

Perceive, Predict, and Plan: Robotic Expeditionary Science in Oceanic Spatiotemporal Fields

by

Victoria Lynn Preston

B.S. 2016, OLIN COLLEGE OF ENGINEERING

S.M. 2019, MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Submitted to the Department of Aeronautics and Astronautics and the Joint Program
in Oceanography and Applied Ocean Science & Engineering in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in Autonomous Systems

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

FEBRUARY 2023

©2023 Massachusetts Institute of Technology and Woods Hole Oceanographic Institution.

All rights reserved.

Authored by: Victoria Lynn Preston
Department of Aeronautics and Astronautics, MIT
Applied Ocean Science & Engineering, WHOI
December 15th, 2022

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

Certified by: Anna P. M. Michel
Associate Scientist with Tenure, Applied Ocean Physics & Engineering, WHOI
Thesis Supervisor

Accepted by: Jonathan P. How
R. C. MacLaurin Professor of Aeronautics and Astronautics, MIT
Chair, Graduate Program Committee

Accepted by: David Ralston
Associate Scientist with Tenure, Applied Ocean Physics & Engineering, WHOI
Chair, Joint Committee for Applied Ocean Science & Engineering

Perceive, Predict, and Plan: Robotic Expeditionary Science in Oceanic Spatiotemporal Fields

by

Victoria Lynn Preston

Submitted to the Department of Aeronautics and Astronautics and the Joint Program in Applied Ocean Science & Engineering on December 15th, 2022 in partial fulfillment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

An improved understanding of our ocean would allow us to characterize the largest habitable biosphere on planet Earth, quantify the geochemical processes that control Earth's climate, and develop responsible regulations for controlling the natural resources stored in its depths. *Expeditionary science* is the art of collecting *in situ* observations of an environment to build approximate models of underlying properties that move us towards this understanding. Robotic platforms are a critical technology for collecting observations of the ocean. Depth-capable autonomous underwater vehicles (AUVs) are commonly used to build static maps of the seafloor by executing pre-programmed surveys. However, there is growing urgency to generate rich data products of *spatiotemporal distributions* that characterize the physics and chemistry of the deep ocean biogeoosphere. In this thesis, the problem of charting dynamic deep sea hydrothermal plumes with depth-capable AUVs is investigated. Effectively collecting samples of geochemical plumes using the operationally preferred strategy of pre-specifying surveys requires access to a dynamics model of the advective currents, bathymetric updrafts, and turbulent mixing at a hydrothermal site. In practice, however, access to this information is unavailable, imperfect, or only partially known, and so a model of plume dynamics must be inferred from observations and subsequently leveraged to improve future sampling performance. As most *in situ* scientific instruments yield point-measurements, considerable uncertainty is placed over the form of the dynamics in purely data-driven solutions. Challenges related to *planning under uncertainty* for geochemical surveys in the deep ocean are addressed in this thesis by embedding scientific knowledge as a strong inductive prior for tractable model learning and decision-making. Algorithmic contributions of this thesis show how plumes can be *perceived* from field data, their fate *predicted* far into the future (e.g., multiple days), and informative fixed trajectories *planned* which place an AUV in the right place at the right time. Scientific assessment of observational data collected with AUV *Sentry* during field trials in the Guaymas Basin, Gulf of California are interwoven with algorithmic analyses, demonstrating how intelligent perception, prediction, and planning enables novel insights about hydrothermal plumes.

Thesis Supervisor: Nicholas Roy

Title: Bisplinghoff Professor of Aeronautics and Astronautics, MIT

Thesis Supervisor: Anna P. M. Michel

Title: Associate Scientist with Tenure, Applied Ocean Physics and Engineering, WHOI

ACKNOWLEDGMENTS

Research is a team sport, and I have profound appreciation for my teammates, coaches, and enthusiastic supporters that have made this work possible.

Since I started graduate school, I have had the great fortune of collaborating closely with Genevieve Flaspoehler, with whom I've talked for hours about algorithms, life, and expeditionary science. I'm so grateful for her significant contributions to this work, as well as her friendship. Claudia Cenedese, Dan Fornari, Pete Girguis, Scott Wankel, Chris German, Michael Jakuba, Guangyu Xu, John Fisher, Thibaut Barreyre, and Xubo Zhang provided invaluable insights about hydrothermalism, the challenges of studying plumes in the deep ocean, and planning under uncertainty; from providing technical mentorship to datasets to sensors to field work opportunities, I have been truly humbled by their expertise, willingness to hear out my quirky ideas, and enthusiasm for robotic tools. Dan Yang and Valentin Peretroukhin participated in developing some of the first prototypes of the algorithmic contributions in this thesis, and their insights and assistance the months before going to sea was critical for the success of the field campaign. Warmest thanks to the AUV *Sentry* team, and especially Sean Kelley, Zac Berkowitz, Justin Fujii, Amanda Sutherland, Joe Garcia, Stefano Suman, and Isaac Vandor, all of whom provided great technical and practical insight into AUV operations, and allowed us to try something a little different on RR2107. I fondly appreciate the camaraderie of the science party and ROV *JASON* team on RR2107, who made my watch shift in the wee hours of the morning fun, explained their projects to me with great patience, and always made sure to remind me when meal times were. The captain and crew on the R/V *Revelle*, who shared their home with me for a few weeks in November, cannot be commended enough for their professionalism and friendliness, curiosity about the science, and willingness to pull off that one final transect experiment when sailing home was just within reach.

Anna and Nick have been key players in all of my doctoral work. It has been a profound privilege to work with two professionals in different fields and getting their unique perspectives on my research; their shared enthusiasm for the work is plain in the ability for this thesis to span multiple disciplines. Anna has opened so many doors for me in the ocean sciences, not least of which have been opportunities to go to sea, and I've been undoubtedly hooked by ocean research and challenges as a result. I have deeply appreciated our conversations about inclusion in the ocean sciences, her attention to my professional development, and her insights on the next frontier of ocean technology and robotics. When my studies at MIT began, Nick was willing to give me some desk space while I was on campus; I am so grateful that he took some interest in my research and has since been an incredible influence in how I think about problem solving. Nick's ability to give high quality feedback, and willingness to receive feedback

in return, has made me feel both professionally and personally valued, which helped me stay engaged when the going got tough. Both Anna and Nick have built incredible labs around them, and I am so grateful to my colleagues in the Robust Robotics Group and the Chemical Sensor Lab, folks who I also count among my friends. I especially want to express appreciation for Beckett Colson, Chris Bradley, Michael Noseworthy, and Martina Stadler for their friendship and feedback on my work throughout our studies together. As part of the research experience in Nick and Anna's groups, I've had the good luck of advising incredible undergraduate students who helped me develop as a mentor and researcher. I am so appreciative of their patience and for giving me the opportunity to work on some really cool projects that otherwise would have never gotten off the ground.

My committee members, Youssef Marzouk and Adam Soule, have been thoughtful contributors to this thesis, and have devoted considerable time towards thinking about my research. I am deeply thankful for their insights and expertise. Thanks also go to Kwesi Rutledge and Valentin Peretrouhkin who are external readers for this work, and are folks whose work and way of thinking about research and research communities I greatly admire.

Finally, a special thanks to my family and to my partner, Bill. My parents shaped me into the independent person I am today, and have modeled hard work, determination, and a get-it-done attitude all my life. Along with my brother, I am so grateful for their enthusiasm and support. While Bill's suggestion for an acronym of the binary sensor filter presented in this thesis ultimately didn't make the cut (it was PHINDS: Procedural Heuristic Integration of Nautical Disparate Sensors, for the curious), I think it's indicative of how much of his own time and energy he's invested in supporting me and becoming familiar with this work. Bill has been a wonderful life partner and research sounding board throughout my graduate school experience, and I'm sincerely grateful for him.

Thank you all,
Victoria

Financial support for my research was provided by the National Defense Graduate Fellowship Program and the MIT Martin Family Society of Fellows for Sustainability. Research activities for the RR2107 cruise were funded by NSF OCE OTIC #1842053, a WHOI Innovation Technology Award, NOAA Ocean Exploration #NA18OAR0110354, and Schmidt Marine Technology Partners Award #G-21-62431.

CONTENTS

1	INTRODUCTION	16
1.1	Hydrothermalism in the Deep Ocean	20
1.2	Challenges for Intelligent Autonomy	23
1.3	Thesis Contributions	28
1.3.1	Perceive	28
1.3.2	Predict	30
1.3.3	Plan	32
1.4	Thesis Overview	33
2	OPERATIONS AT SEA AND FIELD WORK IN GUAYMAS BASIN	34
2.1	Guaymas Basin and Expedition RR2107	34
2.2	Challenges for Robots and Autonomy in the Deep Ocean	37
2.3	The Science Party and Responsibilities	40
2.4	AUV <i>Sentry</i>	42
2.5	ROV <i>JASON</i>	44
2.6	CTD Rosette	48
3	FOUNDATIONAL RELATED WORK	51
3.1	Representing Dynamic Systems	51
3.2	Inverse Problems in Environmental Science	59
3.2.1	Bayesian Inference Techniques	61

3.2.2	Bayesian Representations	64
3.2.3	Scientific Machine Learning	67
3.3	Environmental Sensing as Adaptive Sampling	68
3.3.1	Partially Observable Markov Decision Processes	70
3.3.2	Decision-making under Uncertainty	71
3.3.3	Information-Theoretic Rewards	72
3.4	Vent Prospecting, Odor Localization, and Front Tracking	75
3.5	Hydrothermal Plumes	77
3.5.1	Tracers and Instrumentation	80
4	DISCOVERING HYDROTHERMALISM FROM AFAR	82
4.1	Introduction	83
4.2	Materials and Methods	86
4.2.1	Site Description	86
4.2.2	Sampling Platforms and Instruments	88
4.2.3	Analytical Procedure	93
4.2.4	Transect Design and Execution	93
4.3	Results	95
4.3.1	Methane Measured by Spectroscopic Instruments	95
4.3.2	Methane and Ammonium Measured with the Rosette	96
4.3.3	Turbidity	98
4.3.4	Oxidation Reduction Potential	98
4.3.5	Temperature, Salinity, and Oxygen	100
4.4	Discussion	102
4.4.1	Sensor Cross-Correlations	102
4.4.2	Hydrothermalism Detection via Time-Series Regimes	106
4.4.3	Methane in Deep Sea Exploration	108

4.4.4	Enabling Better Decision-Making for Hydrothermalism Discovery	110
5	PHYSICALLY-INFORMED OPERATIONAL ROBOTIC TRAJECTORIES FOR SCIENTIFIC EXPEDITIONS	112
5.1	Introduction	113
5.1.1	Contributions	117
5.1.2	Advancing Hydrothermal Studies with Robots	118
5.1.3	Closing the Loop: Expedition Logistics	121
5.2	Problem Formulation	123
5.2.1	Scientific Expeditions as a Sequential Decision-Making Problem	123
5.2.2	Scientific Decision-Making with AUV <i>Sentry</i>	125
5.2.3	Charting Hydrothermalism as a POMDP	127
5.3	Methodology	130
5.3.1	Sensor Model for Science Observations	130
5.3.2	PHUMES: Physically-informed Probabilistic Forecasts	131
5.3.3	Trajectory Optimization with Fixed Patterns	137
5.4	Simulation Experiments	140
5.4.1	Evaluation Metrics	143
5.5	PHORTEX Performance	144
5.6	PHUMES Model Validation	144
5.7	Discussion and Future Work	149
5.7.1	Temporal Resolution in PHUMES Forecasts	151
5.8	Conclusion	154
6	CHARTING DEEP-SEA HYDROTHERMAL PLUMES IN THE FIELD	156
6.1	Introduction	157
6.2	Related Work	160
6.2.1	Treating Plume Observations	160

6.2.2	Autonomous Robots Studying Hydrothermalism	162
6.2.3	Using Observations in Scientific Discovery	163
6.3	Methods	164
6.3.1	Treatment of AUV <i>Sentry</i> Science Sensors	164
6.3.2	Opportunistic Sensing Equipment	168
6.3.3	At Sea Operations	171
6.4	Description of the Autonomy Field Work	174
6.4.1	Dives	175
6.4.2	Evaluation of Field Data	176
6.5	Performance of PHORTEX	177
6.6	PHUMES Validation with Basin Observations	179
6.7	PHUMES as a Science Model	181
6.8	Discussion and Future Work	185
6.9	Conclusions	188
7	FUTURE WORK IN EXPEDITIONARY ROBOTS	190
7.1	Technical Challenges and Opportunities in Expeditionary Robotics . .	191
7.1.1	Belief Representations	191
7.1.2	Decision-Making	194
7.2	Physically-Informed Kernel Learning	196
7.3	Measure-Invariant Subspace Learning	201
7.4	Environment Predictability and Planning	205
7.5	Mobile Geochemical Observatories for Atmospheric Emissions	208
7.6	Modeling Anthropogenic Methane in Estuaries	211
7.7	Looking Ahead	213
8	FINAL THOUGHTS	215

A DISCOVERING HYDROTHERMALISM FROM AFAR	219
A.1 Method for Methane Measurement from Niskin Bottles	219
A.2 Leg 2 Niskin Bottle Sample Schedule and Measurements	221
A.3 Normalized Pythia Calibration	223
A.4 Depth-Correction	224
A.5 Description of Plume Model for Transect Design	225
A.6 Code and Data Availability	228
B PHORTEX PERFORMANCE	229
B.1 Convergence of Trajectory Optimizer	229
B.2 MCMC Chain Characteristics of PHUMES	231
B.3 Code Availability	232
C GUAYMAS BASIN FIELD RESULTS	234
C.1 Instrument Data	234
C.2 Acoustically Broadcast Science Data	239
C.3 Code and Data Availability	239
C.4 Autonomy Field Report	240
ACRONYMS	275

LIST OF FIGURES

1.1	Images of hydrothermal vents and illustrative plume structures.	22
1.2	Comparison of informed and uninformed survey of spatiotemporal distribution.	25
2.1	Guaymas Basin, Gulf of California.	35
2.2	AUV <i>Sentry</i>	44
2.3	ROV <i>JASON</i>	45
2.4	Exit velocity estimation with WHOI-MISO cameras.	48
2.5	CTD Rosette.	49
4.1	Overview of transect design for hydrothermal discovery.	87
4.2	Map of transect experiment extent.	88
4.3	Images of SAGE and Pythia methane instruments.	91
4.4	Methane observations collected over the transect.	96
4.5	Methane observations compared to ammonium concentrations.	97
4.6	Turbidity measurements collected during transect.	99
4.7	Oxidation-reduction potential measurements collected during transect.	99
4.8	Depth-corrected oxygen, temperature, and salinity measurements collected during transect.	101
4.9	Global Pearson correlation coefficients between sensors mounted on the rosette and <i>Sentry</i>	103

4.10 Local (rolling) Pearson correlation coefficients for rosette mounted instruments.	104
4.11 Local (rolling) Pearson correlation coefficients for AUV <i>Sentry</i> mounted instruments.	105
4.12 Regime changes in rosette observations.	107
4.13 Regime changes in AUV <i>Sentry</i> observations.	108
5.1 An overview of PHORTEX: P Hysically-informed O perational R obotic T rajectories for E Xpeditions.	116
5.2 PHUMES: P Hysically-informed U ncertainty M odels for E nvironment S patiotemporality	134
5.3 Trajectory optimization in PHORTEX.	138
5.4 Simulated field trial environment.	142
5.5 Naive and PHORTEX-design trajectories.	145
5.6 Evaluation of the PHORTEX trajectories.	146
5.7 Illustration of model learning.	148
5.8 Intersection over Area and Intersection over Union of trained models. .	150
6.1 Example time series (left) and associated detections (right) over the AUV <i>Sentry</i> sensor suite.	167
6.2 Auxiliary data products used in PHUMES.	172
6.3 The operational implementation of PHORTEX at sea.	173
6.4 Study site in the Guaymas Basin, Gulf of California	174
6.5 The four field dives of AUV <i>Sentry</i>	178
6.6 Validation of PHUMES model trained at sea.	180
6.7 PHUMES investigation of neutrally-buoyant plume height and energetic flux.	182

A.1	Fitted calibration curve for measurements of methane observed by Pythia.	223
A.2	Pythia calibrated field data	224
A.3	Functions used for depth normalization	224
A.4	Plume model for transect design	228
B.1	PHORTEX optimization performance	230
B.2	PHUMES simulation chain	232
B.3	PHUMES long chain	233
C.1	AUV <i>Sentry</i> dive 607 data.	235
C.2	AUV <i>Sentry</i> dive 608 data.	236
C.3	AUV <i>Sentry</i> dive 610 data.	237
C.4	AUV <i>Sentry</i> dive 611 data.	238
C.5	Acoustic versus logged OBS messages during AUV <i>Sentry</i> dive 607 . .	239

LIST OF TABLES

2.1	Summary of tiltmeter deployments.	47
6.1	Instruments on AUV <i>Sentry</i> and the criteria used to identify plume fluids for each instrument.	168
6.2	Summary of auxiliary data for PHUMES modeling at sea.	169
6.3	Per-dive statistics for field trials of PHORTEX.	177
A.1	Results of DGEU extraction efficiency calibration experiments.	221
A.2	Schedule of bottle samples during Leg 2 of rosette transect.	222
A.3	Geochemical measurements collected during leg 2 of rosette transect .	222
A.4	Parameter, vent characteristics, and ambient crossflow setting used for transect design.	227

1 INTRODUCTION

It is a wholesome and necessary thing for us to turn again to the earth and in the contemplation of her beauties to know the sense of wonder and humility.

Rachel Carson

The environmental sciences are the multidisciplinary, academic studies which aim to understand the Earth and its processes. *In situ* observational studies, or *expeditions*, serve as the foundation on which scientific discovery and model development are predicated in these fields. With improvements in technology, expeditions have been conducted from the deepest trenches of the ocean to the uppermost atmosphere. Improved reach, in addition to improved observational quality, density, and availability has made it increasingly clear how inextricably entwined Earth’s regulatory processes are, and how crucial a role the ocean plays in these processes. Covering 70% of the Earth’s surface and encompassing 90% of the habitable volume on Earth, the ocean is the largest biosphere on the planet and home to both the largest creatures in Earth and a staggering number of unique microorganisms (Cario *et al.* 2019; Purkis & Chirayath 2022). The plant life supported by the ocean and ocean-coast interfaces are estimated to absorb 50% of all excess carbon dioxide emissions produced by anthropogenic sources, acting as a buffer to global heating (Hori *et al.* 2019). Culturally, the ocean is also entwined with our sense of humanity and development of society—*island and coastal habitats supported the earliest hominids* (Erlandson & Fitzpatrick 2006), traveling the

ocean has shaped trade, conquest, and tradition (Chaudhuri & Chaudhuri 1985; Firth & Naidu 2019; Nunn 2003; Pearson 2003), and the ocean inspires creativity, recreation, and curiosity. Despite the centrality of the ocean to existence (and continued existence) as we know it on Earth, there is so much that has yet to be discovered.

In the last decade, significant effort has been put to finely mapping the seafloor (e.g., through Seabed 2030 (Mayer *et al.* 2018), among other initiatives), and as of mid-2022 nearly a quarter of the seafloor has been mapped bathymetrically (i.e., the depth has been charted) in high resolution (compared to about 6% in 2017)¹. High resolution maps allow us to better understand continental drift and crustal processes, the distribution of natural resources, and the ecosystems which ocean structures can support. Unlike terrestrial environments, which can be largely observed remotely (either by aircraft, or more commonly now, satellite), the deep ocean, cannot be remotely observed due to the conductive properties of water and its tendency to absorb many forms of light and radio energy. Mapping the seafloor requires physically going to sea. The mapping revolution of the ocean is enabled thanks in part² to improved acoustic technology and processing tools, which allow shipboard acoustic sounders to collect high resolution “imagery” of the seafloor while traversing on the surface ocean.

Bathymetry is one piece of the giant puzzle that is understanding the ocean; another piece seeing contemporary scrutiny is geochemistry of the deep ocean. Geochemistry, the study of Earth and other planetary geological systems through chemical principles, enables us to understand the processes which create and sustain the structures that bathymetric maps reveal, and tell us more about local ecosystems and their nutrient and energetic budgets. Geochemical measurements may range from studying the composition of rock samples from the seafloor, to *in situ* observations of dissolved gases in a hydrothermal plume. Modern interests in seafloor mining to remove materials from the

¹As reported by Seabed 2030 at <https://seabed2030.org/mapping-progress>.

²In other part, it has been enabled by an increased economic and political commitment to the ocean, its strategic value, and the resources contained within it.

deep ocean (Thompson *et al.* 2018), and deep ocean carbon sequestration to inject the ocean with excess materials (Teng & Zhang 2018), stand to directly impact the balance of the deep biogeosphere, with implications that have yet to be well understood (or agreed upon) by science (Childs 2020; Fleeger *et al.* 2010; Seibel & Walsh 2001; Sharma 2015; Smith *et al.* 2020; Van Dover 2011). Unlike acoustic surveys, which can be performed from a ship hundreds or thousands of meters above the seabed, geochemical surveys must truly be conducted *in situ* with instruments physically sent to the deep ocean to collect continuous measurements, or physical bottle samples of water or rock specimens retrieved for *ex situ* analysis. There is an ongoing paradigm shift in ocean technology to better enable geochemical studies of the deep ocean: development of novel *in situ* sensors, creation of more depth-capable instruments, and adoption of autonomous technologies for sample collection or expedition planning.

Intelligent autonomous technologies are any systems which can automatically process and analyze a data product to formulate evidence-based decisions which they may then act upon. These systems may be *embodied*, like robotic vehicles, or simply algorithmic. A science party member is an intelligent autonomous agent for an expedition: the scientist will be collecting and processing data throughout the voyage, using that data to coordinate with other science party members to design sampling activities, and managing the deployment of various sensors and platforms. Good algorithmic or robotic contributions relieve the burden of data processing and decision-making on the science party to make their time, the expedition time, and the resources aboard a vessel more effective. For marine geochemical surveys, one of the key challenges for the science party to grapple with is understanding *spatiotemporal distributions*, which are nearly ubiquitous in water column studies. A spatiotemporal distribution is a phenomenon that evolves in space and time; in the deep sea, these may take the form of hydrothermal plumes, hydrate dissolution, sediment transport, or water mass mixing.

Perceiving a spatiotemporal distribution from sparse *in situ* measurements, *predicting* the evolution of the distribution into future expedition dates or sites, and *planning* where and when to take samples of the phenomenon for further inspection, are key challenges for any autonomous agent supporting expeditionary science. This thesis approaches these challenges through an algorithmic lens of *planning under uncertainty*, using scientific knowledge as a foundation on which to shape autonomous technologies that specifically enable deep sea geochemical surveys, with a special focus on charting hydrothermal plumes. Planning under uncertainty is a formal technique for decision-making with incomplete information about a target environment or task. A roboticist implementing a planning under uncertainty architecture has many design choices to make—how data is processed, how the data is used to construct a *belief* about the environment or task, and how that belief is leveraged to pick actions to take. Field and scientific contexts provide constraints and requirements that shape the feasible set of design choices that can be made, and can also inspire algorithmic creativity that may be beyond the pale in theoretical research. In line with this notion is a core tenet of this thesis: field roboticists have a special responsibility to ensure that the solutions that they engineer accomplishes the science that they were designed to support. To be a field roboticist requires a detailed understanding of a particular context and set of scientific queries; it requires assessing the science as well as the engineering. To this end, scientific analyses which enhance our understanding of deep ocean hydrothermalism appear alongside technical evaluations of algorithmic contributions throughout this thesis.

In the rest of this chapter, an overview of deep hydrothermalism and geochemical field operations is provided, in addition to a summary of specific algorithmic challenges and the techniques employed in this thesis to overcome them. A brief contribution statement concludes the chapter.

1.1 HYDROTHERMALISM IN THE DEEP OCEAN

The deep ocean is considered to be any part of the ocean at least 200 m below sea-level³. Volumetrically, the deep ocean is *most* of the ocean, and life has been discovered at every depth, including the Challenger Deep, the deepest point on Earth at nearly 11.000 m (Carlo *et al.* 2019)⁴. A persistent adage in the ocean sciences is that more is known about the surface of other worlds than about our own deep ocean. This is poignantly illustrated by the first observation of hydrothermal vents in 1977 at the Galápagos Rift (Corliss *et al.* 1979), eight years after humans walked on the moon for the first time.

Since 1977, hundreds of vents have been discovered and studied around the world (Beaulieu *et al.* 2013), and been the subject of increasingly urgent conversations about characterizing the deep ocean. Seafloor venting sites, energized by magmatic sources, release fluids 20-400 °C (background deep ocean temperatures are approximately 2 °C), and imbued with minerals, metals, dissolved gases, and other compounds (Jannasch & Mottl 1985; Martin *et al.* 2008). These warm, nutrient-pumping sites in the deep ocean have created oases for chemosynthetic micro- and macro-fauna (Corliss *et al.* 1979), and the venting fluids, called plumes, can deposit minerals and metals over kilometer (basin) scales (Le Bris *et al.* 2019; Resing *et al.* 2015; Scholz *et al.* 2019). Detection and characterization of seafloor hydrothermal venting are critical for understanding fundamental interactions between the deep ocean, its underlying basaltic crust, the deep biosphere, and (bio)geochemical fluxes.

To study hydrothermalism in the deep ocean, rosettes⁵, remotely-operated vehicles (ROVs), human occupied vehicles (HOVs), and autonomous underwater vehicles

³Some literature more specifically claims the deep ocean to be at least 1000 m below the surface. I will use the 200 m definition, unless otherwise stated.

⁴Interestingly, the average depth of the ocean is approximately 3800 m; above sea-level, the average height of land is 840 m.

⁵Rosettes are often a metal cage with instrumentation and Niskin bottles that is attached to a ship via a cable; it can be raised and lowered in the water column with a shipboard winch.

(AUVs), all equipped with specialized *in situ* instrumentation and often bottles for water sampling, are available. ROVs and HOVs have enabled detailed study of venting chimneys and diffusive venting fields on the seafloor, literally “putting eyes” on the structures and physically interacting with them. To study plumes generated by vents, AUVs and rosettes can be used to examine the water column. However, using these technologies to produce detailed studies of plumes is more challenging than studying a vent. Measurable plume signals are naturally variable, turbulent, and ephemeral and navigating in the water column is a logistical challenge due to lack of physical features by which to navigate or localize. Thus, many detections of a plume during a sample-collection dive tend to be serendipitous in practice.

To leverage these detections, laboratory experiments to model how plumes are expected to manifest in the water column have served a critical role in converting field observations to statements about energetic characteristics, nutrient transport, and overall impact in a basin. Among the most widely used models that describe hydrothermal plumes are the Morton, Taylor, and Turner (MTT) model for buoyant plumes derived in the 1950s (Morton *et al.* 1956), and the re-derivation specifically for hydrothermal plumes in the late 1980s by Speer and Rona (Speer & Rona 1989). These idealized (time-averaged) models describe a roughly two part plume structure composed of a buoyant stem and a neutrally-buoyant layer. The buoyant stem is a spatially small expression that describes the fluid that rapidly rises from an originating vent orifice, driven by buoyant forces that result from the difference in density between the warm venting fluid and the ambient seawater. The neutrally-buoyant⁶ layer describes the spatially large spread of vent-derived fluids along the isopycnal of equal density with the ambient seawater (Fig. 5.1⁷). State-of-the-art hydrothermal plume models incorporate time-varying Navier-Stokes models and more complicated fluid structures (e.g., Lavelle

⁶This can be equivalently styled as non-buoyant.

⁷Readers are additionally referred to Yoerger *et al.* 2007 for an illustrative description and figure.

et al. 2013; Xu & Di Iorio 2012). Mathematically, these models describe what has been practically well-understood in observational studies: the spatiotemporal distribution of plumes is instantaneously complicated and on small scales (meters, minutes) driven by compounding, chaotic factors that are difficult to calibrate.

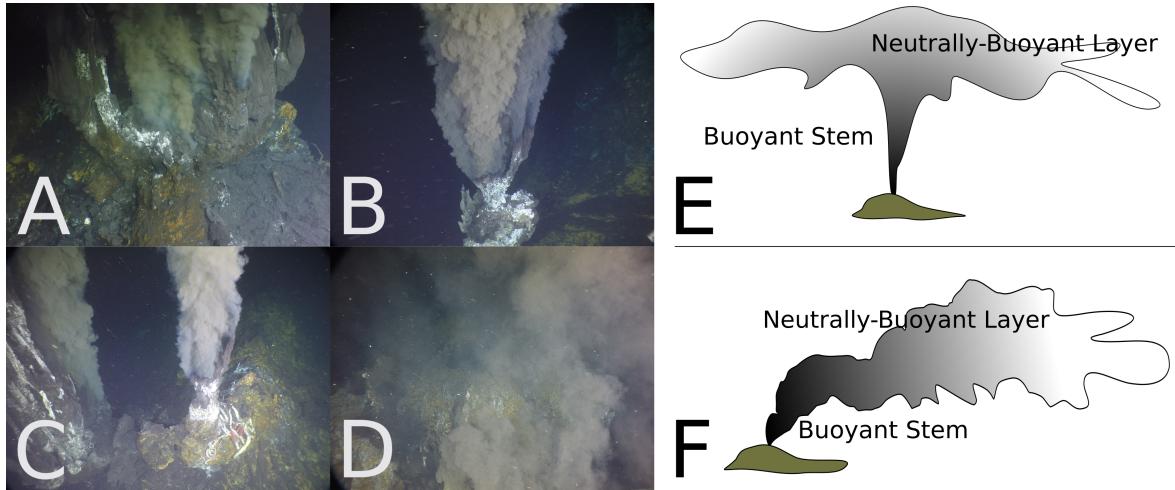


Figure 1.1: Images of hydrothermal vents and illustrative plume structures. Images A-D are different hydrothermal vents observed by remotely operated vehicle (ROV) *JASON* at Guaymas Basin, Gulf of California (Mexico) in November 2021 during a science expedition. A and B show how vents in the basin were formed of many small orifices, creating a complicated chimney structure producing large, turbid plumes. C shows two co-located vent clusters. D is an image of a more diffusive turbid plume. Panels E-F are illustrative examples of plume structures, without crossflow (E) and with crossflow (F). The buoyant stem tends to be a relatively small, coherent structure that rises approximately 100-300 m in a water column (depending on the character of the plume fluid and background seawater). The neutrally-buoyant layer can be spatially large (on the order of kilometers across), but tends to be more diffuse and is found at a height of equal density with the ambient seawater. Crossflow impacts the ultimate rise height of a plume, and the amount of ambient seawater that is entrained (mixed) as the plume evolves.

Field measurements can be used to set the initial conditions or parameters of a numerical model, which can then be used in turn to make claims about characteristics of a plume. This is widely known as solving an *inverse problem*; using observations to find an explanatory set of variables (e.g., initial conditions, numerical parameters). One of the key challenges of solving an inverse problem with observations in the sciences is that the problem is mathematically ill-posed, making it difficult to find the correct, unique

solution due to noise in the observations and the time/space that those observations were taken. For instance, observations along or near the exact centerline of a plume may be significantly more informative about the vent characteristics (Bangian-Tabrizi *et al.* 2022) than samples randomly collected throughout a plume structure (Baker *et al.* 1998). But given that most in-plume field measurements are serendipitous⁸, solving the inverse problem to an acceptable level of accuracy requires grappling with *uncertainty*⁹. Probabilistic representations or analytical models of uncertainty have been used in plume studies to help place confidence intervals over found solutions (Bemis *et al.* 1993; Sohn *et al.* 2019). In classical studies, uncertainty is computed following an expedition and after all data is available. In this thesis, an algorithmic extension to the classical uncertainty formalism is extended to enable inference of plume characteristics while at sea. By computing a notion of uncertainty while at sea, strategic changes to the science activities and instrument deployments can be undertaken to target collection of more informative samples.

1.2 CHALLENGES FOR INTELLIGENT AUTONOMY

For geochemical studies in the water column, AUVs are well-positioned to advance long-term monitoring of and exploration in mesoscale¹⁰ deep ocean environments during *in situ* expeditions. Autonomy for a robotic system can fall on an “agency” spectrum: a robot with full agency can make adjustments to its own behavior; a robot with no agency can execute a pre-determined set of tasks without supervision, but cannot change the tasks while performing. Increasingly, AUVs are being developed for deep

⁸Serendipitous in that while trajectories are planned, actually intersecting with a plume is by chance given unknown dynamical conditions.

⁹Both epistemic (uncertainty in a model) and aleatoric (uncertainty due to inherent randomness). Aleatoric uncertainty in this case comes from the chaotic nature of spatiotemporal distributions (for instance, turbulent flows). Epistemic uncertainty is definitional, as we have uncertainty of the model and access only to noisy observations.

¹⁰Tens of meters to several kilometers

sea research (Kaiser *et al.* 2016; Maki *et al.* 2014; Okamoto *et al.* 2019; Yuh 2000), but their autonomous capabilities are typically limited to executing predetermined hand-designed trajectories such as uniform coverage “lawnmowing” patterns (Camilli *et al.* 2010), falling into the category of robots without agency. This restriction is often applied in order for trajectories to be rigorously checked by engineering teams prior to execution as an operational policy of risk reduction¹¹, and for ease of supervision during execution.

Operating without agency necessarily restricts the class of phenomena that can be effectively studied by expeditionary robots used in the science fleet today. Using non-adaptive surveying strategies, spatiotemporal distributions can be severely undersampled or missed completely (Flasphohler *et al.* 2019; Preston 2019). This can be mitigated when the underlying model of the spatiotemporal dynamics is known, but for reasons explained in the previous section, in field settings a dynamics model must be estimated from data because the environmental condition is initially unknown (Fig. 1.2). This estimation process is challenging, as data that can be collected in real field trials tends to be noisy and *partially observable*—that is, the observations are only at point locations in time and space, and may be indirect measurements of a desired field of interest. Consequently, data-driven modeling techniques used in robotics to describe the environment in which a robot finds itself, such as Gaussian processes (Rasmussen & Williams 2004), neural networks (Cohn 1994; Wang *et al.* 2017), or particle representations (Silver & Veness 2010), require many observations to generalize to a useful model for planning; a luxury that, in the field, is typically not afforded due to limited opportunities for deployments and finite expedition timelines.

Access to numerical models of plume dynamics stands to significantly relieve the burden on data alone to recover a descriptive model of the target environment during an expedition. Scientific machine learning (SML), an emerging subfield of machine learn-

¹¹Risk in the sense of preventing vehicle loss or damage, as these assets are generally rare and so precious in the scientific community.

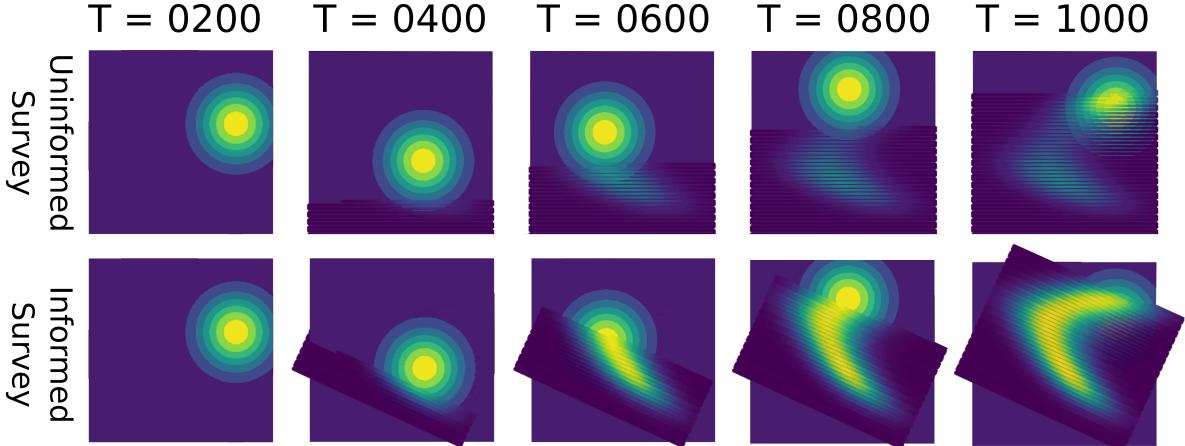


Figure 1.2: Illustrative comparison of an informed and uninformed survey of a spatiotemporal distribution. Several snapshots of an illustrative “bullseye” world are provided, in which two lawnmower trajectories are planned. At each time snapshot, the collected observations by the two surveys are plotted on top of the true underlying phenomenon. In the top row, an uninformed lawnmower starts in the bottom right corner of the world and travels upward. Most samples collected are at the fringes of the bullseye distribution. In contrast, the bottom row shows a lawnmower with all of the same characteristics (resolution, height, width) with an adjusted orientation, informed by the knowledge of the underlying bullseye dynamic. This survey collects a significant number of samples at the center of the bullseye and effectively tracks the phenomenon throughout the duration of the simulated mission.

ing, has shown that leveraging numerical scientific models of physical principles within classically data-driven frameworks (Brunton *et al.* 2016; Jiahouo *et al.* 2021; Kulkarni & Lermusiaux 2019; Mohan *et al.* 2019; Raissi *et al.* 2019; Raissi *et al.* 2018; Sapsis & Lermusiaux 2009) can improve the data efficiency of a learner and the overall quality of the model that is uncovered. To extend SML algorithmic theory to field settings, in which computational time is limited and field observations are corrupted with noise (and partially observable), requires careful selection of both the underlying numerical model and the learning framework that wraps it. Foundational principles (e.g., conservation of mass, non-divergence) and idealized, time averaged models (e.g., MTT (Morton *et al.* 1956)), provide an informative basis for learning that describes the “shape” of a plume distribution, but abstracts the details away to enable tractable computation. Notion-

ally, the idea of using a simplified basis on which to perform inference relates closely to reduced order modeling techniques in dynamical control theory and computational sciences (Burkardt *et al.* 2006; Lucia *et al.* 2004; Salam & Hsieh 2019), or latent space discovery in machine learning (Lu *et al.* 2020; Voynov & Babenko 2020). In these works, the aim is to uncover the smallest set of basis vectors/functions to describe a phenomenon to a desired level of accuracy. Instead of discovering these spaces via data however (which would require too many samples, full-state observations, or an infeasible number of AUV deployments), the “latent space” is directly set by the initial conditions and parameters of a numerical model, and its relationship to the observational space is the model itself. The inference framework that then sits on top of that model can “absorb” uncertainty that arises due to the mismatch between the resolution of the model and the instantaneous measurements that are collected, can be used to describe unmodeled characteristics (i.e., act as a probabilistic closure model), or simply extend the representation of the model to enable computation of information-theoretic measures that can assist with decision-making.

In practice, embedding scientific knowledge into probabilistic models (i.e., belief representations) for planning requires additional algorithmic infrastructure that (1) casts AUV observations into a space that aligns with the solution of the numerical model, and (2) can design operationally satisfactory sampling trajectories that leverage belief effectively. Scientific observations are typically taken by heterogeneous sensors with different operating principles. While some numerical models provide projections of plumes that include observable quantities (Speer & Rona 1989) (e.g., temperature, salinity), many more instead provide a generic “tracer” distribution that does not neatly map to sensors that are available on an AUV. Thus, to fingerprint a plume from data requires characterizing the potentially complicated interrelationships between sensor measurements (Jakuba 2007). Science experts are particularly good at this task, as they

often have years of intuition for the data, experience with specific scientific instruments, and domain knowledge of how plumes have manifested in other studies. In this thesis, tools that assist science experts with identifying hydrothermalism from long AUV surveys (which can be tens of thousands of point observations), and a fully automatic detector informed by science expertise, are presented. The latter is subsequently used as the observation model for the physically-informed belief representation to assist in downstream planning tasks.

A “decision-maker” for the purposes of planning AUV deployments can either be a person, a team of people or an autonomous agent. To enable a decision-maker to make informed choices about a future deployment, the output of the probabilistic model must be interpretable by some means—either semantically for human-readability, or technically for automated optimization via the specification of a reward function. A decision-maker must work under often complicated and changing constraints in the field, which can disrupt or inhibit the planning process. For instance, take the case in which trajectories need to be computed hours before a planned deployment of an AUV in order for them to be rigorously checked prior to execution. Weather, unexpected delays in other science team activities, or emergency AUV maintenance can all change the actual time that the robot is to be deployed. For those plans to still be effective, they must either be timing agnostic or flexible enough to be adapted. As computing universally good plans under any change of conditions is intractable, more heuristic opportunities to design easily interpreted and modifiable plans (potentially by hand) must be undertaken, to ultimately avoid a multi-hour long replanning-rechecking procedure. In this thesis, the probabilistic model provided to a decision-maker is shown to be semantically meaningful for a science team, and can additionally support autonomous planning. A novel algorithmic planner is also proposed, which produces modifiable

plans that respect all AUV and operational policy constraints for a given deployment scenario to improve sample collection for the task of hydrothermal plume charting.

1.3 THESIS CONTRIBUTIONS

This thesis presents both scientific and algorithmic facets of charting deep sea hydrothermal plumes, with discussion grounded by field work conducted at the Guaymas Basin, Gulf of California (Mexico) with AUV *Sentry* in November 2021 during research cruise RR2107. Discussion of hydrothermal plumes and presentation of autonomy frameworks are interwoven in each chapter, following a guiding philosophy in this research that field contexts and algorithmic solutions should closely support one another. To this end, the case for embedding scientific knowledge into the observation model, belief representation, and decision-maker that form the reasoning framework of an intelligent autonomous agent is made. By embedding scientific knowledge, core challenges in planning under uncertainty related to AUV non-agency, partial observability in spatiotemporal distributions, and practical operational restrictions can be overcome to ultimately enable more effective studies. Similarly, by wrapping scientific knowledge in probabilistic representations, novel scientific queries can be supported during and after an expedition, directly advancing classical paradigms in scientific discovery and assisting a science party.

1.3.1 PERCEIVE

Detecting hydrothermalism from sparse, point-observations from long surveys conducted by an AUV requires an intimate understanding of plume physics and deployed instrumentation. There is no “plume-detector” on an AUV. Domain experts bring significant knowledge to the task of fusing multiple heterogeneous sensor observations together to assert locations in which an AUV encountered a plume. How disparate

quantities like methane, oxygen, temperature, salinity, and turbidity manifest generally in the ocean and differently in plumes is a nontrivial estimation problem, especially exacerbated by the unique signature that even hydrothermal vents of the same type (e.g., black smoking chimneys) can have as it relates to the type of magmatic forces under the crust, the crustal quality, and geographic region on Earth. While at sea, classification of hydrothermalism from field observations may be undertaken by a science party in order to inform subsequent sampling tasks. This places an inordinate burden on domain experts while in the field, as each day of an expedition is precious. To assist domain experts in identifying hydrothermalism from AUV data, a set of anomaly detection and temporal analysis tools are presented in Chapter 4 that provide succinct summaries of data and suggested detections for verification or closer study by a science party. These tools are used to demonstrate the effective sensitivity of different instruments at detecting hydrothermalism in the Guaymas Basin, showing that confident detections up to 4-7 km from a venting site are possible in the first study to quantify the extent of hydrothermal expressions in the Northern Basin.

In Chapter 6, the process of identifying hydrothermalism from instruments on AUV *Sentry* is fully automated to support the process of updating a physically-informed belief model of hydrothermalism in the Basin. The proposed sensor model is additionally complemented by work to model other informative environmental properties, including time-varying crossflow (current) and background seawater density stratification (driven by temperature and salinity) from “sensors of opportunity” deployed asynchronously to AUV *Sentry* surveys during other science party activities. Leveraging external sensing is not uncommon in terrestrial or atmospheric field robotics applications, in which satellite, observatory, large field-of-view sensors, or significant prior knowledge are available (Desaraju *et al.* 2015; Everett *et al.* 2019; Heaney *et al.* 2007); however in the deep ocean very little prior information may be known about a particular site save for the location

of a previously discovered vent¹². Thus, developing a methodology for incorporating external sensing for deep sea field operations both alleviates the burden on the *in situ* observations collected by *Sentry* to recover a dynamics model of a plume, and positions the belief model as a universal data aggregator broadly useful for other science activities.

1.3.2 PREDICT

A belief representation encompasses the knowledge of an autonomous agent, and for decision-making, it is useful for this representation to support inference over unseen states. In planning under uncertainty, data-driven probabilistic modeling techniques, like Gaussian processes (Rasmussen & Williams 2004), are used to formulate an agent’s belief over an environment. Data-driven techniques are attractive as they can uncover an explanatory model of an environment without requiring significant prior knowledge of the structure of the environment. They are not silver bullets however; to discover spatiotemporal models of an environment typically requires a significant amount of data or full-state observations, which are unsuited to field contexts. Fortunately, in scientific contexts there is typically access to at least some information about how an environment may be structured. Time-averaged idealized models of plume rise can serve as a layer in an inference framework to provide an inductive bias for a learner to better leverage collected observations and extrapolate to unseen times and places. In this thesis, a novel idealized model for buoyant plume rise in crossflow (i.e., current/advection forces) is applied to the hydrothermal plume setting (Tohidi & Kaye 2016) and coupled with a Bayesian inference framework to update the model from realistic field observations in Chapter 5. The framework, PHUMES: PHysically-informed Uncertainty Models

¹²This is largely because temporal changes in the deep ocean are often difficult to characterize from one-off studies, and there are no universal simulators of the deep ocean as there are for surface features (e.g., ROMS (Moore *et al.* 2011)) to leverage.

for Environment Spatiotemporality, is demonstrated in simulation and in the field (Chapter 6). PHUMES is capable of ingesting data from both AUV *Sentry* and external sensors, and precisely forecasting where plumes will manifest in Guaymas Basin far into the future (at least several days) from only sparse detections common in AUV surveys of spatiotemporal phenomenon.

The choice to use an idealized model that captures the effect of crossflow on plume dynamics is motivated by a desire to improve the predictive power of the probabilistic model for planning sampling trajectories, and choosing such a model for geochemical studies is novel. Classically, *stationary* idealized models that do not consider crossflow are used in post-expeditionary analyses (e.g., Barreyre *et al.* 2012; Mittelstaedt *et al.* 2012; Murch *et al.* 2020) in order to estimate energetic characteristics of a vent. Even in particulate deposition studies, a stationary idealized model is used to characterize a hydrothermal plume expression, and the impact of advective forces is separately modeled with a simple advection-diffusion model applied to a given altitude (typically the neutrally-buoyant height) to simulate any observed lateral transport of plume particulates. In stationary models, estimation of the height of the neutrally-buoyant plume is directly used to solve the inverted problem for vent characteristics; this choice strictly underestimates the energetic characteristics of a plume generated in a basin with even weak crossflow present, as crossflow has the function of pushing the neutrally-buoyant plume height lower in the water column. Moreover, modeling transport using an advective-diffusive model ignores the impact of mixing that occurs along the length of an rising plume under crossflow, leading to overfit model parameter estimates that may have little physically-meaningful interpretation. In Chapter 6, an example of how PHUMES with crossflow-inclusive numerical model can be utilized to estimate hydrothermal characteristics is presented. PHUMES provides both a more expressive estimate of energetic vent characteristics by virtue of capturing the impact

of temporal changes to the environment, and supports analysis of the formation of complicated neutrally-buoyant *intrusions*¹³ observed in vertical transects of the water column. Several new hypotheses on the impact of hydrothermalism in Guaymas Basin and defining characteristics of plume physics are posed via the evidence in the analysis, which could be informative of future expeditions or laboratory experiments.

1.3.3 PLAN

One of the key responsibilities of a science party at sea is to design activities that collect data that can assist in the investigation of scientific queries. Planning strategic sample collection in spatiotemporal environments with only partial information is clearly a difficult problem, which is only exacerbated when operating under constraints that impact the actions or activities that can be performed. In Chapter 5, a trajectory optimization scheme for AUV *Sentry* is proposed called PHORTEX: **P**Hysically-informed **O**perational **R**obotic **T**rajectories for **E**Xpeditions. PHORTEX utilizes forecasts from PHUMES to strategically place *Sentry* at the right place at the right time to encounter a moving plume expression while respecting operational constraints on the robot, including non-adaptivity, altitude limitation, and preferred form for trajectory patterns. In practice, the optimizer produces a set of waypoints that can be ingested by the *Sentry* engineering team safety check processes and *Sentry* mission planner that allows *Sentry* to perform a multi-hour (10-24 hr) dive without intervention that is expected to track and densely survey a plume for a given time of deployment. To be flexible to unexpected changes to deployment times, the trajectories which are designed are *chains* of lawnmower patterns with links in the chain included before and after a given window. Adjusting the trajectory requires only snipping off links that are no longer relevant as a deployment window changes. Additionally, by being lawnmowing trajectories, each

¹³An intrusion is a persistent injection of plume water for a given altitude/depth in a water column.

link in the chain is naturally *exploratory*, and so highly accurate localization or timing is not necessary to make plume intersection likely.

In simulation, it is shown that trajectories designed by PHORTEX can collect significantly more in-plume samples than typical surveying strategies used by AUVs for deep ocean mapping. Importantly, these samples are also more temporally and spatially diverse than serendipitous encounters: a plume is re-encountered throughout a dive, and the structure of the plume is surveyed at multiple distances. In the field (Chapter 6), PHORTEX trajectories are shown to perform at least as well as expert designed trajectories in terms of total sample collection, however temporal and spatial diversity gains are maintained, with important downstream implications for scientific post-expedition analysis. The field deployment is the first iterative deployment of an AUV for deep sea plume charting, illustrating a novel capability for geochemical research and puts over 75% of known vent fields (Beaulieu *et al.* 2013) within reach for strategic surveying.

1.4 THESIS OVERVIEW

This thesis is organized into chapters which expand upon this introduction. Chapter 2 presents the Guaymas Basin field site and operational considerations of deep sea geochemical research for field roboticists. Chapter 3 provides background and discussion of related work across robotics, planning under uncertainty, geochemistry, and oceanography. Chapter 4, Chapter 5, and Chapter 6 present the core contributions of this thesis, with content adapted from Preston *et al.* 2022b and Preston *et al.* 2022a. Discussion of the future of expeditionary robotics and proposed future work is presented in Chapter 7, and the thesis concludes with Chapter 8.

2 OPERATIONS AT SEA AND FIELD

WORK IN GUAYMAS BASIN

How inappropriate to call this planet Earth when it is
quite clearly Ocean.

Arthur C. Clarke

Field robotics is a subfield of robotics research devoted to enabling sophisticated robotic autonomy or control in natural environments for performing real¹ tasks. This thesis presents autonomy tools for the problem of hydrothermal plume charting in the Guaymas Basin, Gulf of California using AUV *Sentry*. This chapter provides a description of the Guaymas Basin and the expeditionary cruise in which *Sentry* was used for plume charting. Additional insights on the general challenge of deploying robots for deep sea research, what going to sea entails, and details on operations that complemented the autonomy study are also discussed.

2.1 GUAYMAS BASIN AND EXPEDITION RR2107

In November 2021, research cruise RR2107 aboard the research vessel (R/V) *Roger Revelle* was conducted at the Guaymas Basin in the Gulf of California (Mexico), which lies between the Baja peninsula and mainland Mexico (Fig. 2.1). The Guaymas Basin

¹As in, the task has consequences and stakeholders associated with the performance of the vehicle and the data it collects.

is an early rifting environment at a depth of 2000 m (Moore & Curran 1982; Scholz *et al.* 2019; Teske *et al.* 2016), meaning that hydrothermal fluids circulate from (geologically) young oceanic crust². Unlike at mid-ocean ridges, magmatic expressions in the Basin are generally doleritic sill intrusions in thick, organic-rich sediments, rather than extrusive basaltic crusts (Lonsdale & Becker 1985; Teske *et al.* 2019). As these intrusions are often shallow, hydrothermal fluids circulating from the sediments are typically biogeochemically rich; iron (Scholz *et al.* 2019), methane and carbon dioxide (Geilert *et al.* 2018), and manganese (Campbell *et al.* 1988) have been particularly noted in previous studies.

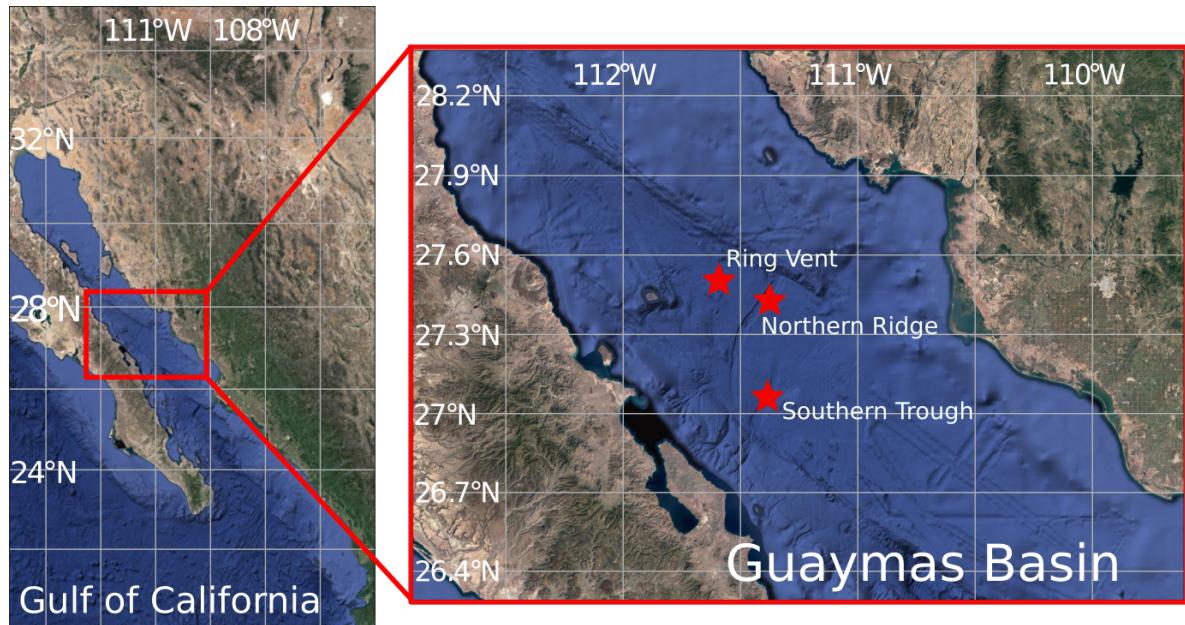


Figure 2.1: **Guaymas Basin, Gulf of California.** The Gulf of California and the Guaymas Basin within it are visualized with satellite imagery and composite satellite and open-source bathymetry (as provided by the GoogleTiles API). The Southern Trough, Northern Ridge, and Ring Vent are marked with stars; the Northern Ridge and Ring Vent were of particular focus during research cruise RR2107.

There are two axial troughs (spreading segments) in the Guaymas Basin, commonly referred to as the Northern and Southern Trough. There is a long documented his-

²Mainland Mexico and Baja California separate at a rate of approximately $5\text{--}6 \text{ cm yr}^{-1}$ (Lonsdale & Becker 1985)

tory of hydrothermal expressions in the Guaymas Basin, with a bulk of these studies focused primarily on hydrothermalism in the Southern Trough (e.g., Lonsdale & Becker 1985; Ondréas *et al.* 2018; Seewald *et al.* 1994; Teske *et al.* 2016; Von Damm *et al.* 1985). In contrast, RR2107 was focused on studying two Northern sites at or near the Northern Trough: Northern Ridge and Ring Vent. At these sites, the expedition had several key objectives: test novel *in situ* instruments to measure dissolved methane (Kapit & Michel 2021a; Kapit & Michel 2021b; Michel *et al.* 2022), test novel *in situ* instruments to measure dissolved inorganic carbon, map the heat distribution in shallow sediments above hydrothermal sills³, collect tubeworm and geological specimens, and collect biological samples of microbiota in hydrothermal plume-fluids to re-construct the structure of a plume microbiome. To enable all science operations, AUV *Sentry*, ROV *JASON*, and standard oceanographic equipment (i.e., Conductivity-Temperature-Depth (CTD) rosette, Niskin carousel, shipboard acoustics) were available.

Ring Vent is a diffusive off-axis venting site approximately 28 km northwest of the Basin spreading center, named for the circular shape of the 1 km diameter site ringed by raised hydrothermally active structure (Teske *et al.* 2019). As Ring Vent is a diffusive site, one of the primary research goals at Ring Vent was to map the thermal gradient in shallow sediments there; as such ROV *JASON* was primarily used.

The Northern Ridge is a recently discovered hydrothermal site (Geilert *et al.* 2018; Soule *et al.* 2018), approximately 1850 m underwater and at the edge of an additionally 300 m deeper graben. The ridge is approximately 600 m long and features several tall sulfide structures 10-25 m in height with active smoking along their bodies. At the top of each sulfide structure is a “chimney” producing smoky, turbid water highly enriched with carbon dioxide, hydrogen, and methane. The chimney vent is composed of a cluster of tens of small orifices (<0.1 m diameter) that create up to an approximately 1.5 m diameter chimney base. Measured with a temperature wand on ROV *JASON*,

³In extension of, e.g., Neumann *et al.* 2022

fluid temperature at the vent orifice is approximately 340 °C and the fluid ventilates rapidly. In contrast, the ambient seawater at the same nominal depth of the chimneys is methane-poor, considerably less turbid, and cold at 4 °C. As vent fluids rise and form plumes at the Northern Ridge, ambient water mixes (entrains) into the plumes at an unknown rate, and mild deep-water currents up to approximately 0.15 m s⁻¹ observed during the expedition advect the fluids. Under these conditions, plume expressions could be transported several kilometers from a known source, and would be expected to rise over 200 m in the water column (Speer & Rona 1989). As an energetic site producing large plumes, the Northern Ridge was the primary site for AUV *Sentry* and near-bottom water-column studies on RR2107, particularly to support tests of the novel *in situ* instruments and collection of microbiota samples. For both of these tasks, charting different regions within a plume is important to test the capabilities of the instruments and to collect biological samples from a diversity of plume-conditions.

2.2 CHALLENGES FOR ROBOTS AND AUTONOMY IN THE DEEP OCEAN

In the Guaymas Basin setting, or any deep ocean research, there are several unique challenges to deploying robotic platforms in contrast to terrestrial applications. Perhaps the most quintessential of these challenges is the conductive nature of water, and its corresponding attenuation of radio frequencies (Qureshi *et al.* 2016). The natural consequence is that ubiquitous terrestrial technologies leveraged by robots and humans alike, such as global positioning satellites (GPS), imaging satellites, and radio-based wireless communication, are not available for underwater navigation, mapping, or communication. The most common alternative for wireless communication underwater are acoustic waves, which can travel hundreds of kilometers and can be used effectively for

ranging (e.g., estimating distance and angle between a source and reflector), but have a significantly reduced data transmission rate (Qureshi *et al.* 2016). Global positioning of an underwater platform is typically done via acoustic ranging with a ship, the ship then having access to GPS to perform an appropriate coordinate transformation for the robot. While accurate, the bit-rate and time-of-flight for information transmission is generally not sufficient for constant localization of a vehicle (let alone also requiring a ship to maintain proximity to a platform, and forgo communicating other types of information), and so many state-of-the-art underwater vehicles use “dead-reckoning” to estimate their position with inertial sensors. The use of Doppler velocity loggers (DVL), downward facing acoustic sensors that estimate velocity relative to another object, has improved the overall accuracy of dead-reckoning for near seafloor (within 150 m) navigation (Fong & Jones 2006; Liu *et al.* 2022; Rigby *et al.* 2006; Stutters *et al.* 2008). To use a DVL requires proximity to the seafloor (within tens of meters) in order to provide a “bottom-lock” useful for navigation; this further restricts an AUV’s navigational freedom.

Beyond navigation and communication, the challenge of working in water necessarily impacts the type of scientific sensing that can also be accomplished. Light, just like radio, is also severely attenuated in water; sunlight typically only penetrates the ocean to about 200 m, and up to 1000 m in the best conditions⁴. This means that vision based technologies, which have enjoyed significant development in robotics for e.g., autonomous driving tasks, are entirely restricted to either shallow depths or very near seafloor studies, in which light sources carried by a platform might be used⁵. Other common optical-based sensors in robotics, such as Lidar, are similarly difficult to use underwater due to attenuation and refraction. This is not to say that there are no optical sensors that can be used underwater. For short distances, transmitted light between a

⁴Depths between 200-1000 m are aptly part of the ocean’s “Twilight Zone” (Martin *et al.* 2020).

⁵Of course, navigating in the water column near no other structure is not necessarily a setting for which vision-based navigation may easily extend, even if there was light.

source and receiver can be used to estimate water turbidity (Bishop 1999). Light can also be used to transmit information between a source and a receiver at a higher bitrate than acoustics over modest distances (tens of meters) (Farr *et al.* 2010; Qureshi *et al.* 2016). Other sensors may use light, but in creative ways—for instance, oxygen optodes detect luminescence of a chemical reaction that corresponds to a measure of dissolved oxygen (Nicholson & Feen 2017), and laser-based spectrometers in depth-capable housing use membrane inlets to accept gaseous samples to analyze (Wankel *et al.* 2010). While many of these technologies are still in their nascent development phase, perhaps one of the most ubiquitous sensors in oceanography (other than acoustic instruments) is the conductivity-temperature-depth (CTD) probe, which uses induction to measure salinity, resistivity to measure temperature, and a pressure sensor to measure depth (Rudnick & Klinke 2007). With these measurements, computation of density is possible, which is the primary driver of water mass mixing in the ocean.

No matter the sensor, water-column sensing and mapping in the deep sea relies on collecting point observations and reconstructing fields of interest from these incredibly sparse data. To address this data sparsity issue, there are several large-scale efforts that contribute to the instrumenting and global understanding of the ocean. Argo, an international network of thousands of small buoyancy-controlled floats which drift in the ocean and take basic *in situ* measurements (e.g., temperature, salinity) of the water column up to 6000 m in depth, is one such example (Jayne *et al.* 2017; Roemmich *et al.* 2009). With a finer degree of control, small glider networks operated by the Ocean Observatories Initiative (OOI)⁶, offer an opportunity for remote targeting of particular regions or depths to study larger-scale phenomenon. Several highly instrumented “observatories”, such as the Martha’s Vineyard Coastal Observatory (Austin *et al.* 2000) or the Endurance Array in the Northeast Pacific (Barth *et al.* 2018), also serve an important role for collecting highly temporally resolved data at specific sites. It is certainly

⁶<https://oceanobservatories.org/marine-technologies/gliders/>, Trowbridge *et al.* 2019

the case that there is a lot of ocean data; what has yet to be standardized is how to leverage this data for enabling targeted research studies, particularly with highly capable robotic platforms. Central to this challenge is that despite a wealth of data, the ocean is truly vast; in ten years of available records from Argo, there are no records of a float in the Guaymas Basin⁷. There are no instrumented arrays in the Gulf of California, and glider or robotics studies conducted there have been part of singular research cruises conducted by disparate research teams, with mixed data discoverability and accessibility.

While initiatives like the UN Ocean Decade⁸ and the NASA EarthData open science program⁹ stand to transform the challenge of finding data, practically, oceanographic researchers are prepared to continue to be self-sufficient for any given expedition. For a roboticist, this requires becoming familiar with prior research at a given site (if available) to contextualize how certain sensors will map to a task at hand, with the scientific instruments on a vehicle and their operating principles, and with the activities of other science team members to effectively share data within the party. Understanding on all of these axes helps to identify any instrumentation gaps for robotic tasks prior to boarding a vessel (so they can be either rectified by preparing new instrumentation, or the impact mitigated by identifying useful proxies), effectively leverage the expertise of the entire science team to inform missions while underway, and clarifies the design of the post-cruise data products to support science objectives.

2.3 THE SCIENCE PARTY AND RESPONSIBILITIES

Before a robot even touches the water, a massive amount of operational overhead is necessary to support deep sea studies. In general, deep ocean research requires using an

⁷As determined via the Euro-Argo data exploration tool, accessed November 20, 2022.

⁸<https://www.oceandecade.org/>

⁹<https://www.earthdata.nasa.gov/esds/open-science>

oceanographic vessel. The R/V *Revelle* is one of several research ships in the University-National Oceanographic Laboratory System (UNOLS) network available through federal support by the United States¹⁰. The R/V *Revelle*, operated by the Scripps Institution of Oceanography at the University of California, San Diego¹¹, is a 273 ft vessel constructed in 1996 and completed a mid-life retrofit in 2020. The *Revelle* houses 21 crew and 37 science staff for any expedition. A 4000 sq ft laboratory space dominates one deck of the ship, containing storerooms, a computer lab, a science freezer, and wet lab (among other spaces). The ship is additionally equipped with acoustic surveying equipment, a rosette bay and winch, chemical hoods and storage, dredges, and weather stations in addition to general office equipment. Any additional pieces of equipment—AUVs, ROVs, specialized sensing equipment, laboratory equipment, computers—are carried on by the science party.

The captain and crew on a research vessel are responsible for the overall safety of the operations, and are key stakeholders in all operations. Marine technicians are crew members that specialize in the scientific instrumentation and infrastructure that belong to a ship, and directly work with the science team to coordinate instrument deployments and perform data collection and curation. The science party itself can be a professionally diverse group of trainees, scientists, and engineers from multiple fields. On RR2107, the science party comprised *Sentry* and *JASON* teams (fourteen personnel), two teams from the Woods Hole Oceanographic Institution (WHOI) comprising engineers, computer scientists, and geochemists (ten personnel), a team from Harvard University and affiliates in biology, geochemistry, and microgenomics (five personnel), and a team from Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE) comprising geodynamicists (two personnel); all teams included graduate students. Everyone on a research vessel not only works together, but must

¹⁰Private research institutions outside of this network, like the Schmidt Ocean Institute, also provide some global-class ships available for deep sea research.

¹¹<https://scripps.ucsd.edu/ships/revelle>

live together for the duration of the expedition. This makes oceanographic research (and any “expeditionary” travel) a highly unique working environment¹² in which every member is necessary for making the experience productive and supportive.

For a field roboticist, being a supportive team member may mean performing more than just *Sentry* or *JASON* operations. Literacy in certain data processing and visualization tools, comfort with network protocols and a command line, and hands-on experience working with certain sensors can prove to be extremely useful to other science activities while aboard. Coordinating with other science teams prior to an expedition can be a way to identify areas in which a computer scientist or a roboticist could be helpful. For instance, prior to RR2107, a desire to monitor *Sentry* science data while it was underway was expressed. Although outside of the direct purview of the autonomy operations, creating a tool that listened to a local network on the ship and displaying data that came over the network in a live visualizer was a relatively straightforward process and provided the science team with a “real-time” view of *Sentry* operations, informative for other activities. In practice, a field roboticist has a lot of tools that they bring with them on an expedition; there is increasing interest in using more of those tools for oceanographic missions as data infrastructure and use of *in situ* sensors accelerates.

2.4 AUV *Sentry*

AUV *Sentry*, operated by the National Deep Submergence Facility (NDSF) at WHOI (Kaiser *et al.* 2016) is a general purpose platform with depth capability up to 6000 m (Fig. 2.2). *Sentry* is the successor of *ABE* (Yoerger *et al.* 1991), the first widely uti-

¹²As late as the 1960s women were barred from working on oceanographic vessels, and this unique working environment can lead to modern inclusivity challenges. Recognition of long-standing challenges to diversity and inclusion in oceanography has inspired new curricula for aspiring chief scientists to design supportive at-sea environments, e.g., <https://cobra.pubpub.org/>

lized AUV for deep ocean research by the U.S. oceanographic community. *Sentry* has completed upwards of 600 dives, and is equipped with three classes of instruments: vehicle sensors, geophysical sensors, and oceanographic sensors. Vehicle sensors are largely those that assist with AUV localization; geophysical sensors include acoustic multibeam, sub-bottom profiler, and magnetometers, which are useful for collecting bathymetric and sub-seafloor maps. For plume studies, the oceanographic sensors are of primary interest, and include a CTD, optical backscatter (OBS) instrument, oxygen optode, and oxidation reduction-potential (ORP) sensor. Each oceanographic sensor is synced to a shared precision clock on *Sentry*, and logged to separate files at the manufacturer recommended sampling frequencies. In addition to the standard sensor suites, *Sentry* can accept some novel instrumentation developed by a science party for an expedition. During RR2107, two novel methane sensors were integrated into *Sentry* to complement the oceanographic sensors for hydrothermal plume charting (Kapit & Michel 2021a; Kapit & Michel 2021b; Michel *et al.* 2022).

The *Sentry* team for an expedition has an appointed expedition leader, who interfaces directly with the Chief Scientist and any other relevant science parties to plan *Sentry* dives. The *Sentry* engineering team is largely responsible for the physical and virtual safety and maintenance of *Sentry* during an expedition. As such, the expedition leader is additionally responsible for confirming the safety of every dive plan and managing pre-dive and post-dive checks. Safety confirmation of a dive plan formally requires passing a simulation dive of a requested trajectory using a set of scripts maintained by the *Sentry* team. This simulation uses bathymetry and marginal environmental estimates to assist in checking a plan for obstacle clearance, in addition to computing possible localization drift (associated with number of turns or degradation of bottom-lock with a DVL) among other heuristics. Less formally, the expedition leader may use contextual knowledge of a location or the vehicle state to deny or suggest changes to

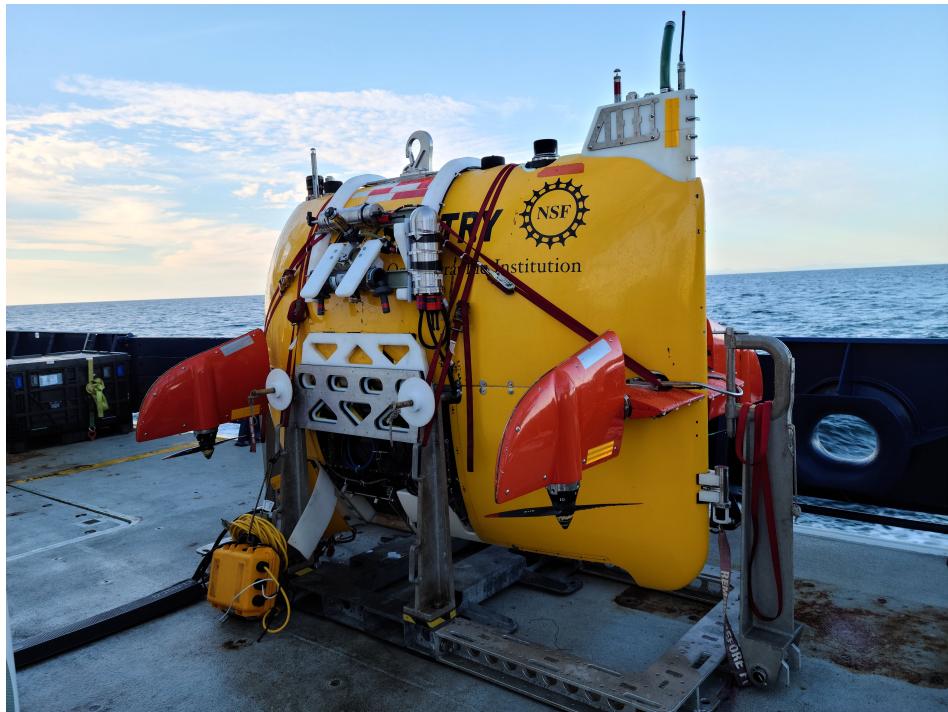


Figure 2.2: **AUV *Sentry*.** A picture of AUV *Sentry* on deck of the R/V *Revelle* during RR2107 activities.

a plan, even if a trajectory passes a safety simulation. For instance, Guaymas Basin has a notably soft seafloor due to sediment deposition. This impacts the quality of DVL estimates, which in turn impacts the quality of *Sentry* localization. This lead to adjusting the altitude restriction for *Sentry* during RR2107 to reduce risk to the vehicle.

2.5 ROV *JASON*

ROV *JASON*, like AUV *Sentry*, is also operated by the NDSF at WHOI (Ballard 1993; Petitt *et al.* 2004; Yoerger & Newman 1986) and a version of it has been in operation since the late 1980s (Fig. 2.3). Distinct from *Sentry*, *JASON* is connected to a ship via a cable which transmits information and power. This allows real-time sensor and visual feedback from the ROV and fine remote control of the vehicle navigation and two manipulator arms. *JASON* is a particularly powerful tool, as it can literally “put

eyes” on the seafloor, and perform physical sample collection tasks that AUVs cannot, including sediment coring, microbial mat retrieval, and macro-flora/fauna and rock sample collection. Standard oceanographic equipment, like CTD, optode, and turbidity sensor are mounted to *JASON* and logged in a similar fashion to *Sentry* equipment. *JASON* can additionally accommodate carrying novel equipment brought by a science party¹³.

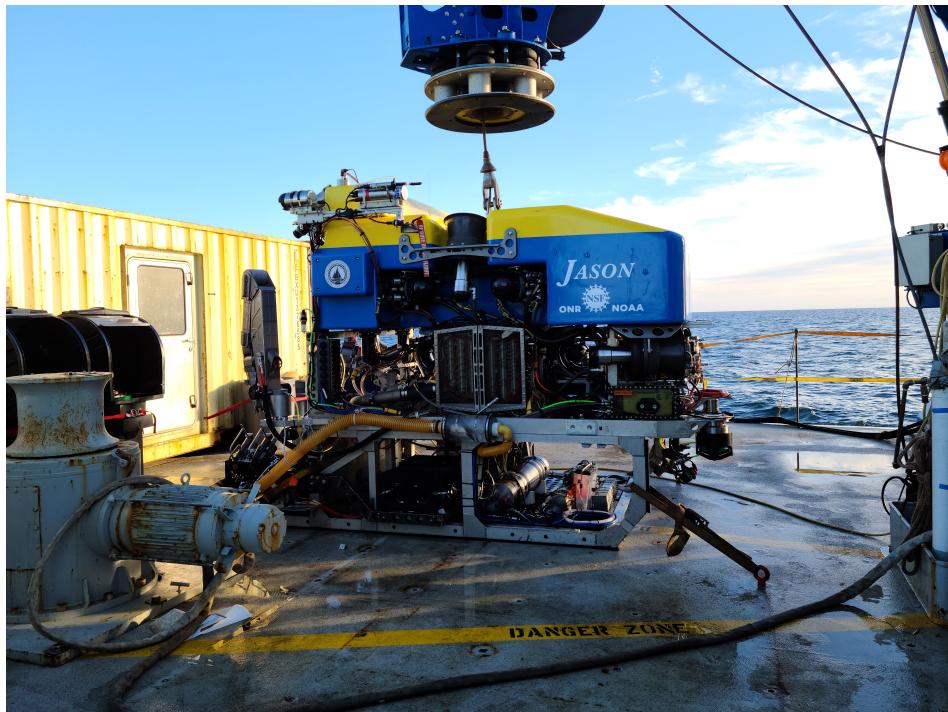


Figure 2.3: **ROV *JASON***. A picture of ROV *JASON* on deck of the R/V *Revelle* during RR2107 activities.

While being operated, three *JASON* engineers act as “pilots” in rotating shifts throughout a dive (which can be over 24 hrs long). These piloting operations occur in the “control van”, a set of shipping containers in which data servers/ship-side data loggers, computer monitors, and annotation stations are set-up. At least three science party members are also required in the control van during operations—a dive science leader and two annotators. The science leader interacts with the pilots and sets the

¹³This requires a pre-cruise certification process for depth-rating.

activities for a dive (e.g., collect this rock, collect some video, navigate to a new location). The annotators listen to the science leader and pilots to keep a running log of science activities and any relevant information about a dive as it occurs. The annotated logs are used post-expedition to search through *JASON* data. The science party is responsible for setting a schedule for science leaders and annotators (a data logger and a video recorder). On RR2107, most members of the science party were assigned at least one 4 hour shift in a 24 hour period in which they would be on call for being in the control van if operations were underway.

AUTONOMY STUDY ACTIVITIES ROV *JASON* was a rich source of external data to assist with the autonomy study presented in this thesis. In particular, *JASON* could physically observe the target vent and deploy additional sensing equipment to compensate for holes in *Sentry* sensing capabilities for the expedition. As described later in Chapter 6 briefly, *JASON* was used to:

- Estimate the fluid temperature at a vent orifice with the temperature wand; measured 340 °C.
- Estimate the size of a chimney and vent via still imagery.
- Deploy MISO cameras at a vent site, used to estimate the fluid exit velocity at a vent via video.
- Deploy current tiltmeters at various locations on the seabed to estimate deep currents at the autonomy test site.

