

Extensions of variational inference, such as stochastic variational inference (SVI)

[49] and automatic differentiation variational inference (ADVI) [50], have been demonstrated to improve standard variational techniques under assumptions of conjugacy or differentiable properties.

Monte Carlo Inference Methods An alternative to variational techniques to estimate the form of the true posterior $p(\theta|y)$ are Monte Carlo (MC) methods [51]. Rather than perform an optimization over analytic functions to get a closed form for the posterior, MC methods approximate estimators (Φ) of a function $\phi(\cdot)$ with respect to the true density $p(y)$ by drawing samples from a proposal density $q(y)$. This work generally assumes that the form of $p(y)$ is known and can be evaluated to within a multiplicative constant, $p(y) = p^*(y)/Z$, but is generally difficult to draw samples from directly (especially in high-dimensional domains). Monte Carlo sampling instead draws samples from a more tractable, known proposal distribution $q(y) = q^*(y)/Z_q$ and evaluates those samples with respect to $p^*(y)$ in order to approximate Φ . One of the most straightforward MC samplers is importance sampling:

1. Draw x_1, \dots, x_N i.i.d. samples from $q(\cdot)$.
2. Calculate weight $w_i = p^*(x_i)/q^*(x_i)$.
3. Calculate estimate $\Phi = \frac{\sum_N w_i \phi(x_i)}{\sum_N w_i}$

Rejection sampling is another MC sampler, and leverages the intuition that some samples generated from $q(y)$ may not align well with $p(y)$. An acceptance criteria for adding a sample x_i to the dataset is defined through the rule $p^*(x_i) > u$ where u is a draw from a uniform distribution with bounds $[0, q^*(x_i)]$.

MC methods generally require that the form of $q(y)$ lie near the form of $p(y)$. In large, complex systems, it is difficult to define a single density that captures these characteristics. Markov chain MC (MCMC) algorithms address this by drawing new samples x' using a proposal density which relies on the state of the previous sample $x^{(t)}$; $q(x', x^{(t)})$. In Metropolis-Hasting MCMC, an acceptance ratio is used to transition between states:

$$a = \frac{p^*(x')q(x^{(t)}; x')}{p^*(x^{(t)})q(x'; x^{(t)})} \quad (2.7)$$

where if $a \geq 1$ the new state x' is accepted and $x^{(t+1)} = x'$; otherwise a new sample x' is drawn from $q(x', x^{(t)})$ and $x^{(t+1)} = x^{(t)}$. Other MCMC samplers, like Gibbs [51], Reversible-Jump [52], and Hamiltonian [53] use other acceptance ratios or special forms of $q(y)$ in order to improve the convergence characteristics, flexibility, and speed of Metropolis-Hastings. In all MCMC samplers, because each new state relies on a density informed by the previous state, a “burn-in” period, in which a potentially large number of samples are drawn, must be used before virtually independent samples are generated. It has been shown that for a large number of samples, the estimator generated with MC methods will converge to the true estimator of the posterior [51].

2.1.2 Bayesian Representations

In order to perform inference, a model is necessary to describe (i.e., the likelihood, prior, or other relationships between latent parameters). This section briefly describes several ways within the scope of Bayesian inference that phenomena may be modeled.

Graphical Models Probabilistic graphical models (PGMs) exploit the conditional independence structure of the latent parameters. Bayesian networks [54] are one particular form of graphical model which has enjoyed adoption in both the environmental sciences [55] and in robotic environmental sensing missions [15] as they allow for specific relationships between latent parameters to be encoded.

Let a Bayesian network be defined as an acyclic graph $\mathcal{G} = (V, E)$ where vertices V represent random variables, and directed edges represent dependencies between two variables indexed $(i, j) \in E$ where i in this example is a *parent* to j . The joint probability of V in the graph can be expressed as the product of all conditional probabilities $\Pr(X_i | \text{parents}(X_i))$ where the parents X_j of node X_i are the connected edges (j, i) . In complex networks, exact inference may be intractable. Variational and MCMC techniques can be used to approximately solve a Bayesian network.

Parametric and Nonparametric Models An alternative to Bayesian networks that allow for exact inference with respect to encoded or learned relationships are parametric models. In a parametric model, a fixed number of parameters are fit to a portion of data, known as the training set. The model’s accuracy is then assessed using an unseen test set of data in a process known as cross-validation. By assessing model accuracy, different numbers or types of parameters can be designed for desired performance. Finite mixture models (e.g., [56]) are one form of parametric model.

Although typically seen as a computational advantage, in the environmental sciences a drawback of parametric models is that model complexity remains the same no matter the size of data available. Selecting a “good” number of parameters, or even *which* parameters to use *a priori* may be difficult for processing real observations, as many phenomenon have complicated structure that may be entirely context dependent. In general, it would be useful if the data could directly inform the parameter-space, rather than have the parameter-space predefined. Nonparametric models were designed specifically to allow model complexity to grow with more data, and can handle a potentially infinite number of parameters.

To do this, nonparametric models capture the notion of representing probability over a collection of distribution functions. The Dirichlet process [57], the Chinese restaurant process [58], and other methods have all been proposed for generating and representing infinite-mixture models for inference. In the environmental sciences and for robotic applications within the environmental sciences, Gaussian Processes (GPs) [16] have received considerable attention [19,35,59–61].

2.1.3 Gaussian Processes

Informally, a GP is used to represent a distribution over functions. Formally, Rasmussen and Williams define a GP in [16] as:

Definition 2.1.1. *A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution.*

For modeling an environmental phenomenon, first let the domain of the inference

target be represented as a d -dimensional compact set $\mathbb{X} \subset \mathbb{R}^d$. The unknown underlying, m -dimensional continuous function is $f : \mathbb{X} \rightarrow \mathbb{R}^m$. For notational simplicity, $m = 1$ in the following discussion. Samples of f can be drawn in a location \mathbf{x} with a noisy sensor: $y = f(\mathbf{x}) + \epsilon$ where $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$ is normally distributed sensor noise.

A GP is fully parameterized by a mean $\mu(\mathbf{x})$ and covariance function $\kappa(\mathbf{x}, \mathbf{x}')$ (also known as a kernel), which take the form:

$$\mu(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})] \tag{2.8}$$

$$\kappa(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - \mu(\mathbf{x}))(f(\mathbf{x}') - \mu(\mathbf{x}'))] \tag{2.9}$$

so that $f(\mathbf{x}) \sim \mathcal{GP}(\mu(\mathbf{x}), \kappa(\mathbf{x}, \mathbf{x}'))$. Given a history $\mathcal{D}_t = \{\mathbf{x}_i, y_i\}_{i=0}^D$ of D observations and observation locations at time t , the posterior belief at a new location $\mathbf{x}' \in \mathbb{X}$ is:

$$g_t(\mathbf{x}') \mid \mathcal{D}_t \sim \mathcal{N}(\mu_t(\mathbf{x}'), \sigma_t^2(\mathbf{x}')), \text{ where} \tag{2.10}$$

$$\mu_t(\mathbf{x}') = \kappa_t(\mathbf{x}')^\top (\mathbf{K}_t + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y}_t, \tag{2.11}$$

$$\sigma_t^2(\mathbf{x}') = \kappa(\mathbf{x}', \mathbf{x}') - \kappa_t(\mathbf{x}')^\top (\mathbf{K}_t + \sigma_n^2 \mathbf{I})^{-1} \kappa_t(\mathbf{x}'), \tag{2.12}$$

where $\mathbf{y}_t = [y_0, \dots, y_{D-1}]^\top$, \mathbf{K}_t is the positive definite kernel matrix with $\mathbf{K}_t[i, j] = \kappa(\mathbf{x}_i, \mathbf{x}_j)$ for all $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_t$, and $\kappa_t(\mathbf{x}') = [\kappa(\mathbf{x}_0, \mathbf{x}'), \dots, \kappa(\mathbf{x}_{D-1}, \mathbf{x}')]^\top$.

In practice, $\mu(\mathbf{x})$ is typically set to 0, and the kernel function is primarily used to encode the relationship between features in the environment. Kernel functions contain hyperparameters, which specify properties like smoothness, correlation strength, periodicity, and scale. Hyperparameters can be set directly through prior knowledge about a phenomenon of interest, or through online optimization techniques (e.g., [62]).

The popularity of GPs for environmental modeling and robotic planning is rooted in the relative ease of training and performing inference with the representation, in addition to the simple but expressive way in which kernel hyperparameters can be utilized to enforce structure over data. In large datasets, GPs can be computationally expensive, as the training and inference time is cubic with respect to number of observations. Sparse GPs [63] can be employed to address scalability concerns.

2.2 Environmental Sensing as a Robotics Problem

Environmental sensing with a robotic platform can be considered an *information-gathering* mission. In such applications, the environmental model is useful for representing the robot’s belief about the state of the target phenomenon with respect to a history of observations, but decision-making requires more information; a planning model is necessary. A planning model specifies the vehicle dynamics, available actions, belief about the physical state of the vehicle, the precise mission objective, and a method for evaluating an action with respect to the mission objective. Information-gathering is most generally modeled as *sequential decision-making* in which actions are evaluated, selected, and executed at each planning iteration. Sequential decision-making processes are considered Markovian when an action selection is conditionally independent of previous actions and observations given the current belief state of the vehicle (with respect to both environmental and state model). Markov decision processes (MDPs) [64, 65] are a useful model for a robotic planning problem, and are described in Sec. 2.2.1. An extension of MDPs to domains in which the state of the world is only partially observable, POMDPs [12], is presented in Sec. 2.2.2.

2.2.1 Markov Decision Processes (MDPs)

A Markov decision process (MDP) is represented as a tuple $(\mathcal{S}, \mathcal{A}, T, R, \gamma, s_0)$ where:

- \mathcal{S} is the set of finite or infinite (in the case of continuous functions) decision states
- \mathcal{A} is the set of finite or infinite (in the case of continuous actions) actions that are available to the vehicle, \mathcal{A}_s is the set of actions available from state s .
- $T : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{P}(\mathcal{S})$ is the transition function which represents the probability density of being in state $s \in \mathcal{S}$, taking action $a \in \mathcal{A}$, and arriving in state $s' \in \mathcal{S}$; $T(s, a, s') = \Pr(\mathcal{S}_{t+1} = s' | \mathcal{S}_t = s, \mathcal{A}_t = a)$. This allows for imperfect dynamics in either the robot control or the modeled environment.

- $R : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ is the reward function, which represents the value of performing some action $a \in \mathcal{A}$ when in state $s \in \mathcal{S}$. Can alternatively be $R : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \mathbb{R}$ if value is awarded by arriving into a state $s' \in \mathcal{S}$ from state $s \in \mathcal{S}$ after taking an action $a \in \mathcal{A}$.
- γ is the discount factor which is applied in infinite-horizon missions.
- s_o is the initial decision state.

A *policy* $\pi : \mathcal{S} \rightarrow \mathcal{A}$ which maps decision states to actions is a solution to an MDP. An optimal policy π^* describes the set of actions to take from any given state that maximize the total (potentially discounted) reward for a h -horizon mission (in which h can be infinity):

$$\pi^* = \arg \max_{\pi} \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) | s_0, \pi \right] \quad (2.13)$$

The optimal policy from state $s \in \mathcal{S}$ can be determined using *value iteration*, which iteratively estimates the value of the optimal policy using the Bellman equation [65]:

$$\begin{aligned} V_{t+1}^*(s) &\leftarrow \max_{a \in \mathcal{A}} \left[\sum_{s' \in \mathcal{S}} T(s, a, s') (R(s, a, s') + \gamma V_t(s')) \right] \\ \pi^*(s) &= \arg \max_{a \in \mathcal{A}} \left[\sum_{s' \in \mathcal{S}} T(s, a, s') (R(s, a, s') + \gamma V^*(s')) \right]. \end{aligned} \quad (2.14)$$

For a threshold ϵ , such that value iteration is terminated when $|V_{t+1}(s) - V_t(s)| < \epsilon$, then $\max_{s \in \mathcal{S}} |V_{t+1}(s) - V^*(s)| < 2\epsilon\gamma/(1 - \gamma)$. Value iteration converges in polynomial time.

2.2.2 Partially Observable MDPs (POMDPs)

At their core, MDPs assume that the state at any time is *fully-observable*, however, in the case of environmental sensing, a collected observation generally does not reveal the full state of the phenomenon. Thus, the underlying function f is *partially-observable*.

Partially-observable Markov decision processes (POMDPs) [12] extends MDPs to partially observable domains, defined as the tuple $(\mathcal{S}, \mathcal{A}, \mathcal{Z}, T, O, R, \gamma, b_0)$ where $\mathcal{S}, \mathcal{A}, R$, and γ are defined as in Sec. 2.2.1 with:

- \mathcal{Z} is the space of all possible observations. May be finite or infinite (in the case of continuous functions).
- $O : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{P}(\mathcal{Z})$ is the observation model, which represents the probability density of observation $z \in \mathcal{Z}$ after executing action $a \in \mathcal{A}$ from state $s \in \mathcal{S}$; $\Pr(O_t = z | \mathcal{S}_t = s, \mathcal{A}_t = a)$. This function can model imperfect sensing.
- b_0 is the prior distribution over the initial state \mathcal{S}_0 ; $b_0 = \Pr(\mathcal{S}_0 = s)$

In partially-observable domains, the state of the world is uncertain. In general, this means that the decision process is no longer Markov, as the optimal policy would no longer be dependent on the state. However, by making decisions based on the *belief* over states rather than making decisions based on the current best estimate of the state, the Markov property can be restored, since the belief state summarizes all the historical observation and action history relevant for policy calculation. Just as in MDPs, the Bellman equation can be used to recursively quantify the value of belief $b_t = \mathcal{P}(S_t)$ over horizon- h under policy $\pi : b_t \rightarrow a_t$ as:

$$V_h^\pi(b_t) = \mathbb{E}[R(s_t, \pi(b_t))] + \gamma \sum_{z \in \mathcal{Z}} V_{h-1}^\pi(b_{t+1}^{\pi(b_t), z}) \Pr(z | b_t, \pi(b_t)), \quad (2.15)$$

where the expectation is taken over the current belief and $b_{t+1}^{\pi(b_t), z}$ is the updated belief after taking action $\pi(b_t)$ and observing $z \in \mathcal{Z}$. The optimal policy π_h^* over horizon- h is the maximizer of the value function over the space of possible policies Π : $\pi_h^* = \arg \max_{\pi \in \Pi} V_h^\pi(b_t)$.

In general, Eq. 2.15 is difficult or intractable to compute in large or continuous state and observation spaces. The same holds for large or continuous-valued MDPs. Thus, approximate solvers are necessary in order to extract the optimal policy for a robotic agent to execute. Sec. 2.3 and Sec. 2.4 discuss in detail the formulation of the reward function and methods for approximately solving MDPs and POMDPs.

2.3 Reward Specification

In an MDP or POMDP formulation, the reward function serves to encode the scientific objective of a mission. For example, if the objective was to drive the vehicle to the maximizer of some distribution, then the reward function may deliver reward only to actions that directly lead to states in the proximity of the maximizer. Some objectives may be easy to encode as a heuristic reward, however in environmental domains, objectives may be less obvious to encode, e.g., “map” a phenomenon, “learn” a distribution, etc. Even with relatively straightforward objectives, like the illustrative maximizer objective above, rewards that take on a “sparse” form (i.e., the reward signal itself is only provided in very few states) are generally hard for a robot agent to optimize well [17]. In either of these cases, it is useful to consider reward functions that can quantify the value of *information content* of a potential observation in order to elicit *explore-exploit* behaviors.

Explore-exploit describes decision sequences which initially allow the agent to *explore* the unknown state space in order to characterize the distribution and build a sophisticated belief over the domain, then *exploit* that knowledge in service of some objective (e.g., converging to a maximizer). Reward functions which encode this trade-off may use hyperparameters to tune the balance in the explore-exploit behavior, whereas other functions may naturally enforce these behaviors. Information measures have been studied as suitable reward signals to encode explore-exploit behavior in several fields, including optimal experimental design, optimal sensor placement, reinforcement learning, and Bayesian optimization. This section provides a brief overview of these fields and core concepts in information measures, with particular focus on information-theoretic measures from Bayesian optimization in Sec. 2.3.2.

2.3.1 Overview of Information Measures

Information measures provide a way to assess the utility of a measurement with respect to the history of observations and ultimate objective. Several widely used information measures have been proposed in multiple fields; for details on information

measures in general, MacKay [66] provides a comprehensive overview.

Optimal Experimental Design In optimal experimental design [67], a typical experiment requires estimating a vector $\mathbf{z} \in \mathbb{R}^d$ from a set of measurements \mathbf{x} and observations \mathbf{y} such that $y_i = x_i^T \mathbf{z} + w_i$ for $i = 1 \dots N$ observations, where w_i is zero-mean Gaussian noise. An estimator is then calculated for \mathbf{z} given the experimental set-up. Commonly, the maximum likelihood estimate (MLE), $\hat{\mathbf{z}}$, is used because it is an unbiased estimator. The estimation error between the $\hat{\mathbf{z}}$ and \mathbf{z} is $E = \hat{\mathbf{z}} - \mathbf{z}$. The covariance matrix of E is ultimately used to characterize informativeness of a measurement by inspecting how the measurements reduce overall error variance.

Several types of optimal design arise based upon different criteria for inspecting variance reduction. Some optimal design regimes used in robotics, learning, and sensing applications are listed:

- *A*-optimal: Minimizes the trace of the inverse covariance matrix, e.g., [68–70]
- *D*-optimal: Minimizes the determinant of the covariance matrix, e.g., [69–71, 71]
- *E*-optimal: Maximizes the smallest eigenvalue of the covariance matrix, e.g., [70]
- *V*-optimal: Minimize the average prediction variance, e.g., [72]

Soft Measures of Information In optimal experimental design, “hard” estimates of a target variable (i.e., explicit estimators) are used to quantify the informativeness of a measurement. In contrast, “soft” estimates use the probability distribution over a target variable. Shannon’s entropy [73] was defined with respect to four key design principles for an information measure, $I(\cdot)$, with respect to probability p :

1. Information measures are continuous, monotonic functions of probability; $I(p)$
2. Information is non-negative; $I(p) \geq 0$
3. If an event has probability 1 (it is certain to occur), then no information is gained; $I(1) = 0$

4. The information gained from observing two independent events is the sum over information gained from each individual event; $I(p_1 p_2) = I(p_1) + I(p_2)$.

From these principles, entropy was defined as:

$$H(X) = - \sum_{i=1}^n \mathcal{P}(x_i) \log_b \mathcal{P}(x_i) \quad (2.16)$$

where $\mathcal{P}(\cdot)$ is a probability density over random variable X with values x_1, \dots, x_n and b is a selected logarithm base. Entropy quantifies the expected log-loss of the distribution over a random variable. In robotic information-gathering missions, it is useful to be able to compare how informative one observation is with respect to another. Relative or conditional entropy follows from Eq. 2.16 to quantify the informativeness of observing a random variable Z after observing another random variable Y :

$$H(Z|Y) = - \sum_{a,b} \Pr(Z = a, Y = b) \log \Pr(Z = a|Y = b) \quad (2.17)$$

This is equivalently interpreted as the expected log-loss of the conditional distribution $\Pr(Z|Y)$ with respect to Z . In kind, mutual information quantifies how much Y reveals about Z , or the average change in log-loss when observing Y :

$$\begin{aligned} I(Z; Y) &= H(Z) - H(Z|Y) \\ &= \sum_{a,b} \Pr(Z = a, Y = b) \log \frac{\Pr(Z = a, Y = b)}{\Pr(Z = a)\Pr(Y = b)} \end{aligned} \quad (2.18)$$

Entropy, conditional entropy, and mutual information have been used extensively in SLAM techniques [74–77], sensor placement [35, 78, 79], optimal navigation [80], and robotic sampling [15, 81].

2.3.2 Rewards in Bayesian Optimization

This thesis particularly draws inspiration from information-theoretic rewards designed in the field of Bayesian optimization (BO). Several quintessential reward measures are

presented in this section; Chapter 4 discusses in detail a state of the art BO reward measure that is incorporated into an adaptive sampling algorithm.

Upper-Confidence Bound The upper-confidence bound (UCB) [82–84] is among the most commonly used reward functions for robotic environmental sampling/sensing missions (e.g., [19, 85–88]). The reward function takes the form:

$$R_{\text{UCB}} = \mu(\mathbf{x}) + \sqrt{\beta_t} \sigma(\mathbf{x}) \quad (2.19)$$

where for a set of queries \mathbf{x} , the UCB reward is the sum of the predictive mean μ and variance σ at the queries, where variance is scaled by a time-dependent hyperparameter. UCB is a *submodular* function, which is defined by Nemhauser et al. [89]:

Definition 2.3.1. *Given a finite set E , a real-valued function f on the set of subsets of E is submodular if $f(A) + f(B) \geq f(A \cup B) + f(A \cap B), \forall A, B \subseteq E$.*

Intuitively, this definition encodes the notion of *diminishing returns*. In the context of environmental sensing, submodularity can be interpreted: “as a robot agent learns more about an environment, new measurements tend to yield less information.”

Submodularity is used in the analysis of UCB-guided algorithms to place bounds and guarantees on overall performance. Srinivas et al. [19] provide detailed analysis of UCB-based reward functions for use in environments represented by GPs, ultimately demonstrating a bound on *regret* for some selection of β_t . Regret is a general performance metric used to quantify the loss in reward from sub-optimal decisions made because the underlying function f is unknown. For robotic and sensor-selection missions, *no-regret* performance implies that as time approaches infinity the accumulated regret goes to 0, and is a popular way of proving useful convergence properties of an algorithm. UCB reward, and UCB variants have been shown to elicit no-regret properties in robotics and sensor selection problem [19, 86, 90].

Probability of Improvement A different measure of sample quality with respect to information is probability of improvement (PI) [84, 91], which is intuitively the

probability measure of whether a proposed query \mathbf{x} will be better than the current best measurement \mathbf{x}^* . For a GP, PI can be written:

$$\gamma(\mathbf{x}) = \frac{f(\mathbf{x}^*) - \mu(\mathbf{x})}{\sigma(\mathbf{x})} \quad (2.20)$$

$$R_{PI} = \Pr(f(\mathbf{x}) \geq f(\mathbf{x}^*)) = \Phi(\gamma\mathbf{x}) \quad (2.21)$$

where Φ is the cumulative density function of a standard normal distribution, f is the unknown function, and $\mu(\cdot)$, $\sigma(\cdot)$ are the predictive mean and variance of f over the queries.

Expected Improvement Expected improvement (EI) [84,92] is a measure of how much better a proposed measurement will be over the current best measurement. For a GP, EI can be written:

$$R_{EI} = \sigma(\mathbf{x})(\gamma(\mathbf{x})\Phi(\gamma(\mathbf{x})) + \epsilon) \quad (2.22)$$

where $\gamma(\mathbf{x})$ is the same form as in Eq. 2.20, and ϵ is normally distributed noise.

Predictive Entropy Search Predictive entropy search (PES) [93] is a relatively recent reward function in BO derived specifically for GPs, and specifically aims to find an optimum in some function f . PES is an extension of Entropy Search, which is derived by Hennig et al. [94]. The intuition of PES is that by using predictive mean and variance of the GP belief state, the conditional entropy between a proposed measurement \mathbf{x} and the predicted optimizer of the GP \mathbf{x}^* can be used to guide useful queries. The reward function takes the form:

$$\begin{aligned} R_{PES} &= H(\Pr(\mathbf{x}^*|D)) - \mathbb{E}_{\Pr(y|D,\mathbf{x})}[H(\Pr(\mathbf{x}^*|D \cup \mathbf{x}, y))] \\ &= H(\Pr(y|D, \mathbf{x})) - \mathbb{E}_{\Pr(\mathbf{x}^*|D)}[H(\Pr(y|D, \mathbf{x}, \mathbf{x}^*))] \end{aligned} \quad (2.23)$$

where D is a history of measurements and observations, and y are observations (i.e.,

the value of f at measurement locations). The second term in Eq. 2.23 arises from noticing that the first term is equivalently the mutual information between \mathbf{x}^* and y given a history D , and that mutual information is a symmetric function. In order to compute the reward function, it is necessary to draw estimates of \mathbf{x}^* . For certain forms of the GP, these samples can be drawn using a spectral density function; details are provided in [93] and briefly described in Chapter 4.

2.4 Decision-Making under Uncertainty

As explained in Sec. 2.2, solving a POMDP or large MDP requires an approximate planning strategy. In the most broad sense, a planner can be characterized as either being *online* or *offline*, which describes at what point in a mission a plan may be generated. This thesis primarily focuses on online planning regimes, which can be further characterized as either *myopic* or *nonmyopic*, referring to how many actions ahead an agent considers when making a decision. In this section, a brief overview of different planners used in robotic environmental sampling and exploration are presented.

2.4.1 Offline Planning

Offline planning approaches specify an execution pattern for an agent prior to a mission, which the agent then executes in open-loop control. Simplistic offline planners perform coverage or monitoring tasks [95,96] in *a priori* known metric environments. Reward functions like “shortest path length” or “minimal energy expenditure” are typical. Offline planning also refers to a system in which many potential plans or contingencies are computed prior to a mission, and during execution one of these plans is selected on-the-fly based on robot state [97].

2.4.2 Online Planning

In contrast to offline planners, online planners are used “in the loop” for vehicle control during mission execution. Online planners may be fully closed-loop, wherein

streaming measurements and observations have direct consequence on robot behavior. Generally, closed-loop planners are used for motion-control, in which obstacle avoidance, perturbation rejection, and navigation are core tasks [98, 99].

To optimize over an information measure, open-loop feedback control (also referred to as partially closed-loop control) can be used [14, 15, 100], in which a planner designs a trajectory to a finite horizon h , the plan is executed using open-loop control, and a subsequent trajectory from the robot’s new state is planned based upon observations gathered. Horizon length further classifies open-loop feedback controllers into *myopic* or *nonmyopic* regimes, in which the former refers to short-horizon (typically $h = 1$) planning, and the latter refers to long-horizon planning.

Myopic Planners

As the name implies, myopic planners are “short-sighted” — the single best action to take at some planning iteration t is selected according to some reward function. Selecting the locally best option at a planning iteration is often referred to as *greedy* planning.

The simplest interpretation of a greedy-myopic planner is:

$$a^* = \arg \max_{a \in \mathcal{A}} R(s, a) \tag{2.24}$$

where the most rewarding action a^* is selected from all actions \mathcal{A} such that the reward function R is maximized with respect to the robot’s current state s . Another form of greedy plan with respect to set notation can be written as in Alg. 5 [101].

Algorithm 1: Greedy Algorithm: An illustrative format for a myopic-greedy planner [101].

Data: reward function R , possible measurements $\mathcal{X} = x_i$, sampling capacity K
Result: sample set $\mathcal{A} \subseteq \mathcal{X}$

- 1 $\mathcal{A} \leftarrow \emptyset$;
- 2 **while** $|\mathcal{A}| \leq K$ **do**
- 3 $\mathbf{x}^* = \arg \max_{x \in \mathcal{X}} R_{\mathcal{A}}(\mathbf{x})$;
- 4 $\mathcal{A} \leftarrow \mathcal{A} \cup \mathbf{x}^*$;
- 5 **return** \mathcal{A}

The performance of a greedy-myopic planner is arbitrarily worse than the optimal plan; however under submodular (Def. 2.3.1) reward functions, such as UCB in GP representations, the greedy-myopic search performance can be bounded. Moreover, it has also been shown to be “no-regret” [19, 86], that is, the average regret per action goes to 0 at the limit of infinite planning iterations. More generally, a bound can also be defined for set-representations given a submodular, *monotone* reward function [101] and greedy-planner in the form of Alg. 5, where a function f is monotone iff $\forall S \subseteq T \subseteq N, f(S) \leq f(T)$. In practice, these performance bounds are loose, and in finite-duration missions agents can be arbitrarily poor-performing.

Nonmyopic Planners

To specifically address improved finite-duration performance, nonmyopic planners simulate taking several actions in order to select the next behavior to execute, optimizing over the expected future reward. Sampling-based methods like RRT [102] are nonmyopic planners which use random samples of target domains to perform computationally efficient, globally optimized path planning. Informative sampling-based planners [103, 104] have also been developed, which allow for optimization of a sampled path over an information-theoretic objective.

By far, the most common form of nonmyopic planner in IPP problems are variations of tree search. The most naive form of this type of planner is brute-force search over a finite horizon, enumerating all possible combinations of action sequences. Perhaps obviously, this method is generally expensive for large or complex discrete systems, and intractable for continuous-domain systems. For problems with discrete states, branch and bound methods [105] have been used under monotone reward functions to prune unpromising pathways and improve overall planning time. For continuous-valued domains, state-of-the-art planners use deterministic discretization [106] or a combination of sampling techniques and particle filter belief representations [87, 107–109] to reduce the size of the possible set of states to search over, and then perform classic tree search techniques. Specific discussion of the challenges of searching over continuous domains can be found in Chapter 4.

Monte Carlo Tree Search (MCTS) [110] has received increased attention in the IPP domain and takes the general form:

1. **Selection** Select a valid action from the robot’s current state
2. **Rollout** Forward simulate taking that action and subsequent actions up to a horizon h
3. Calculate the reward of the simulated trajectory
4. **Backup** Update the value of the root action to be the average reward accumulated by the present simulation and all previous simulations
5. Repeat 1-4 until termination criteria are met
6. Return the most rewarding action

MCTS differs from other tree searches by asymmetrically growing the tree by selecting the most “urgent” node for examining based upon a meta-search heuristic, which examines average accumulated reward for a node, and number of visits for a node (this is the *selection* phase). The meta-search heuristic is known as the *tree policy*, and itself can define the explore-exploit characteristics of the tree search. The *upper confidence bound for trees* (UCT) policy has been shown to converge to the minimax tree as the number of simulations expands to infinity, and the probability of selecting a suboptimal action converges to 0 [110, 111]. This convergence guarantee is particularly attractive as it is agnostic to the specific reward function that is being used to assess the quality of a simulated rollout. This thesis discusses how to adapt MCTS for POMDPs with continuous state and observations spaces in Chapter 4.

2.5 Robots in the Wild

This thesis is primarily interested in using robotic agents to examine environmental phenomenon. Robotic technologies were first used in field campaigns to execute simple, preset missions, e.g., [112]. Many of the early robotic systems were for marine

studies [113] or in efforts to develop extraterrestrial rovers [114]. Presently, applications for robotics can be found in agricultural sciences [115], forestry [116], and climate monitoring [117], to name only a few. This thesis specifically highlights the utility of the ChemYak marine vehicle [118] for biogeochemical studies in Chapter 3. One of the first recognized marine surface vehicle was ARTEMIS [119] which was able to perform simple bathymetry studies. The JetYak [120] developed by Woods Hole Oceanographic Institution (WHOI), the WaveGlider [121] developed by Liquid Robotics, and the Saildrone [122] are among the most prolific surface vehicles for performing science missions.

Complementary developments in the environmental sciences and robotics have driven innovation in both fields. As discussed in Sec. 2.1.1, environmental modeling and probabilistic modeling for planning have considerable overlap. Practically, sensor development, networking, wireless communication, longterm monitoring, and energy consumption are shared concerns in general robotics and environmental engineering. In the subsequent chapters, work at the intersection of the sciences and robotics will be presented as a contribution to both fields.

Chapter 3

Transience in Marine Science: Greenhouse Gas Emissions

The study of water bodies on Earth — oceans, estuaries, rivers, lakes, and the land-water continuum — connects the marine sciences. *Transient phenomenon* are ubiquitous in marine environments, referring to a physical property (that is spatially distributed) which changes over time. For example, on Cape Cod, MA, estuaries and salt marshes (e.g., Little Sippewissett Marsh) are influenced by both fresh and salt waters, which are cyclically mixed by tides, stochastically influenced by weather events and human-influence, and annually impacted by the changing seasons. In an estuary, phenomenon may change on time-scales of hours (in the case of tides) to months (in the case of seasons). In order to completely study biological, chemical, geochemical, and physical aspects of marine environments, transience must be considered.

Canonical methods of studying marine environments have relied on physical sample collection. For example, from a bottle sample of water quantities like dissolved gases and minerals, biological content, and water source (e.g., ground water, storm water) can be examined. However, bottle samples can be difficult to collect in hazardous environments or expensive to collect when specialized equipment is necessary. Moreover, exhaustive bottle collection is generally not practical for logistical reasons, including collection time, sample storage, and timely sample processing. Thus, every bottle sample that can be collected in an environment may be considered “precious”

in a field campaign, and significantly influence conclusions drawn from a study.

Collecting good bottle samples is a hard problem in the environmental and marine sciences, especially when the phenomenon of interest is not obvious and the distribution is unknown. For example, chemical properties of a water body, like the concentration of dissolved gases, may not be detectable to a human observer and often analysis of a water sample will be delayed for weeks or months after collection. In these cases, bottle samples may be distributed uniformly in space over a target region, although the phenomenon of interest may have temporal characteristics, or a spatial distribution, for which this is not well-suited. *In situ* instrumentation can provide near real-time insight about the distribution of a target phenomenon. To use *in situ* instrumentation effectively may require exhaustive and repetitive navigation of a target region. Robotic platforms, particularly unmanned agents, are uniquely suited for performing these tasks. Such agents additionally expand the number of environments that it is possible to observe by going where a piloted vessel or a larger vessel may be unable to navigate for safety or scientific reasons (e.g., shallow waters, at the site of a calving glacier, or in biologically dense areas).

This chapter presents three field campaigns which primarily study the distribution of the greenhouse gases methane (CH_4) and carbon dioxide (CO_2) with an unmanned mobile surface platform, the ChemYak. Section 3.1 motivates the study of CH_4 and CO_2 in the marine sciences, and Section 3.2 presents the ChemYak platform and data processing techniques. Sections 3.3-Section 3.5 present the three field campaigns, which represent near-shore open seas, coastal estuaries, and Arctic estuaries. Section 3.6 discusses transience in natural environments, drawing on the field campaigns to illustrate key challenges in characterizing and sampling such phenomenon with a mobile platform and ultimately motivating adaptive sampling.

3.1 Studying Methane and Carbon Dioxide

Greenhouse gases absorb infrared radiation and contribute to the overall warming of the atmosphere. Methane and carbon dioxide are potent greenhouse gases, and

together with nitrous oxide (N_2O), contribute 80% of total radiative forcing in the atmosphere [123]. Since the industrial revolution, atmospheric levels of these gases have significantly increased from historical levels (in the known geological record), largely due to human activity [123]. CO_2 increased by 40% between 1750 and 2011; and CH_4 increased by 150% in the same time period [123].

3.1.1 Carbon Dioxide

CO_2 is an odorless, colorless gas that is soluble in water [124]. Major sources of anthropogenic CO_2 include burning of fossil fuels, cement production, and agriculture/land-use changes [125]. CO_2 emissions also occur naturally e.g., from soils (generally driven by decomposition, microbial respiration, or plant respiration) [126], and near-shore aquatic environments like estuaries and marshes [127]. Through *gas exchange*, the largest natural contributor of CO_2 into the atmosphere is the surface ocean, driven by the partial CO_2 pressure differential between water and air. Interestingly, the ocean also serves as one of the largest sinks of atmospheric CO_2 , sequestering nearly 48% of all anthropogenic emissions [128]. The complex relationship between CO_2 and seawater is described by three core mechanisms: the solubility pump, the biological pump, and the marine carbonate pump [125].

The solubility pump refers to the relationship between the solubility of CO_2 in water and temperature; in cold water CO_2 is more soluble [129]. Temperature and salinity circulation in the ocean can carry CO_2 from the surface to the depths, effectively sinking CO_2 . However, in upwelling conditions (e.g., warming of interior water), CO_2 can be outgassed at the surface ocean [130]. The biological pump describes the biologically-driven process of carbon cycling. At the surface ocean, phototrophs, like phytoplankton, use atmospheric CO_2 in photosynthesis. When these organisms die, they sink through the ocean column where bacterial decomposition in transit and at the seafloor recycles the carbon. Any carbon that escapes decomposition is sequestered in sediments [131]. The carbonate pump completes the carbon cycle, and describes the process of inorganic carbon recycling in the ocean, centered primarily on calcium carbonate which forms shells and hard exteriors of marine organisms [132].

3.1.2 Methane

CH₄ is a colorless, odorless hydrocarbon that is soluble in water (although much less so than CO₂) [133]. CH₄ is less abundant in the atmosphere than CO₂, but has a significantly higher global warming potential (GWP). For a time horizon of 20 years, the GWP of CH₄ is 87 gCO_{2eq}/gCH₄ and for a 100 year horizon (the standard reporting horizon) is 36 gCO_{2eq}/gCH₄ [123, 134]. Although CH₄ is more potent than CO₂, the overall lifetime in the atmosphere is shorter; approximately 12.4 years [123, 134]. Natural sources of methane include geochemical systems like marine/terrestrial seeps or geothermal vents, wetlands, biogenic sources (like termites), and the ocean [125, 135]. Anthropogenic sources of methane are generally attributed to fossil fuel extraction and use, agriculture, livestock (i.e., ruminants), landfills, and waste treatment [125, 135–138]. More generally, CH₄ is a key component of natural gas [139], and is used in chemical processes like steam methane reforming, which produces bulk hydrogen stock [140].

