

UC Merced

UC Merced Electronic Theses and Dissertations

Title

Digital Twin Enabled Collective Sensing and Steering for Source Determination Problems

Permalink

<https://escholarship.org/uc/item/3vh5d81c>

Author

Hollenbeck, Derek

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, MERCED

Digital Twin Enabled Collective Sensing and Steering for Source Determination Problems

A dissertation submitted in partial fulfilment of the requirements for the degree
Doctor of Philosophy

in

Mechanical Engineering

by

Derek Hollenbeck

Committee in charge:

Professor YangQuan Chen, Chair

Professor Gerardo Diaz

Professor Christopher T. Kello

Professor Ricardo Pinto de Castro

2023

©

Derek Hollenbeck, 2023
All Rights Reserved.

The dissertation of Derek Hollenbeck is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Professor Gerardo Diaz

Professor Christopher T. Kello

Professor Ricardo Pinto de Castro

Professor YangQuan Chen, Chair

University of California, Merced

2023

*To my wife and family,
thank you for always being there and keeping me grounded with your support, love
and understanding throughout my graduate career, without it, this milestone would
not have been possible.*

*To my daughter,
through hard work, persistent effort and grit, anything is possible.*

Contents

List of Figures	vii
List of Tables	xvi
Abbreviations	xviii
Acknowledgment	xxi
Curriculum Vita	xxii
Abstract	xxviii
1 Introduction	1
1.1 Research Motivations	1
1.2 Literature Overview and Challenges	2
1.3 Dissertation Objectives	6
1.4 Dissertation Contributions	7
1.5 Dissertation Outline	8
2 Collective Motion and the Science of Search	9
2.1 The Science of Search	9
2.2 Collective Motion and Steering	10
2.3 A Case Study: On Optimal Tempered Lévy Flight Foraging	11
2.4 A Case Study: Individual and Collective Foraging in Autonomous Search Agents with Human Intervention	20
2.5 A Case Study: Swarm Robotic Source Seeking with Fractional Fluxotaxis	38
2.6 A Case Study: A More Optimal Stochastic Extremum Seeking Control Using Fractional Dithering for A Class of Smooth Convex Function	47
2.7 Chapter Summary	61
3 Leak Detection and Quantification with sUAS	62
3.1 Sensors and Platforms	62
3.2 Leak Quantification Methods	66
3.3 Analysis of Methods	86
3.4 Case Studies: Leak Detection and Localization	89

3.5	Case Studies: Controlled Release Testing and Emission Quantification	102
3.6	Case Studies: Real World Emission Quantification	107
3.7	Chapter Summary	112
4	Towards Smarter Environmental Sensing with Digital Twins	114
4.1	The Digital Twin Framework and POSIM Digital Twin Platform . . .	116
4.2	Methane Odor Abatement Simulator with Digital Twins (MOABS/DT)	123
4.3	A Case Study: Single-sUAS-based Characterization of Ground-To-Air Emissions	128
4.4	A Case Study: Behavior Matching and Localization Using Fixed Sensor Mesh	136
4.5	A Case Study: Multi-sUAS-based Emission Quantification with DE-CVT	139
4.6	A Case Study: Multi-sUAS-based Emission Quantification Using Mod- NGI	146
4.7	A Case Study: Single and Multi-sUAS Emission Quantification Per- formance Assessment using MOABS/DT	152
4.8	Chapter Summary	155
5	Digital Twin Enabled Smart Sensor Placement and Steering	159
5.1	Sensor Placement and Steering Problem	159
5.2	Empirical Observability Gramians	161
5.3	Empirical Identifiability Gramian	162
5.4	Persistence of Excitation and Sufficient Richness	163
5.5	A Case Study: DT-enabled Sensor Placement and Steering with the Empirical Observability Gramian	164
5.6	Chapter Summary	172
6	Conclusions and Future Work	175
	Appendices	177
	List of sUAS-based Field Campaign Missions	177
	Bibliography	179

List of Figures

1.1	This diagram illustrates how the research topics and chapters are connected to the overall scope of the this dissertation.	6
2.1	Searching efficiency of standard Lévy flight for different mean free path λ	14
2.2	Probability density for tempered Lévy distribution with $\rho = 0.5$ for different μ	14
2.3	Searching efficiency $\eta\lambda$ for different order μ and ρ : the nondestructive case with different λ	16
2.4	A typical example of foraging procedure with $\lambda = 1$ and $\mu = 1$	17
2.5	Searching efficiency $\eta\lambda$ for different order μ and ρ : the destructive case with different λ	18
2.6	The actual tempered Lévy density function with $\rho = 0.5$ and $\mu = 1$ and the statistic results of generated jump lengths	18
2.7	Experimental results of searching efficiency for different λ and μ . . .	19
2.8	A typical example of foraging procedure with $\lambda = 0.5$ and $\mu = 1$. . .	19
2.9	A typical example of foraging procedure with $\lambda = 0.5$ and $\mu = 3$. . .	20
2.10	(a) Illustrative view of the task space (not shown to players). The green agent and circle represent the human's agent and their field of view (detailed more below). Area outside their field of view was occluded (greyed out area). The dotted line represents that the human agent moved towards the mouse pointer position, so the human could control movement direction by moving the mouse. (b) Random movement shown to be random angular deviations of movement from each previous heading. (c) Flocking term directed agents to converge towards a similar shared movement trajectory. (d) Distancing term prompted agents to separate from one another when close, and towards each other when further away. (e) Visual Chaining prompted agents to move directly toward an agent flagged as detecting the target.	25
2.11	Example movement trajectories for 2000 ticks for each movement condition: (a) Random; (b) Flocking; (c) Distancing; and (d) Loose Coupling.	26

2.12	Mean trial time per session as a function of movement type without human intervention (left) and with human intervention (right). Mean trial times are divided into their composite search times (red) and consumption times (teal).	27
2.13	Mean targets found per session as a function of movement type without human intervention (left) and with human intervention (right). . . .	28
2.14	Normalized histogram of the number of agents in view during the search time period respective to one autonomous agent without intervention.	30
2.15	Normalized histogram of the number of agents in view during the search time period respective to only human agents.	31
2.16	Entropy values for simulated agents as a function of movement type (Random, Flocking, Distancing, and Loose Coupling) in the experiment (With Human) versus the simulation (Without Human).	31
2.17	Entropy values by movement condition (Random, Flocking, Distancing, and Loose Coupling) for human players (Human) versus simulated agents (Autonomous) in the experiment.	32
2.18	Area search rate for autonomous agents, averaged for each session, and plotted as a function of movement type and human intervention for agents individually (left) and collectively (right).	34
2.19	Area search rate averaged for each session and plotted as a function of movement type and agent type for human players and individual agents with intervention.	35
2.20	Finding and consuming proportions for human players as a function of movement type. The red line indicates the expected proportion to be found and consumed if human performance was no better than that of simulated agents. Proportions above the expected baseline indicate the degree to which humans outperformed simulated agents.	37
2.21	An example of the FO-fluxotaxis algorithm given a meandering plume ($a = b = 0.04$ and $G = 5$). The color of the source seeking agent represents the source seeking strategy (green is FO-fluxotaxis). When the plume signal is lost and neighboring information is available, the algorithm reverts back to the secondary source seeking strategy casting (shown in black).	43
2.22	Simulation agent starting locations (blue), general wind direction $\bar{\mathbf{V}}$ (light blue), and source location $\mathbf{x}_s$ (red) for both simulation cases. . .	43
2.23	Case I: <i>Arrival time</i> performance of different swarm strategies given wind conditions driven by colored noise.	45
2.24	Case I: <i>Distance travelled</i> performance of different swarm strategies given wind conditions driven by colored noise	46
2.25	Case I: <i>Closeness</i> performance of different swarm strategies given wind conditions driven by colored noise.	46
2.26	The performance efficiency, $\mathcal{E}$, of all the swarm strategies.	47
2.27	Control diagram of ESC.	49