While the first two activities are standard for *JASON* studies, the latter two were specific to the operations in RR2107. WHOI-MISO cameras, self-contained depth-capable GoPro devices¹⁴, were mounted to the brow and one of the arms on *JASON*. Set to

¹⁴<https://www2.whoi.edu/site/miso/>

record in 4k, the arm-mounted camera was aligned to be approximately parallel with an active vent and used to record 1-5 min videos of turbulent plume fluids. The video was then processed following a *JASON* recovery using particle imaging velocimetry (Zhang *et al.* 2019) (PIV), which tracks turbulent “parcels” that have high cross-correlation between frames. By tracking many specific parcels over several frames, PIV methods can provide a vector field of velocity estimates, that can then be averaged to provide an estimate of mean flow in a region. Using MATLAB’s open source PIVLab (Thielicke 2014; Thielicke & Sonntag 2021; Thielicke & Stamhuis 2014), a representative vent in the basin recorded during *JASON* dive JD1390 at (27.4018606 N, 111.3991182 W, 1809 m depth) was estimated to have fluid exiting at 0.7-1.33 m s⁻¹ (Fig. 2.4).

Tilt current meters¹⁵ are mounted to the seafloor by one point and are otherwise free to move under crossflow. The pose of the tiltmeter can be directly converted to crossflow magnitude and heading. ROV *JASON* was used to alternately deploy two TCM-3 tiltmeters while performing other operations. Table 2.1 shows information for each deployment. The first deployment of Tiltmeter A was the primary dataset used to generate an estimate of the current function used for the autonomy study presented in Chapter 6.

Tiltmeter	Dive Deploy/Recover	Duration	Location (N,W)	Depth (m)
A	JD1389/-	28* hrs	27.400618, 111.398532	1832
A	-/JD1390	28* hrs	27.400236, 111.396249	1854
B	JD1389/JD1396	6 days, 15 hrs	27.400618, 111.398532	1832
A	JD1392/JD1392	10 hrs	27.414957, 111.387304	1840
A	JD1393/JD1393	20 hrs	27.500116, 111.683227	1732

Table 2.1: **Summary of tiltmeter deployments.** Two tiltmeters were deployed using ROV *JASON* during RR2107 to estimate the magnitude and heading of deep currents. One tiltmeter was deployed for nearly 1 week, while another was deployed intermittently, so data products self-logged on the instrument could be used for autonomy ops. During the first deployment of tiltmeter A, the tiltmeter was moved to better align logically with other *JASON* operations; the timing for this deployment is marked with an asterisk.

¹⁵<https://lowellinstruments.com/>

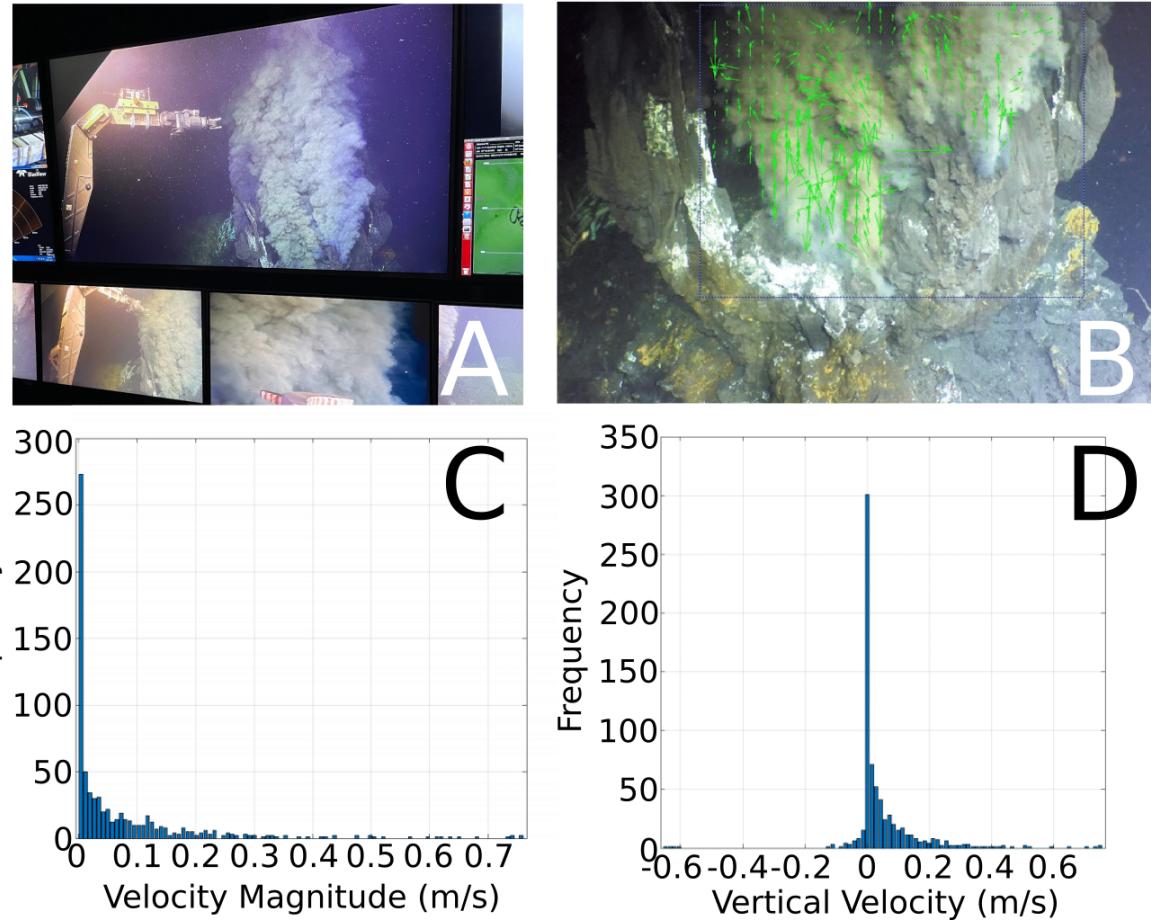


Figure 2.4: Exit velocity estimation with WHOI-MISO cameras. A WHOI-MISO camera was mounted on an ROV *JASON* arm and positioned to film exiting plume fluids from a hydrothermal vent (A). The capture video was processed using a PIV technique (B-D), which provides estimates of fluid velocity. The tails of the velocity field distributions may be indicative of exit velocities near orifices, as most of the imaged field is dominated by turbulent mixing within a meter of the orifices.

2.6 CTD ROSETTE

A CTD rosette (or just rosette, Fig. 2.5), is a standard piece of oceanographic equipment which is typically an instrumented metal cage connected to a ship via a cable and used to collected vertical transects of the water column. Rosettes can also be towed laterally in the water column if lowered and the ship driven in the pattern of the desired transect. Over the cable, data from the instrumentation on a rosette is streamed to a computer

terminal on the ship. The depth is controlled by pulling in or releasing a cable using a winch. The typical workflow is that one or more science team members are responsible for “standing watch” in the computer lab to which data is streamed and displayed, and communicates to a winch operator (always a crew member) via phone with requests to let out or pull in the cable.



Figure 2.5: **CTD Rosette**. A picture of the rosette on R/V *Revelle* during RR2107 activities.

Equipment available on the rosette for RR2107 included a CTD, a fluorometer, an oxygen optode, and a transmissometer (turbidity). All data was logged via a shared clock and in a proprietary format by Sea-Bird Scientific¹⁶, which was converted to CSV files while aboard the *Revelle*. The rosette also had a twelve bottle 10 L Niskin bottle

¹⁶<https://www.seabird.com>

carousel. Niskin bottles are water containers which can be closed (“fired”) to seal water samples for *ex situ* analysis. During rosette operations, Niskin bottles can be closed from the science team computer station by either pre-programming depths before a vertical transect, or by requesting a bottle fire via the data monitoring application. Bottle samples during RR2107 were used to collect microbial samples, dissolved gas samples, and calibrate other *in situ* instrumentation. Eleven deployments (also known as casts) of the rosette were performed, nine of which were standard vertical profiles.

AUTONOMY STUDY ACTIVITIES Within an analytical model for plume rise, density differences are computed in order to estimate buoyant force, which ultimately drives the estimated neutrally-buoyant layer height. While standard stratification (density) profiles for the Atlantic and Pacific ocean basins are widely accepted (Speer & Rona 1989), the Guaymas Basin is a unique semi-closed system in the Gulf of California, and a stratification curve for background seawater in this Basin would lend itself to more accurate estimates of plume characteristics. As an AUV descends, it performs a vertical transect that could be used to compute this local stratification curve. For RR2107 operations, rather than use *Sentry* vertical dives, which were generally altitude limited (thus not getting close to the bottom), multiple high-quality profiles from the rosette were used instead. Using a set of subsampled points from every vertical transect collected by the rosette, a Gaussian process (GP) was fit to the data, and the mean function used as the stratification curve used for autonomy tests.

3 FOUNDATIONAL RELATED WORK

I may remark that the curious transformations many formulae can undergo, the unsuspected and to a beginner apparently impossible identity of forms exceedingly dissimilar at first sight, is I think one of the chief difficulties in the early part of mathematical studies. I am often reminded of certain sprites and fairies one reads of, who are at one's elbows in one shape now, and the next minute in a form most dissimilar.

Ada Lovelace

The research presented in this thesis is built on work spanning multiple fields, including numerical and scientific modeling, planning under uncertainty, robotics, geochemistry, and oceanography. In this chapter, background on topics closely related to the problem of hydrothermal plume charting and autonomous robotic sampling in field environments is broadly provided.

3.1 REPRESENTING DYNAMIC SYSTEMS

Natural environments and the spatiotemporal distributions within them are dynamic systems. Dynamical systems are well represented in the form

$$\dot{x} = f(x, t) \tag{3.1}$$

where the function f may be a nonlinear relationship, and the system may be either an ordinary (ODE) or partial (PDE) differential function.

While analytic forward solutions for some dynamic systems can be found, in most cases these are difficult to compute without significant simplifying assumptions. Numerical solutions, which rely on iterative methods, are more commonly employed. The *finite difference method* (Smith 1985) (FDM) is a popular technique for computing numerical solutions, and which reformulates PDEs into a system of equations that can be solved via matrix operations. To do this requires discretization over the time and space domains of a continuous problem. FDM can be derived from a Taylor series expansion of some function f which is assumed to have proper derivatives:

$$f(x_i) = f_i \tag{3.2}$$

$$f(x_{i-1}) = f_i - \Delta x f'_i + \frac{\Delta x^2}{2} f''_i + \dots \tag{3.3}$$

The goal is then to solve for one of the derivatives of f ; for example the first derivative of the second expression can be rearranged such that:

$$f'(x_i) = \frac{f(x_i) - f(x_{i-1})}{\Delta x} + \frac{\Delta x}{2} f''(x_i) + \dots \tag{3.4}$$

One of the advantages of the FDM is the ability to select the accuracy of the desired approximation. Derived above is a first order accurate approximation of the first derivative, in which only the first term is kept. What is truncated from the solution induces an error on the order $O(\Delta x)$. Second-order accurate methods have an error term on the order $O(\Delta x^2)$, and so on.

In general, adaptive sampling for expeditionary science is focused on predicting time-dynamics of a spatial distribution. Parabolic equations can be used to model these

time-varying systems. The solution to a time-varying PDE/ODE is also known as the *forward problem*: given some parameters and initial condition, find the state of the world at each (discretized) time step. The definition of the problem essentially requires iterative/sequential solvers to be used. For these problems, the iterative or time-stepping methods can be classified as either *implicit* or *explicit* (Biswas *et al.* 2013; Hahn 1991). In an explicit method, only the current state is required to compute the next state; forward Euler is the most widely adopted explicit method, and takes the form:

$$y^{n+1} = y^n + \Delta t f(y^n, t^n) \quad (3.5)$$

where the next state of the system is a function of the current state incremented by the slope (represented by f) computed at the current state. An example, salt diffusion in a vertical water column can be expressed as a parabolic equation with the form:

$$\frac{\partial s}{\partial t} = \frac{\partial}{\partial z} \left(\kappa \frac{\partial s}{\partial z} \right) \quad (3.6)$$

and using the forward Euler scheme, this system can be expanded according to the FDM with the addition of a time step Δt and corresponding increment n :

$$\frac{s_k^{n+1} - s_k^n}{\Delta t} = \frac{1}{\Delta z} (F_k^n - F_{k-1}^n)$$

Substituting the flux terms and rearranging the equation:

$$s_k^{n+1} = \Delta t \left[s_{k+1}^n \left(\frac{\kappa_k}{\Delta z_k \Delta z_{k+1}} \right) + s_k^n \left(\frac{-\kappa_k}{\Delta z_k \Delta z_{k+1}} + \frac{-\kappa_{k-1}}{\Delta z_k^2} + \frac{1}{\Delta t} \right) + s_{k-1}^n \frac{\kappa_{k-1}}{\Delta z_k^2} \right] \quad (3.7)$$

A linear system $Ax = b$ can be defined wherein A takes the form of a tridiagonal matrix of coefficients, x is the current value of the target of interest (e.g., salinity), and b is the value of the target of interest at the next time increment. With this form, a straightforward loop can be used which computes b at each time step and substitutes b for x in the following time step.

Implicit methods, like backwards Euler, take a form which requires knowledge of the future state in order to compute that future state.

$$y^{n+1} = y^n + \Delta t f(y^{n+1}, t^{n+1}) \quad (3.8)$$

Following from the derivation of the implicit form, backwards Euler can simplify to a linear equation $Ax = b$, but where x is the future state and b is the current state. To solve for x an iterative solver (e.g., Jacobi (Forsythe & Henrici 1960), Gauss-Seidel (Usui *et al.* 1994)) can be integrated directly into an outer loop which performs the forward time step.

As compared to explicit methods, implicit methods are more expensive to compute because of a necessary system solve needed at every time step. However, what trade-off exists in speed is compensated for with stability: using various stability techniques (e.g., Von Neumann (Wesseling 1996)) it can be shown that implicit techniques are unconditionally stable for any possible spatial or temporal discretization. In contrast, explicit methods require a constraint on these discretization parameters with respect to the characteristic lengthscale of the phenomenon in order to be numerically stable. To improve stability of explicit methods, advanced time-incrementing methods which can adaptively change the time step based on residual characteristics could be employed.

MODEL ORDER REDUCTION

For very large systems (as in, many discrete states or variables), even iterative methods are too expensive to compute. A subfield of study in numerical computation is therefore focused on *model order reduction*: reducing a large state space into a lower rank embedding (reduced state space, reduced feature space) which can approximate the system dynamics to some selectable error. There are many different model order reduction techniques which can be employed for dynamical systems.

PROPER ORTHOGONAL DECOMPOSITION (POD) POD is perhaps the most common method adopted for fluid studies and reduces the dimensionality of a problem by transforming original unknowns (e.g., salinity in each voxel of a water column) into a new set of variables called modes or principal components (Lassila *et al.* 2014). By virtue of the transform, the first few modes will describe well (i.e., contain most of the “energy” of) the original unknowns. This is accomplished by performing a statistical analysis on “snapshots” of the system computed from several expensive simulations on the full-state of the system (or from observed data, if available). These simulations are used to create a database x and an optimization problem is posed:

$$P^* = \min_P \mathbb{E}[x - Px]^2 \quad (3.9)$$

where a projection operator P needs to be selected such that the error between the original data and transformed data is minimized. In general, the transform that satisfies this optimization is an orthonormal projection in which the modes are the ordered orthonormal eigenvectors of the covariance matrix of x (thus why this is a statistical analysis technique).

Although POD is one of the most popular model order reduction techniques, it suffers from several practical limitations. First, the method relies on access to “snapshots” of

the system in order to learn the basis. On truly large systems this may be prohibitively expensive or impossible; in field settings for the deep sea, access to full snapshots is the latter. Further, POD relies on samples of x being independent, however as a general rule this cannot be assumed in practice. Finally, although the modes may fit a dataset of snapshots well, the modes may not generalize the underlying dynamics of the system well (e.g., it may be incorrect when asked to project outside of the domain of the training data). Some of these concerns have been addressed by computing the covariance matrix via the controllability Gramian of the system, rather than directly from the snapshots, with the idea that the controllability Gramian generalizes the dynamics better than the snapshots do alone (Georges 1995; Zhao & Pasqualetti 2019).

REDUCED BASIS METHODS Like POD, reduced basis methods have an expensive “offline” training period in order to extract lower-rank modes which capture features of the dynamic system and which can be used quickly in online settings (Ohlberger & Rave 2015). In particular, reduced basis methods attempt to find a reduced system of nonlinear equations with a significantly smaller set of unknowns which captures the describes the behavior of the larger system. This is done by applying the Rayleigh-Ritz method (also known as the Galerkin method) on the finite-element form of the eigenvalue problem posed by the PDE, $Ax = \lambda x$ where $A \in \mathbb{C}^{N \times N}$, which yields Ritz pairs $(\tilde{\lambda}_i, \tilde{x}_i)$ which approximate the solution (Noor & Peters 1980). To do so requires the computation of an orthonormal basis $V \in \mathbb{C}^{N \times m}$ where $m \ll N$ which approximates the eigenspace of m eigenvectors. Then, the reduced eigenvalue problem can be posed $Rv_i = \tilde{\lambda}_i v_i$ where $R \leftarrow V^T A V$. The resulting Ritz pairs take the form $(\tilde{\lambda}_i, \tilde{x}_i) = (\tilde{\lambda}_i, V v_i)$.

Selecting the right orthonormal basis V is the critical challenge for this method. In Quarteroni & Rozza 2007, reduced basis methods for solving the Navier-Stokes equations using Lagrangian, Taylor, and Hermite spaces are examined; Lagrangian subspaces

tend to be the most popular selection for reduced basis methods. Historically, reduced basis methods have struggled to describe advective systems which display considerably nonlinear behavior (Ohlberger & Rave 2015; Quarteroni & Rozza 2007).

FOURIER MODES Fourier modes are the result of applying a model order reduction technique on a dynamical system in Fourier space. This form of analysis is popular in works which attempt to place an upper bound on the theoretically finite number of *determining modes* for system, which are a set of parameters used to fully define turbulence or similarly complex structure (Jones & Titi 1993).

LOCALLY LINEAR EMBEDDING Locally linear embedding (LLE) is a dimensionality reduction technique that, similar to POD and reduced basis functions, is an eigenvector method and is typically used to perform manifold transformation (Saul & Roweis 2000). The idea relies on the geometric intuition that points which lie close together on a manifold may be in a locally linear patch, wherein the patch can be described by linear coefficients that allow each data point in the patch to be reconstructed by its neighbors. These weights can be found by minimizing reconstruction errors $E(W) = \sum_i |x_i - \sum_j w_{ij}x_j|^2$ (the simple Euclidean distance between points), equivalently solving the least squares problem. A constraint is additionally placed such that $\sum_j w_{i,j} = 1$. By virtue of the posed constrained optimization problem, the vector of weights for each patch are invariant to rotation, scaling, or translation operations on the data. Thus, these weights can be used to find a valid low-dimensional casting of the original data which preserves their relatedness. This is done by minimizing a new cost function

$$\phi(Y) = \sum_i |y_i - \sum_j w_{ij}y_j|^2 \quad (3.10)$$

where the weights are fixed and the new coordinates Y in some reduced dimensional space are found. *Linear local tangent space alignment* (LLTSA) extends LLE by using

the tangent space of each local geometric patch on a high-dimensional manifold in order to define the tangent spaces for the low-dimensional casting (Zhang *et al.* 2007).

DYNAMIC MODE DECOMPOSITION Dynamic Mode Decomposition (DMD) is a dimensionality reduction technique which reduces a dynamic system into a set of weighted basis functions which are associated with fixed phase/oscillation modes and growth/decay rates in time (Schmid 2010). In linear systems, DMD modes are *composition operators* (normal modes) or Koopman operators (Bruce *et al.* 2019). Explicitly, DMD is tasked with recovering the eigenfunctions of a linear map A , such that $v_{i+1} = Av_i$. To discover A , single value decomposition (SVD) or similar approaches can be applied to a series of snapshots in time for a given system. The advantage of DMD as opposed to other reduction techniques is that it can explicitly represent temporal data; however, it also means that DMD is a less stable methodology since there are few constraints on the computed embedding space (e.g., orthogonality). There are several extended methods associated with finding the DMD basis of a given system (Chen *et al.* 2012): optimized DMD (to reduce sensitivity to noise), optimal mode decomposition (sets the rank of the decomposition), exact DMD (pairwise snapshots), sparsity promoting DMD, multi-resolution DMD, extended DMD (more explicit connection to the Koopman operator), dynamic distribution decomposition (finds the forward transfer operator).

KOOPMAN OPERATORS Koopman operators can be thought of as a set of weighted functions, and are adjoint to the transfer operator (forward simulation in dynamical systems). With respect to dynamical systems, the discovery of Koopman operators (or at least approximation of such) is typically easier than approximation of Lyapunov functions (a “true” composition operator for a system), and so is used as a stand-in. The fundamental expression of the Koopman operator is $\mathcal{K}(g) = g \circ f$ where g is the output map and f is a vector map; the formulation directly states that the Koopman

operator is a linear operator on an infinite-dimensional space of observables. A good overview of the Koopman operator is provided in detail in Bruce *et al.* 2019.

3.2 INVERSE PROBLEMS IN ENVIRONMENTAL SCIENCE

An inverse problem is posed when (possibly hidden) model parameters need to be recovered from (possibly noisy and indirect) observations. For instance, from observations of particulate concentration in the atmosphere, the source of those particulates could be recovered by inverting an advection-diffusion model. In environmental sciences, solving inverse problems from field data can be difficult for several reasons (Arridge *et al.* 2019):

1. The unknown variables and the observables may have different dimensionality.
2. The data may be an incomplete snapshot of the state of a domain of interest.
3. The map between the data and the state space may be rank deficient (uniqueness may not be guaranteed).
4. Observations may be noisy or corrupted.
5. Nonlinear systems are inherently difficult to work with.
6. Most inverse problems are ill-posed; noise in the data can lead to large errors in the model parameter.

Let an inverse problem take the form $y = \mathcal{A}(\theta) + \epsilon$, where $y \in Y$ is the measured data, $\theta \in \Theta$ is a set of model parameters, ϵ is observational noise, and $\mathcal{A} : \Theta \rightarrow Y$ is the *forward operator* which maps the parameter space to the observational space. Examples of \mathcal{A} in environmental science could be the advection-diffusion equations, the Navier-Stokes equations (Euler 1757; Navier 1822; Stokes *et al.* 1851), or a model of buoyant plume rise (Lavelle *et al.* 2013; Speer & Rona 1989).

One method to approach solving inverse problems is to use “knowledge-driven” techniques to place some conditions on the form and quality of the data in order to guarantee a unique solution at a desired accuracy. Regularization methods are a particularly pervasive strategy in many fields (often used in machine learning in the context of preventing overfitting (Srivastava *et al.* 2014)) and can either be explicitly or implicitly applied. In explicit regularization, a term is added to the optimization problem in the form of prior information, penalties, or constraints, which act to impose a unique solution (Benning & Burger 2018; Engl *et al.* 1996; Iglesias *et al.* 2013). Implicit regularization, common in machine learning, takes advantage of training techniques like stochastic gradient descent (Amari 1993; Bottou 2010) and optimization characteristics like epochs or training iterations, to control fitting. A challenge with solving inverse problems with regularization is the requirement to have access to a well-defined forward model. While many environmental models are useful, they are always approximations. Suspending this challenge, the most precise environmental models are typically systems of time-dependent PDEs, which can be computationally expensive to run—approaching intractability for even small systems of fluid equations, for example.

Another method for solving inverse problems is to use fully “data-driven” techniques which learn parameters for a generic representation such that the trained parameters and representation together have useful predictive power over the space of observations. Machine learning methods (e.g., Blanchard & Sapsis 2019b; Chen *et al.* 2019; Follmann & Rosa Jr 2019; Lu *et al.* 2020; Pathak *et al.* 2018) and model order reduction (Section 3.1) or latent space transformations (Bigoni *et al.* 2019; Spantini *et al.* 2018) are common techniques. Transferability/generalizability, interpretation, and data-efficiency are key challenges to adopting data-driven methods for scientific settings in which the trained model may be a desirable scientific product or tool for planning future missions, or in which limited field data is available.

This section will provide an overview of two additional approaches to inverse methods which attempt to blend knowledge-based and data-based techniques. In *Bayesian inference* methods, a notion of uncertainty is used to provide relative estimates of probability over a set of inverse solutions, useful for when data is noisy, the model may be imperfect, and other regularization methods may not be well-suited (Stuart 2010). In *scientific machine learning*, explicit computation of solutions is attempted with hybrid learning-knowledge frameworks, in which knowledge is encoded as a layer to a data-driven process (Baker *et al.* 2019b).

3.2.1 BAYESIAN INFERENCE TECHNIQUES

Bayes' Theorem (Bayes 1763) describes the probability of an event given prior knowledge, evidence, or observations that may be related to the event:

$$\Pi(\theta|y) = \frac{\Pi(y|\theta)\Pi(\theta)}{\Pi(y)} \quad (3.11)$$

where the *posterior distribution* of a set of parameters (event) θ given a data set y is proportional to the *likelihood* of the data given the parameters and the *prior distribution* on the parameters. In practice, exactly solving Bayes' Theorem in inference frameworks is computationally intractable, as it requires computing the *marginal distribution* over the data (the denominator in Eq. (3.11)), which requires an exponential number of computations in the number of latent parameters. Instead, approximate techniques are employed.

VARIATIONAL BAYESIAN INFERENCE Variational Bayesian Inference (often stylized as Variational Bayes) approximates the posterior distribution with a well-behaved function class (Bishop 2006; Wainwright, Jordan, *et al.* 2008), $q^*(\theta) \approx \Pi(\theta|y)$. To identify $q^*(\theta)$

from the set of all Q in the class, an optimization problem over some distance measure f is performed:

$$q^*(\theta) = \arg \min_{q \in Q} f(q(\cdot), \Pi(\cdot|y)). \quad (3.12)$$

The Kullback-Leibler (KL) divergence (Kullback & Leibler 1951), $\text{KL}(\cdot||\cdot)$, is a common choice for the distance metric as it has empirically good performance (Bishop 2006) and leads to a convenient simplification for the optimization problem:

$$\text{KL}(q||\Pi(\cdot|y)) = \log \Pi(y) - \int_{\Theta} q(\theta) \log \frac{\Pi(\theta)\Pi(y|\theta)}{q(\theta)} d\theta \quad (3.13)$$

$$q^*(\theta) = \operatorname{argmax}_{q \in Q} \int_{\Theta} q(\theta) \log \frac{\Pi(\theta)\Pi(y|\theta)}{q(\theta)} d\theta \quad (3.14)$$

where the *evidence lower bound* (ELBO) of the KL divergence can be used in the optimization; this is advantageous as the ELBO only contains well-defined aspects of the model. Choosing a class of well-behaved distributions, Q is another important choice in this methodology. A common choice is the mean field approximation, $Q = \{q : q(\theta) = \prod_{i=1}^n q_i(\theta_i)\}$, which admits low-dimensional representations and guarantees that the set of distributions over parameters θ factorizes. Coordinate ascent approaches can be used to solve the optimization problem with this class of distributions (Wainwright, Jordan, *et al.* 2008). Stochastic variational inference (SVI) (Hoffman *et al.* 2013) and automatic differentiation variational inference (ADVI) (Kucukelbir *et al.* 2017) are extensions of vanilla variation inference that leverage assumptions of conjugacy or differentiable properties to accelerate optimization.

MONTE CARLO METHODS Monte Carlo (MC) methods make use of the law of large numbers and simulation to approximate the posterior $\Pi(\theta|y)$ instead of performing an

optimization over analytic functions (MacKay 1998). Defining a *proposal density* $q(y)$, which is a simplification of the true density $\Pi(y)$, samples from $q(y)$ are drawn and estimators Ψ of a function (simulator) $\psi(\cdot)$ are computed. In general, MC methods require that the form of $\Pi(y)$ is known (and can be evaluated to within a multiplicative constant), but must be approximated by $q(y)$ because it may be difficult to draw samples from directly (e.g., too high-dimensional, not known in an analytic form). One of the most straightforward MC methods is importance sampling (Glynn & Iglehart 1989):

1. Draw \mathbf{x}_1, \dots, x_N i.i.d. samples from $q(\cdot)$.
2. Calculate weight $w_i = \Pi(x_i)/q(x_i)$.
3. Calculate estimate $\Phi = \sum_N w_i \phi(x_i) / \sum_N w_i$.

Rejection sampling (MacKay 1998), another MC sampler, encodes the notion that $q(y)$ may not necessarily align well with $\Pi(y)$. An acceptance criteria for sample x_i is defined with the rule $\Pi(x_i) > u$ where u is a draw from a uniform distribution with bounds $[0, q(x_i)]$. Even with this approach, it is generally required that $q(y)$ lie near the form of $\Pi(y)$ for MC methods; however, in large complex systems it may be difficult to define a single density that captures all of the characteristics of the true underlying distribution. Markov Chain MC (MCMC) methods directly address this issue by drawing new samples x' using a proposal density which is informed by the state of the previous sample $x^{(t)}$, $q(x', x^{(t)})$. In Metropolis-Hastings MCMC (Liu 1996; Metropolis *et al.* 1953), an acceptance ratio is used to transition between samples:

$$a = \frac{\Pi(x')q(x^{(t)}; x')}{\Pi(x^{(t)})q(x'; x^{(t)})} \quad (3.15)$$

in which if $a \geq 1$ the new sample x' is accepted, and $x^{(t+1)} = x'$; otherwise the sample is rejected, a new x' is proposed, and $x^{(t+1)} = x^{(t)}$. Other MCMC sampling methods, like Gibbs (MacKay 1998), Reversible-Jump (Green 1995), and Hamiltonian (Neal *et al.*

2011) use different acceptance ratios of special forms of $q(y)$ in order to improve the convergence characteristics, flexibility, and speed of Metropolis-Hastings. In all MCMC samplers, since each new sample relies on the previous accepted sample, the chain of samples that are accepted must be “burned-in” before a chain of virtually independent samples can be generated and used to compute estimates of Φ . For a sufficiently large number of samples, MC and MCMC methods are guaranteed to converge to the true estimator of the posterior (MacKay 1998).

3.2.2 BAYESIAN REPRESENTATIONS

In the most notional form, Bayesian representations are any (algorithmic) frameworks which can be used to exploit Bayes’ Theorem and related computational approximations by virtue of its form.

GRAPHICAL MODELS Probabilistic graphical models (PGMs) exploit the conditional independence structure of the latent parameters in θ in order to represent complex relationships between those parameters during inference. Bayesian networks (Aguilera *et al.* 2011; Arora *et al.* 2017; Ghahramani 2001) can be defined as an acyclic graph $\mathcal{G} = (V, E)$ where vertices V represent random variables (the latent parameters), and directed edges represent dependencies between two variables indexed $(i, j) \in E$ with i as a parent to j . The joint probability of V in the graph is the product of all conditional probabilities $\Pi(\theta_j | \text{parents}(\theta_j))$. In highly connected or otherwise complex networks, exact inference may be intractable; in these cases variational and MCMC techniques can be used.

PARAMETRIC MODELS Parametric models are a data-driven technique that identify a set of parameters for inference, and then “fit” those parameters to a portion of data known as a training set. The fitting procedure may be solving an inverse problem over

some analytic or probabilistic forward model that converts the set of parameters to the space in which data is available (Puonti *et al.* 2016), or using Expectation-Maximization (EM) to iteratively find local maxima for estimates of the maximum *a posteriori* estimator for parameters (Moon 1996). A test set of data is used to assess the parametric model’s accuracy in a process known as cross-validation. Using accuracy on a withheld test set as a metric, different numbers or types of parameters can be designed for desired performance. One form of parametric model, finite mixture models (Figueiredo & Jain 2002), in the state-of-the-art can automatically tune for number of parameters.

NONPARAMETRIC MODELS In contrast to parametric models, nonparametric models are *infinite* mixtures, allowing for more expressivity and avoiding the need to determine *a priori* the number of parameters to learn. Dirichlet processes (Ferguson 1973), Chinese Restaurant processes (Griffiths *et al.* 2003), and Gaussian processes (Rasmussen & Williams 2004) are all types of nonparametric model.

The latter, GPs, have seen considerable widespread adoption in environmental science and robotic sampling (e.g., Cahill *et al.* 2015; Flaspohler *et al.* 2019; Guestrin *et al.* 2005; Kleiber *et al.* 2012; Krause *et al.* 2008; Luo & Sycara 2018; Ma *et al.* 2017; Marchant *et al.* 2014; Ouyang *et al.* 2014; Srinivas *et al.* 2012; Wan & Sapsis 2017). Informally, a GP is used to represent a distribution over functions. Formally, Rasmussen & Williams 2004 defines a GP as *a collection of random variables, any finite number of which have a joint Gaussian distribution*. Functionally, for modeling the distribution of some environmental phenomenon, let an inference target be represented as a d -dimensional compact set $\mathbb{X} \subset \mathbb{R}^d$ and the unknown underlying distribution be m -dimensional continuous function $f : \mathbb{X} \longrightarrow \mathbb{R}^m$. Samples of f can be drawn in a location \mathbf{x} with a noisy sensor $y = f(\mathbf{x}) + \eta$ where $\eta \sim \mathcal{N}(0, \sigma_n^2)$ is normally distributed sensor noise. A GP is fully specified by a mean function $\mu(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})]$ and covariance (or kernel) function $\kappa(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - \mu(\mathbf{x}))(f(\mathbf{x}') - \mu(\mathbf{x}'))]$. With a history of observations,

$\mathcal{D}_t = \{\mathbf{x}_i, y_i\}_{i=0}^D$ of D observations at time t , the posterior belief $\Pi(\mathbf{x}'|\mathcal{D}_t)$ at a new location $\mathbf{x}' \in \mathbb{X}$ is:

$$\Pi(\mathbf{x}'|\mathcal{D}_t) \sim \mathcal{N}(\mu_t(\mathbf{x}'), \sigma_t^2(\mathbf{x}')) \quad (3.16)$$

$$\mu_t(\mathbf{x}') = \kappa_t(\mathbf{x}')^T (\mathbf{K}_t + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y}_t \quad (3.17)$$

$$\sigma_t^2(\mathbf{x}') = \kappa(\mathbf{x}', \mathbf{x}') - \kappa_t(\mathbf{x}')^T (\mathbf{K}_t + \sigma_n^2 \mathbf{I})^{-1} \kappa_t(\mathbf{x}') \quad (3.18)$$

where $\mathbf{y}_t = [y_0, \dots, y_{D-1}]^T$, \mathbf{K}_t is the positive definite kernel matrix with $\mathbf{K}_t[i, j] = \kappa(\mathbf{x}_i, \mathbf{x}_j) \forall \mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_t$ and $\kappa_t(\mathbf{x}') = [\kappa(\mathbf{x}_0, \mathbf{x}'), \dots, \kappa(\mathbf{x}_{D-1}, \mathbf{x}')]^T$. Typically, $\mu(\mathbf{x})$ is selected to be the zero function, and the kernel function is primarily used to encode the relationship between features in the environment. Training the hyperparameters of a kernel function can be employed using, e.g., Bergstra *et al.* 2011.

For spatiotemporal distributions, notable challenges remain in using GP models. Several kernel functions have been formulated that can model stationary and time-varying distributions (Chen *et al.* 2022; Garg *et al.* 2012; Raissi *et al.* 2018; Singh *et al.* 2010), however it is assumed that there is considerable access to data for training the hyperparameters for these kernels and they have limited predictive capability, which is a core capability necessary for informative planning in natural environments. Recent work embedding numerical models into GP covariance kernels (Raissi *et al.* 2018) and utilizing learned latent spaces (Kingravi *et al.* 2016; Al-Shedivat *et al.* 2017; Sun *et al.* 2018; Wan & Sapsis 2017; Wang *et al.* 2022; Whitman & Chowdhary 2017; Wilson *et al.* 2016a; Wilson *et al.* 2016b; You *et al.* 2017) are promising areas for future adoption of GPs for spatiotemporal settings, but have yet to be demonstrated effectively on realistic data that could be collected in the field (Ober *et al.* 2021).

3.2.3 SCIENTIFIC MACHINE LEARNING

Under the moniker of scientific machine learning (SML), a growing field of research aims to learn governing equations in the sciences from data by combining traditional numerical techniques and state of the art machine learning and probabilistic frameworks. The “combination” of techniques may be as straightforward as adding a PDE or ODE solver as a layer in a neural network (Pakravan *et al.* 2021) or in the loss function at training time (Raissi *et al.* 2019), to as nuanced as defining kernels over pairs of states using finite element discretizations (Raissi *et al.* 2018). Physics-informed neural networks (PINNs) (Raissi *et al.* 2019; Tartakovsky *et al.* 2018) explicitly apply constraints from numerical modeling (e.g., constitutive properties, invariants) to guide the learning problem, and other techniques like PDE-Net (Long *et al.* 2017) structure memory flows through convolution kernels to emulate finite difference operations. Regardless of the technique employed, SML methods are particularly powerful for spatiotemporal modeling because they inject structure to the learning problem, which can selectively control the learned feature space or more quickly converge to reasonable state predictions by imposing constraints on valid next-states according to first principles. One challenge for all of these methods is overcoming high-dimensionality in the state and parameter spaces. Recent work on developing scalable simulations for large environmental systems show that good latent embedding spaces yield statistically equivalent state predictions to the true state of the world (Baddoo *et al.* 2022; Baddoo *et al.* 2021; Mardt *et al.* 2020; Qian *et al.* 2020). For example, Compressed ConvLSTM (Mohan *et al.* 2019) “compresses” large input data with a convolutional autoencoder into a low-dimensional subspace which is passed to a convLSTM and decoded to yield a state prediction. Preliminary results show that the network is able to predict 3D atmospheric turbulence with considerable statistical accuracy.

Although it may not be considered SML canonically, there are many learning frameworks which attempt to recover PDEs from data by using neural networks to function-fit measurements (Berg & Nyström 2019; Kaiser *et al.* 2018) or simultaneously select a model from a library of governing equations and train its parameters (Rudy *et al.* 2017; Sun *et al.* 2020). Reservoir computing networks have seen increased popularity in the sciences because of their ability to handle high-dimensional state spaces gracefully and ability to learn characteristics, like Lyapunov exponent, useful for characterizing traditional PDEs (Pathak *et al.* 2017). Learning stability characteristics about data, rather than performing state prediction, has been demonstrated by other networks which applied constraints on learned latent subspaces (Blanchard & Sapsis 2019b).

Like GPs with sophisticated kernels, SML systems have yet to be rigorously applied to field data for environmental recovery, or in extremely partially-observable domains, let alone for planning sampling trajectories. However, some recent methods in control theory (Chee *et al.* 2022; Gan *et al.* 2020; Jiahao *et al.* n.d.) have shown promise that some SML representations could extend to practical-time operations for robotic tasks, and are worth considering as the field matures in future work.

3.3 ENVIRONMENTAL SENSING AS ADAPTIVE SAMPLING

Adaptive sampling is the art of collecting samples of some *a priori* unknown function to strategically assist in the recovery of the unknown function or assist in performing a specific task by analysis of a previous history of observations. Environmental sensing—the act of collecting observations of some natural phenomenon—can be framed as an adaptive sampling problem when collected observations are used to inform changes to a sensing plan. To determine how measurements should modify behaviors, a planning model must be specified. A planning model defines agent dynamics, available actions, and the mission objective. Informative path planning (IPP) is an approach for approx-

imately solving adaptive sampling problems modeled as *sequential decisions* in which actions are taken, executed, and evaluated over several iterations. Sequential decision-making is considered Markovian when an action selection is conditionally independent of previous history when using an updated belief representation (with respect to both environmental and agent states). Markov decision processes (Bellman 1957; Howard 1960) (MDPs) are a useful model for robotic planning problems, and 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 $s' \in \mathcal{S}$; $T(s, a, s') = \mathbf{Pr}(\mathcal{S}_{t+1} = s' | \mathcal{S}_t = s, \mathcal{A}_t = a)$. 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):

$$\pi^* = \operatorname{argmax}_{\pi} \mathbb{E} \left[\sum_{t=0}^{\inf} \gamma^t R(s_t, a_t) | s_0, \pi \right] \quad (3.19)$$

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 (Bellman 1957):

$$\begin{aligned} V_{t+1}^*(s) &\leftarrow \max_{a \in \mathcal{A}} \left[\sum_{s' \in \mathcal{S}} T(s, a, s') (R(s, a, s') + \gamma V_t(s')) \right] \\ \pi^*(s) &= \operatorname{argmax}_{a \in \mathcal{A}} \left[\sum_{s' \in \mathcal{S}} T(s, a, s') (R(s, a, s') + \gamma V^*(s')) \right]. \end{aligned} \quad (3.20)$$

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.

3.3.1 PARTIALLY OBSERVABLE MARKOV DECISION PROCESSES

In many field operations, measurements that can be collected of an environment or the robot's state are only *partial* observations. Partially-observable Markov decision processes (Kaelbling *et al.* 1998) (POMDPs) extend 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 previously defined, 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 ; $b_0 = \mathbf{Pr}(\mathcal{S}_0 = s)$

In partially-observable domains, the state of the world is uncertain, and so that should mean that the decision process is no longer Markovian, since 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:

$$V_h^\pi(b_t) = \mathbb{E}[R(s_t, \pi(b_t))] + \gamma \sum_{z \in \mathcal{Z}} V_{h-1}^\pi(b_{t+1}^{\pi(b_t), z}) \mathbf{Pr}(z | b_t, \pi(b_t)), \quad (3.21)$$

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 : $\pi_h^* = \operatorname{argmax}_{\pi \in \Pi} V_h^\pi(b_t)$.

3.3.2 DECISION-MAKING UNDER UNCERTAINTY

In general, Eq. (3.21) is difficult or intractable to compute in large or continuous state and observation spaces¹. Thus, approximate solvers are necessary to recover the optimal policy for an agent to execute. In the most broad sense, a solver can be characterized as either being *online* or *offline*, which describes at what point in a mission a plan may be generated.

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

¹The same holds for large or continuous-valued MDPs.

offline planners perform coverage or monitoring tasks (Nam *et al.* 2016; Nikolos *et al.* 2003) 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 (Roa *et al.* 2012). Methods for computing offline plans can include formal optimization, action simulation and scoring (such as in Monte Carlo tree search, particle filters or reinforcement learning) (Arora *et al.* 2017; Raja & Pugazhenthi 2012; Yu *et al.* 2021), or classical search (e.g., probabilistic roadmaps) (Karaman & Frazzoli 2011; Karaman *et al.* 2011).

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 (Esposito & Kumar 2002; Majumdar & Tedrake 2012). Online planners can be either *myopic* (Edwards *et al.* 2005; Vergassola *et al.* 2007) or *nonmyopic* (Arora *et al.* 2017; Browne *et al.* 2012; Kurniawati *et al.* 2008; Lim *et al.* 2016; Meliou *et al.* 2007; Singh *et al.* 2009; Somani *et al.* 2013; Sunberg & Kochenderfer 2018), in reference to how far a horizon is considered in a plan in order to choose an action to take.

3.3.3 INFORMATION-THEORETIC REWARDS

In an MDP or POMDP, the reward function serves to encode the scientific objective of a mission. For strategic sample collection it is typically useful to consider the *information content* of potential observations in order to elicit *explore-exploit* behaviors of the robot. Explore-exploit is a paradigm that describes the phenomenon of *exploring* when little

knowledge is held about an environment, and then transitioning to *exploiting* collected knowledge in order to perform a task. Getting the balance right between exploration and exploitation is a perennial challenge in adaptive sampling, and the choice of the reward function can have further implications for the performance of a planning scheme. For instance, submodular reward functions (i.e., diminishing returns) allow for even greedy-myopic online strategies to have bounded performance (Horel 2016).

Several fields of research have proposed information measures for policy development in robotic sampling tasks. Optimal experimental design (Fedorov 2013) proposes several “criteria” for variance reduction over inference targets:

- *A*-optimal: Minimizes the trace of the inverse covariance matrix (e.g., Carrillo *et al.* 2015b; Kollar & Roy 2008; Sim & Roy 2005)
- *D*-optimal: Minimizes the determinant of the covariance matrix (e.g., Carrillo *et al.* 2015b; Joshi & Boyd 2008; Kollar & Roy 2008)
- *E*-optimal: Maximizes the smallest eigenvalue of the covariance matrix (e.g., Carrillo *et al.* 2015b)
- *V*-optimal: Minimize the average prediction variance (e.g., Cohn 1994)

Optimal experimental design additionally suggests several “soft” measures of information content, including Shannon’s entropy (Shannon & Weaver 1998), conditional entropy, and mutual information all of which have been widely used in robotic simultaneous localization and mapping (SLAM) (Bourgault *et al.* 2002; Burgard *et al.* 1997; Carrillo *et al.* 2015a; Valencia & Andrade-Cetto 2018), sensor placement (Guestrin *et al.* 2005; Krause *et al.* 2008; Papadimitriou *et al.* 2000), and optimal navigation (Daniel *et al.* 2012).

In Bayesian optimization contexts, several information-theoretic rewards are commonly used:

- Upper-Confidence Bound (UCB) (Agrawal 1995; Auer 2002; Snoek *et al.* 2012) of the form $R_{\text{UCB}} = \mu(\mathbf{x}) + \sqrt{\beta}\sigma(\mathbf{x})$ which is the sum of predictive mean μ and variance σ at queries \mathbf{x} . UCB is submodular (Nemhauser *et al.* 1978).
- Probability of Improvement (PI) (Kushner 1964; Snoek *et al.* 2012); a probability measure of whether a query \mathbf{x} will be better than the current best measurement \mathbf{x}^* .
- Expected Improvement (EI) (Jones *et al.* 1998; Snoek *et al.* 2012); a measure of how much better a proposed query \mathbf{x} will be compared to the current best measurement \mathbf{x}^* .
- Predictive Entropy Search (PES) (Hennig & Schuler 2012; Hernández-Lobato *et al.* 2014); a measure of the conditional entropy between a query \mathbf{x} and a predicted optimizer of a distribution $f(\cdot)$, \mathbf{x}^* .

UCB is particularly well-utilized in robotic sampling contexts because of its submodularity property. Srinivas *et al.* 2012 provides a detailed analysis of UCB-based reward functions for use in environments represented by GPs, ultimately demonstrating a bound on *regret* for some selection of belief, 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 (Garivier & Cappé 2011; Srinivas *et al.* 2012; Sun *et al.* 2017).

3.4 VENT PROSPECTING, ODOR LOCALIZATION, AND FRONT TRACKING

In robotics, *plume hunting* is equivalently referred to as vent prospecting, odor mapping, odor localization, source localization, and source seeking. In these works, it is generally assumed that the source *location* is unknown, and through partial observations of emitted gas/odor/plume, the source can be physically discovered (as in, the robot can find and navigate to it) using techniques that can be divided broadly into biologically-inspired heuristic search (e.g., Chen & Huang 2019; Reddy *et al.* 2022) and adaptive informative path planning (e.g., Jakuba 2007; Salam & Hsieh 2019).

Biologically-inspired or heuristic techniques draw (varying-levels of) inspiration from animal or insect behavior in olfactory settings. Such techniques typically include gradient-based algorithms like chemotaxis (Morse *et al.* 1998), or bio-inspired algorithms that directly mimic a particular animal (Edwards 2001; Grasso *et al.* 2000). These techniques are typically reactive and myopic, although they have been demonstrated to be relatively robust in open-world settings. In contrast, adaptive informative path planning can be nonmyopic, and typically attempts to embed knowledge (either heuristically or rigorously) about flow-fields (i.e., advection and diffusion) to assist in plume localization. Such techniques live on a spectrum, from algorithms that resemble biologically-inspired methods, like infotaxis (Vergassola *et al.* 2007), to methods that use model order reduction techniques (like POD) to encode complex numerical models and elucidate spatiotemporal structures in complex data (Peng *et al.* 2014; Salam & Hsieh 2019).

Three field studies have demonstrated the promise of autonomy tools and planners for intelligent autonomous vent localization in the deep sea with the aid of simulation and post-expedition analysis. In Jakuba 2007, a probabilistic occupancy-grid representation is formulated which uses observations of opportunity from a deep-sea vehicle to

estimate the location of a vent. These maps were not tied into the autonomy in any-way during these trials. In a follow-up study presented in Ferri *et al.* 2010, an adaptive surveying strategy was tested using data gathered by a deep sea vehicle (the autonomy was not tested at sea, but verified with field data), leveraging a similar occupancy style representation and allowing a vehicle to place surveys strategically for information gathering. In Branch *et al.* 2020, a direct extension of Ferri et al., simulated hydrothermal expressions of the Juan de Fuca ridge were provided to a glider physically swimming in the Chesapeake Bay to enable “hardware in the loop” tests of a planning methodology that allowed the glider to selectively place finer and finer resolution surveys over an estimated vent location. In full simulation, works using adaptive heuristic planners (Pang 2010; Wang *et al.* 2020) have primarily dominated.

A complement to the vent localization problem is the *front tracking* problem (Chen & Huang 2019), which tasks an agent with collecting samples at a (possibly dynamic) boundary between two or more phenomena. An example of a front could be at the location a river dumps freshwater into a salty bay (McClimans 1988), or at the edge of a warm core ring, which are spun out by the Gulf Stream (Cushman-Roisin *et al.* 1985). Front tracking algorithms are largely focused on classifying observations as being part of one or another water mass and using an estimated (or model-based) gradient between the classes to adapt robot behavior to gather samples at the boundary. Field-deployed examples of the front-tracking problem include using underwater vehicles to target salinity gradients in the surface ocean (Belkin *et al.* 2018) and teams of underwater vehicles and surface vehicles to track ocean fronts (McCammon *et al.* 2021); several simulated studies looking at this problem with respect to chemical plumes (Li *et al.* 2014; Wang *et al.* 2019a) and river plumes (Teixeira *et al.* 2021) have also been demonstrated. While deep ocean plumes have been a motivating context for some studies, front tracking in the deep ocean with robotic technologies has not yet been demonstrated.

It is worth noting that there are two barriers that have contributed to the difficulty of performing autonomous studies in the deep sea for vent localization and front tracking. One is related strongly to the challenge of non-agency in most depth-capable AUVs; indeed, all of the field work discussed in this section with active autonomy was conducted using small gliders with a depth rating of no more than 100 m. The mismatch between the actual autonomous capabilities of the science fleet and the autonomy frameworks being developed is not necessarily bad, especially in thinking about the next decades of oceanographic research and the development of the science fleet in that time², but it does have an impact of the science that can be performed today. The second barrier is the difficulty of accessing the deep sea for non-oceanographers/scientists given how precious and expensive ship and AUV resources are, and the community knowledge necessary to run successful field operations on a ship. Growing interest and ability to invite remote-scientists to sea (i.e., scientists can participate actively in research cruises without being on the ship through high-bandwidth internet links) may extend access to a broader set of researchers in the future, that could possibly enable more opportunities for deploying deep-sea autonomy in the way presented in this thesis.

3.5 HYDROTHERMAL PLUMES

Understanding the physics of plumes is fundamental to interpreting observations gathered during a deep sea geochemical survey. Hydrothermal plumes are typically characterized as buoyancy-driven water masses. On formation at a vent site, emitted fluids are significantly less dense than background seawater (by virtue of being super-heated, with some add-on effects by changes in chemical composition). The less dense water mass rises rapidly in the water column, forming a buoyant stem. As a rule of thumb,

²For instance, the contemporary development of the NUI under-ice vehicle (German & Boetius 2019), or the Mesobot for studying the mid-ocean water column (the Twilight Zone) (Yoerger *et al.* 2018) are promising up-and-coming developments.

a buoyant stem grows in diameter about 1 m for every 10 m vertically traveled. Due to rapid cooling, turbulent mixing, and the natural stratification of ocean water, vent-derived fluids will reach a point of neutral-buoyancy with the background seawater. At this point, the plume forms a nonbuoyant or neutrally buoyant layer which spreads out across the isopycnal that describes the ocean layer of equivalent density. In the Atlantic basin, plume rise height is typically expected to be approximately 300-350 m; in the Pacific basin, this is 150-200 m (Speer & Rona 1989). From the neutrally-buoyant layer, metals, sediment, and other suspended particulates carried by the plume may drop out and be redeposited onto the seafloor, and any persisting chemicals diffused, reacted, or digested by microbes (Dick *et al.* 2013; Scholz *et al.* 2019).

Two general models which have been commonly incorporated in robotic source seeking literature include the Gaussian plume model (Green *et al.* 1980) and the Gaussian puff model (Ludwig *et al.* 1977). These models primarily describe the dispersion envelope of aerosols released as a plume from a coherent source in the atmosphere, modeling the concentration of those aerosols directly as a Gaussian around a plume centerline, which describes the path of the plume in space. These models have largely been used to simulate ground pollution characteristics of smokestack-like sources in open, unstratified environments, and typically assume that the advective crossflow dominates plume movement. In the deep sea, stratified environments are the norm, and buoyancy forces are the primary advective force of plume fluids³ with relatively weak crossflow. These non-trivial differences encourage turning to domain-specific plume models.

Hydrothermal plumes have been mathematically codified perhaps most famously by Morton *et al.* 1956 (MTT) as a system of conservative equations (here for a stratified fluid) in cylindrical coordinates (x, r) with the x -axis vertical with the vent source at the origin:

³Note that in this work the Boussinesq approximation (Van den Bremer & Hunt 2010) is assumed

$$\text{Volume: } \frac{d}{dx}(b^2 u) = 2b\alpha u \quad (3.22)$$

$$\text{Momentum: } \frac{d}{dx}(b^2 u^2) = 2b^2 g \frac{\rho_o - \rho}{\rho_1} \quad (3.23)$$

$$\text{Density deficiency: } \frac{d}{dx}(b^2 u g \frac{\rho_o - \rho}{\rho_1}) = 2b^2 u \frac{g}{\rho_1} \frac{d\rho_o}{dx} \quad (3.24)$$

where α is a proportionality coefficient which represents gross mixing (or entrainment) that occurs at the edge of a plume, $b = b(x)$ is the (symmetric) radius of the plume, $\rho = \rho(x, r)$ is density inside the plume, $\rho_o = \rho_o(x)$ is density outside of the plume, ρ_1 is some reference density such that $\rho_o(0) = \rho_1$, g is acceleration due to gravity, and $u = u(x, r)$ is vertical velocity. These equations have been equivalently expressed in terms of mass, salt, heat, and momentum conservation by Speer & Rona 1989 who usefully decomposes density into components of salinity and temperature that can be directly observed by scientific instruments.

Variations on the time-averaged MTT model, for instance models which consider crossflow (Tohidi & Kaye 2016), are numerous. In addition to time-averaged models, sophisticated simulators of hydrothermal plumes which model time-varying turbulent dynamics are available (Lavelle *et al.* 2013). These models also use conservation properties (momentum, buoyancy, volume), but couch the quantities relative to fluid flow (Navier-Stokes), flow-field non-divergence, boundary reflections, and multiple scales of mixing (viscous, turbulent, diffusive). Solving a highly resolved model is generally computationally expensive; for relatively modest simulated environments of tens of meters on each volumetric axis, several days of computation of a high-performance computing node may yield only a few hours of simulation.

3.5.1 TRACERS AND INSTRUMENTATION

Physically manifested hydrothermal plumes are turbulent, variably warm, and contain particulates and chemicals. As a plume rises and advects, it entrains (mixes) background seawater, diluting a plume expression several orders of magnitude from the originating vent. To identify a plume from observations, consideration of how different *tracers* may manifest as a plume evolves is important. Tracers can be either conservative or non-conservative. Conservative tracers are only impacted by physical advection and diffusion; non-conservative tracers can additionally react, decay, or be consumed in the water column by other processes.

Temperature and salinity are convenient conservative tracers because standard oceanographic equipment can directly observe them (e.g., with a CTD probe). A challenge with interpreting temperature and salinity, however, is that an ocean basin is stratified, and this stratification must be considered when attempting to identify anomalous water masses. For instance, in the Pacific Ocean, neutrally-buoyant plume intrusions tend to be relatively warm compared to ambient water at the same density, whereas intrusions in the Atlantic tend to be cold (Speer & Rona 1989).

Chemicals are generally non-conservative tracers, although the rate of decay or reaction can vary significantly for a given environment and chemical species. Methane is an energetic compound, readily consumed (oxidized) by e.g., microbes (Petersen & Dubilier 2009). Oxygen is generally depleted in hydrothermal fluids, and can be produced and consumed in the deep ocean by microbial activities (Johnson *et al.* 1986; Smith Jr 1985). Turbidity, although closely related to chemical distributions in a hydrothermal plume, can be treated as a conservative tracer. Oxidation reduction-potential (ORP), a measure of reactivity of a water species, or equivalently the relative “age” of a water sample, can similarly be treated as though it were a conservative measurement.

Complicating the measurement of tracers in the deep ocean is the mechanism of the oceanographic sensors that are available. While CTD, ORP, and turbidity measurements can be considered nearly instantaneous, most chemical sensors have response time properties that ought to be considered. For instance, *in situ* methane sensors for the deep sea typically use passive or active equilibration over a membrane to separate gaseous species from water samples for analysis via laser spectrometer. This process can take minutes to hours to reach a steady state, which can prove a challenge in post-processing when deployed on a moving AUV or in a turbulent plume. Often, time-correction is necessary (Bittig *et al.* 2018) which projects a “true” measurement given laboratory (or field) sensor characterization for any given observation. This is important to consider for real-time or post-mission analysis of chemistry data, as these measurements may not yield absolute concentration estimates that can be trivially accepted. This highlights the importance of close collaboration with scientists with familiarity on the working principles of their sensors to enable reasonable (practical-time) analysis of *in situ* analysis to enable informed mission planning at sea.

4 DISCOVERING HYDROTHERMALISM FROM AFAR

I had a blank canvas to fill with extraordinary possibilities, a fascinating jigsaw puzzle to piece together: mapping the world's vast hidden seafloor.

Marie Tharp

To track a spatiotemporal phenomenon requires first being able to sense it. There are two core challenges associated with perceiving hydrothermal plumes: existence (and availability) of technology and interpretation of heterogeneous data streams. With respect to deep-sea capable instrumentation, temperature, pressure, conductivity, and turbidity are all examples of quantities that can be near instantaneously measured by existing state-of-the-art *in situ* sensors. However, for many geochemical quantities such as dissolved greenhouse gases (e.g., carbon dioxide, methane), few (if any) commercial sensors with rapid response times suitable for use on a mobile platform exist. This limitation has had severe impact on the ability for scientists to study phenomena like hydrothermal plumes, as these entities may be difficult to identify only from temperature or conductivity anomalies, but are expected to be significantly geochemically distinctive from background seawater levels of the ocean (Jakuba 2007; Scholz *et al.* 2019). In this chapter, two experimental dissolved methane instruments are used in an ocean

trial at a hydrothermal basin. The utility of methane as a signal for the presence of plume waters is compared against other standard oceanographic equipment.

The second challenge, interpreting heterogeneous data streams, is a problem that arises when the quantity of interest is a conceptual entity, rather than an absolute one. Concretely; temperature can be directly observed with a single instrument, but “plumes” cannot be directly sensed as they are by definition an aggregation of properties distinct from background seawater. So, to observe “plumes” requires interpreting data from multiple heterogeneous sensors in order to identify which robot locations observed plume-derived fluids, and which did not. Heterogeneity in this case refers to the different operating principles and observable quantities that are measured by a suite of scientific sensors. As these sensors may respond to the environment at different time scales, have different sensitivities, and measure quantities which may physically manifest themselves differently in unique spatiotemporal regions of a plume (and may be unique to each plume that is surveyed), it is not straightforward to universally filter these data streams for plume detections. To this end, this chapter presents several methods for detecting “change-points” in data streams that can be used to indicate anomalous features in a data stream unique to hydrothermalism, and which can be broadly applied to many different types of field settings.

The content that proceeds from this point is directly adapted from Preston *et al.* 2022b, which was published in Frontiers Earth Science in 2022. The supplemental information for this publication is also reproduced as Appendix A in this thesis.

4.1 INTRODUCTION

Detecting and characterizing seafloor hydrothermal vents is critical in understanding the fundamental interactions among the geochemical and biological processes on the seafloor, and the fluxes that these processes cause to and from the deep ocean. Since

the first discovery of deep sea hydrothermalism in 1977 (Corliss *et al.* 1979), hundreds of hydrothermal venting sites have been discovered and analyzed (Beaulieu *et al.* 2015). These studies reveal that hydrothermal vents play a major role in ocean-scale elemental and micronutrient budgets (Le Bris *et al.* 2019; Resing *et al.* 2015), serve as nutrient pumps to the deep ocean (Bell *et al.* 2017; Dick *et al.* 2013; Scholz *et al.* 2019; Vic *et al.* 2018), and sustain abundant and unique (e.g., chemosynthetic) forms of complex life (Georgieva *et al.* 2021; Grassle 1987). Hundreds more vent sites are hypothesized to exist and yet remain undiscovered in the deep ocean (Beaulieu *et al.* 2015), limiting efforts to constrain nutrient and energy budgets of the deep ocean, to assess the magmatic budget hypothesis which estimates the global stock of hydrothermal activity, and to understand these novel ecosystems.

Exhaustive search of the seafloor is an impractical method for discovering new vents due to the scale of the ocean environment. Instead, adaptive surveying strategies and novel sensing technologies can be combined to detect hydrothermalism far (over 1 km laterally) from the plume source using water column observations. Hydrothermal plumes form due to a density difference between background seawater and (often significantly) heated vent fluids. The resulting buoyant force creates a coherent rising stem from the vent (the buoyant stem) and a spreading cloud (the neutrally-buoyant layer) at an isopycnal, when the cooling, mixing, hydrothermally-derived fluids reach equivalent density to the ambient background (Morton *et al.* 1956; Speer & Rona 1989). The chemical composition of hydrothermal fluids differs greatly from that of background seawater and the plume-derived fluids near an active vent can be detected using most standard properties (i.e., temperature, salinity, chemical composition, turbidity). However, the spatial expression of the buoyant plume stem is typically no more than a few tens of square meters, making the buoyant stem difficult to localize on a survey. As emitted fluids travel further within the plume, the physically and chemically distinctive nature

of the hydrothermal water mass is rapidly diluted as the plume entrains background seawater. Throughout this advective evolution of the plume, reactive (non-conservative) tracers can be consumed or transformed. Thus, despite the neutrally buoyant layer having a spatial scale extending for several square kilometers, detecting these plume fluids requires innovation in sensing and data analysis.

In this chapter, we discuss the potential for water column-based hydrothermal plume discovery using standard sensing equipment (e.g., CTD, optode, transmissometer) in concert with two novel *in situ* methane instruments installed onboard an autonomous underwater vehicle (AUV) and a towed rosette. We present results from a field deployment at the northern Guaymas Basin in November 2021 and use these results to inform the planning of informative plume transects and the monitoring of real-time instrument responses. Both towed rosettes and AUVs are well-established tools for hydrothermal plume surveys. Rosettes deployed for hydrothermal plume hunting are typically used in either a vertical transect mode, or cast, performed at regularly spaced spatial waypoints along a ship transect, or a “towed” mode, in which the CTD is lowered and pulled through the water by the ship’s motion (Bennett *et al.* 2013; Chin *et al.* 1994). AUVs, by virtue of being untethered from the ship, have the ability to finely control location within the water volume, and can typically operate closer to the seafloor than a towed rosette. Standard sensors mounted on either a rosette or AUV can detect different forms of hydrothermalism. High turbidity several hundred meters from the seafloor may be indicative of a neutrally-buoyant plume generated by a black smoker, whereas changes in oxidation-reduction potential and clear waters near the seafloor may be indicative of diffuse flow. Analyzing these sensors individually and in combination can disambiguate these types of hydrothermalism and elucidate plume structure and characteristics of venting sources on the seafloor.

In 2021, an expedition aboard the R/V *Roger Revelle* (RR2107) with AUV *Sentry* and ROV *JASON*, offered a unique opportunity to examine the emission of hydrothermally derived fluids, their buoyant rise, as well as the evolution and fate of the neutrally-buoyant plume in the mid-water. Here, the results of a targeted lateral transect using chemical sensors mounted on AUV *Sentry* and a towed rosette are presented, in addition to the first field demonstration of novel *in situ* methane instruments. Fig. 4.1 illustrates the overall design of the transect experiment. The results show that methane acts as a reliable indicator of hydrothermal activity in the northern Guaymas Basin on a spatial scale of 1.5-3 km at 100-150 m altitude. Methane performed similarly to standard turbidity sensors in this trial (detection 2.2-3.3 km), more sensitively than oxidation reduction potential, and more clearly than temperature, salinity, and oxygen instruments which readily responded to physical mixing in background seawater. Additionally, the relationships between different sensing modalities are investigated using cross correlation and time-series regime identification, suggesting how these analyses could be used to assist in survey design for future exploratory missions.

4.2 MATERIALS AND METHODS

4.2.1 SITE DESCRIPTION

As introduced in Chapter 2, the Guaymas Basin is a mid-ocean ridge extensional spreading center system, with the unique characteristic of being heavily overlain with high amounts of organic-rich sediment. While the primary spreading center axis trends southwest to northeast, the axis of the spreading center in the more well-studied southern end does not extend linearly northeastward, with the northern end of the axis offset to the northwest. The subseafloor eruption and emplacement of lava into the heavy sediment overburden gives rise to a unique set of hydrothermal characteristics.

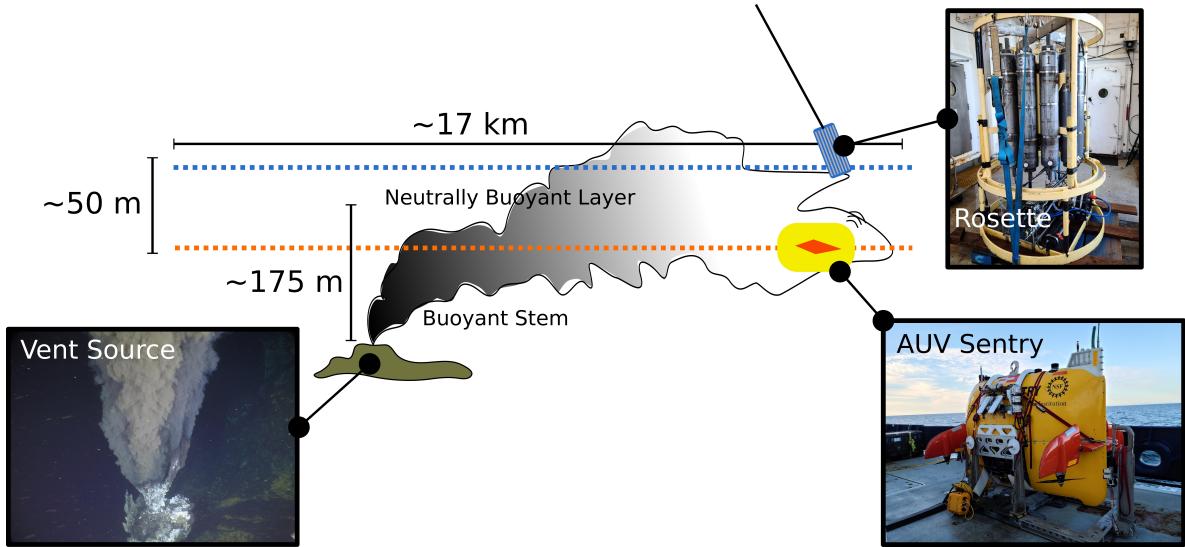


Figure 4.1: Overview of general transect design. Plumes generated by black smoking chimneys at an active hydrothermal ridge in the Northern Guaymas Basin (one example pictured here, taken with an arm mounted MISO camera by ROV *JASON* during RR2107) rise approximately 175 m in the water column and are advected and turbulently mixed with background seawater. AUV *Sentry* and a towed CTD rosette, both equipped with turbidity, oxygen, temperature, salinity, and methane probes, fly trajectories that aim to intersect the lower and upper neutrally buoyant plume layer, respectively. A comparison of the observations collected by both platforms is then used to demonstrate the efficacy of various sensors and algorithmic detection schemes.

Among these, the geochemical composition of the emergent fluids and volatiles is highly enriched in dissolved organic compounds, carbon dioxide (CO_2), hydrogen (H_2), ammonium (NH_4^+), and methane (CH_4) (Seewald *et al.* 1994; Von Damm *et al.* 1985). While the southern end of the basin has been the subject of a long history of geochemical and biological examination (Lonsdale & Becker 1985; Ondréas *et al.* 2018; Seewald *et al.* 1994; Teske *et al.* 2016; Von Damm *et al.* 1985), hydrothermal activity was only recently documented along the northern end of the basin at a 600 m long ridge located at a depth of 1850 m (Geilert *et al.* 2018; Soule *et al.* 2018). Several tall sulfide chimneys 10-25 m in height are located along the ridge, and emit fluids highly-enriched in CO_2 , H_2 , CH_4 among others (Fig. 4.2). The black smoker vents associated with these chimneys consist of clusters of tens of small ($<0.01 \text{ m}^2$) orifices, emitting turbid fluids heated

to over 340 °C, as observed during RR2107 by ROV *JASON*. In this work, the closest identified chimney to the transect trajectories at (27.407489 N, 111.389893 W) is used as a spatial reference point.

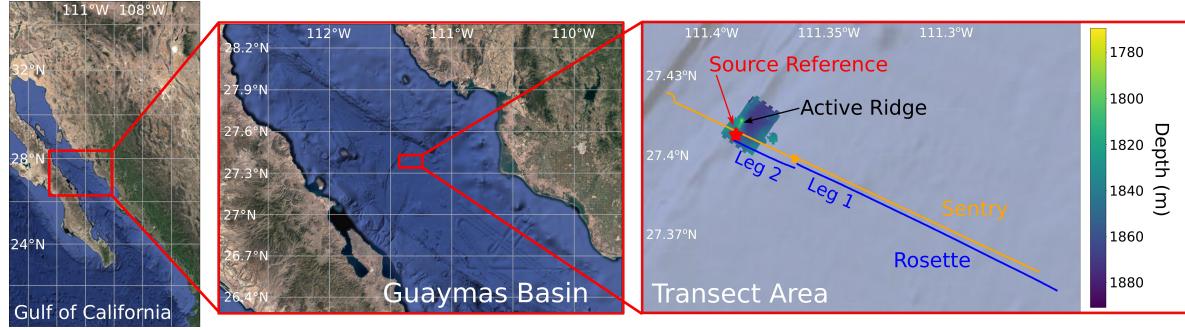


Figure 4.2: Map of the transect experiment extent. AUV *Sentry* and a towed rosette were used to perform coincident several kilometer long trajectories in the Northern Guaymas Basin. The rosette was redeployed mid-trajectory in order to empty the Niskin bottles onboard; this split the rosette trajectory into Leg 1 and Leg 2. The trajectories intersected a region of known hydrothermal activity in the northern basin; a bathymetric relief of this region is overlaid on the far right panel. The red star on the bathymetric relief marks the nearest point of identified hydrothermal activity (black smokers) relative to the trajectories (27.407489 N, 111.389893 W), and is used as a reference point in this work. Imagery is provided by the GoogleTiles API in Cartopy. The bathymetric relief is rendered using data collected by AUV *Sentry* during research cruise RR2107.

4.2.2 SAMPLING PLATFORMS AND INSTRUMENTS

During expedition RR2107, AUV *Sentry* and a towed rosette were deployed to perform a multi-kilometer transect. Two novel *in situ* methane instruments were deployed during the transect, one on *Sentry*, and the other on the towed rosette. Physical water samples collected by the Niskin bottles on the rosette were processed shipboard to measure both methane and ammonium content. To increase the total number of bottle samples that could be collected over the transect, the towed rosette was deployed and recovered twice; we will refer to the rosette transect before the first recovery as “Leg 1” and after re-deployment as “Leg 2.” AUV *Sentry* was placed in a holding pattern when the rosette

was on the ship deck to ensure that spatial measurements between the platforms were temporally comparable.

AUV *Sentry*

AUV *Sentry* executes pre-set trajectories (encoded as a set of waypoints) once underway. During this transect, a starting point at (27.345152 N, 111.253108 W) and ending point at (27.460812 N, 111.527694 W) were given, and a holding pattern was programmed to be executed when the rosette was on the ship deck for sample retrieval after Leg 1. This holding pattern was centered at (27.39592 N, 111.3674 W) and was a lawnmower (back and forth) pattern of approximate dimensions 225 m x 225 m with 15 m resolution. The standard scientific instrumentation deployed on *Sentry* include an oxygen optode (Aanderaa 4330F), an optical backscatter sensor or OBS (Seapoint Turbidity Meter), an oxidation-reduction potential sensor or ORP (NOAA), a CTD (SeaBird SBE49), and 7000 m rated pressure sensor (Paroscientific 8B7000-I). The Pythia instrument (described in Sec. 4.2.2) was additionally installed onto *Sentry* for the transect.

TOWED ROSETTE

During the transect, the rosette was equipped with an ultra-short baseline (USBL) acoustic transceiver to allow the real-time position of the rosette to be tracked with respect to the ship. Scientific instruments mounted on the rosette included a transmissometer (C-Star), a 6000 m rated CTD (SeaBird SBE 911plus), twelve 10 L Niskin sampling bottles, and an oxygen optode (Aanderaa). The SAGE instrument (described in Sec. 4.2.2) was also fixed to the rosette for the transect. Default instrumentation on the rosette was communicated via the winch cable to the rosette watchstander station in the computer lab onboard the ship. Ship speed was set to $\sim 0.5 \text{ m s}^{-1}$ (~ 1 knot) to assist in controlling rosette depth and winch tension. Niskin bottles were fired according to a schedule that favored more bottles near the ridge. A scheduled stop approximately

3 km from the ridge was used to collect samples from twelve full Niskin bottles and re-deploy the rosette to take an additional twelve bottle samples from the stop to the end of the transect.

DISSOLVED METHANE ANALYSIS WITH LASER-BASED SPECTROSCOPY A Los Gatos Research (LGR) Dissolved Gas Extraction Unit (DGEU) and coupled LGR Greenhouse Gas Analyzer (GGA) were used to measure dissolved methane in seawater collected by Niskin sampling bottles fired during the transect. The DGEU uses a membrane contactor for dissolved gas extraction. Extracted gas is pumped from the DGEU to the GGA which uses off-axis integrated cavity output spectroscopy for making 1 Hz, precise (<2 parts per billion) measurements of methane in the measurement range of 0-1000 ppm. Extraction of gas is imperfect by the DGEU, and so an extraction efficiency correction of 2.3-3.3% was applied (for calibration details, see Appendix A.1). Methane measurements in ppm were subsequently converted to nanomolar (nM) using coincident salinity and temperature measurements observed by the rosette CTD. Calibration of the GGA was completed using gas standards from Mesa Gas (Michel *et al.* 2021). During the transect, nine of the twelve bottles from Leg 2 were processed using the DGEU and GGA for methane analysis.

AMMONIUM MEASUREMENT Concentrations of ammonium (NH_4^+) were determined onboard within 6 hours of collection from the Niskin bottles following the OPA method (Holmes *et al.* 1999) in a 1 cm cell using an Aquafluor Field Fluorometer (Turner Designs). Standards were prepared using Milli-Q and surface sea water, and then corrected for matrix effects following Taylor *et al.* 2007. Analytical precision was 5 nM, with a detection limit of 1 nM. Ten of the twelve Niskin bottles were processed in this way during Leg 2 of the rosette transect¹.

¹The two bottles not processed for ammonium during this transect were reserved for other water-intensive analyses.

METHANE SENSORS

Two novel sensors for *in situ* methane observation were deployed on the rosette and AUV *Sentry*. The Sensor for Aqueous Gases in the Environment (SAGE) was deployed on the rosette and a real-time cavity ringdown spectrometer called Pythia, was deployed on AUV *Sentry* (Fig. 4.3). Both instruments were in active development during this cruise, and so we report all measurements from these instruments as normalized observations (this can be interpreted as a sensor “saturation” value) in lieu of calibrated concentrations. For the purposes of the analyses herein, there is no loss of generality in the methods proposed to detect hydrothermalism using these normalized values.