CH₄ is generally removed from the atmosphere by reacting with hydroxide (OH-) radicals in the stratosphere [141, 142]. The products of this chemical reaction are water and CH₃ (known as methyl) [142]. Additionally, some bacteria and archaea have been shown to oxidize methane [143, 144]. In the ocean, large geological stores of methane exist as frozen gas hydrate in shallow ocean sediments and permafrost soils [125, 145]. Additionally, aquatic methanotrophs consume CH₄ at natural seeps or at exposed hydrate ridges [145, 146].

3.1.3 Impacts

CO₂ contributions from natural sources are volumetrically much greater than from anthropogenic sources; nearly half of all emissions are from air-sea exchange, whereas burning fossil fuels accounts for approximately 5% [125, 147]. Similarly, atmospheric CH₄ is dominated by natural sources, although anthropogenic sources of CH₄ have increased at faster rates than CO₂ since 1750 [123]. The influx from anthropogenic sources, however, is enough to perturb the balance between natural sources and sinks.

The most general and direct consequence of elevated greenhouse gases in the atmosphere is known as *global climate change* or *global warming*, and describes the shift in the Earth's climate driven by anthropogenic contribution of greenhouse gases to the atmosphere. Elevated CH_4 and CO_2 concentrations, and global climate change itself, has secondary, compounding impacts on the environment. Rising temperatures have been shown to impact seasonal freeze-thaw cycles and permafrost in the Arctic [123, 148], insect populations [149, 150], and sea-level rise [151]. For these specific examples, consequences include increased greenhouse gas emissions from the Arctic [123, 152, 153], mass extinction events of insects in tropical regions [150], and coastal reshaping that impacts human habitation. Cyclical behaviors, such as increased emissions of gases from Arctic soils or temperature destabilization of gas hydrates in warming seas [154], are of considerable concern.

A second direct consequence of elevated CO_2 in the atmosphere is ocean acidification [155, 156]. The ocean sequesters atmospheric carbon, and the resulting chemical reaction produces small amounts of acid. As the ocean absorbs more carbon available in the atmosphere, the net pH of the water decreases. Ocean acidity has been shown to negatively impact coral colonies and invertebrates which rely on carbonate to form shells [156]. Other direct effects of elevated gas levels include impacts to plant tissues [157] and ozone pollution [158].

The impact of climate change and elevated greenhouse gases has implications for health and safety of humans, human activities, and ecosystems at large. A better understanding of natural and anthropogenic sources, in addition to the efficiency and efficacy of natural sinks and engineered interventions, is a considerable focus in the environmental sciences. The ultimate goals of field work associated with the study of greenhouse gases are to (1) quantify chemical flux, (2) characterize spatial and temporal trends of gas concentration, and (3) identify flux drivers (e.g., river inflow, anthropogenic sources, seasonal warming). Marine environments are particularly important natural sources and sinks, and stand to be disproportionately impacted by climate change and elevated atmospheric CH_4 and CO_2 levels [159]. In the field campaigns presented in this chapter, a ChemYak surface vehicle [118] equipped with *in*

situ instrumentation was used to capture the fine-scale dynamics of CH₄ and CO₂ in the coastal ocean, a tidal estuary and river, and an Arctic estuary.

3.2 The ChemYak Mobile Platform

The ChemYak (Fig. 3-1) is an unmanned, remotely operated mobile platform developed at the Woods Hole Oceanographic Institution (WHOI) for *in situ* measurement of chemical species in the surface ocean and coastal waters [118] (based on the JetYak platform [120]). The chassis of the vehicle is a Mokai* jet-propelled kayak with servo steering control. The vehicle is remotely piloted through a 2.4 GHz radio channel, or can be programmed to follow a pre-specified trajectory using a PixHawk† autopilot, which runs open-source Ardupilot. A central computer onboard the ChemYak logs data from instruments and is connected to a high bandwidth local network to stream data and diagnostic information to a remote computer. The vehicle is equipped with a 10 m profiling winch mounted at the port-aft of the vehicle, and a rigid arm for sensors is mounted at the starboard-aft.

The ChemYak is equipped with a sensor suite capable of measuring CH₄ and CO₂, in addition to other physical and chemical properties of a target environment. The most general configuration of the ChemYak includes a weather station, an oxygen optode, a nitrate sensor, a CTD, and a gas extractor and analyzer.

Weather Station: An Air-Marine‡ 200WX Weather Station is used to record GPS coordinates, relative velocity of the vehicle, wind speed, and air humidity and temperature. GPS coordinates are served by Wide Area Augmentation System (WAAS) satellites, with 3 m accuracy 95% of the time. The weather station is mounted at the rotational center, approximately 1.5 m above the hull of the vehicle. Data are streamed from the weather station through a serial connection to the onboard central computer, and logged to a text file at 1 Hz.

*mokai.com

†pixhawk.org

‡airmar.com

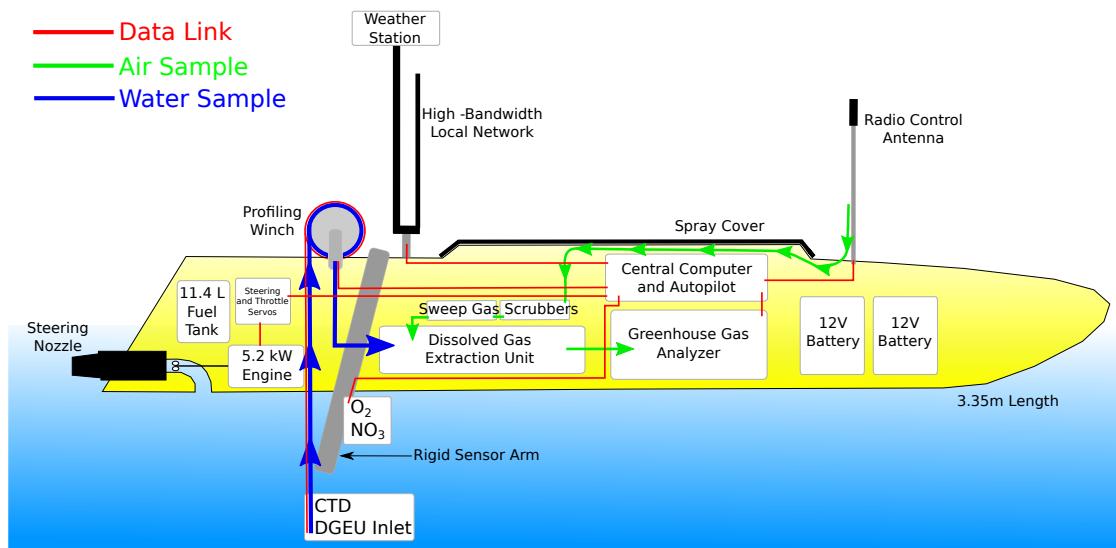


Figure 3-1: **The ChemYak** The ChemYak is a jet-propelled unmanned kayak with servo-controlled steering. The ChemYak can be piloted by a remote human user through a radio link, or can execute predetermined waypoints with a PixHawk autopilot. To measure gas species in the top 10 m of the water column, a winch with a CTD and pump inlet can be lowered. The pump inlet serves a field-portable Los Gatos Research (LGR) Dissolved Gas Extraction Unit (DGEU) and LGR Greenhouse Gas Analyzer (GGA), which uses infrared laser spectroscopy to optically determine the content of an extracted gas sample. An oxygen optode and nitrate sensor are optionally attached to a rigid sensor arm on the starboard-aft of the vehicle at about 0.15-0.25 m depth. A weather station provides GPS coordinates, wind speed, and relative velocity. All instruments log at 1 Hz. On the figure, inlet channels for water samples (blue) and gas samples (green) are indicated; all waste water and gas are pumped and released at the aft of the vehicle. Data connections are marked in red. Electrical connections are not drawn, however all instruments are connected to the battery supply provided by 2 12 V lead-acid batteries.

Oxygen Optode: An Aanderaa* oxygen Optode 4831 is mounted to the rigid arm of the vehicle, approximately 0.15-0.25 m below the water surface. Optode 4831 uses a gas permeable foil to passively extract a gas species to a small sample chamber, and blue light is used to excite a sensing foil. The red light reflection is measured, and oxygen (O_2) content corresponds to the reflection intensity (by applying principles of fluorescence). Dissolved oxygen within 0-1000 μM is measured with a $<0.1 \mu M$ precision and accuracy of $<2 \mu M$ or 1.5%. CH_4 rich environments tend to be hypoxic; adding an optode to the ChemYak configuration provides a second avenue to detect

*aanderaa.com

elevated greenhouse gas content. The instrument can return measurement at 1 Hz and is logged by the central computer on the vehicle.

Nitrate Sensor: In the most general configuration, the ChemYak carries a nitrate (NO_3) sensor, co-located with the optode on the sensor arm. For the purposes of the studies presented here, the nitrate sensor was not used.

CTD: An RBR* Concerto CTD probe (Conductivity-Temperature-Depth) depth-rated for 100m is attached to the end of the profiling winch. A CTD is used to generally observe basic water properties, like salinity and temperature. Data from the CTD can be streamed to the central computer and transmitted to a remote user, providing some feedback with respect to probe depth. The reported depth accuracy is $\pm 0.05\%$ of full-range, with resolution of 0.001% of full-range. Temperature is reported in degrees Celsius with accuracy $\pm 0.002^\circ$, and 0.00005° resolution. Conductivity, which is used to ultimately calculate salinity, is reported with $\pm 0.03 \text{ mS cm}^{-1}$ and resolution 0.001 mS cm^{-1} . The CTD self-logs on an internal computer at 1 Hz.

Gas Analysis: Two instruments are dedicated to the extraction and subsequent analysis of greenhouse gas species in a water sample. The first instrument, a Los Gatos Research[†] (LGR) Dissolved Gas Extraction Unit (DGEU), is used to extract gas from a water sample. The DGEU actively pumps water samples for analysis; the inlet to the pump is attached to the end of the profiling winch, co-located with the CTD. The DGEU also pumps a sweep-gas reference sample from the ambient atmosphere; the sweep-gas inlet is located at the bow of the vehicle, and is scrubbed of CO_2 and water vapor before entering the unit.

The DGEU uses a membrane contactor (Liqui-Cel G420 2.5 x 8) to separate gas and water by using gas permeable tubing and filters. It is generally assumed that the extraction is imperfect, so it is necessary to apply an efficiency correction to better constrain gas content in a sample. Each gas species has its own associated efficiency,

*rbr-global.com

†lgrinc.com

largely due to the differences in solubility between the two species. To estimate this efficiency for a field campaign, laboratory analysis by Nicholson et al. [118], empirical analysis with contemporaneous bottle samples, or analysis of measurement distributions can be used. The selected technique for each field campaign will be indicated in the respective sections.

The separated gas sample is subsequently pumped to a LGR greenhouse gas analyzer (GGA), which is a field-portable infrared spectroscopic instrument (a proprietary off-axis integrated cavity enhanced absorption technique is used). Gas measurements are reported in partial pressure units (ppm). Logging is performed at a frequency of 1 Hz, which corresponds to precision in CH₄ of <2 ppb and CO₂ of <300 ppb. The instrument self-logs. Waste gas is released into the atmosphere at the stern of the vehicle and waste water from the DGEU is expelled into the water body at the aft of the vehicle, near the motor servo nozzle.

3.2.1 Measurement Analysis

In this chapter, all data reported have been quality controlled by hand to remove anomalous events (typically when sensors are out of the water, during the beginning and end of a mission, or while performing periodic filter checks). Although the gas sensor directly reports partial pressure units (ppm), it is standard to report CH₄ in molar quantities (e.g., nM) and CO₂ in atmospheric quantities (e.g., μatm). To perform this conversion, it is necessary to apply the efficiency correction to the raw measurements. This takes the form:

$$p_{dry} = \left(\frac{m - g_{ppm}}{\phi} + g_{ppm} \right) \frac{p_{eq}}{1000} \quad (3.1)$$

where m is the observed measurement from the instrument, g_{ppm} is the reference concentration of the target gas in the atmosphere (e.g., 1.86 ppm for CH₄), p_{eq} is the pressure imposed by the pump during extraction, and ϕ is the efficiency to be applied. p_{dry} is the efficiency-corrected measure, now reported in μatm . For CH₄, an additional step where the corrected value is converted to a molar quantity is necessary.

Principles from [160–162] are implemented by the open-source library *gasex** in order to perform the conversion to molar units. The conversion for CH₄ follows:

$$y = 1.00024T + K_0 \quad (3.2)$$

$$\log K = a_0 + a_1 \frac{100}{y} + a_2 \log \frac{y}{100} + S(b_0 + b_1 \frac{y}{100} + b_2 (\frac{y}{100})^2) - \log V_{CH_4} \quad (3.3)$$

$$C_{CH_4} = \frac{K p_{dry}}{1e - 9} \quad (3.4)$$

where S is the salinity measure in practical salinity units (PSS), T is temperature in degrees Celsius, K_0 is a constant set to 273.15 (0°C in Kelvin), constants a and b are set from [160], and V_{CH_4} is the molar volume of CH₄.

To perform this conversion, contemporaneous salinity and temperature measurements are required. In general, the CTD, gas analyzer, optode, and weather station log on different computers with different clocks. To perform the desired conversion, in addition to assigning a geolocation to every measurement, the instrument data are interpolated onto a common timestamp; the time logged with the GPS coordinates from the weather station is used.

3.3 Wastewater Effluent in Tidal Estuaries

Coastal zones are found at the interface of land and ocean environments, and host the most productive ecosystems on Earth, largely due to a diversity of physical and geochemical processes. The complexity of carbon cycling in these environments makes characterizing and modeling atmospheric flux particularly challenging. For example, rivers are considered sources of CH₄ and CO₂, which are outgassed during turbulent events [163, 164], however salt marshes can be considered a carbon sink through biological activity [165]. When anthropogenic influences are examined with respect to these environments, there are further complications. Estuaries and rivers receive a considerable amount of treated and untreated wastewater. Major facets of research

*github.com/dnicholson/gasex-python

on estuarine waters include characterizing the impacts of synthetic materials (e.g., antibiotics [166]), increased organic materials (e.g., fertilizers and treatment feedstock [167]), and urban developments (e.g., stormwater run-off [168]).

To better understand both natural and anthropogenic carbon cycling in estuaries, this field campaign examines an estuary and river in New England with a treated sewage outfall. The Wareham River in Massachusetts is one of the largest river systems in the Buzzards Bay feeder network*. This study focuses on an approximately 1 km stretch of the Wareham River on which the Wareham Sewerage Department is located and pumps treated effluent. Several other local rivers flow into the Wareham river downstream of the treatment facility, which ultimately opens into Buzzards Bay. The river experiences semi-diurnal tidal cycles, with a water level change of approximately 1.0-1.5 m[†]. At the field site, only small motor-watercraft could traverse the river, and at low tide only personal watercraft (e.g., kayaks) could traverse the shallow portions of the river, and other regions became exposed mudflats. The tidal plane of the river varies from 12 m at its narrowest to 155 m at its widest. The sewage treatment plant is a 1.56 MGD (million-gallon-per-day) BNR (biological nutrient removal) wastewater pollution control facility. The feedstock for bacteria in one treatment stage is methanol, which scientists hypothesize may lead to elevated levels of CH₄ entering the estuary through the treated water. Disruption of the biogeochemical processes of the estuary may have effects to the immediate surrounding area around the treatment facility, as well as downstream. To better determine whether CH₄ is present in wastewater effluent and the extent to which it may be transported by the river, the ChemYak was used to perform dense spatial mapping over the course of a tidal cycle (high-tide to low-tide) in the Summer of 2017.

3.3.1 Overview of Field Work and Analysis

The ChemYak was deployed near the outfall of the treatment facility (41.758022, -70.684128) and piloted by a human user. Over the course of 4-5 hours, 5 transects

*savebuzzardsbay.org/embayments/wareham-river/

†tides.mobilegeographics.com/locations/8728.html

(which will be referenced as Transects A through E), on average 800 m long, were conducted at the end of high tide (Transects A, B), at slack tide (Transect C), and during the falling tide (Transects D, E), where the falling tide pulls water *downstream* towards Buzzards Bay. An oxygen optode, CTD, and gas analyzer were available on the ChemYak to take measurements of the river body, and all instruments were attached (or had inlets attached as is the case for the gas analyzer) to the rigid sensor arm of the vehicle, at approximately 0.15 m depth. Over 11,500 measurements representing approximately 9.5 km of travel are presented. Fig. 3-2 shows the 5 transects undertaken during the field trials overlaid onto a reference map of the region, where the outfall of the treatment plant is marked with a red star. The extraction efficiency applied to convert CH₄ and CO₂ measurements was equivalent to that used in Nicholson et al. [118], and set to 15% and 70% respectively.

3.3.2 Results

The field campaign revealed that elevated CH₄ was present in source water that was likely from the wastewater effluent outfall. Fig. 3-3 shows each transect that was conducted and overlays CH₄, CO₂, O₂, salinity, and temperature onto a representative map. In general, elevated CH₄ levels are present near the outfall, reaching 450 nM. A trail of elevated CH₄ toward Buzzards Bay, likely pulled by the falling tide, is evident in the transects, however CH₄ levels consistently decline to less than 100 nM at the far extent of the field site. This decrease may be indicative of rapid outgassing of CH₄ occurring in the river, or other physical dynamics such as a mixing boundary. Based upon the average salinity (1.63 ± 1.5 PSS) and temperature (25.29 ± 0.4 °C) of the surface waters, the expected equilibrium for the system is approximately 2.6 nM (assuming atmospheric concentration of CH₄ was 1.86 ppm). Elevated CO₂ is also evident with similar patterns to CH₄ and likely derived from effluent waters, with a peak concentration of over 4500 μ atm. At this field site, considerable biological activity was evident, and the waters were measured to have an oxygen content of over 220 μ M. In waters with elevated greenhouse gas content, reduced oxygen concentrations between 120-160 μ M were observed.

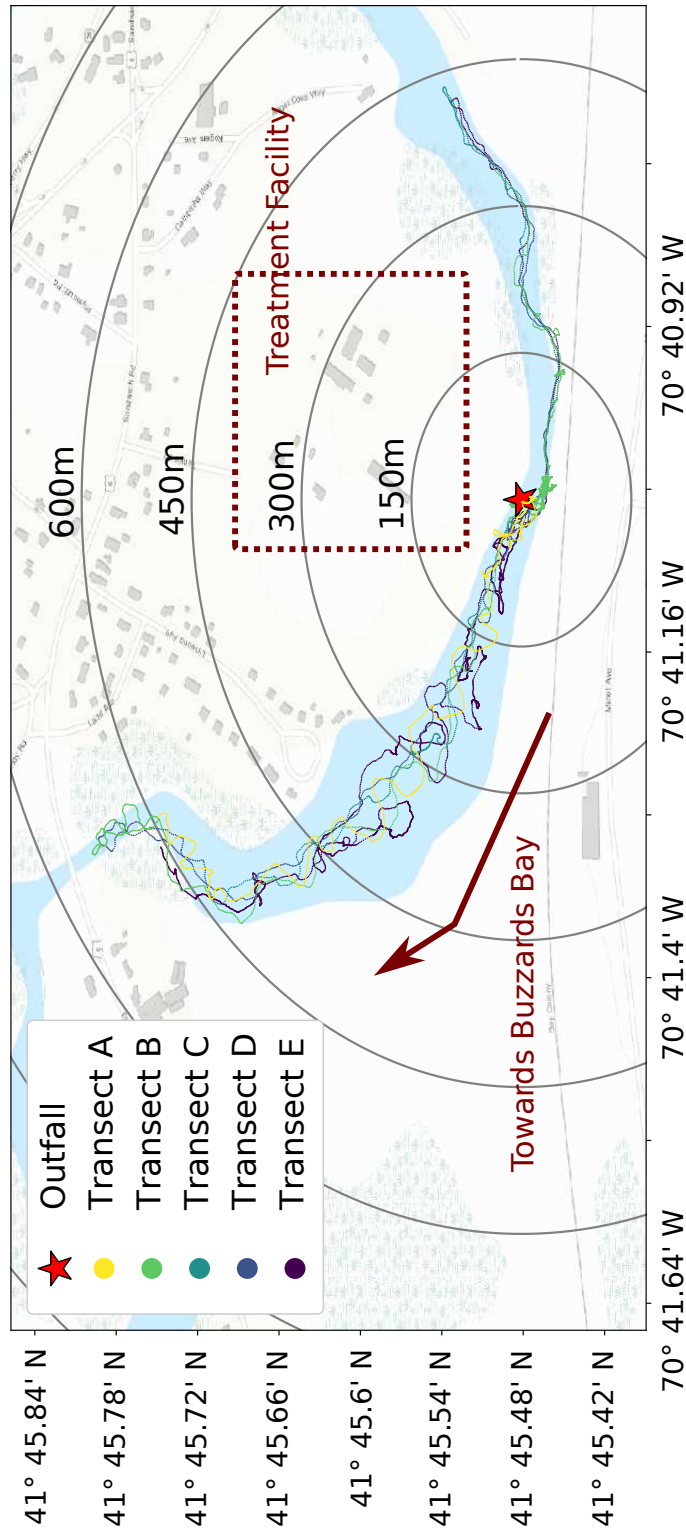


Figure 3-2: **Wareham River field site** A representative map of the Wareham River Estuary field site is provided, with Transects A-E overlaid for reference. The outfall is marked with a red star. To the left of the outfall, annotated with an arrow, the river leads to Buzzards Bay, whereas to the right, the river winds through dense marshland. The site of the treatment facility is also annotated, for reference.

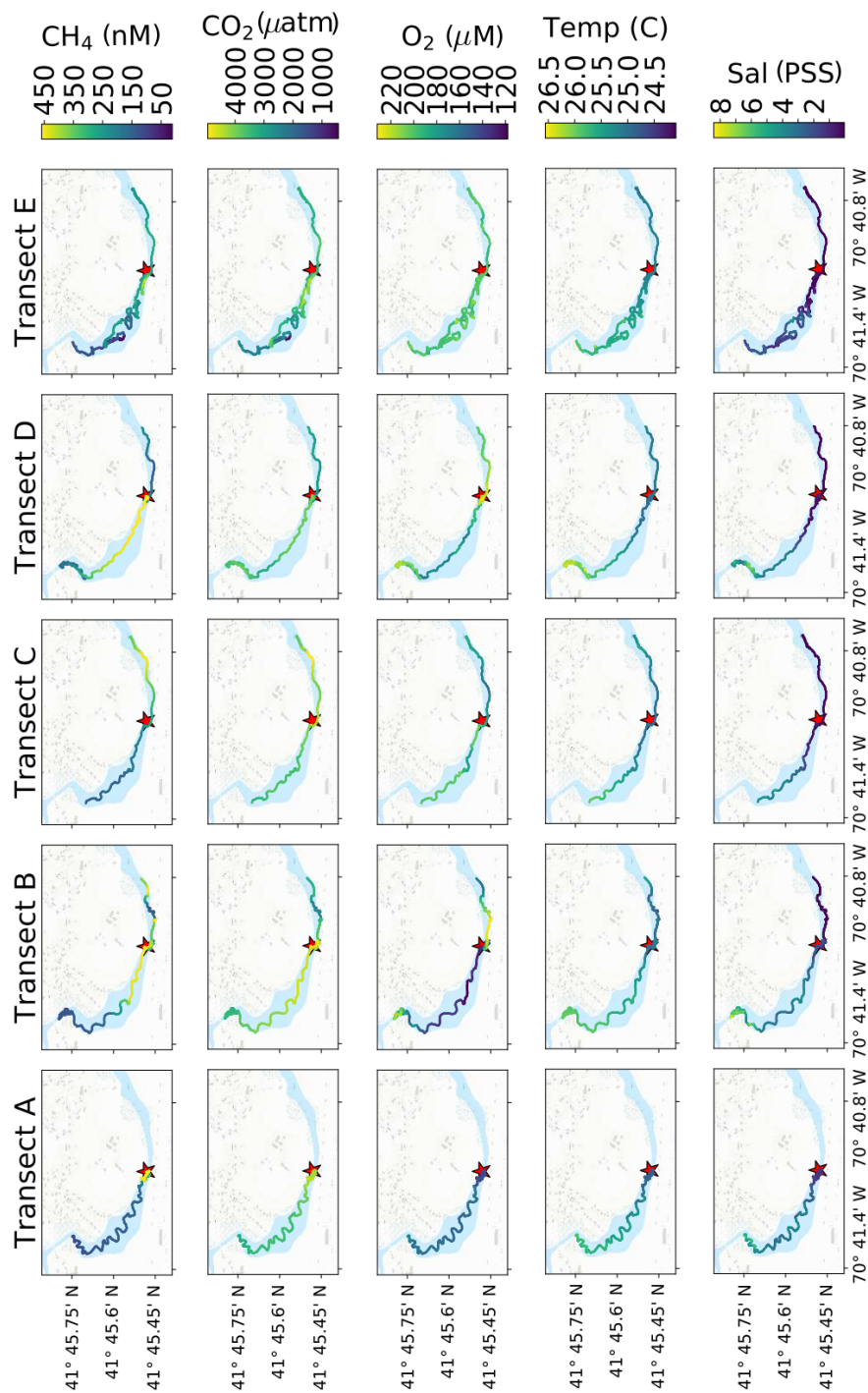


Figure 3-3: **Wareham River ChemYak transects** Individual transects are plotted on a representative map of the field site, colored by observed attributes. A red star marks the site of the outfall. Generally, elevated CH₄ and CO₂ are evident near the outfall and pulled toward Buzzards Bay; likely driven by the falling tide. These regions are also generally associated with lower O₂ readings. A salinity and temperature gradient is also evident, with saltier, warmer water present towards Buzzards Bay. Fresher, colder water appears associated with elevated greenhouse gases, however, the relationship is not necessarily strict.

A salinity and temperature gradient is evident, with saltier, warmer water at the far extent and fresher, colder water near the outfall. Fig. 3-4 shows the relationship between salinity, temperature, and gas concentrations. It appears that water with elevated greenhouse gas content tends to be colder and fresher. From the salinity and temperature plots alone, it is difficult to distinguish treated water from natural fresh water that is being pulled out of the deeper estuary by the tide. The distinguishing feature is the presence of CH_4 and to some extent CO_2 , which both tend to be found downstream of the outfall, except during slack tide, in which an accumulation upstream of gases seems to occur.

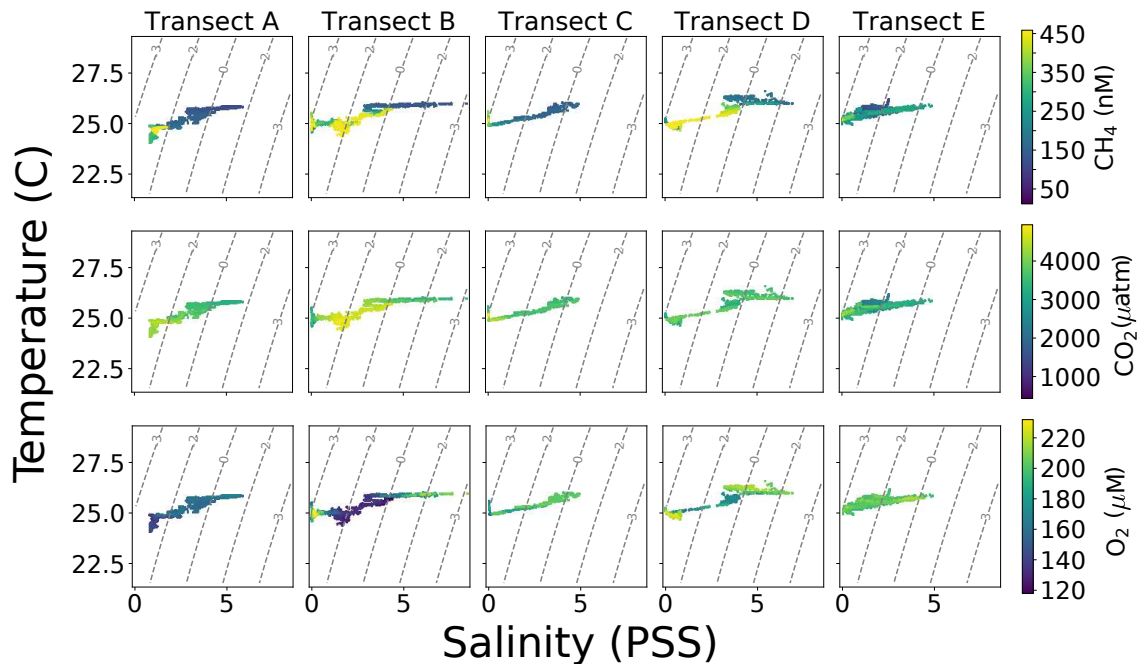


Figure 3-4: Wareham River salinity, temperature, and gas relationships
 These ST-plots highlight the correspondence between fresh, cold water and elevated CH_4 and CO_2 . In general, the field site has an established salinity and temperature gradient, with saltier, warmer water towards Buzzards Bay, and fresher, colder water near the outfall. Water with elevated greenhouse gas content appears to be present in colder, fresher waters, which is partially to be expected from a wastewater treatment plant. It is difficult, however, to distinguish treated water from fresh estuary water.

Fig. 3-5 captures the spatio-temporal variation evident during the field trials, with orange bars indicating points away from Buzzards Bay (upstream), and blue bars indicating points towards Buzzards Bay (downstream). Generally, the temperature

of the region remained stable, with little variation. A temporally variable salinity gradient was measured, with points upstream generally being the freshest observed, and downstream points having higher salinity. The effect of the falling tide is evident; as water is being pulled from upstream, the salinity levels towards Buzzards Bay seems to decrease as freshwater flushes the area. The gas content of the river demonstrates a more complex behavior relative to the tides.

At the end of high-tide (transects A-B), CH_4 is generally stable at the outfall and downstream, but upstream, the concentration of CH_4 grows; at slack tide, CH_4 levels detected upstream of the outfall are the highest observed during the campaign, approximately 450 nM. Similarly, CO_2 concentration also appears to become elevated upstream during slack tide, and O_2 correspondingly seems depleted. As the tide falls, represented by transects D and E, flushing of upstream water occurs, showing rapid decrease of CH_4 and CO_2 concentrations in upstream waters. For CH_4 , concentrations during the flushing event appear to become elevated downstream; but a similar trend does not hold for CO_2 and O_2 . With respect to CH_4 , there appears to be little loss from upstream concentrations to downstream concentrations during the flushing event; in Transect C, 425 nM was recorded in upstream waters 250 m from the outfall, and in transect D, 400 nM was recorded 450 m downstream from the outfall. However, over 600 m away from the outfall, the ChemYak consistently observed concentrations of CH_4 around 100 nM. During the flushing event, CO_2 and O_2 appear to generally decrease and increase, respectively, across the entire field site.

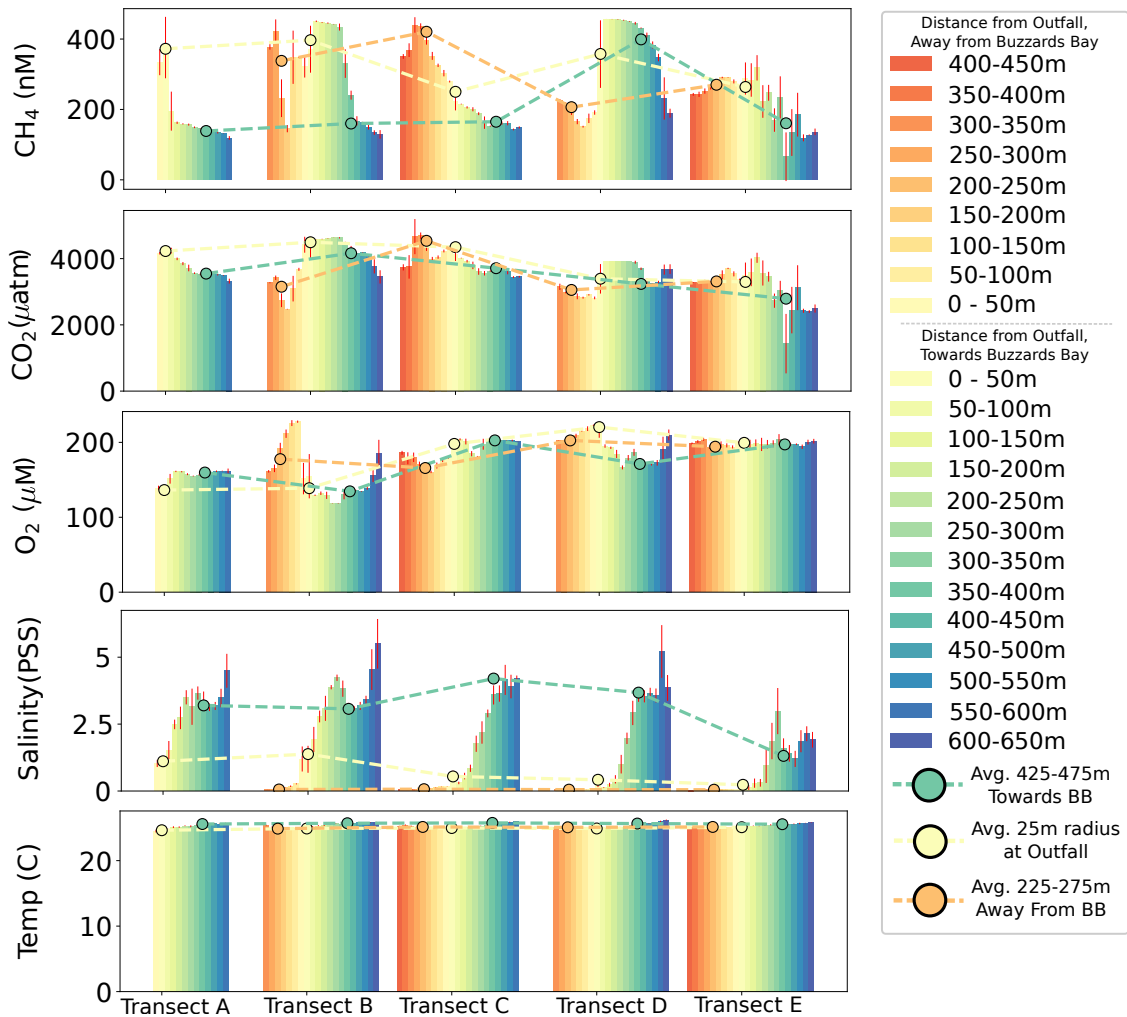


Figure 3-5: **Wareham River spatio-temporal trends** Waters away from Buzzards Bay (upstream) are indicated by orange gradients and waters towards Buzzards Bay are indicated by blue gradients. Averages at 450m downstream, at the outfall, and 250m upstream are shown to highlight trends. Red bars indicate the standard deviation of the measurements. Generally, CO_2 and O_2 demonstrate consistent decreasing and increasing trends, respectively. The salinity downstream decreases as fresh waters from upstream flush the estuary, pulled by the falling tide. Temperature remains generally stable throughout the field site. CH_4 demonstrates a complex spatio-temporal trend that appears indicative of “accumulation” during the end of rising tide and slack tide, and then “flushing” during the falling tide.

3.3.3 Significance and Role of Transience

Together, these observations show that longitudinal transport of CH_4 driven by tidal flushing events occurs at the Wareham River. CH_4 appears to be derived from waters associated with wastewater effluent, which may also contain elevated CO_2 . Rapid outgassing or metabolization of CH_4 may occur within the river, as indicated by rapid decrease of CH_4 concentration towards saltier, warmer waters. In this environment, transience is evident on the order of minutes and hours, associated with the tides and the source of effluent that is constantly pumping into the estuary. Characterizing this transience with sparse bottle samples alone would clearly be challenging, or impossible. The ChemYak, performing simple trajectories along the field site, was able to examine some of these dynamics in coarse detail, and particularly capture a slack tide “accumulation” event, followed by a “flushing” event of the estuary.

Future work in the estuary may include extending the field site to better constrain the extent of lateral CH_4 transport, collecting complimentary bottle samples to confirm (through isotope analysis) the methane sources (known as end-members) present in the sampling region, and quantifying the impact of biological CH_4 consumption in the region versus physical outgassing/venting. Further, observing several full tidal cycles may help to characterize the accumulation-flushing event in the estuary, and aid in the development of models of the region that could be ultimately used to inform pumping policies and new treatment interventions at the facility.

3.4 Methane Bubble Plumes in the Cascadia Margin

Methane seeps, also known as cold seeps, are found throughout the ocean [169,170] at continental margins, geologically active sites (e.g., mud volcanoes), and in hydrate fields. Although a vast number of sites have been identified, the contribution of methane seeps to the global atmospheric carbon budget is poorly constrained. Few measurements of methane bubble plumes from seeps have been recorded at the surface ocean [171]. Mechanisms that limit CH_4 emissions include biologic metabolization/oxidation and dissolution of bubbles at the seafloor, leaving only turbulent

diffusion as a means for gas transport from the deep ocean [172,173]. Estimates of the contribution of CH₄ from deep seeps has been shown to be negligible, however, CH₄ from shallow seep sites, lakes, and reservoirs may contribute a non-negligible amount of atmospheric CH₄, as bubbles can reach the surface in coherent forms [172].