2.28	Plot of fBm showing (left axis) neutral $H = 0.5$, and anti-persistent behavior $H = 0.1$ and (right axis) persistent behavior $H = 0.9$	51
2.29	Step length distribution for different values of H	51
2.30	The PSD shows the behavior in fGn for $H > 0.5$ which $S_{xx}(\omega) \propto \omega^{-(2H-1)}$	52
2.31	The long range dependence can be seen in the heavy tails of the probability distribution function of SαS for various α	53
2.32	Rastrigin function.	56
2.33	The temporal convergence to the global minimum from the initial condition θ_0 , denoted t^* , is shown as a function of fGn and SαS dithering.	58
2.34	The spatial deviation δ around the converged θ is shown as a function of fGn and SαS dithering.	59
2.35	The final converged θ is shown as function of fGn and SαS dithering.	59
2.36	The temporal convergence to the global minimum from the initial condition θ_0 , denoted t^* , is shown as a function of fGn and SαS dithering.	60
2.37	The spatial deviation δ around the converged θ is shown as a function of fGn and SαS dithering.	60
2.38	The final converged θ is shown as function of fGn and SαS dithering.	61
3.1	Payload configuration examples of (a) boom-mounted TDLAS (e.g. see [1] with an OPLS), (b) top-mounted anemometer showcasing the effect of propellers on the streamlines (e.g. see [2] in wind tunnel tests), and (c) bottom-mounted bs-TDLAS or OGI (e.g. see [3] with a RMLD).	65
3.2	Leak detection and quantification methods and modalities.	66
3.3	A diagram depicting the bLS approach (e.g. see [4]).	69
3.4	(a) Depiction of Gaussian plume dispersion with an observer making a stationary measurement downwind. (b) Resulting time-integrated data with a Gaussian fit applied [5].	71
3.5	(left) Polar plot with the wind direction, θ as the radial axis, and the conditionally averaged concentration, $\langle y \theta \rangle$ as the distance from the center. θ_p is the peak wind direction located at the maximum conditionally averaged concentration. (right) Illustration of the wind direction geometry for conversion of θ to crosswind position $\hat{d}$ with the source plume represented by the dashed lines [6].	73
3.6	Demonstration of using sampled flux plane data and applying kriging to it for spatial interpolation [7].	79
3.7	Example of the VFP-PI strategy via a UAV sensing in circular trajectories with (a) being an internal leak producing a net positive flux and (b) being an external leak producing a net zero flux. The color of the arcs are indicative of methane flux strength with green being more negative and red being more positive (see [3] for more details).	81

3.8	Diagram of summary of methods (based on Table 3.2) showing relationship between typical survey time versus sample distance and there associated normalized accuracy, where lower values represent more accurate measurements. (measured using ¹ fixed, ² foot, ³ vehicle, ⁴ manned aircraft, or ⁵ UAV)	90
3.9	Diagram of the complexity ranking of the methods (based on Table 3.1), showing the relationship between the method complexity (red), precision (green), and cost (blue). The precision is normalized on the source estimate multiplied by 10 and the cost is ranked from 0 to 10. (measured using ¹ fixed, ² foot, ³ vehicle, ⁴ manned aircraft, or ⁵ UAV)	90
3.10	Image showing the OPLS emission frequency ($3057\text{ cm}^{-1} = 3.27\text{ }\mu\text{m}$), the OPLS sensor showing the open path optical head, and the integrated OPLS on the 3DR Solo quadcopter hovering and oriented into the wind	92
3.11	The right shows a satellite view of MVPGR test site. The buildings to the north are an abandoned farm house. The red line is 270 m long. On the left, shows the gas diffuser setup and data from a 30 m downwind flight with a 5 SCFH methane release.	93
3.12	Localization back-trajectories from Feb. 2017 Merced Field test. Shown are 4 examples. Small red lines are back-trajectories from individual methane signals. Fat red lines are the averaged back-trajectories for that flight. The yellow is the predicted plume where the leak location is at the sharp end of the yellow plume.	96
3.13	The PG&E Livermore training center. Top: The numbers in red refer to release points that map to the test matrix provided above. Bottom: Actual release points where green are below-ground leaks and yellow are above-ground leaks.	97
3.14	(a) top: early back trajectories (red lines); (b) 2nd from top: mid-flight back trajectories; (c) 3rd from top: late back trajectories; (d) averaged back trajectories and predicted leak location (red square) vs. actual leak location (yellow star).	98
3.15	This diagram depicts the change in searching strategy from Browian motion to Lévy following the work in [8]	100
3.16	(top) Depicts the autonomous searching path where the searcher detects the emission and switches the searching strategy to a Brownian-based random walk, slowing transitioning to a Lévy walk before continuing on the path. (bottom) Schematic of the switching algorithm.	101
3.17	An experimental example applying the quadi-steady state inversion method to localize a real-world leak. In this case, two sources are present but based on the measured data (shown in the reduced image) only one of the sources can be located.	102
3.18	An example curtain flight path for mass balance calculations.	103

3.19	Emission quantification from controlled release tests in the MVPGR February 2017.	104
3.20	Experiment location at Merced Vernal Pools and Grassland Reserve. The yellow star indicates the source, the green triangle is the RMYoung anemometer, and the flight area is shown red dotted line.	105
3.21	DJI Matrice 210 (M210) equipped with a sensitive methane sniffer developed by [9] and attached to the DJI Skyport.	105
3.22	Normalized controlled release estimation using the mass balance method of the August 2018 MVPGR field test.	106
3.23	Example of an off-site emission source affecting the survey and emission quantification.	107
3.24	Example of a survey with an underground source. The initial survey detected a leak near the south-east corner of the site. The successive flights had difficulty detecting the emission source when flying low and slow ($\approx 1\text{m/s}$ or less), whereas, when flying faster the rate of detection increases.	108
3.25	At the MCRWM site a new cell is being developed, such that, the bank shown in the image has no membrane and only a dirt cover.	109
3.26	Two surveys of an active landfill cell without a membrane are conducted where measurements are taken from: (top) near the source; (bottom) further back. The measurements near the source show the emission emanating from the ground whereas when the measurement is taken from further back the emission is subject to more turbulence.	109
3.27	The mobile sUAS measurement platform above an eddy covariance tower in an Alaskan peatland bog.	111
3.28	A top down view of the Alaskan Peatland Experiment (APEX) site, showing the upwind and downwind measurements for two different wind conditions.	112
3.29	(top) Tower-based flux measurement of the APEX site for the days and times of the 2019 September field campaign (bottom-left) Average flux given the cosine angle between the apparent wind direction and the normal vector of the flux plane (bottom-right) The emission quantification values are shown to be higher when the turbulence intensity is small.	113
4.1	Digital twin framework diagram.	117
4.2	MATLAB example of the digital twin with meandering	119
4.3	Scheme for updating the colored noise in the corners interpolating the adjacent boundaries.	120
4.4	Contrast in measurements of two different wind conditions stable (2016) vs unstable (2017) under controlled methane release of 5 standard cubic feet per hour (SCFH) at 30m downwind [10].	121
4.5	The timeseries behavior and sensitivity with tuning parameters θ given one pass through the plume.	124

4.6	A satellite image of the Colorado State University methane emission technology evaluation center (METEC), sub-divided into sites and equipment pads are numbered for simplicity in topology modeling.	125
4.7	(a) Site level emission survey and quantification from METEC using TDLAS equipped drone. (b) The survey begins with a perimeter flight to identify leaking equipment. (c) Once a leak is identified the quantification can be done using downwind and upwind mass balance flights.	126
4.8	Running the DT in parallel with the physical system, the dynamics of the gas filaments are being lofted into the air due to the cylinders on equipment pad 3 at site 1 (top). The physical observations seen from experiment and calculated with the mass balance (bottom) reflect those in the DT.	127
4.9	Example mass balance measurement for one curtain flight.	128
4.10	The concentration time series in ppm during flight curtain 1, where $n_p = 40$ puffs/sec and the symbols τ_{gi} , τ_{pi} , and τ_{lag} represent the i -th gap length, i -th pulse length, and the time lag between the simulation and experimental data [1].	129
4.11	Pattern search method convergence plot for behavior matching of flight 1 given the cost function, $J(\theta)$. The algorithm converged as mesh size tolerance was reached to ≈ 5.813	131
4.12	Response of the single parameter Mittag-Leffler function for $\mathcal{E}_\alpha(-t^\alpha)$ as $\alpha \in [0, 1]$	133
4.13	Here the turbulence intensity, $T_i = \sigma_u/\bar{u}$, of the wind is plotted against the normalized flow rate, which trends toward $\hat{Q}_*/Q = 1$ as $T_i \rightarrow 0$. A boxplot of the source estimation for each flight (set of 10 simulations), given a spatial interpolation method, is shown for each T_i condition observed from the experimental data. A smaller boxplot (shown within each boxplot) contains tails that represent the uncertainty. We can observe 1) several flights agree with experimental data 2) majority of flights underestimate the flow rate in short due to increased T_i and 3) the remaining flights that disagree are likely not capturing vertical plume behavior.	134
4.14	The detection concentrations from the simulation and experimental sensor are compared in a probabilistic histogram. While the distributions show that the majority of indications are near background levels, they also agree with the less likely larger concentrations detected in the experiment.	134
4.15	The detection altitude versus measured concentrations from the simulation and experimental sensor are compared. The distributions match well with experiment and the probability of detection declines rapidly as the altitude increases from 10-15 m.	135