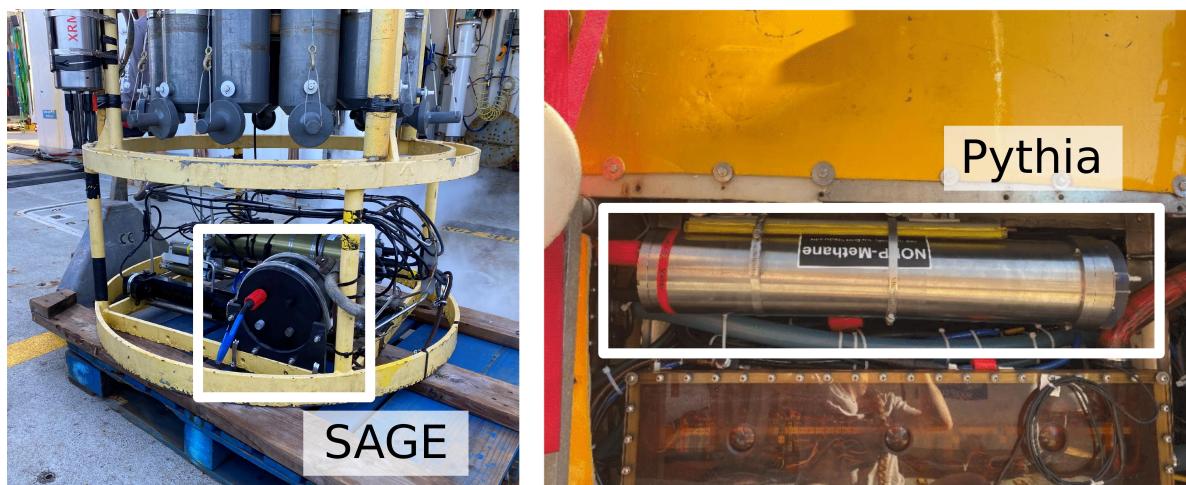


Figure 4.3: **Images of the SAGE and Pythia methane instruments.** The SAGE and Pythia instruments were mounted on the rosette and AUV *Sentry*, respectively.

SAGE SAGE is a dissolved gas sensing technology developed at the Woods Hole Oceanographic Institution (WHOI), and this expedition served as the first at-sea validation of the sensor’s operation. SAGE technology has been previously described in Kapit & Michel 2021a; Kapit & Michel 2021b. Briefly, SAGE is based on infrared absorption spectroscopy performed on extracted gas from seawater via a gas permeable (and water impermeable) membrane. Once the gas enters the sensor, it fills a hollow-core optical fiber (HCF) which also guides light from a laser source tuned to measure the

gas species of interest. The amount of target gas present is determined by measuring the amount of light absorption through the HCF using a photodetector. The prototype version of SAGE deployed on the research cruise was configured to measure methane in the range of 0-10,000 ppm. The resolution of the sensor was <1 ppm and the response time for the deployed configuration was approximately 12 minutes. For the scales of observations reported in this chapter (i.e., <2% of the full scale of the observed signal), the instrument was minimally sensitive to temperature. SAGE is 5.5 in long with a 9 in outer diameter, and the power requirement was 7 W during this field deployment.

PYTHIA Pythia is a novel deep-sea methane sensor developed utilizing real-time cavity ringdown spectroscopy (rt-CRDS) developed by WHOI (Michel *et al.* 2022) and Ring-IR Inc. (Harb *et al.* 2012), and capable of operating to 4000 m depths. Pythia extracts dissolved gas from sea water using a large (113 cm^2) surface area membrane. The extracted sample gas enters an optical cell where it is interrogated by a pulsed mid-infrared Quantum cascade laser (QCL). The laser light is absorbed by methane present in the cell, and the concentration of methane is determined by monitoring the pulsed ringdown signal from the cell using a mercury cadmium telluride (MCT) detector. While the response time of the sensor is slow, on the order of 35 minutes, the sensor is responsive to small (<2 ppm) changes in methane; the temperature sensitivity of Pythia has not yet been characterized. Pythia is ideally suited for long dives in environments in which changes to the methane concentration vary over long temporal and spatial scales. Details on the process for normalizing Pythia observations (which are strongly nonlinear and additionally require time correction) are provided in Appendix A.3. Pythia is 24 in long with a 4.5 in outer diameter, and was operated at a power range between 30-50 W during this field deployment.

4.2.3 ANALYTICAL PROCEDURE

Observations collected by sensors deployed on AUV *Sentry*, including Pythia, were merged into a single dataframe using a common 1 Hz time reference; data were linearly interpolated onto this common time reference if they did not share an exact timestamp. With the exception of the derivative of ORP signal, all data for the purposes of visualization is smoothed using a centered rolling average over 5 minute intervals. Additionally, temperature, oxygen, and salinity measurements are normalized with respect to depth (as these quantities are anticipated to be functions of depth in the weakly stratified deep waters). Depth correction is performed by fitting a linear function to the average observation collected in 20 m wide depth-bins, and computing the residuals of all data points with respect to this line (see Appendix A.4 for plots of the linear functions). Rosette data is treated in the same fashion as *Sentry* data. Down-cast and up-casts are removed from both *Sentry* and rosette data streams for all visualizations.

4.2.4 TRANSECT DESIGN AND EXECUTION

AUV *Sentry* and the rosette were deployed in the basin approximately 16 km from the northern hydrothermal ridge structure, at (27.348152 N, 111.253108 W) with a course of 295° set to intersect the southern part of the ridge (Fig. 4.2). The *Sentry* trackline was placed approximately 200-300 m north of the rosette to avoid any risk of entanglement. *Sentry* was set in altitude hold mode, targeting 120 m from the bottom (this places *Sentry* at a depth of approximately 1750-1700 m, and at the top of its altitude-hold range). Rosette depth was targeted to be approximately 1650-1600 m, controlled primarily by the speed of the ship and length of the winch cable. These depths were designed based on an estimated model of the neutrally buoyant plume layer, as described in Appendix A.5. Leg 1 of the rosette trajectory was terminated at a planned stop at (27.393855 N, 111.364637 W), and Leg 2 was resumed at (27.460812

N, 111.527694 W); see Appendix A.2 for the schedule of bottle samples collected during Leg 2 presented in this manuscript. At the time of the transect, there were no known hydrothermal sites present over the sampling trajectory, save for the northern ridge. Hydrothermal vents in the southern basin were located approximately 40 km further south from the transect starting location (Teske *et al.* 2016).

MODELING TO INFORM TRANSECT DESIGN

The selection of heights for the rosette and AUV *Sentry* was informed by a simple buoyancy model of expected plume characteristics on the ridge, and known operational constraints of AUV *Sentry* (i.e., an absolute floor and ceiling of operation above the bottom). Using an adapted plume crossflow model developed by Tohidi & Kaye 2016 (see Appendix A.5 for more detailed information) with a nominal current crossflow value of 0.1 m s^{-1} , vent temperature of 340°C , and estimated background seawater stratification as per Speer & Rona 1989, a neutrally-buoyant layer was estimated to form between 1570 m and 1750 m. The depths for the rosette (1600-1650 m) and AUV *Sentry* (1700-1750 m) were set given this information in order to target both the upper and lower estimated neutrally buoyant layer (NBL), respectively. The NBL was targeted to increase the likelihood of intersecting plume waters during the transect over a broad, multi-kilometer scope. This is in contrast with targeting the plume buoyant stem, which though significantly easier to distinguish from background seawater, may only have an expression on the order of several square meters.

REAL-TIME DATA FEEDBACK AND WATCHSTANDING

During the transect, data from the standard rosette sensors were available in near-real time at the watchstander station in the shipboard computer lab. This allowed watchstanders to monitor the depth of the rosette and relay requests to the winch operator on deck, and display the data on live-updating visualizers. AUV *Sentry* relayed

occasional data packets up to 128 bytes in length at a rate of approximately 0.01 Hz. These data packets were subsequently graphed on a computer monitor that was linked to the *Sentry* network. A total of 600 messages with information about the standard science instruments on *Sentry*, and 583 messages with information from the Pythia instrument were transferred during the transect.

4.3 RESULTS

4.3.1 METHANE MEASURED BY SPECTROSCOPIC INSTRUMENTS

Elevated methane was observed over a spatial scale of several kilometers, significantly rising as both AUV *Sentry* and the rosette approached the source of known hydrothermalism on the transect (Fig. 4.4). As both methane instruments used on this cruise were in active development, methane observations are reported as normalized values from 0 to 1. A normalized value of 0.5 is used as a conservative threshold for classifying elevated methane measurements. Pythia, mounted on *Sentry*, reached and exceeded this threshold for elevated methane starting at approximately 3 km from the hydrothermal reference point at (27.407489 N, 111.389893 W); SAGE, flying nearly 50 m higher in the water column, reached this threshold starting 1.5 km away. For a less conservative threshold of 0.3, these spatial detection points are reached 6.8 km and 2.2 km away, respectively. SAGE observed a sharp peak of methane just under 1 km from the reference source, with rapid decline of observable methane soon after. In contrast, Pythia reached a methane peak essentially at the 0 km reference point, and shows a gradual decline in methane as *Sentry* descends into a graben just north of the hydrothermal ridge; the rosette was pulled from the water at the ridge. The difference in spatial detection patterns indicated by these instruments may be a function of the different

sensor modalities/sensitivities, the natural structure of the neutrally-buoyant layer, and the relative position of the two platforms within it.

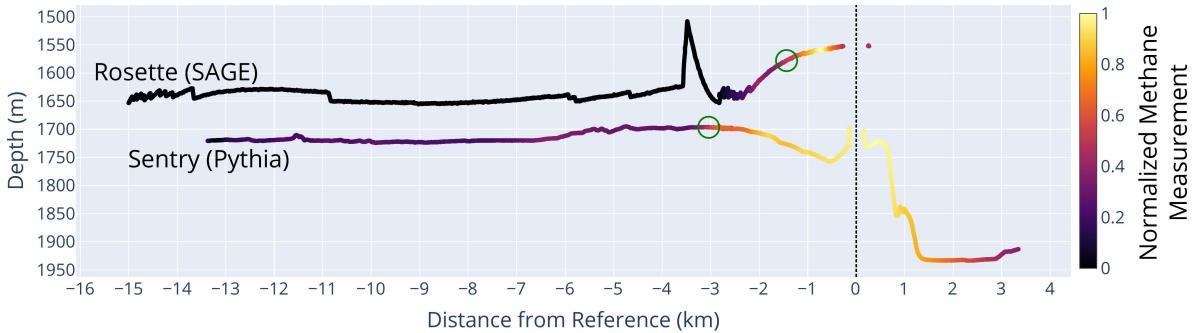


Figure 4.4: Methane observations collected over the transect. Normalized methane values observed with both SAGE (rosette) and Pythia (*Sentry*) over reference distance from (27.407489 N, 111.389893 W). The transect begins at the left of the plot and proceeds to the right. Strong methane anomalies, defined as points above a conservative threshold of 0.5 normalized values, are present starting 3 km from the reference source as observed by Pythia, and 1.5 km as observed by SAGE (open green circles).

4.3.2 METHANE AND AMMONIUM MEASURED WITH THE ROSETTE

Ammonium is a microbial energy source and reduced compound that is produced by the hydrothermal vents at Guaymas Basin. It is expected that ammonium and methane behavior in the basin will behave similarly, providing a “check” on the methane trends observed in methane bottle samples, and recorded by SAGE. Focusing primarily on Leg 2 of the rosette transect, there is a correspondence between methane and ammonium elevation in the approach to the hydrothermal ridge (Fig. 4.5). Methane samples processed directly from Niskin bottles as outlined in Sec. 4.2.2 show a peak methane concentration of 3000-4000 nM (this range is associated with the extremes of calibrated extraction efficiencies valid for the equipment used), approximately 0.75 km from the hydrothermal reference point. Ammonium tracks closely with methane, at 3-4 times smaller concentration, reaching a peak of approximately 1000 nM.

Normalized methane observations by SAGE generally follow the trends shown by the bottle samples, similarly showing a spatial peak at 0.75 km. However, by its nature, SAGE yields a significantly more resolved signal; a small, secondary peak is observed by SAGE at 2 km from the reference point which is essentially missed by the bottle samples. Additionally, by virtue of operating continuously, there is no need for human interaction (unlike for processing bottle samples, which can require time-intensive *ex situ* analysis).

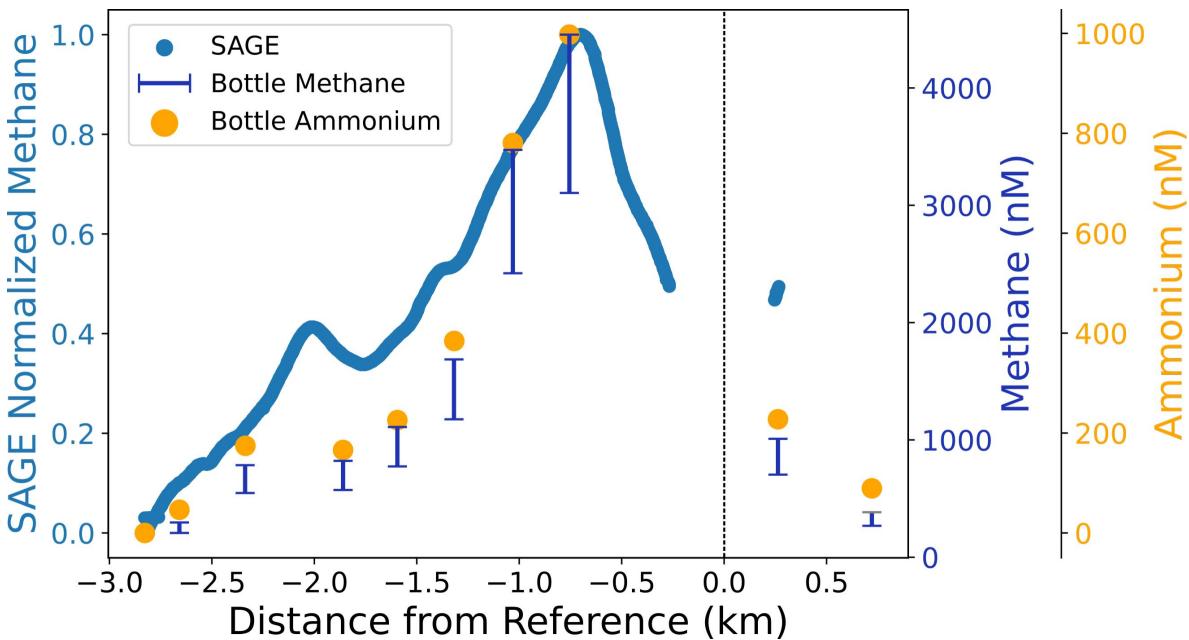


Figure 4.5: **Methane observations compared to ammonium concentrations.** Normalized methane measurements by SAGE plotted with methane measurements taken from Niskin bottle samples (as measured by DGEU/GGA equipment) and ammonium measurements. Bottle methane measurements are reported as a range to reflect sensitivity of the measurement procedure to a calibrated extraction efficiency. All measurements trend towards a peak observation of methane and ammonium 0.75 km from the reference source. SAGE additionally observes a secondary peak approximately 2 km from the source, which is essentially missed by the bottle sample schedule.

4.3.3 TURBIDITY

Turbidity is a commonly used indicator for detecting hydrothermalism from smoking vents; particulate matter produced by smoking vents can remain suspended in the neutrally buoyant layer, acting as a non-conservative tracer for hydrothermalism (Feely *et al.* 1992). In the Guaymas Basin, suspended particulates have been shown to be composed of metals like iron, aluminum, and manganese (Scholz *et al.* 2019) and are primarily mixed into bottom waters from hydrothermal activity. Here, turbidity measurements are reported as normalized values to make direct comparison between the platforms straightforward; in absolute terms, the transmissometer on the rosette reported beam attenuation values between 0-0.2 and the OBS on *Sentry* observed backscatter values between 0.08-0.14. The OBS on *Sentry* encountered an error from the beginning of the dive, potentially caused by a persistent air bubble, until approximately 4.5 km from the ridge reference point; these early measurements are omitted.

Elevated turbidity (defined by a conservative threshold of 0.5 in the normalized data) was observed with the transmissometer on the rosette starting approximately 2.2 km from the reference source and 3.3 km with the OBS on *Sentry* (Fig. 4.6). Even with a less conservative threshold (0.3) these detection points only slightly improve to 2.5 km and 3.4 km respectively. With *Sentry*, a rapid decline in turbidity within tens of meters west of the source reference (positive distance in Fig. 4.6) is observed. This may be indicative of the direction of prevailing crossflow (southeast) in the basin, which would directionally bend a buoyant plume stem and advect the neutrally buoyant layer.

4.3.4 OXIDATION REDUCTION POTENTIAL

AUV *Sentry* carries an ORP sensor; there was no comparable sensor on the rosette. ORP sensors are commonly used in hydrothermal plume hunting, and can be a strong indicator of recently emitted hydrothermal fluids. The derivative of ORP (noted here

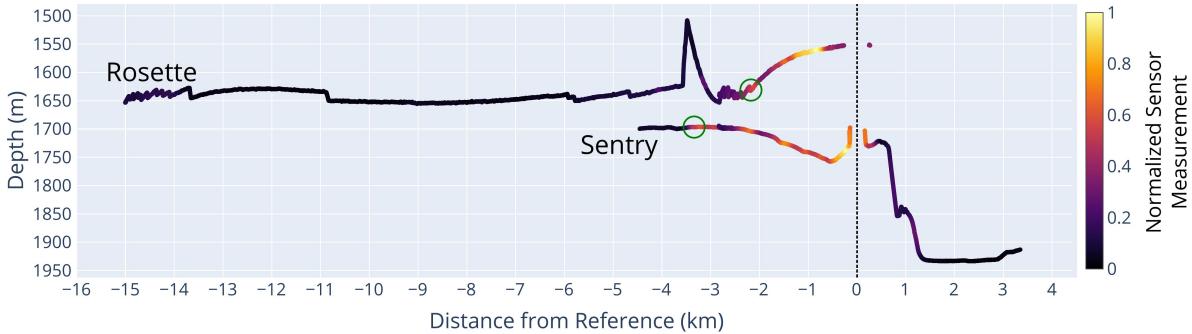


Figure 4.6: Turbidity measurements collected during the transect. Turbidity observed as beam attenuation on the rosette transmissometer and optical backscatter on AUV *Sentry* instruments. *Sentry* encountered a sensor error until approximately 4.5 km from the ridge reference point. After this point, elevated turbidity is detectable throughout the dive, with significant elevations within 3.3 km east of the ridge reference point, dissipating within tens of meters to the west. Elevated turbidity is observed by the rosette 2.2 km from the ridge reference point to the east.

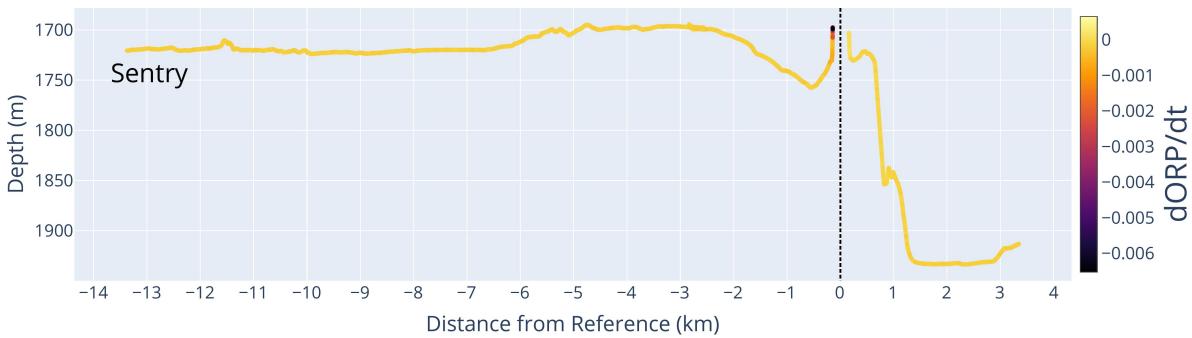


Figure 4.7: Oxidation-reduction potential measurements collected during transect. The derivative of ORP observed by data collected on AUV *Sentry*. Negative slopes are indicative of entering hydrothermal fluids. Only one region of the transect demonstrated a significant reaction to ORP, within 200 m of the reference point.

as $d\text{ORP}/dt$) is particularly used, in which negative $d\text{ORP}/dt$ values typically indicate transition from background water into hydrothermal fluid. During the transect, only one significant $d\text{ORP}/dt$ deviation was observed, within 200 m from the ridge reference point (Fig. 4.7).

4.3.5 TEMPERATURE, SALINITY, AND OXYGEN

Temperature, salinity, and oxygen are expected to be weakly stratified in deep ocean waters, however fluids from hydrothermalism should register as anomalies when present. The magnitude of valid anomalies (i.e., anomalies that positively identify fluids impacted by hydrothermalism) can be exceedingly small; temperature at a vent can be hundreds of degrees Celsius, but anomalies in the water column on the spatial order of only 10 m can be measured as single degrees, and within a neutrally-buoyant plume on the order of hundreds of meters from the source, only register a few hundredths of a degree (Yoerger *et al.* 2007).

Temperature, salinity, and oxygen anomalies are computed according to the process described in Sec. 4.2.3 and the results are shown in Fig. 4.8. Salinity anomalies, although apparently coherent, are reported within the empirical sensor noise for the CTD instruments on both the rosette and *Sentry*. Temperature anomalies on the scale of hundredths of a degree are observed throughout the transect, with two key regions of high temperature anomaly, one located 6-12 km from the reference source, and the other within 3 km of the source. Both the rosette and *Sentry* observe these regions; with *Sentry* observing the first anomaly in a narrower margin between 8-11 km from the reference source. The first region of positive temperature anomaly closely corresponds with marginally fresher water; whereas the region of higher temperature anomaly near the source is not consistently matched in salinity (the rosette observes more salinity content, whereas *Sentry* observes neutral or slightly less salinity content). Oxygen is reported as nominal or slightly depleted within the regions of notable temperature and salinity anomaly.

The first region of interest, far afield from the plume reference point, appears coherent and has similar detection qualities to the near-reference region; however, given the typical expectation of temperature dissipation from hydrothermal sources, it would be

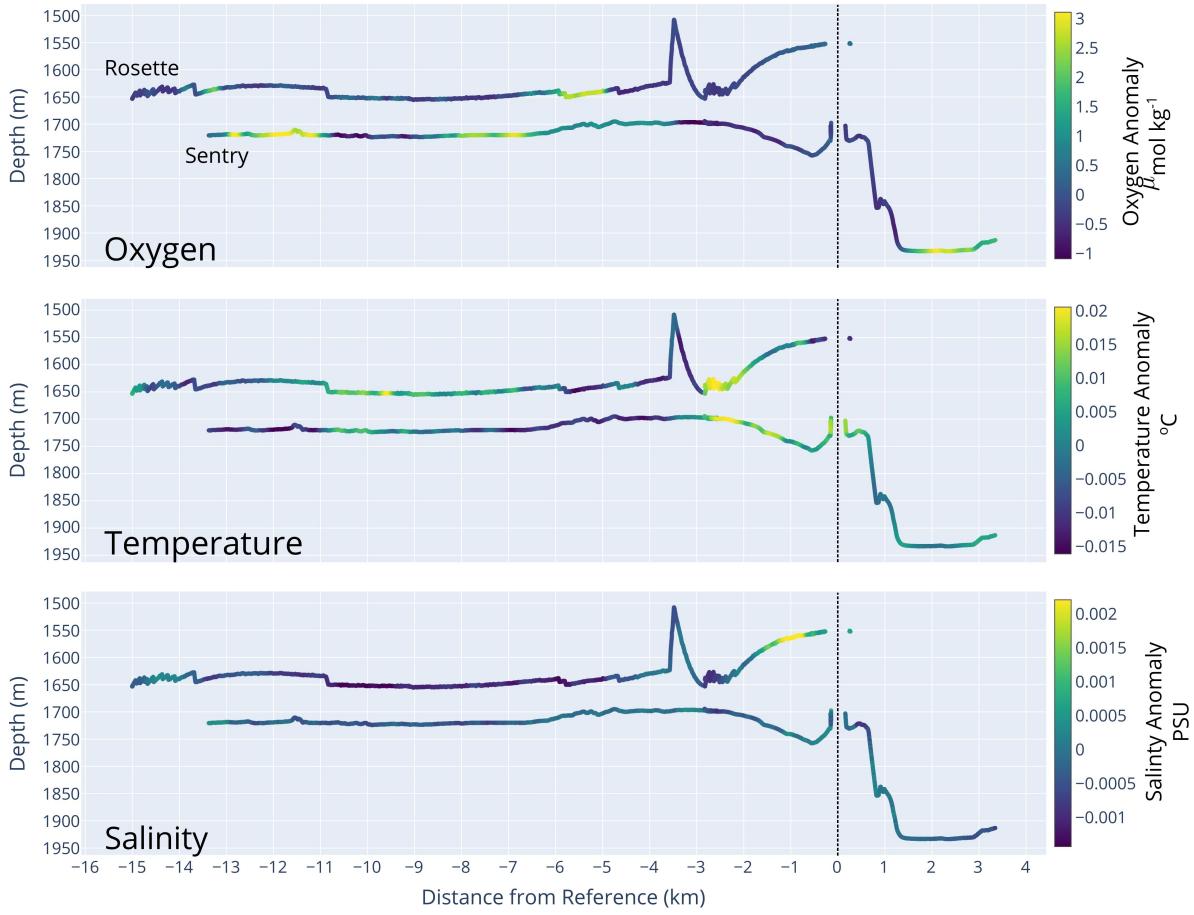


Figure 4.8: Depth-corrected oxygen, temperature, and salinity over reference distance. Two notable regions of high temperature deviation from expected temperature are observed between 6-12 km (rosette; 8-11 km by *Sentry*) and within 3 km of the reference source. The first region of temperature anomaly is closely matched with fresher salinity measurements; whereas salinity is measured as marginally higher near the reference source by the rosette CTD and nominal or lower by the *Sentry* CTD. In both regions, oxygen is nominal or slightly depleted, with regions of notably elevated oxygen at the boundary of these regions.

surprising if this first region were connected with hydrothermalism. The shape of the warm, slightly fresher and oxygen depleted intrusion (laterally broad higher in the water column, and appearing to narrow based on the observations taken by the rosette and *Sentry* approximately 50 m offset in altitude) also does not follow expected patterns in a neutrally buoyant plume layer. Lack of significant methane and turbidity observations in this same region, as presented in Sec. 4.3.1 and Sec. 4.3.3 respectively, additionally casts doubt on hydrothermalism as a driver for this anomaly. Water mass mixing

between the bottom waters, largely sourced from Pacific Deep Waters and the Pacific Intermediate Waters (Bray 1988) may be an alternative explanation, but is out of scope for this study to investigate.

4.4 DISCUSSION

4.4.1 SENSOR CROSS-CORRELATIONS

Successfully detecting hydrothermalism in the deep ocean is a significant challenge, and detection may be most effective using a combination and corroboration of anomalies across multiple sensor inputs (Jakuba 2007). Here, the cross-correlation between sensors mounted on each of the platforms is investigated. Both a global and rolling Pearson correlation coefficient were computed, showing overall correlation trends and situation dependent correlation, respectively.

Fig. 4.9 shows the global correlation among sensors mounted on the rosette individually over Leg 1 and Leg 2, in addition to sensors mounted on *Sentry*. In the absence of significant geochemical features in a target environment, it is expected that no or only weak correlation will be computed globally, as individual sensor noise (which is independent) will dominate the computation; when geochemical structure is present in the environment, it is expected that weak to strong global correlation will be computed as the environment is imposing a (shared) signal across at least a subset of sensors. This is well illustrated by the cross-correlation matrices for the rosette legs, with global coefficients for Leg 1 reporting no correlation between sensors save for a slightly negative correlation between temperature and oxygen, and for Leg 2 reporting weak to strong correlations between all sensors, with notably strong positive correlation between turbidity and methane. Interestingly, in Leg 2 a negative correlation is reported between temperature and methane, and a positive correlation is measured between methane and

oxygen measurements. This runs directly counter to expectations; and also counter with the relationships observed by *Sentry* which marks relationships between methane and temperature as positively correlated, and between methane and oxygen as negatively correlated.

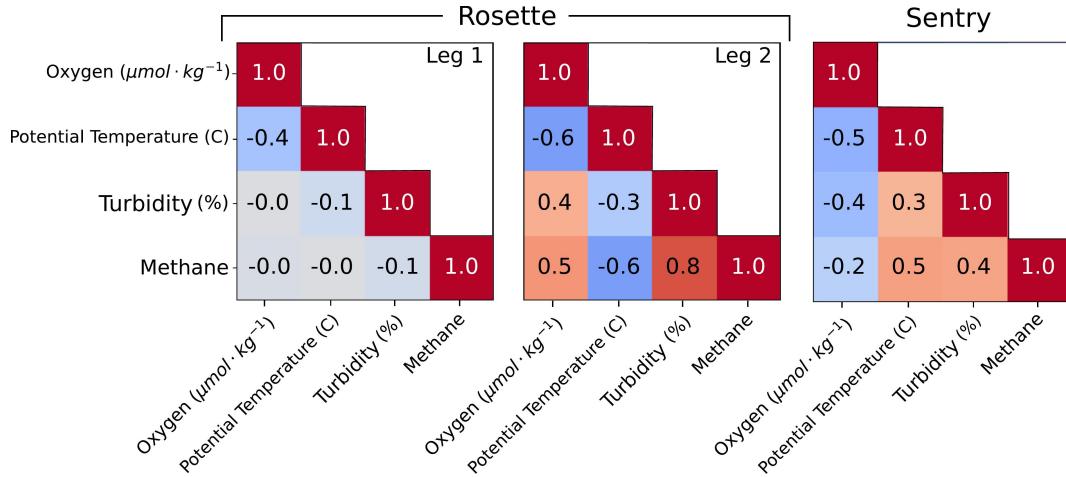


Figure 4.9: Global Pearson correlation coefficient between sensors mounted on the rosette and *Sentry*. Correlation differences between Leg 1 (far from the reference point) and Leg 2 (near the reference point) are indicative of different sensor correlation behaviors with respect to ambient seawater conditions and hydrothermal fluid interception. *Sentry* correlation coefficients reflect an expected relationship between temperature and methane (positive), methane and oxygen (negative), and turbidity and methane (positive) that may be stereotypically associated with hydrothermal fluids. In contrast, the Leg 2 rosette correlation factors do not meet this expectation, despite showing strong overall correlative structure.

The difference between correlative behaviors between the rosette legs, and also between the platforms generally, motivates a finer study of correlation. Fig. 4.10 shows a rolling correlation coefficient computed over a window of 30 minutes for the rosette. Computing local cross-correlations with respect to time, rather than distance, is mathematically more sound², and also aligns directly with how cross-correlative monitoring may be used during live exploration missions. Using the correlative “micro-structure” of rolling coefficients shows regions of possible interest that are greater than nominal (uncorrelated structure). Looking first at the rosette information, Leg 1 nominal

²In the sense that samples are regularly sampled in time, but irregularly sampled in space.

correlation early in the transect is weak or non-existent between most sensors, with exception for oxygen and temperature at around 2:00. Additionally, a coherent region from 04:00-05:00 shows a negative correlation between temperature and turbidity, and positive correlation between temperature and oxygen. In Leg 2, overall more strong, pronounced correlations between sensors are observed, with a distinct period centered in the hour around 10:00 in which correlation between temperature and methane, temperature and turbidity, oxygen and methane, and oxygen and turbidity appear to “flip” compared to the periods of time directly before and after this period, potentially indicating a significant anomalous feature. This time period is well aligned with the spatial proximity of the rosette with the reference source.

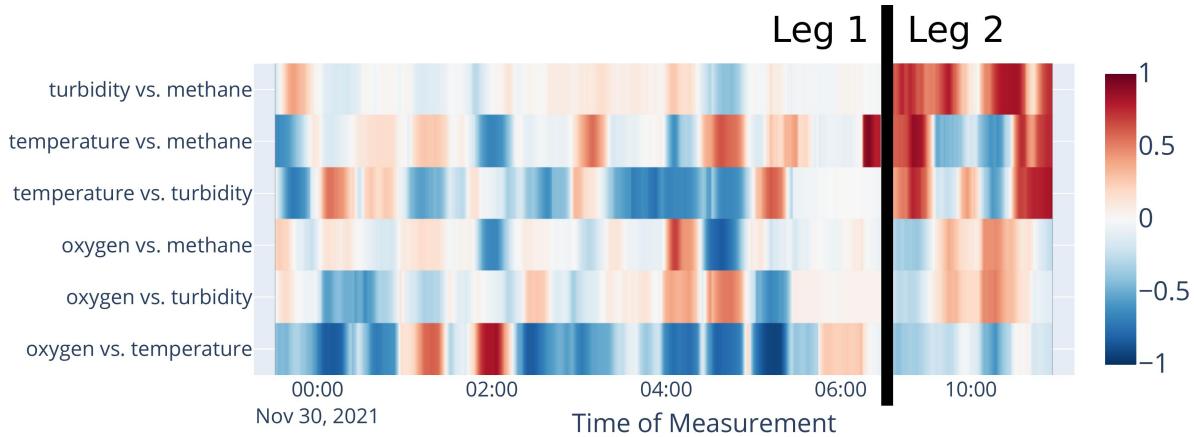


Figure 4.10: Local (rolling) Pearson correlation coefficients between sensors mounted on the rosette over 30 minute windows. Several regions of interest at 2:00 (strong temperature-oxygen correlation), 4:00-5:00 (notably coherent region of stronger correlation across multiple sensors), and during Leg 2 (near the venting source) can be picked out and may indicate anomalous water masses.

Local correlation trends during the *Sentry* transect are reported in Fig. 4.11, and show an intense relationship between oxygen and temperature throughout the dive, with most of the transect reporting a strong negative correlation, save for two regions of positive correlation at 08:00 and again at 11:00. This strong relationship is also reflected in the relationships of temperature and oxygen with methane, being nearly correlative mirrors with respect to methane. During periods in which the turbidity

sensor was operational, a gradual correlative “flip” and intensity increase in correlation between turbidity and oxygen around 11:00 may indicate a structured water mass. This time stamp agrees with the spatial proximity of *Sentry* with the reference source.

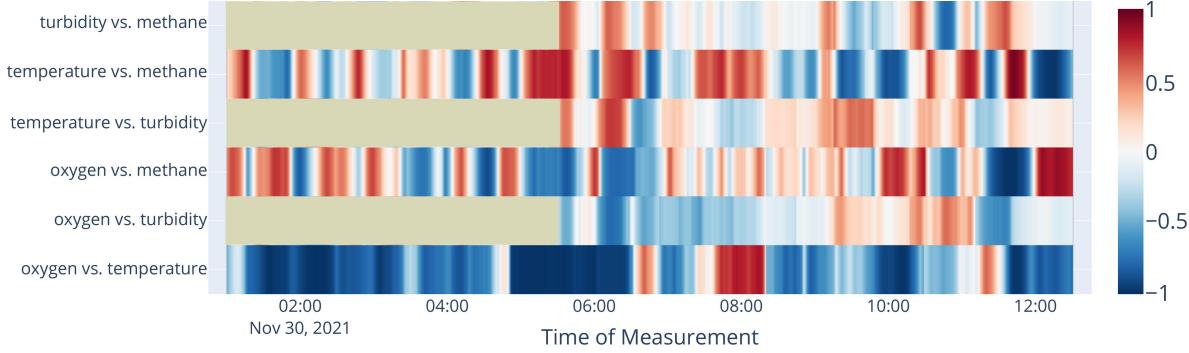


Figure 4.11: Local (rolling) Pearson correlation coefficient between sensors mounted on AUV *Sentry* over 30 minute windows. Temperature and oxygen are strongly negatively correlated throughout the dive, with the exception of 08:00 and 11:00. For times in which the turbidity sensor was operational, the oxygen-turbidity correlative relationship slowly flips from negative to positive, with a peak positive correlation at 11:00. 11:00 agrees with the time that *Sentry* was near the reference source vent.

Correlation alone is not sufficient evidence for the presence of hydrothermal fluids. For instance, some of the coherent regions of positive or negative correlation with methane any time during Leg 1, or early in the *Sentry* transect, are misleading, as the overall methane content of the water was exceedingly small or essentially background. Rolling correlations, coupled with absolute thresholds as reported in Section 4.3, may together be useful tools for indicating transition into new water masses, their absolute properties of which could be used to more closely classify the types of water masses. This correlative study also demonstrates that correlations in expectation (e.g., temperature and methane being positively correlated in hydrothermal fluid) may be reductive assumptions of the complexities of plume evolution within a water column, supporting similar findings e.g., by Cowen *et al.* 2002. For instance, aging plume waters in the neutrally buoyant layer may long have settled to a temperature indistinguishable from background, but still be particulate and possibly gas rich. This motivates additional

study of the “classes” of hydrothermal fluids and their classifying characteristics, which could in turn be used to support studies of microbial evolution and nutrient consumption in plume fluids, or sediment and particulate transport modeling.

4.4.2 HYDROTHERMALISM DETECTION VIA TIME-SERIES REGIMES

As indicated by Sec. 4.4.1, changes in correlative *structure* may be a more useful signal than absolute correlation alone. This notion can be codified as regime changes, which detect inflection points for which a series of observations collected in time may change in typical value, oscillation frequency, or pattern. Here, regime changes are computed over 30 minute detection windows using linearly penalized segmentation (Killick *et al.* 2012) (PELT) as implemented in the `ruptures` Python library (Truong *et al.* 2020) with a radial basis function detection kernel. PELT is a linear-time offline algorithm for selecting changepoints that incrementally performs binary segmentation on a time-series (based on a cost function defined by the detection window and basis function) until all segments are self-consistent; these segments are regimes. In the included figures, we visualize regimes using alternating red and blue color blocks.

In Fig. 4.12 regimes across the entire rosette transect over multiple sensors is illustrated. One initial observation is that the water-mixing anomaly that occurs early in the transect (Sec. 4.3.5) appears to be detected as regime changes in potential temperature, oxygen, and even a correspondence in lowered beam attenuation. Similarly, regime changes in turbidity and methane are early indicators of significant elevation of both of these factors as the rosette intersects with hydrothermal fluids later in the transect. Interestingly, a regime change in oxygen and temperature is evident immediately following the first small peak in methane and turbidity in the absolute measurement data. These peaks, in addition to these regime changes, may together be indicative of mixing plume sources from other hydrothermal vents located along the ridge (that

must travel further than fluids from our reference point) or the mixing of aging plume waters with more recently emitted fluids.

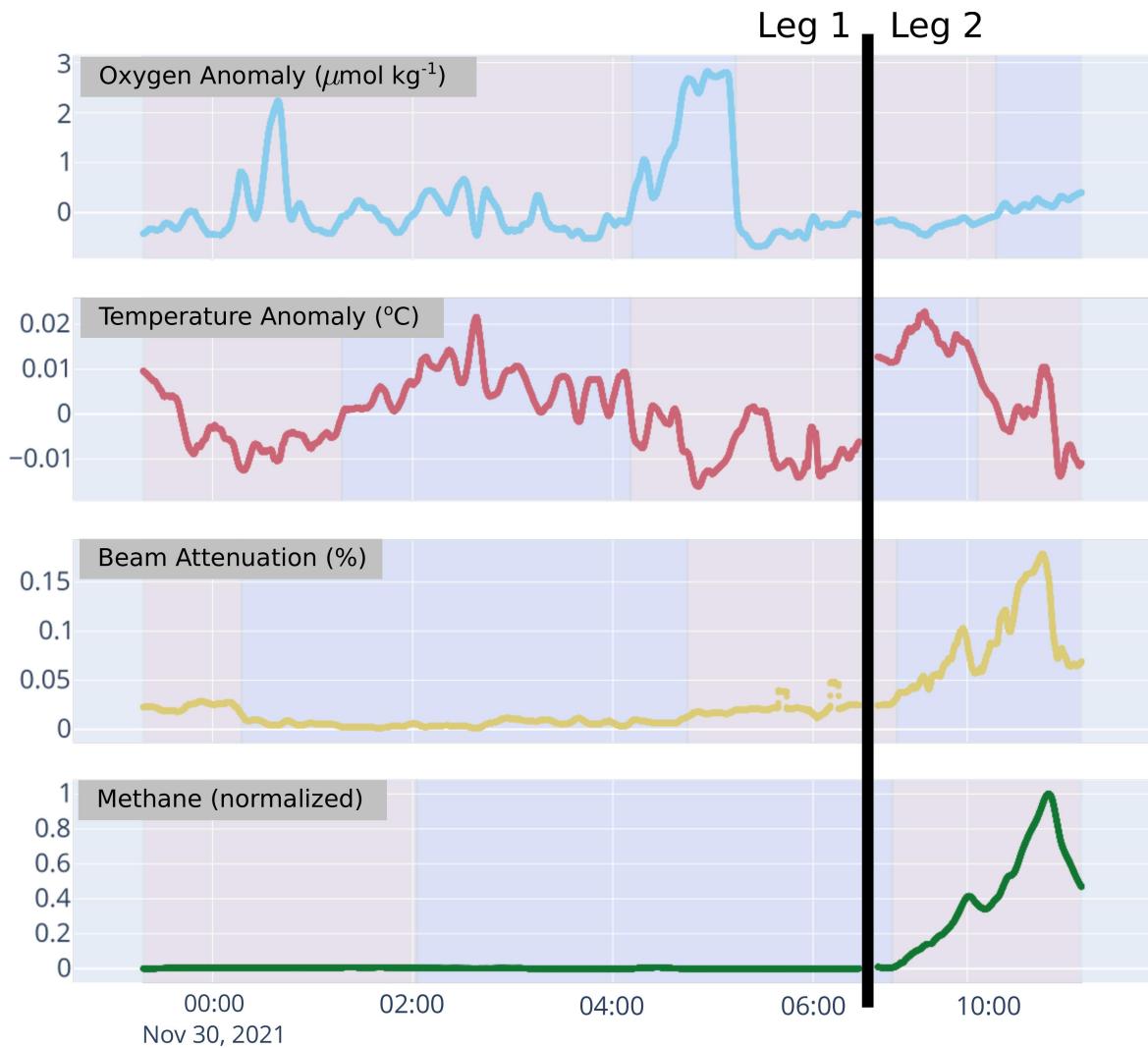


Figure 4.12: **Regime changes in rosette observations.** Regimes, indicated as alternating blue and red regions, detected during the rosette transect with a 30 minute detection window.

With instruments mounted on *Sentry*, in Fig. 4.13, clear “steps” of methane observed by Pythia each mark a regime in that data stream. Some of these steps are nearly coincident with regime changes in turbidity, temperature, and oxygen (particularly the steps at 06:30 and 09:30).

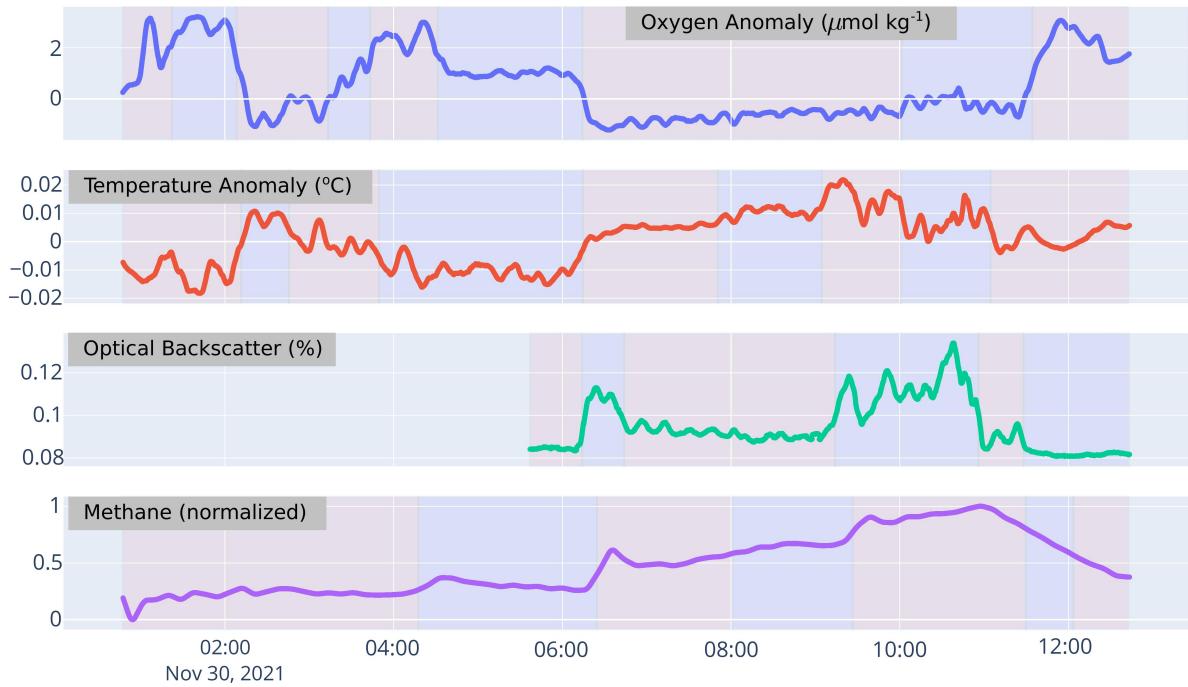


Figure 4.13: Regime changes in AUV *Sentry* observations. Regimes, indicated as alternating blue and red regions, detected during the AUV *Sentry* transect with a 30 minute detection window.

Regimes can be mathematically identified in streaming data, making this a potentially useful method to adopt for real-time hydrothermalism discovery. Coupled with absolute measurements by sensing instruments and rolling correlative structure, identifying water masses across multiple data streams can be done live from streaming data on the ship. The simplicity of the computation and the nature of these analysis techniques as data reduction methods (e.g., regimes can be reported as a single time stamp; cross-correlations over strategic sensor pairs could be reported as a single float) additionally make computing these measures onboard an AUV and reporting them back to watchstanders under data-limited transmission protocols (e.g., acoustic pings) feasible.

4.4.3 METHANE IN DEEP SEA EXPLORATION

AUV and sensor deployments during expedition RR2107 served as an initial proving ground for the SAGE and Pythia *in situ* methane instruments for deep sea exploration,

and the utility of methane as a potential tracer for hydrothermalism discovery. During this transect both instruments observed significantly elevated methane over a span of several kilometers from a known hydrothermal source in Guaymas Basin. Methane proved to be a strong predictor for hydrothermalism that was not easily confounded by physical oceanographic events (e.g., mixing), giving it an advantage over oxygen, temperature, and salinity. Indeed, in this trial, each of the oxygen, temperature, and salinity instruments were impacted by an unknown physical feature not driven by hydrothermalism, but registered as similar scales of expected anomaly. Methane was also shown to be more expressive than ORP, which only registered a possible anomaly long after significant methane measurements were observed. Turbidity was a similarly useful and expressive feature of hydrothermalism in this basin, with similar detection scales to methane during this transect. Notably, for less strict detection criteria (i.e., thresholds) on detection, methane significantly outperformed turbidity in terms of detection scale (positive identification up to 6.8 km away, in contrast to 3.4 km for turbidity). Turbidity and methane together make for a strong pairing for hydrothermalism discovery. While neither one alone is a “universal” proxy for hydrothermal activity—not all hydrothermalism of interest produces particulate heavy smoke (i.e., diffuse flow fields) nor do all vents produce significantly elevated methane—they are complementary indicators which can assist in deep sea exploration for anomalous water masses derived from hydrothermalism.

Collecting high resolution measurements of methane during this transect highlighted the rich structure of dissolved gasses in a neutrally buoyant plume layer over multiple kilometers, with multiple peak detections being possibly indicative of mixing novel and aging hydrothermal fluids, the contribution of multiple sources of hydrothermalism, or complicated internal mixing causing spatiotemporal multimodal distributions of dissolved gas “pockets” throughout the layer. Bottle samples collected on the research

cruise verified the presence and general trend of methane observed by the instruments, but failed to resolve several features that may be of scientific interest. This motivates the use of *in situ* methane sensors for future studies of hydrothermal fluids in the water column.

4.4.4 ENABLING BETTER DECISION-MAKING FOR HYDROTHERMALISM DISCOVERY

Enabling the interpretation of real-time sensor data and adapting scientific missions accordingly are critical future skills for scientific expeditions and exploration in the deep sea. In preparation for this transect, a simple physical model was used to inform the design of the trajectory and monitored progress with live data displays for both the rosette and AUV *Sentry*. While real-time data display for rosettes is now considered standard for oceanographic research, streaming capabilities of scientific data from autonomous platforms like *Sentry* is a relatively new capability. This display infrastructure enabled the science team to make note of the OBS sensor error on *Sentry* while performing the transect, caught a power and logging failure of the Pythia logger upon deployment (which, if left unresolved, would have meant an absence of all methane data associated with *Sentry* for this analysis), and made real-time control and decision-making about the rosette positioning and bottle firing possible. While data presented here was analyzed after the mission, several of these analyses, including rolling correlation and regime detection, could be performed from streaming observations. As a whole, the techniques in this chapter present an opportunity for advancing technical infrastructure on a research vessel in order to enhance decision-making capabilities of the science party and engineering teams, both logistically to better diagnose instrument operation *in situ* and scientifically to enhance data collection.

Real-time data collection and processing could have further implications for embodied intelligence as a tool for scientific expeditions. Using models, inference methods, and streaming data, autonomous agents like AUV *Sentry* could be made capable of performing adaptive decision-making for sample collection. Hydrothermalism discovery has long been a motivating use case for intelligent autonomy at sea (Branch *et al.* 2020; Jakuba 2007; Wang *et al.* 2020; Yoerger *et al.* 2007). This transect experiment demonstrates the utility of simple models for tractable, intelligent planning, motivates the possibility of using methane as an additional, reliable data source for performing autonomous behaviors (e.g., adaptive sampling, tracking), and presents the opportunity to embed simple analytical methods for classifying hydrothermal fluids from sensor streams. Being able to both map the source of hydrothermal plumes with ROVs and chart the evolving nature of fluids in the mid-water with AUVs, would enable an advancement of scientific inquiries that could be pursued with respect to hydrothermalism in the deep ocean. Such queries include the detailed structure of multiple-source plume collision, directly measuring *in situ* the 4D structure of mixing in neutrally buoyant plumes and buoyant plume stems, assessing biological activity supported by plume fluids, tracing the fate of dissolved gasses, and more. This work demonstrates that detection of hydrothermal sources is possible on the scale of several kilometers even from relatively small hydrothermal vents that are present in Guaymas Basin, and has taken initial steps to demonstrate core data infrastructure that can improve human decision-making in hydrothermalism discovery; future work and engagement will be focused on advancing these tools to enable the next generation of scientific inquiry in the deep ocean.

5 PHYSICALLY-INFORMED OPERATIONAL ROBOTIC TRAJECTORIES FOR SCIENTIFIC EXPEDITIONS

Now it would be very remarkable if any system existing in the real world could be exactly represented by any simple model. However, cunningly chosen parsimonious models often do provide remarkably useful approximations...For such a model there is no need to ask the question “Is the model true?”. If “truth” is to be the “whole truth” the answer must be “No”. The only question of interest is “Is the model illuminating and useful?”.

George Box

On expeditions, quality observations of water column properties and their anomalies are strong evidence used by a science party to strategically plan research activities—deployment of certain instrumentation or platforms (where, when, why), allocation of resources to different projects, and scheduling for high-stakes one-off missions. Chapter 4 showed how to parse scientific instrumentation data for hydrothermal plume identification by domain experts to assist in data discoverability and interpretation. These

tools provide descriptive snapshots of what was observed during a mission, but rely on the science experts to extrapolate these snapshots into future mission states. This chapter directly addresses how to recover a predictive model of plume state from partial observations collected from heterogeneous sensors, and formulates an automated routine for leveraging this predictive model to design dense charting surveys with AUV *Sentry*.

5.1 INTRODUCTION

Transient, dynamic phenomena—deep-sea hydrothermal plumes, algal blooms, warm core eddies, lava flows—are of interest in many disciplines of observational science. *Expeditionary science* encapsulates the observational sciences that require *in situ* sample collection of environmental phenomena for scientific discovery and model development. In such cases, the environmental targets are typically impossible to observe using remote sensing (e.g., satellites) either due to desired spatial and temporal resolution, environment adversity (e.g., the deep sea, within closed structures), or the nature of the scientific target of interest and corresponding sensing equipment (e.g., building a taxonomy of algae requires physically processing water samples). Expeditionary science is, by definition, conducted in partially-observable and often dynamic environments; thus performing useful and effective data collection in expeditionary science poses a challenge for human and autonomous decision-makers alike.

In this chapter, mapping the space-time dynamics of deep-sea hydrothermal plumes using an autonomous mobile robot is studied. Deep-sea hydrothermal plumes are a source of chemicals and particulates that play a significant role in biogeochemical cycling in the deep ocean (Dick *et al.* 2013; Le Bris *et al.* 2019; Resing *et al.* 2015; Scholz *et al.* 2019; Vic *et al.* 2018). Understanding the fate of chemicals and particulates in hydrothermal plumes is of significant interest to biogeochemists and physical oceanographers; however,

directly studying plumes in the water column is a substantial challenge. Chiefly, deep-sea plumes have complex dynamics in space and time—advection forces (e.g., deep currents, topographic updrafts), diffusion, and turbulent mixing act on plumes as they rise through the water column. A robot tasked with charting a plume must be able to forecast where and when it will intersect with different regions of the plume in order to collect useful observations of its spatiotemporal structure, but the underlying dynamics of a plume are uncertain as these forces are generally unobserved. The problem is exacerbated by the inherent challenge of sensing a plume using chemical sensors — the chemical distribution of a plume can only be observed with point sensors and no single point measurement is sufficient to locate the plume.

In addition to the technical challenges of determining a sensing strategy for a highly uncertain and dynamic phenomena, the deep-sea environment (>200 m depth) can only be accessed by depth-capable equipment that are often significantly constrained by operational and safety policies. For example, autonomous underwater vehicles (AUVs) in this setting are typically restricted to execute preset trajectories hand-designed by human scientists (e.g., Camilli *et al.* 2010). In this mode, the AUV cannot react to measurements while executing the set trajectory. Such “open-loop” trajectories can result in sparse measurements of a target phenomenon, such as a dynamic plume, or can miss short-lived events entirely (Flasphohler *et al.* 2019; Preston 2019). However, this open-loop concept of operations remains the state-of-the-art in practical deep-sea science because preset trajectories are easy to encode, do not require extensive on-platform computing resources, and result in predictable robot actions that can easily be supervised. Here, constraints introduced by a specific robot, the AUV *Sentry*, are considered. *Sentry* is operated by the National Deep Submergence Facility (NDSF) at the Woods Hole Oceanographic Institution (WHOI) (Kaiser *et al.* 2016) and by operational policy is typ-

ically only permitted to execute pre-determined regular trajectories like “lawnmower” patterns or spirals, making online adaptation impossible.

Enabling depth-capable robots such as *Sentry* to perform expeditionary science and study spatiotemporal phenomena using preset, non-adaptive behaviors requires an autonomy system that can design fixed trajectory patterns strategically to maximize desirable intersections with a dynamic phenomena. To strategize effectively, this autonomy system should learn to forward-simulate the dynamics of the target environment over a long-horizon from a small history of robot deployments and plan subsequent deployments using these predictions. This is fundamentally a *sequential decision-making problem*, and is closely related to informative path planning (IPP) problems, in which a robot selects behaviors according to an information-theoretic reward function using a probabilistic belief of the environment. Existing methods in IPP (e.g., Binney & Sukhatme 2012; Flaspohler *et al.* 2019; Hitz *et al.* 2017; Hollinger & Sukhatme 2013; Levine *et al.* 2010), the related field of experimental design and optimal sensor placement (e.g., Krause *et al.* 2008; Wang *et al.* 2019b), and general decision-making under uncertainty (e.g., Kocsis & Szepesvári 2006; Silver & Veness 2010; Somanı *et al.* 2013; Sunberg & Kochenderfer 2018) have demonstrated that sequential decision making can be applied to sampling scenarios in which online, adaptive behaviors are possible, the phenomenon of interest is static, and/or there is an opportunity to train the belief model from many trials, multiple sensors, or highly adaptive trajectories. The assumptions made in each of these typical scenarios is violated in the expeditionary science sampling problem, and this chapter proposes an autonomy framework, PHORTEX: **P**Hysically-informed **O**perational **R**obotic **T**rajectories for **E**Xpeditions, that addresses each of these constraints with the set of algorithmic choices made for the observational model, belief representation, and planner within the framework (Fig. 5.1).

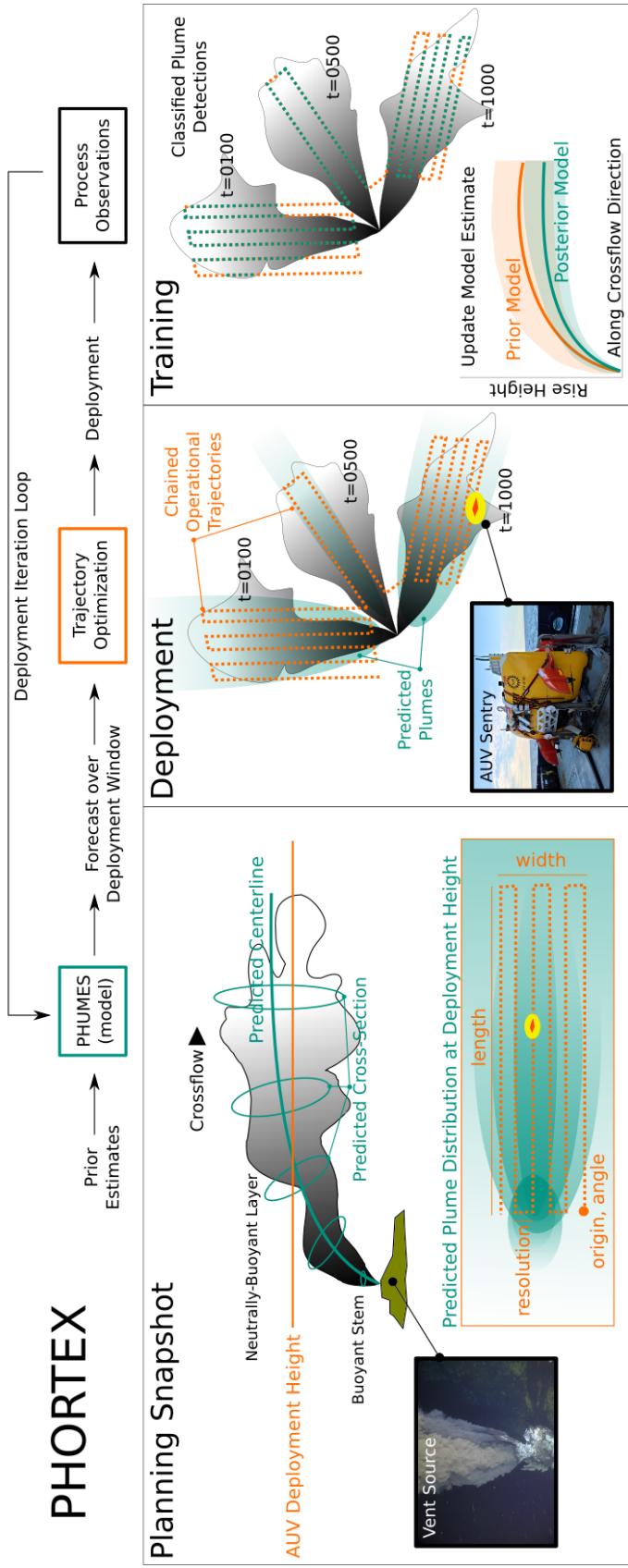


Figure 5.1: An overview of PHORTEX: PHysically-informed Operational Robotic Trajectories for EXpeditions. Over the course of an expedition, AUV *Sentry* may be deployed several times. In preparation for a deployment, the belief model PHUMES: PHysically-informed Uncertainty Models for Environment Spatiotemporality is used to generate a probabilistic forecast of temporally-evolving plume centerlines and cross-sections from estimates of vent characteristics and fluid crossflow (e.g., current) which can be either seeded with prior information from scientific knowledge, data of opportunity from other deployed sensing equipment, or previous *Sentry* deployments. A trajectory optimizer is given the forecasts and for a given altitude that *Sentry* can operate, modifies the parameters of a “lawnmower” trajectory primitive (height, width, origin, orientation) to maximize expected in-plume samples. Several lawnmowers are chained together to form a complete deployment trajectory and “chase” the moving plume distribution throughout the deployment area. *Sentry* executes the trajectory and following a *Sentry* deployment, observation locations are classified as binary plume detections from analysis of several science sensors. This product is then used to update the PHUMES model of plume centerline and cross-section over time. The new PHUMES model is then used to plan the next deployment of *Sentry*.

5.1.1 CONTRIBUTIONS

In this chapter, the autonomy system PHORTEX is formulated to solve the hydrothermal plume charting problem under operational constraints imposed by a state-of-the-art AUV, *Sentry*. *Sentry*, by policy, can only execute pre-defined trajectories while underway, and can only be deployed a small number of times during a given expedition. As modern IPP, plume-hunting, and front-tracking techniques strongly rely on underway adaptive behaviors, these frameworks are extended for *Sentry* by formulating a *deployment-by-deployment* sequential decision-making problem which treats each pre-defined deployment of *Sentry* as a single action in sequence with few steps. Each deployment action is defined as a chain of operationally-approved trajectory primitives (i.e., lawnmower patterns), which are parameterized by a small number of characteristics including their relative size, resolution, and position, and are optimized to collect the most number of expected in-plume detections.

To represent the robot’s knowledge (belief) in order to optimize a given chain for tracking a target plume, a probabilistic model PHUMES: **P**Hysically-informed **U**ncertainty **M**odels for **E**nvironment **S**patiotemporality is introduced, which provides long-horizon forecasts of plume state. As very few deployments of AUV *Sentry* are possible during an expedition, PHUMES must overcome the challenge of sample-efficient dynamics learning from sparse, partial observations. To do so, the existence of well-characterized analytical models for buoyant plume dynamics used in ocean and atmospheric sciences is leveraged to embed a numerical simulator into a Bayesian filtering framework. The use of this simulator creates a strong inductive bias for the dynamics learning problem for a given field site. There are several advantages to this scheme: the forecasts generated by PHUMES are driven by a set of physically-meaningful parameters (e.g., vent temperature, crossflow magnitude) which are interpretable by the science team and over which the science team may have useful prior knowledge; the PHUMES framework

can easily accept data or information external to *Sentry* deployments that map to the physically-meaningful parameters; and forecasts that are generated support computation of many summary statistics (e.g., mean, variance), providing flexibility for defining reward functions for the trajectory optimization scheme.

Simulations demonstrate that *Sentry* using PHORTEX can collect more spatially and temporally diverse plume-derived fluid samples as compared with classical surveying approaches. The diversity of samples corresponds to observing more unique regions of a plume structure, and has important, positive implications for scientific inquiry to be performed on the dataset collected by *Sentry* in post-expedition analyses, to be further discussed in Chapter 6.

The rest of this chapter is organized as follows: in Section 5.2 the hydrothermal plume charting sequential decision-making problem is formally presented as a partially-observable Markov decision process (POMDP). PHORTEX and PHUMES, the algorithmic contributions for approximately solving the hydrothermal plume charting POMDP are presented in Section 5.3. Section 5.4 provides a description of a simulated deployment and experimental results are discussed in both Section 5.5 and Section 5.6. In Section 5.7 future work related to the contributions of this chapter are discussed, and in Section 5.8 some closing remarks are provided.

5.1.2 ADVANCING HYDROTHERMAL STUDIES WITH ROBOTS

Hydrothermal vents, as described in Section 1.1 and Section 3.5, produce warm, chemistry rich fluids that rise rapidly into the water column until reaching a point of neutral buoyancy with the ambient water, where the plume then spreads out in a neutrally-buoyant layer. When studying hydrothermal plumes, recovering location of the venting source, reconstructing the shape of the plume structure, and estimating energetic and chemical flux (e.g., heat output into the ocean), are key objectives.

A wealth of prior work has primarily focused on localizing hydrothermal venting plume sources (e.g., Ferri *et al.* 2010; Jakuba 2007; Kim *et al.* 2020; Mason *et al.* 2020; McGill & Taylor 2011; Nakamura *et al.* 2013; Paduan *et al.* 2018; Wang *et al.* 2020) using a variety of equipment such as ship-based acoustics, towed instrument rosettes, remotely-operated vehicles (ROVs), human-occupied vehicles (HOVs), and autonomous underwater vehicles (AUVs). Specialized seafloor equipment is subsequently deployed at the inferred venting site to e.g., estimate bulk chemical or nutrient flux from the vent or characterize the driving magmatic system underneath the crust. Generally, vent localization approaches use detections of anomalous water masses (as determined from *in situ* sensors) in the water column to constrain the location of a seafloor vent. The localization methods can be fully offline, in which surveys by vehicles like *Sentry* with no adaptive capacity are post-processed and vent locations are inferred (Jakuba 2007; Nakamura *et al.* 2013), or they can be fully online, in which autonomous gliders with adaptive capabilities utilize e.g., gradient descent to seek a plume source (Wang *et al.* 2020). In (Branch *et al.* 2020), an autonomous glider tasked with localizing a vent source adaptively chains uniform coverage trajectories with increasingly fine resolution as the robot position converges on an estimated vent location. This chaining is emulated in PHORTEX, however the selection of trajectories is done completely offline before AUV *Sentry* is deployed. Online strategies for hydrothermal plume hunting almost universally employ glider-type robot platforms, which are typically smaller, payload-limited, and less depth-capable than vehicles like *Sentry*. 90% of known vent fields are deeper than 200 m in the ocean, and over 75% are deeper than 1000 m (Beaulieu *et al.* 2013). Gliders widely accessible to the research community are typically not rated deeper than 1000 m, which means that deep-sea research is reliant on vehicles like *Sentry* and demand advances in offline-suited planning techniques.

Also relevant to this study is “plume hunting” research in robotics, which has been equivalently styled as odor mapping, odor localization, source localization, and source seeking. In these works, the source emits a substance (e.g., gas, radio, acoustic, odor) and through partial observations of the emitted substance, the source is discovered using techniques that can be divided broadly into biologically-inspired heuristic search (e.g., Chen & Huang 2019; Reddy *et al.* 2022) or adaptive IPP (e.g., Salam & Hsieh 2019). Biological or heuristic techniques, including gradient-based algorithms like chemotaxis (Morse *et al.* 1998), or algorithms that directly mimic a particular animal (Edwards 2001), tend to be reactive and myopic. In contrast, adaptive IPP can be nonmyopic, and algorithmically tends to embed knowledge (either heuristically or rigorously) about flow-fields (i.e., advection and diffusion) to assist in plume localization. Such techniques also live on a spectrum, from algorithms that resemble biologically-inspired approaches like infotaxis (Vergassola *et al.* 2007) to methods that use model order reduction techniques to encode complex numerical models (e.g., Navier-Stokes equations) into a belief model to better treat complex data (Peng *et al.* 2014).

While source discovery remains an important area of research, in this chapter the focus is on how science can be advanced at the hundreds of vents that have been successfully identified. Thus, a complementary problem to source discovery is posed: *given a venting source, what impact do the venting fluids have on the local environment?* In this framing, rather than using detections of a plume as a means of source localization, the detections themselves are the valuable data product for scientific inquiry. By placing instruments throughout an evolving plume structure over multiple length- (meter to kilometer) and time- (hours to days) scales to collect dense in-plume measurements, previously intractable questions with respect to microbial lifecycle and transport, carbon cycling, and anomalous water mass formation, can be approached. Work that has used robots to map or chart plume-like structures has been presented as

the “front-tracking” problem (Chen & Huang 2019; Li *et al.* 2014). In this problem, two water masses converge (such as the warm hydrothermal fluid and the cold background seawater), and the goal is to use a robotic vehicle to track the edge of these water masses or stay within a single type of water mass. The importance of both multirobot collaboration and online decision-making in these schemes is essential to their efficacy; the research in this chapter is the first to present a water mass tracking solution within an offline optimization strategy with a single agent, and the first to attempt this for the hydrothermal charting problem.

5.1.3 CLOSING THE LOOP: EXPEDITION LOGISTICS

Oceanographic research expeditions are an undertaking that requires the coordination and collaboration of a science party, external engineering teams that maintain and operate the scientific equipment used during studies, and the captain and crew aboard a research vessel (on which everyone lives and works during operations). Deep-sea (>200 m depth) capable robotic platforms used in oceanic research are assets independently maintained from a ship, and typically requested on a per-expedition basis. AUV *Sentry* may be deployed on tens of expeditions in a given year, with up to 250 days at sea (Kaiser *et al.* 2016). Safety of both people and equipment are held to the highest importance. Further, the critical role of *Sentry* in oceanographic research drives the strict operational policies that dictate *Sentry* deployments to prevent vehicle loss or damage.

It is with this context that *Sentry* deployments are designed by the science party and ultimately approved by the *Sentry* engineering team. In a typical workflow, the science party may provide a set of coordinates or waypoints they generate based on bathymetric maps, prior knowledge, or previous data (when available). The *Sentry* team designs survey trajectories based on these coordinates and respecting basic op-

erational constraints of the vehicle (e.g., speed, minimum/maximum altitude from the seafloor). With approval of the *Sentry* team, science party, and captain, the survey is then executed. A single “dive” of *Sentry* is multiple hours (typically not less than 5 hours, and under 24 hours). At the conclusion of a dive, *Sentry* is recovered from the ocean and data products containing hundreds of thousands of point measurements from multiple heterogeneous sensors are made available to the science team within a few hours after *Sentry* returns to the deck. Depending on the length of the dive, 12-18 hours of vehicle cycling time (e.g., recharging, instrument maintenance, preparation for the next deployment) are required. Based on the length of an expedition and other ongoing research activities, *Sentry* may be deployed only a handful of times.

The complexity of these basic operations for *Sentry* alone, in addition to the burden of coordinating several other ongoing scientific projects happening simultaneously, day-to-day operational changes, and unforeseen discoveries and hurdles, make performing “closed loop science” with robot platforms a challenge while at sea. For hydrothermal plume monitoring, a combination of sensor streams need to be used to make confident plume detections (Jakuba 2007), but information about the exact tidal state, state of the venting source, and background sea characteristics are typically not provided in these products, and can require fusing data from other instruments deployed on a cruise, if available. The planning challenge is further exacerbated when the design of a new mission requires not just deep analysis of the collected data, but forecasting the implications of those data onto a new day, new site, or new objective.

PHUMES and PHORTEX aim to alleviate the burden of closing the loop onboard a research vessel for AUV operations by generating interpretable phenomenon forecasts informed from diverse data streams and trajectories that can be verified by science party members and approved by *Sentry* engineers. Algorithmically, the formulation of PHORTEX as a sequential decision-making framework produces trajectories which are

informed by previous observations, thus literally behaving like a closed-loop controller for robot actions.

5.2 PROBLEM FORMULATION

During scientific expeditions, the objective of a robot is to collect useful measurements, as defined by a given task (e.g., reduce uncertainty over a quantity, find the global optimum in a distribution, track a moving target). For hydrothermal plume charting, the goal is to map or “chart” the spatiotemporal structure of a buoyant plume using a dynamically constrained AUV. Such a chart enables scientists to infer relevant scientific properties of generating vents (e.g., chemical flux) and to create detailed models of deep-sea interactions and nutrient cycling. This section describes how general problems in expeditionary science can be formulated as sequential decision-making problems, the specific constraints of AUV *Sentry* and how they influence the sequential decision-making problem, and finally presents a formal description of hydrothermal charting as a partially observable Markov decision process.

5.2.1 SCIENTIFIC EXPEDITIONS AS A SEQUENTIAL DECISION-MAKING PROBLEM

Expeditionary science requires a robot to make a sequence of decisions to collect scientifically useful measurements of an unknown, partially-observable spatiotemporal environment under operational constraints. Generally, the expeditionary science decision-making problem can be formulated as a partially observable Markov decision process (POMDP). Let $\mathcal{P}(\cdot)$ denote the space of probability distributions over the argument. A finite horizon POMDP can be represented as tuple: $(\mathcal{S}, \mathcal{A}, T, R, \mathcal{Z}, O, b_0, H, \gamma)$, where \mathcal{S} are the states, \mathcal{A} are the actions, and \mathcal{Z} are the observations. At planning iteration

t , the agent selects an action $a \in \mathcal{A}$ and the transition function $T : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{P}(\mathcal{S})$ defines the probability of transitioning between states in the world, given the current state s and control action a . The transition function governs both how the state of the robot will evolve, given a chosen action, and the potentially stochastic evolution of the underlying spatiotemporal environment. After the state transition, the agent receives an observation according to the observation function $O : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{P}(\mathcal{Z})$, which defines the probability of receiving an observation, given the current state s and previous control action a . The reward function $R : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ serves as a specification of the task, assigning the states of the world that are useful for a given scientific objective high reward and others low reward. A POMDP is initialized with belief $b_0 \in \mathcal{P}(\mathcal{S})$ — an initial probability distribution over state — and plans over horizon $H \in \mathbb{Z}^+$ with discount factor $\gamma \in [0, 1]$.

As the robot moves through the world, it selects actions and receives observations. Since the state of the world is not directly observable in a POMDP, the robot maintains a probability distribution over possible states (i.e., belief) and must update this distribution each time it takes an action and receives an observation. Given the transition and observation models, the belief can be updated directly using a Bayes filter (Särkkä 2013):

$$\tau(b_{t-1}, a_{t-1}, z_t) = b_t \triangleq \mathcal{P}(S_t | a_0, z_0, \dots, a_{t-1}, z_{t-1}, z_t) \quad (5.1)$$

$$= \frac{\int_{s \in \mathcal{S}} O(s, a_{t-1}, z_t) T(s, a_{t-1}, s') b_{t-1}(s')}{\int_{s' \in \mathcal{S}} O(z | s', a_{t-1}) \int_{s \in \mathcal{S}} T(s' | s, a_{t-1}) b_{t-1}(s)} \quad (5.2)$$

where $\tau(b, a, z)$ is the updated belief after taking control action a and receiving observation z (Eq. 5.2). Unfortunately, Eq. (5.2) is often intractable to compute directly and an approximate Bayesian inference procedure is required to represent the belief (e.g., a Kalman filter (Welch, Bishop, *et al.* 1995), a particle filter (Silver & Veness 2010), or variational methods (Kucukelbir *et al.* 2017; Wainwright & Mulligan 2002)).

Due to the stochastic, partially observable nature of current and future states, the realized reward in a POMDP is a random variable. Optimal planning is defined as finding a horizon-dependent policy $\{\mathcal{P}_t^* : \mathcal{P}(\mathcal{S}) \rightarrow \mathcal{A}\}_{t=0}^{H-1}$ that maximizes expected reward: $\mathbb{E}\left[\sum_{t=0}^{H-1} \gamma^t R(S_t, \mathcal{P}_t(b_t)) \mid b_0\right]$, where b_t is the updated belief at time t , conditioned on the history of actions and observations. The recursively defined horizon- h optimal value function V_h^* quantifies, for any belief b , the expected cumulative reward of following an optimal policy over the remaining planning iterations: $V_0^*(b) = \max_{a \in \mathcal{A}} \mathbb{E}_{s \sim b}[R(s, a)]$ and

$$V_h^*(b) = \max_{a \in \mathcal{A}} \mathbb{E}_{s \sim b}[R(s, a)] + \gamma \int_{z \in \mathcal{Z}} \mathcal{P}(z \mid b, a) V_{h-1}^*(\tau(b, a, z)) dz, \quad (5.3)$$

for $h \in [1, H - 1]$.

The optimal policy at horizon h is to act greedily according to a one-step look ahead of the horizon- h value function. However, Eq. (5.3) is intractable for large or continuous state, action, or observation spaces and thus the optimal policy must be approximated. Much of the art of practical decision-making uncertainty is making well-designed algorithmic and heuristic choices that enable efficient and robust planning algorithms.