In the Cascadia Margin (Pacific Ocean, Northern California, USA to Vancouver Island, Canada), CH₄ enters the water column from destabilized hydrate fields, and stable hydrates that are exposed on the seafloor and oxidized by bacterial colonies [174,175]. Numerous bubble plumes have been recorded in Cascadia Margin through acoustic studies from depths of 104-2073 m [176,177]. Plumes found along Hydrate Ridge at depths greater than 500 m have been observed using acoustic imaging to remain coherent to approximately 460 m before dissolution [178], and further studies have shown that elevated concentrations of CH₄ have been observed in water samples around 200 m at these sites [179]. The majority of closely studied seep sites have been found at the boundary of the hydrate stability zone, at depths between 600-400 m [177], where negligible flux has been estimated to enter the atmosphere [173,179]. However, shallow plumes sites stand to significantly influence the overall contribution of CH₄ from the Cascadia Margin. Small CH₄ supersaturations have been shown to be expressed at the surface of waters along Hydrate Ridge in addition to elevated surface concentrations near the coast due to upwelling [179,180]. Coastal upwelling may be a primary driver of atmospheric CH₄ contributions from the Cascadia Margin, which has been shown to pull waters from as deep as 100-200 m to the surface. Given shallow plumes in the coastal regions, in addition to deep sea plume penetration to 200 m, it is likely that CH₄ from shallow seep sources is reaching the surface.

Traditional methods of measuring CH₄ in the water column at sea usually employ CTD rosettes, which collect water samples for *ex situ* analysis at predetermined depths. In a single field campaign, the number of CTD casts that can be conducted is relatively small compared to the size of the entire region, and the sparsity makes resolving the distribution of CH₄ in surface layers difficult. Acoustic surveys can identify expressions of seeps on the seafloor, which can help to ultimately target locations for CTD casts, however, surface expressions of the gas are still unresolved.

Using a ChemYak surface vehicle, in conjunction with acoustic surveys and CTD casts, this field campaign investigated representative shallow bubble plumes in the Cascadia Margin, by fully resolving the water column from seafloor to the sea surface. The results from the ChemYak deployments are discussed in detail.

3.4.1 Overview of Field Work and Analysis

In September 2018, field work in the Cascadia Margin was conducted onboard the Schmidt Ocean Institute's R/V Falkor (FK 180824; Hunting Bubbles Cruise). Two representative sites are presented: Yachats (44° 21'N, 124° 10'W) and Stonewall Bank (44° 27'N, 124° 16'W). Both are shallow plume sites, with average depths of 46 m and 68 m respectively.

Shipboard multibeam sonar was used to produce bathymetric maps of the seafloor and look for bubble plumes in the mid-water. From these, scatter anomalies (known as flares) were used to identify and localize bubble seep candidates. CTD rosettes were subsequently cast over-the-side of the vessel in strategic locations to determine whether elevated CH₄ was present in the water column. Typically, CH₄ rich and hypoxic waters were observed from the depth of the cast site to 15-20 m. The ChemYak was launched from the ship in locations with seep activity and elevated concentrations of CH₄ at depth in order to observe fine spatial features in the top 10 m of the water column. The ChemYak was equipped with the gas analyzer, CTD, and oxygen optode as in the previous campaign, and the profiling winch was employed to change the depth of the CTD and inlet tube of the gas analyzer over the course of the mission. A secondary watercraft was deployed to monitor the ChemYak in the open ocean, and to allow a human pilot to better observe sea-state conditions while operating the vehicle.

CH₄ measurements were converted to molar units using an empirical extraction efficiency calculated by inspection of the distribution of measurements gathered, shown in Fig. 3-6. The median of the distribution of all measurements was compared with the expected equilibrium concentration of CH₄ (1.86 ppm, the global atmospheric level of CH₄ in September 2018) to calculate the efficiency correction necessary to

make them equivalent. For the salinity and temperature at the test sites (32.96 PSS and 12.7°C respectively), 1.86 ppm corresponds to a 2.7 nM concentration of CH₄. As a result of this analysis, an efficiency correction of 5.2% was applied. Across the two representative sites, over 16,500 measurements were recorded, and over 20 km traveled by the ChemYak.

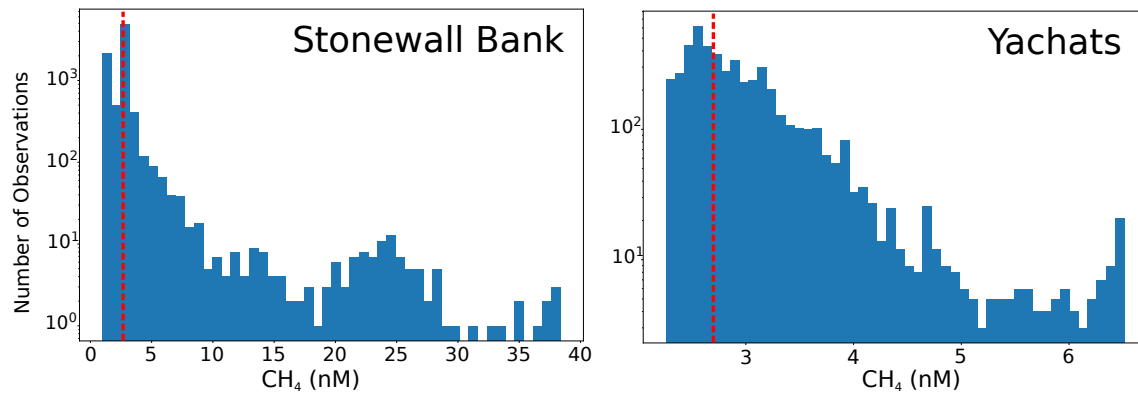


Figure 3-6: **Distribution of corrected observations in Cascadia Margin** To determine the extraction efficiency for the field campaign, the distribution of the CH₄ measurements collected were inspected, and the median concentration set to the hypothesized equilibrium value, 2.7 nM. The equilibrium value is calculated using 1.86 ppm as the atmospheric concentration of CH₄ at the time of sampling, and the average salinity (32.96 PSS) and temperature (12.7°C) between both test regions. The fit of the data for both test sites is shown, where the red dashed line indicates 2.7 nM.

3.4.2 Results

The ChemYak observed elevated CH₄ levels in the surface layers at both representative sites coupled with low-oxygen conditions. Fig. 3-7 and Fig. 3-8 show the geolocated measurements observed by the ChemYak at Stonewall Bank and Yachats, respectively. At Stonewall Bank, up to 35 nM at 10m depth was observed in highly localized, approximately 25m diameter pockets. To the southeast of the measurement site, elevated CH₄ of 3-4 nM was generally observed throughout the top 10m of the water column. Three bubble plumes were detected with the multibeam sonar in the field site; and a CTD cast was performed near the southern most plume site. It is generally difficult to correlate any of the small supersaturations with a particular plume source,

however the presence of plumes in general indicates that there is CH₄ rich and hypoxic waters at depth (which is confirmed by the CTD cast). Low-oxygen conditions at the surface are generally correlated with elevated CH₄ concentrations. Fig. 3-9a shows a tentative relationship between O₂ and CH₄, in which more CH₄ rich waters appear to have slightly lower O₂ content (O₂: 380 μM, CH₄: 3-7 nM) than waters at CH₄ equilibrium (O₂: 387 μM, CH₄: <2.7 nM).

In comparison, CH₄ was found in less supersaturated pockets at Yachats, although localized pockets of 3-7 nM were observed, with a diameter of approximately 50-100 m. The multibeam observed tens of seeps at the seafloor at this location, however it is difficult to distinguish a spatial relationship between these detections and elevated surface expressions. The O₂ and CH₄ at Yachats does not show a clear correlation — most surface CH₄ detections were at equilibrium — however, a potentially anomalous positive correlation between O₂ and CH₄ can be seen for CH₄ concentrations from 3.5-6.5 nM.

A summary of the relationship between depth and CH₄ observed at both sites is provided in Fig. 3-10. At Stonewall Bank, significantly elevated CH₄ was observed at approximately 8 m depth, with occasional peaks between 2-6 m. In the top meter, concentrations of > 7 nM (nearly 3 times the equilibrium concentration) was also observed. Throughout the water column, the average observed CH₄ concentration was distributed around the equilibrium; only at 8 m did the average concentration vary significantly from the equilibrium to 10 nM. At Yachats, elevated CH₄ was observed throughout the water column, with peaks between 2-8 m. In general, the average CH₄ observed at various depths in Yachats was skewed slightly above the equilibrium value (approximately 3.1 nM). At both sites, the majority of observations fall between 2-4 nM.

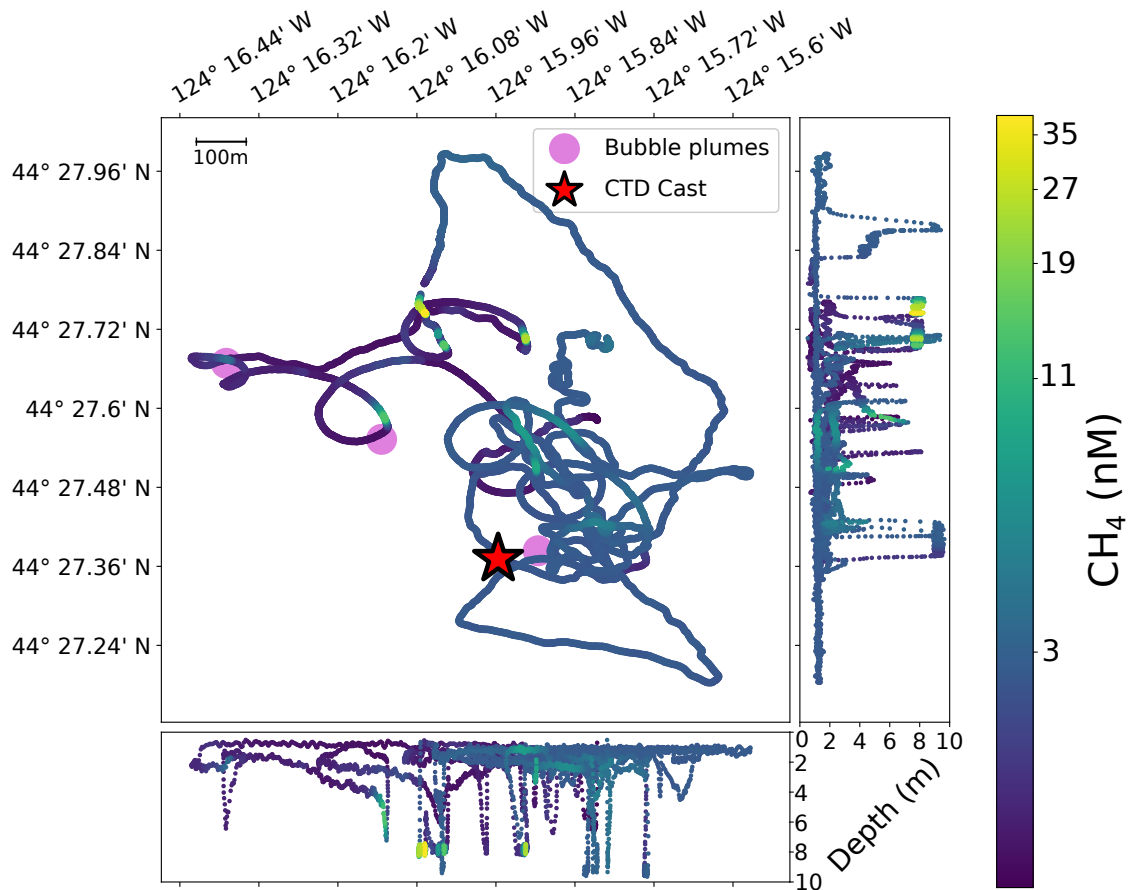


Figure 3-7: **ChemYak tracks at Stonewall Bank** The ChemYak was piloted by a human user given rough GPS coordinates of the bubble plumes (marked with magenta circles) and the CTD cast (marked with a red star). The map, center, shows an overhead view of the path the ChemYak took, and the two side panels at the bottom and right show the depth of the CTD probe and gas analyzer inlet for each sample. Significantly elevated CH_4 is found at several sites, with a peak at 38 nM detected at 8m. CH_4 concentration levels above 2.7 nM were largely observed to the north and southeast of the field site, at depths of 0.5-4m. There does not appear to be a direct spatial correlation between concentrated CH_4 expressions and discovered seafloor seeps. Note: log-scale used for CH_4 concentration.

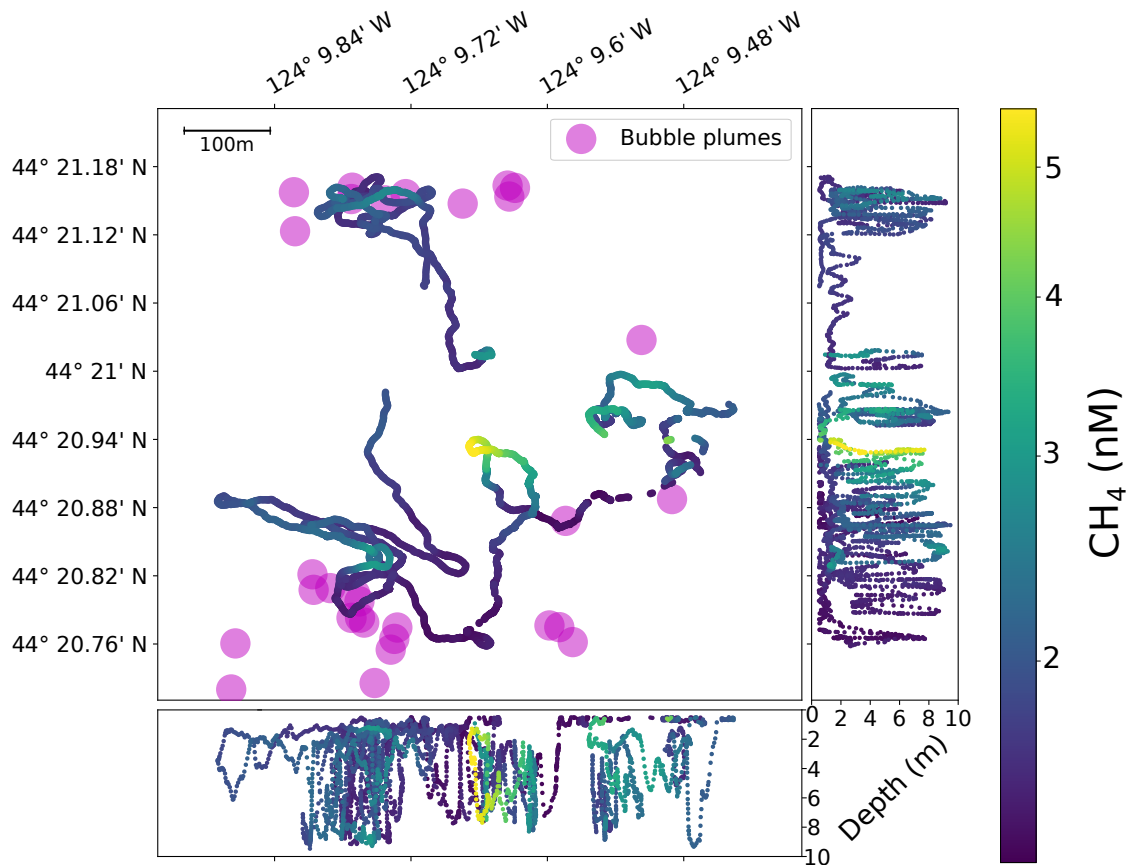


Figure 3-8: **ChemYak tracks at Yachats** The ChemYak was piloted by a human user given rough GPS coordinates of the bubble plumes (marked with magenta circles). The map, center, shows an overhead view of the path the ChemYak took, and the two side panels at the bottom and right show the depth of the CTD probe and gas analyzer inlet for each sample. Relatively little CH₄ elevation was observed at this site as a whole, however several concentrated expressions of 3-7 nM were observed at depths up to the top 1-2m.

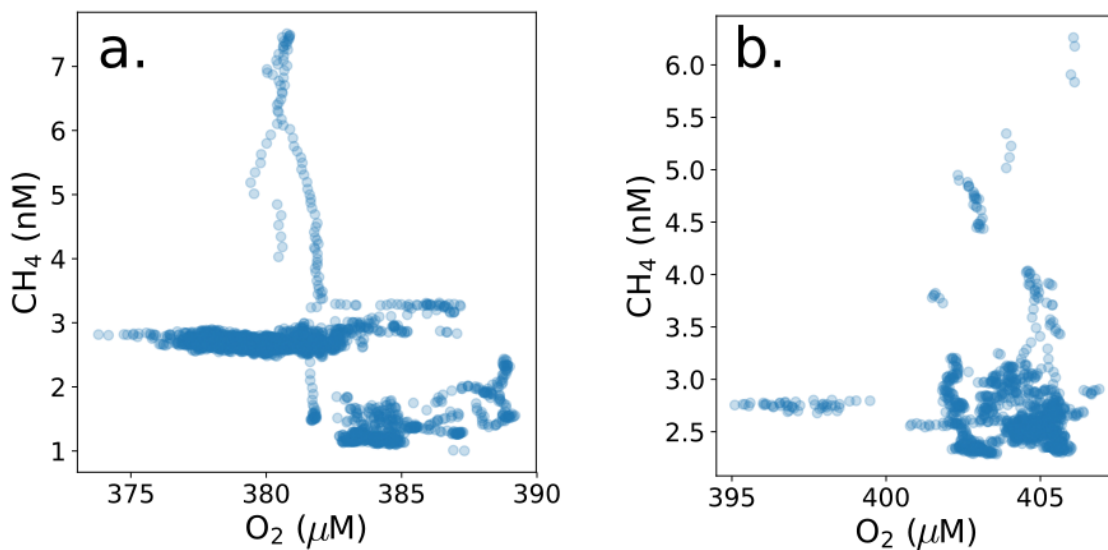


Figure 3-9: **Cascadia Margin surface methane and oxygen** Measurements observed in surface (0.5-1.5 m) waters at (a) Stonewall Bank and (b) Yachats are plotted to compare oxygen and methane. At Stonewall Bank, there are distinct “clusters” shown in the data, which roughly correspond to elevated CH_4 -low O_2 waters and waters that are near CH_4 equilibrium and show higher O_2 content. In Yachats, there is no clear correspondence between CH_4 and O_2 , although an interesting and potentially anomalous rise in O_2 and CH_4 levels appears to be present in some portion of the field site.

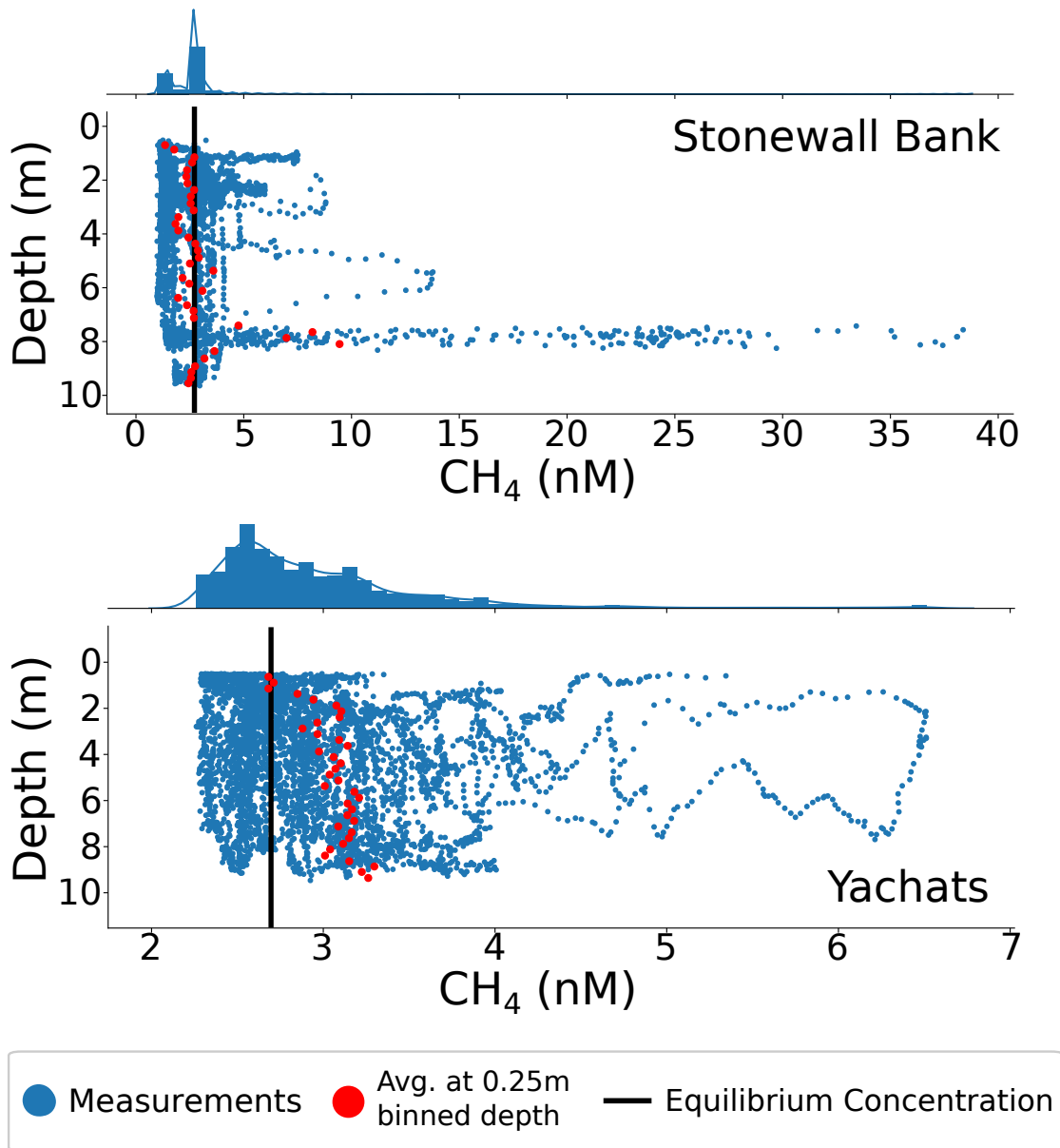


Figure 3-10: **Cascadia Margin methane depth profiles** Measurements from the ChemYak field campaigns at Stonewall Bank and Yachats plotted as CH₄ versus depth, with average CH₄ values at depth binned by 0.25 m increments marked in red, and 2.7 nM equilibrium marked with a black line. A representative density of CH₄ measurements is provided above the scatter plot. At Stonewall Bank, significantly elevated CH₄ was observed at 8 m. Both sites show elevated CH₄ concentrations up to 8 nM in the top meter of the water column, with the majority of observations falling between 2-4 nM.

3.4.3 Significance and Role of Transience

These preliminary results indicate that CH_4 from shallow methane seeps may be reaching the top 10 m of the water column. At Stonewall Bank, evidence of upwelling (elevated CH_4 and depleted O_2) and highly local supersaturated expressions were observed. At Yachats, local supersaturations were observed, and measurements throughout the water column were skewed slightly above the equilibrium value (although O_2 content did not demonstrate a strong relationship).

To get a sense of the potential impact of shallow seeps as a methane source, a conservative estimate for CH_4 flux is presented. The total area of Stonewall Bank and Yachats observed by the ChemYak was approximately 2.87 km^2 . The mean concentration of CH_4 in the waters was observed to be approximately $2.1 \mu\text{atm}$, with a maximum of $4.45 \mu\text{atm}$. During the field trials, a sustained 10 m s^{-1} wind was observed, and is assumed to be a representative measure of wind-speed throughout the year. The atmospheric concentration of CH_4 is given at $1.86 \mu\text{atm}$, with average salinity (32.96 PSS) and temperature (12.7°C) taken from the trials. Under these conditions, a conservative average positive flux for the region is $0.0006 \text{ mol m}^{-2} \text{ y}^{-1}$ with upper bound of $0.0056 \text{ mol m}^{-2} \text{ y}^{-1}$. This ultimately implies a contribution of 29-259 kg y^{-1} of CH_4 from the study sites to the atmosphere. Further work would be required to better constrain this estimate, and a scaling exercise could be applied to estimate the total contribution of shallow seeps to the global carbon budget.

Tracing bubble plumes directly through the water column is a considerable task: ocean currents, the effect of the halocline and mixing boundaries, waves, and atmospheric conditions all impact the journey of CH_4 from sea floor to the atmosphere. In minutes, we found that the concentration of a previously sampled region would be measured by the ChemYak to have a different gas character. Although seep locations were known, there was no clear correlation between sources and surface expressions, likely due to ocean current and surface perturbation by wind and waves. This field campaign motivates future work in the Cascadia Margin and at shallow methane seeps with *in situ* equipment to densely survey and map surface expressions of CH_4 , and

developing regimes that seek and densely sample supersaturation pockets. Additionally, underwater technologies could be used to target and examine methane bubble release events. The outcome of this and future work could ultimately be used to revisit the methane and carbon budget of ocean sources to the atmosphere.

3.5 Spring Freshet River Inflow in Arctic Estuary

Global climate change has significantly impacted freeze-thaw cycles in the Arctic; the ice-free season has lengthened and once permanently frozen land now thaws seasonally [148,181]. Seasonal thawing generally has been associated with outgassing of CH₄, CO₂, and other greenhouse gases, like nitrous oxide (N₂O) [148,152,153,182,183]. Previous studies of Arctic ice sheets, lakes, estuaries, and coastal waters have shown that gas emissions are transient, and vary with the seasons [148,152,153,182,184,185]. However, published measurements in the Arctic are significantly skewed towards summertime measurements under open-water conditions. In order to better constrain the CH₄ and CO₂ budget of coastal Arctic waters, characterizing short-duration outgassing events, such as during the spring freshet, is critical.

For this study, Cambridge Bay, Nunavut, Canada (Fig. 3-11) served as a representative field site in the North American high Arctic to perform dense spatial mapping with the ChemYak. Cambridge Bay (69°07'N, 105°03'W) is a hamlet on the southeastern coast of Victoria Island in the Kitikmeot Region. Generally, the island environment is sedge-moss-meadow and polar semi-desert [186]. Cambridge Bay harbor is fed by the Freshwater Creek estuary; the major riverine influence in the region is Freshwater Creek, which receives discharge from Greiner Lake to the Northeast of the hamlet. The lake, river, and harbor freeze annually, with the estuary freezing to approximately 1.5 m thick by late Spring. Ice acts as a cap for the water body, trapping gases underneath it and preventing air-sea interactions [153,184]. During seasonal ice melt, rapid re-equilibration occurs as the air-sea interface is restored. In Cambridge Bay, river discharge begins in June, leading to the start of ice-free season in late-June through early-July.

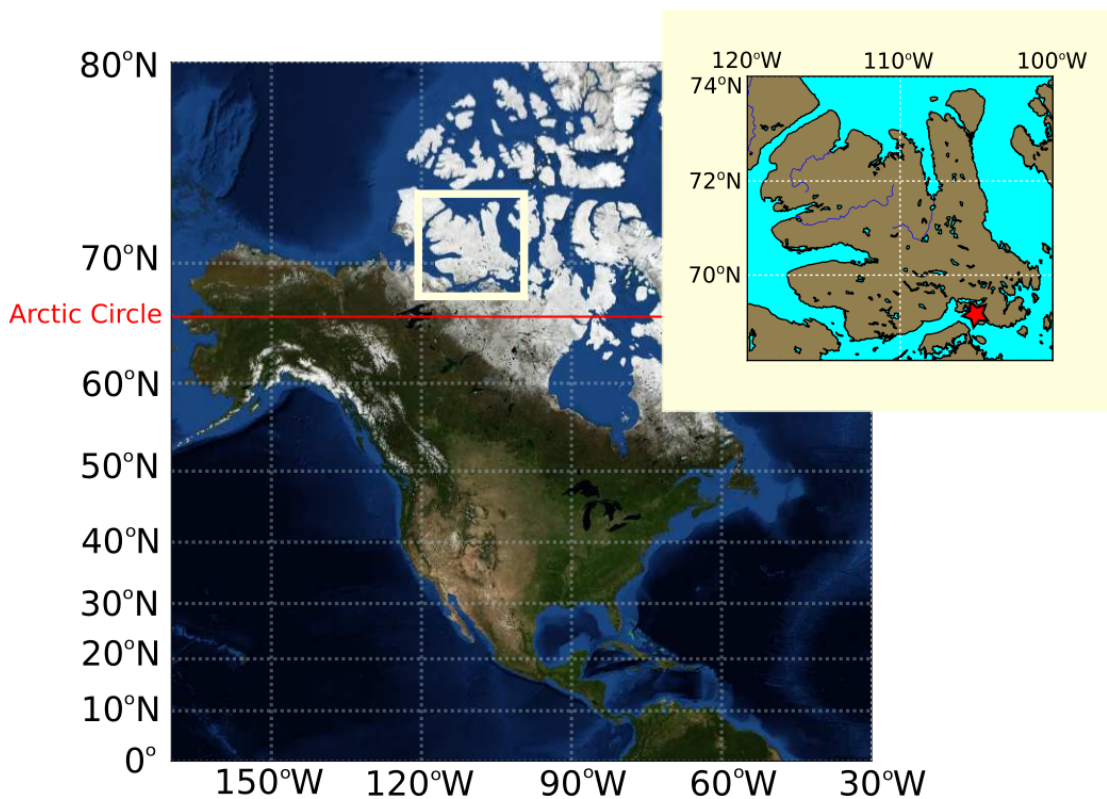


Figure 3-11: **Cambridge Bay reference map** This map shows the location of Cambridge Bay with respect to North America. Cambridge Bay is found on the southeastern coast of Victoria Island (highlighted), which lies about the Arctic circle. In the inset, the location of Cambridge Bay on the island is marked with a red star.

A year-long series of bottle samples from sampling station B1 (69.107556°N, 105.059667°W) and from Freshwater Creek (69.12975°N, 104.99459°W) were collected from Cambridge Bay in 2017-18, consisting of over 30 measurements. These samples indicated that shortly before the harbor became ice-free, the under-ice measurements of CH₄ in Freshwater Creek peaked at 240,000% saturation and in the estuary to 19,000% saturation in late-June and then decreased by over a factor of 100 within four weeks, indicating rapid air-sea ventilation as the ice cover receded. The preliminary measurements indicated a significant and unprecedented outgassing period, however, the temporal resolution of the samples left open questions about the rate of the re-equilibration within the estuary and lateral transport of CH₄ to the coastal ocean. In order to characterize the fine-scale temporal dynamics just as ice melt begins, it is necessary to employ *in situ* instruments for rapid and dense collection.

3.5.1 Overview of Field Work and Analysis

The ChemYak was deployed for a 5-day campaign in late-June through early-July of 2018 in recently ice-free waters in the Freshwater Creek estuary. From 28 June through 02 July, the ChemYak was piloted by a human user from the receding ice edge in Cambridge Bay harbor to the mouth of Freshwater Creek. In addition to the main estuary, a small embayment near the mouth of the river was also surveyed on several days of trials (29 June, 01 July, 02 July). This embayment was fed by a much smaller, minor river. The ChemYak traveled an average of 11 ± 2 km each day, typically between the hours of 15:00-22:00 UTC and over 75,000 observations were collected. The greenhouse gas analyzer and CTD were made available for this study, and the profiling winch was used to make measurements above and below the mixing layer in the estuary. A representative map of the study region and daily trajectories followed by the ChemYak are shown in Fig. 3-12.

An empirical extraction efficiency correction for CH_4 and CO_2 measurements was determined using bottle samples that were collected contemporaneously to ChemYak measurements for each day of the campaign. To compare with bottle samples, ChemYak measurements were sorted by the day the sample was taken, depth proximity to the point at which the Niskin collection bottle was fired (± 0.38 m), and geographic proximity (within 50 m radius) of the logged bottle sample location. An extraction efficiency that allowed for a linear fit of slope 1 between ChemYak measurements and bottle samples was then selected. The outcome of the calibration is shown in Fig. 3-13. An empirical extraction efficiency of 5.09% for CH_4 ($r^2=0.842$) and 50.5% ($r^2=0.717$) for CO_2 was found. One bottle sample was eliminated from the calibration as the temperature/salinity measured in the bottle did not match the temperature and salinity profile with the associated ChemYak measurements; likely caused by a misfiring of the Niskin sample bottle used.

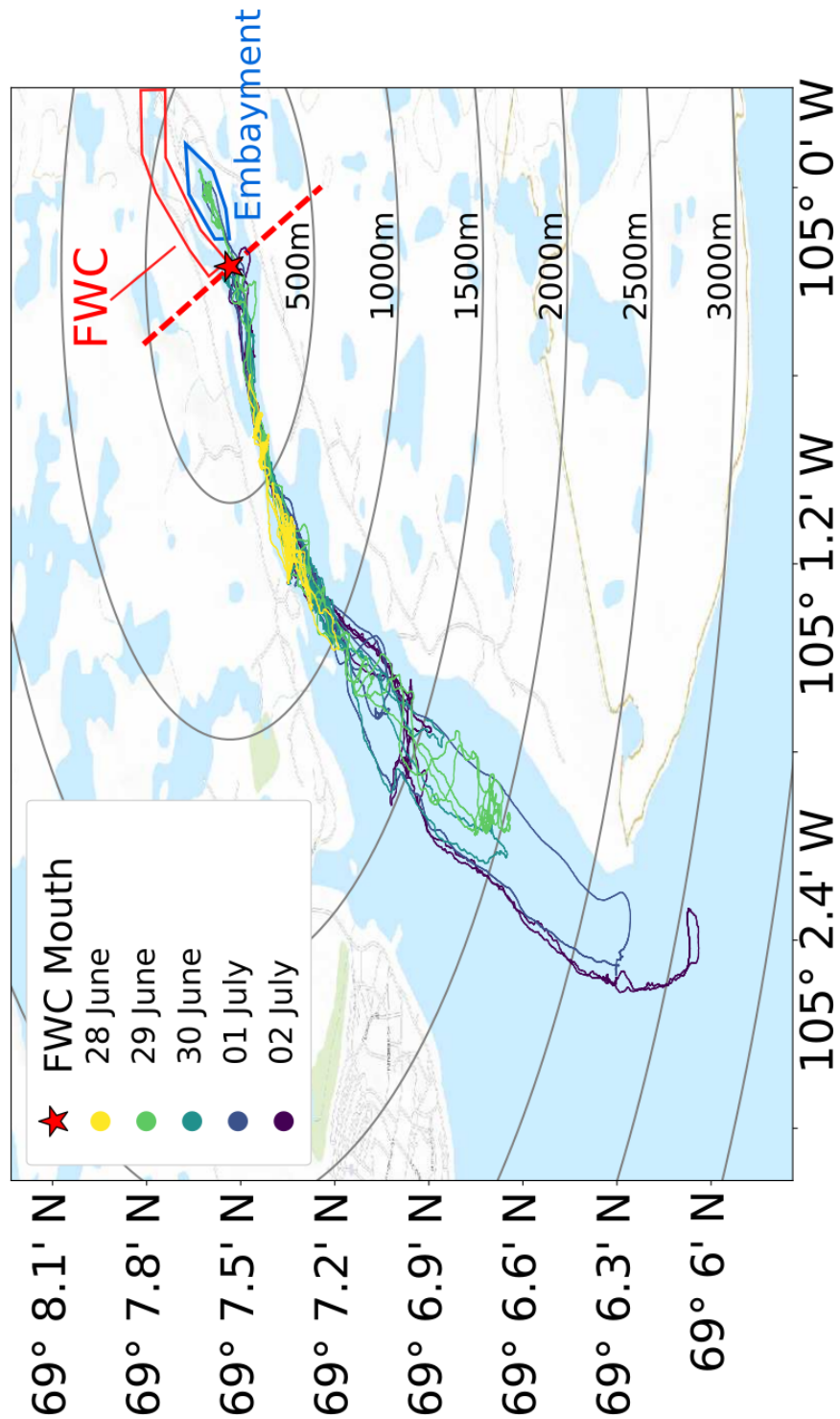


Figure 3-12: **Cambridge Bay field site and ChemYak tracks** The study site with ChemYak trajectories from each day overlaid. The mouth of Freshwater Creek (69.1257°N, 105.0042°W) is marked with a star, and concentric rings at increments of 500 m centered at the mouth are provided for scale. Northeast of the red dashed line lies Freshwater Creek (red box) and a small embayment (blue box) which receives input from a much smaller river.

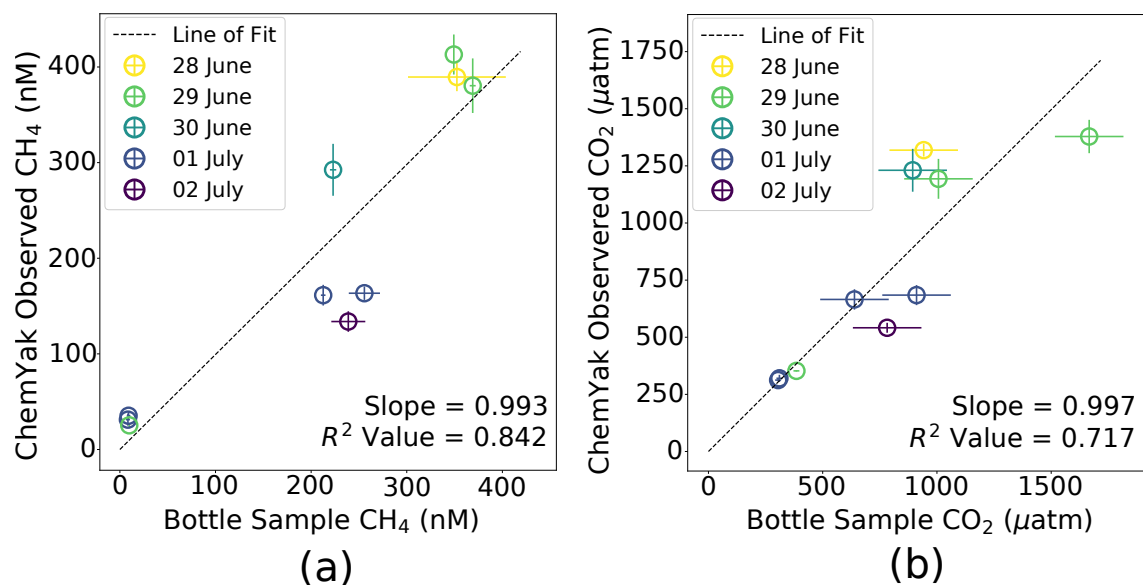


Figure 3-13: **Cambridge Bay calibration results for gas analyzer** The extraction efficiency used to scale ChemYak measurements was determined by comparing ChemYak observations with collected bottle samples. In the plots above, ChemYak observations of CH₄ are calculated using an extraction efficiency of 5.09% ($r^2=0.842$) and for CO₂ using an extraction efficiency of 50.5% ($r^2 = 0.717$). ChemYak measurements are representative of the mean (point coordinate) and standard deviation (error bar) of all observations filtered by location (<50 m radius), depth (± 0.38 m), and day, as recorded for the bottle sample. Error bars reported for bottle samples are determined by processed replicates for CH₄ and uncertainty in total alkalinity, dissolved inorganic carbon, salinity, temperature, and carbonate equilibrium constants for Arctic estuarine environments for CO₂.