4.16	Given an unknown source location (x_s, y_s) the DT is behavior matched to minimize the error e_d in the source location estimation $(\hat{x}_s, \hat{y}_s)$ given a sensor distance described by d_s	136
4.17	The results from a single snapshot of data using IDW is shown on the left and the first four spatially interpolated DMD modes are shown to the right (starting with the top row). The low frequency behavior creates a hotspot that varies less with time, allowing for behavior matching optimization.	138
4.18	The time evolution of the DT behavior matching is shown for various time steps. The DT is shown in blue and the unknown physical system is shown in red.	138
4.19	Given both meandering and no meandering wind conditions, the behavior matching method converges around the true source location. Even though the additional turbulence from the meandering conditions makes the convergence noisy, the estimation of the source location seems to improve.	139
4.20	Preliminary results in source localization error are shown as a function of sensor density. A single source location, initial guess, and mean wind direction are used to vary meandering amplitude $1 \leq G \leq 10$ (where 10 represents larger meandering) for a fixed natural frequency $\omega_n = 0.2$ and damping ratio $\zeta = 0.1$ ($a = \omega_n^2$ and $b = 2\omega_n\zeta$). The dashed line is the distance between adjacent sensors and shows the method provides some improvements in resolution.	140
4.21	This diagram depicts the simulation environment setup and framework implementation flowchart.	141
4.22	These plots show the current position / density reconstruction (left), the ground truth (middle), and the entropy (right) of the DE-CVT (top row) and D-CVT (bottom row) strategies for flux reconstruction.	144
4.23	Given the 2D static case, the estimation error is reduced faster with the DE-CVT. This is desirable behavior since the field may change over time in practical applications.	144
4.24	(left) 3D fugitive gas emission with deployed UAVs performing DE-CVT algorithm (middle-left) Ordinary kriging estimation given current concentration measurements for the UAV's at timestep k (middle-right) The first DMD mode capturing the low frequency behavior of the kriging results and (right) the entropy map computed from the kriging estimation variance.	145
4.25	A snapshot of the instantaneous source rate (red-dotted), the continuous mass rate using 10 s moving average (blue), and true source rate (black).	146
4.26	The simulation sample domain.	146
4.27	This image illustrates the different path planning strategies to sample the plume.	147

4.28	In these boxplots, the estimation performance results for static plumes different plume widths are explored for (a) lawn mower & (b) spiral paths	148
4.29	In these sets of plots, the performance of the NGI (red) and the mod-NGI (blue) at different amplitudes of meandering are shown and the different frequency responses of the algorithms are shown in each row. It can be observed that both algorithms do not perform well given dynamic (e.g. meandering) conditions. The most variability can be seen in the plume width estimates.	149
4.30	Given a moving source these six plots show the effect of the multi-UAS CVT-mod-NGI and single ESC-mod-NGI based algorithms. It can be observed in plots (a)-(b) that as the movement frequency of the source increases the ability of the CVT-mod-NGI algorithm to estimate and track the source diminishes. This behavior is much like the response of a time series signal given a low pass filter. Conversely, as the window of points used to compute the mod-NGI approach is reduced, much like we would expect, the ability to track and estimate the parameters increases. Although the performance in (c) is improved, the estimations become more noisy, and thus there is a trade-off between tracking performance and stable estimations. To illustrate why multi-sUAS systems may be better, an ESC-mod-NGI algorithm is shown in (d), where the ability to estimate parameters is not as robust to meandering and exhibits more difficulty with hyperparameter tuning.	150
4.31	Comparisons between the results from Shah et al.'s NGI, our NGI, the mod-NGI, and the true controlled release flux using the controlled release survey data from [11].	151
4.32	To validate the model the sUAS samples along a fixed path downwind of the source. The measured signal is shown in the top right.	152
4.33	Subsections of the METEC location used for the case studies: No Obstacles (green), Small Obstacles (red), and Large Obstacles (Magenta). The colored points correspond to the plume starting locations.	152
4.34	Results for the Mass Balance and NGI methods averaged over all of the meandering conditions and plume repetitions for each case study and distance. The solid lines indicate the means at the points while the vertical distance from the solid line to the edge of the shaded region represents the standard deviation at that point.	156
4.35	Results for the DE-CVT method for a swarm size of 3 UASs and 9 UASs averaged over all of the meandering conditions and plume repetitions for each case study and distance. The solid lines indicate the means at the points while the vertical distance from the solid line to the edge of the shaded region represents the standard deviation at that point.	157

4.36	Results for the Mod-NGI method for a swarm size of 3 UASs and 9 UASs averaged over all of the meandering conditions and plume repetitions for each case study and distance. The solid lines indicate the means at the points while the vertical distance from the solid line to the edge of the shaded region represents the standard deviation at that point.	158
5.1	The minimum and maximum condition numbers of the Fisher information matrix is bounded by the empirical observability Gramian scaled with the condition number of the measurement noise covariance matrix (see [12] for more details).	162
5.2	A diagram depicting the interconnection between the three major tasks of digital twin enabled smart sensor placement framework.	164
5.3	The domain Ω contains the source ρ , digital twin ρ_{DT} , and the Ω_p plane (functions of $r(t)$ and $\varphi(t)$). The projection from the Lagrangian states to Eulerian states onto $\underline{\Omega}$ are shown by $\hat{\rho}_{DT}$. Inside $\underline{\Omega}$ there may be regions that $r(t)$ cannot go, illustrated in red shaded color. The CVT continuous coverage controller is shown on the right. The combination of the two desired setpoints are used to dynamically steer the centroids in a orbit around the stable point (located in the center of Ω_p). . . .	167
5.4	The phase portrait for different parameters (outlined in 5.28) can be seen here, where the velocity field is shown to divergent/convergent, stretched, and shifted.	169
5.5	The MOABS/DT simulation is first initialized in (Stage 1) by letting the physical and DT plumes develop past the point of measurements, then (Stage 2) initializes the continuous CVT and commands the Ω_p plane to cross the plume. Once a detection is made the sensor placement and steering is activated and begins using the DT projection onto $\underline{\Omega}$ for the empirical observability Gramian computation.	171
5.6	The sensor placement and steering trajectory for each cost function are shown here.	173
5.7	The emission quantification using the mod-NGI method is shown against different cost functions. The best performing metrics have smaller uncertainty and better accuracy, such as with $\det(W^{-1})$	174

List of Tables

3.1	Figures of merit for defining complexity of an estimation method . . .	86
3.2	Summary of methods and their assumptions, operational details, complexity, cost, average precision, and average accuracy are generalized over implementations found in the literature. The average precision and accuracy are given as unitless values, normalized by the true source rate (estimated source rate). (measured using ¹ fixed, ² foot, ³ vehicle, ⁴ manned aircraft, or ⁵ UAV; cost with \$ ≤\$10,000, \$\$ ≤\$50,000, \$\$\$ ≤\$100,000, \$\$\$\$ ≥\$100,000; (·) represents precision normalized on estimated source rate average complexity.	88
3.3	Summary of method advantages and disadvantages along with fields of application	91
3.4	Testing matrix for February 14-15 Merced Vernal Pools controlled release test experiments. The flow rate is given in standard cubic feet per hour (SCFH).	94
3.5	Number of passes needed at Pasquill-Gifford atmospheric stability class and above ground level (AGL) altitude to gain 95% due diligence confidence that a 5 SCFH leak was discovered.	95
3.6	Test Matrix for PG&E Livermore Testing Facility Field Experiments.	98
3.7	Metrics of blind test leaks. The time represents the time to localize all the leaks and the leak size is given in SCFH.	99
3.8	Summary of the performance of the participating technologies in the first phase of the Alberta Methane Field Challenge (AMFC)[13]. Days refers to the number of participation days, rate refers to the number of sites per day, and time is the survey in minutes per site.	110
3.9	The quantification results are compared to tower estimates, which show that the sUAS-based estimates are lower than the tower measurements during the time of the flights. However, the estimates are within a reason considering the variability of the wind and the difficulty in measuring small emission rates.	111
4.1	(top) Definitions of Digital Twin [14] and (bottom) recent Digital Twin platforms and applications.	117