5.2.2 SCIENTIFIC DECISION-MAKING WITH AUV *Sentry*

There are several levels at which an AUV can behave autonomously. At the lowest level, an AUV is given navigation waypoints and executes a closed-loop controller and state estimator to drive to that waypoint; this type of autonomy is commonly implemented and executed on AUV platforms. AUV *Sentry* is capable of autonomously navigating between given waypoints using a closed-loop controller and a state estimator that uses acoustic ranging between the robot and the ship to set latitude, longitude, and depth coordinates. More sophisticated autonomy systems that are aware of scientific

objectives can build upon this waypoint controller to enable sophisticated autonomous behavior.

Underway or online autonomy is one such sophisticated controller: after a waypoint is reached, an autonomy system chooses the next waypoint for a vehicle to target. When operating with underway autonomy, an AUV can act in “closed-loop”, using observations of the environment in real-time to inform subsequent waypoints while underway. Underway autonomy has the potential to greatly increase the utility of the collected data. For example, the AUV may serendipitously encounter plume water while navigating to a waypoint; an underway autonomy system could then attempt to follow chemical gradients to the plume center. However, at present, *Sentry* is not capable of underway decision-making. The lack of underway abilities is both a logistical and policy constraint. Logistically, the robot is computationally limited and solving a POMDP online often requires significant onboard computational resources. *Sentry* relies on a high-latency acoustic link to communicate with the ship, meaning that data from *Sentry* cannot be streamed to an external computing resource on the ship for decision making (science data communication between ship and robot is 0.02 Hz assuming no packet loss, and only a subset of sensor data can be made available in any given packet). Additionally, by policy, *Sentry* trajectories are rigorously vetted before each dive using bathymetric maps of the target region and dynamics validation schemes. Risk aversion to avoid losing or damaging *Sentry* leads to the policy that underway plan changes cannot be part of normal operating procedures.

Thus, to enable sequential decision-making with *Sentry* requires the use of *deployment-by-deployment* autonomy. Unlike underway decision-making, deployment-by-deployment autonomy does not modify the AUV trajectory in real-time, but instead leverages the “down-time” between robot deployments to post-process data with shipboard computers, update a belief model about the environment, and plan a new fixed trajectory for

the next deployment. This form of autonomy honors the strong requirement that each deployment must pass through a rigorous safety and validation check, while enabling adaptive search behavior based on accrued knowledge between deployments. Each planning “step” or iteration in the POMDP framework is an entire deployment of *Sentry*. Although deployment-by-deployment autonomy is less flexible and reactive than underway autonomy, it is a very useful and practical form of autonomy for many applications of scientific robots (e.g., extra-planetary rovers, intermittent monitoring systems) and already fits into the typical way in which science data is used on ships (now it is just the autonomous system making decisions, rather than a person on the science team).

5.2.3 CHARTING HYDROTHERMALISM AS A POMDP

The plume charting POMDP can be formalized as follows:

THE STATE SPACE \mathcal{S} The state space of the plume-charting POMDP consists of the joint continuous states of the environment (i.e., the plume) and the robot. The environment state will be represented by a d -dimensional vector of continuous plume parameters $\mathbf{x}_p \in \mathbb{R}^d$ and a current vector $\mathbf{x}_c \in \mathbb{R}^2$ that contains the heading and velocity of the prevailing crossflow, which vary in time. The robot state will be represented by a vector $\mathbf{x}_r \in \mathbb{R}^3$ that represents the latitude, longitude, and depth of the robot.

THE ACTION SPACE \mathcal{A} The action space of the plume-charting POMDP consists of sequences of parameterized lawnmower pattern (i.e., back-and-forth uniform coverage) trajectory primitives. The selection of the lawnmower as the base primitive was given by *Sentry* operators. By chaining lawnmower trajectories together during a deployment, a relatively expressive action set is available. Each trajectory primitive is parameterized by a set of b real-valued parameters $\theta \in \Theta \subseteq \mathbb{R}^b$. These parameters include scale (height and width that describe the rectangle in which the lawnmower is contained), resolu-

tion (the absolute distance between tracklines of the lawnmower), and global position (latitude-longitude-depth coordinate and planar angle of the origin of the primitive). The robot's action set then consists of sequences of parameterized trajectories, i.e., $\mathcal{A} = \{\theta_1, \dots, \theta_n\}$, $n \in \mathbb{Z}^+$. The number of trajectory objects n and the altitude or depth for which a trajectory will be executed for a given chain is fixed *a priori* to planning.

THE TRANSITION FUNCTION T The transition function $T(s, a, s')$ will be decomposed into a plume transition T_p , a current transition T_c , and a robot transition function T_r .

- The plume state parameters \mathbf{x}_p , e.g., venting characteristics like plume exit velocity or vent temperature, are assumed to be constant and therefore the plume transition function T_p is given by: $T_p(\mathbf{x}_p, a, \mathbf{x}'_p) = \delta_{\mathbf{x}_p=\mathbf{x}'_p} \forall a \in \mathcal{A}, \mathbf{x}_p, \mathbf{x}'_p \in \mathbb{R}^d$. Although it is possible for plume parameters to vary on a timescale relevant to a robotic deployment (over the course of hours (Chevaldonné *et al.* 1991)), the overall impact to gross features of plume rise height, bend angle, and cross-sectional area is essentially negligible, which is reflected in the form of the transition function provided.
- The current transition function T_c is more complex and driven by tidal cycles, local bathymetry, and deep sea currents. A deterministic current transition function will be estimated $T_c(\mathbf{x}_c, a, \mathbf{x}'_c) = \delta_{\mathbf{x}'_c=h(\mathbf{x}_c)} \forall a \in \mathcal{A}, \mathbf{x}_c, \mathbf{x}'_c \in \mathbb{R}^2$, where the function h evaluates the future current magnitude and heading from the present current state, from point observations of current magnitude and heading from a sensor that is not part of the robot (described in detail in Section 6.3.2).
- The robot transition function T_r assumes that the robot's waypoint controller is deterministically able to execute a planned trajectory: $T_r(\mathbf{x}_r, a, \mathbf{x}'_r) = \delta_{\mathbf{x}'_r=g(\mathbf{x}_r, a)}$, where the function g evaluates the goal waypoint of the trajectory given by a . Although there is some uncertainty in the robot's transition, in practice localiza-

tion and control are generally well-resolved for state-of-the-art AUVs and pose uncertainty contributes minimally to the robot’s task execution compared with uncertainty about the plume state.

THE REWARD FUNCTION R The reward function for the plume-charting POMDP encodes the robot’s objective to produce a comprehensive map of the plume. This objective is approximated by rewarding the robot for collecting observations of “plume fluids”, i.e., water that is expected to be derived from hydrothermal vents as indicated by the belief of the environmental state $R([\mathbf{x}_p, \mathbf{x}_c, \mathbf{x}_r]^\top, a) = \mathbb{I}[\text{in_plume}(\mathbf{x}_p, \mathbf{x}_c, \mathbf{x}_r, a)]$.

THE OBSERVATION SPACE \mathcal{Z} The robot carries a variety of scientific and navigational sensors. Here, a sensor model that fuses and converts complex, continuous scientific observations into a simplified measurement of plume content in a given fluid parcel $z_p \in \{0, 1\}$ is used, discussed in Section 5.3.1. By performing this filtering step, the dimensionality and complexity of the observation space is significantly reduced. Outside of the robot, another sensor provides independent observations of current magnitude $z_g \in \mathbb{R}^+$ and heading $z_h \in \{(-\pi, \pi]\}$. Thus the observation space \mathcal{Z} consists of multiple observations of z_p , z_g , and z_h .

THE MEASUREMENT FUNCTION O The measurement function encodes the relationship between the plume parameters and heterogeneous scientific sensors on the robot, the prevailing current with the external sensing system deployed by the science party, and the robot location with the navigation equipment aboard the vehicle and ship. A sensor model described in Section 5.3.1 is used to process scientific sensor data into a measurement that indicates whether a fluid parcel was derived by a plume, and PHUMES (Section 5.3.2) is utilized to map both current data and the simplified plume measurement to plume parameters. It is assumed that the robot position is fully-observable and exactly reported by the navigation equipment.

THE HORIZON H AND DISCOUNT FACTOR γ In deployment-by-deployment autonomy, the horizon H can be set to be equal to the total number of deployments to be conducted during an expedition and the discount factor γ set to 1.0. However, practically, the state of *Sentry* at the end of one deployment often has little or no impact on its achievable reward in the subsequent deployment due to the constrained nature of deployments and the ability for *Sentry* to be released or recovered from a ship at arbitrary coordinates. Under this assumption, the discount factor can be set to zero $\gamma = 0$ to break the finite-horizon sequential decision-making problem into a sequence of horizon-1 planning problems. This reduces the capacity of the planner to reason about long-term, multi-dive information gathering actions, but, as seen in the following sections, computationally simplifies the planning problem.

5.3 METHODOLOGY

To solve the plume-charting POMDP described in Section 5.2.3, this section presents PHORTEX, which first utilizes a physically-informed probabilistic model (PHUMES) to generate forecasts of spatiotemporal distributions of plume fluids and then optimizes chains of trajectory primitives (e.g., lawnmowers) to maximize the total number of observations of those plume fluids. PHORTEX iteratively improves the performance of these trajectory chains for each deployment of AUV *Sentry* using the history of collected observations from the robot’s heterogeneous science sensors.

5.3.1 SENSOR MODEL FOR SCIENCE OBSERVATIONS

In hydrothermal vent hunting literature, it is often assumed that binary measurements which indicate the presence or absence of plume-derived fluids are available for an inference framework (e.g., Saigol *et al.* 2009; Tian *et al.* 2014). In this chapter, this same assumption is applied, and Chapter 6 provides details of practically deriving

this measurements from field data. The result of such a sensor model is to convert multiple, time-stamped sensor observations $s_{t,i} \in \mathbb{R}$, $i = 1, \dots, S$ to a time-stamped binary plume-detection $z_{p,t} \in \{0, 1\}$. These binary plume detections are then used to update PHUMES and plan robot trajectories, as described in the following sections.

Although rather spatially expressive, binary observations are generally difficult to use for estimating temporally evolving crossflow \mathbf{x}_c or characterizing ambient seawater properties, which are quantities useful in formulating a predictive model of dynamics. To alleviate the burden on the binary detections, access to external information available opportunistically from standard science party activities is assumed to estimate these quantities for use in PHUMES. This assumption is well-founded by the ubiquity of current-sensing and water-profiling instruments in oceanographic field settings. What was available during field work in the Guaymas Basin is described in Chapter 6, and here assume access to a noisy model of crossflow magnitude and heading, and a noisy profile of density stratification of the water column to be incorporated in PHUMES, described in the following section.

5.3.2 PHUMES: PHYSICALLY-INFORMED PROBABILISTIC FORECASTS

PHUMES is a modeling approach that can generate predictions of the distribution of a spatiotemporally evolving state from a history of sparse state-space observations. To quickly learn a predictive model of a spatiotemporal phenomenon, PHUMES leverages access to analytical scientific simulators (when available) codified as systems of ordinary differential equations (ODEs). These simulators reduce the dimensionality of the inference problem from the full-state of the environmental phenomenon (e.g., a 4D volume in space and time with binary phenomenon measurement) to the dimensionality of the initial conditions and parameters of the simulator (which can then be used to populate the full-state for planning purposes). The use of ODE systems, as opposed to high-

fidelity numerical simulators using partial differential equations (PDEs), is intentional; the computational requirement of most PDE systems used to model environmental phenomenon at the scales studied during expeditionary missions is intractable. In contrast, ODE systems are less well-resolved, but summarize the structure of an evolving phenomenon in a useful way for positioning a robot to encounter plume fluids and that can be enhanced by a generic probabilistic formulation wrapping the ODEs.

For hydrothermal plume charting, a time-averaged model of buoyant plume evolution through a weakly stratified fluid under crossflow can be used. The model described in Tohidi & Kaye 2016 with modifications for seawater as adapted from Xu & Di Iorio 2012 is used in this research, which for simplicity in notation will be referred to as function $f(\cdot, \cdot)$. The crossflow “bends” the buoyant stem of the plume, and reduces the effective rise height of the plume by introducing more mixing. Using a modified cylindrical coordinate system in which s represents a point along the axis described by the plume centerline and θ describes the vertical angle from the base of the plume, the PHUMES simulator takes the form:

$$\frac{dQ}{ds} = Q \sqrt{\frac{2(1 + \lambda^2)}{M\lambda}} (\alpha \left| \frac{M}{Q} \right| - U_a \cos \theta) + \beta |U_a \sin \theta| \quad (5.4)$$

$$\frac{dM}{ds} - U_a \cos \theta \frac{dQ}{ds} = \frac{FQ}{M} \sin \theta \quad (5.5)$$

$$U \sin \theta \frac{dQ}{ds} + M \frac{d\theta}{ds} = \frac{FQ}{M} \cos \theta \quad (5.6)$$

$$\frac{dF}{ds} = -QN^2 \sin \theta \quad (5.7)$$

$$x_a = \int_0^s \cos \theta ds \quad (5.8)$$

$$h_a = \int_0^s \sin \theta ds \quad (5.9)$$

where $U_a = U_a(z)$ is the ambient crossflow velocity (equivalently T_c in the POMDP formulation), $Q = Q(s, \theta)$ represents the plume specific volume (or heat) flux, $M = M(s, \theta)$ is the specific momentum flux, $F = F(s, \theta)$ is specific buoyancy flux, N is the Brunt-Väisälä frequency, λ is the ratio of the minor and major axis that define the plume cross-sectional ellipse, x_a and h_a represents the Cartesian transform of s and θ within the plume's frame of reference, and α and β are vertical and horizontal entrainment coefficients. To convert abstract notions of buoyancy and momentum flux to more readily interpretable and observable vent characteristics (e.g., vent area, fluid exit velocity), the following relationships can be used:

$$Q_0 = \lambda u_0 \frac{A_0}{\pi} \quad (5.10)$$

$$M_0 = Q_0 u_0 \quad (5.11)$$

$$F_0 = g 10^{-4} (T - T_0) Q_0 \quad (5.12)$$

where A_0 is the vent area, u_0 is the initial fluid velocity leaving the vent, T is the temperature of fluid at the vent, and T_0 is the temperature of ambient seawater at the depth of the vent¹ (where T_0 is provided by access to empirical stratification curves provided by external water profiles collected by the science party). Indeed, each of temperature, area, and exit velocity compose a sufficient set of parameters for representing the initial conditions of any particular plume and plume envelope calculation; these quantities, in addition to the mixing coefficients, form the set of \mathbf{x}_p in \mathcal{S} in the plume-charting POMDP. Correspondingly, U_a and the global heading of the crossflow, Θ_a (not directly modeled in these equations, but can be trivially applied to x_a and x_h to convert plume-reference coordinates to global coordinates), form the parameters in \mathbf{x}_c in \mathcal{S} .

¹Temperature is the dominant component of density, ρ , for deep-sea hydrothermal plumes

With the simulator defined, a specific inference problem is posed: from observations of plume or background waters, what are the generating initial conditions (vent area, vent fluid temperature, vent fluid exit velocity) and seawater properties (horizontal mixing coefficient, vertical mixing coefficient, global current heading, and global current magnitude)? To answer this question, probability distributions over \mathbf{x}_p and \mathbf{x}_c can be placed, initialized with uninformative priors $\mathcal{P}(\mathbf{x}_p)$ and $\mathcal{P}(\mathbf{x}_c)$, and the posterior distributions $\mathcal{P}(\mathbf{x}_p|\mathcal{Z})$ and $\mathcal{P}(\mathbf{x}_c|\mathcal{Z})$ estimated².

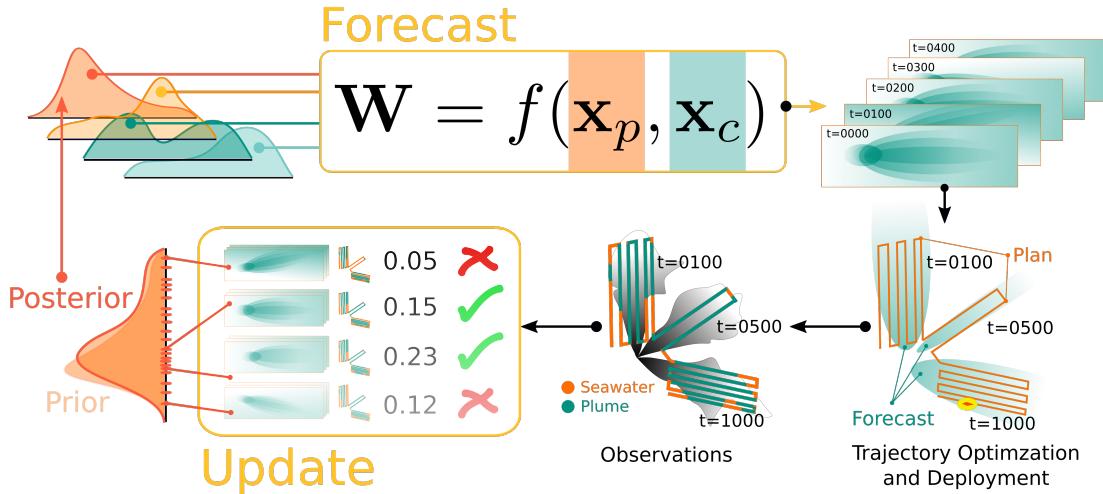


Figure 5.2: PHUMES: PHysically-informed Uncertainty Models for Environment Spatiotemporality. PHUMES is a model for forecasting the evolution of a spatiotemporal distribution trained on partial observations. PHUMES generates forecasts by leveraging an embedded analytical model $f(\cdot, \cdot)$ that approximates the physics-driven evolution of a target distribution. This model is seeded with many samples from distributions placed over initial conditions, physical parameters, or temporal functions (such as \mathbf{x}_p and \mathbf{x}_c here). The composite result of this process is a forecast \mathbf{W} that consists of a mean and variance of phenomenon occupancy in a 3D volume over snapshots of time. This forecast is provided to a trajectory optimizer which sets a deployment trajectory that is executed by a robot. The deployment generates a series of observations, which are then used to update the distributions of the generating distributions via MH-MCMC, which compares the gathered observations with the simulated observations of samples from the generating distributions. The resulting posterior update over the generating distributions is then used for the next planning iteration.

²The inference over \mathbf{x}_p and \mathbf{x}_c can be separated given the observation model available. If instead the sensors were co-located on *Sentry*, inference over the joint posterior $\mathcal{P}(\mathbf{x}_p, \mathbf{x}_c | \mathcal{Z})$ could be done instead.

PHUMES consists of two key phases: forecasting (forward simulation) and updating (inverse problem) (Fig. 5.2). In the forecasting step, samples from the distributions of the initial conditions and seawater properties seed the simulator which is solved many times to create a set of plume-envelope samples in the full state space of the target phenomenon (and trajectory optimizer). Time is discretized over domain-specific key points, and any parameters reliant on time are sampled at those discrete points. The set of composite samples at each time is a “forecast” that is essentially a series of snapshots of the phenomenon. Precisely, PHUMES generates a time-indexed $t \in \mathbb{T}$ composite estimate of the distribution of plume fluid in a 3D volume $\bar{\mathbf{W}}$ by forward simulating time-dependent M samples of the states $x_{p,t}^{(m)} \sim \mathcal{P}(\mathbf{x}_p(t))$ and $x_{c,t}^{(m)} \sim \mathcal{P}(\mathbf{x}_c(t))$ through the plume simulator $f(\cdot, \cdot)$:

$$\mathbf{W}_t = f(x_{p,t}^{(m)}, x_{c,t}^{(m)}) \quad \forall t \in \mathbb{T}. \quad (5.13)$$

The complete forecast \mathbf{W}_t is the robot’s belief b , and since it consists of a set of samples, any statistical measure (the maximum *a posteriori* estimator (MAP), maximum likelihood estimator (MLE), median, etc.) can be computed and used downstream by a trajectory optimizer to approximate the reward function $R([\mathbf{x}_p, \mathbf{x}_c, \mathbf{x}_r]^T, a) = \mathbb{I}[\text{in_plume}(\mathbf{x}_p, \mathbf{x}_c, \mathbf{x}_r, a)] \approx \mathbb{I}[\text{in_plume}(\Phi(\mathbf{W}), a)]$, where $\Phi(\cdot)$ is some statistical computation. For instance, the variance of the forecast \mathbf{S}_W^2 and mean $\bar{\mathbf{W}}$ can be computed and used within an information-theoretic reward function (e.g., upper confidence bound), as contingent on the design of a downstream planner.

After the trajectory optimizer yields a plan, *Sentry* is deployed within several hours of the plan generation (following a safety check by the engineering team). For a single deployment, upwards of 20,000 observations may be available (each deployment is a minimum of 6 hrs in duration, up to 24 hrs, and sensor measurements are logged at 1 Hz). Using the filter described in Section 5.3.1, AUV *Sentry* provides observations

of binary plume detections. Other sensors of opportunity described in Section 6.3.2 provide continuous crossflow magnitude and heading observations. These observations are collated into the sensor model \mathcal{Z} .

At the update step of PHUMES, the probability distributions over \mathbf{x}_p and \mathbf{x}_c are updated from observations \mathcal{Z} . To find $\mathcal{P}(\mathbf{x}_c|\mathcal{Z})$ GP models are trained for crossflow magnitude and heading from external sensor data as referenced in Section 6.3.2. The GP kernel parameters are updated using a maximum-likelihood update following typical procedures (Rasmussen & Williams 2004). For $\mathcal{P}(\mathbf{x}_p|\mathcal{Z})$, a random-walk Metropolis-Hastings Markov Chain Monte Carlo (MH-MCMC) method (Metropolis *et al.* 1953) is used to perform the update. Simulations of deployments are generated by casting the traversed path over solutions to the numerical model $f(\cdot, \cdot)$ seeded with samples from \mathbf{x}_p and \mathbf{x}_c . The output of the simulations is directly compared via a likelihood model with the binary observations of plume waters collected by *Sentry*. To handle binary measurements, the Brier score (Brier *et al.* 1950) is computed over the set of real observations \mathcal{Z} and set of predictive probabilities $\rho(\cdot)$ of the corresponding simulated observations: $\frac{1}{|\mathcal{Z}|} \sum_{i=1}^{|\mathcal{Z}|} (\mathcal{Z}^{(i)} - \rho(f(x_p, x_c)^{(i)}))^2$. In practice, the predictive probabilities $\rho(\cdot)$ are set according to an expected false positive rate and false negative rate for instantaneous sensor measurements established in consultation with instrument experts on the science team; they are set to 0.1 and 0.3, respectively. With the likelihood model applied, an acceptance criteria of the likelihood and evaluated priors over samples of \mathbf{x}_p is defined, and samples probabilistically accepted or rejected accordingly. As MH-MCMC inference method is a chaining procedure, each sample of \mathbf{x}_p selected is informed by the last, and the cumulative distribution of all accepted samples is guaranteed to converge to the true underlying distribution for each of the elements in \mathbf{x}_p for long enough chains. The posterior distribution $\mathcal{P}(\mathbf{x}_p|\mathcal{Z})$ is set as the new sampling distribution for the next forecast to be generated.

While formulated here for a single instance of a plume, it would be trivial to extend PHUMES for inference over or simulation of the generating parameters of multiple co-located vents by jointly updating parameter vectors for each vent. For scalability, choosing a different chaining procedure to accelerate search through a higher dimensional state space at the update step (for instance, Hamiltonian Monte Carlo (Duane *et al.* 1987)), could be advantageous.

5.3.3 TRAJECTORY OPTIMIZATION WITH FIXED PATTERNS

Given the hydrothermal plume-charting POMDP model introduced in Section 5.2 and the probabilistic plume predictions generated by PHUMES, how to select trajectories for AUV *Sentry* that will effectively map the spatiotemporal dynamics of the evolving plume may be considered.

In each deployment, the planner must select an action in the form of a chained lawnmower trajectory (Fig. 5.3). A chained lawnmower is defined by the number, $n \in \mathbb{Z}^+$, of lawnmower trajectories in the chain and the parameters of each individual lawnmower trajectory, $\theta_i \in \Theta$ for $i = 1, \dots, n$. These parameters include the height, width, resolution, origin, and orientation of the pattern and are sufficient to completely specify a lawnmower trajectory. The set Θ is defined to enforce that the lawnmower trajectories are contained within a pre-defined, rectangular safe region and that each lawnmower obeys a time-based budget constraint. As previously mentioned, this constrained action set is dictated by the operational practices of AUV *Sentry*; lawnmower trajectories result largely in *Sentry* traveling in straight lines with few, intermittent turns, which is a beneficial paradigm for the navigational sensors used onboard (i.e., acoustic Doppler Velocity Logger (DVL), inertial sensors). Using lawnmower trajectories has the additional benefit of biasing the vehicle to collect spatially diverse datasets that scientists are accustomed to analyzing.

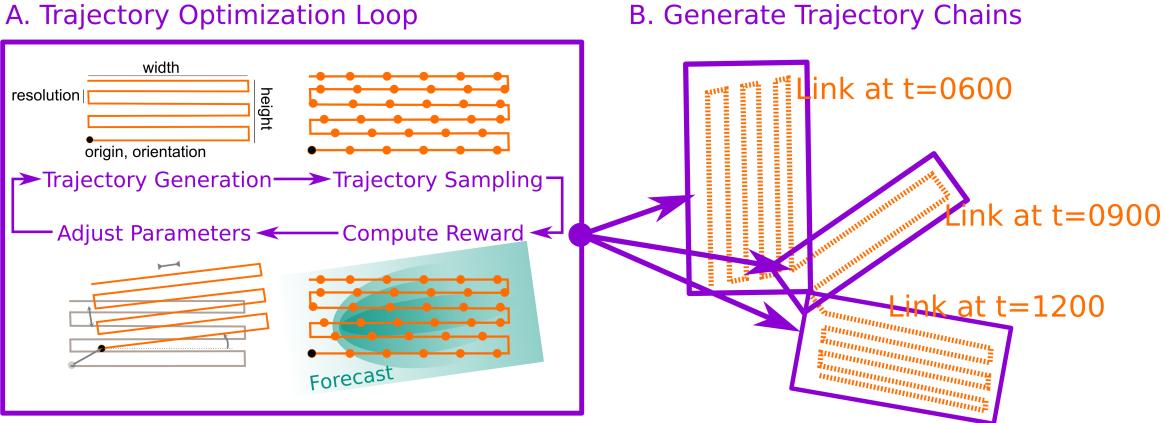


Figure 5.3: Trajectory Optimization in PHORTEX. The trajectory optimizer leverages the PHUMES simulator to select high-reward chains of parameterized lawnmower trajectories. (A) The optimization loop for a single lawnmower object, parameterized by height, width, resolution, origin, and orientation. To evaluate the reward of a specific parameter setting, (1) the lawnmower trajectory object is generated using the specified parameters; (2) the trajectory is uniformly sampled along its length to produce a set of sample locations; (3) the reward of those sample locations is computed using the PHUMES model forecasts of the plume envelope; and (4) the lawnmower parameters are adjusted using gradient-based constrained optimization. (B) This core trajectory optimization loop is used to select parameters for each of a chain of N lawnmowers executed at varying times during the deployment.

The general POMDP value function defined in Eq. (5.3) is reformulated for the elements of the plume-charting POMDP:

$$V_h^*(b) = \max_{\{\theta_1, \dots, \theta_n, n | \theta_i \in \Theta, n \in \mathbb{Z}^+\}} \mathbb{E}_{[\mathbf{x}_p, \mathbf{x}_c, \mathbf{x}_r]^\top \sim b} [R([\mathbf{x}_p, \mathbf{x}_c, \mathbf{x}_r]^\top, \{\theta_1, \dots, \theta_n\})], \quad (5.14)$$

for $h \in [0, H-1]$ and where each $\theta_i \in \Theta$ parameterizes one of the lawnmower trajectories in a length- n sequence of chained trajectories, and b is the planner's belief about the state of the plume, currents, and robot. In this equation, the first of two important planning approximations made by PHORTEX is evident. As mentioned in Section 5.2, the discount factor γ is set to zero, removing the second, recursive portion of the value function. This approximation significantly reduces the complexity of approximating the POMDP value function, allowing each deployment to be optimized myopically. As

deployments of AUV *Sentry* are intermittent, time constrained, and start/end on a ship at arbitrary coordinates, the sequence of deployments are largely decoupled; decisions made in one deployment have very little impact on the achievable reward of the next.

Solving Eq. (5.14) still involves selecting the number of chains n and the joint optimization of all n lawnmower trajectories in the chain. For a standard parameterization of a lawnmower (height, width, resolution, origin, orientation), this results in a challenging high-dimensional, non-convex, constrained optimization problem in which the dimensionality of the optimization problem changes with the number of lawnmowers selected in the chain. In a typical 15-hour deployment of AUV *Sentry* that uses $n = 15$, one-hour chained lawnmowers, this results in a 90-dimensional, non-convex joint optimization problem, which must then further be optimized over the number of lawnmowers n . Optimization is additionally complicated because evaluating the reward function is computationally expensive, requiring PHUMES to produce a prediction of plume probability for the locations sampled by a given lawnmower, and by a lack of analytical gradients for the reward function with respect to the lawnmower parameters (gradients are instead computed numerically).

To simplify the planning problem further to address chaining optimization, a second approximation is necessary. If the number of chained lawnmowers is given, i.e., $n = N$, then is it possible to decompose the joint optimization of all chains in a trajectory into N -independent optimization problems. This approximation breaks a high-dimensional, joint optimization problem into a sequence of much lower-dimensional optimization problems, and is a reasonable approximation if the travel cost between subsequent lawnmowers is not significant.

The final PHORTEX value function, with the two approximations described, is given by the following:

$$V_h^*(b) \approx \max_{\theta_1 \in \Theta} \dots \max_{\theta_N \in \Theta} \mathbb{E}_{[\mathbf{x}_p, \mathbf{x}_c, \mathbf{x}_r]^\top \sim b} [R([\mathbf{x}_p, \mathbf{x}_c, \mathbf{x}_r]^\top, \{\theta_1, \dots, \theta_N\})], \quad (5.15)$$

for $h \in [0, H - 1]$.

Eq. (5.15), which defines multiple, independent, non-convex, constrained optimization problems, is solved using a trust-constrained method in the `scipy` optimization library for a fixed number of iterations (Conn *et al.* 2000). To evaluate the reward function $R(\cdot)$, a trajectory sampler operator is defined $\mathbf{G} : \Theta \rightarrow \mathbb{R}^{3 \times k}$ that takes a trajectory parameter vector as input and produces a set of locations in \mathbb{R}^3 that will be sampled when the robot executes the trajectory, where k is the number of sampled points. In practice, the trajectory sampler \mathbf{G} produces the lawnmower specified by θ and then subsamples uniformly along its length with a fixed spacing. These sample points can then be compared with the plume forecast \mathbf{W} produced by PHUMES to count the number of sample points that are contained within the inferred plume.

5.4 SIMULATION EXPERIMENTS

The performance of PHORTEX is investigated in a simulated environment designed to replicate the field deployment closely. In the simulation, a point robot is tasked with collecting spatially and temporal diverse samples of an advecting plume. Each simulation is a three-dive series in which PHORTEX starts with an uninformative prior over \mathbf{x}_p and executes an initial naive survey (as would occur in a realistic field scenario), then iteratively updates PHUMES with collected observations and uses the trajectory optimizer discussed in Section 5.3.3 to perform two more dives. Ten three-dive simulations for each sampling/planning altitude of 100 m and 150 m are performed in

the same environment. Each single dive in the three-dive sequence is designed to be 12 hrs of simulated time, on a scale similar to the field deployment missions (over 50 acres, or 0.25 km^2).

In the simulator, the underlying analytical model in PHUMES is used to generate a ground-truth environment that closely matches the conditions of a single real-world vent using available data. The simulated environment has vent conditions of 300°C temperature of water at the vent opening, 34.608 PSU salinity of water at the vent opening, 0.8 m^2 orifice area of the vent, and 0.6 m s^{-1} initial fluid velocity at the vent opening. The simulated environment sets the mixing coefficients to 0.15 and 0.2 for horizontal and vertical mixing, respectively. The current function sweeps a generated plume from due North to due East over the course of 12 hours of simulation time, and the magnitude cyclically varies with a beginning and end point of 0.11 m s^{-1} and minimum at 0.04 m s^{-1} . The generating snapshots of the true environment are provided in Fig. 5.4.

In these experiments, the PHUMES model must estimate the vent area, vent fluid velocity, and both mixing coefficients from plume observation in the water column, starting with uninformative priors over each of these targets. A noisy current magnitude and heading function is provided to PHUMES for use in the forecasting and updating step, as would typically be available from sensors of opportunity in the field. For the PHUMES update, 150 samples from an MH-MCMC chain are used to approximate the posterior distributions over the inference targets (this excludes an initial 50 samples of burn-in). In the first simulated dive, the robot executes a best-case uninformed lawnmower trajectory, placed to intersect the sweeping movement of the current. This trajectory is 15 m in resolution, and covers a 500 m by 500 m area, with no rotation. For the second and third dives, trajectories are optimized using PHORTEX for the 12-hour

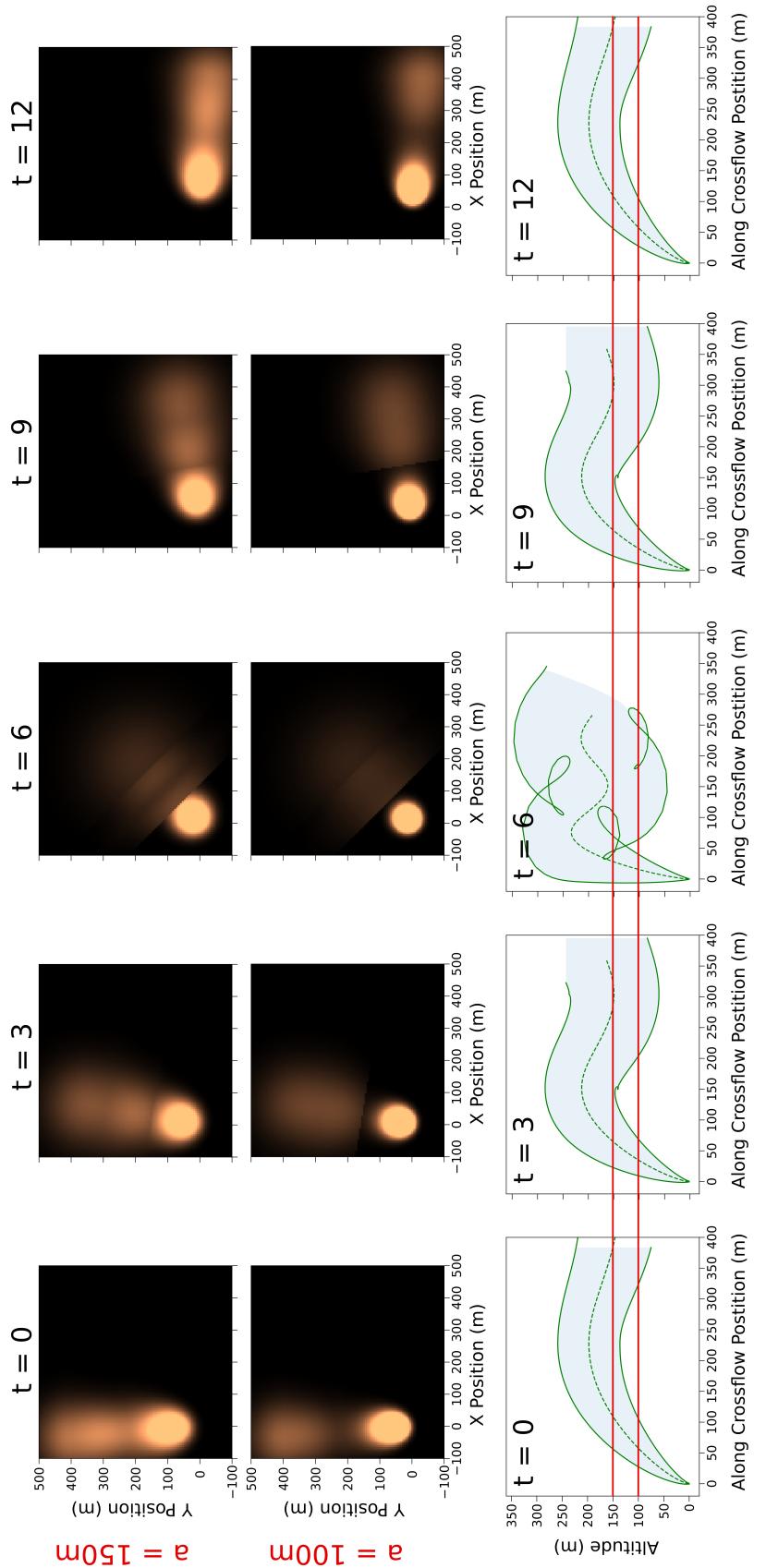


Figure 5.4: Simulated field trial environment. Generated environment for simulation trials at different snapshots of time (in hours) for altitudes of 100 m and 150 m. As the current magnitude and heading changes, the plume expression changes shape and location over 12 hours. Plume intensity is shown in orange in the top plots. The bottom plots show a vertical cross-section of the plume envelope, along the crossflow direction, at different points in the tidal cycle, with the 100 m and 150 m horizontal planes marked.

mission and consist of four, 3-hour long chained lawnmowers³. Each lawnmower in the chain has a fixed resolution of 10 m, and the height, width, origin, and orientation of each lawnmower are optimized to collect the most reward based on a plume forecast using the maximum *a posteriori* (MAP) sample returned by the PHUMES model for the four inference targets. The robot travels at approximately 0.5 m s^{-1} , collecting binary observations every meter that is traveled.

5.4.1 EVALUATION METRICS

The scientific objective of *Sentry* in the field trial is to collect observations of a deep-sea hydrothermal plume that have broad coverage in time and space, and thus are useful downstream for characterizing the space-time dynamics of the plume and other related phenomena (i.e., chemical flux, consumption). To evaluate the performance of PHORTEX for deep-sea plume charting in both simulation and field trials, three key metrics that measure how well *Sentry* collects such samples are:

- **Proportion of positive plume observations:** the number of observations collected in a dive that are classified as in-plume by the binary sensor model (Section 5.3.1). This metric captures how effectively the robot targeted the plume during a deployment.
- **Spatial utilization:** the most distal plume detection and the ratio between the most distal plume detection and the longest distance that the robot traveled from the plume source. This metric captures the spatial coverage of the plume achieved by the robot and the spatial efficiency of the deployment. For example, if detections were made up to 300 m away from the vent, but the robot traveled up to 1 km away, then the survey spent too much time outside of the detectable

³Chain-lengths were dictated by computational constraints placed on the framework to mimic realistic field scenarios, in which only a few hours may be available for plan creation. For more detail on the chaining performance of PHUMES, see Appendix B.

plume region and would not be as effective as a survey that only traveled 200 m away but stayed well within the detectable plume range.

- **Temporal utilization:** the proportion of hours in the dive with at least 10% or more plume detections. This metric quantifies how effective the robot was at *staying in* or *revisiting* the plume over time. Given the long duration of these missions, it is important to use the entire mission window for the task at hand; moreover temporally diverse observations are of scientific interest.

5.5 PHORTEX PERFORMANCE

Fig. 5.5 shows example planned trajectories and Fig. 5.6 shows the distribution of each of the metrics of interest presented in Section 5.4.1—proportion in plume, spatial utilization, temporal utilization—for the three dives at each altitude tested in the simulated trials. These results demonstrate that the PHORTEX optimized trajectories significantly outperform the uninformed baseline, collecting over twice the proportion of samples in-plume, at least doubling the number of hours that the robot spends in a plume, and improving spatial utilization to nearly 100% from 60-70%. For analysis of the convergence characteristics of the optimizer, see Appendix B.

5.6 PHUMES MODEL VALIDATION

The performance of PHORTEX stays consistently high in the second and third dives, suggesting that PHUMES quickly learns a sufficient model for planning from the small number of samples collected by the naive trajectories. To further understand the model learned by PHUMES, qualitatively inspection of the models learned by PHUMES in two exemplar trials are inspected with data collected at 100 m and 150 m altitude during the initial naive survey of the first dive, as presented in Fig. 5.7. In the naive survey,

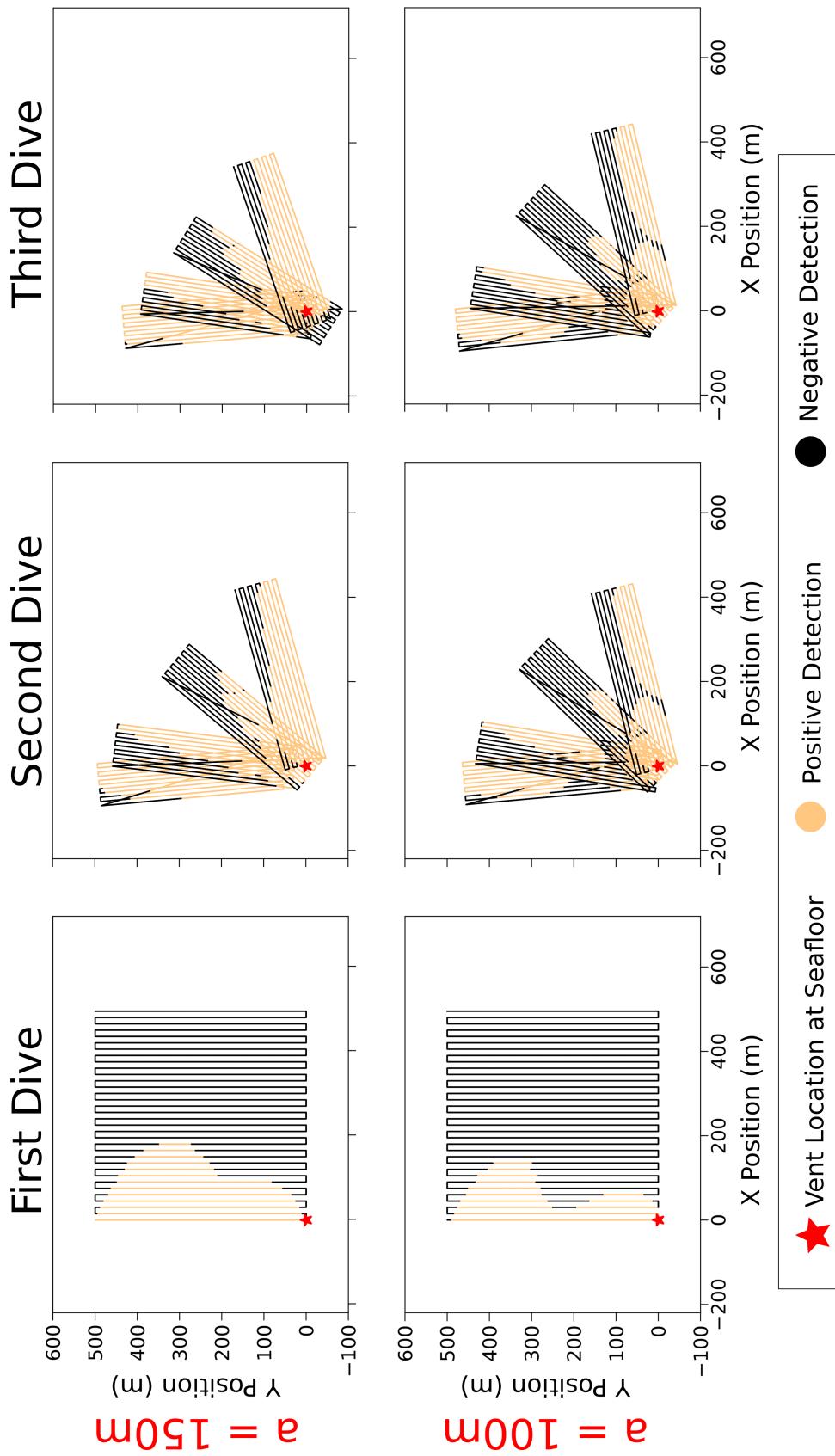


Figure 5.5: **Naive and PHORTEX-designed trajectories.** Trajectory examples for altitudes of 100m and 150m. The first dive is always a naive lawnmower; the second and third dive are PHORTEX designed trajectories. PHUMES is incrementally trained after each dive on the binary plume detections shown in this plot, which are sampled every meter traveled along the trajectory.

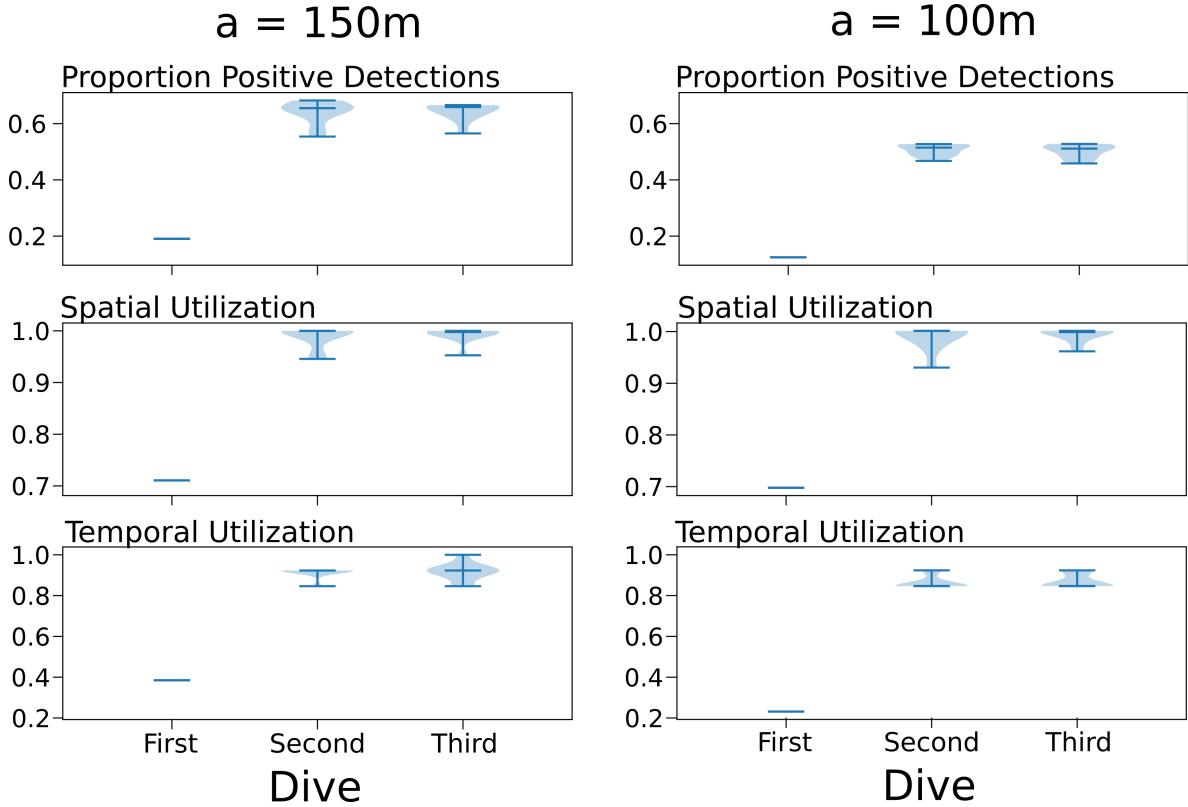


Figure 5.6: Evaluation of the PHORTEX trajectories. The three-dive sequence, consisting of a naive lawnmower followed by two rounds of PHUMES model training and PHORTEX trajectory optimization, are evaluated for proportion of positive plume detections, spatial utilization, and temporal utilization (Section 5.4.1). Note that the y-axis for each plot is in the units of the subtitle labels. The PHORTEX-designed dives show a clear improvement in all three metrics, gathering more spatially and temporally diverse observations of the dynamic hydrothermal plume. Iterative rounds of PHORTEX model-training and trajectory optimization continue to collect a high proportion of scientifically valuable observations.

less than 20% of all detections are positive detections, and all detections only occur in the first 3 hours of the 12 hour mission. At an altitude of 100 m, the robot essentially “skims” the bottom of the neutrally-buoyant plume; at 150 m, the robot is consistently within the range of the neutrally-buoyant plume. Despite these altitude differences, models learned in this example show remarkably similar characteristics—a predicted centerline no more than 25 m off from the true environment’s centerline, and a width that nearly completely envelopes the true plume distribution, which is promising for the context of planning missions. This is in contrast with an illustrative sample from

the uninformative prior, which can arbitrarily produce plume structures that are significantly different in form from the true generating environment. It is also worth noting that positive detections of the plume were made only in the first 3 hours of the 12 hour simulated dive; despite this temporal clustering of detections, the predictive quality of the model forecasts to an unseen time ($t=9\text{hrs}$) remains high. This largely demonstrates the advantage of using an embedded dynamics model to generate predictions of the state space to unseen times.

To quantify the performance of the plume forecast generated by the MAP parameter sample after each trial, the intersection over area (IoA) and intersection over union (IoU) between the true environment and each of the learned models after the naive and first PHORTEX designed simulated dives are computed (Fig. 5.8). A set of 10 parameter samples from the uninformative priors over the inference targets are used to generate a performance distribution representative of the initialized model, to show the breadth of forecast quality before any training. IoA (or recall) provides a number from 0-1 that expresses how many samples predicted by the learned model to be in the plume are in the plume of the true generating environment. This number does not penalize false positives in the model: a value of 1 implies that all points in the model are contained by the environment, and a 0 implies that there are no points contained in the true environment. IoU (or precision) also provides a number from 0-1 that now penalizes false positives: a value of 1 implies perfect alignment between the model and environment, and a 0 implies no alignment. The comparison of these numbers helps to contextualize the performance of model learning.

In Fig. 5.8, the learned models are shown to have a narrower performance window than the baseline samples, and that they generally exhibit very high IoA (up to 1), and a higher IoU (up to 0.9) than the baseline models (up to 0.75). With a high IoA, there is confidence that the learned models are placing predictions in areas where the true

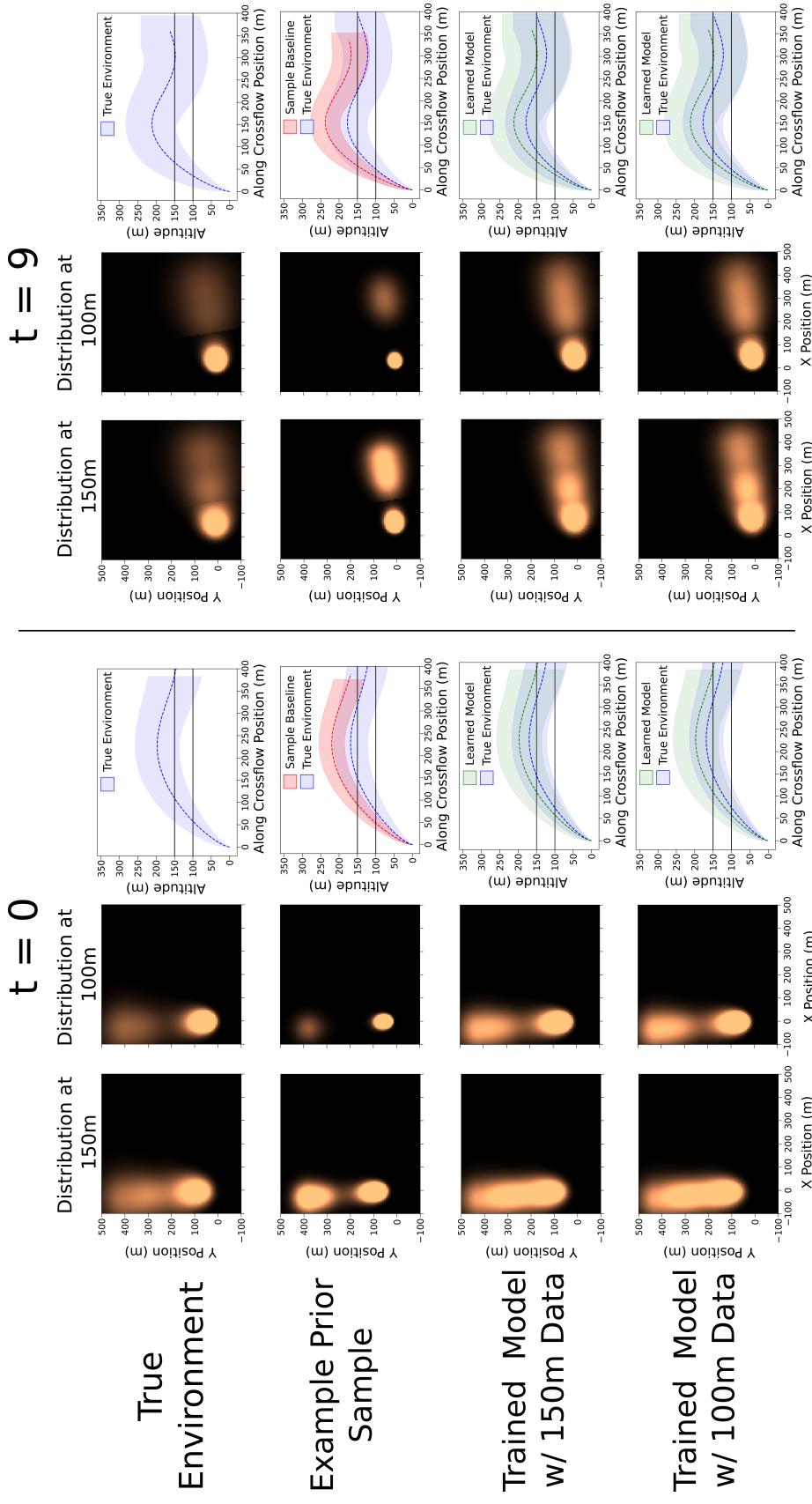


Figure 5.7: Illustration of model learning. Snapshots of the true generating environment are compared with an arbitrary sample from the prior distribution over PHUMES parameters and the learned models from data collected by naive lawnmowers in both 100 m and 150 m simulation trials. In these two exemplar experiments, model learning performance is comparable between the PHUMES models trained on data from different altitudes. The learned model, in comparison to the baseline sample, demonstrates a lower neutrally buoyant stem height, and is wider, and better explaining the data collected at the two sampling heights. Snapshots at different times show that the learned parameters robustly predict future shapes of the plume, even when trained on partial data available from the naive lawnmowers.

plume is present, and with a higher IoU, there is confidence that the structure of the predicted plume regions aligns well with the true environment. Taken together, a very high IoA with medium to high IoU suggest that trajectories planned with the PHORTEX-trained models are very likely to plan for and successfully execute intersections with the targeted plumes, which is advantageous for the scientific task. There is no degradation of performance with different observations available between dives, suggesting that from very little data (a single naive dive), an immediately useful model can be trained.

Notably, there is a distinct difference in the distribution shapes of IoA and IoU between the altitudes across these trials. In particular, training from samples at 150 m appears to be more consistently highly performant (IoA mode is at or near 1; IoU distributions skew towards 0.8) than at the lower altitude, which has a more distributed performance characteristic (with IoA skewed around 0.8, and IoU centered just above 0.6). This has interesting implications for choosing deployment altitudes in practical missions, within the constraints of robot abilities (for instance, AUV *Sentry* cannot swim over a certain altitude and maintain good localization, thus constraining what parts of a plume may be accessible in field deployments). This sets up a direction for future work to characterize the informativeness of different plume regions for model recovery in scientific settings.

5.7 DISCUSSION AND FUTURE WORK

Through this simulation work, PHORTEX is shown to be an autonomy system that can effectively leverage scientific knowledge to enable a deployment-by-deployment mission to chart deep sea hydrothermal plumes. Quantitative gains over typical exploration strategies in terms of number of total in-plume samples, as well as the spatiotemporal diversity of those samples, lends confidence in the ability for this framework to perform in real field settings. As technological advances in robot platforms increase for scientific

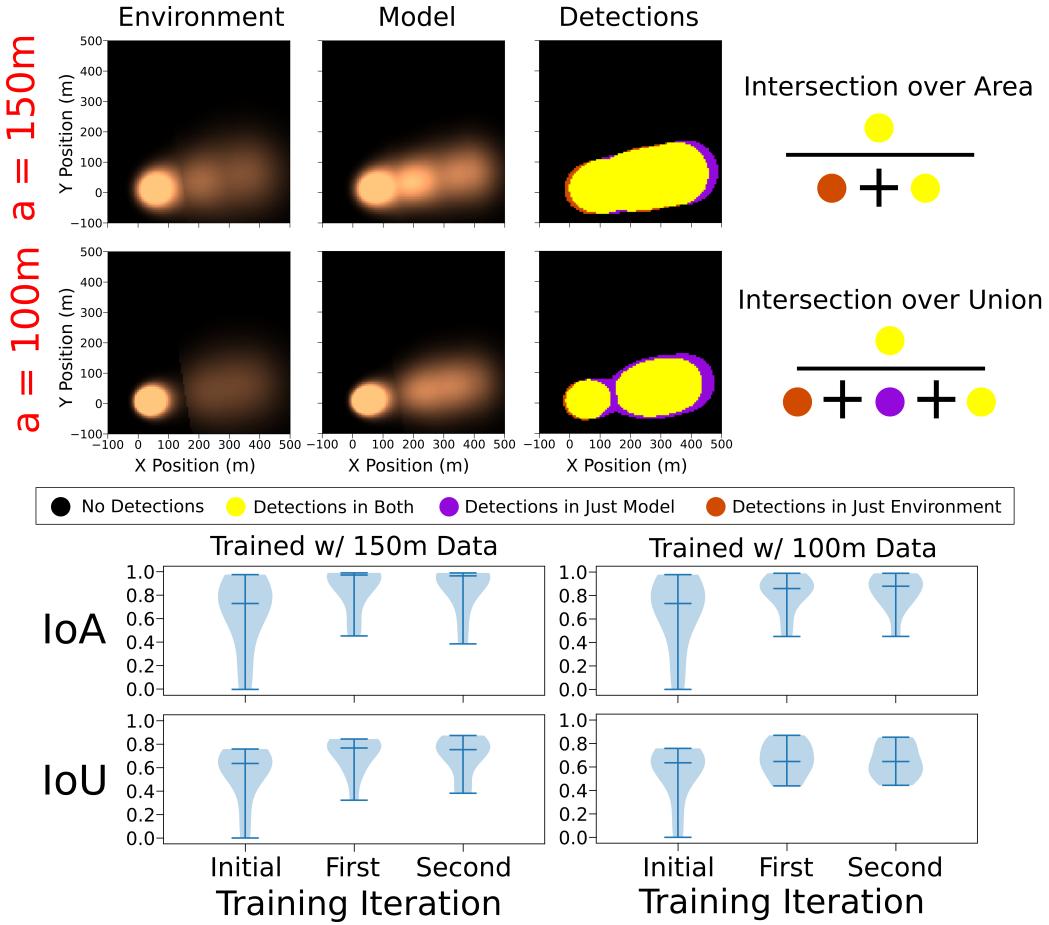


Figure 5.8: Intersection over Area (IoA) and Intersection over Union (IoU) of trained models. For each of trials trained from data at 100 m and 150 m altitudes, the average IoA and IoU are computed (using the method illustrated at the top) for a set of initial model samples (Initial), PHUMES trained on the naive dive (First), and PHUMES trained on the follow-up PHORTEX-designed dive (Second) in simulation. IoA and IoU are both measures from 0-1. IoA does not penalize false positives; a 1 implies that all points predicted to be in the plume in the model are plume samples in the true generating environment. IoU penalizes false positives; a 1 implies perfect agreement between the model and true generating environment. In general, each of the iterative trained dives maintains similar performance; with high IoA and medium-to-high IoU that is consistently higher than initialized samples. Trials trained on data from 150 m tend to have more performant model estimates (IoA near 1, IoU skewed above 0.8) compared to those trials trained on data from 100 m.

contexts, advancing PHORTEX for online settings and adapting multiagent strategies for deep-sea research are attractive next steps. For PHUMES, improving upon the temporal expressivity of the model presented in this chapter, investigating pre-training oppor-

tunities with sophisticated simulators, and adapting state-of-the-art scientific machine learning techniques for a decision-making context, are areas of direct improvement that can be studied. The following sections discuss some of the key challenges in deploying PHORTEX for hydrothermal plume charting, each of which may be avenues for future work.

5.7.1 TEMPORAL RESOLUTION IN PHUMES FORECASTS

The PHUMES forecast provides a sequence of time-averaged plume “snapshots” over which trajectories for a given dive can be planned. By virtue of using the analytical model presented in Section 5.3.2, generating a series of snapshots requires discretizing over time in order to sample a crossflow magnitude and heading, with which the global coordinates of a plume can be computed. This strategy does not capture the effects of advection and mixing on pre-existing (i.e., persistent) plume fluids; the snapshot from $t = 0$ does not influence the snapshot of $t = 1$ because the persistence of plume fluids generated at $t = 0$ are not modeled directly. For the purposes of plume charting in the neutrally-buoyant layer, it could be advantageous to have a more sophisticated model of plume-fluid persistence and/or finer-scale temporal resolution in order to better constrain the spatiotemporal coordinates of a particular observation. For this sophistication to be added, two key innovations would be necessary: a suitable analytical model and a suitable observation scheme.

In settings in which modeling fluid persistence may be useful, a simple advection-diffusion model could be applied on-top of the analytical model used in this work to advect neutrally-buoyant fluids between time discretized snapshots. This would be a simple extension that would introduce another uncertain parameter—rate of diffusion—for inference. Another methodology to explore may be integrating other probabilistic tools to estimate unmodeled characteristics of an environment by the analytical model

(e.g., learned “closure” terms). This methodology may be particularly well-suited in online planning domains, in which forecasts from an analytical model could be used e.g., to set the prior of a GP, and then live observations could be incorporated in real time for course correction while underway. Rather than build upon the model selected for PHUMES here, adapting Gaussian puff models (Ludwig *et al.* 1977) for the deep-sea environment or deriving a minimal set of PDEs to include derivatives in time from full-state models like in Lavelle *et al.* 2013 could be fruitful. In this case, to consider persistence requires additionally modeling non-conservative properties of a plume, such as biological nutrient consumption or particulate deposition. As designing analytical models of these phenomenon is an active area of research, it is obvious that working with domain experts to formulate the right physically-informed model for PHUMES is critical.

The observation scheme for a particular implementation has considerable impact on the sophistication of the inference that can be accomplished. While binary data products are common to compute for vent finding, more finely resolved continuous measurements of the “plume-quality” of a particular water sample could be used. One of the challenges with environmental domains is the access to enough training data to create learned sensors. But perhaps the larger challenge is simply the quality of the data available at large—much of the carbonate and other biogeochemical systems of environmental interest have either a limited or nonexistent selection of (*in situ*) sensors available to measure them. Of the sensors that exist, particularly for deep-sea work, the time-response is on the order of half-an-hour or longer. As the sophistication of sensing equipment improves, so too will the inference abilities of decision-making systems like PHORTEX.

LAWNMOWER TRAJECTORIES FOR DYNAMIC PHENOMENA Lawnmowers and other parameterized trajectories are the foundation of many field robotics deployments. When

used to study spatially-distributed static phenomena, lawnmower trajectories produce intuitive, uniform-coverage maps that are easily interpreted by scientists and domain experts. Studying dynamic phenomena, on the other hand, with lawnmower trajectories can lead to highly counter-intuitive and uninterpretable results. For example, the positive plume detections returned by the standard lawnmower trajectory in Fig. 5.5 only barely resemble the underlying dynamic plume. PHORTEX demonstrates that lawnmower trajectories can still be useful for characterizing dynamic phenomena, when used in combination with probabilistic forecasting models and optimization techniques. However, these methods introduce their own challenges. In PHORTEX, evaluating the reward of a lawnmower trajectory requires generating the trajectory from parameters, sub-sampling that trajectory, and then using the PHUMES model to predict the plume snapshot for a specific point in time and space. Each of these steps can be computationally expensive. To increase the efficiency of the trajectory optimizer during field deployments, time is coarsely discretized and the MAP PHUMES snapshot is only generated once for each lawnmower in the chained trajectories (at the start time of the lawnmower). This substantially improved the speed of the planner, at a loss of targeting accuracy for the moving plume. Developing efficient and accurate techniques that can optimize parameterized trajectories, such as lawnmowers, for dynamic environments is an imperative step in effectively studying spatiotemporal phenomena with mobile robots.

AMBIGUITY CHALLENGES IN INFERENCE Estimating the plume parameters \mathbf{x}_p for deep-sea hydrothermal plumes requires solving an ill-posed inverse problem. The relationship between fluid exit velocity and vent area are significantly entwined in the analytical model proposed in Section 5.3.2 via Eq. (5.10)-Eq. (5.12). There are countably infinite solutions on the two-dimensional manifold describing this relationship for a single target flux value. Ambiguity in inverse problems is a classical problem in

numerical methods, and not easily resolved without setting strong assumptions (i.e., fixing an unknown parameter) or changing the experimental procedure (i.e., collecting more/different types of data). For the purposes of robot trajectory planning, parameter ambiguity is not *necessarily* a problem, as long as the resulting model is sufficient for predicting the plume envelope to strategically place the robot. However, resolving this ambiguity may be important in settings in which the posterior estimates trained by PHUMES are used as a scientific data product in of themselves to make claims about a target environment. To make PHUMES a useful science product (and not just a useful model for planning trajectories), further development that investigates the calibration of uncertainty in posterior estimates and considers possible modifications to the experimental procedure, would be necessary.

5.8 CONCLUSION

This chapter presents PHORTEX: **P**Hysically-informed **O**perational **R**obotic **T**rajectories for **E**Xpeditions. PHORTEX is an autonomy system that plans long-horizon, fixed trajectories to target a partially observable spatiotemporal phenomena by leveraging physics-based dynamical models and Bayesian inference. PHORTEX is motivated by the problem of deep-sea hydrothermal plume charting with AUV *Sentry*, which requires an autonomous agent to map a spatiotemporally evolving plume structure without using underway adaptive capabilities. With these operational constraints in mind, PHORTEX implements a “deployment-by-deployment” autonomy loop (model update, trajectory design, and trajectory execution) for field operations. At its core, PHORTEX consists of a trajectory optimizer and the PHUMES model. The PHUMES model uses an embedded plume physics simulator to solve a Bayesian inverse problem and forecasts discrete snapshots of 3D plume volumes in time. The trajectory optimizer plans parameterized, chained lawnmower-pattern primitives using PHUMES forecasts to create multi-hour

trajectories that maximize expected plume detections. These trajectories are then validated using *Sentry* safety checks and deployed.

PHORTEX was validated in simulation, showing that PHORTEX-designed trajectories can yield obvious gains over naive lawnmower approaches across metrics including total proportion of in-plume samples, and spatiotemporal diversity of those samples. Additionally simulations demonstrated that trained PHUMES models yield insightful and physically-realistic forecasts that closely match an underlying environment, even in times or places that are not directly observed by the robot while collecting training data; PHUMES converges to this good estimate in as little as a single dive in the simulated scenario, agnostic of deployment altitude.

Understanding our dynamic and changing environment is a pressing societal challenge, and algorithmic development in the service of expeditionary science presents many compelling challenges in scientific machine learning, decision-making and the integration of research and practice for field robotics. The contribution of this work is to demonstrate the utility of incorporating domain-specific knowledge into autonomy frameworks for science, provide an example of how scientific knowledge and operational constraints can be formulated into a sophisticated (and deployable) autonomy system, and demonstrate the scientific value of such an approach on the real expeditionary science problem of deep-sea hydrothermal plume charting.

6 CHARTING DEEP-SEA HYDROTHERMAL PLUMES IN THE FIELD

How do you explore the ocean when you're sitting on the deck of a ship, and the average depth of the ocean...is two and a half miles, and we're right there on the surface with these pathetic little tools to try to sample this huge expanse of living blue?

Sylvia Earle

In this chapter, the autonomy tools PHORTEX and PHUMES are used to support a real field operation charting hydrothermal plumes in the Guaymas Basin and perform post-expedition analysis on plume observations. One of the key differentiators between “robot-based experiments” and “field work using robots” is how one can think about the data that is generated. In robot-based experiments, data that is collected is typically designed to assist with assessing the efficacy of a robot platform; the environment may be engineered so the “ground truth” of an environment can be available (e.g., using video capture systems) and robotic measurements may map directly to a metric of interest to easily score performance. In field settings, there is no access to a ground truth, and the sensing equipment on the robot is opportunistic and often a proxy for the true phenomenon of interest. Here, the engineering practicalities of arriving to a useful

and simple sensor model from scientific field data suitable for PHUMES, and deployment logistics for implementing PHORTEX to plan AUV *Sentry* missions is discussed.

An additional aspect of performing fieldwork is that the data product created through robotic deployments is itself a contribution in the context of the science or task that is being conducted. While analyzing the collected data after a field trial can be done agnostic to the autonomy system that collected the data, the PHUMES model directly learns a probabilistic representation over semantically meaningful quantities—fluid velocity at the vent, vent area, and mixing coefficients. There is an opportunity then to investigate the utility of PHUMES for scientific inquiries following a research cruise, and in this chapter several queries are identified that can be supported by the PHUMES representation.

6.1 INTRODUCTION

PHORTEX (with PHUMES embedded), is an autonomy system designed to uncover the dynamics model of hydrothermal plumes from sparse observations, and leverage that dynamics model to plan strategically useful trajectories for AUV *Sentry* to execute during non-adaptive multi-hour dives. In the formulation of PHUMES presented in Chapter 5, it was taken for granted that a binary observation of plume presence was available as a data product. Access to a binary measurement has been generally assumed in vent-hunting or source-seeking literature (Jakuba 2007; Saigol *et al.* 2009; Tian *et al.* 2014). As established in Chapter 4, however, in practice there is no “plume detector” on an AUV, and so this measurement must be approximated from continuous measurements from multiple science sensors. Formulating a filter that identifies plume detections is nontrivial, as variable values of temperature, chemistry, and particulate matter can be found throughout a plume structure. In the buoyant stem of a plume, positive temperature anomalies may be several degrees warmer compared to ambient

seawater; however in the neutrally-buoyant plume layer, these anomalies may be on the scale of sensor noise (only a few hundredths of a degree). Chemistry anomalies may persist longer within a plume, but are subject to unknown and variable rates of microbial digestion or chemical reaction with the environment. Moreover, some chemical signatures may not specifically fingerprint hydrothermal anomalies, and could be equally indicative of other water-mass mixing events in the deep ocean. Finally, some oceanographic properties (e.g., temperature) vary in the water column as a function of depth, and so location of an measurement must be considered. In Section 6.3.1, a method for processing standard scientific equipment on AUV *Sentry* into a binary measurement that is sensitive to each of these challenges is presented that can successfully identify both buoyant stem and neutrally-buoyant plume detections for plume charting.

Separately, binary measurements are not easily informative of other types of structure useful for formulating the PHUMES predictive model. For instance, the current transition function, T_c (Section 5.2.3), could potentially be approximated from binary observations, but it would be significantly more straightforward if point measurements of current magnitude and heading were available. Opportunistic sensor deployments are science party activities which yield data products that can be widely applicable across projects. To measure current, for example, the RR2107 expedition had a pair of current tiltmeters available that could be deployed by ROV *JASON* on the seafloor. The tiltmeters could provide an estimate of current magnitude and heading, just externally to *Sentry*. Section 6.3.2 discusses how PHUMES can incorporate opportunistic science observations from *JASON*, shipboard rosette, and current tiltmeters, acting as a global “aggregator” while at sea to assist in mission design.

The ultimate goal of utilizing PHORTEX is to collect observations that are useful for downstream science tasks. Sample spatial diversity and temporal diversity are particularly key metrics for addressing novel questions about plume manifestation in the

water column: how far do detectable anomalies travel from a vent? what is the vertical structure of the neutrally buoyant layer, and how does it evolve horizontally? how does microbial activity change within a plume structure? These questions have been historically difficult to answer using normal surveying methods, which tend to spatially cluster observations at a vent source, and rarely re-encounter a plume over time. Over four dives with AUV *Sentry*, the performance of PHORTEX is directly quantified with respect to these metrics and compared with reference expert-trajectories in Section 6.5.

One of the key pieces of intuition that enables PHORTEX to perform well at collecting spatially and temporally diverse samples is the use of an analytical model of plume dynamics within PHUMES. The use of idealized models such as Morton *et al.* 1956 or Speer & Rona 1989 for estimating heat (and other energy) fluxes from collected observations at a vent source are common in scientific publications in hydrothermalism (Baker *et al.* 1993; Barreyre *et al.* 2012; Mittelstaedt *et al.* 2012; Ramondenc *et al.* 2006; Wilson *et al.* 1996). While useful, a known limitation of these models is that they fail to consider the impact of crossflow on the observable characteristics of a plume from water column data, strictly underestimating the energetic input of vents. In Section 6.7, the PHUMES model, which considers crossflow since it is strategically useful for sample collection, is shown to be able to perform similar energy analyses, while additionally supporting novel queries about plume structure in the water column.

CONTRIBUTIONS PHORTEX with PHUMES is demonstrated in field trials for hydrothermal plume charting with AUV *Sentry*. In so doing, PHORTEX represents the first iterative offline planning technique deployed for plume charting with deep-sea capable vehicles, illustrating a novel capability for these assets relevant to future research expeditions and putting over 75% of known vent fields (Beaulieu *et al.* 2013) in reach for strategic charting and surveying. In implementing PHORTEX for field trials, a method was developed for processing real *in situ* observations taken by instruments on *Sentry*

into a data product which can be used to indicate whether a particular observation is plume or ambient seawater. The practicalities of using external sensing equipment available during the field trial to benefit the PHUMES formulation is also considered. Through scientific inquiry following the field trials, the PHUMES model is shown to enable energy estimation and neutrally-buoyant plume reconstruction capabilities, establishing an argument for a new paradigm in analyzing scientific data from AUV, ROV, or rosette studies of hydrothermal vents.

6.2 RELATED WORK

6.2.1 TREATING PLUME OBSERVATIONS

Since hydrothermal vents were discovered in 1977 (Corliss *et al.* 1979), studies of hydrothermal vents have richly explored how to describe, measure, and analyze venting structures and the plumes that they produce. Water column detections of plumes are typically shared in the context of vent field discovery and are used to describe the gross characteristics of a venting site (e.g., neutrally-buoyant intrusion height, relative chemical potency, turbidity) (Baker *et al.* 2019a; Caratori Tontini *et al.* 2012; Kim *et al.* 2020). In the discovery context, an important function of water column data is to establish the location of a vent that produces the plume that is observed (e.g., Branch *et al.* 2020; Jakuba & Yoerger 2008; Jakuba 2007), as establishing where a vent is located can be directly compared to the bathymetry of a region, and the characteristic *type* of magmatic activity or hydrothermal formation can be estimated (e.g., on-axis versus off-axis, new or established). Moreover, with a vent location, specialized equipment (like ROV *JASON*) can be deployed for close study of the venting chimney source; estimating the location well keeps these operations efficient and more time can be spent at a site of interest, rather than trying to locate one.

In vent finding, there are two philosophies to approaching sensor models, which are directly tied to the type of autonomous decision-making available. In theoretical odor hunting, or terrestrial/atmospheric applications in which underway autonomy is typically available, a continuous signal of some tracer (e.g., temperature) is generally used. This signal allows a robot to myopically converge to a source (Edwards *et al.* 2005; Mason *et al.* 2020; Morse *et al.* 1998; Reddy *et al.* 2022; Wang *et al.* 2020) or to update an inverse model for source location (Salam & Hsieh 2019; Vergassola *et al.* 2007). While choosing a single tracer is sufficient in this context, when recreating the structure of a plume is a primary objective, it is important to be able to distinguish background water from any type of expression of hydrothermalism.