3.5.2 Results

Over the course of the 5 day field campaign, the ChemYak revealed a strong pycnocline in the estuary that divided fresh, gas-rich water at the surface from salty, cold water. Fig. 3-14 shows the pycnocline at approximately 1.5-2.5 m over the campaign period. Using the mouth of Freshwater Creek as a reference point, elevated CH₄ and CO₂ levels were observed in the surface layers over 2 km towards the receding ice edge in the harbor. This figure also demonstrates the difference between the embayment and river-forced waters. The embayment is composed of very warm, fresh water with gas concentrations that are lower than those measured in nearby surface waters. For example, on 29 June, surface water at Freshwater Creek mouth was observed to have

concentrations of CH₄ around 417±31 nM and CO₂ around 1398±109 μ atm. On the same day, the embayment was observed to have concentrations of approximately 243±41 nM and 790±137 μ atm respectively. This difference illustrates that river inflow is likely driving significant lateral transport of gases into the estuary; as the embayment is somewhat protected from river forcing, the residence time of the water is longer and more gas exchange can occur than in the short-lived river forced waters. During the campaign, river inflow had an average discharge rate of (40 m³ s⁻¹), which would have refreshed the surface layer of the campaign region every 0.6 days.

To further examine the potential for lateral transport in the estuary, Fig. 3-15 shows the spatio-temporal trends of CH₄, CO₂, salinity, and temperature over the campaign in the top 1 m of the estuarine waters. In general, the distribution of CH₄ and CO₂ demonstrated a slight gradient each day, with higher concentrations at the mouth of Freshwater Creek and lower concentrations at the receding ice edge. The presence of a gradient can be indicative of several processes, including outgassing, vertical mixing, or biological oxidation of gas species during transport.

One key event that emerges from the surface data is a warm, outgassing period in the river. Between 01-02 July, the temperature at Freshwater Creek increases nearly a degree, and the observed gas content at the mouth is significantly reduced compared to other sampling days. This is further illustrated in Fig. 3-16 which shows salinity and temperature plots colored by gas concentrations; on 02 July a significant number of warm freshwater samples were collected with decreased gas concentration. Interestingly, the gradient on 02 July is reversed, with higher concentrations of gas observed toward the receding ice edge. This may be indicative of the transport time of waters in the estuary (i.e., during the sample period, the estuary had not been fully flushed by new river waters). A second key event occurred on 30 June, when elevated winds (approximately 10 m s⁻¹) were measured coming from the northeast. Compared to other regions of the estuary on the same day, gas concentrations at the ice edge (to the southwest) were comparably elevated, effectively creating a bimodal distribution with high concentrations at the Freshwater Creek mouth, and high concentrations at the ice edge, in the southwest. This may be indicative of surface waters being pushed

toward the ice edge by the wind, and the ice acting as a barrier for mixing, causing a brief period of accumulation.

Vertical mixing is generally evident by the salinity gradient in the surface waters, and examining the pycnocline highlights interesting spatial trends and ephemeral events in the estuary. As shown in Fig 3-17, regions near the mouth of the river were observed to have a less strict division between the fresh and saline layers. Especially with respect to gas concentrations, there is significantly elevated CH_4 and CO_2 levels as deep as 3-4 m near the mouth of the river, whereas 2-3 km from the river mouth elevated gas concentrations were more strictly observed above 2 m. The nature of the pycnocline at the river mouth may be indicative of turbulent mixing at the interface between the shallow river and the deeper estuarine waters. The mean location of the pycnocline shifted in the water column by approximately 1 m within the span of 1 sampling day (from 29 June to 30 June), which may be indicative of changes in river flow or vertical mixing from wind conditions on 30 June. At the ice edge, the pycnocline remained consistent, on average, between campaign days.

Fig. 3-18 summarizes the temporal trends over each day relative to measurement depth. CH_4 concentrations decreased from 410 ± 20 nM on 28 June to 150 ± 70 nM on 02 July, a rate of -55 nM d^{-1} . CO_2 similarly declined, from 1340 ± 40 μatm to 600 ± 150 μatm , a rate of -146 μatm d^{-1} . Although considerable losses, if river inflow were not a factor (i.e., the water was “stagnant”), the expected observations of CH_4 and CO_2 on 02 July would have been 70 nM and 570 μatm respectively, given the starting conditions on 28 June. Although observed CO_2 concentration nearly match this expectation, CH_4 observations were over double this estimate, indicating continued CH_4 transport by the river from Greiner Lake. At the pycnocline, gas concentrations also declined over the field campaign, and brief temperature and salinity perturbations can be seen on 30 July. The effects of these perturbations in gas concentration are most generally observed by an influx of CO_2 and (lesser, but still present) CH_4 in deeper waters. Outside of these perturbations, deeper waters generally demonstrated stable CH_4 and CO_2 levels, around 50 nM and 380 μatm respectively. These deeper waters became slightly fresher and warmer over the campaign period.

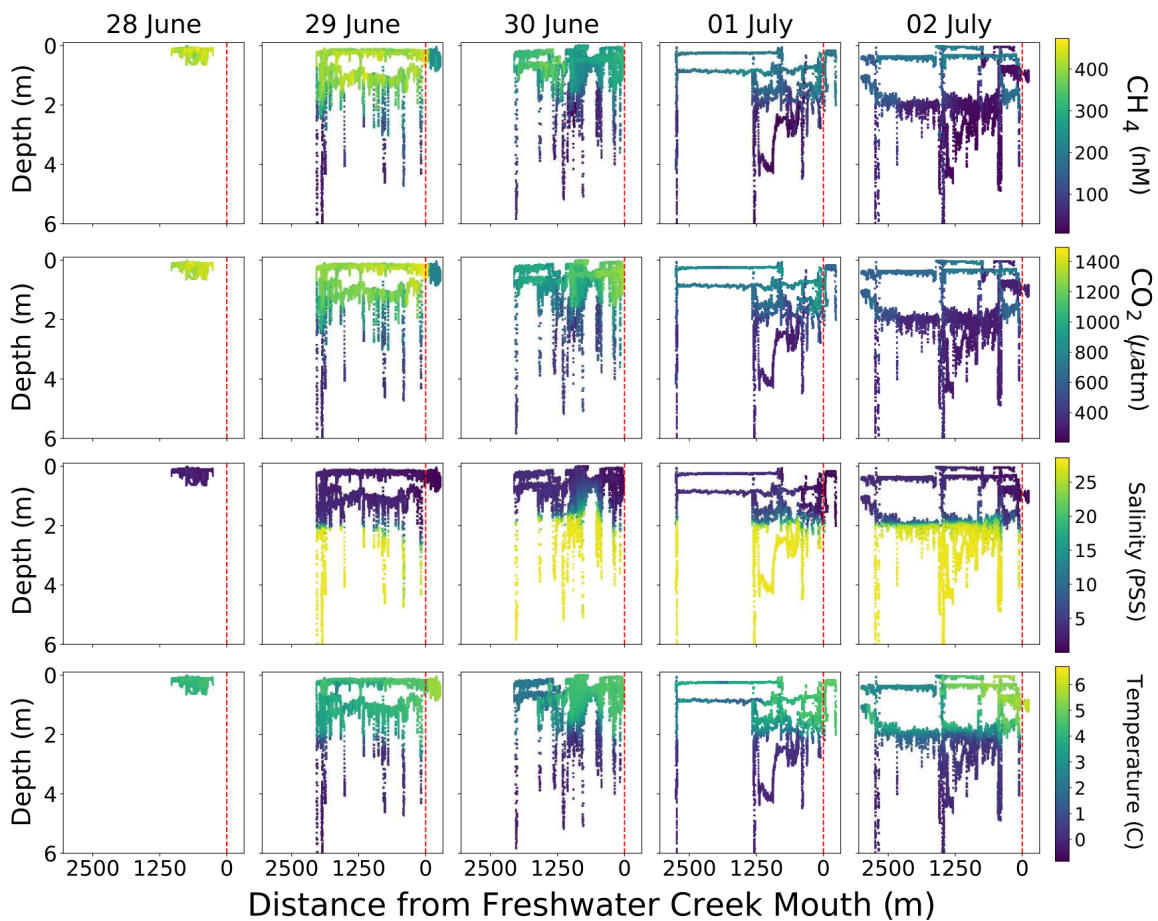


Figure 3-14: **Cambridge Bay depth profiles** Observations made by the ChemYak are plotted by depth versus distance from the Freshwater Creek mouth, where negative distances to the right of the red line represent points northeast of the mouth (a small embayment) and positive distance to the left of the red line represent points southwest of the mouth (downstream). As indicated by the salinity plots, the pycnocline falls between 1.5-2.5m throughout the estuary, and the fresh surface layer was generally higher in both CH₄ and CO₂ concentration than layers below of the pycnocline. The gas concentrations decreased over the multi-day measurement campaign.

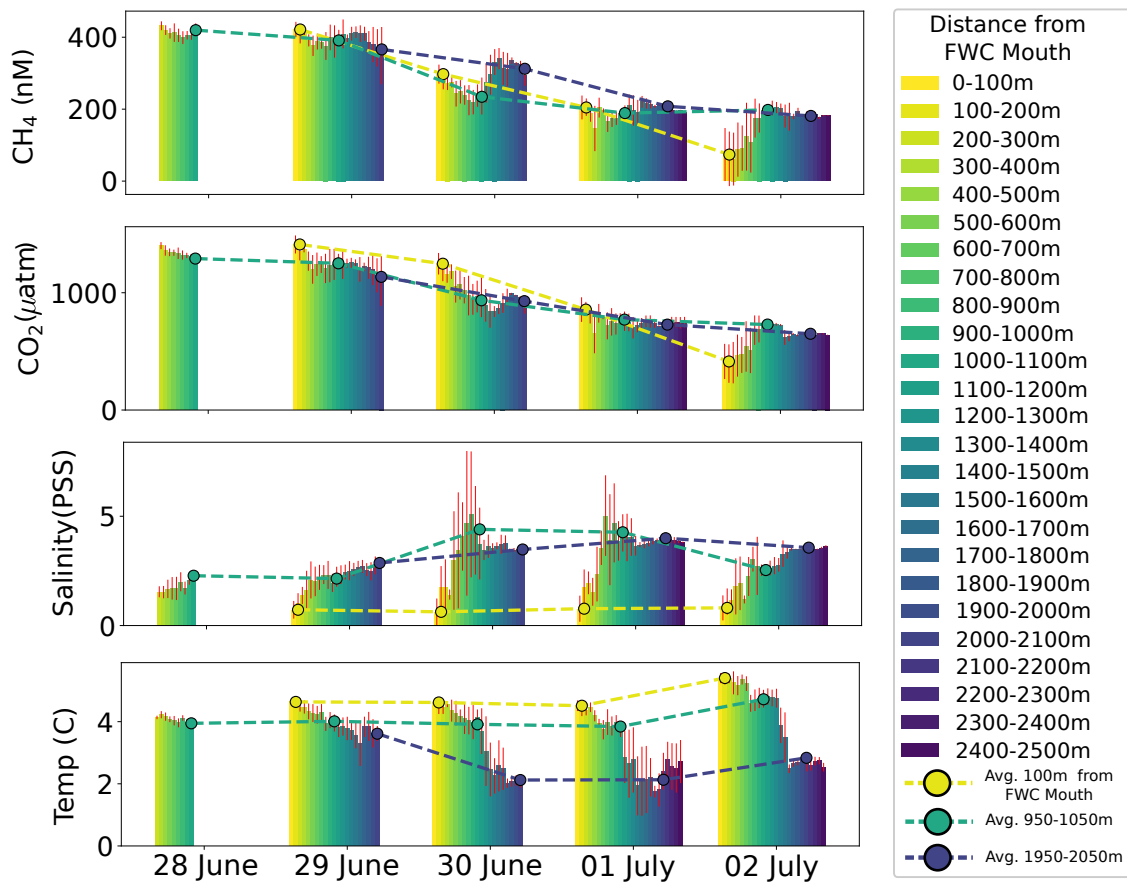


Figure 3-15: **Cambridge Bay spatio-temporal surface trends** Bar graphs showing CH₄, CO₂, salinity, and temperature in the top 1 m of the campaign area, colored by distance binned in 100 m intervals. Trend lines are added to show the average value of measurements in the surface layer at Freshwater Creek mouth (within 100 m), at 1 km towards the ice edge, and at 2 km towards the ice edge. Generally, a gas gradient exists at the surface, with higher concentrations at the mouth of the river and lower concentrations at the ice edge. However, this gradient is disturbed on 30 June with gas accumulation at the ice edge, potentially due to elevated winds from the north-east. On 02 July, the gradient is reversed, simultaneously the river water increased in temperature significantly, potentially indicative of warming outgassing. The high gas concentrations downstream may capture the residence time of water in the estuary (i.e., the river had not yet flushed the estuary). Vertical mixing is indicated by the salinity gradient in the surface waters.

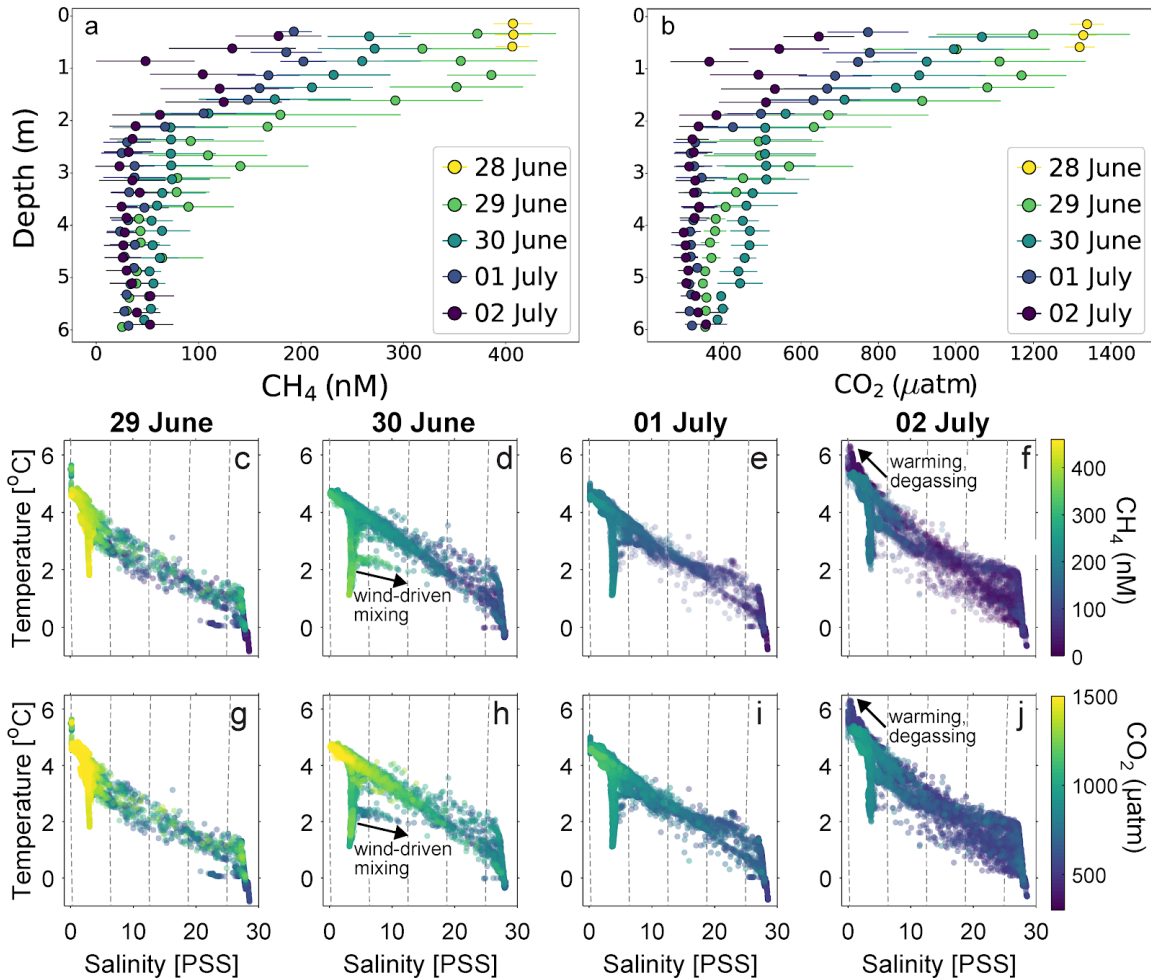


Figure 3-16: **Cambridge Bay salinity, temperature, and gas relationships** a-b. Each day of the measurement campaign is marked with a unique color, and samples collected are binned into 0.25 m increments from the surface to 6 m. Both (a) CH_4 and (b) CO_2 exhibit decreasing trends for each subsequent day, and there is strong stratification between the surface layer and water below 2 m. c-j. Temperature-salinity plots showing changes in (c-f) CH_4 and (g-j) CO_2 concentrations. Two events are captured in the data. One is wind-driven mixing on 30 June when cold, saline water from below the pycnocline mixed with gas-rich waters at the surface. The second is warming degassing on 02 July, when warm, fresh water without significantly elevated gas concentration is observed.

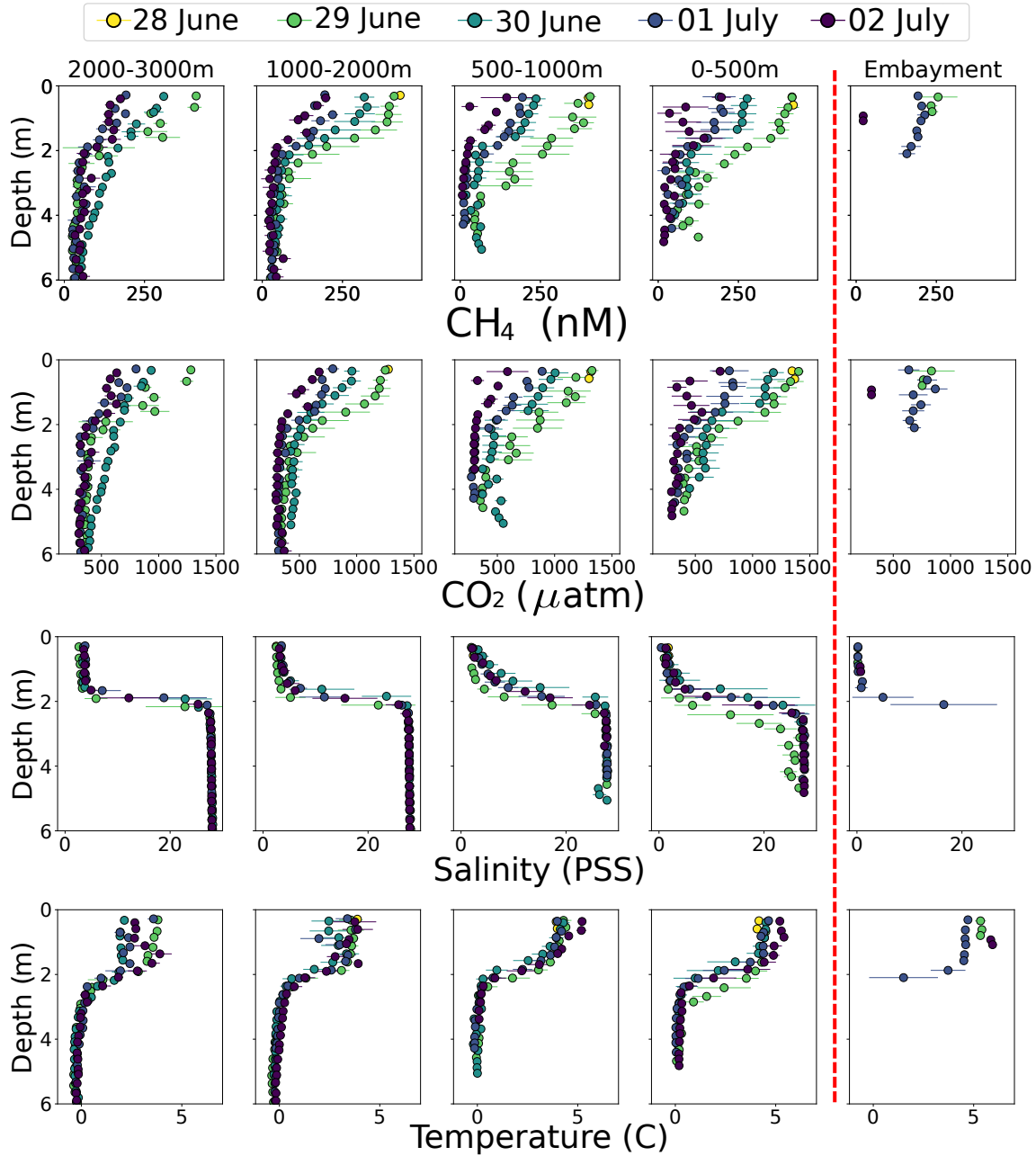


Figure 3-17: **Cambridge Bay depth profiles binned by estuary region** To observe regional variation in the pycnocline between the mouth of Freshwater Creek (0-500 m) the embayment to the northeast, and the ice edge downstream to the Southwest (>500 m), we bin the observations of the ChemYak by day (color), depth (0.25 m), and distance. The embayment was generally warm, fresh water that is uniformly low in CH_4 and CO_2 content in the three days it was observed (29 June, 1 July, and 2 July). At the river mouth, the pycnocline is less pronounced, with elevated gas concentrations observed up to 3-4 m deep. In contrast, towards the ice edge elevated gas content is largely in the top 2 m. Elevated CH_4 and CO_2 was evident below 2 m at the ice edge on 30 June compared to other days, likely due to increased wind speeds driving vertical mixing.

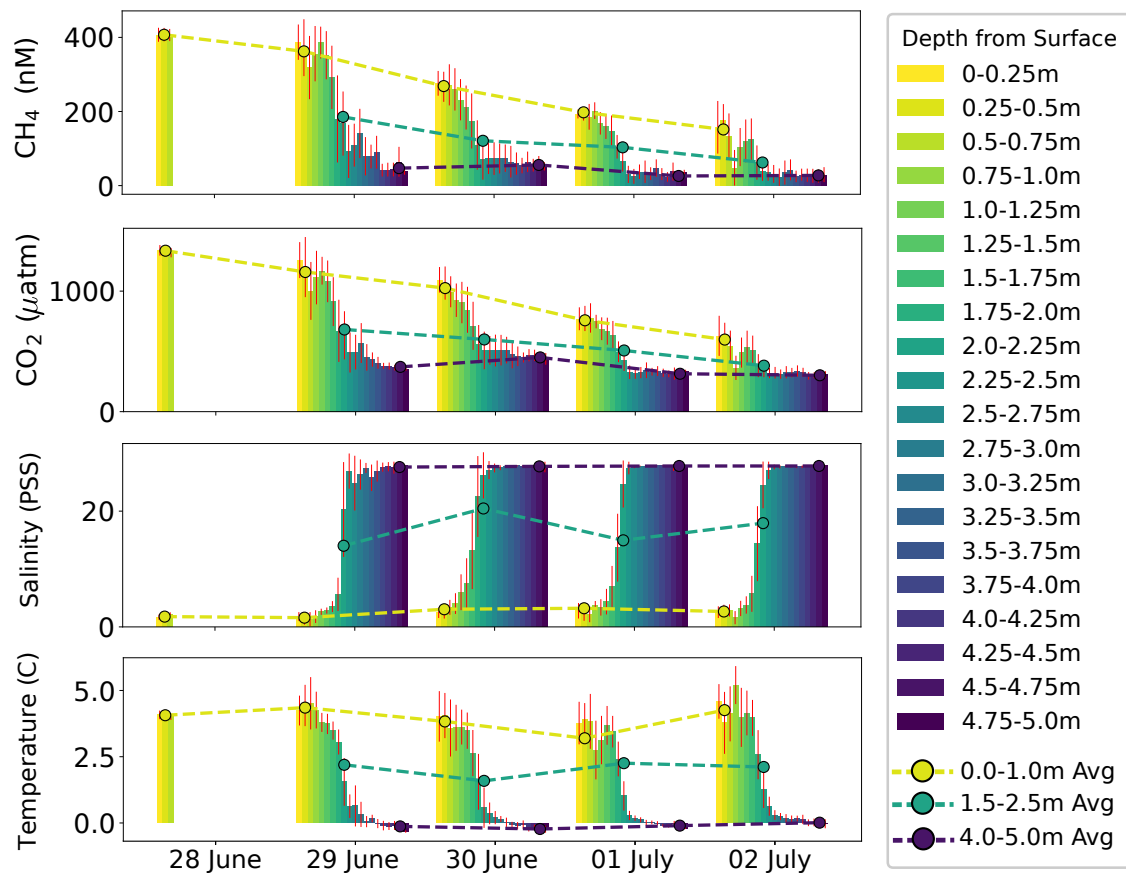


Figure 3-18: **Cambridge Bay spatio-temporal trends of the pycnocline** Bar graphs showing CH₄, CO₂, salinity, and temperature, colored by depth binned in 0.25 m intervals. Trend lines are added to show the average value of measurements in the surface layer (top 1 m), the pycnocline (1.5-2.5 m), and below the pycnocline (4-5 m). On 30 June, elevated wind briefly disturbed the pycnocline leading to vertical mixing, as evidenced by elevated CH₄ and CO₂ levels below the mixed layer, and elevated salinity in the surface layer. Both CH₄ and CO₂ decreased substantially in the surface layer over the 5 day measurement period.

3.5.3 Significance and Role of Transience

River inflow drives CH₄ and CO₂ in the Freshwater Creek estuary during the spring freshet. While there is ice-cover, river water laterally transports greenhouse gases into the estuary and into the harbor. When the ice recedes, the air-sea interface allows for rapid outgassing of methane into the atmosphere. By considering this brief period of flux, total estimates of methane that enters the atmosphere increases 22-fold from 7.4 kg to approximately 640 kg. The contribution in this short time-frame alone could account for as high as 95% of emissions for the estuary. Additionally, due to the lateral transport of gas out of the study region under ice in the harbor, it is hypothesized that significant gas concentrations will continue to build under the ice cover in the harbor to eventually be ventilated to the atmosphere at a later date from within the harbor or as far as the coastal ocean.

Seasonally ice-covered Arctic estuaries receive approximately 10% of global river discharge [187]. If Freshwater Creek is representative of Arctic coastal estuaries, then there may be a significant underestimate of CH₄ and CO₂ contributions from Arctic environments reported using only summertime, low-ice measurements. Additionally, projected increases of river inflow from warming conditions [188] may exacerbate these outgassing events.

Using traditional methods, bottle samples were able to resolve that a major outgassing event occurs in the spring, however, the extent and dynamic of this flux was difficult to constrain. Using the ChemYak, ephemeral lateral transport and vertical mixing events were revealed. Over a 5-day sampling period, the total gas concentration in the surface waters significantly decreased, highlighting the need for dense temporal measurements to constrain the rate of outgassing. Future work that specifically inspects the role of biological processes in the estuary and performs contemporaneous under-ice sampling, can further characterize the carbon cycle in this environment. Additionally, autonomous monitoring regimes to inspect spatio-temporal events could be employed to free human experts to collect and analyze more contemporaneous samples.

3.6 Sampling Transient Phenomenon and Motivating Adaptive Regimes

In each of these field campaigns, transient phenomenon were targets of interest. Using a human-piloted ChemYak aided in the dense collection of *in situ* measurements to reveal complex spatio-temporal structure in post-analysis. Although these campaigns were successful, several key challenges of sampling in transient phenomenon are highlighted.

Tidal Estuary: At Wareham River, Section 3.3, the ChemYak performed several transects over the course of a few hours. Repeated transects revealed fine-scale transience as the spatial distribution of the phenomenon of interest changed significantly with respect to the tides. Although the source of the effluent was known, the distribution of the effluent in the surface waters was unknown and not detectable by a human pilot. This ultimately resulted in a spatial uniform coverage strategy, which resulted in sample sparsity in regions of interest; particularly mixing boundaries and accumulation events. Practically, piloting the ChemYak became challenging during the falling tide, when the waters became shallow enough to make navigation treacherous for the human-occupied chase vehicle.

Open Ocean: At the Cascadia Margin, Section 3.4, the ChemYak revealed small, local supersaturation structures in the water column, however the number of samples of these pockets is relatively few compared to other, less interesting regions. Although a uniform coverage trajectory was manually attempted, piloting the vehicle was made difficult due to waves and lack of physical navigation references.

Arctic River Inflow: In Cambridge Bay, Section 3.5, the ChemYak was used over multiple days. Just as in the tidal estuary, the “source” of the gas signal was known, but the distribution of the phenomenon in the estuary and harbor was unknown and undetectable by a human. In post-processing, it was revealed that there were daily differences in the distribution both vertically and laterally that were indicative of

potentially key environmental influences, like wind, upwelling, or ice melt. If these differences had been known, more observations would have been collected in regions that displayed unique or changed characteristics between deployments.

Challenging and changing physical conditions, in addition to transient behavior in these environments, limited the obvious action set for a human pilot. A naive strategy to improve spatial coverage would be to implement an autonomous uniform coverage search or design a set of repeatable monitoring trajectories. However, these environments illustrate the potential failure of these regimes: sparse sampling of local phenomenon (e.g., supersaturation), missing ephemeral events (e.g., wind-driven vertical mixing, warming outgassing, slack-tide accumulation), and collisions with unknown/changing metric environments (e.g., tidal shallows). An adaptive regime is necessary to optimize a scientific objective with respect to transient phenomenon.

In the next chapter, the *maximum seek-and-sample* (MSS) scientific objective is defined. Using an adaptive sampling framework, optimizing this objective allows for a robotic vehicle to find the globally most interesting region to densely sample in an environment. With respect to these field campaigns, this would be equivalent to tracing the extent of wastewater effluent diffusion, densely sampling supersaturations, or inspecting the Arctic area with most unique/unexpected dynamics. The remainder of the chapter discusses and demonstrates an adaptive sampling regime which optimizes the MSS objective and provides theoretical performance guarantees.

Chapter 4

Adaptive Sampling for Transient Phenomenon

As highlighted in Chapter 3, transience in natural environments can describe both short- and long-term temporal evolution of a phenomenon. In order to draw useful samples of a target environment, consideration of the impact and nature of temporal changes may need to be considered. A pervasive and specific adaptive sampling problem ubiquitous in the sciences is the *maximum seek-and-sample* (MSS) problem, which necessitates consideration of transience. In many environmental and earth science applications, experts want to collect scientifically valuable samples of a maximum (e.g., an oil spill source), but the distribution of the phenomenon is initially unknown. When environmental samples are collected using uniform coverage methods as discussed in Chapter 1, few samples are generally collected of the maximum, known as sample sparsity. Increasing the number of valuable samples at the maximum requires adaptive online planning and execution. To address the MSS problem, this chapter draws on perspectives from *informative path planning* (IPP), *decision making under uncertainty*, and *Bayesian optimization* (BO) to present PLUMES — **Plume Localization under Uncertainty using Maximum-Value information and Search**. PLUMES is an adaptive, flexible algorithm that enables a mobile robot to efficiently localize and densely sample an environmental maximum, subject to practical challenges including action constraints, unknown geometric map and obstacles,

noisy sensors with limited field-of-view, and phenomena transience. Work presented here has been published in brief form [189].

MSS problems are a subset of IPP problems, which require that an agent makes decisions to optimize over an information-theoretic reward function in order to study a target distribution. Typical offline IPP techniques for pure information-gathering that optimize submodular coverage objectives can achieve near-optimal performance [19, 105]. However, in the MSS problem, the value of a sample depends on the unknown maximum location, requiring adaptive planning to enable the robot to select actions that explore to localize the maximum and then seamlessly transition to selecting actions that exploitatively collect valuable samples there. Even for adaptive IPP methods, the MSS problem presents considerable challenges. The target environmental phenomenon is partially observable and most directly modeled as a continuous scalar function. Additionally, efficient maximum sampling with a mobile robot requires consideration of vehicle dynamics, travel cost, and a potentially unknown obstacle map. Handling these challenges in combination excludes adaptive IPP algorithms that use discrete state spaces [14, 15], known metric maps [100, 190], or unconstrained sensor placement [35].

To precisely describe the MSS problem, this chapter defines the MSS POMDP. Partially-observable Markov decision processes (POMDPs) are general models for decision-making under uncertainty that allow the challenging aspects of the MSS problem to be encoded. In the MSS POMDP, the partially observable state represents the continuous environmental phenomenon and a sparse reward function encodes the MSS scientific objective by giving reward only to samples sufficiently close to the global maximum. “Sparsity” with respect to the reward function refers to the small region that is given non-zero or non-trivial reward. Solving a POMDP exactly is generally intractable, and the MSS POMDP is additionally complicated by both continuous state and observation spaces, and the sparse MSS reward function. This presents the two core challenges that PLUMES addresses: performing online search in a belief-space over continuous functions, and overcoming reward function sparsity.

To address the former challenge, PLUMES uses a Gaussian Process (GP) model

to represent belief over the environment state, which is itself a continuous function. A GP representation is a compact probabilistic model with well-characterized analytic methods for online data incorporation and posterior queries. With continuous belief and observations functions, there is an uncountable set of possible GP beliefs which arise as every observation (and resulting state) is unique with probability one. This makes planning over possible actions and their outcomes difficult. To address this difficulty, state-of-the-art online POMDP solvers use deterministic discretization [106] or a combination of sampling techniques and particle filter belief representations [87, 107–109] to reduce the size of the possible set to be tractable. However, efficiently discretizing or maintaining a sufficiently rich particle set to represent the underlying continuous function in MSS applications is itself a challenging problem, and can lead to inaccurate inference of the maximum [191]. Other approaches have considered using the maximum-likelihood observation to make search tractable [85]. However, this assumption can compromise search and has optimality guarantees only in linear-Gaussian systems [192]. Instead, PLUMES uses Monte Carlo Tree Search (MCTS) with progressive widening, referred to as *continuous-observation MCTS*, to limit the growth of the planning tree [22] and retain optimality [21] in continuous environments.

To plan with sparse rewards requires long-horizon information gathering and is an open problem in robotics [17]. To alleviate this difficulty, less sparse heuristic reward functions can be optimized in place of the true reward. These heuristics need to be selected carefully to ensure the planner performs well with respect to the true objective. In IPP, heuristics based on the value of information have been applied successfully [35, 85, 86, 193], primarily using the GP-UCB criteria [18, 19]. This criteria requires setting a hand-tuned parameter, β , that handles the explore-exploit trade-off. For some values of β , UCB heuristics can guarantee convergence to the global maximum at the limit of time, however this chapter will show that using UCB for the MSS POMDP in practice can lead to suboptimal convergence to local maxima due to inefficient explore-exploit regimes in finite duration missions. Instead, PLUMES takes advantage of a parameterless heuristic function from the BO community for state-of-the-art black-box optimization [20], which is referred to as

maximum-value information (MVI). MVI overcomes reward sparsity and encourages long-term information gathering, in addition to efficiently converging to the true reward of the MSS POMDP for more effective exploitative sampling of the maximum.

The contribution of this chapter is the MSS POMDP formalism and the corresponding PLUMES planner, which by virtue of its belief model, information-theoretic reward heuristic, and search framework, enables efficient maximum seek and sample with asymptotic optimality guarantees in continuous environments. PLUMES extends the state-of-the-art in MSS planners by applying a BO heuristic reward function to MSS that alleviates the challenges of the true sparse MSS reward function, and integrating GP belief representations within continuous-observation MCTS. PLUMES is demonstrated in extensive simulation and field trials, showing a statistically significant performance improvement over state-of-the-art baselines.

4.1 Maximum Seek-and-Sample POMDP Formalism

Let the physical region of the environmental domain be a d -dimensional compact set $\mathbb{X}_w \subset \mathbb{R}^d$. Let \mathbb{X}_w contain obstacles with arbitrary geometry and let $\mathbb{X} \subset \mathbb{X}_w$ be the set of reachable points with respect to the robot’s initial pose in the environment. An unknown underlying continuous function $f : \mathbb{X}_w \rightarrow \mathbb{R}$ represents the value of a continuous phenomenon of interest. The objective is to find the unique global maximizer $\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathbb{X}} f(\mathbf{x})$ by safely navigating while receiving noisy observations of this function f . Because f is unknown, there is no access to derivative information or any analytic form.