4.2	Results from mass balance method for 16 curtain flights are shown below along with the average experimental wind speed, $\overline{u}_e$ and the $\cos \beta_e$ between the normal vector of the flux plane. The flow rates (experimental Q_* and simulation $\hat{Q}_*$) are in SCFH, the subscripts denote the spatial interpolation method, and the superscripts are the associated parameters, where K is Kriging, IDW is inverse distance weighting, and KDM is Kernel DM+V.	135
4.3	An overview of the simulation testing scenarios.	153
5.1	Observability metrics that can be used as a minimizing cost function for evaluating a system.	163

Abbreviations

The following abbreviations are used in this dissertation:

ADE	Advection diffusion equation
AJAX	Alpha Jet Atmospheric eXperiment
ALD	Advance Leak Detection
AMFC	Alberta Methane Field Challenge
AMOG	AutoMObile greenhouse Gas
ANOVA	Analysis of variance
ARPA-E	Advanced Research Projects Agency-Energy
AVIRIS-NG	Next generation Airborne visible/infrared imaging spectrometer
bLS	Backward Lagrangian Stochastic dispersion technique
bs-TDLAS	Back scatter tunable diode laser absorption spectrometer
CALMIM	California Landfill Methane Inventory Model
CARB	California Air Resources Board
CASIE	Characterization of Arctic Sea Ice Experiment
CFP	Cylindrical Flux Plane
CRDS	Cavity ring-down spectrometer
DiAL	Differential Absorption LiDAR Method
DT	Digital Twin
EC	Eddy Covariance
EDF	Environmental Defense Fund
EPA	Environmental Protection agency
EPA FLIGHT	EPA Facility Level Informationon Greenhouse gases Tool
ES	Environmental sensing
FEAST	Fugitive Emissions Abatement Simulation Toolkit
FID	Flame Ionization Detector
FLEXPART-WRF	FLEXible PARTicle-Weather Research and Forecasting
FLOM	Fractional lower order moments
FPE	Fokker–Planck equation
FTIR	Fourier Transform Infrared
GADEN	A 3D Gas Dispersion Simulator for Mobile Robot Olfaction in Realistic Environments
GDT	Gauss Divergence Theorem
GDM	Gas distribution mapping
GDMF	Gradient divergence mass flux
GHG	Greenhouse gas
GLM	General linear model
GLM-VFP	General linear model vertical flux plane
GML	Gas Mapping LiDAR™ technology
GSL	Gas source localization
QOGI	Quantitative optical gas imaging
ICOS	Integrated cavity output spectroscopy

IDW	Inverse distance weighted
IMAP-DOAS	Iterative maximum a posteriori differential optical absorption spectroscopy
LDAQ	Leak detection and quantification
LDAR	Leak detection and repair
LES	Large eddy simulation
LiDAR	Light detection and ranging
LGR	Los Gatos Research
LWIR	Long-wave infrared
MCMC	Markov Chain Monte Carlo
MEMS	Micro electrical mechanical systems
METEC	Methane Emission Testing and Evaluation Center
MGGA	Micro greenhouse gas analyzer
MOX, CMOS,MOS	Ceramic Metal oxide sensor
MMC	Stanford EDF Mobile Monitoring Challenge
MMD	Micrometeorological mass difference method
MONITOR	Methane Observation Networks with Innovative Technology to Obtain Reductions
MUST	Mock Urban Test Setting
MWIR	Mid-wave infrared
NDIR	Non-dispersive Infrared
NZMB	Non-zero minimum bootstrap
NGI	Near-field Gaussian plume inversion
NIR	Near infrared
OA-ICOS	Off axis integrated cavity output spectrometer
OGI	Optical gas imaging
OPLS	Open path laser spectrometer
OTM	Other test method
OU	Ornstein-Uhlenbeck
PG	Pasquill-Gifford
PID	Photo-ionization detector
PI-VFP	Path Integrated Vertical Flux Plane
PMT	Pollution mapping tool
PSD	Power spectrum density
PSG	Point source Gaussian
PSG-CS	Conditionally sampled point source Gaussian
PSG-RB	Recursive Bayesian point source Gaussian
PSG-SBM	PSG sequential Bayesian MCMC
QOGI	Quantitative optical gas imaging
RANS	Reynolds averaged Navier Stokes
RB-LSI	Recursive Bayesian least squares inverse
RGB	Red green blue
RMLD	Remote methane leak detector
SDP	Source determination problem

SEBASS	Spatially-Enhanced Broadband Array Spectrograph System
SEM	Surface emission monitoring
SESC	Stochastic extremum seeking control
SLAM	Simultaneous localization and mapping
STE	Source term estimation
sUAS	small Unmanned aircraft system
SWIR	Short-wave infrared
TDLAS	Tunable diode laser absorption spectroscopy
	Tracer dispersion method,
TDM, TCM, ATM	Tracer correlation (or dilution) method,
	Atmospheric tracer method
TIR	Thermal infrared
TSEB	Two source energy balance
UA	Ultrasonic Anemometer
UGGA	LGR Ultraportable GHG analyzer
VCSEL	Vertical cavity surface emitting laser
VFP	Vertical Flux Plane
VRPM	Vertical Radial Plume Mapping Method
WRF	Weather Research and Forecasting Model

Acknowledgment

First and foremost I would like to express my deepest thanks for my family and my wife, without their support I would not be able to make this happen. Secondly, I would like to give recognition to my fearless advisor, Professor YangQuan Chen, who pushes, not just myself, but everyone around him to pursue research excellence. I believe a great deal of my growth and successes are attributed to his mentorship and guidance.

It was through the support of the Mechatronics, Embedded Systems, and Automation Lab that I was able to accomplish so much. I would like to also show my gratitude to all of the MESA Lab members (former, current and future) for all of the collaborations. This includes but not limited to: Gregorio Nunez, Garrett John, Alexis Bonnin, Matthieu Loubens, Tiebiao Zhao, Gouxian Zhang, Yuquan Chen, Haoyu Niu, Jairo Viola, Carlos Rodriguez, Justus Nwoke, Di An, and Rafal Krzysiak.

Throughout my PhD career I was fortunate enough to work with brilliant individuals from outside my home institution and I would like to acknowledge some of them here: Lance Christensen, Francios Rongere, Kristen Manies, Michael Whiticar, Brad Billwiller, and Eric Saczuk.

The text of several sections in Chapter 2 of this dissertation are reprints of articles titled: “On Optimal Tempered Lévy Flight Foraging”; “Individual and Collective Foraging in Autonomous Search Agents with Human Intervention”; “Swarm Robotic Source Seeking with Fractional Fluxotaxis”; and “A More Optimal Stochastic Extremum Seeking Control Using Fractional Dithering for a Class of Smooth Convex Functions”. The text of several sections in Chapter 3 of this dissertation are reprints of articles titled: “Advanced Leak Detection and Quantification of Methane Emissions Using sUAS”; and “Characterization of Ground-to-air Emissions with sUAS Using a Digital Twin Framework”. The text of several sections in Chapter 4 of this dissertation are reprints of articles titled: “A Digital Twin Framework for Environmental Sensing with sUAS”; “MOABS/DT: Methane Odor Abatement Simulator with Digital Twins”; “Digital Twin Behavior Matching of Gas Plumes Using a Fixed Sensor Mesh and Dynamic Mode Decomposition”; “Multi-UAV Method for Continuous Source Rate Estimation of Fugitive Gas Emissions From a Point Source”; “A Modified Near-Field Gaussian Plume Inversion Method Using Multi-sUAS for Emission Quantification”; and “Single and Multi-sUAS Based Emission Quantification Performance Assessment Using MOABS/DT: A Simulation Case Study”. I would also like to acknowledge the financial support I have received over the years from the National Science Foundation (NSF) National Research Traineeship (NRT) Grant No. DGE-1633722, as well as computing time on the Multi-Environment Computer for Exploration and Discovery (MERCED) cluster at UC Merced, which was funded by NSF Grant No. ACI-1429783.