The second philosophical approach assumes limited or no access to autonomous behaviors, and so every observation that is collected must be considered for vent identification. A popular solution has been the development of occupancy-grid style representations (Jakuba & Yoerger 2008; Peng *et al.* 2014), which are updated from binary observations of whether an observation was plume-derived or not. More confident “occupied” regions are assumed to correspond directly with the location of the venting plume. As the occupancy grid representation does not easily extend to the temporal aspect of spatiotemporal charting, it is useful to revisit the binary sensor model to support the form of PHUMES. In Jakuba 2007, plume fluid from hydrothermal vents is identified using temperature, turbidity, ORP, and vertical velocity anomaly¹. Temperature and turbidity are detrended using reference vertical profiles of each quantity. All sensor streams are then processed twice—once to identify buoyant stem detections, and once to identify neutrally-buoyant plume detections. Buoyant stem detections, more indicative of the location of a vent, are identified using an outlier detection technique (Hampel identifier (Liu *et al.* 2004)) and a consensus scheme that requires multiple corresponding outlier detections across the sensors used to positively classify an observation

¹Disturbances in Z-axis acceleration according to vehicle inertial sensors.

as plume fluid. In contrast, neutrally-buoyant plume detections are identified in each individual sensor stream as any observation outside of a statistical confidence interval placed over the data. The treatment of neutrally-buoyant detections in this manner is problematic for plume charting because the method is not capable of rejecting other types of latent structure that may register as small scale anomalies during surveys—for instance, cold anomalies (driven by water mass mixing) are just as likely to be claimed hydrothermal signatures as warm ones. These limitations are addressed directly in Section 6.3.1 by designing instrument-specific filtering techniques, adding geochemical sensors, and applying a corroboration scheme for all detection types.

6.2.2 AUTONOMOUS ROBOTS STUDYING HYDROTHERMALISM

AUVs are popular tools for geological surveys and mapping of the seafloor; for hydrothermal vent studies, that these platforms can also measure geochemical signatures is largely a bonus (Baker *et al.* 2019a; Caratori Tontini *et al.* 2012; Clague *et al.* 2008; Kinsey *et al.* 2011; Kumagai *et al.* 2010; McPhail *et al.* 2010; Ryan *et al.* 2011; Schmid *et al.* 2019). Development and demonstration of new AUV technologies that specialize in hydrothermal studies (e.g., Maki *et al.* 2014; Okamoto *et al.* 2019) demonstrate the general interest in using AUVs for studying hydrothermal sites and extending the capabilities of these vehicles beyond traditional bathymetric mapping. Adding sophistication to the autonomy of AUVs is one such avenue of active research (Branch *et al.* 2020; Mason *et al.* 2020; Saigol *et al.* 2009; Wang *et al.* 2020). Notably, many of these studies either assume some agency of the AUV to adjust trajectories on the fly (Branch *et al.* 2020; Saigol *et al.* 2009; Wang *et al.* 2020) or are primarily focused on processing AUV data products post-analysis to inform the deployment of other equipment (Jakuba & Yoerger 2008). For studies in which “closing-the-loop” is done via online settings, many were performed in simulation, and if deployed, were primarily done in shallow-water set-

tings with gliding vehicles, which have distinctly different operational restrictions to AUV *Sentry*. This chapter provides a unique insight on the practicalities of enabling depth-capable vehicles to perform hydrothermal charting missions.

6.2.3 USING OBSERVATIONS IN SCIENTIFIC DISCOVERY

In the ocean sciences, considerable attention is paid to *stuff* transfer, either from air-sea interactions or between sediments and the deep ocean, where *stuff* may be energy in the form of heat, chemicals, or particulates (among other things). Whenever *stuff* is transferred, there is a change in the relationship between the two environments, with far-reaching implications to other systems. For instance, the ocean is estimated to absorb up to half of excess atmospheric carbon dioxide emissions from anthropogenic activities (Hori *et al.* 2019; Raven *et al.* 2005). This directly global climate-regulation mechanisms, projecting the rate of acidification of the ocean, and modeling interventions that address atmospheric carbon loads. Understanding the load of chemical and energy influx into the deep ocean from underlying magmatic activity helps to shape models of regulatory mechanisms in the deep sea. This understanding of energy and chemical influx in turn could lead to improved future-looking understanding of how deep sea mining, deep carbon sequestration, and similar policy and engineering activities may impact the deep biogeosphere.

To estimate energetic input, the heat flux of a vent is a useful measurement for estimating the thermal energy of a vent (thermal energy is heat flux multiplied by the area of the source vent). To compute heat flux or particulate transport, stationary models of buoyant plume rise are typically inverted, using observations of the location of a neutrally buoyant plume height or noisy estimates of vent characteristics from in-plume samples (if such parameters cannot be directly observed using an e.g., ROV) (Baker *et al.* 1993; Barreyre *et al.* 2011; Barreyre *et al.* 2012; Mittelstaedt *et al.* 2012; Murch *et al.* 2020;

Ramondenc *et al.* 2006; Wilson *et al.* 1996). The strict assumption that there is no effect of crossflow for estimating vent characteristics can lead to an underestimate of heat and particulate input from a vent, as crossflow tends to lower the observable neutrally-buoyant height, and increase plume entrainment of background seawater (Adams & Di Iorio 2020; Tohidi & Kaye 2016).

By failing to consider crossflow when analyzing water column data, there are also limits placed on the ability to reconstruct the particulate/nutrient flux and transport from a vent, in addition to the vertical-horizontal distribution and longevity of plume-derived fluids. Some studies compensate for this shortcoming by applying a simple advective-diffusive model at the observed neutrally-buoyant height to model horizontal transport (Barreyre *et al.* 2011; Murch *et al.* 2020), however extensive laboratory studies on the impact of neutrally-buoyant intrusions (Richards *et al.* 2014) have shown that there is considerable complexity in the formation and persistence of a neutrally-buoyant plume, and so these simplified models may fail to generalize well outside of specific contexts in which they are applied. PHUMES, which embeds a notion of crossflow directly in its formulation, can enable study of the neutrally-buoyant intrusion and energetic/particulate fluxes, and an example investigation is discussed in Section 6.7.

6.3 METHODS

6.3.1 TREATMENT OF AUV *Sentry* SCIENCE SENSORS

Drawing on the work in Jakuba 2007, the oceanographic sensors on *Sentry* are processed into a binary data product to indicate whether *Sentry* was in a plume or in background seawater. In the autonomy study, *Sentry* was equipped the same instrumentation as described in Section 4.2.2 (i.e., temperature, salinity, oxygen, turbidity, and methane

observations) in addition to an oxidation reduction potential (ORP) instrument² to compute a binary detection of plume water. Sensors are internally logged at variable rates, but interpolated and sub-sampled to a fixed 1 Hz frequency with a shared clock time for the purposes of directly comparing the instruments. Each of the sensors has its own characteristic response to the chemistry of plume water. For example, ORP exhibits a large negative spike when first encountering plume water followed by a slow hysteresis back to a nominal values. Measurements of salinity, temperature, and oxygen are expected to be influenced not only by plume water, but background physical mixing in the ocean (Li *et al.* 2020; Preston *et al.* 2022a; Speer & Rona 1989); turbidity, ORP, and methane are signals strongly associated with hydrothermalism and are generally not persistent in background seawater.

To account for the different ways in which instruments respond to plumes, each sensor is processed individually to detect anomalies (both short-lived and persistent, in the case of neutrally-buoyant plumes) in each stream (see Table 6.1). This process is directly informed by the lessons learned in Chapter 4 and to address issues in detecting neutrally-buoyant detections outlined in Section 6.2:

- Salinity and temperature signals must be detrended due to the nontrivial density stratification in the Basin.
- Salinity anomalies are exceedingly small, even at a vent source, and any anomaly may be indicative of hydrothermal activity.
- Oxygen and temperature are susceptible to other forms of physical water mass mixing in the basin.
- For conservative detections, a detected temperature anomaly should be classified as hydrothermalism only if elevated.

²An ORP probe measures the relative reactivity of a water sample.

- Oxygen structure is complicated in the basin, likely as a result of microbial activity. A rolling window is necessary to detect localized anomalies, and either low or high anomalies could be hydrothermally related.
- Turbidity and methane are significantly and obviously elevated only near hydrothermal venting.

Weights are assigned to each sensor based on their individual reliability for identifying plume water, as determined by the science party and consulted experts in preparation for the research expedition. An observation (a vector of continuous valued sensor observations), is then classified as either plume water or background water by computing the weighted sum of the individual sensor binary detections and setting a heuristic threshold (which serves as an “evidence minimum”). A total corroboration score of 4 or more was used to classify an observation as plume water, which ultimately computes to a single binary measurement for each observation. An example of this sensor applied to *Sentry* detections during an at-sea trial in Guaymas Basin is shown in Fig. 6.1.

The result of the sensor model is to convert multiple, time-stamped sensor observations $s_{t,i} \in \mathbb{R}$, $i = 1, \dots, S$ to a time-stamped binary plume-detection $z_{p,t} \in \{0, 1\}$. These binary plume detections are then used to update PHUMES and plan robot trajectories with PHORTEX, as described in Chapter 5. The accuracy of this sensor model is difficult to characterize, as there is no available ground truth in a field setting by which to verify the assigned classifications. Qualitatively, the classifications were reviewed by the science team and verified for their alignment with expert opinions on which label to assign³.

³Experts use context knowledge of an environment, years of experience, and general heuristics over both single sensor and multi-sensor datastreams in order to classify hydrothermalism. As the binary sensor presented here adapts their knowledge and heuristics into a repeatable computation, the alignment between the sensor predictions and expert predictions is indicative that the computation is effectively capturing their expertise.

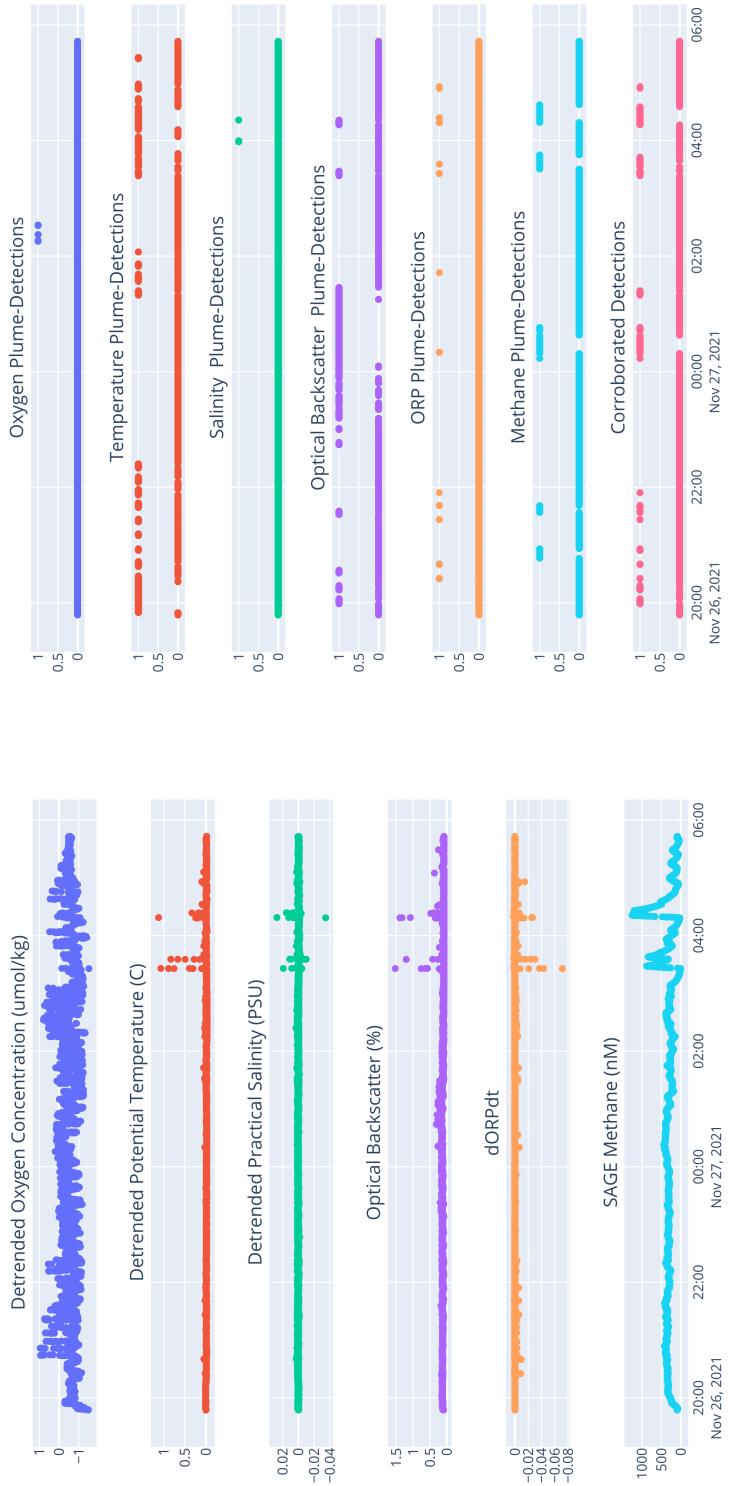


Figure 6.1: Example time series (left) and associated detections (right) over the AUV *Sentry* sensor suite. Note that y-axes are provided in the units indicated by subplot titles. Oxygen, temperature, and salinity measurements are detrended using a linear transformation fit to depth vs. value plots. The time series demonstrates two types of plume detections. The first are “obvious detections” in which most sensors register strong anomalies (this happened twice toward the end of the deployment) and are most strongly associated with buoyant-stem derived fluids. The second are “persistent plume detections” in which the robot traverses through water that is slightly more turbid, warm, or chemical-rich than background water over potentially long horizons (this happened early in the deployment and in the middle). Such detections are most strongly associated with neutrally-buoyant layers. The conservative corroboration detector successfully identifies both forms of plume expressions.

Quantity	Positive Plume Detection Criteria	Weight
Salinity	Detrended practical salinity outside 3 standard deviations of the entire time series in a single dive	1
Temperature	Detrended temperatures above the 75th percentile of entire time series in a single dive	2
ORP	Detections less than -0.005	2
OBS	Optical attenuation above the 75th percentile of entire time series in a single dive	2
Oxygen	Detrended concentrations outside one-hour rolling computation of 3 standard deviations	1
Methane	Normalized concentration above 0.3	2

Table 6.1: **Instruments on AUV *Sentry* and the criteria used to identify plume fluids for each instrument.** The weight is used to indicate relative trustworthiness of a plume detection for each sensor, and is used in a corroboration scheme that sums detections across sensors in order to make a final determination on whether an observation location contained a parcel of plume fluid or consisted of background seawater. Detrended data removes depth-related cross-sensitivity from the measurements; for example, temperature is stratified in the deep ocean, so to ignore the impacts of depth changes in the data stream, those effects are removed by normalizing the data with respect to depth.

6.3.2 OPPORTUNISTIC SENSING EQUIPMENT

Observations from opportunistic external sensors, summarized in Table 6.2 and visualized in Fig. 6.2, were incorporated into the initial conditions, temporal functions, and seawater properties that define the analytical model in PHUMES. PHUMES uses an idealized model of buoyant plume rise under crossflow, the initial conditions of which are the characteristics of the hydrothermal fluids at a vent source. In PHUMES, these characteristics, represented as \mathbf{x}_p , are inference targets. Upon initialization of PHUMES a uniform prior can be placed over each element of \mathbf{x}_p : vent area, fluid velocity at the vent, fluid temperature, fluid salinity, horizontal mixing, and vertical mixing. The bounds on these priors could be informed by previous work or experience by the science team at a given site, consensus in the literature on physically realistic setpoints, or from *in situ* observations of opportunity during other field operations.

Here, access to ROV *JASON* instrumentation was leveraged to set prior bounds on vent area and vent fluid exit velocity, and fix fluid temperature and salinity. *JASON* carries a camera system and temperature wand. Vent area and vent fluid velocity are approximated with images and video captured by ROV *JASON*. Using a 10 cm spaced set of laser points that *JASON* can toggle on and off *in situ*, the vent area is extrapolated from an estimate of vent diameter from pixel-to-distance conversion in still images. Using this method, an area of approximately 1.7 m^2 was estimated for a coherent cluster of many venting orifices, and used to center a uniform distribution over vent area to be updated with PHUMES. Vent exit velocity was estimated by applying particle imaging velocimetry (PIV) (Zhang *et al.* 2019) to 4K video of the turbid fluids at the vent. PIV methods track turbulent parcels that have high cross-correlation values between frames of a video. By tracking many parcels over several frames, PIV yields a vector field of velocity estimates that can be averaged to establish a mean estimate for a region. Using PIVLab, an open-source MATLAB library (Thielicke 2014; Thielicke & Sonntag 2021; Thielicke & Stamhuis 2014), a fluid exit velocity of 0.7 m s^{-1} was estimated and similarly used to set the center of a uniform prior placed on exit velocity for PHUMES. Measuring temperature with an ROV is precise, and so observation of temperature by *JASON*, 340°C , is set as the initial condition for vent fluid temperature in the PHUMES model.

Platform	Instrument	Data Product	PHUMES Incorporation
ROV	Camera	Vent Area	Prior over vent area
ROV	Camera	Fluid Exit Velocity	Prior over fluid exit velocity
ROV	Temperature wand	Vent Temperature	Temperature initial condition
Rosette	CTD probe	Basin Stratification	Reference for buoyancy model
Tiltmeter	Accelerometer	Crossflow	Trained GP for forecasting

Table 6.2: **Summary of auxiliary data.** External equipment and opportunistic data available from other operations during the field expedition that was used to inform the PHUMES model within PHORTEX used for at-sea trials.

In addition to *JASON*, vertical profiles from a rosette of ambient seawater background salinity and temperature were available. Within PHUMES, a reference stratification curve is used to compute depth-dependent buoyancy force. To set this curve in previous simulation work (Chapter 5), a widely accepted set of equations of state was used for Pacific Ocean temperature, Pacific Ocean salinity, and density as described in Speer & Rona 1989. As stratification plays a nontrivial role in the ultimate expected rise height of a plume, access to a better model of stratification for a test site will improve prediction quality. At sea, these equations can be directly approximated for a given water mass using data gathered during standard vertical transects with a rosette. To compute these curves specifically for the Guaymas Basin/Gulf of California, which will be fixed for the purposes of PHUMES, a single rosette profile of temperature and salinity collected early in the expedition were fit by vanilla Gaussian process (GP) models with radial-basis function (RBF) kernels using GPYtorch (Gardner *et al.* 2018) (100 iterations, learning rate 0.1), and the trained mean function of those models was used within PHUMES.

Finally, the crossflow transition function T_c is estimated. Critically there was no sensor on *Sentry* during our expedition that could be used to measure the *in situ* current magnitude and current heading⁴. While it may be possible to estimate T_c solely from the binary observations of the plume, access to an external bottom-mounted tiltmeter on the seafloor during this expedition significantly relieved the burden of this inference process. Using ROV *JASON*, two tiltmeters were intermittently deployed for several days during the cruise, and approximately three days of data were used to compute crossflow magnitude and heading. A maximum crossflow magnitude of approximately 0.13 m s^{-1} was observed, with heading sweeping between north-northwest to southwest. Both magnitude and heading appeared to be semi-cyclic, following a

⁴It is somewhat normal for such a sensor to exist. Acoustic Doppler Current Profilers (ADCPs) are common sensors used with a Doppler velocity logger (DVL) for AUV navigation. Lack of appropriate water column measurements on this cruise was largely a consequence of ongoing vehicle improvement and a yet untested system. For the sake of this work, we can assume that *in situ* point observations of crossflow is generally possible.

time-lagged pattern established by tidal charts produced by Centro de Investigación Científica y de Educación Superior de Ensenada (CISESE) for the time period of the expedition⁵. Time-varying current magnitudes between $0.1\text{--}0.5\text{ m s}^{-1}$ sweeping from the northwest to southwest were previously reported in Scholz *et al.* 2019, corroborating these observations. A GP with RBF kernel was trained for each of current magnitude and heading over hour of the day using GPytorch (magnitude model: 100 training iterations, learning rate 0.5; heading model: 200 training iterations, learning rate 0.1), and the entire trained GP was used within the sampling framework for PHUMES to generate forecasts, allowing for uncertainty in the crossflow functions to be represented in forecast samples.

6.3.3 AT SEA OPERATIONS

PHORTEX was used to enable deployment-by-deployment autonomy during field operations with AUV *Sentry* that fit within the typical workflow of operations at sea (Fig. 6.3). Functionally, the trajectories planned with PHORTEX were provided to the *Sentry* engineering team for extensive safety validation prior to each deployment. If approved by the *Sentry* team, the Chief Scientist, and any other stakeholders, the trajectories were downloaded into the *Sentry* mission planning software as static waypoints. This confirmation process required a lead time of approximately 6 hrs before a scheduled deployment, and 12-18 hrs were generally available between deployments to mechanically service *Sentry* and recharge batteries. The ability of PHORTEX to produce viable trajectories from data within the first 6 hrs that *Sentry* was on-deck following a recovery was critical for keeping this strict timeline. The flexibility in the PHUMES chaining procedure (number of samples to simulate) and the PHORTEX optimizer (optimization steps for a trajectory), allow computation to scale within whatever window is

⁵Charts available from: www.predmar.cicese.mx/calendarios

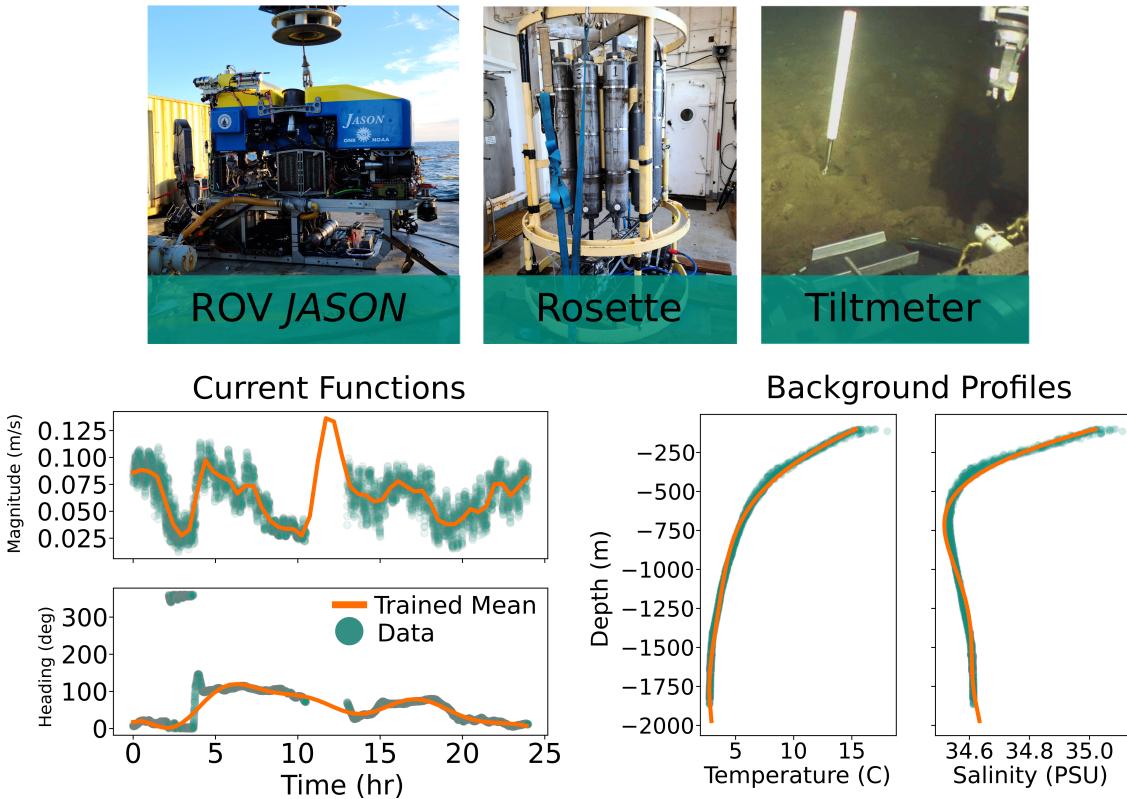


Figure 6.2: **Auxiliary data products used in PHUMES.** External equipment (ROV *JASON*, tiltmeter, and rosette) provided opportunistic data products during the field expedition that were incorporated into PHUMES. The ROV *JASON* was used to determine prior estimates for the plume source parameters. The rosette collected vertical temperature and salinity profiles which are used to compute stratification in the basin. A GP is trained over each of the temperature and salinity data, and the mean is visualized over the data in the lower right panel. The tiltmeter records data of current magnitude and heading; a GP was trained over both functions and is visualized in the lower left panel. Heading is reported in compass-rose orientation. The spike in the predictive mean in the current magnitude function is related to the zero-mean prior assumption of the GP and the lengthscale of the trained kernel in the GP.

available. Given the long lead time between trajectory design and *Sentry* deployment, there were many opportunities for the time of a deployment to change due to e.g., developments in weather, or other science/technology priorities. To be robust to these changes, deployment plans started several hours before and lasted several hours after a given deployment window, and the *Sentry* team could snip the irrelevant links from the chain once a deployment time was known with certainty.

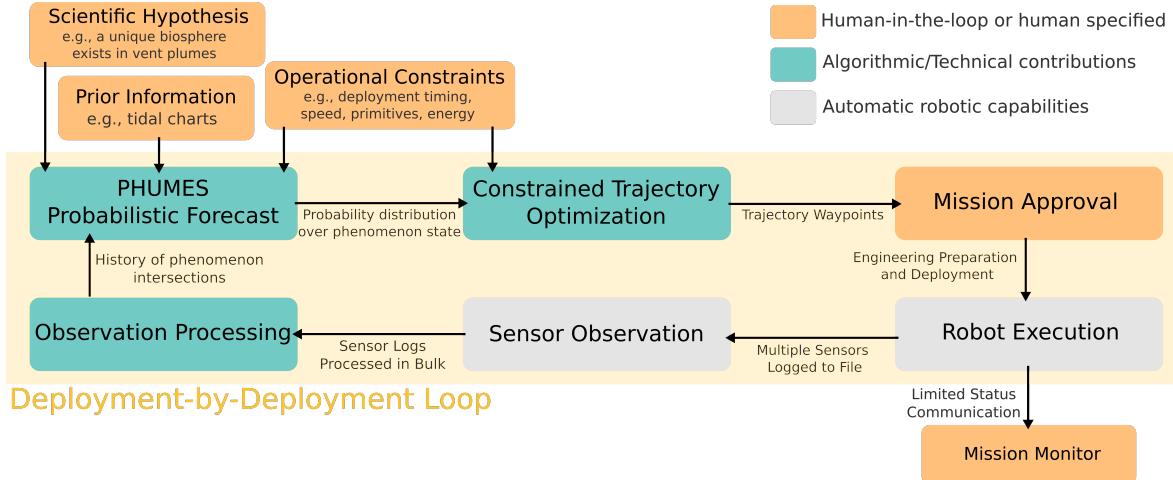


Figure 6.3: **The operational implementation of PHORTEX at sea.** Integration of scientific knowledge, prior information, auxiliary sensor information, and operational constraints was done at the initialization of the PHORTEX deployment-by-deployment loop. Every trajectory generated by PHORTEX was checked by AUV *Sentry* engineers and the science team before execution. *Sentry* status was monitored with an external acoustic tracking system that monitored vehicle location, power, and performance while in acoustic range of the ship. Upon returning to deck, all science sensor observations were downloaded in bulk from the vehicle, and then ingested via the PHORTEX system to plan a new dive.

When *Sentry* completed a dive, it was brought onto deck and within the first hour of being returned, data products from the vehicle were available over the *Sentry* local network. Separate files for each sensor were downloaded, processed in the way described in Section 6.3.1, and used to immediately begin a PHUMES update. In general, waiting until the end of a dive before anything is known about what the oceanographic sensors “saw” while underway is typical. The *Sentry* team usually provides summary reports about the performance of the oceanographic sensors several hours after a dive. This is in sharp contrast with navigational data reporting, which is often updated live on visual monitors while *Sentry* is underway and within acoustic communication range of a ship. Since on-the-fly planning changes are generally not allowable (unless to fix a drifting *Sentry* state estimation), and communication latency is approximately 0.02 Hz, there is not generally support for visualizing the science data while *Sentry* is surveying. For this cruise, the oceanographic sensor data was acoustically relayed to the ship when possible

and a custom live-updating visualization of the key science sensors for hydrothermal mapping was updated. While it did not play a critical role in the operations conducted for PHUMES and PHORTEX, there was considerable excitement and interest in looking at this data. Anecdotally, it was clear that even with significantly sub-sampled data being reported by *Sentry*, hydrothermally driven anomalies could be spotted in the data, potentially making a future for an online version of PHUMES within reach to support ongoing ship-based activities while *Sentry* is underway.

6.4 DESCRIPTION OF THE AUTONOMY FIELD WORK

Testing of PHORTEX during field trials to the Guaymas Basin, described in Section 2.1, largely took place at the Northern Ridge site, targeting a particular venting chimney at decimal coordinates (27.412645 N, 111.386915 W) located at a depth of approximately 1850 m (Fig. 6.4).

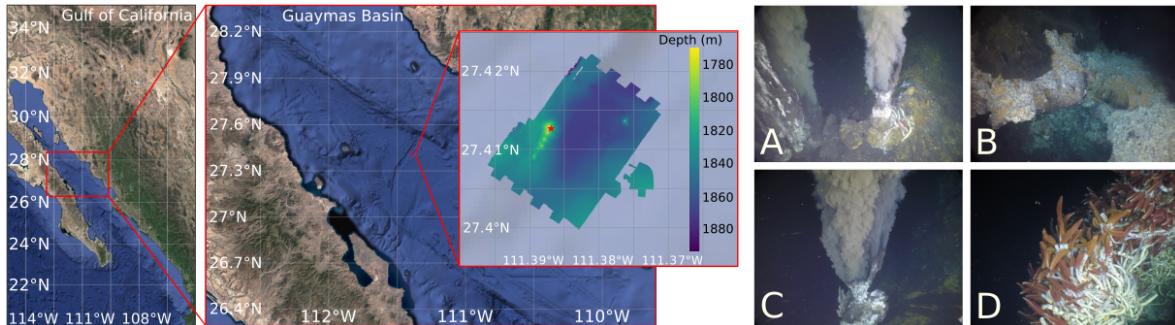


Figure 6.4: **Study site in the Guaymas Basin, Gulf of California.** The inset map is bathymetric data collected by AUV *Sentry* during this expedition and shows the approximately 600 m long ridge in yellow. The red star marks the chimney that was of particular study in this chapter. Pictures A-D show imagery from the ridge and chimney site. A-C show various forms of plume-producing vents located at the chimney and D shows an example of the macrofauna covering the structures along the ridge.

6.4.1 DIVES

Four deployments of AUV *Sentry* were used to study the northern chimney site. These deployments represent a planning spectrum, from fully human-designed surveys to fully PHORTEX designed. The four deployments will be referenced as follows:

- **Dive H-Multi:** human designed, multi-task survey. This was the first deployment of *Sentry* and the survey was designed to both attempt to find plume fluids and to bathymetrically map the local basin area (the map of which would be used as part of the safety check protocol for future deployments). This dive is representative of a standard nested strategy, in which progressively more targeted (finer resolution) surveys are used to study areas of interest. The dive was designed by a human expert who only had access to the approximate location of the target vent. The deployment lasted 21.3 hrs and collected 76,604 observations total.
- **Dive H-Plume:** human designed, plume-charting survey. This was the second deployment of *Sentry* and the survey was hand-designed by the science party onboard the vessel to find and sample plume fluids. The science party had access to the performance of *Sentry* in Dive H-Multi. The strategy was to sweep the basin above areas with known hydrothermal vents, and fly out into the basin in the direction that the plume fluids would be expected to advect. The deployment lasted 21 hrs and collected 75,430 observations total.
- **Dive HP-Plume:** hybrid human and PHORTEX plume-charting survey. This was the third deployment of *Sentry* and the survey consisted of trajectories designed by PHORTEX trained by observations collected only in Dive H-Multi. Two of the trajectory primitives designed by PHORTEX were replaced by “naive” lawnmowers placed over the known vent at two different times in the deployment. The

deployment lasted 22.2 hrs and collected 79,792 observations total. Of these, 8.2 hrs and 29,438 observations were collected via the naive strategy.

- **Dive P-Plume:** PHORTEX plume-charting survey. This was the fourth and last deployment of *Sentry*. The survey was fully designed by PHORTEX using observations only from Dive H-Multi, several days prior to this dive. The deployment lasted 9.9 hrs and collected 35,755 observations total. This deployment is notably much shorter than the other deployments due to increasing time constraints as the expedition was coming to a close. This deployment also used *Sentry* in a “depth-hold” mode: whereas in all other dives *Sentry*’s depth followed the basin terrain, in this experiment the robot held an absolute depth.

6.4.2 EVALUATION OF FIELD DATA

Using the metrics introduced in Section 5.4.1, each of the four dives executed at sea to chart the space-time dynamics of a real hydrothermal plume is evaluated. Each dive took place at different times in the tidal cycle, on different days, and often at different altitudes in the water column, and thus the total plume samples available to collect during each dive is variable. With this in mind, each dive is presented quantitatively and qualitatively, with a special focus on using each dive as a case study of different sampling paradigms. There is no ground-truth available for the deep sea plume-charting problem; each *Sentry* dive is evaluated assuming that the binary detections produced by the method in Section 6.3.1 are honest representations of the presence or absence of hydrothermally derived fluids in the basin.

6.5 PERFORMANCE OF PHORTEX

The results of the field deployment are presented in Table 6.3 and visualized in Fig. 6.5, and demonstrate that PHORTEX performs comparably to science expert-designed trajectories in the proportion of samples that are collected during dives, and importantly improves upon the spatial utilization (increasing both the range of the most distal plume detection and effectively utilizing of the entire explored range). This is most evident in the HP-Plume dive, in which the human designed portion is a lawnmower trajectory placed “on top” of the vent; the PHORTEX-designed trajectory collects samples over twice as far from the plume source. Absolute temporal utilization is similar to human surveys; however the distribution of detections is improved — for human surveys, detections tend to be “bunched” to some window (as in H-Plume in the first half). This is observed most sharply in HP-Plume, in which 90% of positive detections collected by the human-designed survey occurred only in the second of the two lawnmowers, from hours 20-22. In contrast, PHORTEX designed trajectories collected detections more uniformly over the entire PHORTEX portion of the HP-Plume dive. Fig. 6.5 shows the qualitative structure of each dive and showcases the diversity in the resulting datasets.

Dive	Duration	Total Obs.	In-Plume	Spatial Util.	Temporal Util.
H-Multi	21.3 hrs	76,604	22.3%	300 m (19%)	9-17,20-21 (52%)
H-Plume	21 hrs	75,430	10.9%	900 m (64%)	2,5-8,10-11,15-16 (43%)
HP-Plume	22.2 hrs	79,792	41.8%	600 m (100%)	1-3,5,7,11-22 (77%)
(H)	8.2 hrs	29,438	42.3%	250 m (100%)	5,20-22 (49%)
(P)	14 hrs	50,354	41.5%	600 m (100%)	1-3,7,11-19 (93%)
P-Plume	9.9 hrs	35,755	12.8%	450 m (100%)	1,5,8,9 (40%)

Table 6.3: **Per-dive statistics for field trials of PHORTEX.** The spatial utilization is reported as both the most distal plume detection (measured from the plume origin) and the ratio of the most distal plume detection over the farthest distance that the robot traveled from the plume origin. Temporal utilization shows both which hours contain at least 10% positive plume detection and what fraction of the total deployment duration contained such detections. The deployment HP-Plume is broken further into human designed (H) and PHORTEX designed (P) portions for direct comparison.

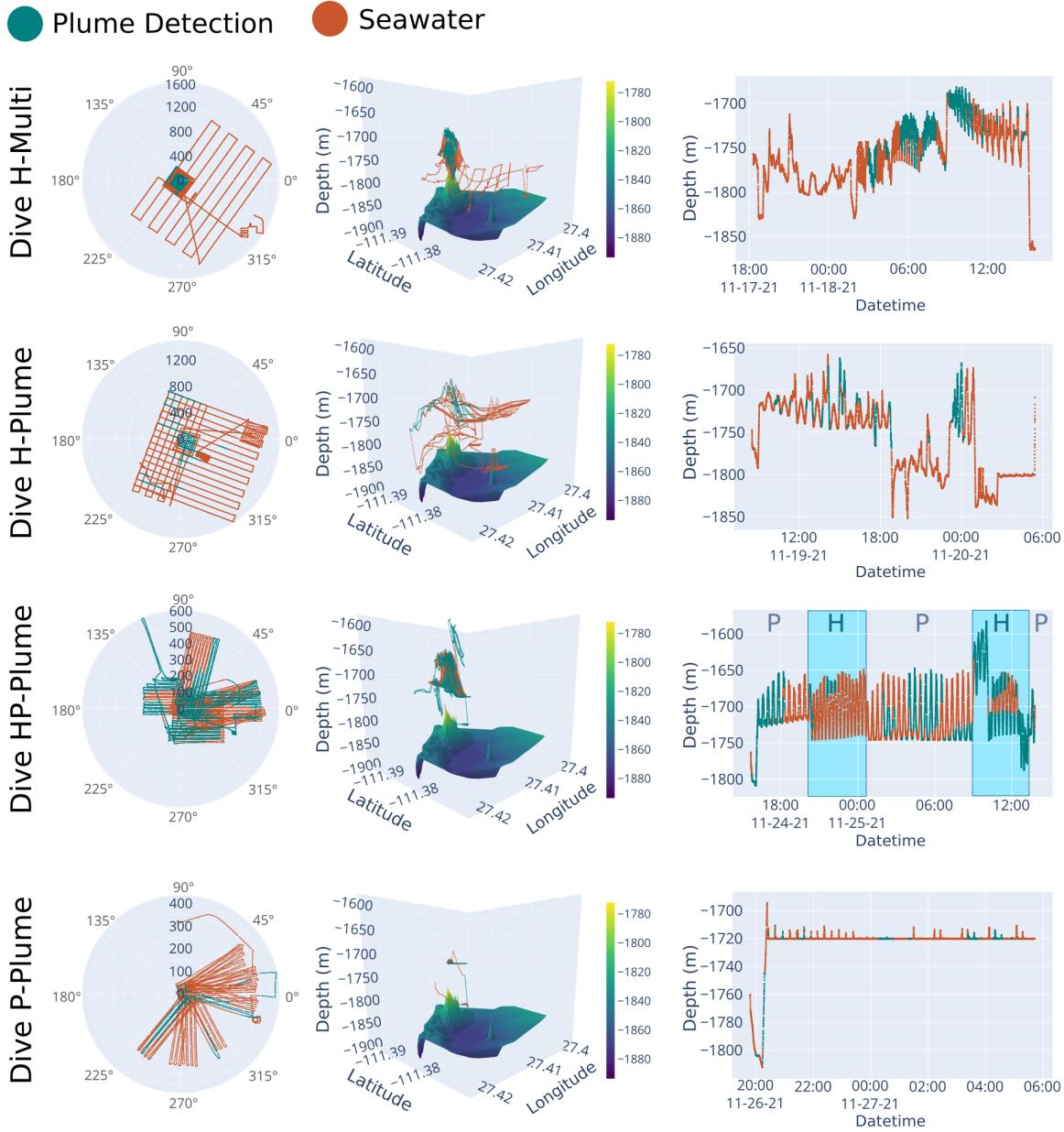


Figure 6.5: The four field dives of AUV *Sentry*. All data is plotted according to its in-plume classification. The first column shows a top-view of the dive trajectories in polar coordinates, in which angle and radius is computed relative to the chimney coordinate of the targeted vent. In the center column, the 3D path of the vehicle over the rendered bathymetric terrain is provided. All but Dive P-Plume were dives conducted in altitude-hold mode with *Sentry*, and so the trajectories show obvious changes in elevation. The final column shows a time series versus depth. In Dive HP-Plume the portions of the dive that were human-designed and PHORTEX-designed are labeled with H and P, respectively. As can be seen in the Dive HP-Plume time series, the two human-designed trajectories have significantly different performance, despite being in locally similar regions of the spatial domain.

For this field deployment, PHORTEX was a useful and practical tool for plume-charting. The performance of trajectories designed with PHORTEX are comparable to those designed by human experts with key improvements in spatial and temporal utilization. It is further worth highlighting that PHORTEX was trained only on data collected during the first dive, H-Multi, and reasonable performance during P-Plume using week-old training data (slightly improved total number of detections over H-Plume given the same initial information, improved spatial utilization, and more evenly encountered plume detections) emphasizes the long-range forecasting ability of the approach. Practically, the automated nature of PHORTEX operationally alleviates significant decision-making burden on a science team and the trajectory-design burden on the *Sentry* team; the ability to ingest data from external sensors and previous *Sentry* missions, and produce trajectories that can be seamlessly ingested by the safety checking system without human intervention is of considerable benefit in the field. Moreover, by virtue of yielding rich context easily interpretable by the science team, the intermediate products of PHORTEX like PHUMES forecasts, are useful for other tasks in field operations, such as deploying other instruments or prioritizing instrument deployment order based on temporal changes in the environment.

6.6 PHUMES VALIDATION WITH BASIN OBSERVATIONS

While there is no external “ground-truth” that can be used to evaluate the performance of PHORTEX, the PHUMES model trained on external and binary *Sentry* observations can be compared with independent snapshots of the vertical distribution of turbidity near the hydrothermal ridge, which can provide a sense of the explanatory power of the PHUMES model. After training, PHUMES estimated the fluid exit velocity from the target chimney to be 0.58 m s^{-1} , the vent area (the opening of the vent from which plume fluids rise) to be 0.82 m^2 , and the vertical and horizontal mixing coefficients to be

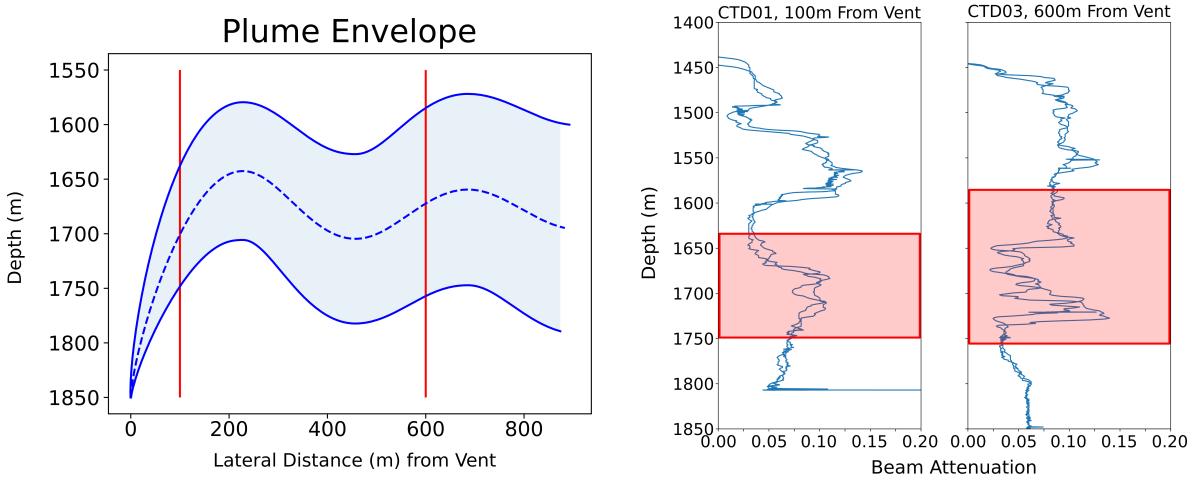


Figure 6.6: Validation of PHUMES model trained at sea. The nominal plume estimate from PHUMES trained on at-sea data and *Sentry* observations with vertical transects of turbidity from shipboard rosette are compared. The plume envelope is the average plume estimated by PHUMES for a nominal crossflow of 0.11 m s^{-1} . Vertical red lines mark 100 m and 600 m laterally from the originating vent, for which vertical CTD casts were conducted. The region that the model estimates containing the lowest plume intrusion in the water column is highlighted on the turbidity transects. There is agreement between the model estimate and the transect data on the location of this turbid layer.

0.15 and 0.19, respectively. Simulating these conditions with an initial vent temperature of 340°C and salinity of 34.908 PSU under a nominal crossflow of 0.11 m s^{-1} , the time-averaged plume height and width are computed. Several vertical profiles of turbidity near vents at the Northern Ridge were collected, and obvious turbid “intrusions” (layers for which there is a significantly elevated signal) were visible. In Fig. 6.6 two vertical transects, one conducted about 100 m from a known vent, and one conducted 600 m from the same vent, are shown and compared with the plume envelope projects by PHUMES. There is good alignment between the model and the lower intrusion visible in the vertical transects. This is indicative that the learned PHUMES model is capable of uncovering the structure of the hydrothermal plume and lends confidence that the model is informative for planning sample trajectories that will intersect with plume fluids.

6.7 PHUMES AS A SCIENCE MODEL

As demonstrated through both simulation (Chapter 5) and these field results, PHUMES can learn an explanatory model of the plume structure in a target environment useful for planning PHORTEX trajectories. Here, the utility of PHUMES as a model for scientific inquiry is investigated, focusing on two key areas: investigation of energy flux and descriptive summary of the neutrally-buoyant plume. Fig. 6.7 summarizes the results of the investigation.

As established in Section 6.2, one of the key reported metrics when studying hydrothermalism is the energetic characteristic of a venting source and the plume itself. Reports of estimated heat flux from hydrothermal vents vary; within the Guaymas Basin, a recent study has shown a wide range of hydrothermal energetic expressions within the soils of the basin (Geilert *et al.* 2018). In the study, specialized equipment was available to probe the soils directly to estimate diffusive flux. For chimney studies, access to similar equipment or ROVs can assist in directly measuring fluid exit velocity and area and the corresponding flux values can be constrained. While access to such equipment was available on the cruise, the formation of the Guaymas chimneys—as many closely clustered, small orifices—make natural measurements of area and fluid velocity challenging, since complicated plume merging is taking place. Instead, it is more useful to understand the *effective* area and fluid exit velocity, which will need to be inferred from water column data despite access to the venting sources.

For water column observations, energy estimates are often computed using a stationary buoyant plume model, e.g., Morton *et al.* 1956 or Speer & Rona 1989, which do not incorporate crossflow. To investigate how PHUMES may be used for similar studies, the relationship between the learned parameters \mathbf{x}_c and \mathbf{x}_p , observations of the neutrally-buoyant layer intrusions in the water column, and energy fluxes (heat, momentum, and buoyancy) must be established. In general, flux at a vent orifice is computed using tem-

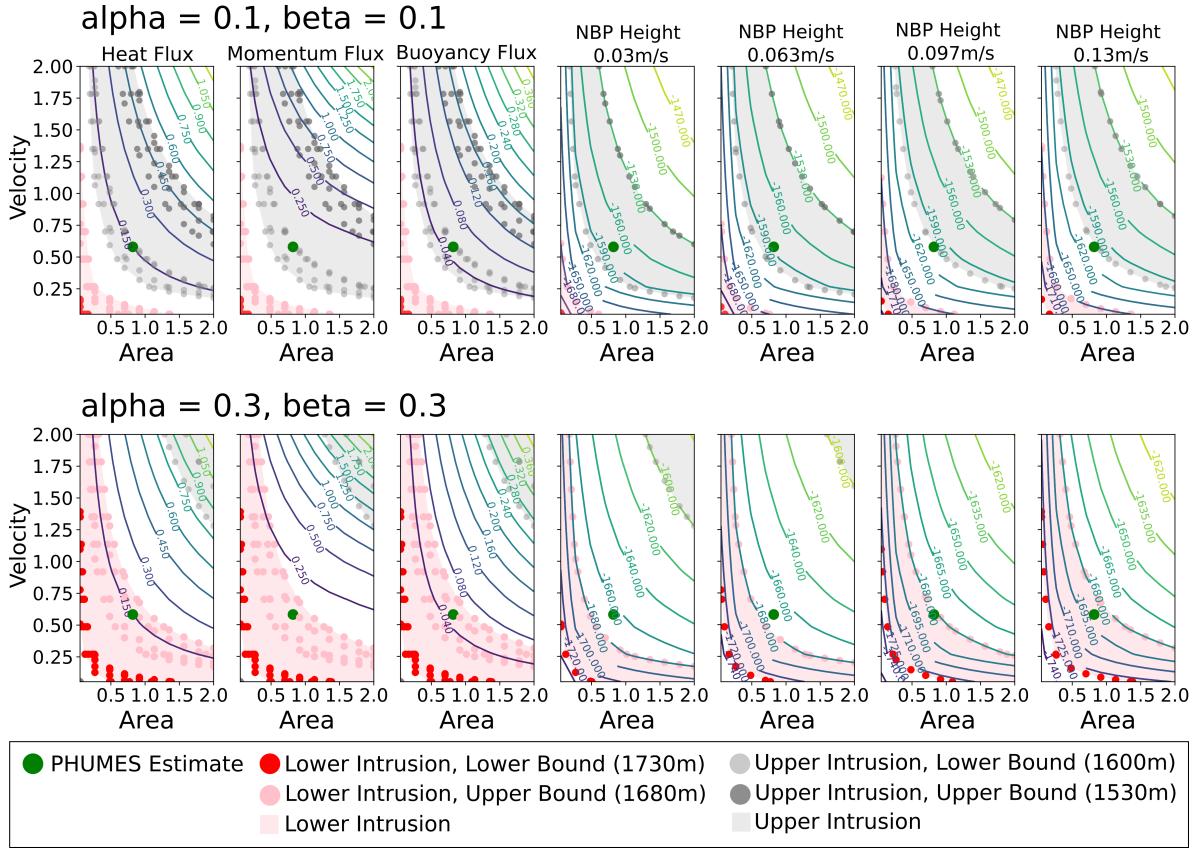


Figure 6.7: PHUMES investigation of neutrally-buoyant plume height and energetic flux. The heat, momentum, and buoyancy fluxes (first three columns) for a hypothetical vent are computed for different vent area and vent fluid exit velocity settings (axes) with two settings of mixing coefficients α and β (each row). Projected onto the flux plots are samples of velocity and area that align with particular neutrally-buoyant height depth values, selected from the vertical transects in Fig. 6.6 that correspond with a deeper/lower intrusion (in pink), and shallower/higher intrusion (in gray). The velocity-area samples are computed for different crossflow values, as specified by the four left-hand columns. As entrainment values directly impact the neutrally-buoyant height and are related to the strength of crossflow, this leads to variation in the estimated velocity-area state space. The green dot in all plots represents the fixed velocity and area estimate learned by PHUMES, as reference. In these plots, it is clear that a single fixed set of mixing coefficients may be insufficient to explain two distinctive inclusions, as the variation with current covers a relatively small margin for both coefficient settings. This implies that there may be alternative explanations.

perature, vent area, and vent fluid exit velocity; these values are agnostic to crossflow and mixing entrainment, as they are estimated to be the values immediately at the vent source. In Fig. 6.7, the contours which describe heat, buoyancy, and momentum flux with respect to a set of vent area and vent fluid exit velocity for a fixed temperature are shown. As is evident, for any flux curve, there is a countably infinite set of velocity and area settings that could satisfy a given flux estimate.

To begin to better explore the solution space, more information is necessary and directly enabled by PHUMES is consideration of crossflow. For the PHUMES model, both crossflow and degree of mixing are linked to the neutrally-buoyant plume height, which is observable by AUV or rosette transects. To understand the relationship between crossflow and neutrally-buoyant height, Fig. 6.7 shows how estimates of fluid exit velocity and area can change with current magnitude and two particular settings of entrainment coefficients. Using the vertical transects in Fig. 6.6 as a reference, two neutrally-buoyant intrusion layers can be identified—a lower intrusion at 1680-1730 m depth, and an upper intrusion at 1530-1600 m depth—and are mapped onto the plots for each current-entrainment setting. To explain either intrusion with a fixed setting of exit velocity, area, and entrainment, would require picking a set of values that can explain all temporal conditions of a particular intrusion region colored on the plots (for instance, any point in the center of the composite gray region could be selected). This selection could be used to estimate a constant flux value that well describes the entire intrusion state space, or to specify confidence bounds on a flux estimate.

What stands out in Fig. 6.7 however, is that there is no combination of area, exit velocity, current, or entrainment coefficients that can fully describe the existence of *both* intrusion layers. This is indicated by the failure of the pink and gray regions to overlap temporally for any fixed setting. If that overlap existed, it may strongly indicate a particular explanatory model of the vent characteristics; being able to tie

complicated vertical structure to energetic contributions of a vent is a unique feature of the type of analysis PHUMES enables. In absence of this overlap in the particular case of the Guaymas vent studied here, attention can be focused towards investigating other explanations for the presence of two intrusion layers: time/current-varying mixing coefficients (for instance, in these plots, temporal overlap between pink and gray regions could occur at low current magnitudes with small mixing coefficients and high current magnitudes with large mixing coefficients); the contribution of external hydrothermal expressions to the particular vent being studied (the ridge had multiple large chimneys at different depths, it is feasible that the intrusion for a different chimney will manifest at a different place in the water column); or unmodeled forcing. To perform these investigations, physical water samples collected by the rosette could be analyzed from the different intrusion layers and their chemical “fingerprints” compared to determine whether they may have been generated by the same vents or to understand their relative age to one another. Continuous observations from *Sentry*, as collected by PHORTEX, could be used to more closely study the lower-intrusion distribution which may hint at possible unmodeled forces (e.g., the plume density is the same as background but still quite warm, which could lead to future density/buoyancy changes that manifest in a change in depth) or help to characterize the statistics of turbulent mixing observed (to inform the entrainment coefficients).

This section showed that PHUMES, in addition to providing a probability distribution over the vent and seawater characteristics, can also serve as a rich test-bed for investigating plume structure. Performing this type of analysis on a ship could assist in designing additional experiments or targeted sample collection that could directly drive at answering the new set of questions posed: How can two intrusions exist in the water column? What are the statistics of turbulence in the plume? What is the character of neutrally-buoyant plume waters relative to the ambient seawater? Answering these

questions would not only drive at more accurate estimates of energetic flux, but also help inform where the shortcomings of the underlying analytical model may be, and over time and further experimentation addressed by formulating a new idealized model (just as PHUMES is a response to consistent shortcomings in stationary buoyant rise models, today).

6.8 DISCUSSION AND FUTURE WORK

There is a significant desire for embodied intelligence and assistive decision-making in environmental exploration and expeditionary science. PHORTEX fundamentally relies on human expertise to inform the scientific models used within PHUMES, generate useful reward functions, set trajectory primitives, and complete operationally safe and robust deployments while in the field. Relieving the burden on these human agents — whether by creating aggregated data products or proposing multiple field missions with explanations — could lead to significant gains in the short-term for expeditionary science tasks while robot technology matures. In the most simple case of this on the research cruise in this study, the real-time visualizer of acoustically transferred science data between *Sentry* and the ship was sufficient for science experts to identify trends in robot performance with respect to the plume charting task. The capability of viewing real-time data, which to many academic and industrial roboticists may seem obvious or straightforward, is not yet pervasive or standard in the sciences or on state-of-the-art vessels and autonomous platforms. Research efforts on improving data infrastructure, data visualization, real-time signal processing, human decision-making, and supervised autonomy promise to be extremely impactful to the expeditionary sciences.

This chapter shows that PHORTEX scales and fits within a practical deployment-by-deployment mission in the field for hydrothermal plume charting. Data collected in the field demonstrated that PHORTEX assists in collecting at least as many positive

detections as science team experts, with gains in spatial utilization and temporal diversity; even on days old data. To implement PHORTEX in the field required careful consideration of how to incorporate science data of opportunity, which is beyond the norm for planning under uncertainty problems and typical IPP solvers. The following section discusses several key areas for expanding IPP in similar ways for real field trials, and concludes with a discussion of how both PHORTEX and PHUMES, as modular frameworks, can serve as a methodological starting point for more general expeditionary science settings.

COMPENSATING FOR ONBOARD SENSING LIMITATIONS OF AUVs Leveraging historical or remote sensing equipment is well established for environmental studies in which satellite, fixed observatory, or historical observations are available for a human expert to use. However, as established in Chapter 2, in oceanic environments, such observational equipment is not widely distributed and often needs to be independently deployed by a science team or by a robot explorer. In this study, a tiltmeter with ROV *JASON* was deployed to compensate for one such shortcoming for observing the deep advecting currents in the basin. One of the key advantages of using the PHUMES formulation, with a well-described analytical model at its core, is the ease by which many different types of information can be incorporated naturally and opportunistically. Black-box belief representations (e.g., neural learners, Gaussian Processes) could struggle to similarly take advantage of asynchronous, opportunistic data without intentionally planning for all possibilities or without considering a modular/hierarchical approach to formulating belief. Access to external equipment to a robot additionally invites advances in decision-making as well, as multiagent (Li *et al.* 2014; Luo & Sycara 2018; Ouyang *et al.* 2014; Salam & Hsieh 2019) or task and motion-planning frameworks could be applied to expeditionary planning at large (Timmons *et al.* 2019).

SCIENTIFIC IMPLICATIONS OF THE COLLECTED DATA Tens of thousands of *in situ* observations were collected in the four field dives that were executed in this study. These data can be directly used in external scientific frameworks for investigating hydrothermalism expressions in the Guaymas Basin. Most directly, *in situ* observations of plume detections further than 300 m will assist biochemists in mapping the fate of biologically digested chemicals and nutrients in hydrothermal fluids that rise through the water column. Coupled with physical bottle samples that the science team collected, the *Sentry* data from these dives will fill in the blanks between the sparse measurements. Given the rarity of scientific expeditions on the scale of cruise RR2107 and the ability to perform targeted sampling enabled by PHORTEX, the data set collected is generally a contribution for the larger oceanographic community. Future work will expand on the investigation in Section 6.7 to estimate the rate of spread of fluid that intrudes into different strata of the water column, ultimately impacting estimates of the overall transport of particulates, chemicals, and nutrients into the larger basin ecosystem, and uniquely supported by the PHUMES model.

EXTENDING PHORTEX AND PHUMES FOR OTHER EXPEDITIONARY CONTEXTS PHORTEX and PHUMES are formulated as modular frameworks, and in different expeditionary contexts each of the design decisions — trajectory optimization scheme, definition of the reward function, and analytical model at the heart of PHUMES— could be replaced directly. PHORTEX is formulated in this work as a deployment-by-deployment sequential decision-making framework that enables offline optimization of operationally-constrained trajectories. Fundamentally, this framework is general enough to extend to any robotic system which may not have access to adaptive behaviors, such as sub-sea AUVs and extraplanetary rovers. For online settings, PHUMES is a model which can support the computation of information-theoretic reward functions and so online belief-based search (e.g., Arora *et al.* 2017; Flaspohler *et al.* 2019; Sun *et al.* 2017; Sun-

berg & Kochenderfer 2018) common in adaptive sampling and IPP literature could be pursued instead. PHUMES itself is a Bayesian inference model that centers around a particular choice for numerical simulator. To extend to other scientific settings, a different numerical simulator can be selected. This requires some initial knowledge of how a particular target environment may evolve; this knowledge could be partial (as in, only knowing that certain properties may be conserved), approximate (as is presented in this article as an idealized model of plume dynamics), or complete (as in, having a full-fidelity simulator of a target environment). Scientific expeditions in the ocean and other marine environments, as well as atmospheric studies, are particularly well-suited for formulation with PHUMES to inform mobile robot trajectories, given the wealth of numerical simulators which exist to describe these environments.

6.9 CONCLUSIONS

Using mobile robots to perform science in the deep-sea is predicated upon both state-of-the-art autonomy systems, such as PHORTEX, and on the effective use of resources within the operational constraints of an oceanographic research vessel. Beyond the core PHORTEX method, this chapter presented a variety of technical and operational approaches to augment the performance of PHORTEX and the science party during field deployments. For instance, a filtering technique was formulated which processed *Sentry* data into a useful binary measurement of “plume occupancy”. Non-robotic platforms and instruments were leveraged to augment the PHUMES learning procedure, which was enabled by using an underlying dynamical model in PHUMES. PHORTEX is modular and can be easily adapted to other domain-specific and expeditionary science tasks. The binary pseudo-sensor can be replaced with any discrete or continuous observation model; the scientific model leveraged within PHUMES could be trivially swapped for another ODE or highly simplified PDE system (well-suited for e.g., ecological/population

studies, fluid or thermal environments); and the reward function and trajectory optimization scheme can be modified based on the operational constraints of a target platform.

During research cruise RR2107, PHORTEX was used in a sequence of four deployments of AUV *Sentry*. The experimental results demonstrate that PHORTEX can collect at least as many plume observations as surveys designed by expert human scientists and showed improved spatial and temporal diversity in samples for any given deployment, even when trained on days-old data. The field deployment marked the first demonstration of an iterative offline planning technique for plume charting with deep-sea capable vehicles. The far-reaching impact of the demonstration is to assert that for the hundreds of deep sea vents only accessible by operationally restricted vehicles like *Sentry*, strategic charting of complex spatiotemporal structure is within reach for future scientific studies. With this capability, novel questions about nutrient transport, water column ecosystems, and the fundamental structure of hydrothermal plumes can be more directly studied. In line with this ultimate goal, PHUMES was shown to be a useful investigatory tool for plume structure and energetic characteristics. This chapter asserts several novel hypotheses for the formation of complex neutrally-buoyant intrusions in the water column at Guaymas Basin informed by this investigation, and established a future path forward for additional analysis.

7 FUTURE WORK IN EXPEDITIONARY ROBOTS

How is it possible for us to know how to wisely use and protect our planet if we do not know what resources we have, their interactions with each other, and their interactions with humans? We need to know and fully understand our ocean so that we may thrive in harmony with nature now and in perpetuity.

Katy Croff Bell

Modern tools in machine learning and artificial intelligence, robotics, and planning under uncertainty have yet to be fully leveraged for ocean and environmental sciences. In a statement to the Subcommittee on Environment of the Committee on Science, Space, and Technology of the U.S. House of Representatives in 2019 (Bell 2019), Dr. Katy Croff Bell identified three key areas for strategic development in ocean exploration: maximizing efficiency of discovery, developing a spectrum of exploration, and attention to big data challenges. Robotic technologies, and roboticists, have the tools to make meaningful impact in each of these areas. AUVs are uniquely well-suited compared to ROV or HOV technologies for underwater exploratory tasks, particularly in mapping and targeted water column surveying. Other autonomous vehicles—such as ocean surface drones, aerial drones, or terrestrial vehicles—can support a diversity of

scientific studies that can collectively link the deep ocean to the troposphere. Modern machine learning is well-positioned to assist in complex analysis of Earth data.

To understand the ocean and Earth, it is imperative that cross-disciplinary collaborative professional, academic, and public projects are undertaken. In this chapter, the possible role of a roboticist on those teams is considered. Starting with a brief overview of challenges which arise in creating autonomous systems for expeditionary sciences, five projects that build on the work presented in this thesis are pitched; three oriented towards advancing the state of the art in robotic technologies and two rooted in open questions about geochemical distributions in natural environments outside of the deep sea. Taken together, these projects aim to illustrate the many ways in which a roboticist (both theoretical and applied) can engage with impactful research towards understanding our planet.

7.1 TECHNICAL CHALLENGES AND OPPORTUNITIES IN EXPEDITIONARY ROBOTICS

In this thesis, an autonomy framework for deep sea hydrothermal plume charting was presented. The constituent parts of the framework—perception, prediction, and planning—each represent rich areas for further development. Here, specific technical challenges related to *belief representations* and *decision-making* are presented for consideration.

7.1.1 BELIEF REPRESENTATIONS

Partial observability and treatment of observations is a core challenge in any expeditionary robotics problem. A belief representation useful for scientific endeavors must

be able to approximate the uncertainty over scientifically-relevant quantities and make forecasts of future environmental states when trained on realistic field data.

HETEROGENEOUS OBSERVATION MODELS Robots used in environmental studies typically carry heterogeneous observational equipment (e.g., point chemical sensors, cameras, acoustic sonar). Optimizing sample collection to address scientific hypotheses requires fusing these different sensing modalities together and implementing complex observational models that link domain knowledge about sensor data to the state of a scientific phenomenon. Embedding expert knowledge into fused observational models, modeling sensor importance to a particular task, and reasoning across different sensors with distinct spatial and temporal resolutions (e.g., Sarkar *et al.* 2014) are all active challenges. A barrier to advancing development of field-oriented science models is a lack of accessible simulation environments and education for non-science experts that sufficiently model phenomenon to a level at which reasoning about sensor models is useful. This is in contrast with other domains in robotics, like self-driving (e.g., KITTI dataset (Geiger *et al.* 2012)), manipulation (e.g., YCB dataset (Calli *et al.* 2017)), or indoor navigation (e.g., Unity, Gazebo, or other simulators). Opportunities to develop e.g., “geochemical playgrounds” for simulation of oceanic and atmospheric environments would be a contribution to the community.

PISTEMIC AND ALEATORIC UNCERTAINTY Reducing epistemic uncertainty of a spatiotemporal environment requires access to a model of the underlying dynamical system, or a data-driven technique that can uncover it. Extracting physically-meaningful quantities from observational data is typically performed post-expedition using computationally expensive numerical models “tuned” by observations. While this lends itself well to Bayesian inference formulations, it is intractable for practical decision-making. Data-driven techniques for model discovery (Raissi *et al.* 2019) may be arguably more

tractable, but generally suffer small-data challenges. Developing models that overcome the challenges of efficiently characterizing spatiotemporal dynamics from streaming, sparse observations would generally improve expeditionary robotics. Additionally, there is a unique opportunity to enable computation of proxies for aleatoric uncertainty, which are well-described in spatiotemporal environments with measures of chaotic motion (e.g., Lyapunov exponents) inferred from data (Blanchard & Sapsis 2019a). The implication that aleatoric uncertainty can be estimated has yet to be utilized to, e.g., assess the attainable resolution of a model or set planning horizons.

SCIENTIFIC KNOWLEDGE AS INDUCTIVE BIAS The kernel of a GP, the loss function in a neural network, or the activation functions between layers in a deep network can all be viewed as forms of inductive bias in a learning problem. For data-driven discovery of spatiotemporal dynamics, improving sample efficiency by leveraging opportunities to inject scientific knowledge to alleviate the learning burden is an open problem. While canonical numerical models of spatiotemporal phenomena are too computationally expensive to directly incorporate into e.g., GP kernels, the physical principles that underlie these models can be more easily summarized. “Physically-informed” data-driven probabilistic representations, have been demonstrated outside of expeditionary robotics (Raissi *et al.* 2019) and initially explored in this thesis with PHUMES. Some additional adaptive sampling work within IPP (Salam & Hsieh 2019) shows rich opportunities for analyzing and extending these methods for larger environments and longer planning horizons.

LOW-DIMENSIONAL STATE EMBEDDINGS Expressing a spatiotemporal environment completely would require an exceedingly large, high-dimensional representation. Model order reduction (MOR) techniques reduce the dimensionality of spatiotemporal systems to a set of weights and vectors that sufficiently describe patterns in the dynamics.

Uncovering low-dimensional state embeddings from partially-observed expedition data is a general challenge (Spantini *et al.* 2018); uncovering a *useful* embedding for a specific decision-making problem is additionally challenging (Pacelli & Majumdar 2019). Access to such an embedding would reduce the computational burden of representing belief in large environments for planning.

7.1.2 DECISION-MAKING

The combination of high-dimensional and continuous state, action, and observation spaces make expeditionary science problems, formulated as POMDPs, challenging even for state-of-the-art solvers. Additional challenges related to the formulation of information-theoretic rewards for scientific hypotheses, robust planning under model-environment mismatch, and decision-interpretability are highlighted here.

ROLLOUT-BASED PLANNING WITH EXPENSIVE BELIEF MODELS State-of-the-art planners for POMDP problems often make use of rollout-based planning in tree search frameworks; continuous search variables are handled using strategies such as progressive widening or scenario sampling (Sunberg & Kochenderfer 2018). However, these planners require extensive online simulations for each rollout performed. Forward-simulating the dynamics and observational models for complex, spatiotemporal phenomena can be computationally intensive, which often limits the feasible look-ahead horizon in real-time operations on computationally-limited robotic platforms. Planners that selectively or adaptively perform expensive rollouts, automatically adjust the planning horizon based on the dynamics of the environmental system, or make use of continuous, offline planners would enable improved decision-making for expeditionary science.

ABSTRACTIONS FOR PLANNING Another promising direction is the development of abstract planning domains for expeditionary robotic problems. Instead of planning

over a set of low-level, continuous control actions, planners could make use of high-level, abstract actions. These planning abstractions may come from human scientists or could be learned directly using recent developments in reinforcement learning and macro-action discovery (Liu *et al.* 2017).

INFORMATION REWARDS AND TASK-DRIVEN EXPLORATION Due to partial observability and stochastic dynamics in spatiotemporal contexts, a decision-maker must operate with significant and often growing state uncertainty. However, not all state uncertainty impacts task performance and uniform information gathering strategies can be inefficient. Understanding the value of information for accomplishing a task is a known challenge for planning under uncertainty and this is particularly true for expeditionary robotics. Recent works that develop heuristic information rewards (Flaspohler *et al.* 2019) or task-driven value of information metrics (Flaspohler *et al.* 2020) begin to build the tools necessary for expeditionary robotic planning.

ROBUST PLANNING UNDER MODEL MISMATCH AND UNCERTAINTY Scientific models, whether data-driven or based on physical principles, are always imperfect representations of a robot’s environment. Model mismatch or uncertainty in key model parameters leads to discrepancies between the environmental predictions that a robot uses during planning and its real-time observations (Singh *et al.* 2018). Planning robot trajectories that entirely miss a phenomenon due to overconfidence in an incorrect model is detrimental to scientific objectives. Planners must develop policies or trajectories that are robust to model mismatch and uncertainty, or are guaranteed to perform as well as a simple, naive data collection strategy.

INTERPRETABLE AND OPERATIONAL DECISION-MAKING Decision-making algorithms must interface with and are constrained by a variety of stakeholders, including scientists, robot operators, and engineers. For example, when deploying an AUV from an

oceanographic research vessel, the decision-making algorithm must account for ship scheduling, timing delays, weather, and multi-platform operations. This requires developing flexible planners that can understand and account for these complex constraints (Timmons *et al.* 2019). Additionally, stakeholders are often concerned with robot safety and data quality. Producing plans that are interpretable for scientific and operational stakeholders is key for building trust and confidence in scientific autonomy.

7.2 PHYSICALLY-INFORMED KERNEL LEARNING

MOTIVATION In order to investigate scientific queries or take informative samples, a robotic vehicle must be equipped with a model. In expeditionary science contexts, the true, continuous, underlying dynamic is generally unknown, and must be approximated from partial observations in order to form a useful belief representation to use in path planning missions. Computing a belief from point measurements is tricky: naive treatment of the data can cause misleading conclusions to be drawn about the underlying dynamic and can ultimately lead to poor robotic behaviors. In this thesis, PHUMES used a Bayesian filtering framework that was placed over a simplified model of hydrothermal plume dynamics. This level of resolution was useful for predicting the location of plume masses slowly advected over a multi-hour mission to assist in sampling through an iterative, offline planning strategy. In settings for which higher spatial or temporal resolution are necessary, however, PHUMES may not easily extend. For instance, in online planning settings, it would be useful to take advantage of faster inference techniques and a more highly resolved temporal model for deciding on sampling actions.

PROPOSED PROJECT Gaussian processes (Rasmussen & Williams 2004) (GPs) have received considerable attention for overcoming partial observability in static scalar fields for environmental sampling (Hitz *et al.* 2017; Marchant *et al.* 2014; Zhang & Sukhatme

2007). GPs place a distribution over continuous functions and have an analytic closed-form update suitable for streaming observations, making the framework well-suited for planning under uncertainty problems. GPs have been used to represent spatiotemporal distributions by using specialized kernel functions that allow time to be treated as an additional dimension (Garg *et al.* 2012; Marchant *et al.* 2014; Singh *et al.* 2010). Although powerful, these kernels are difficult to define and non-intuitive to interpret, which is particularly problematic for science applications in which the model may be used not just by a robot, but by scientists in decision-making during expeditions. Recent work embedding scientific models directly into a kernel structure (Raissi *et al.* 2018) has demonstrated serious scalability challenges (similar to PHUMES) for online inference settings.

To address scalability and interpretability concerns, the proposed project is to develop a *physically-informed kernel learner* (PIKL) that embeds scientific knowledge into the process of designing a kernel for a GP. A desirable learned kernel will be expressive and explainable with respect to a dynamics system, provide insightful uncertainty updates as real observations are drawn, and effectively represent temporal dynamics (e.g., be able to extrapolate.)

BACKGROUND AND CONSIDERATIONS Two promising architectures for PIKL are inspired by state-of-the-art machine learning/scientific machine learning research. *Deep kernel learning* (Wilson & Nickisch 2015) (DKL) is a neural architecture that places a GP as the final layer of an (arbitrary) artificial neural network. *Knowledge-based neural ordinary differential equations* (Jiahao *et al.* 2021) (KNODEs) train a neural “compensator” for expertly identified ODE models.

The advantage of using a neural architecture is one of pattern discovery—neural networks are universal approximators (given enough data), and can be used to discover expressive features that can be linked to task performance (through a loss function, for

instance). In the DKL setting, a neural network is used to transform an observation space into a latent space, over which a GP with a generic kernel can be placed. The GP then provides a notion of uncertainty over the output of the architecture. The kernel can be expressed as:

$$k(x_i, x_j | \theta) \longrightarrow k(h(x_i, w), h(x_j, w) | \theta, w) \quad (7.1)$$

where $h(\cdot, \cdot)$ is a non-linear mapping given by the neural network with parameters w . By maximizing the log marginal likelihood \mathcal{L} of the GP, both the weights w and GP parameters θ can be trained. Using the chain rule, it can be shown that $\frac{\partial \mathcal{L}}{\partial \theta} = \frac{\partial \mathcal{L}}{\partial K} \frac{\partial K}{\partial \theta}$ and $\frac{\partial \mathcal{L}}{\partial w} = \frac{\partial \mathcal{L}}{\partial K} \frac{\partial K}{\partial h(x, w)} \frac{\partial h(x, w)}{\partial w}$ where K is the covariance matrix. Work presented in Wilson & Nickisch 2015 demonstrates that the covariance matrix K can be approximated to have a form which is linear in the number of inputs to manipulate, allowing for considerable computational speed-up over typical strategies that use GPs, which is attractive for robotics applications. Additionally, it was demonstrated in Al-Shedivat *et al.* 2017 that DKL can support neural architectures that learn temporal relationships and recurrent structures in data, which is relevant for modeling natural spatiotemporal environments.

To develop PIKL with a DKL backbone, one of the key design choices is the preceding neural architecture before the GP layer. In general, it would be advantageous to both use scientific knowledge of a system to pre-train PIKL in simulation and to embed scientific principles into the neural network framework. The latter is important because once out in the field, the entire architecture will be tuned from data and it would be useful to place some principled constraints on the way in which the architecture changes for data efficiency and explainability reasons (e.g., useful to have a strong inductive bias). Physically-informed neural networks (Lu *et al.* 2020; Mohan *et al.* 2019; Raissi *et al.* 2019) have been shown to improve data efficiency at training time in scientific machine learning works, but have yet to be demonstrated in field contexts, leaving a

large opportunity for this project to experiment with and extend these architectures for practical expeditions.

Designing the PIKL architecture requires both consideration of how it can learn an underlying function and how it can operationally be deployed during field missions. A DKL-architecture offers some amount of flexibility in its formulation for deployment experimentation. For instance, the neural layers could be “frozen” while a robot is underway (as in, the weights are fixed to a pre-trained value) while the kernel parameters are adjusted online to fit realistic observations as they are collected; following a deployment the entire architecture could be bulk-updated from all observations. In yet another experiment, the neural layers and the GP layer could be asynchronously updated during long-term missions.

In contrast to DKL, KNODEs do not natively support a representation of uncertainty, but explicitly are designed to embed expert knowledge (i.e., scientific models) into a neural architecture. KNODES make use of vanilla neural ordinary differential equations (Chen *et al.* 2018) (NODEs) are deep neural network models that use the intuition that the relationship between residual layers in a residual neural network (ResNet) is identical to an Euler discretization of a continuous dynamical system. This allows a chain of residual blocks in a ResNet to be equivalently viewed as an ODE initial value problem, and so a blackbox ODE solver can be used to recover the solution of the last layer. The advantage of this formulation is that the chain of residual blocks no longer needs to be rigidly defined; a differential solver can evaluate the hidden unit dynamics at arbitrary resolution based on a desired accuracy of the result. Even more generally, this means that any system which can be “truly” modeled as an ODE can be learned using a ResNet.