The process of navigating and generating observations is modeled as the MSS POMDP: an 8-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{Z}, T, O, R, \gamma, b_0)$:

- \mathcal{S} : continuous state space of the robot and environment
- \mathcal{A} : discrete set of action primitives
- \mathcal{Z} : continuous space of possible observations
- $T: \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{P}(\mathcal{S})$, the transition function, i.e., $\Pr(S_{t+1} = s' \mid S_t = s, A_t = a)$

- $O: \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{P}(\mathcal{Z})$, the observation model, i.e., $\Pr(Z_{t+1} = z \mid S_{t+1} = s, A_t = a)$
- $R: \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$, the reward of taking action a when robot's state is s , i.e., $R(s, a)$
- γ : discount factor, $0 \leq \gamma \leq 1$
- b_0 : initial belief state of the robot, $b_0 \in \mathcal{P}(S_0)$

where $\mathcal{P}(\cdot)$ denotes the space of probability distributions over the argument.

The Bellman equation is used to recursively quantify the value of belief $b_t = \mathcal{P}(S_t)$ over a finite horizon h under policy $\pi : b_t \rightarrow a_t$ as:

$$V_h^\pi(b_t) = \mathbb{E}[R(s_t, \pi(b_t))] + \gamma \int_{z \in \mathcal{Z}} V_{h-1}^\pi(b_{t+1}^{\pi(b_t), z}) \Pr(z \mid b_t, \pi(b_t)) dz, \quad (4.1)$$

where the expectation is taken over the current belief and $b_{t+1}^{\pi(b_t), z}$ is the updated belief after taking action $\pi(b_t)$ and observing $z \in \mathcal{Z}$. The optimal policy π_h^* over horizon- h is the maximizer of the value function over the space of possible policies Π : $\pi_h^* = \arg \max_{\pi \in \Pi} V_h^\pi(b_t)$. Unfortunately, Eq. 4.1 is intractable to compute in continuous state and observation spaces; an approximation is necessary. Using a sequential decision-making structure, PLUMES performs receding-horizon, online search to approximate the optimal policy for the MSS POMDP online:

1. Conditioned on b_t , approximate the optimal policy π_h^* for finite horizon h and execute the action $a = \hat{\pi}_h^*(b_t)$.
2. Collect observations $z \in \mathcal{Z}$, according to O .
3. Update b_t to incorporate these new observations; repeat.

In the following sections, the specific choice of belief model, planning algorithm, and heuristic reward function that PLUMES uses to solve the MSS POMDP is presented.

4.2 Representing Belief with Gaussian Processes

The robot’s pose \mathbf{x}_t at planning iteration t is assumed to be fully observable (e.g., GPS or local positioning system is available) and the unknown phenomenon f is only partially observable for any time step. The full belief-state can be represented as a tuple b_t , of robot metric state \mathbf{x}_t and environment belief (i.e., belief about the form of function f) $g_t = \mathcal{P}(f)$ at time t . Because f is a continuous function, the belief g_t cannot be represented as a distribution over discrete states, as is standard in POMDP literature [12]. As an alternative, PLUMES uses a Gaussian process (GP) [16], to represent g_t conditioned on a history of past observations. A GP is parameterized by mean $\mu(\mathbf{x})$ and covariance function $\kappa(\mathbf{x}, \mathbf{x}')$, which allow for prior knowledge to be encoded. For the purposes of this work, a 0 mean function is assumed. For spatio-temporal phenomena, kernels that incorporate knowledge of the current time or planning iteration are used, which allow the predictive mean and variance of the spatial distribution to change temporally [194].

The robot traverses a location \mathbf{x} and gathers observations $z \in \mathcal{Z}$ of f subject to sensor noise σ_n^2 , such that $z = f(\mathbf{x}) + \epsilon$ with $\epsilon \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_n^2)$. Given a history $\mathcal{D}_t = \{\mathbf{x}_i, z_i\}_{i=0}^D$ of D observations and observation locations at planning iteration t , the posterior belief at a new location $\mathbf{x}' \in \mathbb{X}$ is computed:

$$g_t(\mathbf{x}') \mid \mathcal{D}_t \sim \mathcal{N}(\mu_t(\mathbf{x}'), \sigma_t^2(\mathbf{x}')), \text{ where} \quad (4.2)$$

$$\mu_t(\mathbf{x}') = \kappa_t(\mathbf{x}')^\top (\mathbf{K}_t + \sigma_n^2 \mathbf{I})^{-1} \mathbf{z}_t, \quad (4.3)$$

$$\sigma_t^2(\mathbf{x}') = \kappa(\mathbf{x}', \mathbf{x}') - \kappa_t(\mathbf{x}')^\top (\mathbf{K}_t + \sigma_n^2 \mathbf{I})^{-1} \kappa_t(\mathbf{x}'), \quad (4.4)$$

where $\mathbf{z}_t = [z_0, \dots, z_{D-1}]^\top$, \mathbf{K}_t is the positive definite kernel matrix with $\mathbf{K}_t[i, j] = \kappa(\mathbf{x}_i, \mathbf{x}_j)$ for all $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_t$, and $\kappa_t(\mathbf{x}') = [\kappa(\mathbf{x}_0, \mathbf{x}'), \dots, \kappa(\mathbf{x}_{D-1}, \mathbf{x}')]^\top$.

4.2.1 Kernels for Natural Phenomenon

Using a kernel function, relationships in space and time can be encoded, e.g., spatial diffusivity of a chemical plume, period of a seasonal or tidal cycle. Representative

kernels for modeling natural phenomenon are described in this section.

Radial Basis Function (RBF) One of the most common kernel functions used with GPs is the radial basis function (RBF), also known as the squared exponential kernel, which takes the form:

$$\kappa_{RBF}(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2l^2}\right) \quad (4.5)$$

where σ^2 is a variance parameter, and l is a lengthscale parameter. In the RBF kernel, variance acts as a scale factor, and directly describes the range at which measurements may fall around the mean. The lengthscale describes the smoothness of a distribution, and captures the relatedness of points which lie “near” each other in the distribution. Fig. 4-1 illustrates three distributions drawn from three GPs with RBF kernels that have different parameter settings.

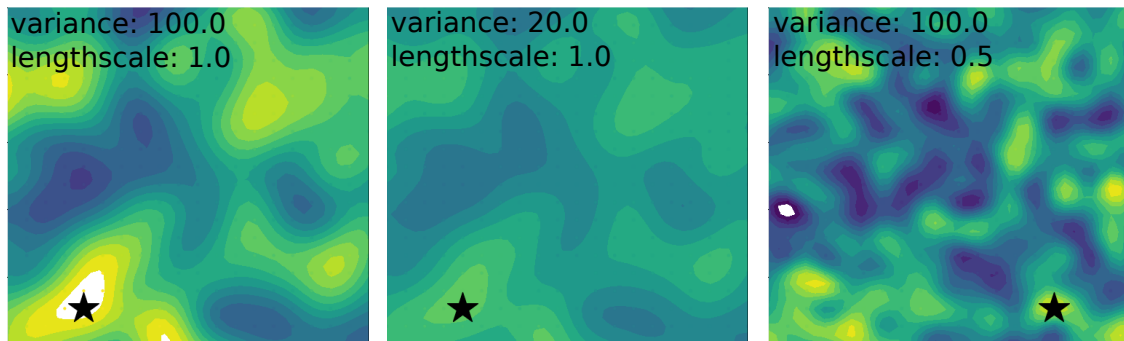


Figure 4-1: **GP with RBF kernels** Shown are several draws of a GP endowed with RBF kernels with varying lengthscale and variance hyperparameters. Generally, lengthscale corresponds to the diffusivity or smoothness of the distribution, whereas variance captures the range at which values may fall around the mean. The maximum in each distribution is marked with a black star. All plots share the same color-scale.

An RBF kernel can be used to model random-walk behavior in time by simply extending the kernel from 2-dimensional inputs (e.g., Cartesian coordinates) to 3-dimensional inputs, in which the third dimension is time. Fig. 4-2 shows several snapshots of an RBF kernel with 3-dimensional input to show the slight stochasticity in the spatial distribution over time.

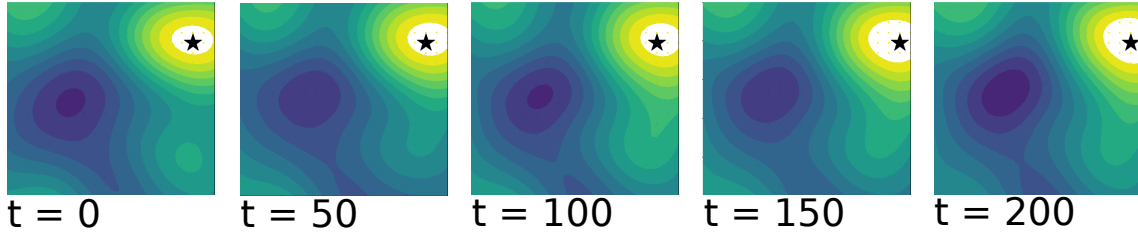


Figure 4-2: **GP with time-varying RBF kernel** In this example, several snapshots of a dynamic phenomenon, modeled with an RBF kernel in 3-dimensions (spatial and temporal) is shown with spatial lengthscale 2.5, temporal lengthscale 30, and variance 100. This type of model is compelling when the time-variation of a phenomenon is stochastically changing and normally distributed about some general average or mean. The maximum is noted with a star. All images render with the same color-scale.

Periodic Another common kernel function is the periodic kernel, which relates points in a distribution according to a period. This kernel captures the idea that information at a point in space may yield information about points which are a set distance away or at a set time in the future. The kernel takes the form:

$$\kappa_{PER}(\mathbf{x}, (\mathbf{x}')) = \sigma^2 \exp\left(-\frac{\sin^2\left(\frac{\pi\|\mathbf{x}-\mathbf{x}'\|}{F}\right)}{2l^2}\right) \quad (4.6)$$

where σ^2 is the variance, F is the period of the cycle, and l is the lengthscale of the process. Fig. 4-3 shows a series of distributions drawn from a GP with periodic kernel of varying parameters. Like the RBF kernel, extending a periodic kernel to time simply requires adding another dimension to the kernel.

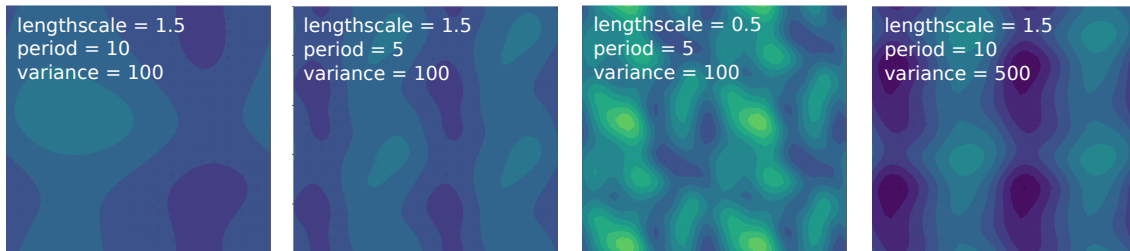


Figure 4-3: **GP with Periodic kernel** These representative distributions are drawn from a GP with periodic kernel. Like in the RBF kernel, variance controls the scale of distribution around the mean and lengthscale controls the smoothness of the distribution. The period controls the frequency with which cyclic relationships appear. All plots share a color-scale.

With only the RBF and periodic kernels, a large number of processes can be

described. Composite kernels can be formed by adding or multiplying kernels together [16], which could allow spatial phenomenon to be distributed according to an RBF kernel, but the scale of the phenomenon to cycle according to time. Fig. 4-4 shows snapshots of a Periodic-RBF kernel ($\kappa_{PER} + \kappa_{RBF}$) to illustrate the complex relationships that can be captured from simple kernel composition.

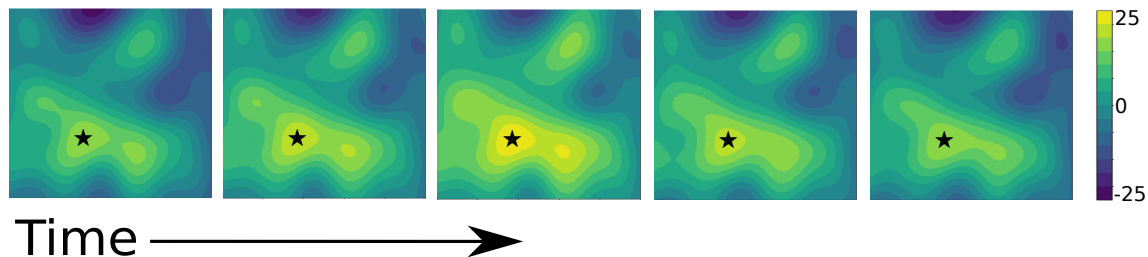


Figure 4-4: **GP with compositional Periodic-RBF kernel** Several snapshots in time of a phenomenon simulated with a GP using a Periodic-RBF kernel is shown. Such a kernel allows for cyclic patterns to be captured in the value of the distribution. The maximum in the distribution is marked with a star. All plots share a color-scale.

With respect to natural phenomenon, periodic kernels or composition kernels with periodic features can be used to describe tides, seasons, diurnal cycles, biological cycles, and many more processes which follow a regular (or predictable) cyclic schedule.

Custom Kernels In some environments, a specific dynamic may be known about the phenomenon, e.g., a moving ship leaking fuel, a river flow in a channel, the trajectory of a migratory species. Kernels can be designed with specific dynamic functions. Marchant et al. [85] demonstrate their adaptive sampling algorithm in a world drawn from a specific spatio-temporal function. It is adapted here as an illustration of the flexibility of kernel and GP formulations:

$$f(\mathbf{x}, T) = \exp\left(-\frac{\mathbf{x}_0 - c_0 - \sin\left(\frac{T}{F_0}\right)}{l_0}\right) \exp\left(-\frac{\mathbf{x}_1 - c_1 - \sin\left(\frac{T}{F_1}\right)}{l_1}\right) \quad (4.7)$$

$$\kappa_{DYN}((\mathbf{x}, T), (\mathbf{x}', T')) = \sigma^2(f^2(\mathbf{x}, T)) \cdot (f^2(\mathbf{x}', T'))$$

where $f(-, -)$ describes the dynamic function of a moving source, parameterized by

$c_{\{0,1\}}$, $F_{\{0,1\}}$, and $l_{\{0,1\}}$ which represent the center point of the circular motion, the period of the rotation cycle, and the lengthscale of the spatial source in Cartesian coordinates $(0, 1)$, respectively. σ^2 is the kernel variance. Fig. 4-5 shows several frames of a dynamic target that can be described with a GP using this kernel.

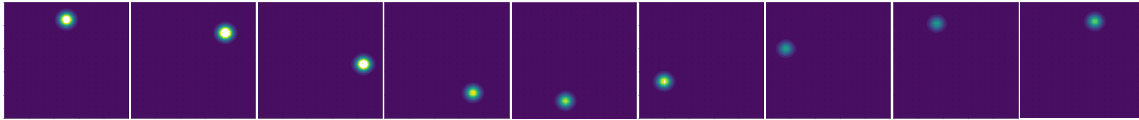


Figure 4-5: **GP with dynamics model embedded in kernel** Several snapshots in time of a phenomenon simulated using the kernel defined in Eq. 4.7 is shown. The path of the source follows a roughly circular trajectory around the center of the environment.

4.3 Planning in Continuous State, Observation Spaces

PLUMES selects high-reward actions with receding-horizon search over possible belief states. This search requires a simulator that can sample observations and generate beliefs given a proposed action. The GP model, which represents the belief over the continuous function f , can simulate drawing continuous observations from proposed actions by sampling from the Gaussian distribution defined by Eq. 4.3 & 4.4.

PLUMES uses continuous-observation MCTS to overcome the challenges of planning in continuous state and observation spaces. Continuous-observation MCTS has three stages: *selection*, *forward simulation*, and *back-propagation*. Each node in the tree is a tuple of robot pose and GP belief, $b_t = \{\mathbf{x}_t, g_t\}$. Additionally, each node is provided a label distinguishing *belief nodes* and *belief-action nodes*. The root of the tree is a belief node which contains the entire history of physically realized actions and observations through the current planning iteration. Through selection and simulation, belief and belief-action nodes are alternately added to the tree (Fig. 4-6).

From the root, a *rollout* begins with the *selection* stage, in which a belief-action child is selected according to the Polynomial Upper Confidence Tree (PUCT) policy [21]. The PUCT value $\hat{Q}_{aug}^*(b_t, a)$ is the sum of the average heuristic rewards (i.e., MVI) from all previous simulations and a term that favors less-simulated children:

$$\hat{Q}_{aug}^*(b_t, a) = \hat{Q}^*(b_t, a) + \sqrt{\frac{N(b_t)^{e_d}}{N(b_t, a)}}, \quad (4.8)$$

where $\hat{Q}^*(b_t, a)$ is the average heuristic reward of choosing action a with belief b_t in all previous rollouts, $N(b_t)$ is the number of times the node b_t has been simulated, $N(b_t, a)$ is the number of times that particular action from node b_t has been selected, and e_d is a depth-dependent parameter*.

Once a child belief-action node is selected, the associated action is *forward simulated* using the generative observation model O , and a new belief node is generated $b_{t+1} = \{\mathbf{x}_{t+1}, g_{t+1}\}$ as though the action were taken and samples observed. The simulated observations are drawn from the belief-action node's GP model g_t , and the robot's pose is updated deterministically based on the selected action.

*Refer to Table 1 of Auger et al. [21] for parameter settings.

- ① Forward simulation in belief tree with PUCT and prog. widening.
- ② MVI reward computation and backpropagation from leaf to root.

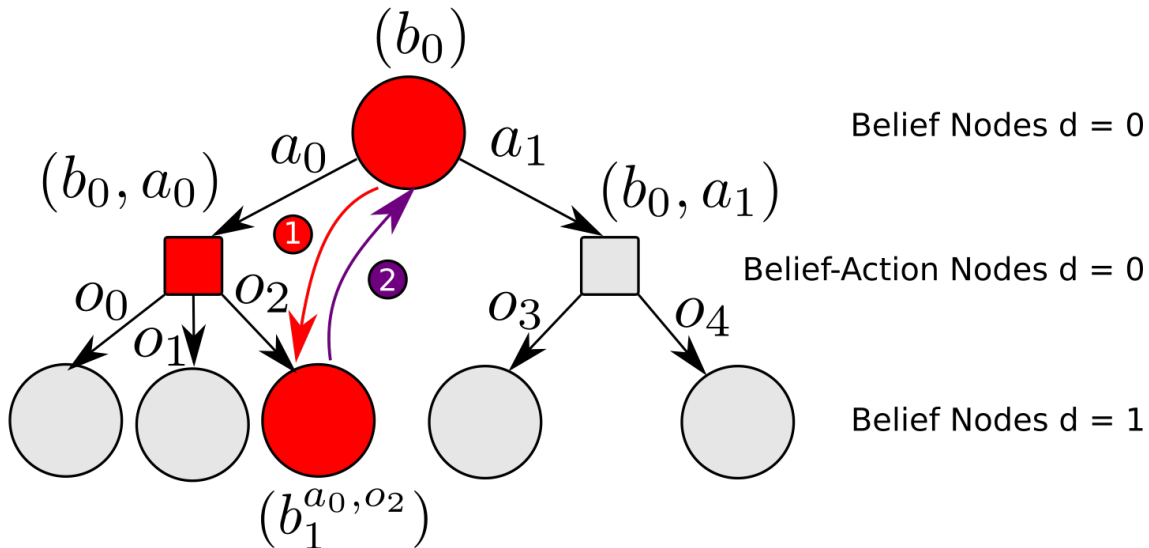


Figure 4-6: **Continuous-observation MCTS** Illustrated to horizon $h = 1$, the tree consists of alternating belief and belief-action nodes. Action decisions are made at belief nodes and random belief transitions according to the observation function occur at belief-action nodes. Note that belief-action nodes have a varying number of children due to progressive widening and unequal simulation (not visualized) due to PUCT policy. Reprinted from [189].

Since the observations in a GP are continuous, every sampled observation is unique with probability one. *Progressive widening*, with depth-dependent parameter* α_d incrementally grows the tree by limiting the number of belief children from each belief-action node. When growing the tree, b_{t+1} is either chosen to be the least visited child if $\lfloor N(b_t, a)^{\alpha_d} \rfloor = \lfloor (N(b_t, a) - 1)^{\alpha_d} \rfloor$, or otherwise a new child with observations simulated from b_t . By limiting the search tree width and incrementally adding explored children, progressive widening avoids search degeneracy in continuous environments.

Once a sequence of actions has been rolled out to a horizon h , the accumulated heuristic reward is *propagated upward* from the leaves to the tree root. The average accumulated heuristic reward and number of queries are updated for each node visited in the rollout. Rollouts continue until the computation budget is exhausted. The most visited belief-action child of the root node is executed. Pseudo-code for continuous-observation MCTS is provided in Algorithms 2 & 3.

Algorithm 2: Pseudo-code for continuous-observation MCTS

```

1 def Main(pose, world, budget, horizon):
2   root = (pose, world, label = 'belief', queries = 0);
3   while budget not exhausted do
4     children  $\leftarrow$  GetActions (root);
5     child  $\leftarrow$  SelectChild (children);
6     current_node = child;
7     i = 0;
8     while i < horizon do
9       if current_node.label is 'belief' then
10        children  $\leftarrow$  GetActions (current_node);
11        current_node  $\leftarrow$  SelectChild (children);
12        i += 1;
13      end
14      if current_node.label is 'belief-action' then
15        current_node, reward = Simulate (current_node);
16        total_reward += reward;
17      end
18    end
19    BackProp (current_node, total_reward);
20  end
21  return best child of root.children according to child.queries;

```

*Refer to Table 1 of Auger et al. [21] for parameter settings.

Algorithm 3: Pseudo-code for supporting functions of continuous-observation MCTS

```
1 def GetActions(node):
2   | node.children  $\leftarrow$  valid actions from node.pose;
3   | label all node.children as ‘belief-action’ nodes;
4   | return node.children;
5 def SelectChild(nodes):
6   | puct_values = [];
7   | for node in nodes do
8   |   | if node.queries is 0 then
9   |   |   | return node;
10  |   | end
11  |   | puct_values.append( $\frac{\text{child.reward}}{\text{child.queries}} + c\sqrt{\frac{\text{node.parent.queries}^{e_d}}{\text{node.queries}}}$ );
12  |   | end
13  |   | return node with largest corresponding value in puct_values;
14 def Simulate(node):
15  | if node.children is not None then
16  |   | if  $\text{round}(\text{len}(\text{node.children})^\alpha) == \text{round}(\text{len}(\text{node.children})-1)^\alpha$  then
17  |   |   | return child with fewest queries;
18  |   | end
19  |   | end
20  |   | world  $\leftarrow$  node.world updated with locations at node.action and samples of
21  |   |   | node.world at node.action;
22  |   |   | child  $\leftarrow$  (pose = node.action [-1], world, label = ‘belief’, queries = 0);
23  |   |   | reward  $\leftarrow$  Heuristic (node.action, node.world);
24  |   |   | add child to node.children;
25  |   |   | return child, reward;
26 def BackProp(node, reward):
27  | node.queries += 1;
28  | node.reward += reward;
29  | if node.parent is None then
30  |   | return;
31  |   | end
32  |   | else
33  |   |   | BackProp (node.parent, reward);
34  |   |   | end
35  |   | end
```

4.3.1 Analysis of Continuous Observation MCTS

Continuous-observation MCTS within PLUMES provides both practical and theoretical benefits. Practically, progressive-widening directly addresses search degeneracy by visiting belief nodes multiple times even in continuous observation spaces, allowing for a more representative estimate of their value. Theoretically, PLUMES can be shown to select asymptotically optimal actions; analysis from Auger et al. [21] for PUCT-MCTS with progressive widening in MDPs is extended here for PLUMES. Note: the optimality of actions selected by continuous-observation MCTS with PUCT and progressive-widening modifications is characterized with respect to any reward function bounded in $[0, 1]$.

Lemma 4.3.1. *Given a POMDP of the form $M = (\mathcal{S}, \mathcal{A}, \mathcal{Z}, T, O, R, \gamma, b_0)$, derive an equivalent belief-state MDP, with a state for each of the uncountably infinite beliefs of M , such that $\tilde{M} = (\tilde{\mathcal{S}}, \mathcal{A}, \tilde{T}, \tilde{R}, b_0)$. In the derived belief-state MDP, the transition function \tilde{T} is the probability density of taking action $a \in \mathcal{A}$, observing $o \in \mathcal{Z}$ according to the current belief-state b_t and observation function O , and transitioning to belief-state b_{t+1} , while the reward function \tilde{R} is the expected reward R of the original POMDP with respect to the current belief.*

Then the value function $Q_h(\tilde{s}, a)$ of the derived MDP \tilde{M} is equal to that of the POMDP M , $Q_h(b, a)$, for any policy π and finite horizon h , and the optimal policy of the belief-state MDP is the optimal policy for the original POMDP.

Proof. The observation that the belief-state is fully-observable to the robot, and that stochastic observations lead to stochastic transitions in the continuous belief-state immediately provides this well known reduction found in e.g., Kaelbling et al. [12], and applications of which appear in e.g., Lemma 1 of Silver et al. [108]. \square

Given the reduction from POMDP to belief-state MDP, the convergence proof of fully-observable PUCT-MCTS with PW to partially observable problems with continuous observation functions and belief-state spaces can be adapted. By using GPs to perform belief updates in each stage of PLUMES, the potential challenges of applying MCTS to continuous, partially observable problems discussed in [195] is overcome.

Definition 4.3.1. *Exponentially sure in n .* A property depending on integer n is exponentially sure in n if there exists positive constants C, h , such that the probability the property holds is at least: $1 - C \exp(-hn)$ [21].

Theorem 4.3.2. *For the right choice of PUCT exploration e_d and progressive widening α_d parameters, the derived belief-state MDP value function estimated using continuous observation PUCT-MCTS with PW converges to the value function of the optimal policy, as number of forward simulations approaches infinity. The bias of the value function decreases polynomially with the number of visits to a belief-action node, such that: $\left| \hat{Q}_h^*(\{\mathbf{x}_t, b_t\}, a) - Q_h^*(\{\mathbf{x}_t, b_t\}, a) \right| \leq \frac{C}{N(\{\mathbf{x}_t, b_t\}, a)^{\gamma_d}}$ exponentially surely, for depth-dependent constants $C > 0$ and γ_d .*

Proof. The proof of this theorem follows directly from Lemma 4.3.1, which reduces the original POMDP with continuous observations and belief-state, to a belief-state MDP with stochastic transitions in a continuous state space; and from Theorem 1 in Auger et al. [21], which provides convergence results for MCTS in continuous MDPs with PUCT and double PW. Although only discrete action spaces are considered in this work, the method presented by Auger et al. additionally searches over continuous action spaces. For finite action spaces with only PUCT exploration, every child belief-action node will be visited infinitely often as the number of simulations approaches infinity, and Lemma 3 of their analysis is satisfied. \square

Corollary 4.3.2.1. *After n simulations, the value of the most simulated action from the root belief node differs from the optimal value by less than $O(n^{-1/c})$, exponentially surely in n , for constant $c > 0$ that depends on the horizon h (see Lemma 4.3.1 and [21]).*

This result provides asymptotic and finite-time performance guarantees for the continuous-observation MCTS proposed in PLUMES, implying that PLUMES selects actions according to the optimal policy with high probability, without exploring the entire infinitely large search tree.

4.4 Maximum-value Information Heuristic Reward

The true state-dependent reward function for the MSS POMDP would place value on collecting sample points \mathbf{x} within an ϵ -ball of the true global maximum \mathbf{x}^* :

$$R(f, \mathbf{x}) = \mathbb{1}_{\|\mathbf{x}-\mathbf{x}^*\|<\epsilon}, \quad (4.9)$$

where ϵ is determined by the scientific application. Optimizing this sparse reward function directly is challenging, so PLUMES approximates the true MSS reward by using the maximum-value information (MVI) heuristic reward [20]. MVI initially encourages exploration behavior, but ultimately rewards exploitative sampling near the inferred maximum.

The belief-dependent MVI heuristic reward $\tilde{R}(b_t, \mathbf{x})$ quantifies the expected value of having belief b_t and collecting a sample at location $\mathbf{x} \in \mathbb{X}$. MVI reward quantifies the mutual information between the random variable Z , representing the observation at location \mathbf{x} , and Z^* , the random variable representing the value of the function f at the global maximum:

$$\tilde{R}(b_t, \mathbf{x}) = I(\{\mathbf{x}, Z\}; Z^* | b_t), \quad (4.10)$$

where $Z^* = \max_{\mathbf{x}' \in \mathbb{X}} f(\mathbf{x}')$. To compute the reward of collecting a random observation Z at location \mathbf{x} under belief b_t , PLUMES approximates the expectation over the unknown Z^* by sampling from the posterior distribution $z_i^* \sim p(Z^* | b_t)$ and uses Monte Carlo integration with M samples following Wang et al. [20]:

$$\begin{aligned} \tilde{R}(b_t, \mathbf{x}) &= H[\mathbf{Pr}(Z | \mathbf{x}, b_t)] - \mathbb{E}_{z' \sim \mathbf{Pr}(Z^* | b_t)} [H[\mathbf{Pr}(Z | \mathbf{x}, b_t, Z^* = z')]], \\ &\approx H[\mathbf{Pr}(Z | \mathbf{x}, b_t)] - \frac{1}{M} \sum_{i=0}^M H[\mathbf{Pr}(Z | \mathbf{x}, b_t, Z^* = z_i^*)]. \end{aligned} \quad (4.11)$$

Each entropy expression $H[\cdot]$ can be respectively approximated as the entropy of a Gaussian random variable with mean and variance given by the GP equations

(Eq. 4.3 & 4.4), and the entropy of a truncated Gaussian, with upper limit z_i^* and the same mean and variance.

To draw samples z_i^* from the posterior $p(Z^* | b_t)$, spectral sampling [196] can be employed. Spectral sampling draws a function \hat{f} , which has analytic form and is differentiable, from the posterior belief of a GP with a stationary, shift-invariant covariance function [20,93]. The restriction on kernel properties is a result of Bochner’s theorem [197], which provides that for a valid (i.e., positive definite) shift-invariant kernel, then the Fourier dual of the kernel is also a probability distribution. To complete the evaluation of Eq. 4.11, $z_i^* \sim p(Z^* | b_t)$ can be computed by applying standard efficient global optimization techniques (e.g., sequential least squares programming, quasi-Newton methods) to find the global maximum of the sampled \hat{f} from the normalized spectral density.

In practice, spectral sampling requires deriving the spectral density of a kernel analytically, and the restriction on stationary kernels generally limits the flexibility of the approach. An alternative technique which does not place a restriction on the form of the kernel function was suggested by Wang et al. [20], which approximates z_i^* using a Gumbel distribution, $\mathcal{G}(a, b)$. The Gumbel distribution is generally used to describe the distribution of extrema in samples of various root distributions, and in this work is used such that $\Pr(Z^* < z) \approx \mathcal{G}(a, b)$. Parameters a and b of the distribution are set according to [20]. PLUMES makes use of Gumbel sampling for all 3-dimensional and non-RBF kernel functions.

Once z^* is drawn, MVI reward can be expressed as:

$$\tilde{R}(b_t, \mathbf{x}) \approx \frac{1}{M} \sum_{i=0}^M \frac{\gamma_{z_i^*}(\mathbf{x}) \phi(\gamma_{z_i^*}(\mathbf{x}))}{2\Phi(\gamma_{z_i^*}(\mathbf{x}))} - \log(\Phi(\gamma_{z_i^*}(\mathbf{x}))) \quad (4.12)$$

where $\gamma_{z_i^*}(\mathbf{x}) = \frac{z_i^* - \mu_t(\mathbf{x})}{\sigma_t(\mathbf{x})}$, $\mu_t(x)$ and $\sigma_t(x)$ are given by Eq. 4.3 & 4.4, and ϕ and Φ are the standard normal PDF and CDF. For actions that collect samples at more than one location, the reward of an action $\tilde{R}(b_t, a)$ is the sum of rewards of the locations sampled by that action. Pseudo-code for calculating MVI reward is provided in Algorithm 4.

Algorithm 4: Calculating MVI Reward

Data: query as locations to calculate reward, world as current GP belief state,
 M number of z^* samples to draw
Result: reward of query as determined by MVI
/* draw samples of z_i^* by spectral or Gumbel sampling */;

```
1 for  $i \in \text{range}(M)$  do
2   if using spectral sampling then
3     compute  $\phi(\mathbf{x})$  by Bochner's Theorem [20, 197];
4     compute  $\tilde{a}^T \leftarrow \mathcal{N}(\nu, \Sigma)$  where  $\nu$  and  $\Sigma$  can be found from
       world.mean( $\mathbf{x}$ ) =  $\nu^T \phi(\mathbf{x})$  and world.kernel =  $\phi(\mathbf{x})^T \Sigma \phi(\mathbf{x}')$ ;
5      $i = 0$ ;
6     while  $i < M$  do
7       draw  $\tilde{f}_i$  from  $\tilde{a}^T \phi(\mathbf{x})$ ;
8        $z_i^* \leftarrow \max_{\mathbf{x} \in \mathbb{X}} \tilde{f}_i$ ;
9        $i += 1$ ;
10    end
11  end
12  else
13    draw  $p(\mathbf{x})$  from  $\mathcal{N}(\mathbf{x} - \text{mean}(\text{world}(\mathbf{x})), \text{std}(\text{world}(\mathbf{x})))$ ;
14     $q_1, q_2 \leftarrow$  interquartile range of  $p(\mathbf{x})$ ;
15     $m \leftarrow$  median of  $p(\mathbf{x})$ ;
16     $b \leftarrow \frac{q_1 - q_2}{\log(\log(1.3)) - \log(\log(4))}$ ;
17     $a \leftarrow m + b \log(\log(2))$ ;
18     $i = 0$ ;
19    while  $i < M$  do
20      draw  $r$  from Uniform[0, 1];
21       $z_i^* = a - b \log(-\log(r))$ ;
22       $i += 1$ ;
23    end
24  end
25 end
/* now approximate reward of query based on samples of  $z_i^*$  */;
```

26 return $\frac{1}{M} \sum_{i=0}^M \frac{\gamma_{z_i^*}(\text{query}) \phi(\gamma_{z_i^*}(\text{query}))}{2\Phi(\gamma_{z_i^*}(\text{query}))} - \log(\Phi(\gamma_{z_i^*}(\text{query})))$

MVI initially favors collecting observations in areas that have high uncertainty due to sampling maxima from the initial uniform GP belief. As observations are collected and uncertainty diminishes, the sampled maxima converge to the true maximum and reward concentrates locally at this point, encouraging exploitative behavior. This contrasts with the Upper Confidence Bound (UCB) heuristic, which distributes re-

ward proportional to predictive mean $\mu_t(\mathbf{x})$ and weighted variance $\sigma_t(\mathbf{x})$ of the current GP belief model (Eq. 4.3 & 4.4): $\tilde{R}_{\text{UCB}}(b_t, \mathbf{x}) = \mu_t(\mathbf{x}) + \sqrt{\beta_t}\sigma(\mathbf{x})$. As the robot explores, UCB reward converges to the underlying phenomenon, f . The difference in convergence characteristics between MVI and UCB can be observed in Fig. 4-7.

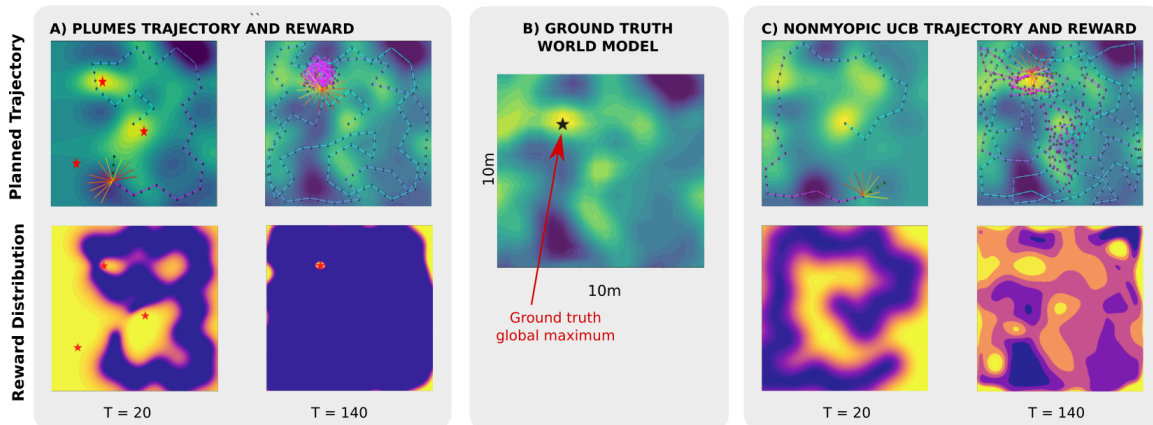


Figure 4-7: **Convergence of MVI vs UCB heuristic** The true environmental phenomenon with the global maximum marked by a star is shown in the center; high regions are colored yellow and low regions blue. In (A,C), the robot trajectory and corresponding reward functions are shown early (20 actions) and later (140 actions) in a mission. On the top row, snapshots of the robot belief state with planned trajectories are shown, with recent actions colored pink and earlier actions colored blue. Red stars mark maxima sampled by MVI. In the bottom row, the corresponding reward function is shown, with high-reward regions colored yellow and low reward regions colored purple. By the end of the mission, MVI clearly converges to placing reward only at the global maximum, which in turn leads to efficient convergence of the robot. In contrast, the reward landscape resulting from canonically used UCB converges to the underlying function, causing the UCB planner to uniformly tour high-valued regions of the environment. Reprinted from [189].