Curriculum Vita

Derek Hollenbeck
Mechanical Engineering
University of California, Merced
dhollenbeck@ucmerced.edu

Education

- **University of California, Merced.** Merced, CA, USA. (2016 – 2023).
Ph.D. in Mechanical Engineering.
- **University of California, Merced.** Merced, CA, USA. (2013 – 2016).
B.Sc. in Mechanical Engineering.
- **Merced College.** Merced, CA, USA. (2012 – 2013).
Mathematics and Physics emphasis
- **Lassen Community College.** Susanville, CA, USA. (2010 – 2012).
A.S. Mathematics and Physical Science

Publications

28. **Hollenbeck, D.**, Zheng, K., Zulevic, D., & Chen, Y. (2023). Swarm Robotic Source Seeking with Fractional Fluxotaxis. *in Proceedings of 2023 International Conference on Fractional Derivatives and Applications*. Ajman, Jordan.
27. **Hollenbeck, D.**, An, D., Krzysiak, R., & Chen, Y. (2023). Towards Cognitive Battery Monitoring on Hybrid VTOL Fixed-Wing sUAS with Maximized Safe Endurance. *In 2023 International Conference on Control, Mechatronics, and Automation (ICCMA)*. (Planned)
26. An, D., Krzysiak, R., **Hollenbeck, D.**, & Chen, Y. (2023). Battery-health-aware UAV Mission Planning Using a Cognitive Battery Management System. *In 2023 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE. (Accepted)

25. An, D., Krzysiak, R., **Hollenbeck, D.**, & Chen, Y. (2023). Long Endurance Site-Specific Management of Biochar Applications Using Unmanned Aircraft Vehicle and Unmanned Ground Vehicle. *In 2023 International Federation on Automatic Control (IFAC)* (Accepted)
24. **Hollenbeck, D.**, Zulevic, D., & Chen, Y. (2022). Single and Multi-sUAS Based Emission Quantification Performance Assessment Using MOABS/DT: A Simulation Case Study. *In 2022 18th IEEE/ASME International Conference on Mechatronics and Embedded Systems and Applications.*
23. An, D., **Hollenbeck, D.**, Cao, K., & Chen, Y. (2022). Soil Methane Emission Suppression Control Using Unmanned Aircraft Vehicle Swarm Application of Biochar Mulch - A Simulation Study. *Journal of Information and Intelligence*
22. An, D., Liu, Z., **Hollenbeck, D.**, & Chen, Y. (2022). A Greenhouse Gas Proximity Sensing Using Chemiresistive Strip and Miniaturized Radar Array. *In 2022 10th International Conference on Control, Mechatronics and Automation.*
21. **Hollenbeck, D.**, & Chen, Y. (2022). A Digital Twin Framework For Environmental Sensing with sUAS. *Journal of Intelligent & Robotic Systems* 105 (1), 1-15.
20. An, D., **Hollenbeck, D.**, Gao, S., & Chen, Y. (2022). A Field Study of Soil Biochar Treatment Response Using Small Unmanned Aerial Systems (sUAS). *In 2022 International Conference on Unmanned Aircraft Systems (ICUAS). IEEE.*
19. **Hollenbeck, D.**, Zulevic, D., & Chen, Y. (2022). A Modified Near-Field Gaussian Plume Inversion Method Using Multi-sUAS for Emission Quantification. *In 2022 International Conference on Unmanned Aircraft Systems (ICUAS). IEEE.*
18. Viola, J., Rodriguez, C., **Hollenbeck, D.**, and Chen, YQ. (2022). *Outliers in Control Engineering*, 237. (Book Chapter)
17. **Hollenbeck, D.**, & Chen, Y. (2021). Digital Twin Behavior Matching of Gas Plumes Using a Fixed Sensor Mesh and Dynamic Mode Decomposition. *In ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers Digital Collection.*
16. **Hollenbeck, D.**, Zulevic, D., & Chen, Y. (2021). Advanced Leak Detection and Quantification of Methane Emissions Using sUAS. *Drones*, 5(4), 117.
15. **Hollenbeck, D.**, Zulevic, D., & Chen, Y. (2021). MOABS/DT: Methane Odor Abatement Simulator with Digital Twins. *In 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI) (pp. 378-381). IEEE.*

14. **Hollenbeck, D.**, & Chen, Y. (2021). Multi-UAV Method for Continuous Source Rate Estimation of Fugitive Gas Emissions From a Point Source. *In 2021 International Conference on Unmanned Aircraft Systems (ICUAS) (pp. 1308-1313). IEEE.*
13. Schloesser, D. S., **Hollenbeck, D.**, & Kello, C. T. (2021). Individual and Collective Foraging in Autonomous Search Agents With Human Intervention. *Scientific Reports, 11(1), 1-13.*
12. Viola, J., **Hollenbeck, D.**, Rodriguez, C., & Chen, Y. (2021). Fractional-Order Stochastic Extremum Seeking Control with Dithering Noise for Plasma Impedance Matching. *In proceedings of CCTA.*
11. Fatehiboroujeni, S., **Hollenbeck, D.**, Mishra, A., & Goyal, S. (2021) Post-buckling Stability of a Cantilever Beam with Cubic Non-Linearity in Constitutive Laws. *In ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers Digital Collection. (Accepted)*
10. **Hollenbeck, D.**, Manies, K., Chen, Y., Baldocchi, D., Euskirchen, E. S., & Christensen, L. (2020). Evaluating a UAV-based Mobile Sensing System Designed to Quantify Ecosystem-based Methane. *In AGU Fall Meeting Abstracts (Vol. 2020, pp. A115-0007).*
9. Niu, H., **Hollenbeck, D.**, Zhao, T., Wang, D., & Chen, Y. (2020). Evapotranspiration Estimation with Small UAVs in Precision Agriculture. *Sensors, 20(22), 6427.*
8. **Hollenbeck, D.**, & Chen, Y. (2020). Characterization of Ground-to-air Emissions with sUAS Using a Digital Twin Framework. *In 2020 International Conference on Unmanned Aircraft Systems (ICUAS) (pp. 1162-1166). IEEE.*
7. **Hollenbeck, D.**, & Chen, Y. (2020). A More Optimal Stochastic Extremum Seeking Control Using Fractional Dithering For A Class of Smooth Convex Functions. *IFAC-PapersOnLine, 53(2), 3737-3742.*
6. Schloesser, D., **Hollenbeck, D.**, & Kello, C. T. (2020). Social Foraging in Groups of Search Agents with Human Intervention. *In CogSci.*
5. Whiticar, M., **Hollenbeck, D.**, Billwiller, B., Salas, C., & Christensen, L. (2019). Application of the BC GHGMapperTM Platform for the Alberta Methane Field Challenge (AMFC). Geoscience BC Summary of Activities, 2020-02.
4. **Hollenbeck, D.**, Dahabra, M., Christensen, L. E., & Chen, Y. (2019). Data Quality Aware Flight Mission Design for Fugitive Methane Sniffing Using Fixed Wing sUAS. *In 2019 International Conference on Unmanned Aircraft Systems (ICUAS) (pp. 813-818). IEEE.*

3. **Hollenbeck, D.**, Oyama, M., Garcia, A., & Chen, Y. (2019). Pitch and Roll Effects of On-board Wind Measurements Using sUAS. *In 2019 International Conference on Unmanned Aircraft Systems (ICUAS) (pp. 1249-1254). IEEE.*
2. Chen, Y., **Hollenbeck, D.**, Wang, Y., & Chen, Y. (2018). On Optimal Tempered Lévy Flight Foraging. *Frontiers in Physics*, 6, 111.
1. **Hollenbeck, D.**, Nunez, G., Christensen, L. E., & Chen, Y. (2018). Wind Measurement and Estimation with Small Unmanned Aerial Systems (sUAS) Using On-board Mini Ultrasonic Anemometers. *In 2018 International Conference on Unmanned Aircraft Systems (ICUAS) (pp. 285-292). IEEE.*

Awards, Fellowships, and Grants

- CITRIS Aviation Prize (inaugural, Fall 2021)
- UC Merced ME Bobcat Fellowship (Fall 2021)
- UC Merced ME Bobcat Fellowship (Sum 2021)
- NSF NRT - Extended Intelligent Adaptive Systems Fellowship (Fall 2020 - Sum 2021)
- NSF NRT - Extended Intelligent Adaptive Systems Fellowship (Fall 2019)
- NSF NRT - Extended Intelligent Adaptive Systems Fellowship (Spring 2019 - Sum 2019)
- NSF NRT - Intelligent Adaptive Systems Fellowship (Spring 2017 - Sum 2018)
- NSF NRT - Interdisciplinary Computational Graduate Education (Spring 2017)
- UC Merced Outstanding Student Award (undergraduate, Spring 2016)

Presentations

- ICUAS 2022 Sum 2022 (title: A Modified Near-Field Gaussian Plume Inversion Method Using Multi-sUAS for Emission Quantification)
- CITRIS Research Exchange Spring 2022 (title: Long Endurance Edge-AI Platform For Research Opportunities and Data Gathering)
- PhD Dissertation Qualifying Exam Spring 2022 (title: Digital Twin Enabled Collective Sensing and Steering for Source Determination Problems)
- MESA Lab Research Group Meeting Fall 2021 (title: The empirical gramian and why it is important)