The correspondence of NODEs and classical numerical modeling makes the NODEs representation attractive for studying dynamic environmental systems. Since formu-

lation, NODEs have been used as a forward simulator for reduced-order models of hydrodynamics (Dutta *et al.* 2021), detonation in engines (Koch 2021), and turbulence forecasting (Portwood *et al.* 2019; Shankar *et al.* 2020). Technical extensions for NODEs to learn stiff systems (Kim *et al.* 2021) and nonhomeomorphic/noncontinuous flows (Dupont *et al.* 2019), among many other variations, have also extended its capabilities to more than just “true” ODE systems. As NODEs are still a general learning framework, complex dynamical systems are still difficult to recover in a data efficient way; leveraging “expert ODEs” within NODEs would be useful to apply a physically-informed inductive bias during training and testing, and this is what inspired KNODEs.

In brief, a KNODE is a model $\hat{f}(\mathbf{x}, t, \tilde{f}(\mathbf{x}, t), \theta)$ where $\mathbf{x}(t) \in \mathbb{R}^n$ is an n -dimensional state vector for some time t , $\tilde{f}(\mathbf{x}, t)$ is expert knowledge of an environment (e.g., an ODE model of fluid flow), and \hat{f} is a function approximator for a true underlying function f , parameterized by θ . How \tilde{f} is incorporated is not rigorously restricted, however in many publications leveraging the framework a linear coupling of the output of \tilde{f} and a generic neural network is used. Using a neural network to compensate for unmodeled dynamics in \tilde{f} is notionally similar to *closure modeling* common in computational fluid dynamics, in which turbulent correlations are “added back in” to Reynolds- or time-averaged simulations that typically only model gross structures (Durbin 2018).

In robotics contexts, KNODEs have been utilized for augmenting model predictive controllers for multirotors, using a small number of training trajectories. In these settings, an expert model of the multirotor dynamics was used as expert knowledge (Chee *et al.* 2022). While the training and testing were performed separately in these studies, an obvious advantage of the KNODEs framework is that “falling back” to the expert knowledge may be a reasonable strategy (in non-control contexts; for instance in estimating the dynamics of an environment), and could enable asynchronous training of the neural layer online.

While KNODEs effectively learn to compensate for unmodeled dynamics when expert knowledge is either incorrect or incomplete, this is not necessarily equivalent to learning an uncertainty model that can be leveraged for informative planning. For instance, a KNODE can express how much a model may be under performing in a control task, but not necessarily *why* nor provide relative weights on the source of the error that could be mapped to the observational space. This leaves open an opportunity for future development. Imagined directions for extending KNODEs for use in environmental adaptive sampling could be (1) changing the coupling mechanism of a neural net and the expert knowledge and (2) placing a GP layer into a NODE formulation at the compensator. Changing the coupling mechanism between the neural network and expert knowledge will change what the neural network is compensating for. For instance, using the network to explicitly learn coefficients for the expert model is significantly different than using the network to learn a function that adjusts the entire output of the expert model. When then coupled with a GP layer at the compensator, for example, notions of uncertainty can be tied to ways in which the neural network is performing compensation and what the implications are for task performance, useful for downstream planning tasks.

7.3 MEASURE-INVARIANT SUBSPACE LEARNING

MOTIVATION In some environmental settings, the expert knowledge which may be available for incorporation into PIKL may be too computationally demanding to use effectively. This especially arises in fluid flows, in which systems of PDEs that are numerically stiff are often found. However, applying a generic belief representation over the state space for a practical environment is also expensive, due to the high dimensionality of the data (in the sense of having tens of thousands of *in situ* measurements

of several types of instruments, and trying to represent the state of every voxel within a 3D volume over time).

To reduce the dimensionality of the inference problem for hydrothermal plume physics in this thesis, a simplified time-averaged approximate model was used instead of high-fidelity expert model. When well-known simplified models for a system do not exist, or the simplified model approximation unsatisfactorily compromises on spatial or temporal resolution for a task, then leveraging numerical data reduction techniques may be useful. Performing model order reduction (MOR) generally requires time-series samples of a generating high-fidelity model, which are then summarized as a set of “modes” and “weights” that represent energetic features of the data. A summary of many common MOR techniques is provided in Chapter 3. Unfortunately, MOR techniques tend to be ill-suited for representing nonlinear systems, smoothing away local structures (e.g., spatially varying diffusion coefficients) that may be relevant for adaptive sampling tasks. A good reduced order model would be able to reflect both global and relevant local information that is useful for computing task-relevant measures of uncertainty.

PROPOSED PROJECT One perspective on the form of a low-dimensional subspace is that it should exhibit *measure invariant* qualities; that is, a statistical measure in the high-dimensional space should have correspondence to a measure in the low-dimensional one. Measure-invariant dynamic systems are defined by the tuple (X, \mathcal{B}, μ, T) where X is a set, \mathcal{B} is a σ -algebra over X , $\mu : \mathcal{B} \rightarrow [0, 1]$ is a probability measure, and $T : X \rightarrow X$ is a transformation which preserves μ . By the Krylov-Bogolyubov theorem (Kryloff & Bogoliouboff 1937) (which places some assumptions on the form of X and T), an invariant probability measure is admissible and recoverable from transformations on the data. Considerable work has shown that this theorem can be applied to samples of trajectories drawn from systems of differential equations that describe natural phenomena (Stannat 2011; Touzé *et al.* 2021; Vizzaccaro *et al.* 2022). The implication of this is that a transform

exists which could cast high-dimensional states onto a lower-dimensional manifold and which can preserve probabilistic measures that could be leveraged in a real-time (or practical-time) IPP sampling problem.

The proposed project is to develop a measure-invariant subspace learner (MISL) which identifies the transform (with slight modification on former notation) $T : X \rightarrow Y$ that relates measure $\mu \in \mathcal{P}(X)$ and measure $\nu \in \mathcal{P}(Y)$ such that:

$$\nu(B) = \mu(T^{-1}(B))$$

for all ν -measurable sets B .

BACKGROUND AND CONSIDERATIONS Learning invariant subspaces has been proposed in several domains. Among the most common approaches is *similarity preserving embeddings* which is an implicit method that uses a kernel to map data in a Euclidean space into a Hilbert kernel space (Schölkopf *et al.* 1997). Designing good kernels is relatively challenging, and some argue that since the method is implicit, temporal patterns are either not captured or are not interpretable in the subspace (Deng *et al.* 2020). *Shift-invariant representations* can be produced using sparse coding and dictionary learning methods which enable temporal trends to be captured, but fail to capture local structure in data (Lewicki & Sejnowski 1999). Finally, *time alignment* (Deng *et al.* 2020) aims to minimize error between pairs of samples which admits temporal trends and local structure, but requires a distance measure which may be difficult to represent or optimize to be defined.

The challenges in choosing a good formal method invite creativity in finding an invariant subspace. One promising technique that could be targeted by the MISL project comes from within Bayesian inference. Transport maps and optimal projectors have been used to compute subspaces suitable for solving inverse problems for dynamic sys-

tems (Cui *et al.* 2014; Spantini *et al.* 2018; Spantini *et al.* 2015). Computing these operators can be difficult, however work in *approximate* optimal transport using Bayesian inference (Bigoni *et al.* 2019; Spantini *et al.* 2019; Spantini *et al.* 2018; Villani 2008; Zahm *et al.* 2022) shows significant computational savings and advantages over other MOR techniques in terms of expressivity and data requirements. To uncover a good transport map that has measure-invariant qualities requires identifying the statistical measures that ought to be retained. In expeditionary science problems, the sampling objective generally is motivated by scientific hypotheses, which may focus on different features of a spatiotemporal distribution. For example, monitoring chemical sources in an environment only requires information about where the sources are, whereas quantifying gaseous flux additionally requires information about the temporal characteristics of each source and where emitted chemicals are advected. The “optimal” low-dimensional subspace to be uncovered by a map will target only those features necessary for computing the target metric.

Task-aware, or goal-aware, subspaces can be seen as a specific instance of invariant subspaces, in which, for a task-relevant measure, the lowest-dimensional invariant subspace *is* task-aware. Goal-aware learning was recently demonstrated for robotics problems (Nair *et al.* 2020) in which a learner was rewarded for computing a subspace that best captured features relevant to a specific task that was to be later executed by a robot (offline and using full observations). By using task-relevant metrics to additionally guide subspace identification through approximate map building, structure can be injected into the learning problem making it possibly more tractable and more generalizable when utilized in field settings with real data.

7.4 ENVIRONMENT PREDICTABILITY AND PLANNING

MOTIVATION To gather useful observations of a spatiotemporal environment for a given task in a finite mission duration requires nonmyopic planning. State of the art IPP strategies typically employ belief-space planning trees, in which a robot’s belief about the environment is used to simulate potential future outcomes of action sequences in order to “score” (according to a reward heuristic) the single next action that leads to higher future reward (Arora *et al.* 2017; Morere *et al.* 2018; Sunberg & Kochenderfer 2018). These methods typically select a fixed finite horizon h over which to “rollout” scenarios in order to compute the score measure. Practically, $h \leq N$ where N is the total number of actions that could be executed in a single mission. The trouble with fixing h for adaptive sampling problems is that it ignores how information density of actions changes over time. For example, typically very little information is known about an environment at the beginning of a mission, and any action is likely to be as good as any other action to take; simulating long chains of sequences at this point could be considered wasted computation. As observations are drawn it becomes more strategic to plan trajectories in order to effectively explore or exploit knowledge. However, if there comes a point in which the world is known very well, then simulating long trajectories at every planning step may again be wasted computation, since an open-loop plan executed to some horizon may do just as well.

The concept of *reachable belief spaces* (Kurniawati *et al.* 2008) has been used to quantify the number of belief states that can be reached from an initial starting configuration. Recent work (Flaspohler *et al.* 2020), leverages this idea to modify the behavior of a robotic agent (selecting between closed loop and open loop trajectories) in order to save planning computation. This idea is generally appealing for robotic adaptive sampling, however there is an additional difficulty with respect to natural spatiotemporal systems: even if the underlying function is known completely, there may be emergent

stochasticity/chaotic behavior. For instance, while the Navier Stokes equations can be written exactly, for some selections of the Reynolds number of a fluid the solution is entirely predictable (laminar flow), and for others, completely chaotic (turbulent flow). Thus, the reachable belief space cannot be computed directly from uncertainty over the environmental model; it must also be computed with respect to the implications that model may have about chaotic behaviors.

PROPOSED PROJECT One method for solving this problem when there is access to a model is drawing samples from a distribution of parameters of that model, and forward simulating these samples in order to characterize the stochasticity of the environment. However, computationally, this is an intractable task for practical applications in IPP, as most sophisticated forward simulators that can capture the impacts of turbulence/stochasticity are numerically slow to solve, even in small systems. This project proposes developing a set of learned predictability measures to be used as meta-heuristics for nonmyopic search. *Predictability* in this case is the notion of how deterministic or chaotic a dynamic system may be. Intuitively, characterizing the predictability of an environment will signal times or locations of increased stochasticity which may require different treatment in planning than during more predictable times or locations. One way predictability could be used in a planning framework is in the setting of a planning horizon—optimized long-horizon exploitative trajectories may be more fruitful when predictability is high, whereas when predictability is low, any short-horizon action may be as good as another. Another way to use predictability could be in the reward function for a planner. The utility of predictability as a reward signal would be to directly optimize trajectories that complement different, complex scientific objectives about the form of an environment (for instance, an objective to reduce uncertainty about the *statistics* of turbulence in an environment may favor sampling in the most

chaotic regions, whereas an objective to localize effluent sources may favor the most predictable regions).

BACKGROUND AND CONSIDERATIONS One of the critical challenges of this research is computing a notion of predictability. For dynamical systems which can be described by differential equations, the degree to which a system can be considered predictable (alternatively, chaotic) is typically quantified by one of several measures, including *Lyapunov exponents* (Wolf *et al.* 1985) and the *Kolmogorov-Sinai measure-theoretic entropy* (Frigg 2004).

Kolmogorov-Sinai entropy (KSE) is a measure of trajectory divergence. For measure-invariant systems, KSE can be computed by dividing the dynamical system into piecewise partitions $Q = \{Q_1, \dots, Q_k\}$ and defining a measure of entropy on the partition Q as: $H(Q) = -\sum_{m=1}^k \mu(Q_m) \log(\mu(Q_m))$. The measure-theoretic entropy of the system with respect to a partition follows: $h_\mu(T, Q) = \lim_{N \rightarrow \infty} \frac{1}{N} H(\vee_{n=0}^N T^{-n} Q)$, and the KSE ultimately takes the form:

$$h_\mu(T) = \sup_Q h_\mu(T, Q) \quad (7.2)$$

Lyapunov exponents, which are more commonly used in numerical analysis of dynamic systems, are also a measure of the rate of separation of infinitesimally close trajectories. In phase space, given an initial separation of trajectories as $\delta \mathbf{Z}_0$, the rate of divergence is

$$|\delta \mathbf{Z}(t)| \approx e^{\lambda t} |\delta \mathbf{Z}_0| \quad (7.3)$$

where λ is the Lyapunov exponent. In practice, there is a Lyapunov exponent associated with each dimension of the phase space. For example, in an n -dimensional space $\dot{x}_i = f_i(x)$ let the Jacobian for this system take the form $J_{ij} = \left. \frac{df_i(x)}{dx_j} \right|_{x(t)}$. The tangent vectors

of a trajectory, Y , describe how a change at $x(0)$ propagates to point $x(t)$, and evolve according to $\dot{Y} = JY$. Finally, let $\Lambda = \lim_{t \rightarrow \infty} \frac{1}{2t} \log(Y(t)Y^T(t))$. The eigenvalues of Λ give the Lyapunov exponent spectrum. The *maximal Lyapunov exponent* (MLE) is the largest exponent for a given system; positive MLEs imply chaotic dynamics.

To recover KSE or Lyapunov exponents from field data requires a belief representation that natively supports their estimation. If a measure invariant subspace can be discovered, as proposed in MISL, the KSE can be computed directly. To compute Lyapunov exponents, however, requires some additional stipulations on a belief representation or learned subspace. For instance, typical MOR techniques tend to smooth away and erase instabilities in a dynamical system, making computation of Lyapunov exponents essentially impossible. Thus, when learning a descriptive subspace, preserving features in the data useful to exponent computation is necessary. Recent work has shown for low-dimensional dynamical systems that the “instability” subspace can be approximated from data (Blanchard & Sapsis 2019a; Blanchard & Sapsis 2019b) and can explicitly support Lyapunov characterization. Applying similar restrictions placed on the learning framework in this work on a belief model like MISL could be an interesting initial starting point.

7.5 MOBILE GEOCHEMICAL OBSERVATORIES FOR ATMOSPHERIC EMISSIONS

MOTIVATION Just as hydrothermal vents in the deep sea reveal information about crustal processes and magmatic activity, so to does the gas chemistry emitted from terrestrial volcanic vents and geothermally active fumaroles (Fischer *et al.* 2015; McGonigle *et al.* 2008; Smith *et al.* 2009). Several gas species, like water vapor (H_2O), hydrogen sulfide (H_2S), hydrogen chloride (HCl), carbon dioxide (CO_2), and sulfur dioxide (SO_2)

can be used to predict eruptive events (Fischer *et al.* 2015; Fischer & de Moor 2019; Hernández *et al.* 2001; Smith *et al.* 2009) as their presence and abundance indicates magmatic upwelling events. Remote sensing of these gases e.g., from satellites, can be used to observe large emissive events, however, such techniques lack the temporal resolution and spatial sensitivity necessary for monitoring (McGonigle *et al.* 2008). Observatories, like those operated by the United States Geological Service (USGS) use multiple sensing modalities to monitor high-risk areas (*USGS operates five U.S. volcano observatories* 2019). However, while GPS masts or tilt-sensors are widely used, singular geochemical sensing stations are typically reserved for the most high-risk areas due to cost (Smith *et al.* 2009).

PROJECT PROPOSAL Cheap and distributable multigas sensors are in active development at Woods Hole Oceanographic Institution among other institutions (Galle *et al.* 2021; Kaliszewski *et al.* 2021; Stix *et al.* 2018). To get an approximate picture of the geochemical structure of venting at a volcanically dynamic site, a network of these chemical sensors could be deployed. Exhaustively instrumenting a vent field is generally not feasible¹, however strategically using a few sensors which can be occasionally moved, or are themselves mobile and can adapt to the environment, would make deploying such a network a distinct possibility.

In this project, a team of aerial and ground vehicles for ground-to-atmospheric monitoring of gaseous volcanic emissions is envisioned for development. Equipped with small, multigas sensors, ground vehicles can be used to inspect vents and strategically monitor key productive sites over potentially long (days/weeks) periods of time, to characterize their temporal stability. Aerial teammates can provide intermittent coverage over a larger area and monitor gaseous fluxes, as well as identify regions in which

¹Due, for instance, in the risk placed on technicians deploying and maintaining these instruments; and the sheer scale of many environments

an interesting vent may be likely. Coordination by the ground and air teams will ultimately yield a multi-resolution dataset of a volcanic field, which can be analyzed for e.g., long-term changes in gaseous content and estimating net gaseous and energetic output from an entire instrumented site.

BACKGROUND AND CONSIDERATIONS Flying geochemical observatories for studying volcanoes have been previously proposed and built (Galle *et al.* 2021; McGonigle *et al.* 2008; Stix *et al.* 2018; Xi *et al.* 2016), however these systems have only been used to measure short-lived eruptive events. Similarly, remotely-operated (or autonomous waypoint-following) ground vehicles have been used to study actively erupting sites (Muscato *et al.* 2012), but never in a long term context. To date, no study has been conducted which uses an intelligent agent to characterize long-term characteristics in geothermal volcanic fields. Such a system would provide insight about volcanic emissions that would otherwise be difficult or impossible to retrieve with current sensing modalities.

The problem of long-term monitoring in dynamic environments poses significant research questions for both intelligent sampling development and scientific machine learning. In general, learning the underlying dynamic of a spatiotemporally-evolving system is a hard problem; exploiting it in order to answer varied scientific queries adds further complexity; and adding multi-fidelity sensing and heterogeneous teams makes the problem an interesting avenue for algorithmic research. Functionally, this is a potent planning under uncertainty problem. To begin to approach designing an autonomy framework for this problem, inspiration in other academic fields may prove useful. For instance, in sea-surface temperature research, satellite data and *in situ* measurements from buoys are often combined into predictive models with notions of uncertainty (Babaei *et al.* 2020). Application of these techniques outside of ocean environments would be scientifically novel, let alone potentially useful for planning. In

planning, state-of-the-art results in multirobot coordination in search and rescue contexts, in which intermittent communication, collaborative mapping, and heterogeneous capabilities are considered, may similarly find a novel application in volcanic monitoring (Queralta *et al.* 2020; Rizk *et al.* 2019).

7.6 MODELING ANTHROPOGENIC METHANE IN ESTUARIES

MOTIVATION 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 methane and carbon dioxide, which are outgassed during turbulent events (Cole *et al.* 2007; Stanley *et al.* 2016); however salt marshes can be considered a carbon sink through biological activity (Cai 2011). 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. While untreated wastewater is obviously problematic, even treated wastewater can be hiding dangerous pollutants. In illustration, consider a study of the Wareham River in Massachusetts² in which a 1.56 million-gallon-per-day biological nutrient removal wastewater control center cleans and dumps water. The feedstock for bacteria in one of the treatment stages is methanol. When that methanol is not completely digested by the bacteria, it is carried by the cleaned wastewater into the river. Significant methane plumes were confirmed to be present in the river near the wastewater outfall, and hypothesized to be rapidly ventilated or digested within the

²savebuzzardsbay.org/embayments/wareham-river/

estuary (Preston 2019). This is alarming as methane is a potent greenhouse gas, and disrupting natural food chains in an estuary has compounding impacts on overall river health; it is also alarming because this is an essentially “invisible” problem as the water meets regulatory cleanliness standards and methane is an odorless, colorless gas. Major facets of research on estuarine waters influenced by wastewater include characterizing the impacts of synthetic materials (e.g., antibiotics (Roberts & Thomas 2006)), increased organic materials (e.g., fertilizers and treatment feedstock (Valiela *et al.* 2016)), and urban developments (e.g., stormwater run-off (Geedieke *et al.* 2018)).

PROPOSED PROJECT It is imperative that monitoring solutions and models for biogeochemical transport of methane and other greenhouse gases and pollutants that enter coastal estuaries be developed for effective policy-based interventions at wastewater treatment plants and other anthropogenic outfalls. Such a systems would characterize the fate of geochemicals: where they are transported to aqueously, how much may ventilate into the atmosphere, and what is consumed by biological processes. To assist in developing these models, high-resolution measurements of dissolved gases, particulates, and other tracers (e.g., salinity, temperature) are necessary.

Robotic platforms, from small underwater vehicles to surface vehicles to aerial vehicles, could all carry instrumentation that looks at the river-to-atmosphere continuum while carrying *in situ* equipment. The distinct advantage to using robotic platforms in this setting is related to the character of the environment; some river systems may only be practically navigable by small (kayak-sized) water craft, requiring scientists or technicians to perform relatively labor-intensive sample collection if they were to do this themselves. As collecting data over multiple days, weeks, and seasons would be beneficial, using robotic technologies that can execute consistent tasks with minimal supervision over long monitoring horizons would significantly change the paradigm of estuary sampling today. This project envisions the development of robotic monitoring

tools for estuary science, and automated data science tools for processing the big data that is generated by such studies.

BACKGROUND AND CONSIDERATIONS Estuarine research has historically used numerical models of fluid flows (Geyer & MacCready 2014) (of varying resolution) and sparse observations of pollutants (e.g., bottle samples) (Rheuban *et al.* 2019) to generate estimates of tracer transport and fate. *In situ* sensors suitable for these studies have yet to be fully leveraged for geochemical estuarine research, and they represent a large departure from classical approaches. Hesitancy in part comes from “data paralysis” of going from a few sparse measurements to tens of thousands of observations, as typical methods used to invert fluid models may be brittle or computationally intractable to perform over so much data. Moreover, there is not yet an agreed standard for parsing bulk *in situ* observations of estuarine chemicals. A roboticist may be well-suited to assist in these areas as fluency in developing and interpreting belief representations (data summaries) is a natural tool to address these issues. Algorithmically, similar challenges as other expeditionary contexts persist—namely dealing with partial observations of spatiotemporal environments, making this not just a straightforward application of out-of-the-box methods.

7.7 LOOKING AHEAD

Expeditionary science motivates a set of interesting robotics, modeling, and decision-making problems that are not well-addressed by current state-of-the-art methods. There is increasing need to enable better *in situ* methodologies for monitoring and characterizing large spatiotemporal environments as accelerated climate change transforms delicate carbon budgets, ecosystem networks, and weather patterns. Robotic technologies, both

theoretical and applied, can play a transformative role in understanding these evolving trends and assessing policy interventions.

8 FINAL THOUGHTS

I believe that every engineer has a responsibility to make the world a better place. We are gifted with an amazing power to take people's wishes and make them a reality.

Ayanna Howard

In November 2021, AUV *Sentry* was used to chart hydrothermal plumes in the first field demonstration of a physically-informed autonomous mission planner for deep sea geochemical surveys. This thesis presents the algorithmic underpinnings and preliminary scientific results. Fundamental to the research program presented here is the notion that robots and robotic algorithms are tools that are best selected when the context of a problem domain is fully defined. Expeditionary science, the art of collecting useful *in situ* observations of natural phenomena, provides a context that requires tools that can parse complex, sparse observations into interpretable summaries of spatiotemporal distributions, reason over uncertainty to generate environmental forecasts, and leverage sophisticated sensing platforms to the fullest of their abilities.

Several tools are presented in this thesis. In Chapter 4, a set of anomaly detection and temporal analysis methods are presented that succinctly summarize complex, heterogeneous data from *in situ* geochemical sensors to assist a science party in identifying hydrothermal expressions. Using these tools demonstrated that hydrothermalism could be confidently identified far from a generating vent in the Guaymas Basin (4-7 km) in the first study to quantify the extent of hydrothermal intrusion in the water column in

the Northern Basin. These tools were later adapted in Chapter 6 to create an automatic hydrothermalism detector for use in a robotic decision-making framework.

That framework, presented in Chapter 5 and demonstrated in the field in Chapter 6, consisted of a belief representation, PHUMES: **P**Hysically-informed **U**ncertainty **M**odels for **E**nvironment **S**patiotemporality, and an algorithmic decision-maker, PHORTEX: **P**Hysically-informed **O**perational **R**obotic **T**rajectories for **E**Xpeditions. PHUMES leveraged a novel time-averaged model of buoyant plumes in stratified environments with crossflow to generate forecasts over the location of plume fluid from a history of observations. The formulation of PHUMES positioned the belief representation as a universal data aggregator not just for *Sentry* operations, but all other science activities in the water column, enabling PHUMES to produce forecasts with predictive power several days into the future. By virtue of the choice of analytical model, PHUMES was able to examine the structure of neutrally-buoyant plumes observed in the Basin, and generate novel hypotheses about their formation and characteristics that would otherwise be impossible to resolve using classical models of buoyant rise in the deep ocean.

With PHUMES in hand, PHORTEX generated trajectories that enabled AUV *Sentry* to track moving hydrothermal plumes without adaptive capabilities. Using chained lawnmower trajectory patterns, PHORTEX trajectories “fanned out” over the course of a multi-hour dive to effectively re-encounter and survey plume structures, collecting spatiotemporally diverse datasets that sharply contrast with human-designed surveys which tend to bias detections of hydrothermalism to within a small distance of a generating vent. With operational changes made on the fly while at sea, the ability for PHORTEX to design trajectories that could robustly sample a plume demonstrates its utility for future field campaigns.

In Chapter 7, this thesis makes a direct appeal for consideration by the robotics community of expeditionary science contexts, asserting that expeditionary science pro-

vides interesting constraints and requirements that challenge the state-of-the-art tools currently available today. Moreover, roboticists are uniquely equipped to impact the strategic and urgent vision for improving scientific understanding of the Earth, the ocean, and the processes within them. Several possible project settings are defined in detail, presenting both algorithmic and contextual advances ripe for development. Uniting all of these projects is a persistent need for belief representation tools that can usefully describe spatiotemporal distributions which can only be partially observed. The key insight of this thesis is that scientific models can be embedded as principled priors and bias in learning frameworks. Not only does embedding knowledge allow for tractable recovery of spatiotemporal dynamics, but the belief representations themselves can directly support scientific inquiry by virtue of being interpretable, grounded by accepted principles, and familiar to scientist stakeholders. For advances in scientific machine learning, stochastic model learning, and physically-informed belief representations to make an impact, it is imperative that the systems in which these algorithms are to be deployed—the environments, people, platforms, and operations—be considered, and ideally involved through collaborative projects.

In closing, the research in this thesis is the culmination of collaborative efforts between scientists and engineers to do something that has never been done before, but which will hopefully be done hundreds of times more, and better, into the future. It is a distinct privilege of a field roboticist to have the ability to stretch between academic silos and work with diverse teams to make sophisticated autonomy while getting their boots wet on the deck of a ship. It also comes with a distinct responsibility to prove that their tools not only work, but that they address the need presented by sponsors/stakeholders, collaborators, and communities. Showing that a need is addressed is not necessarily as straightforward as showing that some cost function is minimized; it requires following up, pursuing the science, and creating adjustments that can make all the difference in

a tool being used again. Being a meaningful part of a team makes engaging with the process of demonstrating impact natural and not onerous; necessary and not frivolous. Responsibility to one another and towards being stewards of Earth will be potent drivers for the development of robotic expeditionary science in the years to come.

A DISCOVERING HYDROTHERMALISM FROM AFAR

A.1 METHOD FOR METHANE MEASUREMENT FROM NISKIN BOTTLES

A Los Gatos Research (LGR) Dissolved Gas Extraction Unit (DGEU) and Greenhouse Gas Analyzer (GGA) were used to process water collected by Niskin bottle samples during the transect, and report methane concentration estimates to be compared to the *in situ* observation of normalized methane by SAGE mounted on the rosette. Measurements of methane made by the GGA are reported as the stabilized parts per million (ppm) reading provided by the instrument after consuming 3-5 L of seawater from each Niskin bottle, and are converted to nanomolar (nM) values by first computing the partial pressure of methane, and then computing molarity by estimating the solubility constant of methane using coincident measurements of salinity and temperature of the seawater at time of bottle sample collection as measured by the rosette CTD. The conversion from partial pressure to molarity is done using the `gasex` Python library, publicly hosted at <https://github.com/boom-lab/gasex-python>.

To transform GGA measurements in ppm to partial pressure, the DGEU cell pressure is used, such that $\text{ppm} \times \text{cell pressure} = \text{partial pressure}$. Additionally, gas extraction

inefficiency is taken into consideration at this step; the DGEU does not perfectly extract gas across the membrane during sampling. Extraction efficiency is used to scale the GGA measurement of methane prior to computing the partial pressure estimate by

$$\left[\frac{x_{obs} - x_{ref}}{\lambda_{eff}} + x_{ref} \right] \frac{p_{cell}}{1000} = x_{pp} \quad (\text{A.1})$$

where x_{obs} is the ppm measurement made by the GGA, x_{ref} is a methane reference value (the atmospheric concentration of methane, typically between 1.86-1.99 ppm), λ_{eff} is the extraction efficiency, p_{cell} is the cell pressure in millibar, and x_{pp} is the estimated partial pressure value, in μatm .

The extraction efficiency used in this manuscript was estimated by laboratory calibrations to be between 2.3-3.3%, consistent across different water temperatures and different test tank concentrations. In the laboratory calibration procedure, methane was bubbled in a temperature-controlled tank which was stirred before two discrete samples were taken using 60 mL syringes filled with 40 mL of water, and 20 mL of pure nitrogen gas. A DGEU, connected to the GGA, was then used to take water from the target tank, and ppm measurements by the GGA were recorded when measurements stabilized; this was done with two different DGEUs, which we label A and B. To estimate “ground truth” partial pressure of methane in the tank, the syringe samples were shaken for 2 minutes to extract the dissolved gas content, and the water drained. The samples were then processed within 24 hours on a gas chromatography instrument (Shimadzu GC-14B), run alongside a set of standards processed every 5 minutes. The measurements from the processed syringes (DGEU influent) were used as x_{pp} in Eq. A.1, the GGA observations as x_{obs} , the value 1.99 ppm used as x_{ref} , and 495 mbar as p_{cell} . The relevant data from these calibrations is available in Tab. A.1. DGEU A was the instrument used in the transect field mission as presented in this manuscript.

DGEU	Temperature (C)	Influent (μatm)	GGA Methane (ppm)	Efficiency
A	4.7	299.13	21.82	3.29%
A	4.7	512.03	6.44	0.4%
A	4.7	588.25	41.04	3.29%
B	4.7	299.13	17.77	2.62%
B	4.7	512.03	25.68	2.29%
B	4.7	588.25	30.09	2.37%
A	9.9	267.14	17.02	2.80%
A	9.9	403.45	27.84	3.18%
A	9.9	856.89	55.77	3.11%
B	9.9	267.14	12.72	2.00%
B	9.9	403.45	18.63	2.05%
B	9.9	856.89	36.99	2.02%
A	14.8	18.64	2.78	2.22%
A	14.8	1549.18	101.26	3.17%
A	14.8	1640.81	100.41	2.97%
B	14.8	18.64	2.63	1.80%
B	14.8	1549.18	78.93	2.46%
B	14.8	1640.81	68.43	2.01%

Table A.1: Results of DGEU extraction efficiency calibration experiments.

A.2 LEG 2 NISKIN BOTTLE SAMPLE SCHEDULE AND MEASUREMENTS

This manuscript presents methane and ammonium measurements collected by Niskin bottles during Leg 2 of the rosette trajectory. Table A.2 provides the schedule of Niskin bottle firing performed during Leg 2, and Table A.3 provides all data associated with those bottles collected and presented in Chapter 3 of this thesis. The range of methane nM values is provided by converting GGA methane ppm measurements as described in Sec. A.1 for the conservative range of valid DGEU extraction efficiency values.

Bottle	Time	Location	Depth (m)
1	2021-11-30 09:10:03	27.3951N 111.3649W	1648.62
3	2021-11-30 09:30:03	27.3956N 111.3665W	1625.67
5	2021-11-30 09:47:01	27.3967N 111.3696W	1639.25
7	2021-11-30 09:47:05	27.3967N 111.3696W	1639.05
9	2021-11-30 10:07:00	27.3985N 111.3740W	1598.32
11	2021-11-30 10:17:02	27.2994N 111.3765W	1580.5
13	2021-11-30 10:27:01	27.4005N 111.3791W	1568.27
15	2021-11-30 10:27:04	27.4005N 111.3791W	1568
17	2021-11-30 10:37:20	27.4016N 111.2818W	1558.64
19	2021-11-30 10:46:59	27.4027N 111.3845W	1553.92
21	2021-11-30 11:07:05	27.4051N 111.3900W	1547
23	2021-11-30 11:33:00	27.4082N 111.3971W	1545.4

Table A.2: Schedule of bottle samples during Leg 2 of rosette transect.

Bottle	CH ₄ (ppm)	CH ₄ (nM)	NH ₄ ⁺ (nM)	Temp. (C)	Salinity (PSU)
1	—	—	0.00	2.8334	34.6104
3	9.29	207-296	46.35	2.8578	35.6095
5	21.6	547-785	—	2.8458	34.6107
7	—	—	174.48	2.8461	34.6108
9	22.54	573-821	165.99	2.8659	34.6101
11	29.82	774-1110	225.87	2.8719	34.6096
13	44.36	1176-1686	—	2.8734	34.6099
15	—	—	384.28	2.8733	34.6098
17	89.45	2421-3473	780.53	2.8849	34.6105
19	114.27	3105-4454	997.45	2.8968	34.6111
21	27.29	704-1009	227.54	2.8835	34.6087
23	11.5	268-384	89.29	2.8964	34.6075

Table A.3: Geochemical measurements associated with the schedule of bottle samples during Leg 2 of rosette trajectory. Note that methane expressed in nM is computed using coincident temperature and salinity measurements during the transect as measured by rosette CTD, and extraction inefficiency of the DGEU is compensated for as described in Sec. A.1.

A.3 NORMALIZED PYTHIA CALIBRATION

The Pythia instrument provides a significantly nonlinear output reference value when measuring methane. We correct for this nonlinearity using a reference curve computed in the laboratory before normalizing the measurements as reported in this manuscript. The reference curve was created using a temperature-fixed (3°C) tank and closed equilibration chamber, in which methane standards were bubbled until fully equilibrated before being measured by the instrument. Stable measurements by Pythia (which has a response time of approximately 35 minutes) were then recorded at different chamber concentrations. The calibration curve that results is a piece-wise linear function, shown in Fig. A.1.

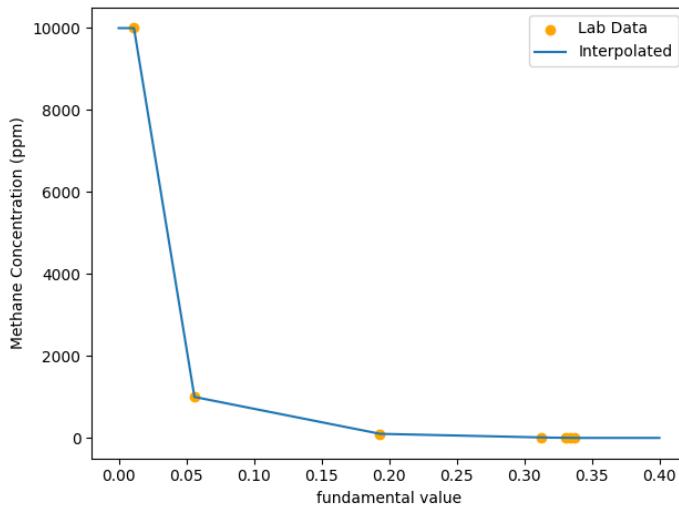


Figure A.1: Fitted calibration curve for measurements of methane observed by Pythia.

Compensation of Pythia's time response was also performed on post-calibrated data using the methodology described in Miloshevich *et al.* 2004 with a smoothing window of 5 minutes, and subsampling at a quarter of the time delay window. This methodology is sensitive to noise in the signal, which motivates the extreme sub-sampling that is performed. Fig. A.2 shows the effect of smoothing, time-correction, and conversion on the direct signal recorded by Pythia before normalization.

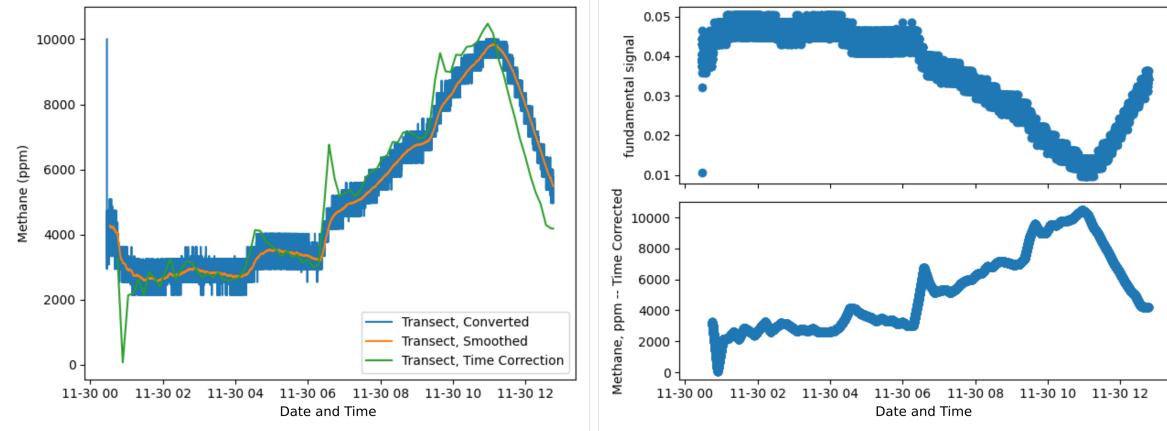


Figure A.2: Calibration curve, smoothing, and time correction applied to Pythia observations during the transect, before reported normalization in the manuscript.

A.4 DEPTH-CORRECTION

Temperature, salinity, and oxygen are expected to be weakly stratified in the deep ocean. To remove these effects from data collected by AUV Sentry and the rosette, we fit a line to the average observations collected within binned 20 m intervals of observed depth for each platform separately. Separately computing the correction for each instrument additionally controls for small discrepancies in calibration between the platforms. Fig. A.3 compares these lines with the observations collected.

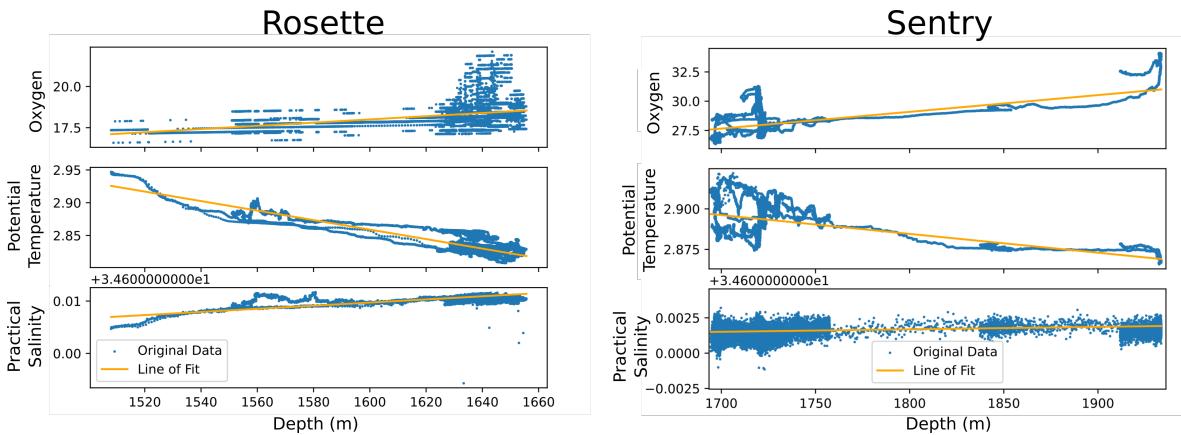


Figure A.3: Linear functions are fit to data collected for oxygen, temperature, and salinity instruments on each platform separately. A residual value is then computed for each observation.

A.5 DESCRIPTION OF PLUME MODEL FOR TRANSECT DESIGN

We adapted an idealized buoyant bent-plume model proposed by Tohidi & Kaye 2016 for atmospheric bent plumes in a weakly stratified fluid in order to inform at what heights to deploy AUV Sentry and the rosette during the transect. We rewrite the system of equations provided in Tohidi & Kaye 2016 as follows:

$$E = \alpha \left| \frac{M}{Q} - u \cos(\theta) \right| + \beta |u \sin(\theta)| \quad (\text{A.2})$$

$$\frac{dQ}{ds} = QE \sqrt{\frac{2(1 + \lambda^2)}{M\lambda}} \quad (\text{A.3})$$

$$\frac{dM}{ds} = u \cos(\theta) \frac{dQ}{ds} + \frac{FQ}{M} \sin(\theta) \quad (\text{A.4})$$

$$\frac{d\theta}{ds} = \left(\frac{FQ}{M} \cos(\theta) - u \sin(\theta) \frac{dQ}{ds} \right) \frac{1}{M} \quad (\text{A.5})$$

$$\frac{dF}{ds} = -QN^2 \sin(\theta) \quad (\text{A.6})$$

$$\frac{dX}{ds} = \cos(\theta) \quad (\text{A.7})$$

$$\frac{dZ}{ds} = \sin(\theta) \quad (\text{A.8})$$

where E is a mixing entrainment coefficient which considers both vertical and horizontal mixing and is weighted by parameters α and β , u is the crossflow velocity which can be a function of depth and time, λ is a parameter which modifies the ellipse which describes the plume envelope, Q is specific volume flux, M is specific momentum flux, F is specific buoyancy flux, θ is plume centerline trajectory angle, s is the plume centerline trajectory, X is distance along a coordinate axis aligned with the plume centerline, Z is

height with respect to plume source along a vertical axis, and N^2 is the Brunt-Väisälä frequency, computed with respect to the density gradient at the reference depths of the source and plume height.

The system of equations essentially yields a “snapshot” of a plume envelope at some moment in time. For time-varying crossflows, multiple snapshots can be computed for different moments in time (different crossflow orientations and magnitudes) and chained together in a common coordinate reference system in order to track a plume trajectory. For the purposes of determining which heights to deploy AUV Sentry and the rosette for the transect, we compute a prototypical envelope and use the estimated bent nonbuoyant plume height to set the transect depths/altitudes.

The initial conditions for solving this system of ordinary differential equations are set via estimates of vent characteristics including exit velocity, temperature, salinity, and area. Specifically:

$$Q_o = \lambda V_v \frac{A_v}{\pi} \quad (\text{A.9})$$

$$M_o = Q_o V_v \quad (\text{A.10})$$

$$F_o = -g10^{-4}(T_v - T_z)Q_o \quad (\text{A.11})$$

$$\theta_o = \frac{\pi}{2} \quad (\text{A.12})$$

where V_v is exit velocity at the vent orifice, A_v is the vent orifice area, T_v is the temperature at the orifice area, and T_z is the expected temperature of ambient seawater at the estimated vent depth. Note that initial buoyancy flux is primarily driven by temperature changes, as we anticipate this to be the major driver of density gradients at our measurement scale. Expected salinity gradients could be similarly considered.

Estimated vent characteristics and crossflow were selected based on empirical observations of the deep sea vents located along the northern Guaymas Basin ridge and observations of current magnitude collected by a current tiltmeter deployed by ROV Jason during several days of the research cruise. Table A.4 lists the settings for planning the transect selected for these characteristics. Background salinity and temperature profiles were computed according to standard Pacific Ocean temperature and salinity functions as described in Speer & Rona 1989; additionally the equation of state for computing density profile from salinity and temperature measurements was used also as defined in Speer & Rona 1989. The prototypical plume is computed with a source located at 1850 m depth.

Parameter	Assignment	Description
λ	1.0	Ratio of elliptical axes of the plume envelope
V_v	0.58 m s^{-1}	Exit velocity of fluids at vent orifice
A_v	0.82 m^2	Area of vent orifice
T_v	340°C	Temperature of fluids at vent orifice
α	0.15	Longitudinal shear-driven mixing coefficient
β	0.19	Transverse shear-driven mixing coefficient
u	0.1 m s^{-1}	Magnitude of crossflow

Table A.4: Parameter, vent characteristics, and ambient crossflow setting used for transect design.

The prototypical plume envelope computed in this manner estimates a nonbuoyant plume depth between 1570-1750 m (Fig. A.4). AUV Sentry is altitude limited in order to keep a fix on the ocean floor for navigation; it is set to its maximum altitude of 120 m in order to intersect with the bottom of the estimated nonbuoyant layer; this corresponds to a depth of approximately 1700 m throughout the basin. The rosette can be arbitrarily fixed to a height, but so as not to interfere with AUV Sentry operations and to sample a different point in the estimated nonbuoyant layer, a depth of 1650-1600 m was targeted.

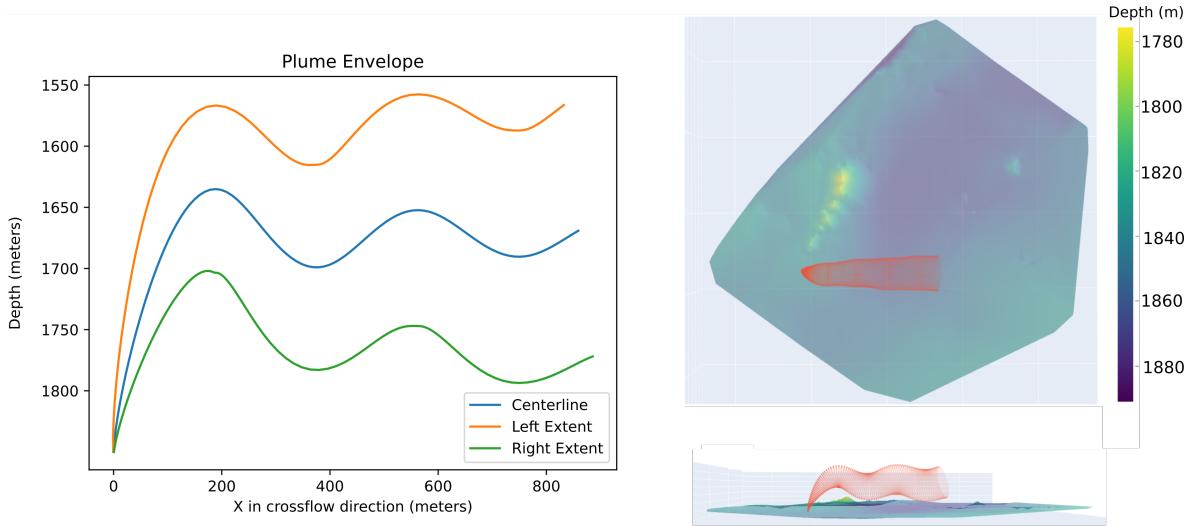


Figure A.4: A prototypical plume estimate according to the modified buoyant plume model in crossflow. The same envelope is plotted with respect to absolute depth (with a source located at 1850 m) on the left, and illustratively in the context of the hydrothermal ridge on the right.

A.6 CODE AND DATA AVAILABILITY

Ship-based and towed rosette data can be found in the Rolling Deck to Repository (R2R) at <https://doi.org/10.7284/909325>. AUV Sentry, SAGE, and Pythia data presented in this study are available on the Woods Hole Open Access Server (WHOAS) at <https://hdl.handle.net/1912/29403> with DOI 10.26025/1912/29403. Software and data analysis tools are available on Github at <https://doi.org/10.5281/zenodo.6789105>.

B PHORTEX PERFORMANCE

B.1 CONVERGENCE OF TRAJECTORY OPTIMIZER

The trajectory optimization scheme presented in Section 5.3.3 uses a gradient-based optimizer with trust-bounded constraints in order to set the parameters of a set of lawnmower trajectories. In general, this is a difficult optimization problem, as analytical gradients that map the defining parameters of a lawnmower trajectory (orientation, size, and location) are not available, and must be numerically approximated. Moreover, by defining trust constraints, jumping between minima could be made more difficult, as the approximated gradients may lead through unsafe (un-trusted) space. In practice, this meant that initializing each lawnmower in a chain to be near a “good trajectory” was important. In the case of charting hydrothermal plumes, a “good” trajectory would be one approximately aligned with the estimated crossflow. As PHUMES provides complete access to this information, it is generally easy to seed the trajectories to be near a performant minima. This leads to relatively fast convergence of the optimizer for each element in the chain. In Fig. B.1 there is an example a snapshot along the path to convergence for each link in a chain visualized, and the convergence plot. As is evidenced in the plot, a long travel time between chains is accumulated; this is likely because “flipping” a trajectory to shorten this distance would be a prohibitively difficult step in the nonlinear, constrained space that the optimizer operates.

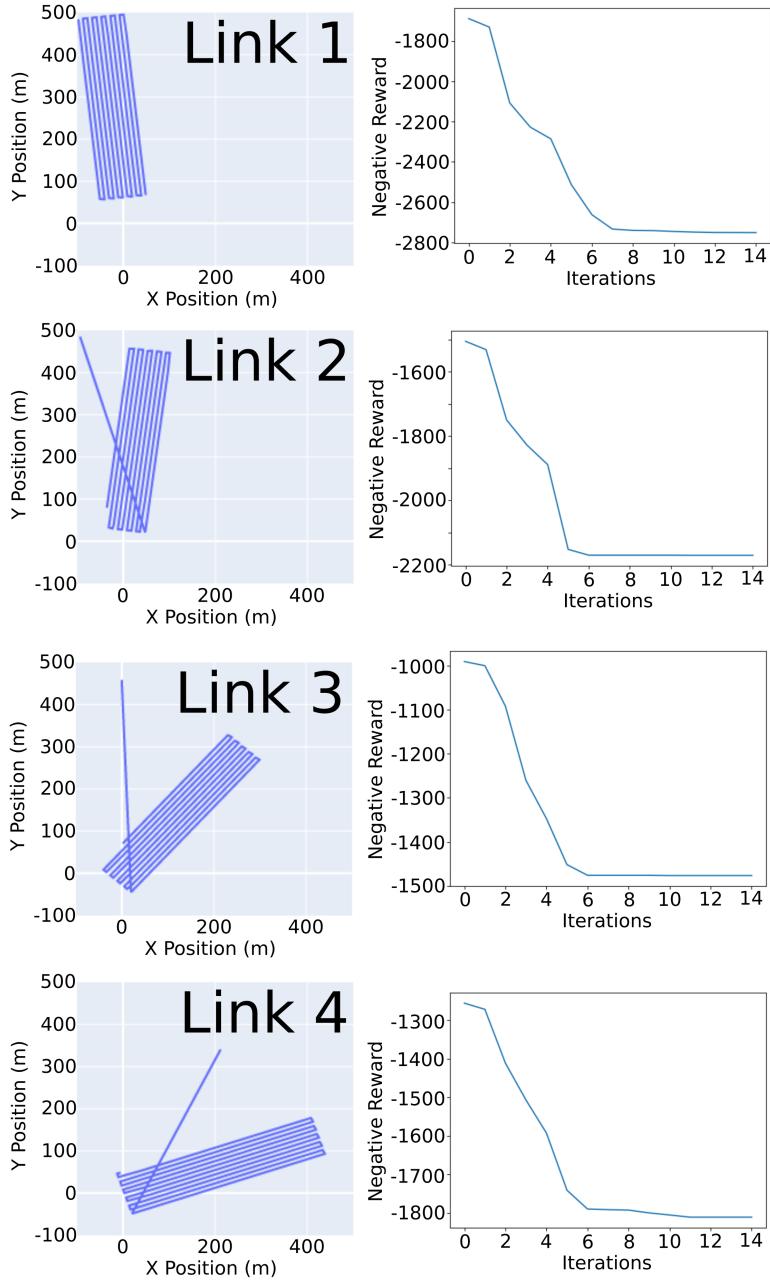


Figure B.1: Example snapshots of each link in a PHORTEX trajectory from a Chapter 5 simulation trial, and their corresponding convergence plots. While trajectories show good agreement with the current function and size of a plume expression as modeled by PHUMES, a long travel time between chains is accumulated; this is likely because “flipping” a trajectory to shorten this distance would be a prohibitively difficult step in the nonlinear, constrained space that the optimizer operates.

B.2 MCMC CHAIN CHARACTERISTICS OF PHUMES

For computational reasons, it is typically not possible to run the chains for thousands of samples while in the field, and the chains must be arbitrarily cut short in order to generate a plan. For this reason, it is interesting to look at the chaining properties. For a simulation trial, as described in Chapter 5, we show the chains and probability densities for each of vent area, vent fluid velocity, and entrainment coefficients in Fig. B.2. As evidenced by the multimodal area plot, we observe that the relationship between the parameters is complicated; in general, an inverse problem like the one we have posed for PHUMES is a challenge to infer. What is promising however is that probability densities are being concentrated in areas that make sense (as observed by a mode at 0.8 for area), and that the *combinations* of learned parameters are sensible — a large area should be coupled with a slow or medium velocity; a high velocity should be paired with a small area — based on the relationships between parameters and how they manifest in the ultimate shape of the plumes in the 3D Cartesian space.

We can also look at how convergence may change for longer chains. In an exemplar trial (Fig. B.3), a chain is run for 650 samples, and the last 500 samples are used to compute densities. The acceptance rate stays approximately the same, at about 34%. However, the distributions have generally improved with respect to placing more density at the parameters that defined the true underlying environment. As evidenced in the area and velocity sampling plots, there is still some exploration of the state space being performed by the chains; this suggests that true convergence of these chains may either take a long time, or the information content of the data that is available to explore the state space leads to ambiguous “wells” which the chains would “bounce” between for long durations. This suggests that other Monte Carlo techniques, such as Hamiltonian Monte Carlo (Duane *et al.* 1987), may be of interest in adaptation of this work, to accelerate chain exploration.

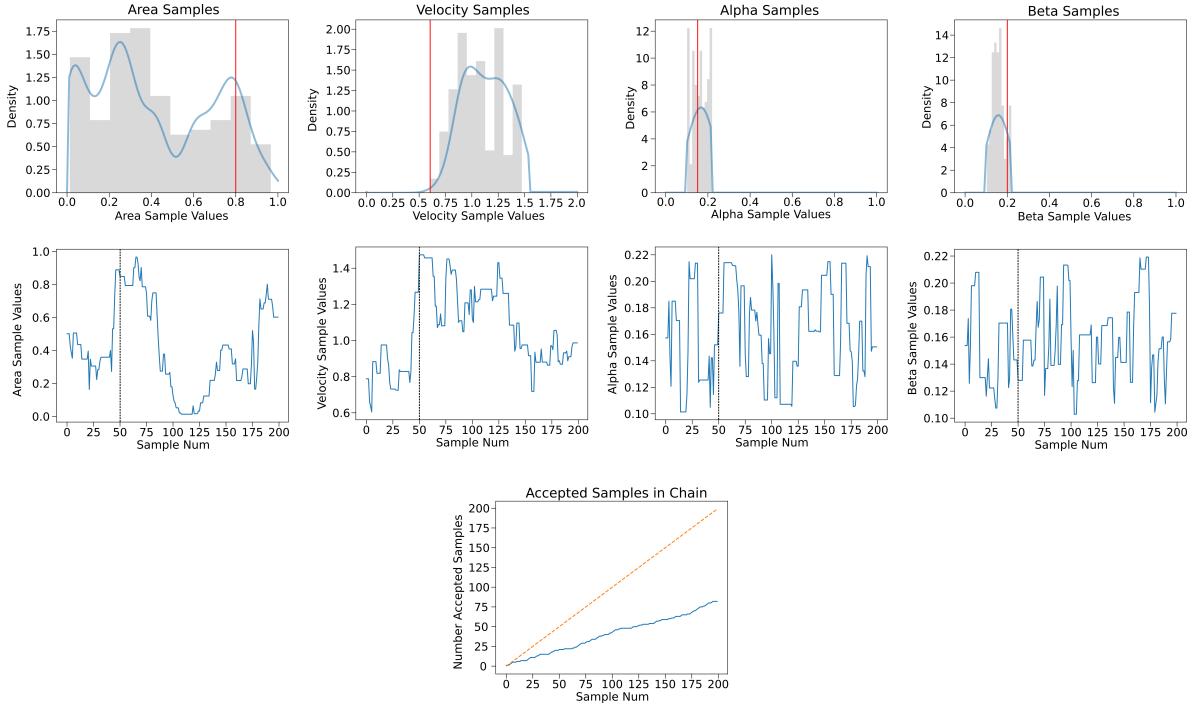


Figure B.2: An example chain from a simulation trial as described in Chapter 5. The last 150 samples in each chain is used to compute the distributions in blue; all samples are shown in gray in the top plots. Red lines in the top plots indicate the generating environment value for the simulation. The chain efficiency is approximately 37.5%.

B.3 CODE AVAILABILITY

PHUMES and PHORTEX is available through the Expeditionary Robotics organization on GitHub, <https://github.com/expeditionary-robotics>.

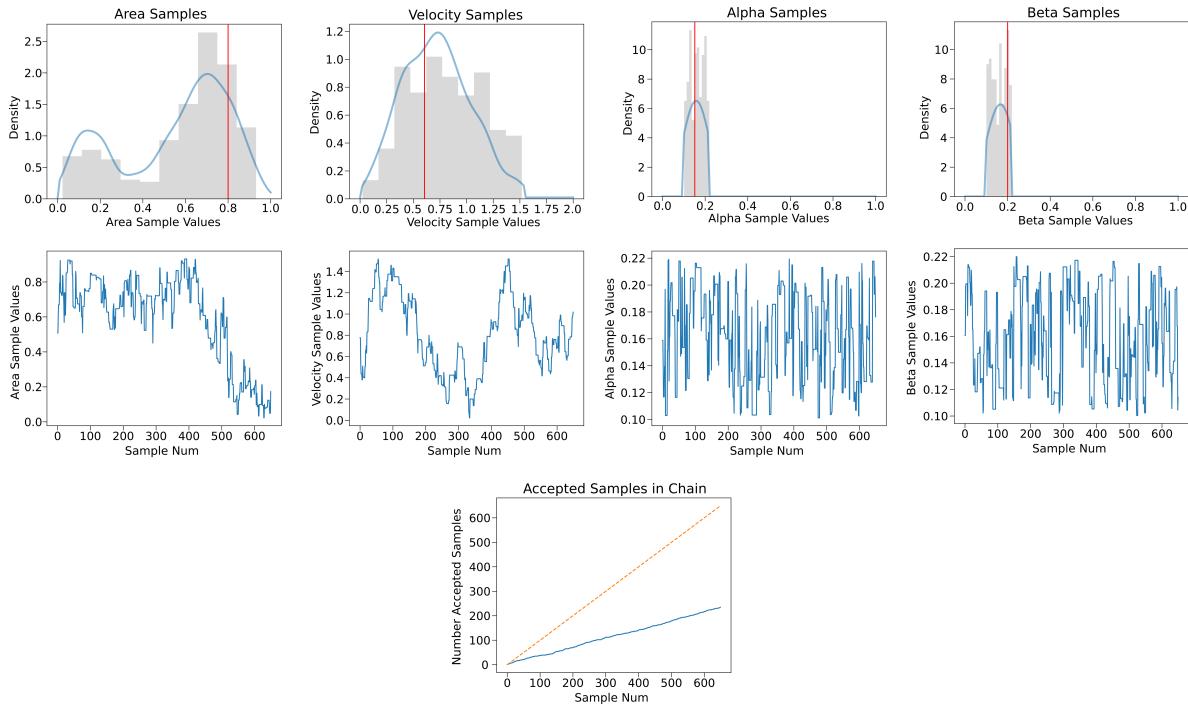


Figure B.3: An example chain from a simulation trial as described in Chapter 5, but run for 650 samples; the last 500 are used to compute the distributions in blue with all samples shown in gray in the top plots. Red lines in the top plots indicate the generating environment value for the simulation. The chain efficiency is approximately 34%.

C GUAYMAS BASIN FIELD RESULTS

C.1 INSTRUMENT DATA

In the main body of the text, the binary detections are reported for each of four AUV *Sentry* dives. Here, the continuous data is plotted on overhead polar charts for reference.

Sentry607

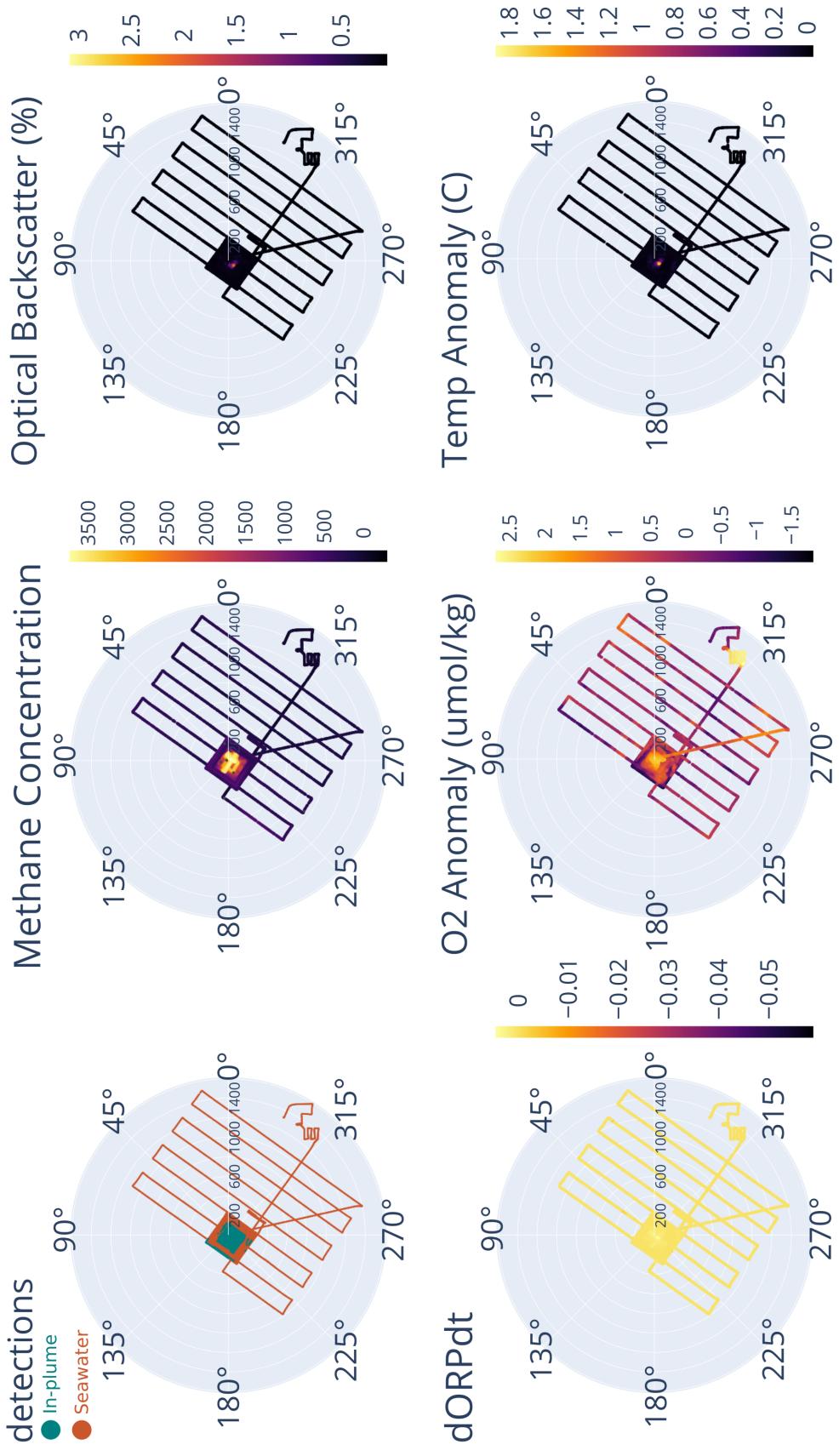


Figure C.1: **AUV *Sentry* dive 607 data.** The binary detections are plotted with methane value from the Pythia methane instrument (time-corrected and converted data to nM), turbidity, ORP, detrended oxygen, and detrended temperature.

Sentry608

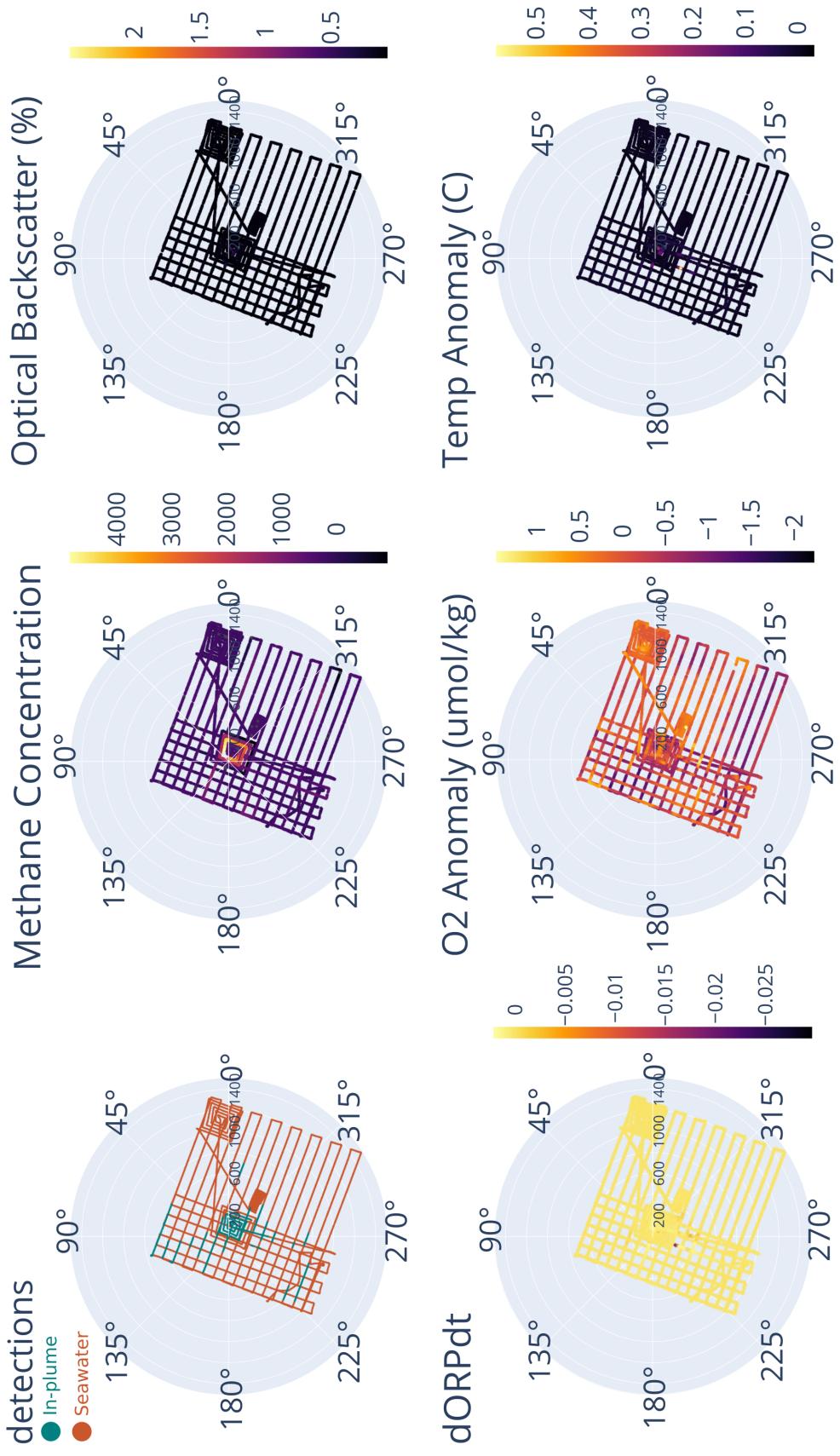


Figure C.2: **AUV *Sentry* dive 608 data.** The binary detections are plotted with methane value from the Pythia methane instrument (time-corrected and converted data to nM), turbidity, ORP, detrended oxygen, and detrended temperature.

Sentry610

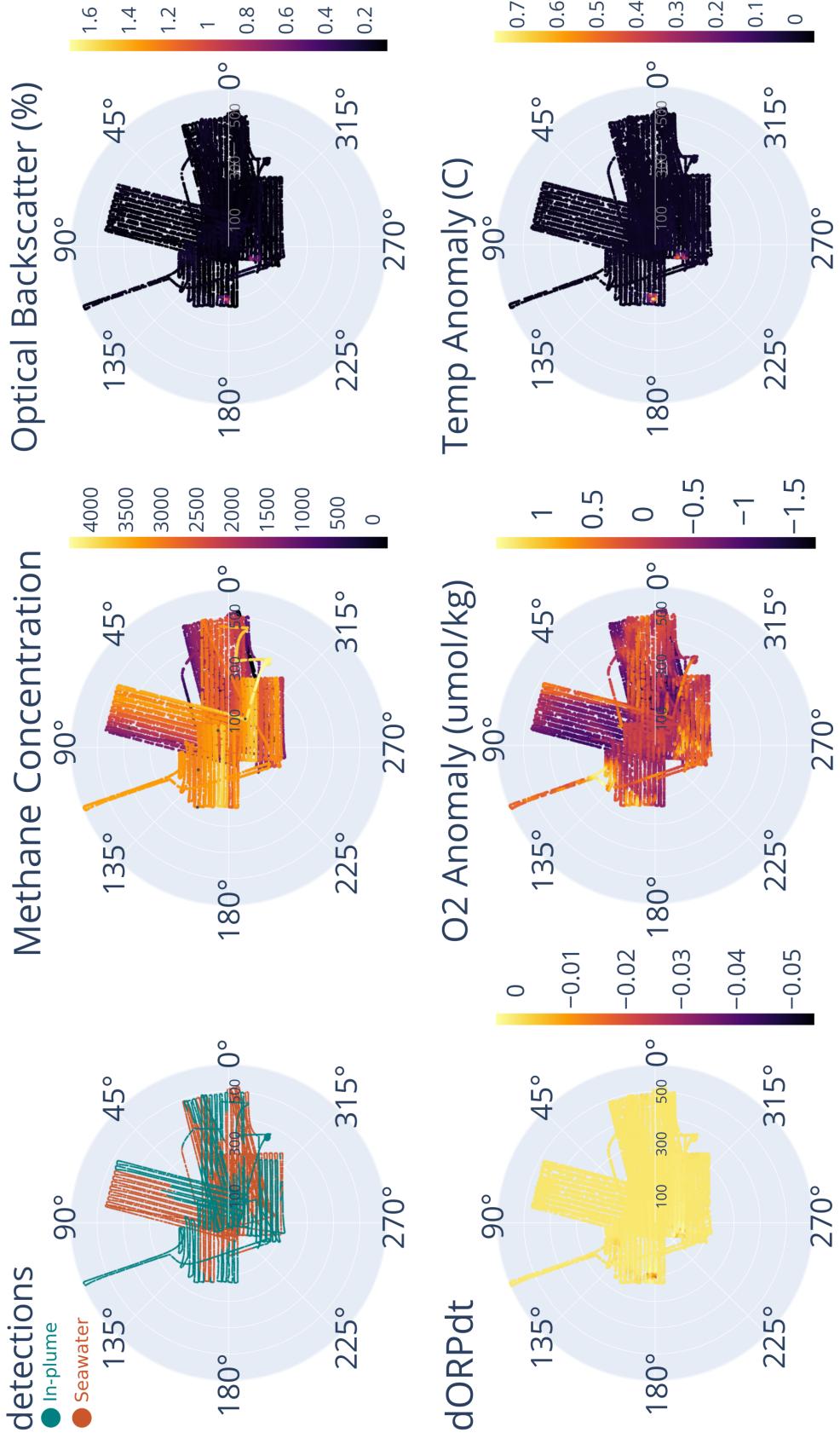


Figure C.3: AUV *Sentry* dive 610 data. The binary detections are plotted with methane value from the Pythia methane instrument (time-corrected and converted data to nM), turbidity, ORP, detrended oxygen, and detrended temperature.

Sentry611

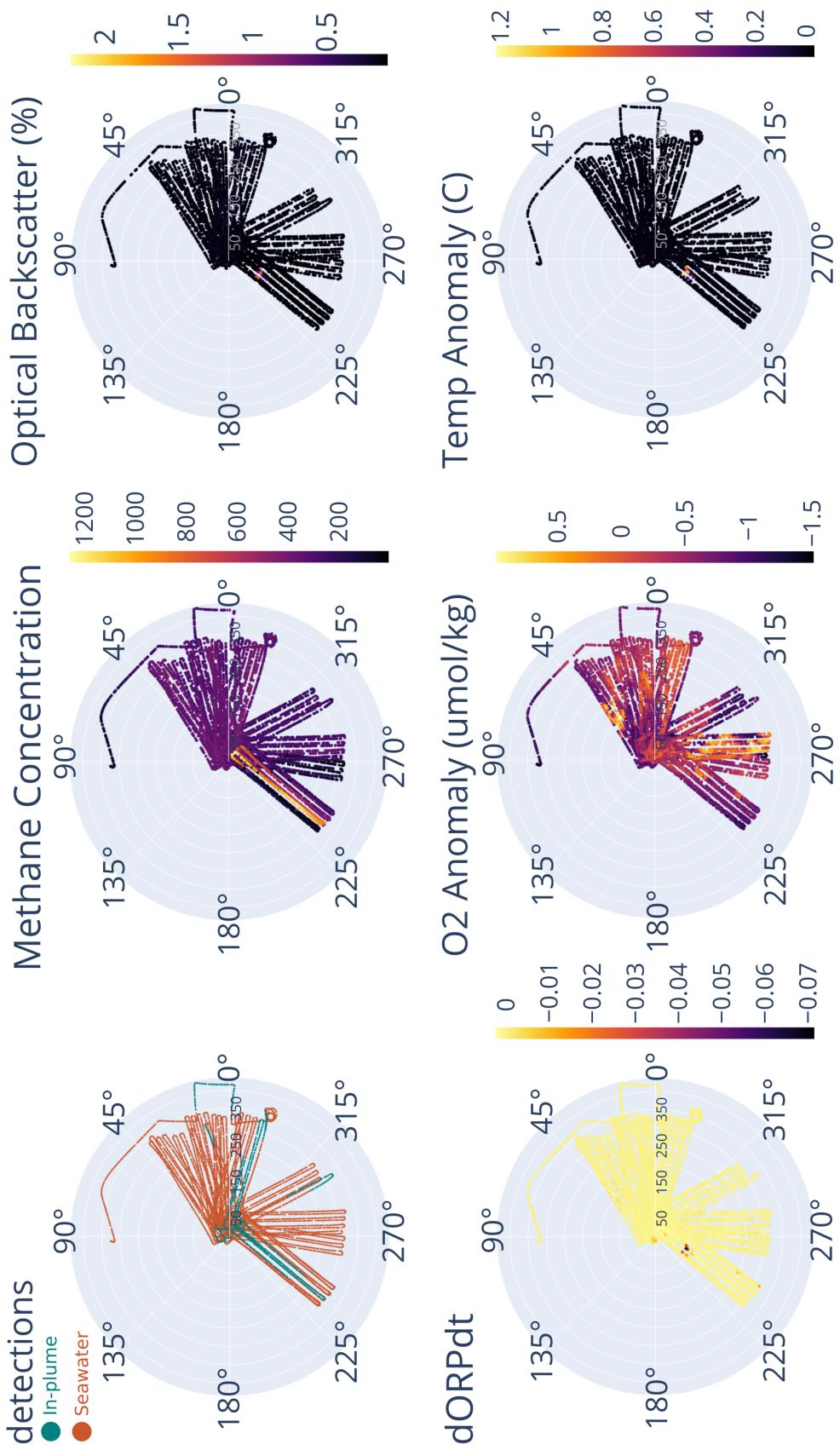


Figure C.4: **AUV *Sentry* dive 611 data.** The binary detections are plotted with methane value from the SAGE methane instrument (time-corrected and converted data to nM), turbidity, ORP, detrended oxygen, and detrended temperature.