4.5 Results in Static Environments

Using a GP belief, continuous-observation MCTS, and the MVI reward heuristic, PLUMES is a novel algorithm in the field of IPP. To compare the performance of PLUMES with state-of-the-art baselines, PLUMES is applied to “static” environments with and without obstacles. This situation generally mimics the scenario in which the target phenomenon transience takes place over durations that are longer than a robotic mission (e.g., seasonal changes versus minutes-long mission). Even

without transience, the problem of optimizing over a static, complex, multimodal, and unknown spatial field is a nontrivial problem, and is the standard demonstration for IPP algorithms. Results in which the spatial distribution varies with respect to the timescale of a mission will be presented in Section 4.6.

4.5.1 Evaluation

Target phenomenon are generated by drawing a (seeded) random sample from a GP with known mean and covariance function. Convex environments are used to illustrate MSS problems in marine and atmospheric domains, in which a mission generally takes place in a geographically bounded, obstacle-free region. Nonconvex environments are used to model terrestrial applications, in which physical barriers or safety “no-go” zones may be placed throughout the target region. Nonconvex environments are simulated by placing rectangular obstacles throughout a bounded mission region, and the robot is endowed the ability to “sense” the obstacles in order to generate safe, valid trajectories at a given planning iteration. Examples of convex and nonconvex environments are shown in Fig. 4-8.

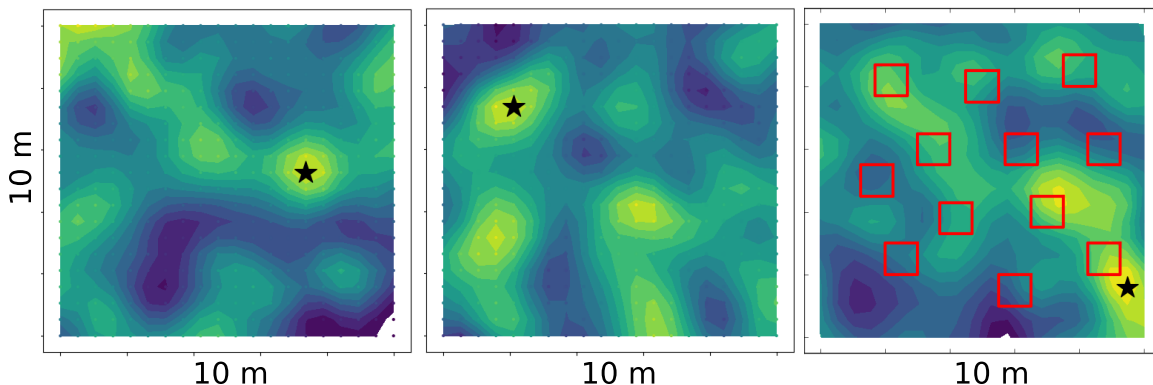


Figure 4-8: **Overview of simulation environments** The multimodal simulated $10\text{m} \times 10\text{m}$ environments. Yellow regions are high-valued; blue regions are low-valued. The global maximum is marked with a star. The left and center environments represent convex-worlds (Section 4.5.2), while the right environment is representative of a non-convex world (Section 4.5.3). Reprinted from [189].

PLUMES is compared against three baselines used in environmental surveying: non-adaptive lawnmower-coverage (Boustro., an abbreviation of boustrophedonic [7]),

greedy myopic planning with UCB reward (UCB-Myopic) [86], and nonmyopic planning with traditional MCTS [110] that uses the maximum-likelihood observation and UCB reward (UCB-MCTS) [85]. The performance of UCB planners has been shown to be sensitive with respect to the β value [85]. In order to avoid subjective tuning, a time-varying β_t that is known to enable no-regret UCB planning [19, 86] is selected. PLUMES uses continuous-observation MCTS with hyperparameters presented in Auger et al. [21] and MVI with spectral sampling [93] in these trials.

To evaluate the mission performance of all planners, the accumulated MSS reward (Eq. 4.9) is reported, which directly corresponds to the number of scientifically valuable samples collected within an ϵ -ball of the true maximum. Several other metrics commonly used in IPP to evaluate posterior model quality are also reported: overall environmental posterior root mean-squared error (RMSE) and error in posterior prediction of \mathbf{x}^* at the end of a mission (\mathbf{x}^* error). A Mann-Whitney U non-parametric significance test [198] is used to report statistical significance ($p = 0.05$ level) in performance between PLUMES and baseline algorithms over repeated trials.

4.5.2 Bounded Convex Environments

PLUMES and baseline planners on a point robot are applied to 50 unique $10\text{ m} \times 10\text{ m}$ simulated environments with multimodal phenomena drawn randomly from a GP prior with a squared-exponential covariance function and zero mean ($l = 1.0$, $\sigma^2 = 100.0$, $\sigma_n^2 = 1.0$ [1%]) (see Fig. 4-8). The action set for the vehicle consisted of ten viable trajectories centered at the robot’s pose with path length 1.5 m, and samples were collected every 0.5 m of travel. Mission lengths were budgeted to be 200 m. Nonmyopic planners rolled out to a 5-action horizon and were allowed 250 rollouts per planning iteration. Summary simulation results are presented in Table 4.1 and MSS reward is illustrated in Fig. 4-9.

Table 4.1: Accumulated True MSS Reward (Eq. 4.9), RMSE, and \mathbf{x}^* Error, Reported as Median (Interquartile Range) for static, convex environments. Asterisks denote baselines whose difference in performance is statistically significant compared to PLUMES. Reprinted from [189].

	Convex Simulation Trials $\epsilon = 1.5$ m, 50 trials			Field Trial $\epsilon = 10$ m, 1 trial
	MSS Reward	RMSE	\mathbf{x}^* Error	MSS Reward
PLUMES	199 (89)	3.8 (9.2)	0.21 (0.23)	524
UCB-MCTS	171 (179)*	3.7 (9.6)	0.24 (0.29)	-
UCB-Myopic	148 (199)*	3.6 (9.2)	0.33 (3.25)	-
Boustro.	27 (3)*	2.7 (10.4)	0.26 (0.46)	63

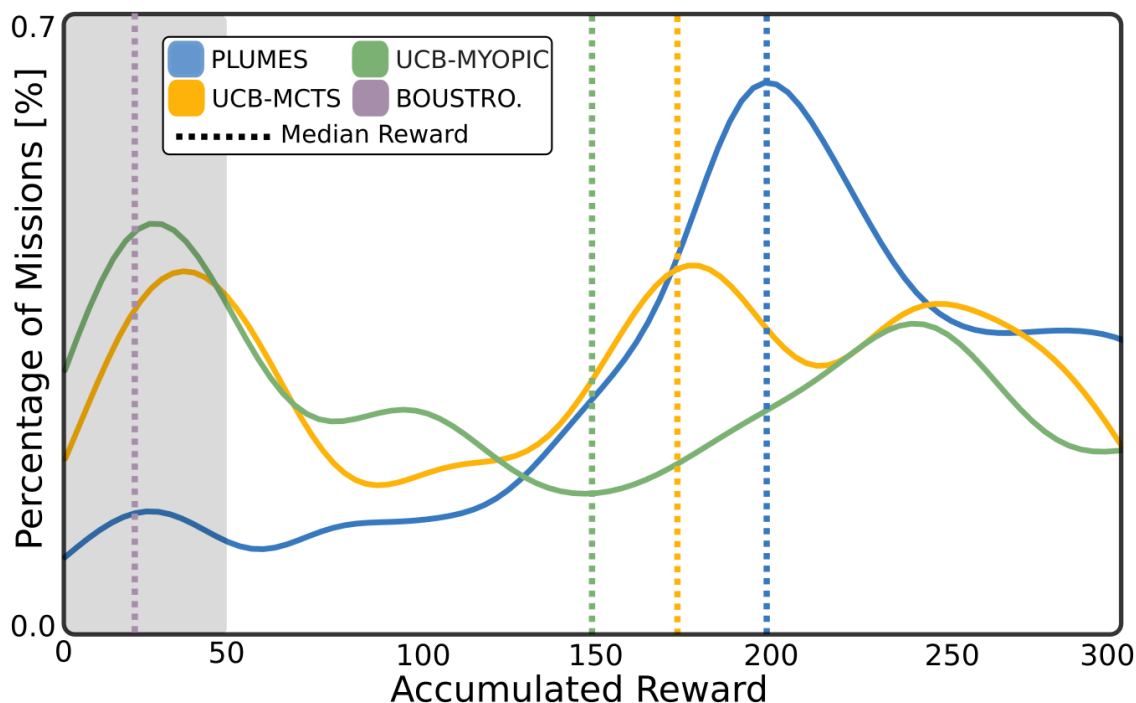


Figure 4-9: **Distribution of accumulated MSS reward in 50 static convex simulations** Accumulated MSS reward is calculated for each trial and the distribution for each planner is plotted as a kernel density estimate (solid line). The dashed lines represent the median accumulated reward for each planner (reported in Table 4.1). The gray area of the plot indicates a low performance region where the planner collected < 50 samples near the maximum. PLUMES has a single mode near 200, whereas both UCB-based methods are multi-modal, with nontrivial modes in the low performance region. Reprinted from [189].

In these trials, PLUMES accumulated significantly (0.05-level) more reward than baselines. The distribution of accumulated reward (Fig. 4-9) shows that PLUMES has a single dominating mode near reward 200 and few low-performing missions (reward <50). In contrast, both UCB-based methods have distributions which are multimodal, with non-trivial modes in the low-performance region. Boustro. collected consistently few scientifically valuable samples, clearly demonstrating sample sparsity in regions of interest. PLUMES additionally achieved statistically indistinguishable levels of posterior RMSE and \mathbf{x}^* error compared to baselines (Table 4.1), indicating that PLUMES preserves exploration performance of the state of the art.

To demonstrate PLUMES in real-world field conditions, maximum search was performed in the Bellairs Fringing Reef, Barbados in January 2019 with a custom-built autonomous surface vehicle (ASV), with the objective of localizing the most exposed coral head. Fig. 4-10 illustrates the motivating application.

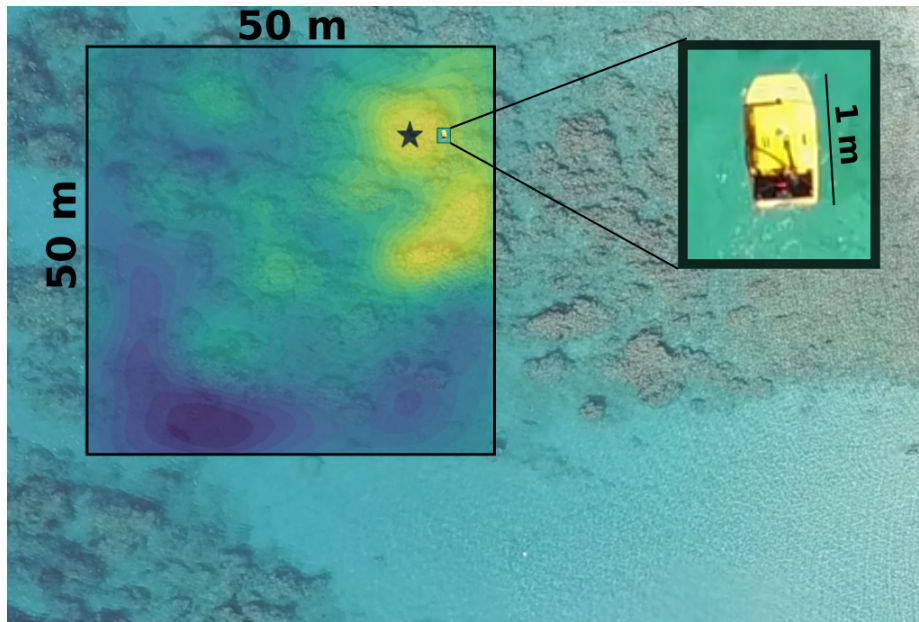


Figure 4-10: **Coral head localization with an autonomous surface vehicle (ASV)**: The objective of the ASV was to find and measure at the most exposed (shallowest) coral head in a region of Bellairs Fringing Reef, Barbados. Overlaid on the aerial photo is the *a priori* unknown bathymetry of the region (yellow is shallow, blue is deep). Equipped with an acoustic point altimeter, the ASV had to explore to infer the location of the maximum (marked with a star) and then measure at that coral colony. Reprinted from [189].

In a reef environment, scientists may be interested in collecting high-resolution images of coral species [199] or to study exposure effects on coral organisms, e.g., ultraviolet radiation [200]. Generally, the best images come from coral heads that are located in more shallow waters and the best target for imaging is the most exposed coral colony. If the location of this colony is unknown, then the ASV must use a “cheap” proxy sensor — in this case an acoustic altimeter to measure depth — to identify the location of the target coral head by taking observations and inferring the location of the maximum (i.e., most exposed coral head). Due to time and resource constraints, only one trial of two planners was feasible on the physical reef; Boustro. was selected to compare against PLUMES as it is one of the most canonical surveying strategies in the marine sciences.

The ASV (1 m \times 0.5 m) had holonomic dynamics and a downward-facing acoustic point altimeter (Tritech Micron Echosounder) with returns at 1 Hz. Ten valid 10 m straight paths radiating from the location of the ASV were used in the action set. The environment was bounded with a 50 m by 50 m geofence. Localization and control was provided by a PixHawk Autopilot with GPS and internal IMU; the fused state estimate was empirically suitable for the desired maximum localization accuracy ($\epsilon = 10$ m). The budget for each mission was 1000 m, which took approximately 45 minutes to travel. A RBF kernel was trained on altimeter data from a dense dataset collected the day before the trials (parameters $l = 2.01$, $\sigma^2 = 0.53$, $\sigma_n^2 = 0.02$ [26%]) to provide useful hyperparameters. Note the high noise in the inferred GP model, as well as the relatively small length-scale in the 2500 m² field site. Although the true topography of the environment was unknown, the trained GP on the dense data collection (shown in Fig. 4-11) served as a “ground-truth” measurement for the purposes of discussion.

PLUMES successfully identified the same coral head to be maximal as that inferred from the GP trained on the dense dataset, as indicated by accumulated reward in Table 4.1, overcoming the challenges of moving in ocean waves, noisy altimeter measurements, and highly multimodal environment. Additionally, the posterior prediction \mathbf{x}^* error (error in the location of the global maximizer) was only 1.78 m while Boustro. reported 8.75 m error due to its non-adaptive sampling strategy.

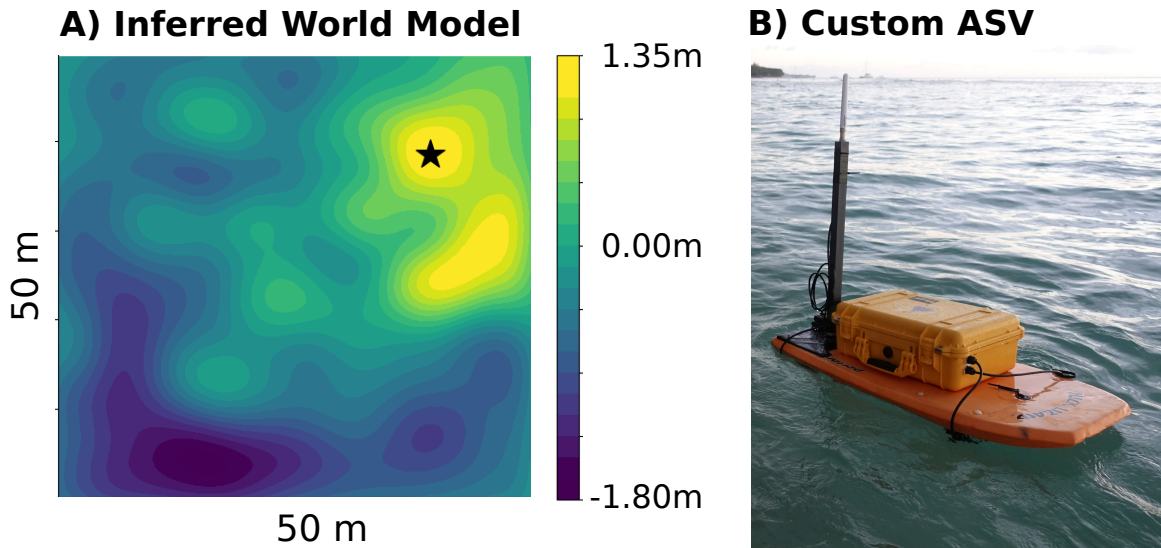


Figure 4-11: **Coral head map and ASV** (A) The ground truth bathymetric map inferred from all collected data, mean corrected in depth. Yellow represents shallower depths, and blue is deeper. The global maximum (most exposed coral head) is marked with a black star. (B) The custom ASV used to traverse the 2500 m² region. Reprinted from [189].

4.5.3 Non-Convex Environments

PLUMES, UCB-Myopic, and UCB-MCTS were deployed in 50 simulated trials with the same simulated phenomenon as in Section 4.5.2, with the addition of 12 block obstacles placed uniformly around the world in known locations (see Fig. 4-8). Boustro. was not used as a baseline in these trials because of the non-generality of the offline approach to complex and potentially unknown geometries. Summary results are presented in Table 4.2 and MSS reward for the simulated trials is inspected in Fig. 4-12.

Table 4.2: Accumulated True MSS Reward (Eq. 4.9), RMSE, and \mathbf{x}^* Error, Reported as Median (Interquartile Range). Asterisks denote baselines whose difference in performance is statistically significant compared to PLUMES. Reprinted from [189].

	Non-convex Simulation Trials $\epsilon = 1.5$ m, 50 trials			Dubins Car Trials $\epsilon = 1.5$ m, 5 trials
	MSS Reward	RMSE	\mathbf{x}^* Error	MSS Reward
PLUMES	206 (100)	3.6 (2.1)	0.25 (0.56)	159 (74)
UCB-MCTS	115 (184)*	3.6 (1.5)	0.27 (1.18)	52 (17)
UCB-Myopic	86 (102)*	3.4 (1.0)	0.23 (0.34)	42 (66)

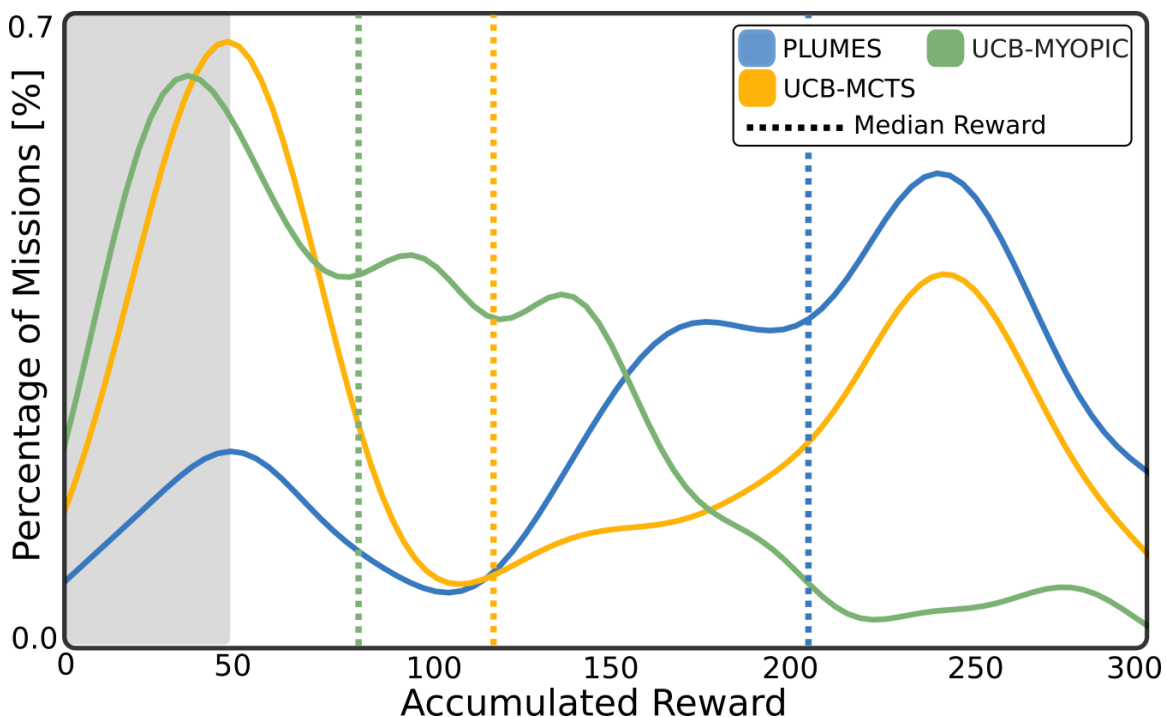


Figure 4-12: **Distribution of accumulated MSS reward in 50 static, non-convex mission simulations** Accumulated MSS reward distribution (solid line) and median (dashed line, reported in Table 4.2) for each planner. The gray area of the plot indicates a low performance region (reward < 50). PLUMES has few low-performing missions and a primary mode near reward 250. The primary mode of both UCB-based methods is in the low performance region due to convergence to suboptimal local maxima. Reprinted from [189].

As indicated in Table 4.2, PLUMES accumulated significantly more (0.05-level) MSS reward than UCB-MCTS and UCB-Myopic. The distribution of reward across the trials is visualized in Fig. 4-12. Like in the convex-world, PLUMES has a primary mode between reward 200-250, while the UCB-based planners have a primary mode in

the low-performance region (reward < 50). There was no significant difference between planners with respect to RMSE or \mathbf{x}^* error. The fact that PLUMES maximized the true MSS reward while achieving statistically indistinguishable error highlights the difference in exploitation efficiency between PLUMES and UCB-based methods.

The simulation experiments assume that a geometric map is known *a priori*. However in practical applications, like indoor gas leak detection, access to a map may be limited or unavailable. A scenario to model MSS in an environment with unknown obstacles is performed with a nonholonomic car equipped with a laser range-finder in a cluttered room. The vehicle must build a map online of the cluttered indoor environment (Fig. 4-13). A simulated chemical phenomenon from a GP ($l = 0.8$, $\sigma^2 = 100.0$, $\sigma_n^2 = 2.0$ [2%]) was used to provide noisy observations to the vehicle at 1 Hz. The action set for the vehicle consisted of eleven 1.5 m Dubins curves projected in front of the vehicle, one straight path behind the vehicle, and a “stay in place” action. Results for five trials are shown in Table 4.2 and illustrate that PLUMES accumulates more MSS reward than baselines, indicating robust performance in a more realistic physical environment and difficult navigation conditions.

As a whole, the simulation and robot trials in convex and nonconvex environments demonstrate the utility of PLUMES compared to canonical and state-of-the-art baselines in a diverse set of environments with challenging practical conditions. For high-stakes scientific deployments, the consistent convergence and sampling performance of PLUMES is critical and beneficial.

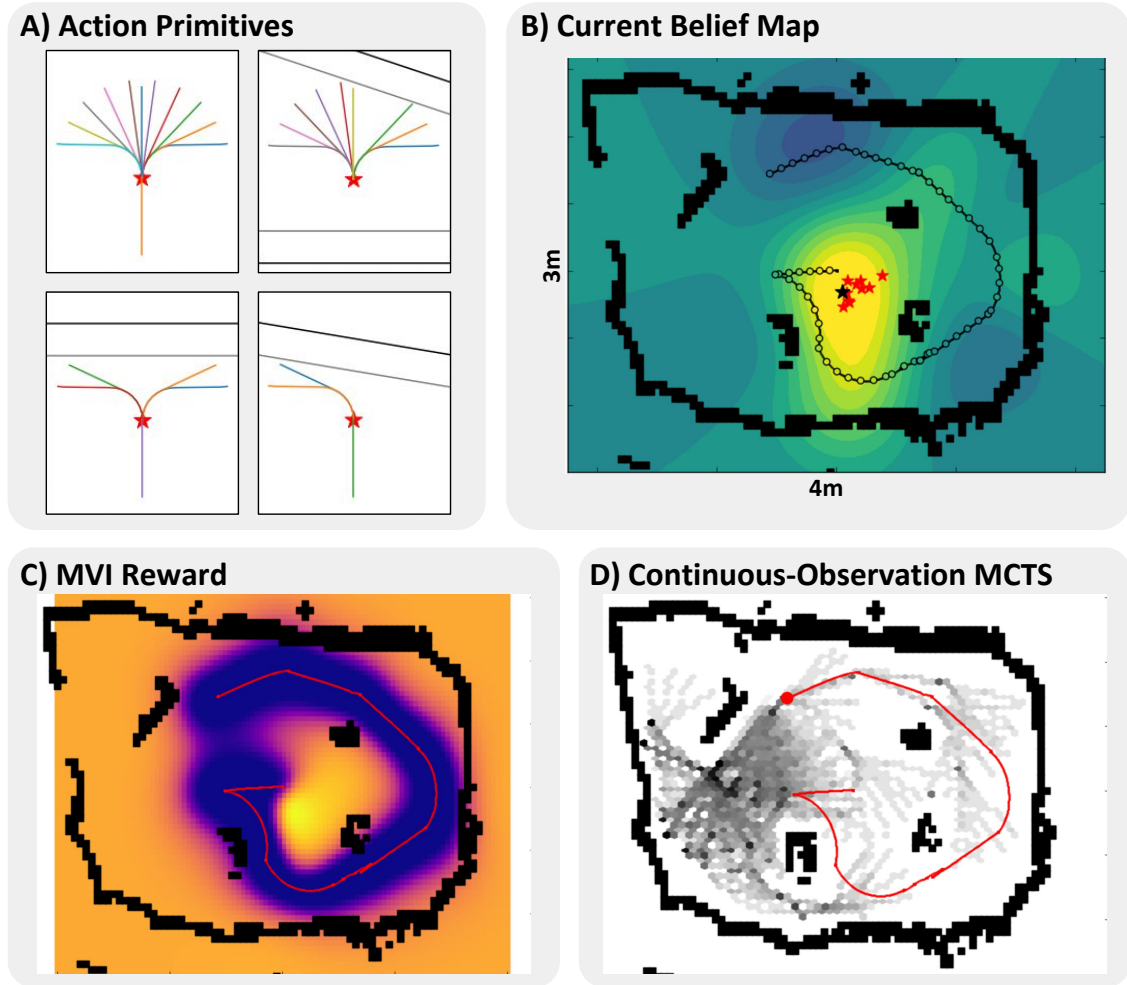


Figure 4-13: **Snapshot of unknown non-convex map scenario:** (A) shows examples of how the action-primitives change based upon obstacle detection (black lines) and safety padding (gray lines). (B-D) show a planning iteration of PLUMES, starting with the current belief map and obstacle detections (B). The MVI heuristic is illustrated in (C) where lighter regions are higher value. (D) shows the rollout visibility of continuous-observation MCTS where darker regions are visited more often. Areas of high reward are generally visited more often by the search as the tree expands. Reprinted from [189].

4.6 Results in Dynamic Environments

The theoretical results of PLUMES directly extend to time-variant systems in which time is a third dimension that can be modeled with the GP belief. This section primarily illustrates the behavior of PLUMES in several representative scenarios in which transience is either unknown/unmodeled, modeled as a third dimension, or modeled with respect to a specific dynamic. These scenarios will be referred to as *maximum-value cycling*, *temporal random-walk*, and *target tracking*. As a point of comparison, the UCB-MCTS baseline is presented. Previous work by Marchant et al. [85] has discussed the use of UCB-MCTS to learn dynamic phenomenon, but demonstrated that there was significant sensitivity in performance with the selection of the explore-exploit parameter β and the depth of the tree. With this in mind, the results are generally used to highlight the performance of PLUMES “out of the box” versus one tuned version of UCB-based systems. In these trials, the same β_t for UCB is used as in Sec. 4.5, and all nonmyopic planners are given the ability to simulate 7 actions ahead, with 150 rollouts per planning iteration. PLUMES uses MVI reward that leverages Gumbel sampling to infer the value of the maximum.

4.6.1 Transience Rejection in Maximum-Value Cycling

In the maximum-value cycling scenario, the spatial distribution of a phenomenon does not change in time, but the mean of the distribution fluctuates according to some underlying period. This would be akin to long-term missions in the Bellairs Fringing Reef under the influence of tides — the maximal coral head is constant in time, but the proxy depth measures are influenced by tides. With respect to maximum seeking, although the spatial distribution does not change, the fluctuation in mean value observations poses a non-trivial challenge. For some arbitrary point in time, the maximum value may be much lower than previous measurements of the maximum, or could even be lower than historical values measured in other locations.

The environmental phenomenon in these trials is modeled by an RBF-Periodic kernel (Sec. 4.2.1) with RBF spatial parameters ($l = 1.5$, $\sigma^2 = 100$) and periodic

temporal parameters ($l = 50, \sigma^2 = 10$). Observations were drawn with noise variance ($\sigma_n^2 = 0.1$). As in the static simulations, PLUMES is endowed with a two-dimensional RBF kernel, with spatial parameters ($l = 1.5, \sigma^2 = 500, \sigma_n^2 = 0.1$); no explicit knowledge about the temporal variation is provided; instead, the variance term is set so that it could potentially “absorb” the cyclic nature of the maximum values. Practically, this is useful when the transient phenomenon may not have a regular or known period such as in systems with episodic fluctuations.

Fig. 4-14 demonstrates the behavior of PLUMES and UCB-MCTS over snapshots of a mission in a representative trial. As showcased by the result, MVI reward is still able to be confidently placed at the location of the global maximum, despite fluctuations. In the calculation of MVI reward, the high variance encoded in the kernel influences sample draws of z^* , which rely on posterior estimates of state with respect to historical values. This in turn allows draws of z^* to be distributed according to the variance that is representative of the periodicity-driven scale of the maximum value, and ultimately still lead to convergence. For UCB-based planners, the same applies, but because variance factors directly into the calculation of the reward heuristic, much more exploration is required before the reward function can converge well to the underlying phenomenon.

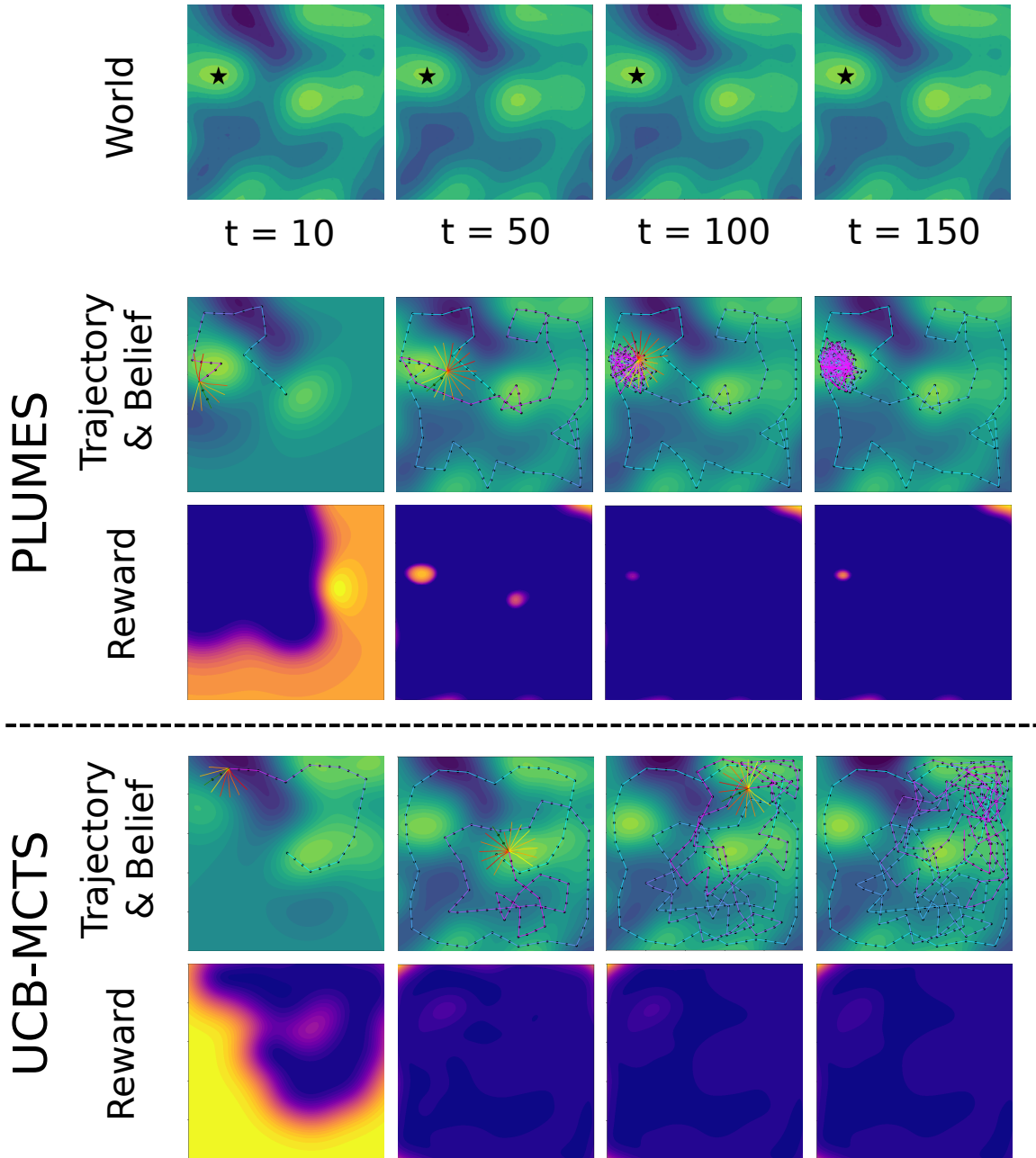


Figure 4-14: **Robust performance under unknown transience** This figure shows the performance of PLUMES and UCB-MCTS in a world with a cyclic mean function. Across the top, the true world at the specified time is pictured, with the maximum marked with a star. Under each planner, the robot's trajectory and reward distribution at the specified time are pictured, wherein the samples along the trajectory are plotted as black dots, and the most recent trajectory selection is colored in pink, and the oldest in blue. High value is colored in yellow in the reward plots. MVI reward in PLUMES is able to converge on the true maximum efficiently, leading to compelling exploitative sampling. UCB baselines generally require more exploration before converging, and the myopic planner showcases a classic failure case of local maxima convergence.

4.6.2 Transience Incorporation

In the temporal random walk scenario, the spatial distribution of the phenomenon of interest stochastically changes, according to a three-dimensional RBF kernel (Eq. 4.5). In natural environments, this may be indicative of short-term transient events like surface-waves on the ocean, or sediment transport. For these illustrative trials, a kernel with spatial parameters ($l = 1.5, \sigma^2 = 100$) and temporal parameters ($l = 100, \sigma^2 = 100$) was used. Observations were drawn with noise parameter $\sigma^2 = 0.1$.

Unlike in the previous scenario, the location of the maximum may change between modes over time. An ideal agent would be able to either predict the change in the maximum location, or perform occasional monitoring actions in order to exploit the time-accurate maximum for sample collection. This requires the agent to directly consider the transience of the phenomenon. A GP kernel with parameters set based on those of the illustrative model is provided to the robot agent, just as was done in simulation trials in Sec. 4.5. A representative scenario which demonstrates monitoring behavior from PLUMES is shown in Fig. 4-15.

Through transience consideration, an agent should be able to track between time-dependent maximum in the environment. Fig. 4-16 shows MSS reward (Eq. 4.9) accumulation for PLUMES and UCB-MCTS for 10 representative trials. As evidenced in both accumulated performance and the representative trial, PLUMES is able to implicitly track the modes of the environment, and transition between modes as posterior predictive estimates of MVI reward change over time; eliciting directed monitoring behavior and accumulating more reward than UCB-MCTS. Indeed, UCB strategies suffer from consideration of the variance, which will generally grow over time everywhere in the world, encouraging extensive exploratory actions over the whole domain.

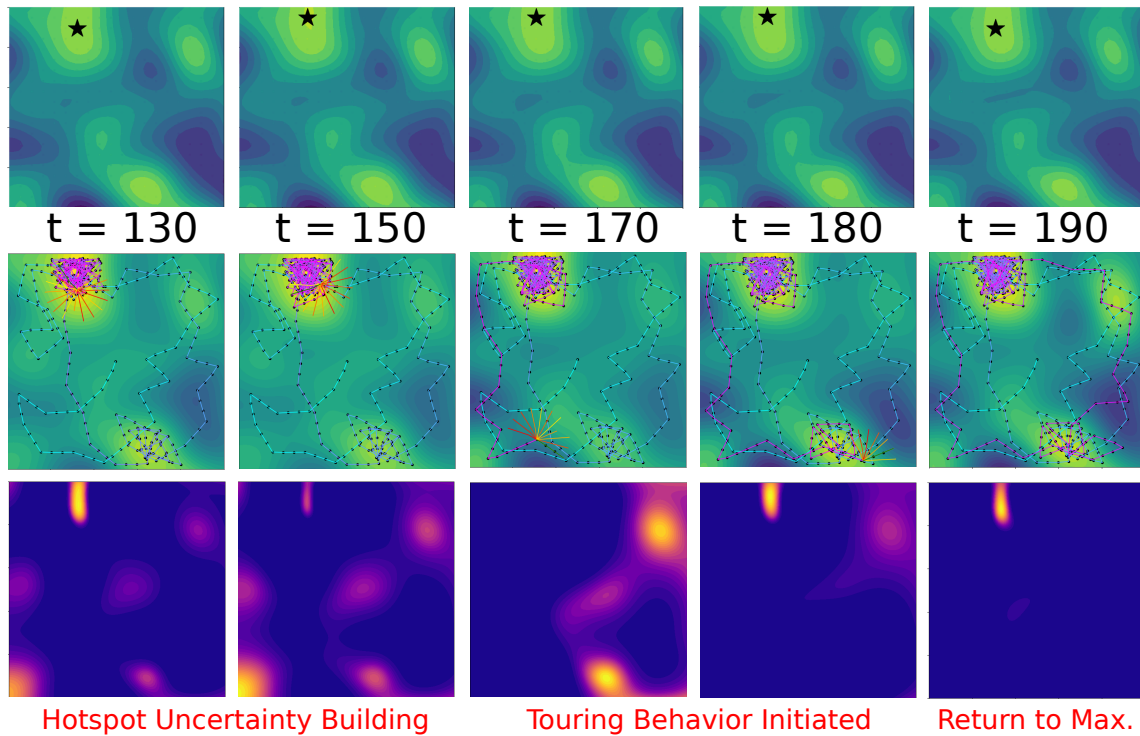


Figure 4-15: **Emergent monitoring behavior with modeled transient phenomena** This figure shows the emergent monitoring behavior of PLUMES in a representative trial. Across the top the true world at the specified time is pictured, with the maximum marked with a star. The robot's trajectory and reward distribution at the specified time are pictured underneath, wherein the samples along the trajectory are plotted as black dots, and the most recent trajectory selection is colored in pink, and the oldest in blue. High value is colored in yellow in the reward plots. Towards the end of the mission, the agent engages in a hotspot monitoring behavior in which modes that were previously seen are revisited, as the uncertainty in the stochastic transition drives the potential reward of returning to the locations up.