- MSNDC 2021 Sum 2021 (title: Post-buckling Stability of a Cantilever Beam with Cubic Non-linearity in Constitutive Laws)
- ICUAS 2021 Sum 2021 (title: Multi-UAV Method for Continuous Source Rate Estimation of Fugitive Gas Emissions from a Point Source)
- DTPI 2021 Sum 2021 (title: MOABS/DT: Methane Odor Abatement Simulator with Digital Twins)
- MESA Lab Research Group Meeting Spring 2021 (title: Swarm Robotic Source Seeking with Fractional Fluxotaxis and Seasonal Meandering)
- MESA Lab Research Group Meeting Fall 2020 (title: Physicomemetics, Chemical Plume Tracing, and Fluxotaxis (I))
- MESA Lab Research Group Meeting Fall 2020 (title: MOAB/DT: Towards Smart Swarming for Methane Emission Sensing and Understanding/Quantification)
- ICUAS 2020 Workshop Tutorial Sum 2020 (title: A digital twin framework on using drones as edge devices for environmental monitoring)
- ICUAS 2020 Workshop Tutorial Sum 2020 (title: Methane sensing drones with machine learning)
- UC Agriculture and Natural Resources (UCANR) DroneCamp Spring 2020 (title: Methane emission drone swarm sensing to better understand permafrost thawing dynamics.)
- MESA Lab Research Group Meeting Fall 2019 (title: Fractional Order Modeling of Atmospheric Dispersion)
- MESA Lab Research Group Meeting Spring 2019 (title: UAV Gas Sensing, a Digital Twin, and PET (POSIM Experimental Testbed))
- Invited Talk – Applied Math Optimization Research Group Spring 2019 (title: tbd) NSF NRT Retreat Meeting Spring 2019 (title: Collective Foraging Among Loosely Coupled Agents)
- Invited Talk – NSF NRT Fall Kickoff Meeting Fall 2018 (title: On Collective Behavior of Foraging Vultures and Individual Impact)
- NSF NRT Retreat Meeting Spring 2018 (title: Collective Foraging Among Loosely Coupled Agents)
- Invited Talk – Applied Math Optimization Research Group Fall 2017 (title: Fractional Calculus Enhanced Source Seeking of Fugitive Methane)

- Guest Lecture ME190 – Sum 2017 (topics: Communication Basics, Long Distance Drones, UTM)
- Guest Lecture ME190 – Long Distance Drones Spring 2017
- Guest Lecture ME190 – Long Distance Drones Fall 2016
- Guest Lecture ME190 – Unmanned Aerial System Traffic Management Fall 2016

Teaching

- ENGR 057 - Circuit Theory Fall 2022
- ENGR 057 - (IOR) Circuit Theory Sum 2022
- ENGR 065 – Statics and Dynamics Fall 2021
- ME 142 – Mechatronics Spring 2020
- ME 143 – Unmanned Aircraft Systems Fall 2018
- ENGR 057 – Circuit Theory Sum 2017
- ENGR 190 – Capstone Design Spring 2017
- ENGR 057 – Circuit Theory Fall 2016

Abstract

Digital Twin Enabled Collective Sensing and Steering for Source Determination Problems

by: **Derek Hollenbeck**

Mechanical Engineering

University of California, Merced. 2023.

Committee chair: Professor YangQuan Chen

Motivated by climate change and the global warming potential of methane (86 times more potent than CO_2), this dissertation focuses on the source determination problem using collective sensing and Digital Twins. Recently, Digital Twins have been developed to provide better performance assessment, fault prognosis and predict future behavior of complex systems. The term ‘collective’, refers to the group of mobile sensors that, as a whole, provide more information than a single mobile sensor can. The mitigation of methane emissions into the atmosphere is important to focus on in reducing the effects of global warming in the near term. In order to mitigate emissions, the leaks have to first be detected and assessed before they can be repaired. Many of these emissions can be modeled as a point source governed by partial differential equations (PDE), which, solutions are typically time-stepped into the future (i.e. the forward problem). In many cases, the emission plume is subject to turbulence which requires the use of turbulence models, such as large eddy simulations (LES), to compute. In both cases, the computational requirements and run-time can prevent real-time or near real-time analysis. Considering hybrid modeling approaches (e.g. deterministic and stochastic), the forward behavior matched Digital Twin model can be computed in near-real time and used for improving emission quantification methodologies as well as perform optimization (e.g. sensor placement and mobile or fixed-location sensing / actuation policy). The dissertation is broken into four main parts: the first part is on source seeking based optimization using random search, collective foraging, Fluxotaxis, and Extremum Seeking Control; the second part is on the application of leak detection and quantification with sUAS (including: sensors, platforms, and methods) as well as controlled release and real world field campaigns; the third part is on Digital Twins (POSIM and MOABS/DT) and how to use them for environmental sensing, method development, and performance evaluation case studies; the last part is on the sensor placement problem and how the observability Gramian combined with Digital Twins, can be used for smarter collective sensing and steering.

Chapter 1

Introduction

In this chapter, the research motivations behind this work are introduced, then an overview of the literature and the research challenges are discussed, before concluding with the dissertation objectives, contributions and outline.

1.1 Research Motivations

The research motivations behind this dissertation work, in general, aims to answer the most basic of research questions: so what? and who cares? The “so what” component of this research, is motivated by the changing climate, which is due in part to emitting greenhouse gases (GHG) into the atmosphere (e.g. carbon dioxide or CO_2 , and methane CH_4). In particular, we are mainly focusing on methane, as it is a powerful GHG that has a global warming potential (GWP) 86 times greater than that of CO_2 in a 20 year time period. Why is this important to focus on? As the global temperature increases, local climates start to see greater frequencies of abnormal events. In some cases, such as with some natural ecosystems, the increased temperature can ultimately increase the release of methane back into the atmosphere, creating a positive feedback loop. Natural ecosystems such as with thawing of Alaskan permafrost bogs, which was initially thought to be a carbon sink (e.g. net flux of CH_4 and CO_2 goes into the ground), may no longer be the case.

One way to combat reverse the effects of climate change, is to look at mitigating methane emissions. The first reason is that methane has a residence time in the atmosphere of around 10 years, as compared to 100 years with carbon dioxide. Therefore, by reducing methane today, there will be greater impact in the global warming the near term [15]. The methane emissions can be mitigated through upgrading and repairing infrastructure as well as changing local federal and state policies to add incentives for reduction in emissions. This pathway requires the use of leak detection and repair (LDAR) programs to be put in place, such as those proposed by the Environmental Protection Agency (EPA) which utilizes Method 21 for the oil & gas industries.

As Peter Drucker once said, “You can’t manage what you can’t measure”, puts in

place the importance of being able to mitigate, by being able to measure simultaneously. Ultimately, to mitigate methane emissions, several technical problems need to be solved first, including: 1). the sensor technology used to detect and measure the quantity of interest (i.e. methane, in this case); 2). where to place the sensor (e.g. using mobile sensing along a specified path or with a fixed sensor position), such that, the most amount of information is attained; and 3). what to do with the data, and how to make sense of it (e.g. quantification, mapping, etc.). In this dissertation we will mainly focus on the latter two.

The 2020 overview of GHG emissions, according to the EPA, show that methane contributes 11% to the total emissions. The top three methane emissions categories include: oil & gas 32%, enteric fermentation 27%, and landfills 17%. With the recent submitting of EPA's November 11, 2022 supplemental proposal which outlines changes in the leak monitoring requirements in the oil & gas industry based optical gas imaging monitoring, including minimum screening frequency and minimum detection thresholds. The addition of alternative technology approval process, expands the possibilities of meeting methane mitigation goals. The use of software like FEAST [16] have been looked at to make the equivalent comparison between alternative methods and the approved OGI method [17]. This advancement provides a pathway for new methodologies and technologies to be implemented in the mitigation process and is another source of motivation in this dissertation to develop new methods of quantification to improve on speed, accuracy, and cost of surveys.

With the policy advancements in conducting LDAR in oil & gas industries, the same policy advancements need to happen in other industries as well (e.g. landfills, dairy farms, waste water treatment plants, etc.), further promoting methane mitigation goals, as well as providing a pathway to develop new techniques for effectively measuring these sites with their unique characteristics.

1.2 Literature Overview and Challenges

Within the scope of this dissertation there are several components and sub-components contained within, such as: source determination problems, source seeking, leak detection & quantification (LDAQ), and Digital Twins (DT).