C.2 ACOUSTICALLY BROADCAST SCIENCE DATA

During the cruise, a custom live-plotter for acoustically transmitted oceanographic sensing equipment from AUV *Sentry* was developed. The open source code for this plotter can be accessed at https://github.com/expeditionary-robotics/auv_listener. While acoustically transmitted data is severely sub-sampled, the plots can be useful for showing general trends. Here, an example of the data that was transmitted by *Sentry* during dive 607 and the absolute data from that dive are shown in Fig. C.5.

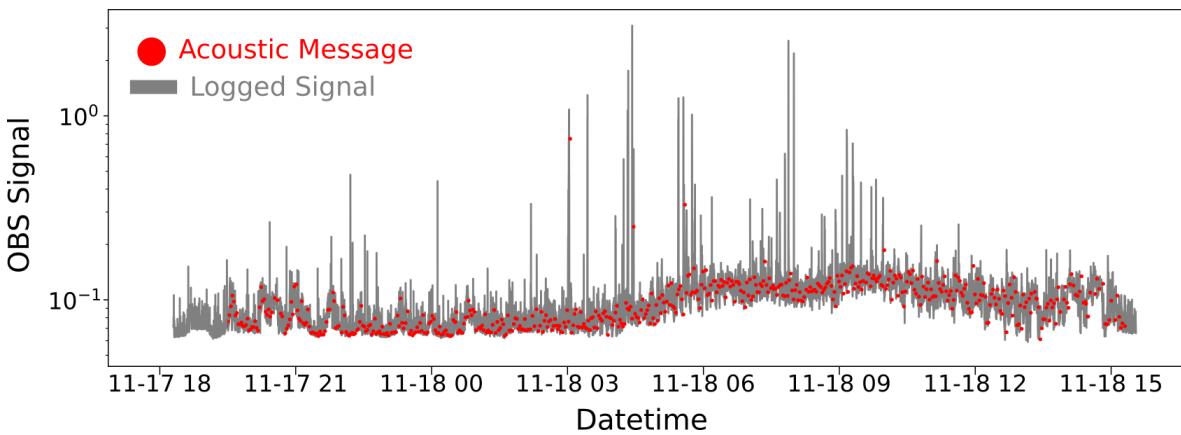


Figure C.5: **Acoustic versus logged OBS messages during AUV *Sentry* dive 607.**

Red points mark the acoustic messages that were transmitted over the acoustic *Sentry* network, plotted, and logged to a shipside computer. Most of the gross structure of the OBS data stream was captured, potentially indicating that the acoustic messages, although severely subsampled, may be sufficient for real-time analysis of vehicle performance.

C.3 CODE AND DATA AVAILABILITY

Rosette data can be found in the Rolling Deck to Repository (R2R) at <https://doi.org/10.7284/909325>. Data analysis tools used for processing *Sentry* data are available on Github at <https://doi.org/10.5281/zenodo.6789105>. *Sentry*, ROV *JASON*, and tiltmeter data will be made available at a future date. Please contact the author with questions.

C.4 AUTONOMY FIELD REPORT

During RR2107, every science team generated a field report of activities as a reference for post-cruise analysis. While much of the report by the autonomy team on the RR2107 has been included throughout the main thesis body, the report as it was generated on the ship, is included here for completeness and is part of the RR2107 overall cruise report.

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

Summary

During cruise RR2107 (November 13 – December 4, 2021; on station November 17 – November 30, 2021), AUV SENTRY, ROV JASON, and a CTD Rosette were used together to map hydrothermal plume surveying, modeling, and tracking in the mid-water (with special focus on 250m altitude from bottom). Buoyant plumes from hydrothermal activity inject significant energy into deep waters and can serve as a mechanism for biogeochemical transport. However, mapping the extent and composition of hydrothermal plumes in the water column is challenging due to their large spatial O(10km) scales and temporal dynamics. This work was aimed at validating and refining scientifically informed probabilistic models of plume phenomenon and using these models to plan trajectories to discover and map plume waters using an autonomous underwater vehicle. Preliminary results indicate that probabilistic, dynamical models are useful for forecasting hydrothermal plume expressions in the water column and can be used effectively to target AUV trajectories. In future work, we will explore how data collected by heterogeneous vehicle teams (AUVs, ROVs, and CTDs) can be used jointly to estimate plume source parameters from mid-water observations and the scientific implications of the resulting models of plume dispersion and nutrient deposition.

Motivation and Scope

To improve expeditionary robotics for science, robots can make use of models of an environment's spatiotemporal dynamics. By explicitly constructing and refining a model the physical environment and using it to inform robotic exploration, we hope to enable more efficient convergence to and sampling of ephemeral, transient targets of interest. For this model-driven approach to be useful for real-time robotic exploration, we need to develop computationally efficient dynamics model of plume phenomena and measures of model uncertainty that can be refined via data-driven inference. In our research, we propose a scientifically-informed probabilistic model that builds a Bayesian inference framework on top of idealized analytical models of plume dynamics. We then optimize sampling trajectories over predictions of plume dynamics from this model to collect scientifically useful observations.

At the hydrothermal vent sites in Guaymas Basin, there is scientific interest in understanding how key compounds, such as methane and ammonium, are distributed and spread throughout the basin ecosystem by plume dynamics. To understand this question, we aim to construct a probabilistic model of hydrothermal plume dynamics in a stratified medium with current-based crossflow. This model can be used to inform sampling strategies of the ROV Jason, CTD casts, and Niskin bottle firings to target chemically diverse observations and samples of plume waters. Additionally, for the autonomous underwater vehicle Sentry, this plume model can be used alongside an optimization-based trajectory planner to produce surveys that target specific portions of the plume phenomena as they evolve due to change currents and ocean conditions. These three sensing modalities have complementary strengths for plume mapping: the ROV JASON enables direct observation of the plume sources; CTD casts and

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking

Prepared by: Victoria Preston, Genevieve Flaspoehler

Niskin bottles enable quantification of the vertical structure of the plume high in the water column; and the AUV Sentry can map the spatial extent and time dynamics of the plume in much higher resolution. By combining the information provided by these platforms, we aim to provide a complete picture of plume evolution, from source parameters to large-scale non-buoyant plume dispersion that can be used to inform scientific models of nutrient cycling and deposition in Guaymas Basin.

Goals

Our goals for the expedition included:

1. Estimate plume source parameters using the JASON ROV and tiltmeter data
2. Use these parameter estimates to form a prior model of plume phenomena and plan surveys over the plume site
3. Update model parameter estimates from Sentry and CTD observations and plan a second, adaptive survey over the plume site.
4. Demonstrate long-distance plume discovery and mapping using standard scientific sensors and experimental methane sensors

Research Activities

Overview of Methodology

To generate sampling trajectories for mid-water plume surveys requires several modules that work together. In this work, we complement a science-informed probabilistic plume modeling technique and trajectory optimization algorithm with several data-processing pipelines. This section provides detailed descriptions of what was tested on RR2107.

Science-Informed Probabilistic Plume Modeling

Morton, Taylor, and Turner's seminal 1956 work demonstrates that hydrothermal plumes can be well described by conservation of momentum, buoyancy, and mass fluxes. Plumes consist of two conceptual parts: a *buoyant stem* and *non-buoyant plume/intrusion layer/neutrally-buoyant intrusion*. The buoyant stem is a positively buoyant structure in the water column that is taller than it is wide (indeed, buoyant stems are well known to only grow a meter for every 10m in height gain as a rule of thumb). The non-buoyant plume forms at the point in which the hydrothermal fluid reaches equal density to the ambient background, and in theory expands infinitely within that uniform density layer.

We use analytical models of idealized plumes in order to ground a probabilistic methodology that allows us to represent uncertainty over our model initial conditions, as well as temporal dynamics. Specifically, we implement three types of model: a *Stationary* plume model based on the work of Speer and Rona (1989) who describe hydrothermal plume models for Atlantic and Pacific basins and assumes

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspohler

no crossflow (i.e., current); a *Crossflow* plume model based on the work of Tohidi and Kaye (2016) who propose a bent-plume model under conditions of crossflow; and a *Multiplume* model, which wraps multiple independent instances of Stationary or Crossflow models and reasons about how those plumes may intersect.

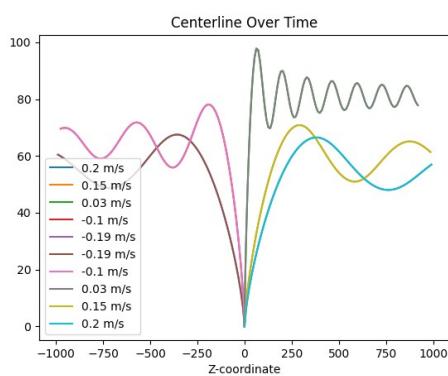


Illustration 1: Crossflow plume centerline under different current magnitude and heading conditions.

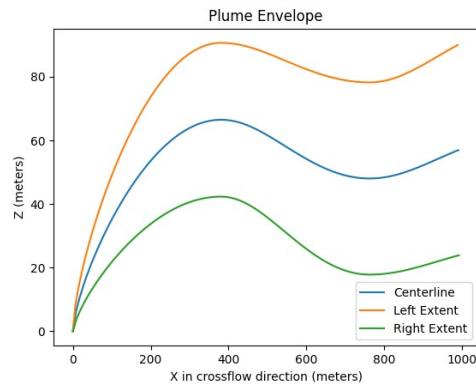


Illustration 2: Vertical cross-section of Crossflow plume model envelope.

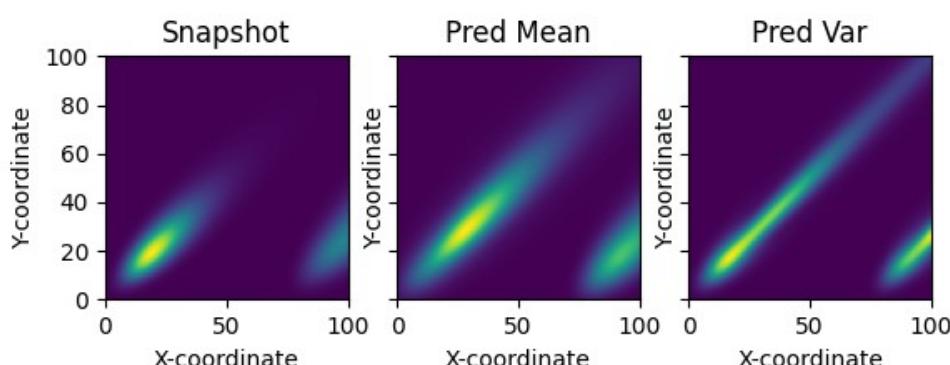


Illustration 3: Example of multiplume model with two crossflow plumes at [0,0] and [75,0]) for some top-down view of a fixed depth. Shows mean and variance from probabilistic samples.

For the purposes of planning in this cruise, we primarily use a Crossflow model, described by a system of coupled ordinary differential equations:

$$E = \alpha \left| \frac{M}{Q} - U \cos(\theta) \right| + \beta |U \sin(\theta)|$$

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

$$\frac{dQ}{ds} = Q \sqrt{\frac{2(1-\lambda)}{M\lambda}} E$$

$$\frac{dM}{ds} = U \cos(\theta) \frac{dQ}{ds} + F \frac{Q}{M} \sin(\theta)$$

$$\frac{d\theta}{ds} = \left(F \frac{Q}{M} \cos(\theta) - U \sin(\theta) \frac{dQ}{ds} \right) \frac{1}{M}$$

$$\frac{dF}{ds} = -Q N_2 \sin(\theta)$$

$$\frac{dx}{ds} = \cos(\theta)$$

$$\frac{dz}{ds} = \sin(\theta)$$

where E is an entrainment coefficient (roughly a mixing value) with coefficients α and β , Q is specific volume flux, M is specific momentum flux, θ is plume centerline angle, F is specific buoyancy, s is a notion of distance along the plume centerline, x and z are planar Cartesian plume centerline coordinates (within the plume crossflow direction frame of reference), U is crossflow magnitude, and N_2 is the Brunt-Väisälä frequency. The latter is computed by comparing plume density and background density profiles.

The initial conditions of buoyancy, momentum, and volume fluxes are set by plume source characteristics of *temperature*, *salinity*, *orifice area*, an *exit velocity*. For the purposes of this cruise, we assume source temperature and salinity are effectively known (end members were previously characterized on cruise NA090 in 2017 and directly measured with ROV JASON on this cruise), but we represent orifice area and exit velocity as probabilistic distributions. The prior distributions for these parameters are set to be uninformative uniform priors over the conservative observed bounds for these parameters (exit velocity: 0.1m/s to 1.3m/s; area: 0.15m² to 3.14m²). Additionally, we are uncertain about the mixing characteristic in the system, and so represent coefficients α and β in the entrainment computation as uninformative uniform priors spanning literature indicated ranges (α : 0.1 to 0.2 and β : 0.1 to 0.8).

Uncertainty over these conditions and parameters is largely driven by the difficulty in directly observing or quantifying them. In order to better estimate these parameters (and therefore better track the plume), we use indirect, in situ plume observations. We employ a Bayesian framework (Metropolis-Hastings) to draw samples of our unknown initial conditions and parameters, push these through our idealized model, and compare the output directly with actual binary detections of plume influenced waters (process described in **Plume Detection**) in the mid-water, with some probability of accepting or rejecting samples as a function of model-observation agreement and likelihood of a sample.

Uncertainty in temporal evolution of the plume (e.g., advective current magnitude and heading) is not explicitly updated in the system tested on the cruise, however current (and background profiles of temperature and salinity) are mean functions of data-trained probabilistic models as described in **Current Modeling** and **Background Profile Characterization** which are updated from direct point observations of those quantities in a separate process.

Plume Detection

Our probabilistic plume model uses positive and negative plume water detections at given locations and times to refine the plume parameter estimates used. Water originating from hydrothermal plumes at Guaymas basin is generally hotter, saltier, more turbid, and richer in compounds like methane and ammonia compared to background water. To build a plume detector, we fuse various scientific sensors that can detect and quantify the differences between plume and background water to produce binary plume detections. Our plume detector is based on the method of corroborated anomalies proposed by Jakuba (2007) and uses seven data streams:

- ORP: the oxidative reductive potential, used as the natural log of the time derivative
- PT: potential temperature
- PS: practical salinity
- OBS: the optical backscatter, used as the natural log
- O2: [partial pressure of oxygen], used as the natural log of the time derivative
- NOPP: methane concentration, measured via laser-based spectroscopy
- HCF: methane concentration, measured via hollow core fiber

After a dive, each data stream is subsampled (interpolated) onto the 1 Hz navigational data, so that each timestamp has an associated location and observation from each science sensor. We apply post-processing steps to several data streams (ORP, OBS, O2) to correct for non-stationarity and specific sensor response characteristics, as noted above. Then, for each data stream, we compute the mean and standard deviation over a 200 sample sliding window and classify a point as anomalous if it lies outside of a ± 2 standard deviation interval. Finally, we classify a sample location as being “in” or “out” of a plume by requiring consensus among multiple sensors within a sliding window of 200-samples. For consensus, we “weigh” each sensor based on sensor reliability and relevance to plume detections. Some measurements, such as the value of the ORP probe, exhibit large anomalies in the presence of plume water; others, such as the practical salinity, are less strong indicators and have lower signal to noise ratios. We therefore allow for sensors to be counted with non-uniform weights when evaluating consensus. Our implementation uses the following sensor weights: ORP: 2, PT: 1.5, PS: 1.5, OBS: 2, O2: 2, NOPP: 1, HCF: 1. A “weight” of 6 is used to classify a particular observation as plume water.

Current Modeling

Deep sea currents of magnitude 0.1-0.5m/s have been observed in Guaymas Basin (Scholz et al., 2019) variably from Northwest to Southwest headings and are largely driven by tidal forces. To capture the influence of currents on the mixing and advection of hydrothermal plumes, we include current magnitude and course in our plume model computation. These current models are driven by data collected from two sources: tiltmeters placed at a fixed location on the ocean floor by the ROV JASON and downward facing DVL measurements collected by the AUV SENTRY. During the cruise, tiltmeter data was primarily used; DVL data will be used later to assess and complement this data stream.

Using the tiltmeter data, we applied a data-driven modeling approach to estimate current magnitude and course functions utilizing Gaussian Processes (GPs), a non-parametric Bayesian model. GPs are fully defined by a mean function and a covariance kernel, which describes the way in which data may be inter-related. We trained two GPs with a zero-mean and radial basis function (RBF) kernels individually on subsampled tiltmeter estimates of magnitude and heading (magnitude model: 100 training iterations, learning rate of 0.5; heading model: 200 training iterations, learning rate of 0.1). The mean magnitude and heading function from the trained GPs was then used within our predictive plume framework for the purposes of planning trajectories on this cruise.

Background Profile Characterization

For the purposes of this work, we assume that salinity and temperature most strongly contribute to the density profile of the hydrothermal plumes and background seawater. To compute the background salinity and temperature profiles of our target sites, we train two GPs (zero-mean and RBF kernel) on sub-sampled salinity and temperature data collected by CTD Rosette casts conducted on the cruise. Both GPs are trained with 100 iterations with a learning rate of 0.1. The trained mean function for both temperature and salinity are used within our model and planning framework for this cruise.

Exit Velocity Estimation

The steady-state buoyant plume model we use depends critically on the properties of the hydrothermal vent source, including the total buoyancy and momentum flux. To estimate these source parameters accurately, we require an estimate of the water velocity as it exits the plume source. There are two primary strategies for measuring plume exit velocity: invasive and non-invasive. Invasive measurements use instruments that can be placed directly in the exiting flow and provide a velocity estimate via, i.e., quantifying the rotation rate of a mechanical spinner. Non-invasive measurement strategies use sensors such as RGB cameras or ADCP to estimate fluid velocity without directly interacting with plume fluid. Passive methods have been shown to provide accurate estimates of plume exit velocity, with a slight bias towards underestimation of velocity in experiments (Zhang et al., 2019).

Among passive methods, particle imaging velocimetry (PIV) is often the simplest to implement, requiring only an onboard RGB camera, and is well suited for plume waters that differ visibly from background water, such as in the black smoker sites in at Guaymas Basin. PIV methods track turbulent parcels that have high cross-correlation values between frames. By tracking many specific parcels over several frames, PIV methods can provide a vector field of velocity estimates that can then be averaged

to provide mean flow in a specific region. Using the 4K video data collected by a MISO camera, we applied MATLAB's open-source PIVLab software to estimate mean fluid velocity at the orifice of each vent site. The MISO camera was aligned using JASON's manipulator to ensure that the camera face was parallel to the flow exiting the vent source and approximately 5m from the plume. The 10cm-spaced laser points from JASON were used to estimate the scale of features in the camera image.

We additionally collected several opportunistic datasets of the water exiting the plume source with JASON's BlueView multibeam sonar sensor and onboard science cameras. We plan to use these datasets to explore acoustic strategies for passive exit velocity estimation.

Trajectory Optimization

Given a probabilistic plume model, the second key component of our autonomous decision-making stack is the trajectory optimization method, which uses the plume model to select Sentry trajectories that have high probability of intersecting plume water.

The trajectory optimizer consists of three primary components:

- **Parameterized trajectories:** Parameterized trajectories form the basis of the trajectory optimizer. A parameterized trajectory is a function that takes in a set of trajectory parameters, e.g., length, width, location, orientation, and resolution specifying a lawnmower parameter, and produces a set of waypoints that define a Sentry trajectory. In our experiments, we leveraged both lawnmower (e.g., standard boustrrophedonic surveying patterns) and spiral-shaped parameterized trajectories.
- **Reward function:** The reward function quantifies which locations are useful for SENTRY to collect observations from. This reward function must encode the scientific objectives of a mission, such as desire to collect observations of locations that have high probability of containing plume waters or desire to collect a diverse set of observations from different concentrations of plume waters. The reward function can query our probabilistic plume model to estimate the probability of detecting plume water at various locations and times along a trajectory. In our experiments, we used a reward function that encouraged SENTRY to spend time in regions with high probability of plume detection.
- **Optimization method:** The optimization method uses both the parameterized trajectories and the reward function to produce a final trajectory to accomplish the scientific objective encoded in the reward function and respects vehicle constraints, such as time budget and safety regions. Starting from an initial guess of trajectory parameters, the optimizer uses numerical gradient-based methods to find parameters that result in maximal reward. We use the Trust Region method with constraints in Python's scipy library and run for 25 iterations or until convergence.

The final output from the trajectory optimizer is a sequence of lawnmower trajectories, each of which have been optimized to maximize the probability of detecting plume water for a given start-time and

plume model forecast. This trajectory is loaded into Sentry's mission planning software as a sequence of waypoints and executed over a 14-18 hour mission.

Vehicles, Equipment, and Instrumentation

To complete this work, we primarily made use of the AUV SENTRY and its typical science instrument suite for mid-water flights over survey targets. Our work was additionally supplemented by excursions with ROV JASON and opportunistic CTD Rosette casts.

UDP Acoustic Communications Monitoring

To listen to USBL location pings during all missions with SENTRY and relevant CTD transects, and monitor AUV SENTRY science instruments during all SENTRY missions, we developed UDP listeners within the SENTRY network architecture that subscribed to various sockets and shunted messages received to sorted files. This allowed us to diagnose progress during missions in terms of plume-intersections, and allowed for opportunistic small changes in missions (or informed decision-making when mission timing needed to change). This also helped to catch instrument errors and send power reset commands to instruments. Acoustic messages for specific queues were received every 40-120 seconds, depending on the number of active acoustic queues. Occasional dropout of messages due to distance from ship were also observed.

AUV SENTRY

SENTRY is equipped with an optical backscatter (OBS) unit, optode, oxidation reduction potential (ORP) probe, and CTD which served as our primary sensing targets. Specific to this cruise, we additionally had access to water-tracking acoustic returns from the onboard DVL used for navigation in bottom-lock mode, 2 1L Niskin bottles, 2 experimental instruments (NOPP and HCL, used mutually exclusively), and a downward-looking MISO GoPro camera.

SENTRY was used to fly in both altitude-hold and bottom-lock mode from 30m-120m in (average) altitude. Niskin bottles were fired in both scheduled and opportunistic fashions. The MISO camera was deployed on two dives, once in a snapshot-photo mode, and another in 4k video shooting mode. The relevant configuration for our dives with SENTRY:

- sentry607 (V/G planned) – standard science suite, NOPP
- sentry608 (V/G planned) – standard science suite, NOPP, Niskins
- sentry609 (scrubbed, redive by SENTRY team) – standard science suite, HCF, Niskins, MISO
- sentry610 (V/G planned) – standard science suite, NOPP, water track, Niskins
- sentry611 (V/G planned) – standard science suite, HCF, water track, Niskins, depth-hold mode
- sentry612 (collaboratively planned) – standard science suite, NOPP, water track, Niskins, MISO
- sentry613 (SENTRY team planned) – standard science suite, NOPP, water track

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

ROV JASON

JASON is equipped with a CTD, optode, and OBS mounted on the chassis, a temperature probe that can be maneuvered using the manipulator, and multiple cameras (including a brow and arm mounted MISO and a 4k science camera). JASON is also equipped with several physical samplers, including two Niskin bottles and high-temperature titanium major samplers.

Our primary objective for JASON was to characterize the plume source properties. The high temperature wand was inserted into the largest plume orifice to measure end member temperature. A peak reading of 340C was observed near the Chimney #1 black smoker site. No salinity end member was directly observed, however previous work has reported slightly elevated salinity (on the order of 0.3 PSU) salinity from Guaymas vents.

Additionally, we were opportunistically able to observe the height of the plume nonbuoyant layer as JASON descended to depth, via a spike in the OBS sensor and visible black smoke in the JASON cameras. Smoke was visible to the human eye starting at 1480m and extending to 1700m. This is largely corroborated by CTD Rosette casts. Relevant dives to our work include:

- JD1388-JD1389 (tiltmeter deployment, MISO Arm camera, MISO Brow camera, temperature probe)
- JD1390 (tiltmeter deployment, MISO Arm camera, MISO Brow camera)
- JD1392 (tiltmeter deployment, MISO Brow camera)
- JD1393 (tiltmeter deployment, MISO Brow camera)
- JD1394 (MISO Brow camera)
- JD1395-JD1396 (tiltmeter recovery, MISO Brow camera, MISO Arm camera)
- JD1398 (MISO Brow camera)
- JD1400 (MISO Brow camera)

CTD Rosette Casts and Transect

CTD Rosettes are equipped with a SBE43 CTD and oxygen sensor, a transmissometer, a fluorescence sensor (traded with HCF for casts 06, 08, 09, 10, 11), and a bottle carousel. Salinity and temperature profiles were used within our framework to establish background profiles. Cast information:

- CTD01 – (27.40836, -111.38910); within 200m of Chimney 1 at the graben-ridge site
- CTD02 – (27.41014, -111.38818); shallow water transect through the OZM at the graben-ridge
- CTD03 – (27.40964, -111.38232); 600m NE of the graben-ridge site
- CTD04 – (27.50896, -111.68166); Ring vent

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

- CTD05 – (27.50908, -111.68156); Ring vent
- CTD06 – (27.40796, -111.38834); slow HCF cast SE of graben-ridge site
- CTD07 – (27.40438, -111.37812); 1.2km SE from graben-ridge site
- CTD08 – (27.40834, -111.38924); Repeat of CTD06 cast
- CTD09 – (27.20106, -111.39796); 22km south of graben-ridge site
- CTD10 – NW transect 16km from graben-ridge to within 3km
- CTD11 – NW transect starting from CTD10 stop across the graben-ridge, to about 3-4km past

In the last two CTD casts (CTD10 and CTD11), the CTD Rosette was towed at ~1650m depth.

Absolute position of the CTD cast was tracked with a USBL beacon. This transect will be used as part of a future model validation procedure.

NOPP and HCF

Experimental instruments NOPP and HCF were interchangeably used on AUV SENTRY dives and CTD casts (for HCF); the datastreams of which we have incorporated into our binary plume detection algorithm. Instrument data is interpolated onto the timestamp of the AUV SENTRY or CTD Rosette cast for use in our system.

Tiltmeter

Two TCM-3 tiltmeters from Lowell Instruments were deployed using the ROV JASON. One tiltmeter (tiltmeter A, serial number: 2108300) was placed and recovered over multiple dives, providing dive “snapshots” of current. Another tiltmeter (tiltmeter B, serial number: 2110300) was deployed at a site on the graben-ridge and left over multiple days and recovered at the end of initial operations in N. Guaymas). A lead weight with an eye-bolt was used to anchor the tiltmeters. Deployment locations and durations indicated:

- Tiltmeter A; deployed JD1389; recovered JD1390 total duration 28hrs; (27.4006177, -111.3985321, 1832m initially; moved to 27.4002362, -111.3962494, 1854m); primary dataset used in cruise predictive modeling
- Tiltmeter B; deployed JD1389; recovered JD1396 total duration 6days 15hrs; (27.4006177, -111.3985321, 1832m)
- Tiltmeter A; deployed JD1392; recovered JD1392 total duration 10hrs; (27.4149571, -111.3873036, 1840m)
- Tiltmeter A; deployed JD1393; recovered JD1393 total duration 20hrs; (27.5001163, -111.6832265, 1732m); sunk into the mud at Ring Vent and literally read nothing

RR2107

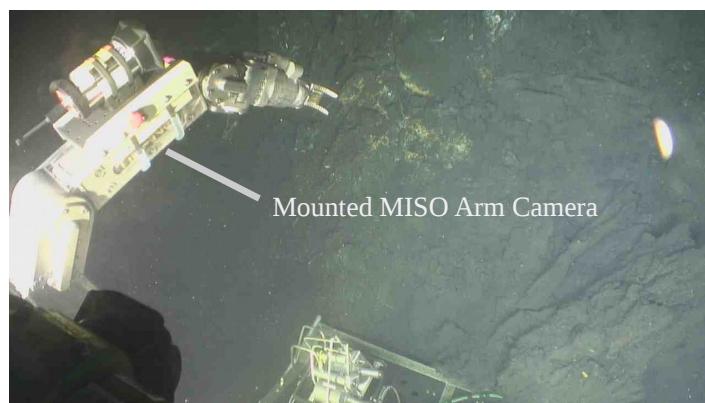
Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler



MISO

Two MISO GoPro cameras were used on ROV JASON and AUV SENTRY operations. The “Arm” camera was used on both ROV JASON and AUV SENTRY in order to film black smokers and image seafloor transits (respectively). The “Brow” camera was used primarily to take pictures during ROV JASON dives. Specifically:

- Brow Camera – ROV JASON – All dives except JD1397
 - All modes set to 12MP 1pic/5seconds mode
- Arm Camera – ROV JASON – Dives JD1388, JD1389, JD1390, JD1395, JD1396
 - All modes set to 4k video
- “Arm” Camera – AUV SENTRY – Dives sentry609, sentry612
 - sentry609 – 12MP 1pic/2sec mode; quality poor (no dimmer used)
 - sentry612 – 4k video mode; quality decent (no dimmer used)



RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspohler

Site Descriptions

Dives took place in four locations, the N. Guaymas ridge/graben-ridge, the N. Guaymas Ring Vent, S. Guaymas, and a long-transect from between S. Guaymas back to the N. Guaymas ridge/graben-ridge.

SENTRY flew in all locations; JASON also examined all locations save for the S to N transect.

- **North Guaymas Ridge:** This was the primary site for Jason and Sentry operations. The site consists of a ridge approximately 600m in length. Six separate high-temperature venting sites have been observed along the ridge along features that are ~45-75m high. Of these, Chimney 1 (located on the far southern end of the ridge with coordinates (27.40926223, -111.38931794), orange arrow) and Chimney 2 (located on the far northern end of the ridge with coordinates (27.412645, -111.386915), red arrow) were well characterized by the JASON ROV. Both chimneys are black smokers and characterized by source fluids of ~350C temperature anomaly and heightened salinity. These two chimneys were the focus of Sentry's missions and adaptive mapping was focused on Chimney 1.
- **North Guaymas Ring Vent:** The ring vent site is a smaller, low-temperature diffuse flow site with several features arranged in a ring shape. AUV SENTRY flew a tight survey over a “blow

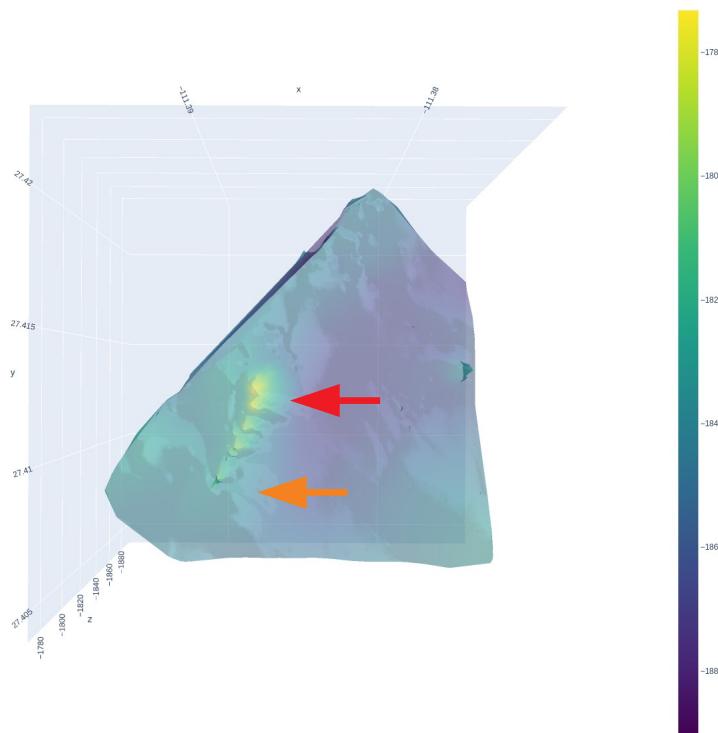


Illustration 4: Overview of graben-ridge bathymetry. Red arrow: Chimney 2; orange arrow: Chimney 1.

RR2107
 Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
 Prepared by: Victoria Preston, Genevieve Flaspoehler

hole” observed by ROV JASON as well as a survey over the entire ring vent. ROV JASON performed a slow heat-flow probe survey of ring vent, and examined structure to the SW.

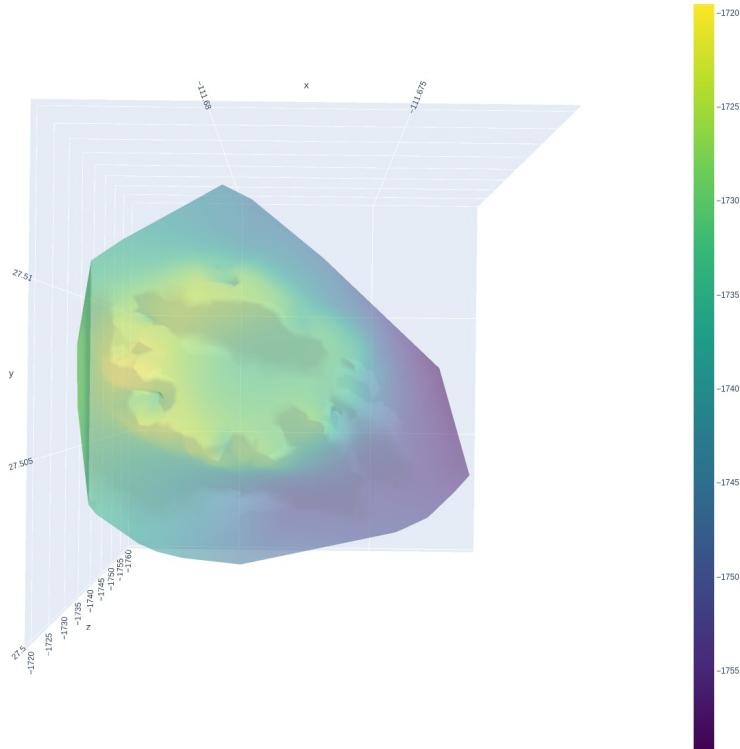


Illustration 5: Overview of Ring Vent bathymetry.

- **South Guaymas Vent Clusters:** S. Guaymas has been previously well studied and marked. SENTRY flew over two clusters of noted vents in the northern part of the site, including a cluster with Rebecca’s Roost, and Cathedral Hill. JASON examined bacterial mats to the South. We did not fly any adaptive trajectories at this site, but will investigate applying multiplume models to the data.
- **South Guaymas to North Guaymas Transect:** The southern and northern Guaymas hydrothermal sites are separated by $\sim 0(50\text{km})$ of flat basin with little known hydrothermal activity. As a final exploratory mission, Sentry flew a transect from a midway point between the Southern and Northern Guaymas sites ($\sim 22\text{km}$ southeast of the Northern Guaymas Chimney 1), ending the trajectory in the graben feature to the northwest of Chimney 1.

Preliminary Results Science-Informed Probabilistic Plume Modeling

As discussed in *Overview of Methodology: Science-Informed Probabilistic Plume Modeling*, we trained a model on data from sentry607 over Chimney 2 with 625 parameters samples (125 samples used for chain burn-in) to use in planning several of the future sentry dives. The prior and posterior initial condition maximum likelihood expectations were as follows:

- Prior Exit Velocity: 0.699m/s > Posterior Exit Velocity: 0.58m/s
- Prior Source Area: 1.654m² > Posterior Source Area: 0.82m²
- Prior Alpha: 0.148 > Posterior Alpha: 0.15
- Prior Beta: 0.451 > Posterior Beta: 0.19

Notably, in the training process chain mixing is of serious concern. Ideally, more samples would be used to estimate each of these parameters. Due to pushing these samples through the analytical model, the composite distribution of plume detections which maps to initial condition samples is relatively complex, and examining this distribution effectively takes time. Moreover, the only signal to compare a model to real observations is a binary notion of “in” and “out” of a plume (as discussed in *Overview of Methodology: Plume Detection*), which is helpful, but by no means necessarily “expressive” in the way that continuous-valued observations may be.

Plume Detection

Our binary plume detector using corroborated anomalies was applied the Sentry scientific data for all cruise dives. Initial results for dives sentry607, sentry608, sentry610, and sentry611 are shown below, with plume detections shown in red and non-detections shown in blue. By requiring six sensors to confirm a plume detection, our classifier is quite conservative, detecting plume water only when there was very strong buoyant stem signal near the location of the source (yellow arrow). This is useful when fitting our plume MTT model from observations, as these strong detections help to quickly constrain the entainment rate and source parameters. In the future, we plan to explore detector performance as a function of the number of corroborations required and anomaly detection algorithm.

RR2107
Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

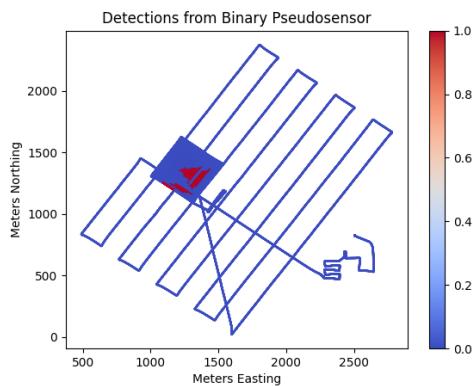


Illustration 6: sentry607

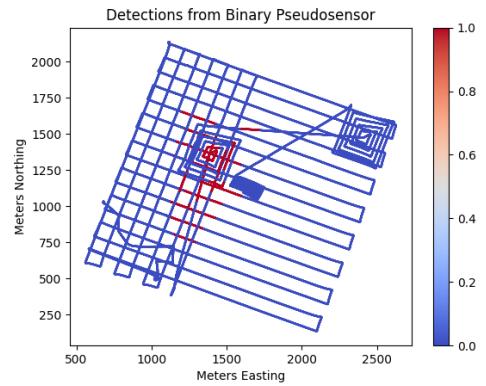


Illustration 7: sentry608

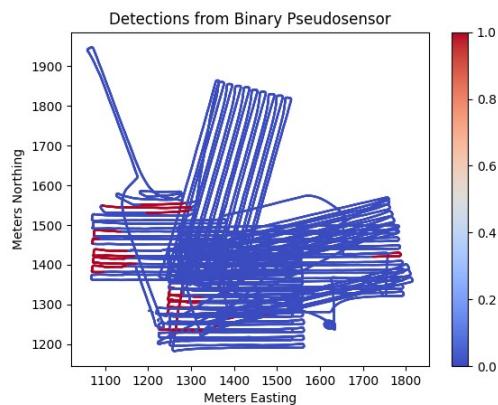


Illustration 8: sentry610

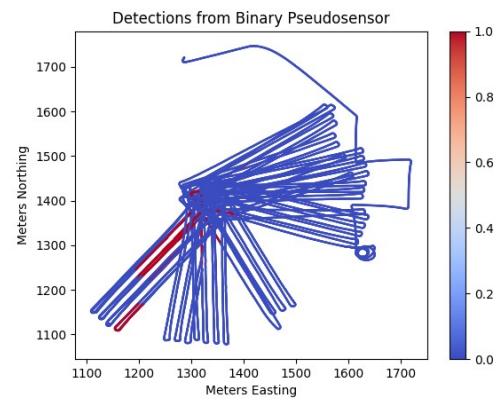


Illustration 9: sentry611

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

Current Modeling

Current magnitude and course functions trained using the GP method described in *Overview of Methodology: Current Modeling* demonstrated a 12hr periodic cycle that tracked with predictive tidal information for the city of Guaymas as computed by CICESE.

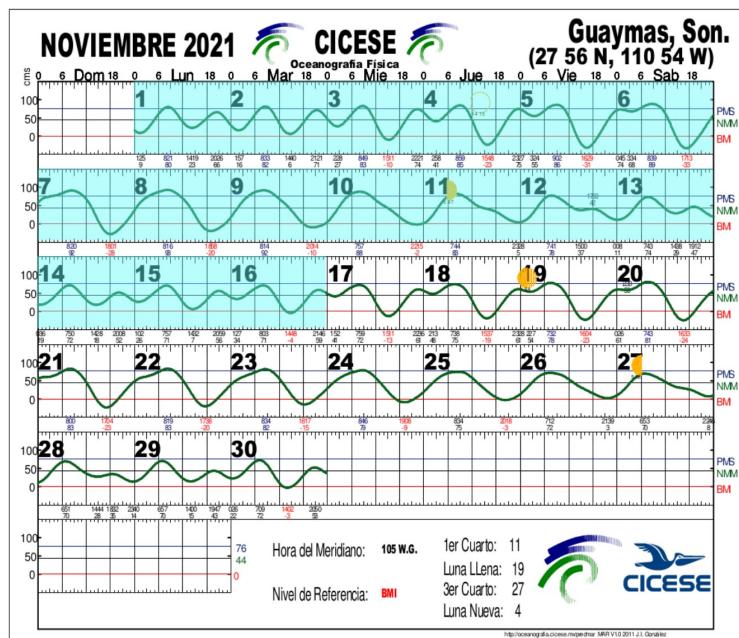


Illustration 10: Reference tidal chart; cruise dates unhighlighted.

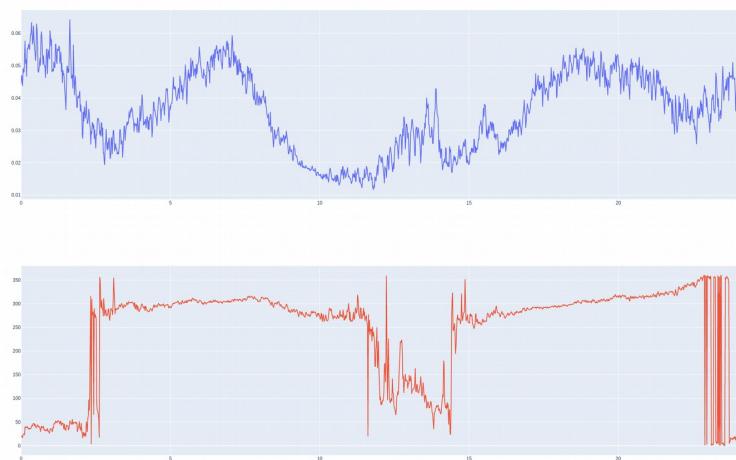


Illustration 11: Magnitude (top) and Heading (bottom) of tiltmeter B data flattened and averaged over a 24 hour span.

Background Profile Characterization

A GP for each of salinity (PSU) and potential temperature (C, ITS90) was trained with a subsample (every 100 samples of spatially ordered data) of all CTD casts (excluding the transect) in order to get a basic profile for use in our system. GPs were trained with learning rate 0.1 and 200 iterations. Agreement between the GP mean and the sampled data was good.

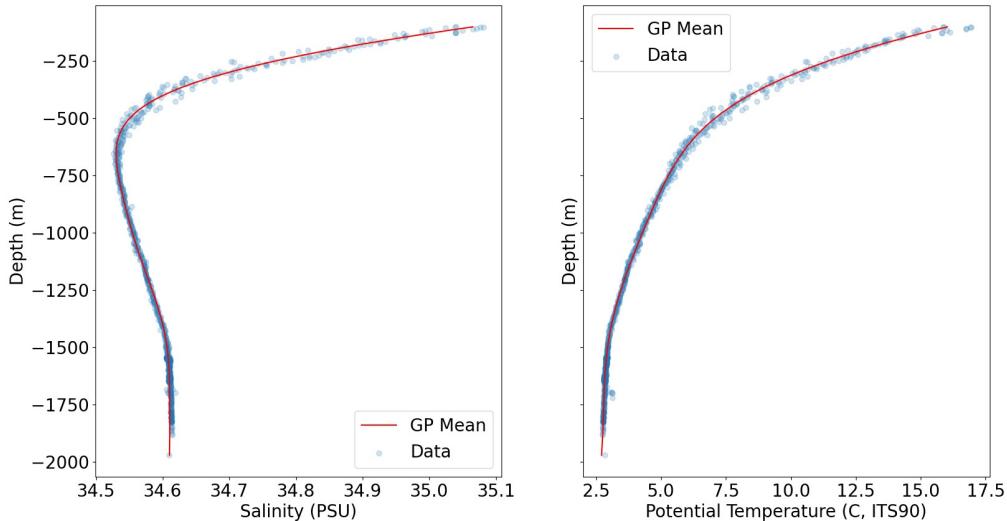


Illustration 12: GP mean of salinity and temperature profiles, with training data plotted.

Exit Velocity Estimation

To ground exit velocity estimates from our model, we have started the process of using a non-invasive PIV technique. Using PIVLab through MatLab, we have analyzed some footage from the Arm MISO camera taken during JD1390 at (27.4018606, -111.3991182, 1809m). Using the JASON 10cm lasers to calibrate the spatial scale in the image, and using auto-calibration in PIVLab to set processing parameters, we found exit velocities estimates well above 0.7m/s (up to 1.33m/s); although over an entire plume region (1-2m wide, 1m tall plume structure visible), the net upward velocity was estimated to be only 0.04m/s.

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking

Prepared by: Victoria Preston, Genevieve Flaspoehler

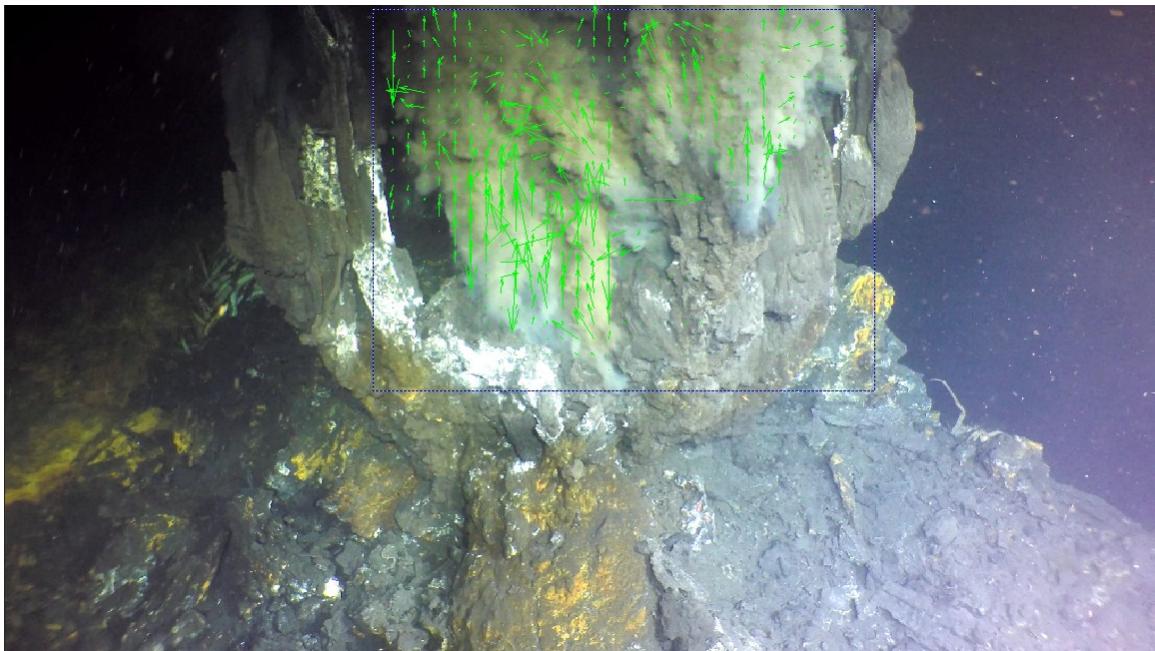


Illustration 13: Example image from MISO Arm Camera of black smoker plume with ROI and PIVLab computed vectors in place.

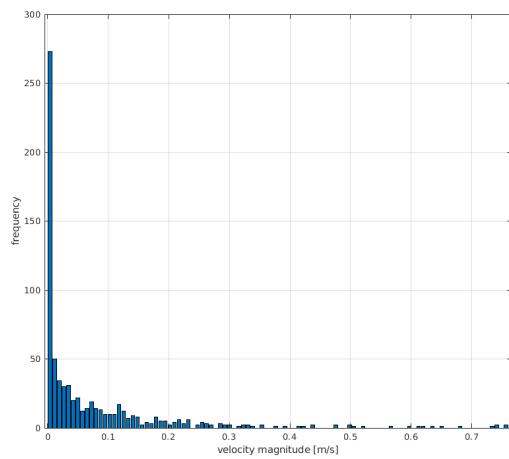


Illustration 14: Histogram of velocity magnitude detected in above frame.

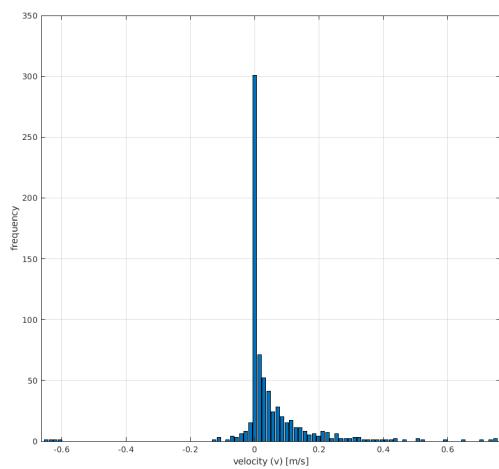


Illustration 15: Histogram of vertical velocities detected in the frame above. Long tail potentially indicates source magnitudes.

Trajectory Optimization

Starting on sentry610, optimized trajectories based on our scientifically-informed probabilistic models were used. A complete list of trajectories optimized by our system is below. Note that all origins are given in meters easting and northing of reference point (27.4, -111.4):

- sentry610
 - lawnmower 1 – origin (1379.1730, 1354.6755); orientation 75.5288
 - lawnmower 3 – origin (1301.266, 1305.1493); orientation 14.5283
 - lawnmower 4 – origin (1275.264, 1319.538); orientation 0.5
 - lawnmower 6 – origin (1389.952, 1403.596); orientation 109.073
- sentry611 (all lawnmowers generated for possible dive window; only subset used)
 - lawnmower 1 – origin (1293.265, 1394.562); h 74.947m; w 315.493m; orientation 31.522
 - lawnmower 2 – origin (1293.420, 1394.695); h 73.370m; w 327.982m; orientation 33.139
 - lawnmower 3 – origin (1291.652, 1395.314); h 75.008; w 300.045; orientation 21.446
 - lawnmower 4 – origin (1292.102, 1396.304); h 74.977; w 327.908; orientation 13.468
 - lawnmower 5 – origin (1298.069, 1378.418); h 74.924; w 300.704; orientation 0.944
 - lawnmower 6 – origin (1292.006, 1399.954); h 74.997; w 299.991; orientation 343.466
 - lawnmower 7 – origin (1295.508, 1403.481); h 74.987; w 299.942; orientation 297.347
 - lawnmower 8 – origin (1298.171, 1404.231); h 75.004; w 309.752; orientation 270.896
 - lawnmower 9 – origin (1302.756, 1402.395); h 75.035; w 299.887; orientation 230.484
 - lawnmower 10 – origin (1303.977, 1403.934); h 74.749; w 299.972; orientation 230.423
 - lawnmower 11 – origin (1299.914, 1406.229); h 75.003; w 299.999; orientation 260.315
 - lawnmower 12 – origin (1300.734, 1405.879); h 75.000; w 300.000; orientation 256.986
 - lawnmower 13 – origin (1301.007, 1403.149); h 74.999; w 299.982; orientation 247.109

Notably, we see that orientation and origin are among the more variable in the optimizer. This is likely because for a given depth, current magnitude and heading have an outsized effect on the plume location, but doesn't necessarily change (at the magnitudes we see) the overall size of the plume.

It is also notable from an operations standpoint that planning these trajectories is sensitive to dive schedules. Working closely with the SENTRY team and Chief Sci was critical for assembling trajectories that were timely and correct. To assist with this in the last optimized-trajectory dive, we

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

assembled a list of many trajectories, and then selected the relevant subset when the dive schedule became more concrete.

Future Investigation

An incredible amount of data was collected during the cruise relevant to the further development of scientifically-informed probabilistic models for spatiotemporal plume surveying, monitoring, and tracking, as well as sample optimization and decision-making. Our next steps will be to thoroughly assess the techniques used in the cruise from all collected SENTRY dives, and then investigate how well these methods would have performed under other sensing modalities (e.g., if only CTD Rosette casts were available, if only JASON operations data was used, if only two SENTRY dives would be available at a site, etc.). This validation and development pipeline will involve both using there data directly as well as further simulation of plumes for bulk assessment.

Future work that builds on the techniques in this cruise will involve developing a new type of probabilistic science-informed model that rather than takes a hierarchical approach as done here, instead embeds science models more implicitly within a probabilistic representation. Further decision-making work will involve more completely considering multiple levels of operations planning on a ship. For instance, in this work we focused specifically on placing a single vehicle in the right place at the right time, but we could extend this to consider when and where a CTD Rosette cast could be placed relative to a vehicle dive, weigh which type of vehicle or what information is necessary to better place a CTD Rosette or the AUV SENTRY, or increase general adaptability of single vehicle operations.

Several “engineering” tasks are of interest to pursue as well, including further extraction of current information from DVL water track pings and further using PIVLab to analyze both MISO Arm camera and JASON science camera data. Opportunistic Blueview multibeam data of plume flares was also captured at source vents and would similarly be interesting to compare against the visual methods.

While we anticipate a field robotics publication of this cruise work, and that the data will be used to advance algorithmic contributions, we also anticipate scientific contributions with respect to the scientifically-informed probabilistic model development anticipated after the cruise, and in collaboration with several outside scientists. Further, we anticipate assisting with several projects that intersected with our work during the cruise, including ammonia/ammonium and methane mapping in plume structures with A. Michel and S. Wankel. Finally, this data will support several undergraduate research projects advised by Victoria, as well as be components to both Victoria and Genevieve’s PhD thesis work.

References

Jakuba M. *Stochastic mapping for chemical plume source localization with application to autonomous hydrothermal vent discovery*. Doctoral Thesis, Massachusetts Institute of Technology and Woods Hole Oceanographic Institution (2007).

Morton B. R., Taylor G., Turner J.S. *Turbulent gravitational convection from maintained and instantaneous sources*. Royal Society Publishing 234 A (1956).

Scholz F., Schmidt M., Hensen C., Eroglu S., Geilert S., Gutjahr M., and Liebetrau V. *Shelf-to-basin iron shuttle in the Guaymas Basin, Gulf of California*. Geochemica et Cosmochimica Acta 261 (2019), p. 76-92. <https://doi.org/10.1016/j.gca.2019.07.006>.

Speer K., and Rona P. *A model of Atlantic and Pacific hydrothermal plume*. Journal of Geophysical Research 94 C5 (1989), p. 6213-6220.

Tohidi A., and Kaye N. *Highly buoyant bent-over plumes in a boundary layer*. Atmospheric Environment 131 (2016), p. 97-114. <https://dx.doi.org/10.1016/j.atmosenv.2016.01.046>.

Zhang X., Lin J., and Houshuo J. *Time-dependent variations in vertical fluxes of hydrothermal plumes at mid-ocean ridges*. Marine Geophysical Research 40.3 (2019), p. 245-260.

Appendix: Cruise On Station Notes

[20211117] CTD01 0430-0630 ship time; see turbidity spikes with some temp variance from 1800-1600m; site co-located with Chimney #1 in archival information from NA090.

[20211117] Sentry607 0900 ship time in water, multibeam until 1745 ship time, then starts waffles 1815 ship time at 60m in alt hold mode. See ORP dips at UTC 1915, 2015, 2100, 2150, 2330, 20211118 0100, 0215, big excursion at 0303, and so on.

[20211117] JD1388 aborted because of manip ground fault. Dive scrapped until Nov 18, 1000 ship time.

[20211117] Recommended that we talk to Shannon Walker at NOAA about plumes, point-casts with an ORP over diffuse flows, etc.

[20211118] Sentry607 ends 0420 ship time and on deck at 1000 ship time. Data IP is 100.124.34.66 (ship) and 192.168.100.9 (sentry).

[20211118] JD1389 dive aborted after tiltmeters dropped.

[20211118] Sentry608 dive planning handed to Zac at 1800 ship time; request for Niskin fires at (27.40782, -111.3866) and (27.4111, -111.3820)

[20211118] Sentry608 in water at 2300 ship time

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

[20211118] JD1390 to go in at 0000 20211119

[20211119] JD1390 watch notes: on descent notice at 1520m it's very smoky, at 1440m we are not in the plume. Tiltmeters are already deployed; pick up A at 1050UTC and place it at 1210UTC.

[20211120] Recommended to check out the work of Mary Lilley, Dave Butterfield, Gretchen Ugreen, Shannon Walker, and Tamara Baumgarten who are all "plume people"

[20211121] JD1391 watch notes: smoke seems to be blowing east observed at 0934UTC

[20211121] Current estimation is a challenge; there isn't too much trustworthy current data yet; going to attempt to train current and heading functions as GPs from data that has been collected from tiltmeter A. Will use scipy interpolate function to extract curves from the GP.

[20211121] Planning sentry609 trajectories for 1600 – 2000 potential dive window; want to train model on SENTRY data from last dive.

[20211121] Sentry609 aborted due to servo error.

[20211122] Working in PIVLab on MISO Arm camera data from JD1390. 100 frame snippet grabbed from plume film. Apply PIVLab to restricted region (Load Video > Select ROI > Calibrate > Train PIV Model > Analyze > Statistics). Observe as estimate of 0.33m/s up to 1.277m/s exit velocity estimate, with 0.2m/s mean.

[20211123] Training models from sentry607 data:

- CrossflowModelICU8_sentry607: Bayesian; 200 iters, 1e-15 thresh, 50 burnin, every 10 obs; priors V, A, alph, bet = (0.699, 1.654, 0.148, 0.451); posterior V, A, alph, bet = (0.66, 1.52, 0.154, 0.3). 43 samples accepted. Definitely seems like more samples in the chain are necessary.
- CrossflowModelESTX_sentry607: Bayesian; 500 iters, 1e-15 thresh, 100 burnin, every 10 obs; priors same as previous; posterior V, A, alph, bet = (0.44, 0.692, 0.146, 0.219). 77 samples accepted.

[20211123] Sentry610 dive plan to interleave adaptive and naive lawnmowers together based on current estimates drawn from tiltmeter data.

[20211123] Sentry610 mission looked good, but it was much more behind our timing estimates than expected.

[20211124] Training models from sentry607 data:

- CrossflowModel5WNG_sentry607: Grid; 10 samples/param (10^4 total evaluations); only gets to 5 iterations over many hours; very little movement from initial priors, the means of all of them.

[20211124] Note that sentry610 track needs to be adjusted in the 3rd lawnmower due to drift at 0154UTC

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

[20211125] JD1395 watch notes: observe the current moving due west at 0755UTC; smoke lens detected at 1510m at (27.410189, -111.388381); smoke layer looks like it exists from 1510m-1620m.

[20211125] Sentry611 mission looking ok, asking Joe to fly the vehicle as high as possible.....some comms errors later and rest of mission flown at 120m.

[20211125] JD1395 watch notes: observe temperature at vent to be 330C at (27.413214, -111.38686, 1790m). At heading 104 see a chimney to the left at 1153UTC.

[20211125] Training models from sentry607 data:

- CrossflowModel071D_sentry607: Grid-Search; 5 samples per parameter (total of 625 combos); prior same as before; posterior V, A, alph, bet = (0.1030, 0.1612, 0.1155, 0.1359).
- CrossflowModelH727_sentry607: Bayesian; 625 samples with 125 burnin; same prior as before; posterior V, A, alph, bet = (0.58, 0.82, 0.15, 0.19) --- electing to use this model moving forward

[20211125] sentry611 things are behind at 0432 ship time; 200m failed so flying at 120m, running out of time so skipping rest of naive lawnmower to get in a few tracklines in the last adaptive path.

[20211126] working on extracting current information from DVL. Prelim results show good agreement on magnitude, but heading is messy, potentially due to the really small magnitude of everything.

[20211126] planned sentry612 survey of S. Guaymas over two venting mounds

[20211127] wrote a haiku: flying through the plume; firing all of the Niskins; the bottles were closed

[20211128] discussing a long transect with SENTRY and CTD through the plume given far-CTD still saw the plume...

[20211129] Suggested to us to look us some of the work of Matthew Albert at Scripps who thinks about the transport and impact of tailings plumes in the deep sea

[20211129] hypothesize that SENTRY should see fallout “rain” before CTD sees the plume top if we start far enough away; unclear if we have

[20211129] see SENTRY obs climb very high then drop; schmutz or real signal?

[20211130] CTD transect recovered

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

Original Dive Planning Document

SENTRY Dives on Graben Ridge

Dive 1: Wednesday, Nov 17 – Model Constraining Dives

Allows for model seeding (current confirmation, initial condition setting)

Hrs 0-7: Multi-beam survey @ 60m

We can use science data from this mission, but specifications to be coordinated by Anna/Zac

Hrs 7-15: Reverse pyramid waffle-spirals (ideally depth hold)

4x lawnmowers over Chimney 2 at 100mx100m, 10m resolution, 30m altitude (2 waffles)

2x lawnmowers over Chimney 2 at 200mx200m, 10m resolution, 60m altitude (1 waffle)

1x spiral over Chimney 2 at 400mx400m, 20m resolution, 90m altitude

Navigate back to center (Chimney 2) and ascend at end of dive

Dive 2: Date TBD – Model Validation Dives

Allows for confirmation of current model and initial conditions

- Coordinate with Anna, but time in the dive to fly low and slow over the plumes along the ridge in order to best-test instruments
Hrs 2-14: Medium-resolution extent-constraining spirals
- V&G provide a set of chained spirals with center-offsets that track the heading of the current

Dive 3: Date TBD – Planning Optimization Dives

Allows for a trial of the trajectory optimizer with validated model

- Hrs 0-2: Super low ridge fly-by
- Coordinate with Anna, but time in the dive to fly low and slow over the plumes along the ridge in order to best-test instruments
Hrs 2-18: High resolution plume tracking
 - V&G provide a set of chained lawnmowers at medium resolution potentially at multiple heights (to be determined from previous dives) that attempt to closely track the plume maxima over time
 - Navigate back to Chimney 2 and ascend at end of dive

Further dives:

- If further dives on site are available, further validate models (similar to Dive 2 structure) and attempt online adjustments (similar to Dive 3 structure with semi-real time shifts allowed)

SENTRY Dive(s) on Ring Vent or Sonora

Dive A: Date TBD

Have initial data from model confirmation and attempt to transfer model + trajectory training to new environment. Allows opportunity for compelling online planning.

Hrs 0-2:

- 2x lawnmowers over Chimney 2 at 100mx100m, 10m resolution, 20m altitude (1 waffle)
- 2x lawnmowers over Chimney 2 at 100mx100m, 10m resolution, 40m altitude (1 waffle)

Hrs 2-14:

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

- Medium-plume tracking
 - V&G provide a set of chained lawnmowers at medium to low resolution potentially at multiple heights (to be determined) that attempt to closely track the plume maxima over time
 - Potential for online shifts
- Hrs 15-17:
 - 2x lawnmowers over Chimney 2 at 100mx100m, 10m resolution, 20m altitude (1 waffles)
 - 2x lawnmowers over Chimney 2 at 100mx100m, 10m resolution, 40m altitude (1 waffle)

JASON at All Sites

Dive 1: Wednesday, Nov 17 (and on first dive at any new site)

Allows us to seed models with good prior on initial conditions

Start of Dive (Hours 0-1): Deploy the 2 tiltmeters at coordinates planned with operations

Initial suggestion: TM A: (27.40246, -111.3835) [plane in front of ridge][deploy after descent], TM B: (27.40757, -111.390) [any location along ridge][deploy during ops]

Hrs 1-2: 10-20 minutes of temperature probe directly (or as close as is safe) in the largest orifice of Chimney #2 ; Align camera and ensure that green lasers are on and visible at the orifice or in the plume (where the temperature probe was deployed). Use lasers to measure orifice diameter. Then, take 10-20 minutes of MISO camera data focused on Chimney #2 (27.41265607, -111.38690875, ~1845m depth)

Hrs 2-8: A ridge survey. For each plume below (or as many as possible), repeat the above process to gather ~10 minutes of video and in plume data.

End of dive (final hour): Recover TM A at coordinates (27.40246, -111.3835)

Chimney Coordinates

Reference	Label	Lat	Lon	Est Depth
NA090-20171022	tall_peak_smoker	27.41030475	-111.38840791	1840
NA090-20171022	black_smoker_chimney	27.41038474	-111.38847751	1840
NA090-20171022	slender_black_smoker	27.40926223	-111.38931794	1840
NA090-20171022	huge_sulfide_wall_chimney	27.40916004	-111.38932667	1840
NA090-20171022	large_diffuse_flow_chimney	27.408978	-111.389397	1850
NA090-20171022	NA090_021, Chimney #1	27.407473	-111.389862	1850
NA090-20171022	NA090_111, Chimney #2	27.41265607	-111.38690875	1845
Herc Track	Southwest Center	27.407739247	-111.390099146	NA
Herc Track	Southwest 2 Center	27.4090877667	-111.389352359	NA
Herc Track	Chimney Field	27.4105224206	-111.388347357	NA
Herc Track	Northeast 2 Center	27.4111370942	-111.388086109	NA
Herc Track	Northeast Center	27.4126149693	-111.387381041	NA

Dive 2-Final:

Allows us to have a continuously updated current model

Start of dive (Hours 0-1): Deploy the new tiltmeter

End of dive (final hours): Recover the second deployed tiltmeter

When possible:

- Station-keeping in the plume waters – *using camera feed and on-board sensors, attempt to keep JASON near the center of plume expression; this allows us to collect a time-series that may be of interest for model validation*
 - Ideally done during slack current (coordinate timing with V & G)
 - Various altitudes between as low as possible (~5m) and 30m
 - As long as possible [ideally 30 minutes minimum]
 - One of V & G should be present during this maneuver
- Short plume profiles – *ascend to ~200m altitude from the plume source at constant location every 1-3 hours; this allows us to collect a profile-time-series that is of interest for model development*
 - Drive ~60m off plume and then ascend to 200m
 - One of V & G should be present during these maneuver

SENTRY Science Dive Plans

Science Dive Description: Sentry 608 – Graben Ridge

Goal: Constrain the non-buoyant plume height (and bending angle) of advecting plumes in the basin and capture the “current sweep” of plume waters within the basin.

Motivation: Estimating current from observational data alone is an interesting challenge and would add to the robustness of the modeling approach proposed (e.g., if we can do it without tiltmeters or show that we could estimate as well as tiltmeters, that would be of algorithmic and modeling interest). This dive will also allow us to better constrain unknown plume initial conditions by attempting to hit the plume at the non-buoyant layer, which should better reveal aspects of plume buoyancy/momentun and help to contextualize buoyant stem measurements.

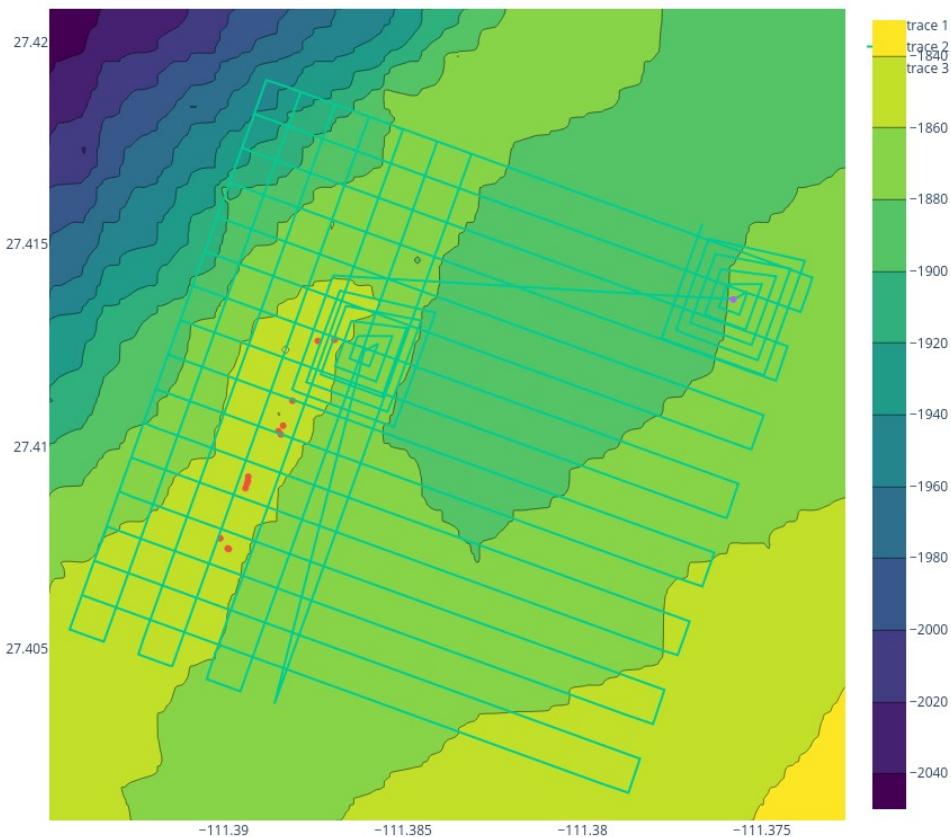
Description:

- **Phase 1: Lawnmower over plain**
 - Perform a 1500mx1600m lawnmower starting to the southwest of the ridge at 100m, oriented at the same angle as the ridge/graben at 120m (to be potentially adjusted while in flight)
 - Starting position: 27.40625, -111.394
 - Approximately 10hrs
 - **Note:** Lawnmower designed to cover the entire ridge and plain up to the “Knob” at 27.41364, -111.3758. Some of the graben is also covered. This lawnmower is to be used to capture “current sweep” and non-buoyant layer edges
 - **NISKIN FIRING:** For capturing out of plume water (or lightly plume influenced water) this is likely the time of opportunity for firing the Niskin(s) on Sentry
- **Phase 2: Lawnmower over ridge**
 - Perform a 600mx1600m lawnmower at 100m resolution with long-legs oriented along-ridge at 45m altitude
 - Starting position: 27.41898, -111.3888
 - Approximately 4.5hrs
 - **Note:** Designed to capture contributions from multiple sources along the ridge and capture behind-ridge advection of plume water
 - **NISKIN FIRING:** For highly influenced plume water, this is likely the time of opportunity for firing the Niskin(s) on Sentry
- **Phase 3: Spiral over Chimney 3**
 - Perform a 200mx200m spiral at 100m altitude at 30m resolution just to the southeast of Chimney 2 site.
 - Approximately 1.5hrs
 - **Note:** Designed to capture nonbuoyant advection from a site with a lot of previous Sentry data
- **Phase 4: Spiral over Knob**
 - Transit from Chimney 2 to the Knob and perform a 200mx200m spiral at 30m altitude at 30m resolution centered with the feature

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

- Approximately 1.5hrs
- **Note:** Included for scientific interest over whether there is interesting hydrothermal activity/pluming at the site
- **Holding Pattern:**
 - A 200mx200m 10m resolution holding pattern at 65m is included at the end of the dive, with potential for opportunistic placement
- **Opportunistic Holding Patterns:**
 - There is some interest in occasionally halting Sentry into a ballast test and adjusting vertical profiles to capture short vertical plume structures as possible. Likely to be done during the holding pattern if time, or opportunistically during the spirals.



Science Dive Description: Sentry 609 – Adaptive Graben Ridge

Note: Mission never executed b/c technical issues

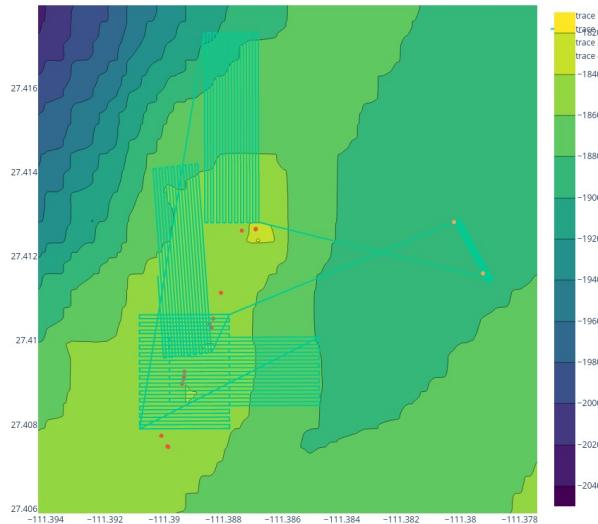
Goal: Using previously collected scientific data (background profiles, tiltmeter data, source measurements), forecast the envelope of the plume and design adaptive lawnmowers that exploit these models to spend more time high in the buoyant plume. “Naively designed” lawnmowers are included as benchmarks.

Motivation: Our measurement informed MTT forward model enables us to forecast plume evolution in space and time. This mission will demonstrate that we can use this model to design adaptive lawnmowers that spend more time in buoyant plume waters, compared to naively designed benchmark lawnmowers. We will additionally use these models to collect bottle samples of buoyant plume water of different ages/distances from the plume source.

Description:

- **Phase 1: Adaptive Lawnmower over Chimney #1**
 - Perform a 180mx500mx10m lawnmower oriented at 0 degrees and offset slightly to the east of Chimney 1, designed to intersect plume water advected by currents at 0 degrees (standard convention).
 - Starting position: 27.40926223, -111.38931794 at 120m
 - Approximately 3.58hrs
 - **Note:** This lawnmower is designed to map plume water current from west to east.
- **Phase 2: Naive Lawnmower over Chimney #1**
 - Perform a 300mx300mx10m “naive” lawnmower oriented at 0 degrees and centered over Chimney 1.
 - Starting position: 27.40926223, -111.38931794 at 120m
 - Approximately 3.55hrs
 - **Note:** This lawnmower is naively designed to map plume water over Chimney #1
- **Phase 3: Video Survey**
 - Transect over to the knobletts and perform a 20x200x5m lawnmower for an exploratory video survey
 - Starting position: 27.41282, -11.38022 at 5m
 - Approximately 0.56hrs
 - **Note:** Exploratory photo mosaic of interesting features in bathy using MISO camera
 - **Note:** Included for scientific interest over whether there is interesting hydrothermal activity/pluming at the site
- **Phase 4: Adaptive Lawnmower over Chimney #2**
 - Perform a 500mx180mx10m lawnmower oriented at 90 degrees and offset slightly to the north of Chimney 2, designed to intersect plume water advected by currents at 90 degrees (standard convention).
 - Starting position: 27.41236, -111.3861 at 200m
 - Approximately 3.58hrs
 - **Note:** This lawnmower is designed to map plume water with currents from south to north.

- **NISKIN FIRING:** For capturing out of plume water (or lightly plume influenced water) firing a Niskin at 27.41685, -111.388
- **Phase 5: Naive Lawnmower over Chimney #1 (second time)**
 - Repeat an identical 300mx300mx10m “naive” lawnmower oriented at 0 degrees and centered over Chimney 1.
 - Starting position: 27.40926223, -111.38931794 at 120m
 - Approximately 3.55hrs
 - **Note:** This lawnmower is naively designed to map plume water over Chimney #1.
- **Phase 6: Adaptive Lawnmower over Chimney #1 (second time)**
 - Perform a 500mx180mx10m lawnmower oriented at 96 degrees and offset slightly to the north of Chimney 1, designed to intersect plume water advected by currents at 95 degrees (standard convention).
 - Starting position: 27.40926223, -111.38931794 at 120m
 - Approximately 3.58hrs
 - **Note:** This lawnmower is designed to map plume water current from south to north.
 - **NISKIN FIRING:** For capturing out of plume water (or lightly plume influenced water) firing a Niskin at 27.41391, -11.3901
- **Holding Pattern:**
 - A 200mx200m 10m resolution holding pattern at 65m is included at the end of the dive, with potential for opportunistic placement
- **Opportunistic Holding Patterns:**
 - There is some interest in occasionally halting Sentry into a ballast test and adjusting vertical profiles to capture short vertical plume structures as possible. Likely to be done during the holding pattern if time, or opportunistically.



Science Dive Description: Sentry 610 – Adaptive Graben/Ridge, Chimney 2

Goal: Using previously collected scientific data (background profiles, tiltmeter data, source measurements), forecast the envelope of the plume and design adaptive lawnmowers that exploit these models to spend more time high in the buoyant plume. “Naively designed” lawnmowers are included as benchmarks.

Motivation: Our measurement informed MTT forward model enables us to forecast plume evolution in space and time. This mission will demonstrate that we can use this model to design adaptive lawnmowers that spend more time in buoyant plume waters, compared to naively designed benchmark lawnmowers. We will additionally use these models to collect bottle samples of buoyant plume water of different ages/distances from the plume source.