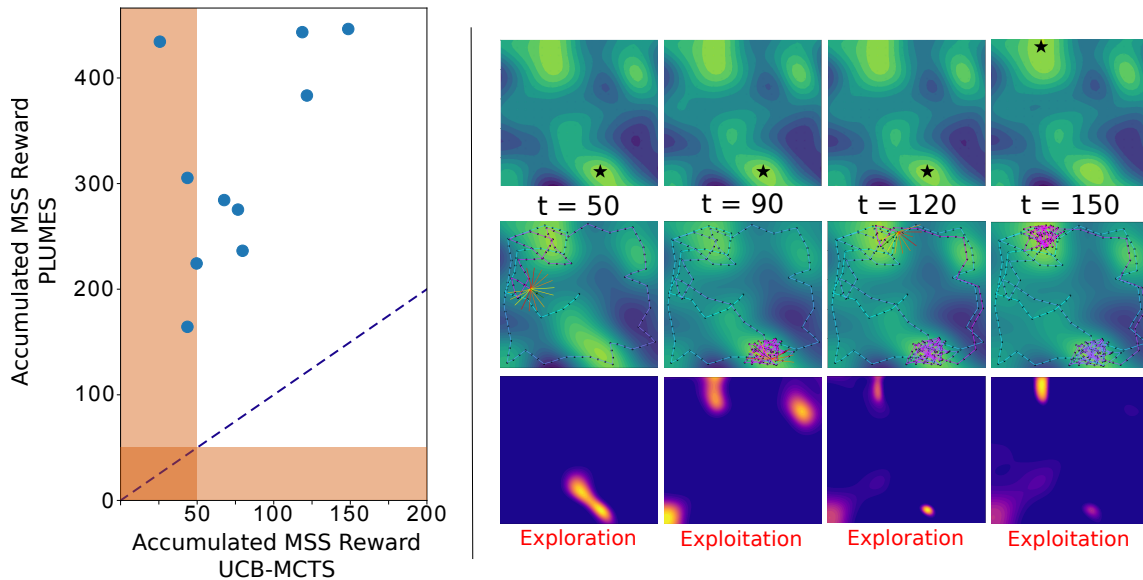


Figure 4-16: **Comparison of PLUMES and UCB-MCTS behavior in known transience** On the left, this figure shows the performance of PLUMES and UCB-MCTS with respect to accumulated MVI reward over 10 represented missions. The dashed line has slope 1, points falling above the line indicate that PLUMES collected more reward for that specific mission than UCB-MCTS. Red-regions indicate low-performance areas, when fewer than 50 samples were collected of the time-varying maximum. Generally, PLUMES outperforms UCB-MCTS in convergence to maximum; UCB-MCTS tends to take many more exploratory actions before converging. On the right, an illustrative PLUMES trial, in which the maximum changes locations, is shown at the top, where the maximum is marked with a star. The world model and reward function for the corresponding time are shown below. As in static regimes, PLUMES initially explores the environment and converges to its belief of the maximum. As time passes, uncertainty about the other high modes in the environment grows, and PLUMES explores these modes again. In this case, this led to convergence to the new maximum location.

4.6.3 Target Tracking

In the target tracking scenario, a singular mode moves in a circle, represented by a custom kernel (Eq. 4.7). The mode begins at the center-top of the environment, and completes a full revolution in 78 time steps. The mode itself has a detectable radius of approximately 1.5 m. Refer to Fig. 4-5 for an illustration.

Similar to the temporal random walk, direct consideration of the phenomenon transience is required. In the trivial case, the robot has complete insight to the trajectory of the phenomenon, and convergence to the mode and tracking can occur in even myopic planning. Fig. 4-17 illustrates how perfect phenomenon knowledge results in explicit near-optimal behavior of the vehicle.

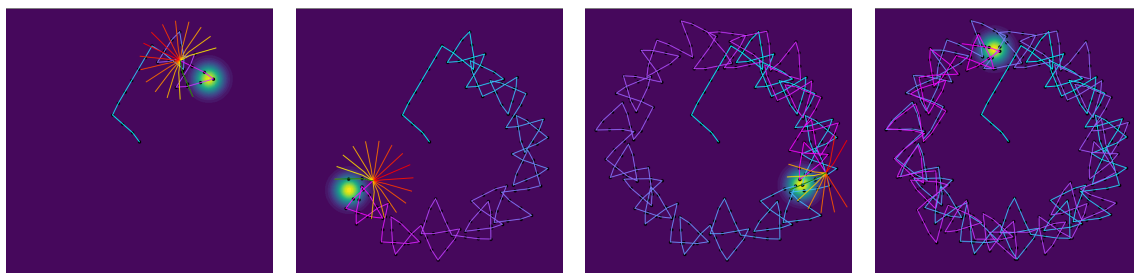


Figure 4-17: **Trivial dynamic tracking case** When any planner has the perfect kernel for the phenomenon, even myopic planners (pictured) can trivially find and track the mode.

However, in many cases, only partial knowledge about a phenomenon may exist, and only intuition about the other factors can be encoded. To illustrate this process, the robot agent is given knowledge of the approximate period of the phenomenon (it is given the period of 80 time intervals for a full revolution), and the center of the rotation. The distance of the mode from the center, the number of modes, and the lengthscale of the mode are unknown. To approximately model the distribution of the system, the following model was used to convert Cartesian and time vectors to time-dependent polar coordinates (r, Θ) :

$$\begin{aligned}
\mathbf{r} &= \sqrt{(\mathbf{x} - c_x)^2 + (\mathbf{y} - c_y)^2} \\
\Theta &= 2\pi \frac{\mathbf{t}}{F} + \arctan\left(\frac{\mathbf{y} - c_y}{\mathbf{x} - c_x}\right) \\
\mathbf{x}_p &= \mathbf{r} \cos(\Theta) \\
\mathbf{y}_p &= \mathbf{r} \sin(\Theta)
\end{aligned} \tag{4.13}$$

where c represents the rotation center, F represents the estimated period of rotation, \mathbf{x} and \mathbf{y} are input coordinate vectors, and \mathbf{t} is input time vector. An RBF kernel (Eq. 4.5) is then placed over the polar coordinates with parameters ($l = 1.5$, $\sigma^2 = 100$). Observations were drawn with noise $\sigma_n^2 = 2.0$. This kernel generally captures that a measurement at some coordinate in time will be correlated with a point along an arc.

By using an approximate kernel and adding significant noise to measurements, the act of finding and tracking the maximum is non-trivial, as uncertainty must be sufficiently reduced everywhere before the maximum can converge to the moving target and confounding measurements will cause uncertainty to fluctuate throughout the mission. Fig. 4-18 shows the convergence-monitoring behavior that PLUMES exhibits with the approximate model. As a point of comparison, UCB-MCTS is also shown for this scenario, as a similar task was presented by Marchant et al. [85] for this algorithm. Unlike the MVI heuristic reward, UCB reward more significantly suffers from the measurement noise, increasing the attractiveness of exploratory behaviors in the mission.

As a whole, these three representative simulation results demonstrate the emergent behaviors of agents using PLUMES in different forms and treatments of transience. They also further reveal the robustness of the MVI heuristic under noisy observations or underlying variance, in addition to the utility of nonmyopic planning in general for strategic monitoring behaviors.

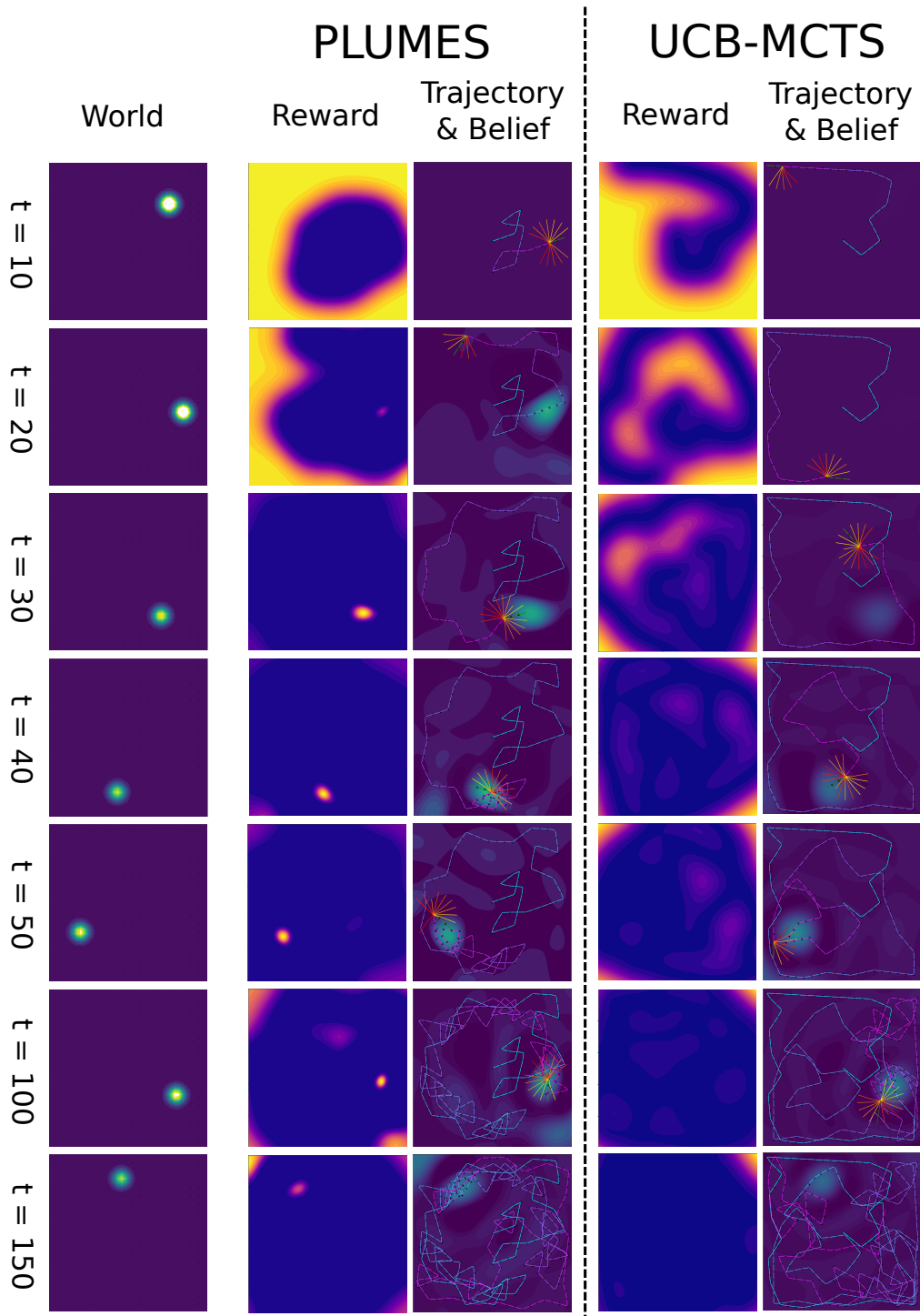


Figure 4-18: **Tracking an unknown moving source** Using an approximate kernel, PLUMES initially explores the world, encountering the maximum during the initial exploration to place a prior on the location. After reducing uncertainty in the world, the agent converges and tracks with the maximum. As time passes, uncertainty in the world state grows and brief exploratory behaviors emerge. The agent re-converges with the maximum by the end of the mission. In contrast, the agent using UCB-MCTS tends towards more exploratory actions throughout the mission.

4.7 Discussion

Online planning methods for robotic maximum seek-and-sample are critical in a variety of contexts, including general environmental monitoring (scientific inquiry, reconnaissance) and disaster response (oil spill, gas leak, radiation). For partially observable environments that can be modeled using a GP, PLUMES is a novel approach for global maximum seek-and-sample that additionally extends to transient phenomenon.

With respect to the constituent elements of PLUMES, both MVI and continuous-observation MCTS provide considerable advantages for adaptive sampling regimes in natural environments. This work presents MVI as an empirically suitable alternative to the canonical GP-UCB heuristic in MSS solvers. MVI is both naturally adaptive and avoids using a hand-tuned parameter to balance exploration and exploitation, instead sampling potential global maxima from the robot’s full belief state to manage exploration and exploitation. This method is additionally robust to transience with respect to placing value at the true maximum. In contrast, heuristic functions like UCB place reward on all high-valued or highly uncertain regions, leading to unnecessary exploration and limiting the time available to exploit knowledge of the true maximum. In transient regimes, this inefficiency is exacerbated by growing uncertainty in previously explored regions over time, making tuning the exploration parameter difficult. Ultimately, the MVI heuristic allows PLUMES to collect exploitative samples, while still achieving in static environments the same overall level of posterior model accuracy (shown by RMSE) as UCB-based planners. Continuous-observation MCTS allows PLUMES to search nonmyopically over belief-spaces on continuous functions without discretization or maximum-likelihood assumptions. Logistically, this allows for navigation in the presence of obstacles and empirically improves finite-time convergence to global maximum.

This work primarily focused on exploiting knowledge about a transient phenomenon in order to select good GP models for belief representation. In scenarios in which nothing is known about a phenomenon of interest, an extension of this work which performs online GP kernel hyperparameter learning [201] could be used. For some

phenomenon and scientific objectives, a GP may not be the best representation. Incorporation of scientific models into PLUMES, in place of or in conjunction with GP belief models, is an open line of potential research. Additionally, it was generally assumed that a discrete action was sufficient for navigation in an environment, such that all reachable space could be accessed by some combination of available actions. In many natural environments, particularly convex environments, such action sets are generally sufficient, however for some nonconvex environments, continuous actions spaces in the spirit of, e.g., Morere et al. [202], would allow increased flexibility and path sophistication.

Recalling the field campaigns from Chapter 3, PLUMES is most immediately applicable in environments like the shallow seep fields in the Cascadia Margin, where ephemeral supersaturations are the target of interest. PLUMES can be used without modification to effectively, autonomously conduct similar campaigns. For operations like those in the Wareham River where the location of the maximum is already known (i.e., the outfall pump), or like those in Cambridge Bay where there is no clear maximum in the environment, PLUMES could be used to optimize over a meta-distribution which encodes more abstract scientific quantities. For example, rather than optimize over raw measurements, PLUMES could be used to optimize a distribution which represent flow velocities, spatial gradients, or co-dependent phenomena.

Ultimately, this chapter formalizes the MSS POMDP and presents PLUMES, an adaptive planning algorithm that employs continuous-observation MCTS and maximum-value information reward to perform efficient maximum-seeking in partially observable, continuous environments. PLUMES outperforms canonical coverage and UCB-based state-of-the-art methods with statistical significance in challenging simulated and real-world conditions (e.g., multiple local maxima, unknown obstacles, sensor noise) and demonstrates robust performance under transient conditions. Maximum seek-and-sample is a critical task in environmental monitoring for which PLUMES, with theoretical convergence guarantees, strong empirical performance, and robustness under real-world conditions, is well-suited.

Chapter 5

Conclusions

The study of natural phenomena forms the core of the environmental and earth sciences. To examine scientific questions about natural environments, *in situ* observations and samples are required to draw conclusions. Sample collection can be logistically time-consuming, difficult, and dangerous for a human operator, and choosing where to draw representative or valuable samples is an open technical challenge as the distribution of the phenomenon may be unknown or ephemeral. Robotic technologies, in conjunction with *in situ* instrumentation, are poised to disrupt sample methodologies in the environmental sciences due to the many-fold more observations that can be collected in longer, unsupervised missions. This presents an opportunity to both demonstrate the utility of robotic platforms in various field conditions and set new standards in sampling regimes.

This thesis presented in Chapter 3 scientific findings from three field campaigns that used an unmanned surface vehicle equipped with gas sensors to examine two greenhouse gases (CH_4 and CO_2) in unique marine environments. As a whole, the results of these campaigns serve as a compelling case study for robotic technologies, and highlighted a key challenge that should be addressed by sampling regimes: phenomenon transience. In many natural environments, a target of interest is subject to time-variation; whether the phenomenon itself dynamically moves or is influenced by changing environmental conditions (e.g., tides, seasons).

To begin to address sampling regimes in natural environments, this thesis con-

siders a specific scientific sampling objective, the *maximum seek-and-sample* (MSS) problem. A scientist poses the MSS problem when seeking a gaseous plume source or searching for the most nutrient rich soils to collect a time-consuming physical sample. Canonically, uniform-coverage strategies have been adopted in the environmental and earth sciences in order to collect representative observations, however in the MSS problem and in environments with transient phenomenon, these methods tend to be sample sparse in the region of interest. To address sample sparsity, an intelligent and adaptive regime is necessary. Chapter 4 formalizes the MSS problem as a partially-observable Markov decision process (POMDP) and proposes the PLUMES adaptive sampling algorithm to approximately solve it, demonstrating better performance than state of the art baselines in a variety of static, dynamic, convex, and nonconvex simulations and field trials.

5.1 Thesis Contributions

This thesis offers contributions in both the marine sciences and in adaptive sampling.

5.1.1 Marine Sciences

With respect to the marine sciences, the three field campaigns each present novel results in their domains. Logistically, the ChemYak [118] vehicle played a critical role in the collection of all measurements, and this thesis illustrates the use of the vehicle in rivers, open seas, and Arctic environments.

The Wareham River Estuary field study in Sec. 3.3 examined the impact of wastewater effluent on a tidal salt-water estuary in Massachusetts. The bacteria used to treat the water at one point of processing required methanol feedstock to be injected in the water, and the campaign was motivated by the question of whether the bacteria were fully metabolizing the stock before the water was pumped into the estuary. The results of this work demonstrated that CH_4 concentrations in the treated water were elevated at the wastewater pump outfall, and both CO_2 and CH_4 were longitudinally transported by the river from the outfall, driven by tidal cycles. The

spatial and temporal resolution of gas transport generally goes unmeasured in wastewater effluent studies and these observations show the strong influence of the ebbing tide on longitudinal transport of greenhouse gases in the estuary.

Sec. 3.4 presented results from the Cascadia Margin, an area with natural CH₄ seeps from hydrate instability and bacterial activity [174, 175]. Previous studies [179, 180] hypothesized that supersaturations on surface waters may be possible in shallow water sites, driven by bubble transport and coastal upwelling. Evidence of these supersaturations in the top 10 m of the water column were presented and observed at two independent locations, in addition to generally elevated CH₄ levels in surface waters above active venting sites. Based on the ChemYak measurements, a conservative estimate of the potential quantity of CH₄ that reaches the atmosphere from the campaign sites is 20-250 kg y⁻¹. These results add to the small body of literature which specifically estimates contributions from shallow seeps.

Sec. 3.5 presented an extended field campaign in Cambridge Bay, Nunavut, Canada to capture a major CH₄ and CO₂ outgassing event during the annual spring freshet in which the ice cap over the Freshwater Creek estuary recedes. The results not only revealed extremely elevated gas concentration in river-derived waters, but also showed that considerable lateral transport of gas to the coastal ocean likely occurs in the early freshet and rapid outgassing to the atmosphere occurs in recently ice-free waters. With typical sampling methods, this event would largely be under-constrained or potentially missed, as summer-time, ice-free measurements are the standard in biogeochemical studies of Arctic regions. Analysis from measurements drawn in this study were used to estimate that 95% of greenhouse gas emissions from Cambridge Bay over the course of an entire year likely occurs during the spring freshet alone. This work has far-reaching implications about the overall contribution of atmospheric CH₄ and CO₂ from Arctic regions and compels future studies which continue to resolve transient phenomena in Arctic waters.

5.1.2 Informative Path Planning and Adaptive Sampling

This thesis further presents the MSS POMDP formalism, and derives the PLUMES algorithm which utilizes Gaussian process (GP) belief representations of natural phenomenon, maximum value-information (MVI) reward heuristic, and continuous-observation Monte Carlo tree search (MCTS) to optimize over the MSS problem. GP belief states offer a compact representation over the uncountably infinite number of possible realizations of a phenomenon’s distribution with respect to a history of continuous observations. By leveraging knowledge of a scientific phenomenon (e.g., diffusivity, trajectory), this work demonstrates that the GP kernel function can encode this insight for efficient prediction in complex, multimodal, and potentially transient environments.

The true reward of the MSS problem — value placed only at the global maximum — is difficult to optimize over as the signal is sparse. State-of-the-art methods use the Upper Confidence Bound (UCB) heuristic to find the global maximum of a distribution, however this reward function converges to the underlying distribution of a phenomenon and can, in practice, yield suboptimal convergence to local maxima in finite time missions. Moreover, UCB has a hand-tuned parameter to control the explore-exploit behavior of a vehicle, and this parameter can be difficult to tune in one-shot missions. The MVI reward heuristic, first presented as a black-box optimizer in Bayesian Optimization [20], has no tunable parameter and instead converges to place value (after sufficient exploration) to the true global maximum. This leads to efficient seek and sample behavior in finite duration missions. This thesis further demonstrates that calculating MVI is possible in time-varying distributions, and can lead to attractive monitoring behaviors in difficult, transient regimes.

Finally, continuous-observation MCTS uses progressive-widening to search over continuous state and observation spaces, which are typically difficult to search over because every state is unique with probability 1. This thesis extends performance guarantees for fully-observable MDPs solved with continuous-observation MCTS [21] to the partially observable domain.

The demonstration of PLUMES and state-of-the-art baselines in static and dynamic environments, convex and nonconvex metric spaces (with known and unknown obstacles), and in the presence of a variety of illustrative transience shows that PLUMES is a robust tool for robotic exploration, mapping, and sample collection. Together with the technical content, this work serves as a comprehensive presentation of PLUMES for the MSS problem.

5.2 Future Work

Several directions for future efforts in developing robotic solutions for environmental and earth science objectives include: *novel representations of scientific phenomenon*, *multi-objective missions*, *long-term monitoring regimes*, and *multi-agent fleets*.

5.2.1 Representing Scientific Phenomenon for Planning

GPs are primarily leveraged in this thesis to represent target phenomena. The advantage of a GP lies in the compact representation of a state (through mean and kernel function) and closed form data incorporation and posterior inference. For complex phenomenon, GPs may not be expressive enough or simply may not have features that can be described using a Gaussian relationship. To improve the flexibility of GPs, kernel learning [203, 204] is one potential avenue for development, in which using observations collected online or through a training dataset offline, the hyperparameters of a kernel function can be optimized to better conform the kernel to the relationships inherent in the data. However, alternative representations for some environmental phenomenon will still be required. A rich history of science modeling and analytical computation, e.g., [205], presents an opportunity to incorporate first principle specifications with probabilistic or learning frameworks for robotic decision-making. Bayesian networks [206], neural networks and genetic programming [207], topic modeling [208], and other methods are among some representations already under consideration in the field.

5.2.2 Multi-Objective Missions

This thesis primarily examined scientific objectives which require optimizing over a single “task” or target. However, in some applications, multiple phenomena may be of interest and the science objective may be more general (e.g., mapping several phenomenon, finding and taking note of “interesting” phenomena, exploration). Multi-objective missions describe these applications, and pose a significant set of challenges to address, including: learning relationships between phenomena, optimal planning in multi-dimensional spaces, compact belief representation, encoding “curiosity” onto a platform. Some work in this space has examined heuristics [209], Petri nets [210], and particle swarm optimization [211] among other techniques to consider multiple objectives that must be met by a single agent. Work specific to multi-objective planning with respect to information-measures is an open area.

5.2.3 Longterm Monitoring

Some realistic objectives of science missions may require multiple deployments or longterm monitoring (over weeks or months) to detect events, characterize temporal trends, or map large fields. Some monitoring objectives may include MSS, hotspot coverage, mapping, exploration, gradient characterization, and others. Longterm monitoring requires data management solutions, consideration of power requirements, and physical robustness of the platform and instruments. From a decision-making and modeling perspective, longterm monitoring presents an opportunity to consider how scale of an optimization problem changes potential robotic behaviors. For example, in a small world it may be feasible for a robot to move from one end to the other multiple times in a mission, but in large environments such transit may be forbidden. This requires consideration of how to perform no-regret actions in local regions, long-range strategic planning, and periodic mission reformulation based on incremental data updates. Additionally, longterm monitoring may benefit from incorporation of remote data streams (e.g., satellites) and other novel sensing technologies to improve strategic planning.

5.2.4 Multi-Agent Systems

Robot swarms or multi-agent fleets have direct utility in the environmental sciences: coordinated agents could distribute workload in longterm monitoring applications, allow for specialized task-management for multi-objective missions, and provide a dense temporal “snapshot” of a spatial phenomenon. All of the challenges of using a single agent, in addition to developing communication, coordinate, and planning frameworks can be addressed by work in this domain.

5.3 Final Thoughts

This thesis proposes that natural phenomenon are an interesting context for algorithm development in robotics, and presents an argument for robotic systems to be incorporated into environmental and earth science practices. To tackle difficult questions in the environmental and earth sciences, massive amounts of spatial and temporal data are needed, and this data cannot be collected through standard sampling techniques alone. Remote sensing through satellites has offered unprecedented global resolution of phenomenon, but mid- to small-scale phenomenon, and calibration of these sensors, require *in situ* examination. New technologies are immediately necessary, and robotics is well-suited to respond. Within robotics, using science applications as a means of inspiration for technical innovation requires consideration of both theoretical and practical aspects of intelligent design to create robust agents which can elicit community trust, guarantee efficacy, and produce inherently useful products (e.g., raw observations, belief models/representations). In the pursuit of resolving difficult science questions and creating sophisticated robotic technologies, multidisciplinary and diverse perspectives, like those presented in this thesis, will be necessary.

Bibliography

- [1] Dean Roemmich, Gregory C. Johnson, Stephen Riser, Russ Davis, John Gilson, W. Brechner Owens, Silvia L. Garzoli, Claudia Schmid, and Mark Ignaszewski. The Argo Program: Observing the global ocean with profiling floats. *Oceanography*, 22(2):34–43, 2009.
- [2] R. Krishfield, J. Toole, A. Proshutinsky, and M.-L. Timmermans. Automated ice-tethered profilers for seawater observations under pack ice in all seasons. *J. Atmospheric Ocean. Technol.*, 25(11):2091–2105, 2008.
- [3] D. Yoerger, J. Newman, and J.J. Slotine. Supervisory control system for the JASON ROV. *J. Ocean. Eng.*, 11(3):392–400, 1986.
- [4] Lawrence H. Keith. *Environmental sampling and analysis: a practical guide*. Routledge, 2017.
- [5] Netta Shalev, Juraj Farkaš, Jan Fietzke, Martin Novak, Jan A Schuessler, Philip A.E. Pogge von Strandmann, and Philip B. Törber. Mg isotope interlaboratory comparison of reference materials from earth-surface low-temperature environments. *Geostand. Geoanalytical Res.*, 42(2):205–221, 2018.
- [6] Taku Umezawa, Carl Brenninkmeijer, Thomas Röckmann, Carina Van Der Veen, Stanley C. Tyler, Ryo Fujita, Shinji Morimoto, Shuji Aoki, Todd Sowers, Jochen Schmitt, et al. Interlaboratory comparison of $\delta^{13}\text{C}$ and δD measurements of atmospheric CH_4 for combined use of data sets from different laboratories. *Atmospheric Meas. Tech.*, 11:1207–1231, 2018.
- [7] Howie Choset and Philippe Pignon. Coverage path planning: The boustrophedon cellular decomposition. In *Proc. Field Service Robot.*, pages 203–209, 1998.
- [8] Ryan N. Smith, Mac Schwager, Stephen L. Smith, Daniela Rus, and Gaurav S. Sukhatme. Persistent ocean monitoring with underwater gliders: Towards accurate reconstruction of dynamic ocean processes. In *IEEE Int. Conf. Robot. Autom.*, pages 1517–1524. IEEE, 2011.
- [9] Samuel M. Smith, Stanley E. Dunn, Thomas L. Hopkins, Karl Heeb, and Tom Pantelakis. The application of a modular AUV to coastal oceanography: case study on the ocean explorer. In *Proc. IEEE/MTS Int. Conf. Oceans*, volume 3, pages 1423–1432. IEEE, 1995.

- [10] Edward Tunstel, John M. Dolan, Terrence Fong, and Debra Schreckenghost. Mobile robotic surveying performance for planetary surface site characterization. In *Performance evaluation and benchmarking of intelligent systems*, pages 249–268. Springer, 2009.
- [11] Stephen Williams, Lonnie T. Parker, and Ayanna M. Howard. Terrain reconstruction of glacial surfaces: Robotic surveying techniques. *IEEE Robot. Autom. Mag.*, 19(4):59–71, 2012.
- [12] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artif. Intell.*, 101(1/2):99–134, 1998.
- [13] Omid Madani, Steve Hanks, and Anne Condon. On the undecidability of probabilistic planning and infinite-horizon partially observable Markov decision problems. In *Proc. Innovative Applicat. Artif. Intell.*, pages 541–548, 1999.
- [14] Zhan Wei Lim, David Hsu, and Wee Sun Lee. Adaptive informative path planning in metric spaces. *Int. J. Robot. Res.*, 35(5):585–598, 2016.
- [15] Akash Arora, P. Michael Furlong, Robert Fitch, Salah Sukkarieh, and Terrence Fong. Multi-modal active perception for information gathering in science missions. In *Proc. Int. Symp. Auton. Robots*, pages 1–27, 2017.
- [16] Carl E. Rasmussen and Christopher K.I. Williams. *Gaussian processes for machine learning.*, volume 14. MIT Press MIT Press, 2004.
- [17] W.D. Smart and L. Pack Kaelbling. Effective reinforcement learning for mobile robots. In *Proc. IEEE Int. Conf. Robot. Autom.*, volume 4, pages 3404–3410, 2002.
- [18] Emile Contal, David Buffoni, Alexandre Robicquet, and Nicolas Vayatis. Parallel Gaussian process optimization with upper confidence bound and pure exploration. In *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*, pages 225–240. Springer, 2013.
- [19] Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias W. Seeger. Information-Theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting. *IEEE Trans. Inf. Theory*, 58:3250–3265, May 2012.
- [20] Zi Wang and Stefanie Jegelka. Max-value entropy search for efficient Bayesian optimization. In *Proc. 34th Int. Conf. Mach. Learn.*, volume 70, pages 3627–3635, 2017.
- [21] David Auger, Adrien Couetoux, and Olivier Teytaud. Continuous upper confidence trees with polynomial exploration–consistency. In *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*, pages 194–209, 2013.

- [22] Adrien Couëtoux, Jean-Baptiste Hooek, Nataliya Sokolovska, Olivier Teytaud, and Nicolas Bonnard. Continuous upper confidence trees. In *Proc. Int. Conf. Learn. Intell. Optim.*, pages 433–445, 2011.
- [23] Bin Zhang and Gaurav S. Sukhatme. Adaptive sampling for estimating a scalar field using a robotic boat and a sensor network. In *Proc. IEEE Int. Conf. Robot. Autom.*, pages 3673–3680. IEEE, 2007.
- [24] Mohammad Rahimi, Mark Hansen, William J. Kaiser, Gaurav S. Sukhatme, and Deborah Estrin. Adaptive sampling for environmental field estimation using robotic sensors. In *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, pages 3692–3698. IEEE, 2005.
- [25] Mohammad Rahimi, Richard Pon, William J. Kaiser, Gaurav S. Sukhatme, Deborah Estrin, and Mani Srivastava. Adaptive sampling for environmental robotics. In *Proc. IEEE Int. Conf. Robot. Autom.*, volume 4, pages 3537–3544. IEEE, 2004.
- [26] Wenhao Luo and Katia Sycara. Adaptive sampling and online learning in multi-robot sensor coverage with mixture of Gaussian processes. In *Proc. IEEE Int. Conf. Robot. Autom.*, pages 6359–6364. IEEE, 2018.
- [27] Yew Teck Tan, Abhinav Kunapareddy, and Marin Kobilarov. Gaussian process adaptive sampling using the cross-entropy method for environmental sensing and monitoring. In *Proc. IEEE Int. Conf. Robot. Autom.*, pages 6220–6227. IEEE, 2018.
- [28] Dieter Fox, Wolfram Burgard, Frank Dellaert, and Sebastian Thrun. Monte Carlo localization: Efficient position estimation for mobile robots. *Proc. Innovative Applicat. Artif. Intell.*, 1999(343-349):2–2, 1999.
- [29] Cyrill Stachniss, Giorgio Grisetti, Dirk Hähnel, and Wolfram Burgard. Improved Rao-Blackwellized mapping by adaptive sampling and active loop-closure. In *Proc. Workshop Self-Organization of Adaptive behavior*, pages 1–15, 2004.
- [30] Tomás de J Mateo Sanguino and Francisco Ponce Gómez. Toward simple strategy for optimal tracking and localization of robots with adaptive particle filtering. *IEEE/ASME Trans. Mechatron.*, 21(6):2793–2804, 2016.
- [31] Hao Lang, Tiancheng Li, Gabriel Villarrubia, Shudong Sun, and Javier Bajo. An adaptive particle filter for indoor robot localization. In *Ambient Intelligence-Software and Applications*, pages 45–55. Springer, 2015.
- [32] M. Zwicker, W. Jarosz, J. Lehtinen, B. Moon, R. Ramamoorthi, F. Rousselle, P. Sen, C. Soler, and S.E. Yoon. Recent advances in adaptive sampling and reconstruction for Monte Carlo rendering. In *Computer Graphics Forum*, volume 34, pages 2:667–2:681. Wiley Online Library, 2015.

- [33] Toshiya Hachisuka, Wojciech Jarosz, Richard Peter Weistroffer, Kevin Dale, Greg Humphreys, Matthias Zwicker, and Henrik Wann Jensen. Multidimensional adaptive sampling and reconstruction for ray tracing. In *ACM Trans. Graphic.*, volume 27, pages 33:1–33:10. ACM, 2008.
- [34] M.D. Zakirul Alam Bhuiyan, J. Wu, G. Wang, T. Wang, and M.M. Hassan. e-sampling: Event-sensitive autonomous adaptive sensing and low-cost monitoring in networked sensing systems. *ACM Trans. Auton. Adapt. Sys.*, 12(1):1–29, 2017.
- [35] Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies. *J. Mach. Learn. Res.*, 9:235–284, 2008.
- [36] Yuan Ren, Yu Ding, and Faming Liang. Adaptive evolutionary Monte Carlo algorithm for optimization with applications to sensor placement problems. *Stat. Comput.*, 18(4):375–390, 2008.
- [37] Paul N. Edwards. *Representing the global atmosphere: Computer models, data, and knowledge about climate change*, volume 1. MIT Press Cambridge, MA, 2001.
- [38] Iraj Javandel, Christine Doughty, and Chin-Fu Tsang. Groundwater transport: Handbook of mathematical models. Technical report, Lawrence Berkeley Lab., CA (USA), 1984.
- [39] Dale B. Haidvogel and Aike Beckmann. *Numerical ocean circulation modeling*. World Scientific, 1999.
- [40] John Wainwright and Mark Mulligan. *Environmental modelling*. Wiley Online Library, 2002.
- [41] Frédéric Hourdin, Thorsten Mauritsen, Andrew Gettelman, Jean-Christophe Golaz, Venkatramani Balaji, Qingyun Duan, Doris Folini, Duoying Ji, Daniel Klocke, Yun Qian, et al. The art and science of climate model tuning. *Bull. Amer. Meteor. Soc.*, 98(3):589–602, 2017.
- [42] Christopher K. Wikle. Hierarchical models in environmental science. *Int. Stat. Rev.*, 71(2):181–199, 2003.
- [43] Manouchehr Amini, Kim Mueller, Karim C. Abbaspour, Thomas Rosenberg, Majid Afyuni, Klaus N. Møller, Mamadou Sarr, and C. Annette Johnson. Statistical modeling of global geogenic fluoride contamination in groundwaters. *Environ. Sci. Technol.*, 42(10):3662–3668, 2008.
- [44] Noel Cressie. The origins of kriging. *Math. Geol.*, 22(3):239–252, 1990.
- [45] Alberto Elfes. Using occupancy grids for mobile robot perception and navigation. *Computer*, 22(6):46–57, 1989.