1.2.1 Source Determination Problems

The source determination problem is made up of three general categories, namely, source localization, source rate estimation, and source term estimation (or parameter estimation). The first category is source localization, which generally includes two types of approaches: finding the source location through inference (does not require the sensor to travel to the source location explicitly) and source seeking. The second, is source rate estimation, which utilizes mobile sensing data through optimization, statistical inference, and physical quantities (such as flux) to approximate the source

rate. The last category is parameter estimation, which aims to estimate the parameters associated with model of the system (e.g. source rate, location, dispersion, etc.). It should be noted that source rate estimation, and emission (or leak) quantification are synonymous in this work and are a subset of the parameter estimation problem.

1.2.2 Leak Detection and Quantification

An applied form of the source determination problem is leak detection & quantification (LDAQ). In order to mitigate emissions, leaks have to first be detected and quantified. The quantification can be done in many ways but results are often presented as a mass flow rate or flux estimate. These estimates can be made, typically, using methods based on static measurements, on foot, by vehicle, manned aircraft, and by satellite. Static measurements consist of: (1) Eddy Covariance (EC) towers: A footprint modeling technique that looks at the turbulent exchange with the environment and utilizes meteorological conditions with precision concentration measurements to estimate the flux, (2) Chambers (autochambers): An enclosed chamber is placed over a target piece of land and is sampled with a syringe occasionally (to be analyzed at a later time, typically with gas chromatography) or dynamically sampled within a closed loop (such as GASMET's Fourier Transform Infrared (FTIR) analyzer [18]).

Measurements made on foot consist of handheld sensors that are used with survey equipment and the surface. Surface emission monitoring (SEM) is typically a routine operation for landfills, done quarterly, to maintain compliance with local regulations to account for emissions lost from gas control systems. SEM is a point-based scanning technique that can take on the order of a week or so to complete. The concentrations are measured with devices, such as the flame ionization detector (FID) (regulated by EPA's guidance and Method 21), and are integrated along sub divided grids looking for elevated methane levels (greater than 500 ppmv) [19].

Landfill emissions are generally calculated using inventory-type estimates. Measurements by vehicle consist of methods, such as the Tracer Correlation Method (TCM) and the environmental protection agency (EPA)'s other test method 33A (OTM33A) [5]. In [20], they explored vehicle-based advanced leak detection (ALD) with a cavity ring-down spectrometer (CRDS) from Picarro and determined that five to eight drives will capture a majority of leaks ($> 90\%$) as well as indicate detection limitations (such as wind and soil conditions and variations in methane enhancements making quantifying emissions difficult).

In [21], a vehicle mounted CRDS (Picarro G2301 and G4302) used empirical formulation to turn elevated concentration levels (or leak indications), C , to emission rates, Q (based on the work of [22], $\ln C = -0.988 + 0.817 \ln Q$) and used Gaussian plume model to quantify emissions from site-level emissions in Utrecht and Hamburg, Germany. In work by [23], 6650 sites were evaluated using inventory and inverse point source Gaussian measurements, and they found that the methane inventory was underestimated by a factor of 1.5.

A series of campaigns were carried out utilizing TCM and downwind mobile measurements to explore the accuracy of different TCM approaches as well as compare CRDS with FTIR instrumentation in multiple source separation [24]. Measurements from manned aircraft have been done using FID, mounted on both fixed wing (Piper (Vero Beach, FL, USA) Seneca or Piper Navajo twin engine) and helicopter (Bell (Fort Worth, TX, USA) 206 Long Ranger) to detect liquid hydrocarbons from pipelines [25]. In [26; 27], they utilize the next generation Airborne Visible/Infrared Imaging Spectrometer (AVIRIS-NG) to retrieve methane, carbon dioxide, and water vapor.

In [28], AVIRIS-NG was used to generate the VIsTA-CA geospatial dataset to provide a comparison to the attribution of sources with California Air Resources Board (CARB) Pollution Mapping Tool (CARB PMT) and the U.S. Environmental Protection Agency (EPA) Facility Level Information on Greenhouse gases Tool (EPA FLIGHT). In [29], they investigate the uncertainty for estimating urban fluxes by an aircraft-based mass balance approach. They assess the sensitivity of the estimated city-wide CO_2 and CH_4 fluxes for several flight experiments, including the regional background concentration, depth of the convective boundary layer, magnitude of the wind speed, and type of interpolation technique.

In [30], they utilized a Sky Arrow Environmental Research Aircraft to measure emissions from multiple landfills and combined steady state Gaussian models to distinguish the emission coefficients for each individual site. A Bridger Photonics Gas Mapping LiDAR (GML) system was deployed on a Cessna 172 and blindly evaluated where the detection limits were as low as 1 kg/h depending on the wind conditions [31]. This method was also introduced to the Fugitive Emissions Abatement Simulation Toolkit (FEAST) [16] and shown to be comparable to OGI-based methods at equivalent survey frequencies for the detection and repair of emissions. Optical Gas Imaging (OGI) was explored in [32; 33] and the effectiveness was evaluated in [33]. NASA's Alpha Jet Atmospheric eXperiment (AJAX) and the AutoMOBILE greenhouse Gas (AMOG) surveyor were used to fuse airborne and ground-based data together (as part of the GOSAT-COMEX Experiment) using an anomaly approach instead of the typical mass balance approach [34]. Measurements by satellite have been explored in [35], where a ResNet-50 was trained from ESA's Sentinel-2 data and labeled with a U-Net to detect smoke plumes.

1.2.3 Digital Twins

What is a DT? Some of the earliest work of using physical twins (or hardware twins) can be seen as early as National Aeronautics and Space Administration's (NASA) Apollo program [36]. In this case there existed at least two space vehicles or rovers at once. This allowed for the continuous testing of code, maneuvers, and conditions before sending them to the deployed version. Software can also be introduced to the hardware twin, as in the 'iron bird' which combined a physical interface of a plane cockpit with an aerodynamics simulation for training purposes. The simulation was used for a visual aide as well as to compute the meteorological conditions and re-

sulting forces on the plane [36]. Some of the early reports of the current concept of DT (as digital equivalence of a physical product) showed up in a product lifecycle management course at the University of Michigan in 2003 taught by Michael Grieves. Since then a white paper was written in reference to virtual factory replication [37]. It is explained that the DT in this scenario contains three parts: (1) physical products in real space (2) virtual products in virtual space and (3) and connections that tie together the real and virtual products. The use cases for DT are: (1) conceptualization - visualization of the physical and virtual products, (2) comparison - compare physical product behavior with the virtual product to improve results and (3) collaboration - solutions found can be applied immediately to other factories (or systems).

In the literature, there are many papers dedicated to DT's modeling systems of components for applications in manufacturing, smart cities, industry 4.0, and robotics to name a few [36; 38; 39; 40; 41; 42]. However, there is a gap for environmental sensing applications such as in the source determination problem explored here. This largely has to do with the difficulties in implementing models governed by partial differential equations (PDE) that are very high dimensional.

1.2.4 Research Challenges

There are many challenges within this dissertation associated with each aspect of the mobile sensing, LDAQ, and digital twin modeling. From the sensing point of view, there are still many hurdles, such as: the cost of the sensor, given the desired sensor response & relaxation time, the stability of the sensor (is it linear? does it drift?). Low cost sensing is very desirable but also has its limits in sensitivity including slow response times, making them not suitable for many applications of detection and quantifying fugitive emissions.

Once the sensing part has been solved, the detection and quantification can be studied. This includes questions such as where to best detect? or how to trace the plume? and which quantification strategy to choose? In most cases the last case requires simulation-, optimization-, or mass balance-based approaches, which, can vary in precision, accuracy, sample distance and sample time. The required resources to perform these approaches may vary too. For example, the simulation-based approaches may require substantial time to setup and run on the super computer.

Using hybrid modeling for the forward simulation, provides possibilities for computing in near real-time. Behavior matching this forward model to the physical systems data, forms the DT. However, the behavior matching strategies require rich data and utilize optimization, regression, and parameter estimation techniques. Thus the natural question that follows is, how do we obtain information rich data for this behavior matching to be efficient and effective? In many instances, the physical systems have disturbances (e.g. turbulence, or meandering) and the sensors have noise. For ill-posed problems, the presence of noise can magnify the high frequency modes into the solution [43].

Once the DT has been behavior matched, how can we best deploy the DT to

improve existing detection and quantification methodologies as well as create new ones? This can ultimately reduce the number of field campaigns and associated costs, making DTs even more desirable.