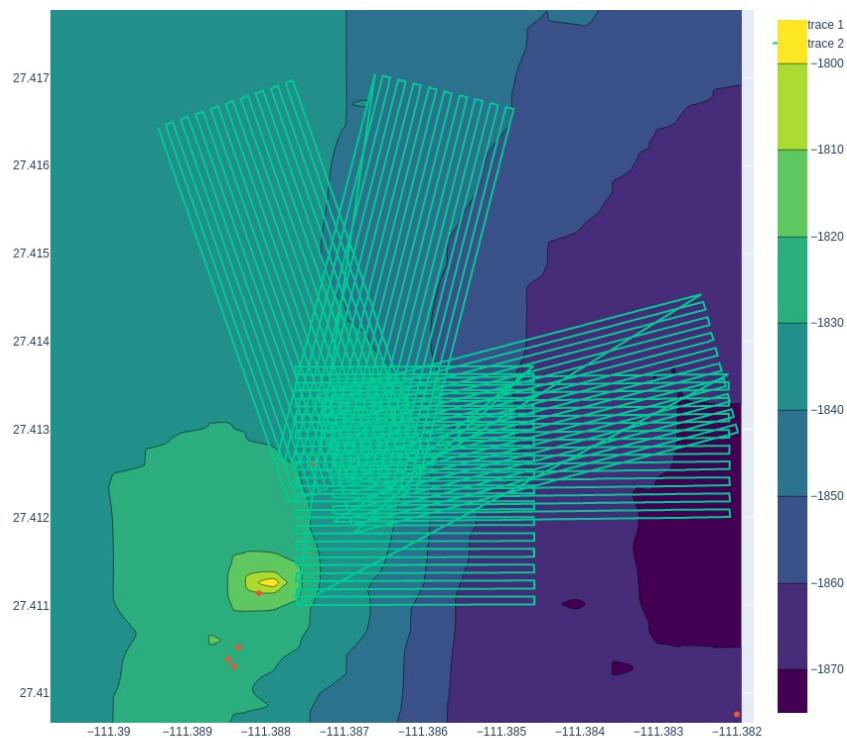
Description:

- **Phase 1: Adaptive Lawnmower over Chimney #2**
 - Perform a 180mx500mx10m lawnmower oriented at 75.5288 degrees and offset slightly to the east of Chimney 2, designed to intersect plume water advected by currents.
 - Starting position: 1379.1730m easting, 1354.6755m northing from (27.4, -111.4) at 120m
 - Approximately 3.0hrs
 - **NISKIN FIRING:** A Niskin will be opportunistically fired in order to collect distal plume water
- **Phase 2: Naive Lawnmower over Chimney #2**
 - Perform a 300mx300mx10m “naive” lawnmower placed above Chimney 2
 - Starting position: 150m more west and 150m more south than Chimney 2 locations at 120m
 - Approximately 3.00hrs
 - **Note:** This lawnmower is naively designed to map plume water over Chimney #2
- **Phase 3: Adaptive Lawnmower over Chimney #2**
 - Perform a 180mx500mx10m lawnmower oriented at 14.5283 degrees and offset slightly to the north of Chimney 2, designed to intersect plume water advected by currents.
 - Starting position: 1301.266m E, 1305.1493m N from (27.4, -111.4) at 120m
 - Approximately 3hrs
- **Phase 4: Adaptive Lawnmower over Chimney #2**
 - Perform a 180mx500mx10m lawnmower oriented at 0.5 degrees and offset slightly to the north of Chimney 2, designed to intersect plume water advected by currents.
 - Starting position: 1275.264m E, 1319.538m N from (27.4, -111.4) at 120m
 - Approximately 3hrs
- **Phase 5: Naive Lawnmower over Chimney #2 (second time)**
 - Repeat an identical 300mx300mx10m “naive” lawnmower centered over Chimney 2.
 - Starting position: 150m more west and 150m more south than Chimney 2 locations at 120m
 - Approximately 3.00hrs
 - **Note:** This lawnmower is naively designed to map plume water over Chimney #2
 - **Note:** This was supposed to be flown at 200m, but bottom tracking was difficult, so the track was moved to 120.0m altitude.
 - **Note:** Due to the height of flights, this trajectory was started much later than assumed. To get to the next lawnmower, several waypoints of this trajectory were canceled.

RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flasphohler

- **Phase 6: Adaptive Lawnmower over Chimney #2**
 - Perform a 500mx180mx10m lawnmower oriented at 109.073 degrees and offset slightly to the north of Chimney 2, designed to intersect plume water advected by currents
 - Starting position: 1389.952m E, 1403.596m N of (27.4, -111.4)
 - Approximately 3hrs
 - **NISKIN FIRING:** A Niskin will be opportunistically fired in order to collect distal plume water
 - **Note:** Only a few tracklines were able to executed in this trajectory due to timing.
- **Holding Pattern:**
 - The last trajectory will be shifted to be used as a holding pattern when necessary.
- **Opportunistic Holding Patterns:**
 - There is some interest in occasionally halting Sentry into a ballast test and adjusting vertical profiles to capture short vertical plume structures as possible. Likely to be done during the holding pattern if time, or opportunistically.



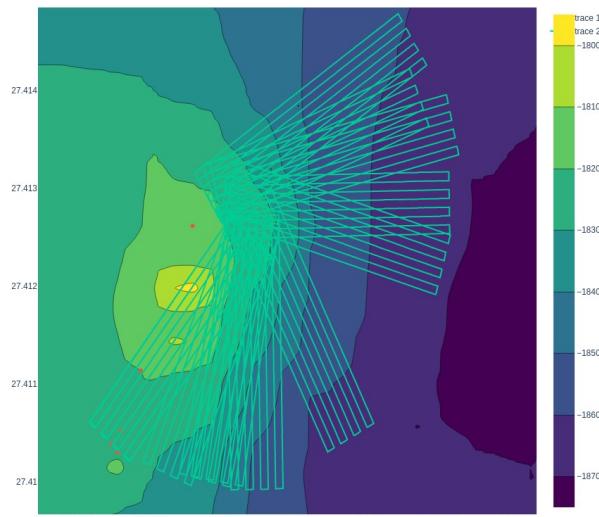
Science Dive Description: Sentry 611 – Data-Informed Adaptive Graben Ridge at Chimney 2

Goal: Using previously collected scientific data (background profiles, tiltmeter data, source measurements), as well as Sentry observations, forecast the envelope of the plume and design adaptive lawnmowers that exploit these models to catch high resolution snapshots of the plume bend.

Motivation: Our measurement informed MTT forward model enables us to forecast plume evolution in space and time. We have updated the *prior* version of this model with *in situ* observations from previous Sentry flights. This mission will demonstrate that we can use this model to design adaptive lawnmowers that spend more time in buoyant plume waters, compared to naively designed benchmark lawnmowers, and that these models can be used within an iterative data-informed regime. We will additionally use these models to collect bottle samples of buoyant plume water of different ages/distances from the plume source.

Description:

- **Radially fanned lawnmowers that track the current**
 - 7-8 1hr long lawnmowers of various lengths will be used to gather near-plume measurements near the Chimney 2 source.
 - **NISKIN FIRING:** will be opportunistically fired
 - **Note:** This trajectory will be done in *depth-hold mode* ranging from 60-140m above the bottom; this is to attempt to hit the plume's nonbuoyant "underlayer" over time.
- **Holding Pattern:**
 - The last trajectory will be shifted to be used as a holding pattern when necessary.
- **Opportunistic Holding Patterns:**
 - There is some interest in occasionally halting Sentry into a ballast test and adjusting vertical profiles to capture short vertical plume structures as possible. Likely to be done during the holding pattern if time, or opportunistically.



RR2107

Midwater Hydrothermal Plume Surveying, Modeling, and Tracking
Prepared by: Victoria Preston, Genevieve Flaspoehler

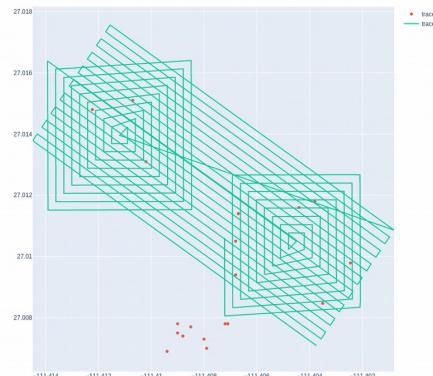
Science Dive Description: Sentry 612 – Exploratory Survey of S. Guaymas

Goal: From markers placed by previous science parties at S. Guaymas, this survey is intended to provide a snapshot of multiple-plume expressions possible within the region, specifically concentrated in two areas.

Motivation: As a well-studied site, collecting mid-water chemistry transects of the region offers a unique perspective, scientifically. Algorithmically, as a site with known, clustered large vents, this is an interesting case for studying probabilistic multiplume models.

Description:

- **Video Transect (designed by Sentry team)**
 - A short transect at 5m to be conducted in a flat region in order to test 4k video capabilities of a MISO camera
- **Regional lawnmower**
 - 530m x 1300m at 30m resolution, oriented 145deg from horizontal centered between two clusters of known markers
 - 40m altitude
 - Approx. 8.9hrs
- **Spiral over northern cluster**
 - Centered at (27.01396, -111.4112)
 - 300m x 300m with 30m resolution
 - 100m altitude
 - Approx. 3.9hrs
 - **NISKIN FIRING:** A Niskin will be fired upon arriving to the center of the spiral trajectory
- **Spiral over southern cluster (Rebecca's Roost, Cathedral Hill, etc.)**
 - Centered at (27.0105, -111.4045)
 - 300m x 300m with 30m resolution
 - 100m altitude
 - Approx. 3.9hrs
 - **NISKIN FIRING:** A Niskin will be fired upon arriving to the center of the spiral trajectory



ACRONYMS

AUV	Autonomous Underwater Vehicle
CTD	Conductivity Temperature Depth probe
FDM	Finite Difference Method
FEM	Finite Element Method
GP	Gaussian process
HOV	Human Occupied Vehicle
MCMC	Merkov Chain Monte Carlo
MISO	Multidisciplinary Instrumentation in Support of Oceanography
MOR	Model Order Reduction
NDSF	National Deep Submergence Facility
OBS	Optical backscatter instrument
ODE	Ordinary Differential Equation
ORP	Oxidation-reduction potential instrument
PDE	Partial Differential Equation
PHORTEX	Physically-informed Operational Robotic Trajectories for Expeditions
PHUMES	Physically-informed Uncertainty Models for Environment Spatiotemporality

PIV	Particle Imaging Velocimetry
POD	Proper Orthogonal Decomposition
POMDP	Partially-observable Markov decision-process
R/V	Research Vessel
ROV	Remotely Operated Vehicle
SML	Scientific Machine Learning
UNOLS	University-National Oceanographic Laboratory System
WHOI	Woods Hole Oceanographic Institution

BIBLIOGRAPHY

1. Adams, I. and Di Iorio, D. Turbulence properties of a deep-sea hydrothermal plume in a time-variable cross-flow: field and model comparisons for Dante in the Main Endeavour Field. *Journal of Geophysical Research: Oceans*, 2020.
2. Agrawal, R. Sample mean based index policies by $O(\log n)$ regret for the multi-armed bandit problem. *Adv. Appl. Probab.* **27**, 1995.
3. Aguilera, P. A., Fernández, A., Fernández, R., Rumí, R., and Salmerón, A. Bayesian networks in environmental modelling. *Environmental Modelling & Software* **26**, 2011.
4. Amari, S. Backpropagation and stochastic gradient descent method. *Neurocomputing* **5**, 1993.
5. Arora, A., Furlong, P. M., Fitch, R., Sukkarieh, S., and Fong, T. *Multi-modal active perception for information gathering in science missions* in *Proc. Int. Symp. Auton. Robots* 2017.
6. Arridge, S., Maass, P., Öktem, O., and Schönlieb, C.-B. Solving inverse problems using data-driven models. *Acta Numerica* **28**, 2019.
7. Auer, P. Using confidence bounds for exploitation-exploration trade-offs. *J. Mach. Learn. Research* **3**, 2002.
8. Austin, T. *et al.* *The Martha's Vineyard Coastal Observatory: A long term facility for monitoring air-sea processes* in *OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings (Cat. No. 00CH37158)* **3**. 2000.
9. Babaee, H., Bastidas, C., DeFilippo, M., Chryssostomidis, C., and Karniadakis, G. A multifidelity framework and uncertainty quantification for sea surface temperature in the Massachusetts and Cape Cod Bays. *Earth and Space Science* **7**, 2020.
10. Baddoo, P. J., Herrmann, B., McKeon, B. J., and Brunton, S. L. Kernel learning for robust dynamic mode decomposition: linear and nonlinear disambiguation optimization. *Proceedings of the Royal Society A* **478**, 2022.
11. Baddoo, P. J., Herrmann, B., McKeon, B. J., Kutz, J. N., and Brunton, S. L. Physics-informed dynamic mode decomposition (piDMD). *arXiv preprint arXiv:2112.04307*, 2021.
12. Baker, E. T., Massoth, G. J., Feely, R. A., Cannon, G. A., and Thomson, R. E. The rise and fall of the CoAxial hydrothermal site, 1993–1996. *Journal of Geophysical Research: Solid Earth* **103**, 1998.

13. Baker, E. T., Massoth, G. J., Walker, S. L., and Embley, R. W. A method for quantitatively estimating diffuse and discrete hydrothermal discharge. *Earth and Planetary Science Letters* **118**, 1993.
14. Baker, E. T. *et al.* Posteruption enhancement of hydrothermal activity: A 33-year, multieruption time series at Axial Seamount (Juan de Fuca Ridge). *Geochemistry, Geophysics, Geosystems* **20**, 2019.
15. Baker, N. *et al.* *Workshop report on basic research needs for scientific machine learning: Core technologies for artificial intelligence* technical report (USDOE Office of Science (SC), Washington, DC (United States), 2019).
16. Ballard, R. D. The MEDEA/JASON remotely operated vehicle system. *Deep Sea Research Part I: Oceanographic Research Papers* **40**, 1993.
17. Bangian-Tabrizi, A., Zhang, X., and Jaluria, Y. Solution to inverse natural convection problem using experimental data. *International Journal of Heat and Mass Transfer* **189**, 2022.
18. Barreyre, T., Soule, S. A., and Sohn, R. A. *Dispersal of volcaniclasts during deep-sea eruptions: Settling velocities and entrainment in buoyant seawater plumes* 2011.
19. Barreyre, T. *et al.* Structure, temporal evolution, and heat flux estimates from the Lucky Strike deep-sea hydrothermal field derived from seafloor image mosaics. *Geochemistry, Geophysics, Geosystems* **13**, 2012.
20. Barth, J. A. *et al.* Warm blobs, low-oxygen events, and an eclipse: The Ocean Observatories Initiative Endurance Array captures them all. *Oceanography* **31**, 2018.
21. Bayes, T. An essay towards solving a problem in the doctrine of chances. *Philosophical transactions of the Royal Society of London*, 1763.
22. Beaulieu, S. E., Baker, E. T., and German, C. R. Where are the undiscovered hydrothermal vents on oceanic spreading ridges? *Deep Sea Research Part II: Topical Studies in Oceanography* **121**, 2015.
23. Beaulieu, S. E., Baker, E. T., German, C. R., and Maffei, A. An authoritative global database for active submarine hydrothermal vent fields. *Geochemistry, Geophysics, Geosystems* **14**, 2013.
24. Belkin, I., de Sousa, J. B., Pinto, J., Mendes, R., and López-Castejón, F. *A new front-tracking algorithm for marine robots in 2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV)* 2018.
25. Bell, J. B., Woulds, C., and Oevelen, D. v. Hydrothermal activity, functional diversity and chemoautotrophy are major drivers of seafloor carbon cycling. *Scientific reports* **7**, 2017.
26. Bell, K. L. C. *Envisioning the Future of Ocean Exploration* (eds Subcommittee on Environment Committee of Science, S. and Technology, U. H. o. R.) 2019.

27. Bellman, R. A Markovian decision process. *J. Math. Mechanics*, 1957.
28. Bemis, K. G., Von Herzen, R. P., and Mottl, M. J. Geothermal heat flux from hydrothermal plumes on the Juan de Fuca Ridge. *Journal of Geophysical Research: Solid Earth* **98**, 1993.
29. Bennett, S. A. *et al.* Trophic regions of a hydrothermal plume dispersing away from an ultramafic-hosted vent-system: Von Damm vent-site, Mid-Cayman Rise. *Geochemistry, Geophysics, Geosystems* **14**, 2013.
30. Benning, M. and Burger, M. Modern regularization methods for inverse problems. *Acta Numerica* **27**, 2018.
31. Berg, J. and Nyström, K. Data-driven discovery of PDEs in complex datasets. *Journal of Computational Physics* **384**, 2019.
32. Bergstra, J.S., Bardenet, R., Bengio, Y., and Kégl, B. *Algorithms for hyperparameter optimization* in *Proc. Adv. Neural Inf. Process. Syst.* 2011.
33. Bigoni, D., Zahm, O., Spantini, A., and Marzouk, Y. Greedy inference with layers of lazy maps. *arXiv preprint arXiv:1906.00031*, 2019.
34. Binney, J. and Sukhatme, G. S. *Branch and bound for informative path planning* in *Proc. IEEE Int. Conf. Robot. Autom.* 2012.
35. Bishop, C. M. *Pattern recognition and machine learning* chap. 10. Springer, 2006.
36. Bishop, J. K. Transmissometer measurement of POC. *Deep Sea Research Part I: Oceanographic Research Papers* **46**, 1999.
37. Biswas, B., Chatterjee, S., Mukherjee, S., and Pal, S. A discussion on Euler method: A review. *Electronic Journal of Mathematical Analysis and Applications* **1**, 2013.
38. Bittig, H. C. *et al.* Oxygen optode sensors: principle, characterization, calibration, and application in the ocean. *Frontiers in Marine Science* **4**, 2018.
39. Blanchard, A. and Sapsis, T. P. Analytical description of optimally time-dependent modes for reduced-order modeling of transient instabilities. *SIAM Journal on Applied Dynamical Systems* **18**, 2019.
40. Blanchard, A. and Sapsis, T. P. Learning the tangent space of dynamical instabilities from data. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **29**, 2019.
41. Bottou, L. In *Proceedings of COMPSTAT 2010* Springer, 2010.
42. Bourgault, F., Makarenko, A. A., Williams, S. B., Grocholsky, B., and Durrant-Whyte, H. F. *Information based adaptive robotic exploration* in *IEEE/RSJ Int. Conf. Intell. Robot. Syst.* **1**. 2002.
43. Branch, A. *et al.* *Demonstration of autonomous nested search for local maxima using an unmanned underwater vehicle* in *2020 IEEE International Conference on Robotics and Automation (ICRA)* 2020.

44. Bray, N. Water mass formation in the Gulf of California. *Journal of Geophysical Research: Oceans* **93**, 1988.
45. Brier, G. W. *et al.* Verification of forecasts expressed in terms of probability. *Monthly weather review* **78**, 1950.
46. Browne, C. B. *et al.* A survey of Monte Carlo tree search methods. *IEEE Trans. Comput. Intell. AI Games* **4**, 2012.
47. Bruce, A. L., Zeidan, V. M., and Bernstein, D. S. *What is the Koopman operator? a simplified treatment for discrete-time systems* in *2019 American Control Conference (ACC)* 2019.
48. Brunton, S. L., Proctor, J. L., and Kutz, J. N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences* **113**, 2016.
49. Burgard, W., Fox, D., and Thrun, S. *Active mobile robot localization by entropy minimization* in *Proc. Euromicro Workshop Advance. Mob. Rob.* 1997.
50. Burkardt, J., Gunzburger, M., and Lee, H.-C. POD and CVT-based reduced-order modeling of Navier–Stokes flows. *Computer methods in applied mechanics and engineering* **196**, 2006.
51. Cahill, N., Kemp, A. C., Horton, B. P., and Parnell, A. C. Modeling sea-level change using errors-in-variables integrated Gaussian processes. *The Annals of Applied Statistics* **9**, 2015.
52. Cai, W.-J. Estuarine and coastal ocean carbon paradox: CO₂ sinks or sites of terrestrial carbon incineration? *Annu. Rev. Mar. Sci.* **3**, 2011.
53. Calli, B. *et al.* Yale-CMU-Berkeley dataset for robotic manipulation research. *The International Journal of Robotics Research* **36**, 2017.
54. Camilli, R. *et al.* Tracking hydrocarbon plume transport and biodegradation at Deepwater Horizon. *Science* **330**, 2010.
55. Campbell, A. C., Gieskes, J. M., Lupton, J. E., and Lonsdale, P. F. Manganese geochemistry in the Guaymas Basin, Gulf of California. *Geochimica et Cosmochimica Acta* **52**, 1988.
56. Caratori Tontini, F. *et al.* Crustal magnetization of Brothers volcano, New Zealand, measured by autonomous underwater vehicles: Geophysical expression of a submarine hydrothermal system. *Economic Geology* **107**, 2012.
57. Cario, A., Oliver, G. C., and Rogers, K. L. Exploring the deep marine biosphere: challenges, innovations, and opportunities. *Frontiers in Earth Science* **7**, 2019.
58. Carrillo, H., Dames, P., Kumar, V., and Castellanos, J. A. *Autonomous robotic exploration using occupancy grid maps and graph SLAM based on Shannon and Rényi entropy* in *Proc. IEEE Int. Conf. Robot. Autom.* 2015.

59. Carrillo, H., Latif, Y., Rodriguez-Arevalo, M. L., Neira, J., and Castellanos, J. A. *On the monotonicity of optimality criteria during exploration in active SLAM* in *IEEE Int. Conf. Robot. Autom.* 2015.
60. Chaudhuri, K. N. and Chaudhuri, K. *Trade and civilisation in the Indian Ocean: an economic history from the rise of Islam to 1750* Cambridge University Press, 1985.
61. Chee, K. Y., Jiahao, T. Z., and Hsieh, M. A. KNODE-MPC: A knowledge-based data-driven predictive control framework for aerial robots. *IEEE Robotics and Automation Letters* **7**, 2022.
62. Chen, K. K., Tu, J. H., and Rowley, C. W. Variants of dynamic mode decomposition: boundary condition, Koopman, and Fourier analyses. *Journal of nonlinear science* **22**, 2012.
63. Chen, R. T. Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. *Neural ordinary differential equations* presentation. 2019.
64. Chen, T. Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K. *Neural ordinary differential equations* in *Advances in neural information processing systems* 2018.
65. Chen, W., Khardon, R., and Liu, L. AK: Attentive kernel for information gathering. *arXiv preprint arXiv:2205.06426*, 2022.
66. Chen, X.-x. and Huang, J. Odor source localization algorithms on mobile robots: A review and future outlook. *Robotics and Autonomous Systems* **112**, 2019.
67. Chevaldonné, P., Desbruyères, D., and Le Haître, M. Time-series of temperature from three deep-sea hydrothermal vent sites. *Deep Sea Research Part A. Oceanographic Research Papers* **38**, 1991.
68. Childs, J. Extraction in four dimensions: time, space and the emerging geopolitics of deep-sea mining. *Geopolitics* **25**, 2020.
69. Chin, C. S. *et al.* In situ observations of dissolved iron and manganese in hydrothermal vent plumes, Juan de Fuca Ridge. *Journal of Geophysical Research: Solid Earth* **99**, 1994.
70. Clague, D. *et al.* Abundance and distribution of hydrothermal chimneys and mounds on the Endeavour Ridge determined by 1-m resolution AUV multibeam mapping surveys in *AGU Fall Meeting Abstracts* **2008**. 2008.
71. Cohn, D. A. Neural network exploration using optimal experiment design in *Proc. Adv. Neural Inf. Process. Syst.* 1994.
72. Cole, J. J. *et al.* Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems* **10**, 2007.
73. Conn, A. R., Gould, N. I., and Toint, P. L. *Trust region methods* SIAM, 2000.
74. Corliss, J. B. *et al.* Submarine thermal springs on the Galapagos Rift. *Science* **203**, 1979.

75. Cowen, J. P., Wen, X., and Popp, B. N. Methane in aging hydrothermal plumes. *Geochimica et Cosmochimica Acta* **66**, 2002.
76. Cui, T., Martin, J., Marzouk, Y. M., Solonen, A., and Spantini, A. Likelihood-informed dimension reduction for nonlinear inverse problems. *Inverse Problems* **30**, 2014.
77. Cushman-Roisin, B., Heil, W. H., and Nof, D. Oscillations and rotations of elliptical warm-core rings. *Journal of Geophysical Research: Oceans* **90**, 1985.
78. Daniel, C., Neumann, G., and Peters, J. *Hierarchical relative entropy policy search* in *Artif. Intell. Stat.* 2012.
79. Deng, H. *et al.* Invariant subspace learning for time series data based on dynamic time warping distance. *Pattern Recognition* **102**, 2020.
80. Desaraju, V. R. *et al.* Vision-based landing site evaluation and informed optimal trajectory generation toward autonomous rooftop landing. *Autonomous Robots* **39**, 2015.
81. Dick, G. J. *et al.* The microbiology of deep-sea hydrothermal vent plumes: ecological and biogeographic linkages to seafloor and water column habitats. *Frontiers in Microbiology* **4**, 2013.
82. Duane, S., Kennedy, A. D., Pendleton, B. J., and Roweth, D. Hybrid Monte Carlo. *Physics letters B* **195**, 1987.
83. Dupont, E., Doucet, A., and Teh, Y. W. *Augmented neural odes* in *Advances in Neural Information Processing Systems* 2019.
84. Durbin, P. A. Some recent developments in turbulence closure modeling. *Annual Review of Fluid Mechanics* **50**, 2018.
85. Dutta, S. *et al.* Data-driven reduced order modeling of environmental hydrodynamics using deep autoencoders and neural ODEs. *arXiv preprint arXiv:2107.02784*, 2021.
86. Edwards, P. N. *Representing the global atmosphere: Computer models, data, and knowledge about climate change* chap. Changing the atmosphere: Expert knowledge and environmental governance. MIT Press Cambridge, MA, 2001.
87. Edwards, S., Rutkowski, A. J., Quinn, R. D., and Willis, M. A. *Moth-inspired plume tracking strategies in three-dimensions* in *Proceedings of the 2005 IEEE international conference on robotics and automation* 2005.
88. Engl, H. W., Hanke, M., and Neubauer, A. *Regularization of inverse problems* Springer Science & Business Media, 1996.
89. Erlandson, J. M. and Fitzpatrick, S. M. Oceans, islands, and coasts: Current perspectives on the role of the sea in human prehistory. *Journal of Island & Coastal Archaeology* **1**, 2006.

90. Esposito, J. M. and Kumar, V. *A method for modifying closed-loop motion plans to satisfy unpredictable dynamic constraints at runtime* in *Proc. IEEE Int. Conf. Robot. Autom.* **2**, 2002.
91. Euler, L. Principes généraux du mouvement des fluides. *Mémoires de l'académie des sciences de Berlin*, 1757.
92. Everett, M., Miller, J., and How, J. P. *Planning beyond the sensing horizon using a learned context* in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* 2019.
93. Farr, N., Bowen, A., Ware, J., Pontbriand, C., and Tivey, M. *An integrated, underwater optical/acoustic communications system* in *OCEANS'10 IEEE SYDNEY* 2010.
94. Fedorov, V. V. *Theory of optimal experiments* Elsevier, 2013.
95. Feely, R. A., Massoth, G. J., Baker, E. T., Lebon, G. T., and Geiselman, T. L. Tracking the dispersal of hydrothermal plumes from the Juan de Fuca Ridge using suspended matter compositions. *Journal of Geophysical Research: Solid Earth* **97**, 1992.
96. Ferguson, T. S. A Bayesian analysis of some nonparametric problems. *The annals of statistics*, 1973.
97. Ferri, G., Jakuba, M. V., and Yoerger, D. R. A novel trigger-based method for hydrothermal vents prospecting using an autonomous underwater robot. *Autonomous Robots* **29**, 2010.
98. Figueiredo, M. A. T. and Jain, A. K. Unsupervised learning of finite mixture models. *IEEE Transactions on pattern analysis and machine intelligence* **24**, 2002.
99. Firth, S. and Naidu, V. *Understanding Oceania: celebrating the University of the South Pacific and its collaboration with the Australian National University* ANU Press, 2019.
100. Fischer, T. P. *et al.* Temporal variations in fumarole gas chemistry at Poás volcano, Costa Rica. *Journal of Volcanology and Geothermal Research* **294**, 2015.
101. Fischer, T. P. and de Moor, M. J. Explosions at Poás, Costa Rica: a case for long-term, high-rate geochemical volcano monitoring to detect precursory signals. *AGUFM* **2019**, 2019.
102. Flaspohler, G., Preston, V., Michel, A. P., Girdhar, Y., and Roy, N. Information-guided robotic maximum seek-and-sample in partially observable continuous environments. *IEEE Robotics and Automation Letters*, 2019.
103. Flaspohler, G., Roy, N. A., and Fisher III, J. W. Belief-dependent macro-action discovery in POMDPs using the value of information. *Advances in Neural Information Processing Systems* **33**, 2020.
104. Fleeger, J. W. *et al.* The response of nematodes to deep-sea CO₂ sequestration: A quantile regression approach. *Deep Sea Research Part I: Oceanographic Research Papers* **57**, 2010.

105. Follmann, R. and Rosa Jr, E. Predicting slow and fast neuronal dynamics with machine learning. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **29**, 2019.
106. Fong, D. A. and Jones, N. L. Evaluation of AUV-based ADCP measurements. *Limnology and Oceanography: methods* **4**, 2006.
107. Forsythe, G. E. and Henrici, P. The cyclic Jacobi method for computing the principal values of a complex matrix. *Transactions of the American Mathematical Society* **94**, 1960.
108. Frigg, R. In what sense is the Kolmogorov-Sinai entropy a measure for chaotic behaviour? Bridging the gap between dynamical systems theory and communication theory. *British Journal for the Philosophy of Science*, 2004.
109. Galle, B. *et al.* A multi-purpose, multi-rotor drone system for long-range and high-altitude volcanic gas plume measurements. *Atmospheric Measurement Techniques* **14**, 2021.
110. Gan, L. K., Zhang, P., Lee, J., Osborne, M. A., and Howey, D. A. Data-driven energy management system with Gaussian process forecasting and MPC for interconnected microgrids. *IEEE Transactions on Sustainable Energy* **12**, 2020.
111. Gardner, J., Pleiss, G., Weinberger, K. Q., Bindel, D., and Wilson, A. G. Gpytorch: Blackbox matrix-matrix gaussian process inference with GPU acceleration. *Advances in neural information processing systems* **31**, 2018.
112. Garg, S., Singh, A., and Ramos, F. *Learning non-stationary space-time models for environmental monitoring* in *Twenty-Sixth AAAI Conference on Artificial Intelligence* 2012.
113. Garivier, A. and Cappé, O. *The KL-UCB algorithm for bounded stochastic bandits and beyond* in *Proc. 24th Annual Conf. Learn. Theory* 2011.
114. Geedicke, I., Oldeland, J., and Leishman, M. R. Urban stormwater run-off promotes compression of saltmarshes by freshwater plants and mangrove forests. *Sci. Total Environ.* **637**, 2018.
115. Geiger, A., Lenz, P., and Urtasun, R. *Are we ready for autonomous driving? The KITTI vision benchmark suite* in *2012 IEEE conference on computer vision and pattern recognition* 2012.
116. Geilert, S. *et al.* On the formation of hydrothermal vents and cold seeps in the Guaymas Basin, Gulf of California. *Biogeosciences* **15**, 2018.
117. Georges, D. *The use of observability and controllability gramians or functions for optimal sensor and actuator location in finite-dimensional systems* in *Proceedings of 1995 34th ieee conference on decision and control* **4**. 1995.
118. Georgieva, M. N. *et al.* The history of life at hydrothermal vents. *Earth-Science Reviews* **217**, 2021.

119. German, C. and Boetius, A. Robotics-based scientific investigations at an ice-ocean interface: first results from Nereid Under Ice in the Arctic. *Ocean Worlds* **4** *2168*, 2019.
120. Geyer, W. R. and MacCready, P. *The estuarine circulation* 2014.
121. Ghahramani, Z. In *Hidden Markov models: applications in computer vision* World Scientific, 2001.
122. Glynn, P. W. and Iglehart, D. L. Importance sampling for stochastic simulations. *Management science* **35**, 1989.
123. Grassle, J. F. In *Advances in Marine Biology* Elsevier, 1987.
124. Grasso, F. W., Consi, T. R., Mountain, D. C., and Atema, J. Biomimetic robot lobster performs chemo-orientation in turbulence using a pair of spatially separated sensors: Progress and challenges. *Robotics and Autonomous Systems* **30**, 2000.
125. Green, A., Singhal, R., and Venkateswar, R. Analytic extensions of the Gaussian plume model. *Journal of the Air Pollution Control Association* **30**, 1980.
126. Green, P. J. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**, 1995.
127. Griffiths, T., Jordan, M., Tenenbaum, J., and Blei, D. Hierarchical topic models and the nested Chinese restaurant process. *Advances in neural information processing systems* **16**, 2003.
128. Guestrin, C., Krause, A., and Singh, A. P. *Near-optimal sensor placements in Gaussian processes* in *Proc. 22nd Int. Conf. Mach. Learn.* 2005.
129. Hahn, G. A modified Euler method for dynamic analyses. *International Journal for Numerical Methods in Engineering* **32**, 1991.
130. Harb, C. C. *et al.* Pulsed quantum cascade laser-based CRDS substance detection: real-time detection of TNT. *Opt. Express* **20**. DOI: [10.1364/OE.20.015489](https://doi.org/10.1364/OE.20.015489), 2012.
131. Heaney, K. D., Gawarkiewicz, G., Duda, T. F., and Lermusiaux, P. F. Nonlinear optimization of autonomous undersea vehicle sampling strategies for oceanographic data-assimilation. *Journal of Field Robotics* **24**, 2007.
132. Hennig, P. and Schuler, C. J. Entropy search for information-efficient global optimization. *J. Mach. Learn. Res.* **13**, 2012.
133. Hernández-Lobato, J. M., Hoffman, M. W., and Ghahramani, Z. *Predictive entropy search for efficient global optimization of black-box functions* in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.* 2014.
134. Hernández, P. *et al.* Carbon dioxide degassing by advective flow from Usu volcano, Japan. *Science* **292**, 2001.
135. Hitz, G., Galceran, E., Garneau, M.-È., Pomerleau, F., and Siegwart, R. Adaptive continuous-space informative path planning for online environmental monitoring. *J. Field Robot.* **34**, 2017.

136. Hoffman, M. D., Blei, D. M., Wang, C., and Paisley, J. Stochastic variational inference. *J. Mach. Learn. Res.* **14**, 2013.
137. Hollinger, G. A. and Sukhatme, G. S. *Sampling-based motion planning for robotic information gathering* in *Proc. Robot. Sci. Syst.* **3**. 2013.
138. Holmes, R. M., Aminot, A., Kérouel, R., Hooker, B. A., and Peterson, B. J. A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Canadian Journal of Fisheries and Aquatic Sciences* **56**, 1999.
139. Horel, T. *Notes on greedy algorithms for submodular maximization* 2016.
140. Hori, M., Bayne, C. J., and Kuwae, T. In *Blue carbon in shallow coastal ecosystems* Springer, 2019.
141. Howard, R. A. *Dynamic programming and Markov processes* John Wiley, 1960.
142. Iglesias, M. A., Law, K. J., and Stuart, A. M. Ensemble Kalman methods for inverse problems. *Inverse Problems* **29**, 2013.
143. Jakuba, M. and Yoerger, D. R. *Autonomous search for hydrothermal vent fields with occupancy grid maps* in *Proc. of ACRA* **8**. 2008.
144. Jakuba, M. V. *Stochastic mapping for chemical plume source localization with application to autonomous hydrothermal vent discovery* PhD thesis (Massachusetts Institute of Technology, 2007).
145. Jannasch, H. W. and Mottl, M. J. Geomicrobiology of deep-sea hydrothermal vents. *Science* **229**, 1985.
146. Jayne, S. R. *et al.* The Argo program: present and future. *Oceanography* **30**, 2017.
147. Jiahao, T. Z., Chee, K. Y., and Hsieh, M. A. *Online Dynamics Learning for Predictive Control with an Application to Aerial Robots* in *6th Annual Conference on Robot Learning*
148. Jiahao, T. Z., Hsieh, M. A., and Forgoston, E. Knowledge-based learning of non-linear dynamics and chaos. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **31**, 2021.
149. Johnson, K. S., Beehler, C. L., Sakamoto-Arnold, C. M., and Childress, J. J. In situ measurements of chemical distributions in a deep-sea hydrothermal vent field. *Science* **231**, 1986.
150. Jones, D. A. and Titi, E. S. Upper bounds on the number of determining modes, nodes, and volume elements for the Navier-Stokes equations. *Indiana University Mathematics Journal*, 1993.
151. Jones, D. R., Schonlau, M., and Welch, W. J. Efficient global optimization of expensive black-box functions. *J. Global Optim.* **13**, 1998.
152. Joshi, S. and Boyd, S. Sensor selection via convex optimization. *IEEE Trans. Signal Process.* **57**, 2008.

153. Kaelbling, L. P., Littman, M. L., and Cassandra, A. R. Planning and acting in partially observable stochastic domains. *Artif. Intell.* **101**, 1998.
154. Kaiser, C. L. *et al.* *The design and 200 day per year operation of the autonomous underwater vehicle Sentry in 2016 IEEE/OES Autonomous Underwater Vehicles (AUV)* 2016.
155. Kaiser, E., Kutz, J. N., and Brunton, S. L. Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proceedings of the Royal Society A* **474**, 2018.
156. Kaliszewski, M. *et al.* The multi-gas sensor for remote UAV and UGV missions: Development and tests. *Sensors* **21**, 2021.
157. Kapit, J. and Michel, A. P. Dissolved gas sensing using an anti-resonant hollow core optical fiber. *Applied Optics* **60**, 2021.
158. Kapit, J. and Michel, A. P. *Measurement of dissolved gases using a hollow core optical fiber and capillary membrane inlet* in *CLEO: Applications and Technology* 2021.
159. Karaman, S. and Frazzoli, E. Sampling-based algorithms for optimal motion planning. *The international journal of robotics research* **30**, 2011.
160. Karaman, S., Walter, M. R., Perez, A., Frazzoli, E., and Teller, S. *Anytime motion planning using the RRT* in *IEEE Int. Conf. Robot. Autom.* 2011.
161. Killick, R., Fearnhead, P., and Eckley, I. A. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association* **107**, 2012.
162. Kim, J. *et al.* Discovery of active hydrothermal vent fields along the Central Indian Ridge, 8–12° S. *Geochemistry, Geophysics, Geosystems* **21**, 2020.
163. Kim, S., Ji, W., Deng, S., Ma, Y., and Rackauckas, C. Stiff neural ordinary differential equations. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **31**, 2021.
164. Kingravi, H. A., Maske, H. R., and Chowdhary, G. *Kernel observers: Systems-theoretic modeling and inference of spatiotemporally evolving processes* in *Advances in Neural Information Processing Systems* 2016.
165. Kinsey, J. C. *et al.* *Assessing the deepwater horizon oil spill with the sentry autonomous underwater vehicle in 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* 2011.
166. Kleiber, W., Katz, R. W., and Rajagopalan, B. Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resources Research* **48**, 2012.
167. Koch, J. Data-driven surrogates of rotating detonation engine physics with neural ordinary differential equations and high-speed camera footage. *Physics of Fluids* **33**, 2021.

168. Kocsis, L. and Szepesvári, C. *Bandit based Monte-Carlo planning* in *Proc. Euro. Conf. Mach. Learn.* 2006.
169. Kollar, T. and Roy, N. Trajectory optimization using reinforcement learning for map exploration. *Int. J. Robot. Res.* **27**, 2008.
170. Krause, A., Singh, A., and Guestrin, C. Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies. *J. Mach. Learn. Res.* **9**, 2008.
171. Kryloff, N. and Bogoliouboff, N. La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire. *Annals of mathematics*, 1937.
172. Kucukelbir, A., Tran, D., Ranganath, R., Gelman, A., and Blei, D. M. Automatic differentiation variational inference. *J. Mach. Learn. Res.* **18**, 2017.
173. Kulkarni, C. S. and Lermusiaux, P. F. Advection without compounding errors through flow map composition. *Journal of Computational Physics* **398**, 2019.
174. Kullback, S. and Leibler, R. A. On information and sufficiency. *The annals of mathematical statistics* **22**, 1951.
175. Kumagai, H. *et al.* Hydrothermal plumes imaged by high-resolution side-scan sonar on a cruising AUV, Urashima. *Geochemistry, Geophysics, Geosystems* **11**, 2010.
176. Kurniawati, H., Hsu, D., and Lee, W. S. *SARSOP: Efficient point-based POMDP planning by approximating optimally reachable belief spaces*. in *Robotics: Science and systems* 2008.
177. Kushner, H. J. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *J. Basic Eng.* **86**, 1964.
178. Lassila, T., Manzoni, A., Quarteroni, A., and Rozza, G. In *Reduced Order Methods for modeling and computational reduction* Springer, 2014.
179. Lavelle, J., DiIorio, D., and Rona, P. A turbulent convection model with an observational context for a deep-sea hydrothermal plume in a time-variable cross flow. *Journal of Geophysical Research: Oceans* **118**, 2013.
180. Le Bris, N. *et al.* Hydrothermal energy transfer and organic carbon production at the deep seafloor. *Frontiers in Marine Science* **5**, 2019.
181. Levine, D., Luders, B., and How, J. In *AIAA Infotech at Aerospace 2010* AIAA, 2010.
182. Lewicki, M. S. and Sejnowski, T. J. *Coding time-varying signals using sparse, shift-invariant representations* in *Advances in neural information processing systems* 1999.
183. Li, G. *et al.* Increasing ocean stratification over the past half-century. *Nature Climate Change* **10**, 2020.

184. Li, S., Guo, Y., and Bingham, B. *Multi-robot cooperative control for monitoring and tracking dynamic plumes* in *2014 IEEE International Conference on Robotics and Automation (ICRA)* 2014.
185. Lim, Z. W., Hsu, D., and Lee, W. S. Adaptive informative path planning in metric spaces. *Int. J. Robot. Res.* **35**, 2016.
186. Liu, H., Shah, S., and Jiang, W. On-line outlier detection and data cleaning. *Computers & chemical engineering* **28**, 2004.
187. Liu, J. S. Metropolized independent sampling with comparisons to rejection sampling and importance sampling. *Statistics and computing* **6**, 1996.
188. Liu, M., Sivakumar, K., Omidshafiei, S., Amato, C., and How, J. P. *Learning for multi-robot cooperation in partially observable stochastic environments with macro-actions* in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* 2017.
189. Liu, P., Hou, D., Wang, K., and Wang, B. In *Advances in Guidance, Navigation and Control* Springer, 2022.
190. Long, Z., Lu, Y., Ma, X., and Dong, B. PDE-net: Learning PDEs from data. *arXiv preprint arXiv:1710.09668*, 2017.
191. Lonsdale, P. and Becker, K. Hydrothermal plumes, hot springs, and conductive heat flow in the Southern Trough of Guaymas Basin. *Earth and Planetary Science Letters* **73**, 1985.
192. Lu, P. Y., Kim, S., and Soljačić, M. Extracting interpretable physical parameters from spatiotemporal systems using unsupervised learning. *Physical Review X* **10**, 2020.
193. Lucia, D. J., Beran, P. S., and Silva, W. A. Reduced-order modeling: new approaches for computational physics. *Progress in aerospace sciences* **40**, 2004.
194. Ludwig, F., Gasiorek, L., and Ruff, R. Simplification of a Gaussian puff model for real-time minicomputer use. *Atmospheric Environment (1967)* **11**, 1977.
195. Luo, W. and Sycara, K. *Adaptive sampling and online learning in multi-robot sensor coverage with mixture of gaussian processes* in *IEEE Int. Conf. Robot. Autom.* 2018.
196. Ma, K.-C., Liu, L., and Sukhatme, G. S. *Informative planning and online learning with sparse Gaussian processes* in *2017 IEEE International Conference on Robotics and Automation (ICRA)* 2017.
197. MacKay, D. J. In *Learning in graphical models* Springer, 1998.
198. Majumdar, A. and Tedrake, R. In *Proc. Workshop Algorithmic Foundations Robot.* Springer, 2012.
199. Maki, T., Sato, Y., Matsuda, T., Shiroku, R.-T., and Sakamaki, T. *AUV Tri-TON 2: An intelligent platform for detailed survey of hydrothermal vent fields* in *2014 IEEE/OES Autonomous Underwater Vehicles (AUV)* 2014.

200. Marchant, R., Ramos, F., and Sanner, S. *Sequential Bayesian optimisation for spatial-temporal monitoring* in *Proc. 13th Conf. Uncertainty Artif. Intell.* 2014.
201. Mardt, A., Pasquali, L., Noé, F., and Wu, H. *Deep learning Markov and Koopman models with physical constraints* in *Mathematical and Scientific Machine Learning* 2020.
202. Martin, A. *et al.* *The oceans' twilight zone must be studied now, before it is too late* 2020.
203. Martin, W., Baross, J., Kelley, D., and Russell, M. J. Hydrothermal vents and the origin of life. *Nature Reviews Microbiology* **6**, 2008.
204. Mason, J. C. *et al.* Evaluation of AUV search strategies for the localization of hydrothermal venting, 2020.
205. Mayer, L. *et al.* The Nippon Foundation—GEBCO seabed 2030 project: The quest to see the world's oceans completely mapped by 2030. *Geosciences* **8**, 2018.
206. McCammon, S. *et al.* Ocean front detection and tracking using a team of heterogeneous marine vehicles. *Journal of Field Robotics* **38**, 2021.
207. Mc Climans, T. Estuarine fronts and river plumes. *Physical processes in estuaries*, 1988.
208. McGill, K. and Taylor, S. Robot algorithms for localization of multiple emission sources. *ACM Computing Surveys (CSUR)* **43**, 2011.
209. McGonigle, A. *et al.* Unmanned aerial vehicle measurements of volcanic carbon dioxide fluxes. *Geophysical research letters* **35**, 2008.
210. McPhail, S. *et al.* *Challenges of using an AUV to find and map hydrothermal vent sites in deep and rugged terrains* in *2010 IEEE/OES Autonomous Underwater Vehicles* 2010.
211. Meliou, A., Krause, A., Guestrin, C., and Hellerstein, J. M. *Nonmyopic informative path planning in spatio-temporal models* in *AAAI* **10**. 2007.
212. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. Equation of state calculations by fast computing machines. *The journal of chemical physics* **21**, 1953.
213. Michel, A. P., Preston, V. L., Fauria, K. E., and Nicholson, D. P. Observations of shallow methane bubble emissions from Cascadia Margin. *Frontiers in Earth Science* **9**, 2021.
214. Michel, A. P. *et al.* *Gas sensing in the deep ocean: advancing our ability to chemically explore* in. 2022.
215. Miloshevich, L. M., Paukkunen, A., Vömel, H., and Oltmans, S. J. Development and validation of a time-lag correction for Vaisala radiosonde humidity measurements. *Journal of Atmospheric and Oceanic Technology* **21**, 2004.

216. Mittelstaedt, E. *et al.* Quantifying diffuse and discrete venting at the Tour Eiffel vent site, Lucky Strike hydrothermal field. *Geochemistry, Geophysics, Geosystems* **13**, 2012.
217. Mohan, A., Daniel, D., Chertkov, M., and Livescu, D. Compressed convolutional LSTM: An efficient deep learning framework to model high fidelity 3D turbulence. *arXiv preprint arXiv:1903.00033*, 2019.
218. Moon, T. K. The expectation-maximization algorithm. *IEEE Signal processing magazine* **13**, 1996.
219. Moore, A. M. *et al.* The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part I–System overview and formulation. *Progress in Oceanography* **91**, 2011.
220. Moore, D. G. and Curray, J. R. Geologic and tectonic history of the Gulf of California. *Initial Reports of the Deep Sea Drilling Project* **64**, 1982.
221. Morere, P., Marchant, R., and Ramos, F. *Continuous state-action-observation POMDPs for trajectory planning with Bayesian optimisation* in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.* 2018.
222. Morse, T. M., Lockery, S. R., and Ferrée, T. C. Robust spatial navigation in a robot inspired by chemotaxis in *Caenorhabditis elegans*. *Adaptive Behavior* **6**, 1998.
223. Morton, B., Taylor, G. I., and Turner, J. S. Turbulent gravitational convection from maintained and instantaneous sources. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 1956.
224. Murch, A. P. *et al.* Volcaniclastic dispersal during submarine lava effusion: the 2012 eruption of Havre Volcano, Kermadec Arc, New Zealand. *Frontiers in Earth Science* **8**, 2020.
225. Muscato, G., Bonaccorso, F., Cantelli, L., Longo, D., and Melita, C. D. Volcanic environments: Robots for exploration and measurement. *IEEE Robotics & Automation Magazine* **19**, 2012.
226. Nair, S., Savarese, S., and Finn, C. Goal-Aware Prediction: Learning to Model What Matters. *arXiv preprint arXiv:2007.07170*, 2020.
227. Nakamura, K. *et al.* Discovery of a new hydrothermal vent based on an underwater, high-resolution geophysical survey. *Deep Sea Research Part I: Oceanographic Research Papers* **74**, 2013.
228. Nam, L., Huang, L., Li, X., and Xu, J. *An approach for coverage path planning for UAVs* in *IEEE 14th Int. Workshop Advance. Motion Control* 2016.
229. Navier, C.-L. M. H. Sur les lois du mouvement des fluides, en ayant égard à l’adhésion de leurs molécules. *Bull. Soc. Philomath*, 1822.
230. Neal, R. M. *et al.* MCMC using Hamiltonian dynamics. *Handbook of Markov chain Monte Carlo*, 2011.

231. Nemhauser, G. L., Wolsey, L. A., and Fisher, M. L. An analysis of approximations for maximizing submodular set functions—I. *Math. Program.* **14**, 1978.
232. Neumann, F. *et al.* Heat flow and thermal regime in the Guaymas Basin, Gulf of California: Estimates of conductive and advective heat transport in EGU General Assembly 2022 2022.
233. Nicholson, D. P. and Feen, M. L. Air calibration of an oxygen optode on an underwater glider. *Limnology and Oceanography: Methods* **15**, 2017.
234. Nikolos, I. K., Valavanis, K. P., Tsourveloudis, N. C., and Kostaras, A. N. Evolutionary algorithm based offline/online path planner for UAV navigation. *IEEE Trans. Syst., Man, Cybern. B, Cybern.* **33**, 2003.
235. Noor, A. K. and Peters, J. M. Reduced basis technique for nonlinear analysis of structures. *Aiaa journal* **18**, 1980.
236. Nunn, P. D. Nature-society interactions in the pacific islands. *Geografiska Annaler: Series B, Human Geography* **85**, 2003.
237. Ober, S. W., Rasmussen, C. E., and van der Wilk, M. *The promises and pitfalls of deep kernel learning in Uncertainty in Artificial Intelligence* 2021.
238. Ohlberger, M. and Rave, S. Reduced basis methods: Success, limitations and future challenges. *arXiv preprint arXiv:1511.02021*, 2015.
239. Okamoto, A., Seta, T., Sasano, M., Inoue, S., and Ura, T. Visual and autonomous survey of hydrothermal vents using a hovering-type AUV: Launching Hobalin into the Western Offshore of Kumejima Island. *Geochemistry, Geophysics, Geosystems* **20**, 2019.
240. Ondréas, H., Scalabrin, C., Fouquet, Y., and Godfroy, A. Recent high-resolution mapping of Guaymas hydrothermal fields (Southern Trough). *Bulletin de la Société Géologique de France* **189**, 2018.
241. Ouyang, R., Low, K. H., Chen, J., and Jaillet, P. *Multi-robot active sensing of non-stationary Gaussian process-based environmental phenomena* in *Proc. Int. Conf. Auton. Agents. Multiagent Sys.* Int. Foundation Auton. Agents Multiagent Sys., 2014.
242. Pacelli, V. and Majumdar, A. *Task-driven estimation and control via information bottlenecks* in *2019 International Conference on Robotics and Automation (ICRA)* 2019.
243. Paduan, J. B. *et al.* Discovery of hydrothermal vent fields on Alarcón Rise and in Southern Pescadero Basin, Gulf of California. *Geochemistry, Geophysics, Geosystems* **19**, 2018.
244. Pakravan, S., Mistani, P. A., Aragon-Calvo, M. A., and Gibou, F. Solving inverse-PDE problems with physics-aware neural networks. *Journal of Computational Physics* **440**, 2021.

245. Pang, S. *Plume source localization for AUV based autonomous hydrothermal vent discovery* in *OCEANS 2010 MTS/IEEE SEATTLE* 2010.
246. Papadimitriou, C., Beck, J. L., and Au, S.-K. Entropy-based optimal sensor location for structural model updating. *J. Vib. Control* **6**, 2000.
247. Pathak, J., Hunt, B., Girvan, M., Lu, Z., and Ott, E. Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. *Physical review letters* **120**, 2018.
248. Pathak, J., Lu, Z., Hunt, B. R., Girvan, M., and Ott, E. Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **27**, 2017.
249. Pearson, M. N. *The Indian Ocean* Routledge, 2003.
250. Peng, L., Lipinski, D., and Mohseni, K. Dynamic data driven application system for plume estimation using UAVs. *Journal of Intelligent & Robotic Systems* **74**, 2014.
251. Petersen, J. M. and Dubilier, N. Methanotrophic symbioses in marine invertebrates. *Environmental Microbiology Reports* **1**, 2009.
252. Petitt, R., Bowen, A., Elder, R., Howland, J., and Naiman, M. *Power system for the new Jason ROV in Oceans' 04 MTS/IEEE Techno-Ocean'04 (IEEE Cat. No. 04CH37600)* **3**. 2004.
253. Portwood, G. D. *et al.* Turbulence forecasting via neural ode. *arXiv preprint arXiv:1911.05180*, 2019.
254. Preston, V. L. *Adaptive sampling of transient environmental phenomena with autonomous mobile platforms* Massachusetts Institute of Technology, 2019.
255. Preston, V., Flaspohler, G., Fisher, J., Michel, A., and Roy, N. Physically-informed operational robotic trajectories for scientific expeditions. *Preprint*, 2022.
256. Preston, V. *et al.* Discovering hydrothermalism from afar: in situ methane instrumentation and change-point detection for decision-making. *Frontiers in Earth Science* **10**. DOI: [10.3389/feart.2022.984355](https://doi.org/10.3389/feart.2022.984355), 2022.
257. Puonti, O., Iglesias, J. E., and Van Leemput, K. Fast and sequence-adaptive whole-brain segmentation using parametric Bayesian modeling. *NeuroImage* **143**, 2016.
258. Purkis, S. and Chirayath, V. Remote sensing the ocean biosphere. *Annual Review of Environment and Resources* **47**, 2022.
259. Qian, E., Kramer, B., Peherstorfer, B., and Willcox, K. Lift & Learn: Physics-informed machine learning for large-scale nonlinear dynamical systems. *Physica D: Nonlinear Phenomena* **406**, 2020.
260. Quarteroni, A. and Rozza, G. Numerical solution of parametrized Navier–Stokes equations by reduced basis methods. *Numerical Methods for Partial Differential Equations: An International Journal* **23**, 2007.

261. Queralta, J. P. *et al.* Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision. *Ieee Access* **8**, 2020.
262. Qureshi, U. M. *et al.* RF path and absorption loss estimation for underwater wireless sensor networks in different water environments. *Sensors* **16**, 2016.
263. Raissi, M., Perdikaris, P., and Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* **378**, 2019.
264. Raissi, M., Perdikaris, P., and Karniadakis, G. E. Numerical Gaussian processes for time-dependent and nonlinear partial differential equations. *SIAM Journal on Scientific Computing* **40**, 2018.
265. Raja, P. and Pugazhenthi, S. Optimal path planning of mobile robots: A review. *International journal of physical sciences* **7**, 2012.
266. Ramondenc, P., Germanovich, L. N., Von Damm, K. L., and Lowell, R. P. The first measurements of hydrothermal heat output at 9°50' N, East Pacific Rise. *Earth and Planetary Science Letters* **245**, 2006.
267. Rasmussen, C. E. and Williams, C. K. *Gaussian processes for machine learning*. MIT Press MIT Press, 2004.
268. Raven, J. *et al.* *Ocean acidification due to increasing atmospheric carbon dioxide* The Royal Society, 2005.
269. Reddy, G., Murthy, V. N., and Vergassola, M. Olfactory sensing and navigation in turbulent environments. *Annual Review of Condensed Matter Physics* **13**, 2022.
270. Resing, J. A. *et al.* Basin-scale transport of hydrothermal dissolved metals across the South Pacific Ocean. *Nature* **523**, 2015.
271. Rheuban, J. E., Doney, S. C., McCorkle, D. C., and Jakuba, R. W. Quantifying the effects of nutrient enrichment and freshwater mixing on coastal ocean acidification. *Journal of Geophysical Research: Oceans* **124**, 2019.
272. Richards, T. S., Aubourg, Q., and Sutherland, B. R. Radial intrusions from turbulent plumes in uniform stratification. *Physics of Fluids* **26**, 2014.
273. Rigby, P., Pizarro, O., and Williams, S. B. *Towards geo-referenced AUV navigation through fusion of USBL and DVL measurements* in *OCEANS 2006* 2006.
274. Rizk, Y., Awad, M., and Tunstel, E. W. Cooperative heterogeneous multi-robot systems: A survey. *ACM Computing Surveys (CSUR)* **52**, 2019.
275. Roa, M. A., Argus, M. J., Leidner, D., Borst, C., and Hirzinger, G. *Power grasp planning for anthropomorphic robot hands* in *IEEE Int. Conf. Robot. Autom.* 2012.
276. Roberts, P. H. and Thomas, K. V. The occurrence of selected pharmaceuticals in wastewater effluent and surface waters of the lower Tyne catchment. *Sci. Total Environ.* **356**, 2006.

277. Roemmich, D. *et al.* The Argo Program: Observing the global ocean with profiling floats. *Oceanography* **22**, 2009.
278. Rudnick, D. L. and Klinke, J. The underway conductivity–temperature–depth instrument. *Journal of Atmospheric and Oceanic Technology* **24**, 2007.
279. Rudy, S. H., Brunton, S. L., Proctor, J. L., and Kutz, J. N. Data-driven discovery of partial differential equations. *Science Advances* **3**, 2017.
280. Ryan, J. *et al.* A high-resolution survey of a deep hydrocarbon plume in the Gulf of Mexico during the 2010 Macondo blowout. *Geophysical Monograph Series* **195**, 2011.
281. Saigol, Z. A., Dearden, R., Wyatt, J. L., and Murton, B. J. *Information-lookahead planning for AUV mapping* in *IJCAI 2009*.
282. Salam, T. and Hsieh, M. A. Adaptive sampling and reduced-order modeling of dynamic processes by robot teams. *IEEE Robotics and Automation Letters* **4**, 2019.
283. Sapsis, T. P. and Lermusiaux, P. F. Dynamically orthogonal field equations for continuous stochastic dynamical systems. *Physica D: Nonlinear Phenomena* **238**, 2009.
284. Sarkar, S., Sarkar, S., Virani, N., Ray, A., and Yasar, M. Sensor fusion for fault detection and classification in distributed physical processes. *Frontiers in Robotics and AI* **1**, 2014.
285. Särkkä, S. *Bayesian filtering and smoothing* **3**. Cambridge university press, 2013.
286. Saul, L. K. and Roweis, S. T. An introduction to locally linear embedding. *unpublished. Available at: <http://www.cs.toronto.edu/roweis/lle/publications.html>*, 2000.
287. Schmid, F. *et al.* Physico-chemical properties of newly discovered hydrothermal plumes above the Southern Mid-Atlantic Ridge (13°-33° S). *Deep Sea Research Part I: Oceanographic Research Papers* **148**, 2019.
288. Schmid, P. J. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics* **656**, 2010.
289. Schölkopf, B., Smola, A., and Müller, K.-R. *Kernel principal component analysis* in *International conference on artificial neural networks* 1997.
290. Scholz, F. *et al.* Shelf-to-basin iron shuttle in the Guaymas Basin, Gulf of California. *Geochimica et Cosmochimica Acta* **261**, 2019.
291. Seewald, J. S., Seyfried Jr, W. E., and Shanks III, W. C. Variations in the chemical and stable isotope composition of carbon and sulfur species during organic-rich sediment alteration: an experimental and theoretical study of hydrothermal activity at Guaymas Basin, Gulf of California. *Geochimica et Cosmochimica Acta* **58**, 1994.

292. Seibel, B. A. and Walsh, P. J. Potential impacts of CO₂ injection on deep-sea biota. *Science* **294**, 2001.
293. Shankar, V. *et al.* *Learning non-linear spatio-temporal dynamics with convolutional Neural ODEs* in *Third Workshop on Machine Learning and the Physical Sciences (NeurIPS 2020)* 2020.
294. Shannon, C. E. and Weaver, W. *The mathematical theory of communication* University of Illinois press, 1998.
295. Sharma, R. Environmental issues of deep-sea mining. *Procedia Earth and Planetary Science* **11**, 2015.
296. Al-Shedivat, M., Wilson, A. G., Saatchi, Y., Hu, Z., and Xing, E. P. Learning scalable deep kernels with recurrent structure. *The Journal of Machine Learning Research* **18**, 2017.
297. Silver, D. and Veness, J. *Monte-Carlo planning in large POMDPs* in *Proc. Adv. Neural Inf. Process. Syst.* 2010.
298. Sim, R. and Roy, N. *Global a-optimal robot exploration in SLAM* in *Proc. IEEE Int. Conf. Robot. Autom.* 2005.
299. Singh, A., Krause, A., and Kaiser, W. J. *Nonmyopic adaptive informative path planning for multiple robots* in *Proc. 21st Int. Joint Conf. Artif. Intell.* 2009.
300. Singh, A., Ramos, F., Whyte, H. D., and Kaiser, W. J. *Modeling and decision making in spatio-temporal processes for environmental surveillance* in *2010 IEEE International Conference on Robotics and Automation* 2010.
301. Singh, S., Chen, M., Herbert, S. L., Tomlin, C. J., and Pavone, M. *Robust tracking with model mismatch for fast and safe planning: an SOS optimization approach* in *International Workshop on the Algorithmic Foundations of Robotics* 2018.
302. Smith Jr, K. Maerozooplankton of a deep sea hydrothermal vent: In situ rates of oxygen consumption 1. *Limnology and Oceanography* **30**, 1985.
303. Smith, C. R. *et al.* Deep-sea misconceptions cause underestimation of seabed-mining impacts. *Trends in Ecology & Evolution* **35**, 2020.
304. Smith, G. D. *Numerical solution of partial differential equations: finite difference methods* Oxford university press, 1985.
305. Smith, J. G. *et al.* Volcano monitoring. *Geological Monitoring*, 2009.
306. Snoek, J., Larochelle, H., and Adams, R. P. *Practical Bayesian optimization of machine learning algorithms* in *Proc. Adv. Neural Info. Process. Syst.* 2012.
307. Sohn, R. A. *et al.* Observations and modeling of a hydrothermal plume in Yellowstone Lake. *Geophysical Research Letters* **46**, 2019.
308. Somani, A., Ye, N., Hsu, D., and Lee, W. S. *DESPOT: Online POMDP planning with regularization* in *Proc. Adv. Neural Inf. Process. Syst.* 2013.

309. Soule, S. A. *et al.* Exploration of the Northern Guaymas Basin. *Oceanography* **31**, 2018.
310. Spantini, A., Baptista, R., and Marzouk, Y. Coupling techniques for nonlinear ensemble filtering. *arXiv preprint arXiv:1907.00389*, 2019.
311. Spantini, A., Bigoni, D., and Marzouk, Y. Inference via low-dimensional couplings. *The Journal of Machine Learning Research* **19**, 2018.
312. Spantini, A. *et al.* Optimal low-rank approximations of Bayesian linear inverse problems. *SIAM Journal on Scientific Computing* **37**, 2015.
313. Speer, K. G. and Rona, P. A. A model of an Atlantic and Pacific hydrothermal plume. *Journal of Geophysical Research: Oceans* **94**, 1989.
314. Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. W. Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. *IEEE Trans. Inf. Theory* **58**, 2012.
315. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* **15**, 2014.
316. Stanley, E. H. *et al.* The ecology of methane in streams and rivers: patterns, controls, and global significance. *Ecol. Monographs* **86**, 2016.
317. Stannat, W. Stochastic partial differential equations: Kolmogorov operators and invariant measures. *Jahresbericht der Deutschen Mathematiker-Vereinigung* **113**, 2011.
318. Stix, J. *et al.* Using drones and miniaturized instrumentation to study degassing at Turrialba and Masaya volcanoes, Central America. *Journal of Geophysical Research: Solid Earth* **123**, 2018.
319. Stokes, G. G. *et al.* On the effect of the internal friction of fluids on the motion of pendulums, 1851.
320. Stuart, A. M. Inverse problems: a Bayesian perspective. *Acta numerica* **19**, 2010.
321. Stutters, L., Liu, H., Tiltman, C., and Brown, D. J. Navigation technologies for autonomous underwater vehicles. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **38**, 2008.
322. Sun, S. *et al.* Differentiable compositional kernel learning for Gaussian processes. *arXiv preprint arXiv:1806.04326*, 2018.
323. Sun, W. *et al.* No-regret replanning under uncertainty in Proc. IEEE Int. Conf. Robot. Autom. 2017.
324. Sun, Y., Zhang, L., and Schaeffer, H. NEUPDE: Neural network based ordinary and partial differential equations for modeling time-dependent data in Mathematical and Scientific Machine Learning 2020.

325. Sunberg, Z. N. and Kochenderfer, M. J. *Online algorithms for POMDPs with continuous state, action, and observation spaces* in *Proc. 24th Int. Conf. Automated Plan. Schedul.* 2018.
326. Tartakovsky, A. M., Marrero, C. O., Perdikaris, P., Tartakovsky, G. D., and Barajas-Solano, D. Learning parameters and constitutive relationships with physics informed deep neural networks. *arXiv preprint arXiv:1808.03398*, 2018.
327. Taylor, B. W. *et al.* Improving the fluorometric ammonium method: matrix effects, background fluorescence, and standard additions. *Journal of the North American Benthological Society* **26**, 2007.
328. Teixeira, D., de Sousa, J. B., Mendes, R., and Fonseca, J. *3D tracking of a river plume front with an AUV* in *OCEANS 2021: San Diego–Porto 2021*.
329. Teng, Y. and Zhang, D. Long-term viability of carbon sequestration in deep-sea sediments. *Science advances* **4**, 2018.
330. Teske, A. *et al.* Characteristics and evolution of sill-driven off-axis hydrothermalism in Guaymas Basin—the Ringvent site. *Scientific reports* **9**, 2019.
331. Teske, A. *et al.* The Guaymas Basin hiking guide to hydrothermal mounds, chimneys, and microbial mats: Complex seafloor expressions of subsurface hydrothermal circulation. *Frontiers in Microbiology* **7**, 2016.
332. Thielicke, W. The flapping flight of birds. *Diss. University of Groningen*, 2014.
333. Thielicke, W. and Sonntag, R. Particle Image Velocimetry for MATLAB: Accuracy and enhanced algorithms in PIVlab. *Journal of Open Research Software* **9**, 2021.
334. Thielicke, W. and Stamhuis, E. PIVlab—towards user-friendly, affordable and accurate digital particle image velocimetry in MATLAB. *Journal of open research software* **2**, 2014.
335. Thompson, K. F., Miller, K. A., Currie, D., Johnston, P., and Santillo, D. Seabed mining and approaches to governance of the deep seabed. *Frontiers in Marine Science* **5**, 2018.
336. Tian, Y. *et al.* A behavior-based planning strategy for deep-sea hydrothermal plume tracing with autonomous underwater vehicles in *OCEANS 2014-TAIPEI 2014*.
337. Timmons, E., Ayton, B., Camilli, R., Williams, B., *et al.* Automated experiment selection and scheduling in the Costa Rica subduction zone in *AGU Fall Meeting Abstracts* **2019**. 2019.
338. Tohidi, A. and Kaye, N. B. Highly buoyant bent-over plumes in a boundary layer. *Atmospheric Environment* **131**, 2016.
339. Touzé, C., Vizzaccaro, A., and Thomas, O. Model order reduction methods for geometrically nonlinear structures: a review of nonlinear techniques. *Nonlinear Dynamics* **105**, 2021.

340. Trowbridge, J. *et al.* The ocean observatories initiative. *Frontiers in Marine Science*, 2019.
341. Truong, C., Oudre, L., and Vayatis, N. Selective review of offline change point detection methods. *Signal Processing* **167**, 2020.
342. *USGS operates five U.S. volcano observatories* 2019.
343. Usui, M., Niki, H., and Kohno, T. Adaptive Gauss-Seidel method for linear systems. *International Journal of Computer Mathematics* **51**, 1994.
344. Valencia, R. and Andrade-Cetto, J. In *Mapping, Planning and Exploration with Pose SLAM* Springer, 2018.
345. Valiela, I., Owens, C., Elmstrom, E., and Lloret, J. Eutrophication of Cape Cod estuaries: Effect of decadal changes in global-driven atmospheric and local-scale wastewater nutrient loads. *Mar. Pollut. Bull.* **110**, 2016.
346. Van den Bremer, T. and Hunt, G. R. Universal solutions for Boussinesq and non-Boussinesq plumes. *Journal of fluid mechanics* **644**, 2010.
347. Van Dover, C. L. Tighten regulations on deep-sea mining. *Nature* **470**, 2011.
348. Vergassola, M., Villermaux, E., and Shraiman, B. I. 'Infotaxis' as a strategy for searching without gradients. *Nature* **445**, 2007.
349. Vic, C., Gula, J., Roullet, G., and Pradillon, F. Dispersion of deep-sea hydrothermal vent effluents and larvae by submesoscale and tidal currents. *Deep Sea Research Part I: Oceanographic Research Papers* **133**, 2018.
350. Villani, C. *Optimal transport: old and new* Springer Science & Business Media, 2008.
351. Vizzaccaro, A., Opreni, A., Salles, L., Frangi, A., and Touzé, C. High order direct parametrisation of invariant manifolds for model order reduction of finite element structures: application to large amplitude vibrations and uncovering of a folding point. *Nonlinear Dynamics* **110**, 2022.
352. Von Damm, K., Edmond, J. t., Measures, C., and Grant, B. Chemistry of submarine hydrothermal solutions at Guaymas Basin, Gulf of California. *Geochimica et Cosmochimica Acta* **49**, 1985.
353. Voynov, A. and Babenko, A. *Unsupervised discovery of interpretable directions in the GAN latent space* in *International conference on machine learning* 2020.
354. Wainwright, J. and Mulligan, M. *Environmental modelling* Wiley Online Library, 2002.
355. Wainwright, M. J., Jordan, M. I., *et al.* Graphical models, exponential families, and variational inference. *Foundations and Trends in Machine Learning* **1**, 2008.
356. Wan, Z. Y. and Sapsis, T. P. Reduced-space Gaussian process regression for data-driven probabilistic forecast of chaotic dynamical systems. *Physica D: Nonlinear Phenomena* **345**, 2017.

357. Wang, J.-W., Guo, Y., Fahad, M., and Bingham, B. Dynamic plume tracking by cooperative robots. *IEEE/ASME Transactions on Mechatronics* **24**, 2019.
358. Wang, L., Pang, S., and Xu, G. *3-dimensional hydrothermal vent localization based on chemical plume tracing* in *Global Oceans 2020: Singapore–US Gulf Coast* 2020.
359. Wang, Y., Long, M., Wang, J., Gao, Z., and Philip, S. Y. *PredRNN: Recurrent neural networks for predictive learning using spatiotemporal LSTMs* in *Advances in Neural Information Processing Systems* 2017.
360. Wang, Z., Xing, W., Kirby, R., and Zhe, S. *Physics Informed Deep Kernel Learning* in *International Conference on Artificial Intelligence and Statistics* 2022.
361. Wang, Z., Li, H.-X., and Chen, C. Reinforcement learning-based optimal sensor placement for spatiotemporal modeling. *IEEE transactions on cybernetics* **50**, 2019.
362. Wankel, S. D. *et al.* New constraints on methane fluxes and rates of anaerobic methane oxidation in a Gulf of Mexico brine pool via in situ mass spectrometry. *Deep Sea Research Part II: Topical Studies in Oceanography* **57**, 2010.
363. Welch, G., Bishop, G., *et al.* An introduction to the Kalman filter, 1995.
364. Wesseling, P. von Neumann stability conditions for the convection-diffusion equation. *IMA journal of Numerical Analysis* **16**, 1996.
365. Whitman, J. and Chowdhary, G. *Learning dynamics across similar spatiotemporally-evolving physical systems* in *Conf. on Rob. Learn.* 2017.
366. Wilson, A. G., Hu, Z., Salakhutdinov, R. R., and Xing, E. P. Stochastic variational deep kernel learning. *Advances in Neural Information Processing Systems* **29**, 2016.
367. Wilson, A. G., Hu, Z., Salakhutdinov, R., and Xing, E. P. *Deep kernel learning* in *Artificial Intelligence and Statistics* 2016.
368. Wilson, A. and Nickisch, H. *Kernel interpolation for scalable structured Gaussian processes (KISS-GP)* in *International Conference on Machine Learning* 2015.
369. Wilson, C. *et al.* Hydrothermal anomalies in the Lucky Strike segment on the Mid-Atlantic Ridge (37° 17' N). *Earth and planetary science letters* **142**, 1996.
370. Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. Determining Lyapunov exponents from a time series. *Physica D: nonlinear phenomena* **16**, 1985.
371. Xi, X. *et al.* Constraining the sulfur dioxide degassing flux from Turrialba volcano, Costa Rica using unmanned aerial system measurements. *Journal of Volcanology and Geothermal Research* **325**, 2016.
372. Xu, G. and Di Iorio, D. Deep sea hydrothermal plumes and their interaction with oscillatory flows. *Geochemistry, Geophysics, Geosystems* **13**, 2012.

373. Yoerger, D. R. *et al.* *Mesobot: An autonomous underwater vehicle for tracking and sampling midwater targets in 2018 IEEE/OES autonomous underwater vehicle workshop (AUV)* 2018.
374. Yoerger, D. R., Bradley, A. M., and Walden, B. B. The Autonomous Benthic Explorer (ABE): A deep ocean AUV for scientific seafloor survey. *Woods Hole Oceanographic Institution, Advanced Engineering Laboratory Project Summaries*, 1991.
375. Yoerger, D. R. and Newman, J. B. In *ROV86: Remotely operated vehicles* Springer, 1986.
376. Yoerger, D. R. *et al.* Autonomous and remotely operated vehicle technology for hydrothermal vent discovery, exploration, and sampling. *Oceanography* **20**, 2007.
377. You, J., Li, X., Low, M., Lobell, D., and Ermon, S. *Deep Gaussian process for crop yield prediction based on remote sensing data* in *Thirty-First AAAI conference on artificial intelligence* 2017.
378. Yu, T. *et al.* Combo: Conservative offline model-based policy optimization. *Advances in neural information processing systems* **34**, 2021.
379. Yuh, J. Design and control of autonomous underwater robots: A survey. *Auton. Robot.* **8**, 2000.
380. Zahm, O., Cui, T., Law, K., Spantini, A., and Marzouk, Y. Certified dimension reduction in nonlinear Bayesian inverse problems. *Mathematics of Computation* **91**, 2022.
381. Zhang, B. and Sukhatme, G. S. *Adaptive sampling for estimating a scalar field using a robotic boat and a sensor network* in *Proceedings 2007 IEEE International Conference on Robotics and Automation* 2007.
382. Zhang, T., Yang, J., Zhao, D., and Ge, X. Linear local tangent space alignment and application to face recognition. *Neurocomputing* **70**, 2007.
383. Zhang, X., Lin, J., and Jiang, H. Time-dependent variations in vertical fluxes of hydrothermal plumes at mid-ocean ridges. *Marine Geophysical Research* **40**, 2019.
384. Zhao, S. and Pasqualetti, F. Networks with diagonal controllability Gramian: Analysis, graphical conditions, and design algorithms. *Automatica* **102**, 2019.