- [46] James S. Clark. Why environmental scientists are becoming Bayesians. *Ecol. Lett.*, 8(1):2–14, 2005.
- [47] Martin J. Wainwright, Michael I. Jordan, et al. Graphical models, exponential families, and variational inference. *Foundations and Trends in Machine Learning*, 1(1–2):1–305, 2008.
- [48] Christopher M. Bishop. *Pattern recognition and machine learning*. springer, 2006.
- [49] Matthew D. Hoffman, David M. Blei, Chong Wang, and John Paisley. Stochastic variational inference. *J. Mach. Learn. Res.*, 14(1):1303–1347, 2013.
- [50] Alp Kucukelbir, Dustin Tran, Rajesh Ranganath, Andrew Gelman, and David M. Blei. Automatic differentiation variational inference. *J. Mach. Learn. Res.*, 18(1):430–474, 2017.
- [51] David J.C. MacKay. Introduction to Monte Carlo methods. In *Learning in graphical models*, pages 175–204. Springer, 1998.
- [52] Peter J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4):711–732, 1995.
- [53] Radford M. Neal et al. MCMC using Hamiltonian dynamics. *Handbook of Markov chain Monte Carlo*, pages 113–162, 2011.
- [54] Zoubin Ghahramani. An introduction to hidden Markov models and Bayesian networks. In *Hidden Markov models: applications in computer vision*, pages 9–41. World Scientific, 2001.
- [55] P.A. Aguilera, A. Fernández, R. Fernández, R. Rumí, and A. Salmerón. Bayesian networks in environmental modelling. *Environ. Model. Soft.*, 26(12):1376–1388, 2011.
- [56] Mario A.T. Figueiredo and Anil K. Jain. Unsupervised learning of finite mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(3):381–396, 2002.
- [57] Thomas S. Ferguson. A Bayesian analysis of some nonparametric problems. *Ann. Stat.*, pages 209–230, 1973.
- [58] Thomas L. Griffiths, Michael I. Jordan, Joshua B. Tenenbaum, and David M. Blei. Hierarchical topic models and the nested Chinese restaurant process. In *Proc. Adv. Neural Inf. Process. Syst.*, pages 17–24, 2004.
- [59] William Kleiber, Richard W. Katz, and Balaji Rajagopalan. Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resour. Res.*, 48(1), 2012.

- [60] Niamh Cahill, Andrew C. Kemp, Benjamin P. Horton, Andrew C. Parnell, et al. Modeling sea-level change using errors-in-variables integrated Gaussian processes. *Ann. Appl. Stat.*, 9(2):547–571, 2015.
- [61] Ruofei Ouyang, Kian Hsiang Low, Jie Chen, and Patrick Jaillet. Multi-robot active sensing of non-stationary Gaussian process-based environmental phenomena. In *Proc. Int. Conf. Auton. Agents Multi-Agent Sys.*, pages 573–580. International Foundation for Autonomous Agents and Multiagent Systems, 2014.
- [62] James S. Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for hyper-parameter optimization. In *Proc. Adv. Neural Inf. Process. Syst.*, pages 2546–2554, 2011.
- [63] Ralf Herbrich, Neil D. Lawrence, and Matthias Seeger. Fast sparse Gaussian process methods: The informative vector machine. In *Proc. Adv. Neural Inf. Process. Syst.*, pages 625–632, 2003.
- [64] Ronald A. Howard. *Dynamic programming and Markov processes*. John Wiley, 1960.
- [65] Richard Bellman. A Markovian decision process. *J. Math. Mechanics*, pages 679–684, 1957.
- [66] David J.C. MacKay. Information-based objective functions for active data selection. *Neural Comput.*, 4(4):590–604, 1992.
- [67] Valerii Vadimovich Fedorov. *Theory of optimal experiments*. Elsevier, 2013.
- [68] Robert Sim and Nicholas Roy. Global a-optimal robot exploration in SLAM. In *Proc. IEEE Int. Conf. Robot. Autom.*, pages 661–666. IEEE, 2005.
- [69] Thomas Kollar and Nicholas Roy. Trajectory optimization using reinforcement learning for map exploration. *Int. J. Robot. Res.*, 27(2):175–196, 2008.
- [70] Henry Carrillo, Yasir Latif, Maria L. Rodriguez-Arevalo, José Neira, and José A. Castellanos. On the monotonicity of optimality criteria during exploration in active SLAM. In *IEEE Int. Conf. Robot. Autom.*, pages 1476–1483. IEEE, 2015.
- [71] Siddharth Joshi and Stephen Boyd. Sensor selection via convex optimization. *IEEE Trans. Signal Process.*, 57(2):451–462, 2008.
- [72] David A. Cohn. Neural network exploration using optimal experiment design. In *Proc. Adv. Neural Inf. Process. Syst.*, pages 679–686, 1994.
- [73] Claude E. Shannon and Warren Weaver. *The mathematical theory of communication*. University of Illinois press, 1998.
- [74] Wolfram Burgard, Dieter Fox, and Sebastian Thrun. Active mobile robot localization by entropy minimization. In *Proc. Euromicro Workshop Advance. Mob. Rob.*, pages 155–162. IEEE, 1997.

- [75] Henry Carrillo, Philip Dames, Vijay Kumar, and José A. Castellanos. Autonomous robotic exploration using occupancy grid maps and graph SLAM based on Shannon and Rényi entropy. In *Proc. IEEE Int. Conf. Robot. Autom.*, pages 487–494. IEEE, 2015.
- [76] Frederic Bourgault, Alexei A. Makarenko, Stefan B. Williams, Ben Grocholsky, and Hugh F. Durrant-Whyte. Information based adaptive robotic exploration. In *IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, volume 1, pages 540–545. IEEE, 2002.
- [77] Rafael Valencia and Juan Andrade-Cetto. Active pose SLAM. In *Mapping, Planning and Exploration with Pose SLAM*, pages 89–108. Springer, 2018.
- [78] Carlos Guestrin, Andreas Krause, and Ajit Paul Singh. Near-optimal sensor placements in Gaussian processes. In *Proc. 22nd Int. Conf. Mach. Learn.*, pages 265–272. ACM, 2005.
- [79] Costas Papadimitriou, James L. Beck, and Siu-Kui Au. Entropy-based optimal sensor location for structural model updating. *J. Vib. Control*, 6(5):781–800, 2000.
- [80] Christian Daniel, Gerhard Neumann, and Jan Peters. Hierarchical relative entropy policy search. In *Artif. Intell. Stat.*, pages 273–281, 2012.
- [81] Sue Zheng, Jason Pacheco, and John Fisher. A robust approach to sequential information theoretic planning. In *Proc. Int. Conf. Mach. Learn.*, pages 5936–5944, 2018.
- [82] Rajeev Agrawal. Sample mean based index policies by $o(\log n)$ regret for the multi-armed bandit problem. *Adv. Appl. Probab.*, 27(4):1054–1078, 1995.
- [83] Peter Auer. Using confidence bounds for exploitation-exploration trade-offs. *J. Mach. Learn. Research*, 3(Nov):397–422, 2002.
- [84] Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. Practical Bayesian optimization of machine learning algorithms. In *Proc. Adv. Neural Info. Process. Syst.*, pages 2951–2959, 2012.
- [85] Roman Marchant, Fabio Ramos, and Scott Sanner. Sequential Bayesian optimization for spatial-temporal monitoring. In *Proc. 13th Conf. Uncertainty Artif. Intell.*, pages 553–562, 2014.
- [86] Wen Sun, Niteesh Sood, Debadepta Dey, Gireeja Ranade, Siddharth Prakash, and Ashish Kapoor. No-regret replanning under uncertainty. In *Proc. IEEE Int. Conf. Robot. Autom.*, pages 6420–6427, 2017.
- [87] Hanna Kurniawati and Vinay Yadav. An online POMDP solver for uncertainty planning in dynamic environment. In *Robotics Research*, pages 611–629. Berlin, Germany: Springer, 2016.

- [88] Jnaneshwar Das, Frédéric Py, Julio B.J. Harvey, John P Ryan, Alyssa Gelene, Rishi Graham, David A. Caron, Kanna Rajan, and Gaurav S. Sukhatme. Data-driven robotic sampling for marine ecosystem monitoring. *Int. J. Robot. Research*, 34(12):1435–1452, 2015.
- [89] George L. Nemhauser, Laurence A. Wolsey, and Marshall L. Fisher. An analysis of approximations for maximizing submodular set functions—i. *Math. Program.*, 14(1):265–294, 1978.
- [90] Aurélien Garivier and Olivier Cappé. The KL-UCB algorithm for bounded stochastic bandits and beyond. In *Proc. 24th Annual Conf. Learn. Theory*, pages 359–376, 2011.
- [91] Harold J. Kushner. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *J. Basic Eng.*, 86(1):97–106, 1964.
- [92] Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *J. Global Optim.*, 13(4):455–492, 1998.
- [93] José Miguel Hernández-Lobato, Matthew W. Hoffman, and Zoubin Ghahramani. Predictive entropy search for efficient global optimization of black-box functions. In *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, pages 918–926, 2014.
- [94] Philipp Hennig and Christian J. Schuler. Entropy search for information-efficient global optimization. *J. Mach. Learn. Res.*, 13(Jun):1809–1837, 2012.
- [95] Ioannis K. Nikolos, Kimon P. Valavanis, Nikos C. Tsourveloudis, and Anargyros N. Kostaras. Evolutionary algorithm based offline/online path planner for UAV navigation. *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, 33(6):898–912, 2003.
- [96] L.H. Nam, L. Huang, X.J. Li, and J.F. Xu. An approach for coverage path planning for UAVs. In *IEEE 14th Int. Workshop Advance. Motion Control*, pages 411–416. IEEE, 2016.
- [97] Maximo A. Roa, Max J. Argus, Daniel Leidner, Christoph Borst, and Gerd Hirzinger. Power grasp planning for anthropomorphic robot hands. In *IEEE Int. Conf. Robot. Autom.*, pages 563–569. IEEE, 2012.
- [98] Anirudha Majumdar and Russ Tedrake. Robust online motion planning with regions of finite time invariance. In *Proc. Workshop Algorithmic Foundations Robot.*, pages 543–558. Springer, 2012.
- [99] Joel M. Esposito and Vijay Kumar. A method for modifying closed-loop motion plans to satisfy unpredictable dynamic constraints at runtime. In *Proc. IEEE Int. Conf. Robot. Autom.*, volume 2, pages 1691–1696. IEEE, 2002.

- [100] Amarjeet Singh, Andreas Krause, and William J. Kaiser. Nonmyopic adaptive informative path planning for multiple robots. In *Proc. 21st Int. Joint Conf. Artif. Intell.*, pages 1843–1850, 2009.
- [101] Thibaut Horel. Notes on greedy algorithms for submodular maximization, 2016.
- [102] Sertac Karaman, Matthew R. Walter, Alejandro Perez, Emilio Frazzoli, and Seth Teller. Anytime motion planning using the RRT. In *IEEE Int. Conf. Robot. Autom.*, pages 1478–1483. IEEE, 2011.
- [103] Geoffrey A. Hollinger and Gaurav S. Sukhatme. Sampling-based motion planning for robotic information gathering. In *Proc. Robot. Sci. Syst.*, volume 3, pages 5:3931–5:3938. Citeseer, 2013.
- [104] Daniel Levine, Brandon Luders, and Jonathan How. Information-rich path planning with general constraints using rapidly-exploring random trees. In *AIAA Infotech at Aerospace 2010*, page 3360. AIAA, 2010.
- [105] Jonathan Binney and Gaurav S. Sukhatme. Branch and bound for informative path planning. In *Proc. IEEE Int. Conf. Robot. Autom.*, pages 2147–2154, 2012.
- [106] Chun Kai Ling, Kian Hsiang Low, and Patrick Jaillet. Gaussian process planning with Lipschitz continuous reward functions: Towards unifying Bayesian optimization, active learning, and beyond. In *Proc. 30th AAAI conf. Artif. Intell.*, pages 1860–1866, 2016.
- [107] Adhiraj Somani, Nan Ye, David Hsu, and Wee Sun Lee. DESPOT: Online POMDP planning with regularization. In *Proc. Adv. Neural Inf. Process. Syst.*, pages 1772–1780, 2013.
- [108] D. Silver and J. Veness. Monte-Carlo planning in large POMDPs. In *Proc. Adv. Neural Inf. Process. Syst.*, pages 2164–2172, 2010.
- [109] Zachary N. Sunberg, Christopher J. Ho, and Mykel J. Kochenderfer. The value of inferring the internal state of traffic participants for autonomous freeway driving. In *Proc. IEEE Amer. Control Conf.*, pages 3004–3010, 2017.
- [110] Cameron B. Browne, Edward Powley, Daniel Whitehouse, Simon M. Lucas, Peter I. Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. A survey of Monte Carlo tree search methods. *IEEE Trans. Comput. Intell. AI Games*, 4(1):1–43, Mar. 2012.
- [111] Levente Kocsis and Csaba Szepesvári. Bandit based monte-carlo planning. In *Proc. Euro. Conf. Mach. Learn.*, pages 282–293. Springer, 2006.
- [112] Ben Allen, Roger Stokey, Tom Austin, Ned Forrester, Rob Goldsborough, Mike Purcell, and Chris von Alt. REMUS: a small, low cost AUV; system description, field trials and performance results. In *Proc. MTS/IEEE Int. Conf. Oceans*, volume 2, pages 994–1000. IEEE, 1997.

- [113] Junku Yuh. Design and control of autonomous underwater robots: A survey. *Auton. Robot.*, 8(1):7–24, 2000.
- [114] John Bares, Martial Hebert, Takeo Kanade, Eric Krotkov, Tom Mitchell, Reid Simmons, and William Whittaker. Ambler: An autonomous rover for planetary exploration. *Computer*, 22(6):18–26, 1989.
- [115] Arno Ruckelshausen, Peter Biber, Michael Dorna, Holger Gremmes, Ralph Klose, Andreas Linz, Florian Rahe, Rainer Resch, Marius Thiel, Dieter Trautz, et al. Bonirob—an autonomous field robot platform for individual plant phenotyping. *Precis. Agric.*, 9(841):1, 2009.
- [116] John Billingsley, Arto Visala, and Mark Dunn. Robotics in agriculture and forestry. *Springer Handbook of Robotics*, pages 1065–1077, 2008.
- [117] Arko Lucieer, Darren Turner, Diana H. King, and Sharon A. Robinson. Using an unmanned aerial vehicle (UAV) to capture micro-topography of Antarctic moss beds. *Int. J. Appl. Earth Obs. Geoinform.*, 27:53–62, 2014.
- [118] David P. Nicholson, Anna P.M. Michel, Scott D. Wankel, Kevin Manganini, Rebecca A. Sugrue, Zoe O. Sandwith, and Samuel A. Monk. Rapid mapping of dissolved methane and carbon dioxide in coastal ecosystems using the ChemYak autonomous surface vehicle. *Environ. Sci. Technol.*, 52(22):13314–13324, 2018.
- [119] Thomas W. Vaneck, Claudia D. Rodriguez-Ortiz, Mads C. Schmidt, and Justin E. Manley. Automated bathymetry using an autonomous surface craft. *Navigation*, 43(4):407–419, 1996.
- [120] Peter Kimball, John Bailey, Sarah Das, Rocky Geyer, Trevor Harrison, Clay Kunz, Kevin Manganini, Ken Mankoff, Katie Samuelson, Thomas Sayre-McCord, et al. The WHOI JetYak: An autonomous surface vehicle for oceanographic research in shallow or dangerous waters. In *IEEE/OES Auton. Underwater Vehicle Symp.*, pages 1–7. IEEE, 2014.
- [121] Justin Manley and Scott Willcox. The wave glider: A persistent platform for ocean science. In *Proc. MTS/IEEE Int. Conf. Oceans*, pages 1–5. IEEE, 2010.
- [122] Christian Meinig, Noah Lawrence-Slavas, Richard Jenkins, and Heather M. Tabisola. The use of Saildrones to examine spring conditions in the Bering Sea: Vehicle specification and mission performance. In *Proc. MTS/IEEE Int. Conf. Oceans*, pages 1–6. IEEE, 2015.
- [123] Thomas F. Stocker, Dahe Qin, Gian-Kasper Plattner, Melinda Tignor, Simon K. Allen, Judith Boschung, Alexander Nauels, Yu Xia, Vincent Bex, Pauline M. Midgley, et al. Climate change 2013: The physical science basis, 2013.
- [124] *Carbon dioxide*. Encyclopaedia Britannica, Inc., 5th edition, 2018.

- [125] Susan Solomon, Dahe Qin, Martin Manning, Kristen Averyt, and Melinda Marquis. *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC*, volume 4. Cambridge University Press, 2007.
- [126] Y. Kuzyakov. Sources of CO₂ efflux from soil and review of partitioning methods. *Soil Biol. Biochem.*, 38(3):425–448, 2006.
- [127] A.V. Borges, Bruno Delille, and M. Frankignoulle. Budgeting sinks and sources of CO₂ in the coastal ocean: Diversity of ecosystems counts. *Geophys. Res. Lett.*, 32(14), 2005.
- [128] Christopher L. Sabine, Richard A. Feely, Nicolas Gruber, Robert M. Key, Kitack Lee, John L. Bullister, Rik Wanninkhof, C.S.L. Wong, Douglas W.R. Wallace, Bronte Tilbrook, et al. The oceanic sink for anthropogenic CO₂. *Science*, 305(5682):367–371, 2004.
- [129] H. Teng, S.M. Masutani, C.M. Kinoshita, and G.C. Nihous. Solubility of CO₂ in the ocean and its effect on CO₂ dissolution. *Energy Convers. Manag.*, 37(6-8):1029–1038, 1996.
- [130] J.A. Raven and P.G. Falkowski. Oceanic sinks for atmospheric CO₂. *Plant Cell Environ.*, 22(6):741–755, 1999.
- [131] Uta Passow and Craig A. Carlson. The biological pump in a high CO₂ world. *Mar. Ecol. Prog. Ser.*, 470:249–271, 2012.
- [132] Steven Emerson and John Hedges. *Chemical oceanography and the marine carbon cycle*. Cambridge University Press, 2008.
- [133] *Methane*. Encyclopaedia Britannica, Inc., 5th edition, 2018.
- [134] Paul Balcombe, Jamie F. Speirs, Nigel P. Brandon, and Adam D. Hawkes. Methane emissions: choosing the right climate metric and time horizon. *Environ. Sci.: Process. Impacts*, 20(10):1323–1339, 2018.
- [135] Stefanie Kirschke, Philippe Bousquet, Philippe Ciais, Marielle Saunois, Josep G. Canadell, Edward J. Dlugokencky, Peter Bergamaschi, Daniel Bergmann, Donald R. Blake, Lori Bruhwiler, et al. Three decades of global methane sources and sinks. *Nat. Geosci.*, 6(10):813, 2013.
- [136] J. Bogner, Kurt Spokas, Elizabeth Burton, R. Sweeney, and V. Corona. Landfills as atmospheric methane sources and sinks. *Chemosphere*, 31(9):4119–4130, 1995.
- [137] Ralph J. Cicerone and James D. Shetter. Sources of atmospheric methane: measurements in rice paddies and a discussion. *J. Geophys. Res.*, 86(C8):7203–7209, 1981.

- [138] Adam R. Brandt, G.A. Heath, E.A. Kort, Francis O’Sullivan, Gabrielle Pétron, Sarah M. Jordaan, P. Tans, Jennifer Wilcox, A.M. Gopstein, Doug Arent, et al. Methane leaks from North American natural gas systems. *Science*, 343(6172):733–735, 2014.
- [139] Keith A. Kvenvolden. A review of the geochemistry of methane in natural gas hydrate. *Org. Geochem.*, 23(11-12):997–1008, 1995.
- [140] L. Barelli, G. Bidini, F. Gallorini, and S. Servili. Hydrogen production through sorption-enhanced steam methane reforming and membrane technology: a review. *Energy*, 33(4):554–570, 2008.
- [141] M.A.K. Khalil and R.A. Rasmussen. Sources, sinks, and seasonal cycles of atmospheric methane. *J. Geophys. Res.*, 88(C9):5131–5144, 1983.
- [142] D.H. Ehhalt and U. Schmidt. Sources and sinks of atmospheric methane. *Pure Appl. Geophys.*, 116(2-3):452–464, 1978.
- [143] Yuri A. Trotsenko and John Colin Murrell. Metabolic aspects of aerobic obligate methanotrophy. *Adv. Appl. Microbiol.*, 63:183–229, 2008.
- [144] Katharina F. Ettwig, Baoli Zhu, Daan Speth, Jan T. Keltjens, Mike S.M. Jetten, and Boran Kartal. Archaea catalyze iron-dependent anaerobic oxidation of methane. *Proc. Natl. Acad. Sci. U.S.A.*, 113(45):12792–12796, 2016.
- [145] Joshua F. Dean, Jack J. Middelburg, Thomas Röckmann, Rien Aerts, Luke G. Blauw, Matthias Egger, Mike S.M. Jetten, Anniek E.E. de Jong, Ove H. Meisel, Olivia Rasigraf, et al. Methane feedbacks to the global climate system in a warmer world. *Rev. Geophys.*, 56(1):207–250, 2018.
- [146] Katrin Knittel, Tina Lösekann, Antje Boetius, Renate Kort, and Rudolf Amann. Diversity and distribution of methanotrophic archaea at cold seeps. *Appl. Environ. Microbiol.*, 71(1):467–479, 2005.
- [147] Corinne Le Quéré, Robert Joseph Andres, T. Boden, Thomas Conway, Richard A. Houghton, Joanna I. House, Gregg Marland, Glen Philip Peters, Guido Van der Werf, Anders Ahlström, et al. The global carbon budget 1959–2011. *Earth Syst. Sci. Data*, 5(2):1107–1157, 2012.
- [148] Donatella Zona, Beniamino Gioli, Róisín Commane, Jakob Lindaas, Steven C. Wofsy, Charles E. Miller, Steven J. Dinardo, Sigrid Dengel, Colm Sweeney, Anna Karion, et al. Cold season emissions dominate the Arctic tundra methane budget. *Proc. Natl. Acad. Sci. U.S.A.*, 113(1):40–45, 2016.
- [149] Jeffery S Bale, Gregory J Masters, Ian D. Hodkinson, Caroline Awmack, T. Martijn Bezemer, Valerie K. Brown, Jennifer Butterfield, Alan Buse, John C. Coulson, John Farrar, et al. Herbivory in global climate change research: direct effects of rising temperature on insect herbivores. *Glob. Change Biol.*, 8(1):1–16, 2002.

- [150] Curtis A. Deutsch, Joshua J. Tewksbury, Raymond B. Huey, Kimberly S. Sheldon, Cameron K. Ghalambor, David C. Haak, and Paul R. Martin. Impacts of climate warming on terrestrial ectotherms across latitude. *Proc. Natl. Acad. Sci. U.S.A.*, 105(18):6668–6672, 2008.
- [151] Martin Vermeer and Stefan Rahmstorf. Global sea level linked to global temperature. *Proc. Natl. Acad. Sci. U.S.A.*, 106(51):21527–21532, 2009.
- [152] Guillaume Lamarche-Gagnon, Jemma L. Wadham, Barbara Sherwood Lollar, Sandra Arndt, Peer Fietzek, Alexander D. Beaton, Andrew J. Tedstone, Jon Telling, Elizabeth A. Bagshaw, Jon R. Hawkings, et al. Greenland melt drives continuous export of methane from the ice-sheet bed. *Nature*, 565(7737):73, 2019.
- [153] Jan Karlsson, Reiner Giesler, Jenny Persson, and Erik Lundin. High emission of carbon dioxide and methane during ice thaw in high latitude lakes. *Geophys. Res. Lett.*, 40(6):1123–1127, 2013.
- [154] Arne Biastoch, Tina Treude, Lars H. Rüpke, Ulf Riebesell, Christina Roth, Ewa B. Burwicz, Wonsun Park, Mojib Latif, Claus W. Böning, Gurvan Madec, et al. Rising Arctic Ocean temperatures cause gas hydrate destabilization and ocean acidification. *Geophys. Res. Lett.*, 38(8), 2011.
- [155] Peter M. Haugan and Helge Drange. Effects of CO₂ on the ocean environment. *Energy Convers. Manag.*, 37(6-8):1019–1022, 1996.
- [156] Scott C. Doney, Victoria J. Fabry, Richard A. Feely, and Joan A. Kleypas. Ocean acidification: the other CO₂ problem. *Annu. Rev. Mar. Sci.*, 1:169–192, 2009.
- [157] M. Francesca Cotrufo, P. Ineson, and A. Scott. Elevated CO₂ reduces the nitrogen concentration of plant tissues. *Glob. Change Bio.*, 4(1):43–54, 1998.
- [158] J.J. West, A.M. Fiore, L.W. Horowitz, and D.L. Mauzerall. Global health benefits of mitigating ozone pollution with methane emission controls. *Proc. Natl. Acad. Sci. U.S.A.*, 103(11):3988–3993, 2006.
- [159] Ove Hoegh-Guldberg and John F. Bruno. The impact of climate change on the world’s marine ecosystems. *Science*, 328(5985):1523–1528, 2010.
- [160] Denis A. Wiesenburg and Norman L. Guinasso Jr. Equilibrium solubilities of methane, carbon monoxide, and hydrogen in water and sea water. *J. Chem. Eng. Data*, 24(4):356–360, 1979.
- [161] Sachio Yamamoto, James B. Alcauskas, and Thomas E. Crozier. Solubility of methane in distilled water and seawater. *J. Chem. Eng. Data*, 21(1):78–80, 1976.

- [162] R.F. Weiss. Carbon dioxide in water and seawater: the solubility of a non-ideal gas. *Mar. Chem.*, 2(3):203–215, 1974.
- [163] Jonathan J. Cole, Yves T. Prairie, Nina F. Caraco, William H. McDowell, Lars J. Tranvik, Robert G. Striegl, Carlos M. Duarte, Pirkko Kortelainen, John A. Downing, Jack J. Middelburg, et al. Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems*, 10(1):172–185, 2007.
- [164] Emily H. Stanley, Nora J. Casson, Samuel T. Christel, John T. Crawford, Luke C. Loken, and Samantha K. Oliver. The ecology of methane in streams and rivers: patterns, controls, and global significance. *Ecol. Monographs*, 86(2):146–171, 2016.
- [165] Wei-Jun Cai. Estuarine and coastal ocean carbon paradox: CO₂ sinks or sites of terrestrial carbon incineration? *Annu. Rev. Mar. Sci.*, 3:123–145, 2011.
- [166] Paul H. Roberts and Kevin V. Thomas. The occurrence of selected pharmaceuticals in wastewater effluent and surface waters of the lower tyne catchment. *Sci. Total Environ.*, 356(1-3):143–153, 2006.
- [167] Ivan Valiela, Caroline Owens, Elizabeth Elmstrom, and Javier Lloret. Eutrophication of Cape Cod estuaries: Effect of decadal changes in global-driven atmospheric and local-scale wastewater nutrient loads. *Mar. Pollut. Bull.*, 110(1):309–315, 2016.
- [168] Ina Geedicke, Jens Oldeland, and Michelle R. Leishman. Urban stormwater run-off promotes compression of saltmarshes by freshwater plants and mangrove forests. *Sci. Total Environ.*, 637:137–144, 2018.
- [169] Karine Olu, Erik E. Cordes, Charles R. Fisher, James M. Brooks, Myriam Sibuet, and Daniel Desbruyères. Biogeography and potential exchanges among the Atlantic equatorial belt cold-seep faunas. *PloS One*, 5(8):e11967, 2010.
- [170] Lisa A. Levin. Ecology of cold seep sediments: interactions of fauna with flow, chemistry and microbes. In *Oceanogr. Mar. Biol.*, pages 11–56. CRC Press, 2005.
- [171] William S. Reeburgh. Oceanic methane biogeochemistry. *Chem. Rev.*, 107(2):486–513, 2007.
- [172] D.F. McGinnis, J. Greinert, Y. Artemov, S.E. Beaubien, and A. Wüest. Fate of rising methane bubbles in stratified waters: How much methane reaches the atmosphere? *J. Geophys. Res.*, 111(C9), 2006.
- [173] Nicholas J. Grant and Michael J. Whiticar. Stable carbon isotopic evidence for methane oxidation in plumes above Hydrate Ridge, Cascadia Oregon Margin. *Global Biogeochem. Cycles*, 16(4):71–1, 2002.

- [174] E. Suess, M.E. Torres, G. Bohrmann, R.W. Collier, D. Rickert, C. Goldfinger, P. Linke, A. Heuser, H. Sahling, K. Heeschen, et al. Sea floor methane hydrates at Hydrate Ridge, Cascadia Margin. *Geophys. Monograph*, 124:87–98, 2001.
- [175] M.E. Torres, K. Wallmann, A.M. Tréhu, G. Bohrmann, W.S. Borowski, and H. Tomaru. Gas hydrate growth, methane transport, and chloride enrichment at the southern summit of Hydrate Ridge, Cascadia margin off Oregon. *Earth Planet. Sci. Lett.*, 226(1-2):225–241, 2004.
- [176] Robert W. Embley, Susan G. Merle, Tamara Baumberger, Nicole Raineault, and Marta E. Torres. Extensive authigenic carbonate hardgrounds on the Cascadia continental margin. In *AGU Fall Meeting Abstracts*, 2018.
- [177] H. Paul Johnson, Una K. Miller, Marie S. Salmi, and Evan A. Solomon. Analysis of bubble plume distributions to evaluate methane hydrate decomposition on the continental slope. *Geochem., Geophys., Geosyst.*, 16(11):3825–3839, 2015.
- [178] Katja U. Heeschen, Anne M. Tréhu, Robert W. Collier, Erwin Suess, and Gregor Rehder. Distribution and height of methane bubble plumes on the Cascadia Margin characterized by acoustic imaging. *Geophys. Res. Lett.*, 30(12), 2003.
- [179] Katja U. Heeschen, Robert W. Collier, Marie A. de Angelis, Erwin Suess, Gregor Rehder, Peter Linke, and Gary P. Klinkhammer. Methane sources, distributions, and fluxes from cold vent sites at Hydrate Ridge, Cascadia Margin. *Global Biogeochem. Cycles*, 19(2), 2005.
- [180] Gregor Rehder, Peter W. Brewer, Edward T. Peltzer, and Gernot Friederich. Enhanced lifetime of methane bubble streams within the deep ocean. *Geophys. Res. Lett.*, 29(15):21–1, 2002.
- [181] Philippe Ciais, Christopher Sabine, Govindasamy Bala, Laurent Bopp, Victor Brovkin, Josep Canadell, Abha Chhabra, Ruth DeFries, James Galloway, Martin Heimann, et al. Carbon and other biogeochemical cycles. In *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 465–570. Cambridge University Press, 2014.
- [182] Carolina Voigt, Maija E. Marushchak, Richard E. Lamprecht, Marcin Jackowicz-Korczyński, Amelie Lindgren, Mikhail Mastepanov, Lars Granlund, Torben R. Christensen, Teemu Tahvanainen, Pertti J. Martikainen, et al. Increased nitrous oxide emissions from Arctic peatlands after permafrost thaw. *Proc. Natl. Acad. Sci. U.S.A.*, 114(24):6238–6243, 2017.
- [183] Keith A. Kvenvolden, Marvin D. Lilley, Thomas D. Lorenson, Peter W. Barnes, and Elizabeth McLaughlin. The Beaufort Sea continental shelf as a seasonal source of atmospheric methane. *Geophys. Res. Lett.*, 20(22):2459–2462, 1993.

- [184] Blaize A. Denfeld, Helen M. Baulch, Paul A. del Giorgio, Stephanie E. Hampton, and Jan Karlsson. A synthesis of carbon dioxide and methane dynamics during the ice-covered period of northern lakes. *Limnology and Oceanography Letters*, 3(3):117–131, 2018.
- [185] Allan R. Phelps, Kim M. Peterson, and Martin O. Jeffries. Methane efflux from high-latitude lakes during spring ice melt. *J. of Geophys. Res.*, 103(D22):29029–29036, 1998.
- [186] T. Bliss. *Tundra ecosystems: a comparative analysis*, chapter 1.2 North American and Scandinavian tundras and polar deserts. CUP Archive, 1981.
- [187] Aiguo Dai and Kevin E. Trenberth. Estimates of freshwater discharge from continents: Latitudinal and seasonal variations. *J. Hydrometeorol.*, 3(6):660–687, 2002.
- [188] R.W. Macdonald, Z.A. Kuzyk, and S.C. Johannessen. It is not just about the ice: a geochemical perspective on the changing Arctic Ocean. *J. Environmental Studies Sci.*, 5(3):288–301, 2015.
- [189] Victoria Preston, Genevieve Flaspohler, Anna Pauline Miranda Michel, Yogesh Girdhar, and Nicholas Roy. Information-guided robotic maximum seek-and-sample in partially observable continuous environments. *Robot. Autom. Lett.*, 2019.
- [190] Syed Talha Jawaid and Stephen L Smith. Informative path planning as a maximum traveling salesman problem with submodular rewards. *Discr. Appl. Math.*, 186:112–127, 2015.
- [191] Patrick Dallaire, Camille Besse, Stephane Ross, and Brahim Chaib-draa. Bayesian reinforcement learning in continuous POMDPs with Gaussian processes. In *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, pages 2604–2609, 2009.
- [192] Robert Platt, Russell Tedrake, Leslie Kaelbling, and Tomas Lozano-Perez. Belief space planning assuming maximum likelihood observations. In *Proc. Robot. Sci. Syst. Conf.*, 2010.
- [193] Gregory Hitz, Enric Galceran, Marie-Ève Garneau, François Pomerleau, and Roland Siegwart. Adaptive continuous-space informative path planning for on-line environmental monitoring. *J. Field Robot.*, 34(8):1427–1449, 2017.
- [194] Amarjeet Singh, Fabio Ramos, Hugh Durrant Whyte, and William J. Kaiser. Modeling and decision making in spatio-temporal processes for environmental surveillance. In *Proc. IEEE Int. Conf. Robot. Autom.*, pages 5490–5497, 2010.
- [195] Zachary N. Sunberg and Mykel J. Kochenderfer. Online algorithms for POMDPs with continuous state, action, and observation spaces. In *Proc. 24th Int. Conf. Automated Plan. Schedul.*, 2018.

- [196] Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Proc. Adv. Neural Inf. Process. Syst.*, pages 1177–1184, 2008.
- [197] Salomon Bochner. *Lectures on Fourier integrals*. Princeton University Press, 1959.
- [198] Henry B. Mann and Donald R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.*, 18:50–60, 1947.
- [199] Sandeep Manjanna, Nikhil Kakodkar, Malika Meghjani, and Gregory Dudek. Efficient terrain driven coral coverage using Gaussian processes for mosaic synthesis. In *Proc. 13th IEEE Conf. Comput. Robot Vis.*, pages 448–455, 2016.
- [200] Anastazia T. Banaszak and Michael P. Lesser. Effects of solar ultraviolet radiation on coral reef organisms. *Photochem. Photobiol. Sci.*, 8(9):1276–1294, 2009.
- [201] Ananth Ranganathan, Ming-Hsuan Yang, and Jeffrey Ho. Online sparse Gaussian process regression and its applications. *IEEE Trans. Image Process.*, 20(2):391–404, Feb. 2011.
- [202] Philippe Morere, Roman Marchant, and Fabio Ramos. Continuous state-action-observation POMDPs for trajectory planning with Bayesian optimisation. In *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, pages 8779–8786, 2018.
- [203] Kai-Chieh Ma, Lantao Liu, Hordur K. Heidarsson, and Gaurav S. Sukhatme. Data-driven learning and planning for environmental sampling. *J. Field Robot.*, 35(5):643–661, 2018.
- [204] Ransalu Senanayake, Anthony Tompkins, and Fabio Ramos. Automorphing kernels for nonstationarity in mapping unstructured environments. In *Proc. Conf. Robot. Learn.*, pages 443–455, 2018.
- [205] M. Kanakidou, J.H. Seinfeld, S.N. Pandis, I. Barnes, F.J. Dentener, M.C. Facchini, R. Van Dingenen, B. Ervens, A. Nenes, C.J. Nielsen, et al. Organic aerosol and global climate modelling: a review. *Atmospheric Chem. Phys.*, 5(4):1053–1123, 2005.
- [206] Laura Uusitalo. Advantages and challenges of Bayesian networks in environmental modelling. *Ecol. Model.*, 203(3-4):312–318, 2007.
- [207] Nitin Muttill and Kwok-wing Chau. Neural network and genetic programming for modelling coastal algal blooms. *Int. J. Environ. Pollut.*, 2006.
- [208] Yogesh Girdhar, Philippe Giguere, and Gregory Dudek. Autonomous adaptive exploration using realtime online spatiotemporal topic modeling. *Int. J. Robot. Res.*, 33(4):645–657, 2014.

- [209] Itziar Landa-Torres, Diana Manjarres, Sonia Bilbao, and Javier Del Ser. Underwater robot task planning using multi-objective meta-heuristics. *Sensors*, 17(4):762, 2017.
- [210] Daniele Calisi, Alessandro Farinelli, Luca Iocchi, and Daniele Nardi. Multi-objective exploration and search for autonomous rescue robots. *J. Field Robot.*, 24(8-9):763–777, 2007.
- [211] Ellips Masehian and Davoud Sedighizadeh. A multi-objective PSO-based algorithm for robot path planning. In *IEEE Int. Conf. Ind. Technol.*, pages 465–470. IEEE, 2010.