1.3 Dissertation Objectives

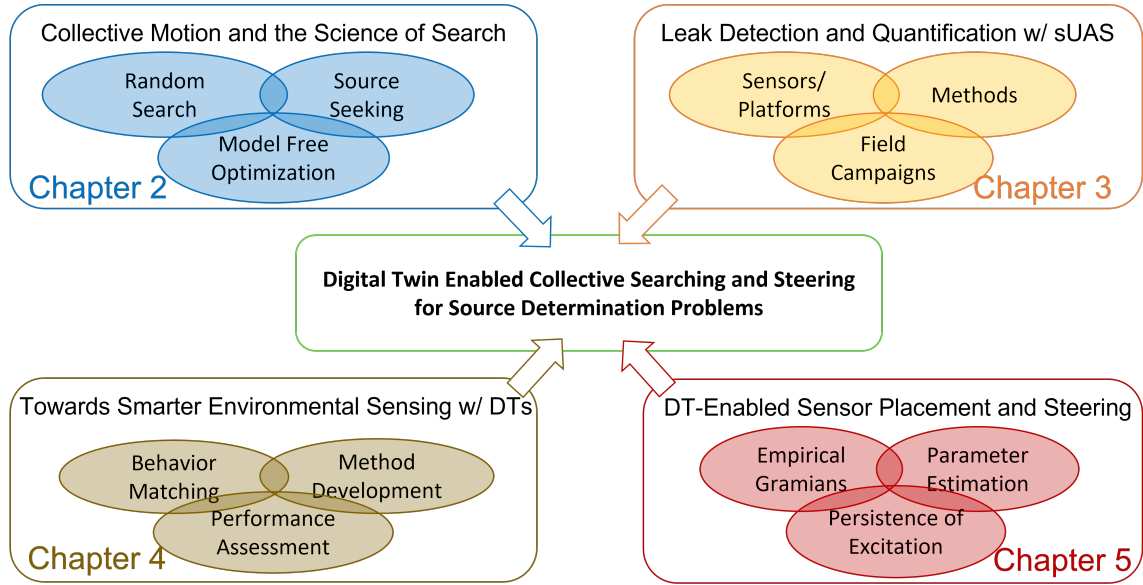


Figure 1.1: This diagram illustrates how the research topics and chapters are connected to the overall scope of the this dissertation.

The objectives of this dissertation are to investigate the source determination problems through collective sensing and DTs with applications in methane emission quantification using sUAS of point and area sources, such as natural (e.g. wetlands, vernal pools) and anthropogenic (e.g. oil & gas infrastructure, landfills, dairy farms) sources. To accomplish these objectives, we divide the dissertation into the following specific parts (also see Fig. 1.1). The key objectives are listed here as:

- Develop random search and source seeking methods for finding and tracing a chemical plume back to the source;
- Develop and validate source determination problems (via optimization, inverse, inference, and mass-balance approaches), including spatial-temporal gas distribution modeling;
- Develop swarm sensing via multi-sUAS mobile sensor networks for improving source determination problem methods;
- Develop DTs and a framework for modeling, control, and inference of parameters through behavior matching;

- Utilize data-driven reduced order modeling and control, including empirical Gramians (such as the observability Gramian) for sensor placement problem.

In Chapter 2, finding the location of a source is approached mainly from a random search [44; 45] and source seeking point of view [46]. The random search / random walk problem, is studied in the foraging context when there is little or no priors with single- and multi-agent searchers. The source seeking problem is explored using heuristic and physics-based techniques (e.g. casting, chemo-, anemo-, and fluxotaxis) as well as adaptive nonlinear control [47]. In Chapter 3, LDAQ is explored from an application point of view using sUAS. The emission source rate can be estimated using simulation, optimization, mass balance, imaging, and correlation-based methods. This dissertation focuses on optimization- and mass-balance-based techniques to estimate the source rate which includes several field campaigns with sUAS [48; 7]. In Chapter 4, method development and performance for emission quantification are explored through the Plume Odor Simulator (POSIM) DT [1] as well as the Methane Odor Abatement Simulator with DTs (MOABS/DT) based on the ARPA-E Methane Emission Technology Evaluation Center (METEC) [49; 50]. Localization using DTs are also explored through behavior matching [51]. In Chapter 5, the focus is on solving the sensor placement problem by using control theoretic-based techniques for state/parameter estimation using empirical observability Gramians.

1.4 Dissertation Contributions

The main contributions of this dissertation are in connecting the four major Chapter topics, around the source determination problem, to provide a pathway for improving existing LDAQ strategies as well as allow for the development of new methodologies. Through these contributions, a list of research questions were posed and explored, such as: (Chapter 2) What is the optimal way to randomly search for a source? Can fractional operators be used to improve existing source seeking approaches? (Chapter 3) What is the current state of the art and best way to quantify emission sources using sUAS-based approaches? (Chapter 4) Can DTs be used to analyze existing methodologies, and provide performance assessments for new methodologies? Can swarm sensing be used to produce more accurate, precise, and faster source emission estimates? (Chapter 5) Can optimal sensor placements be used to better inform and estimate source emission rates? What methodology can be used to improve information gain, making sensing smarter and more adaptive (i.e. intelligent)? Can control-theoretic-based observability be used to inform sensor placement while also improve the accuracy of source emission estimates?

The following contributions can be obtained from this dissertation:

- Fractional calculus based techniques were developed and used in evaluating the efficiency of random search and source seeking applications.
- Application review of traditional and advanced LDAQ techniques and methods

with a focus on sUAS. Including, multiple field campaigns in controlled release and real world scenarios.

- A Digital Twin framework for environmental sensing that can be used as a performance assessment tool in method development. Including, development of flux plane quantification methods using optimization techniques (e.g. Modified NGI).
- Development of sensor placement / deployment methods for improved quantification using control-theoretic observability concept. Including data-driven DT enabled smart sensor placement and steering for LDAQ.

1.5 Dissertation Outline

The dissertation is outlined as follows: Chapter 2 presents random search and source seeking optimization; Chapter 3 presents the application of sUAS-based LDAQ; Chapter 4 presents the environmental sensing with DT framework, applications, and new LDAQ methodology; Chapter 5 presents the sensor placement problem, the empirical gramian, and the DT-enabled smart sensor placement and steering framework; and Chapter 6, will conclude the dissertation.

Chapter 2

Collective Motion and the Science of Search

Search can be defined in several ways, such as, random search (e.g. with no previous or current information known), a priori random search (with some prior information known), or even statistical-based search (e.g. Bayesian approaches). Search can also be thought of as an optimization given some unknown function. Therefore the main focus of this chapter will be to explore several case studies aligned to the topics of random search (using single [44] and collective foraging [45] strategies), source seeking (swarm-based [46]), and model-free optimization (from nonlinear control perspective [47]). The chapter will begin with some basic concepts in the science of search, and collective motion before going into the case studies.

2.1 The Science of Search

The term ‘search’ refers to the act of looking ‘carefully’ for something. Figuratively speaking, the ‘science’ happens when asking how/when to look ‘carefully’. Given a domain $\Omega \subseteq \mathbb{R}^n$, the searcher’s goal is to navigate from some initial starting point $\mathbf{x}(0) = \mathbf{x}_0$ towards some target point $\mathbf{x}^*$ (e.g. looking for a food source, or prey, etc.) to maximize some cost function. The searcher may be given some information (or a priori) of the physical system (e.g. areas that are rich in targets), such that, the searcher’s control step $\mathbf{u}$ will depend on some probability distributions related to the a priori of the physical system. The search step can typically be implemented by letting,

$$\mathbf{l} = \begin{cases} l \cos(\theta) \\ l \sin(\theta) \end{cases}, \quad (2.1)$$

where l and θ are drawn from some predefined distribution $p_l(l)$ and $p_\theta(\theta)$, respectively. In most cases, such as with Brownian motion, the turn angle θ is drawn from a uniform distribution $p_\theta(\theta) \sim \mathcal{U}(-\pi, \pi)$ and the jump length l can be drawn from a Normal distribution, $p_l(l) \sim \mathcal{N}(0, \sigma_l)$ with standard deviation σ_l . If the turn angle is

biased towards zero (relative to the previous heading), the random walk is referred to as a correlated random walk. In addition to the turn angle and jump length, the time step between jumps can be drawn from a distribution as well (this is the case for continuous time random walks (CTRW)). Therefore, one of the major challenges of search becomes finding the optimal choice of these spatial/temporal distributions, such that the search dynamics,

$$\mathbf{x}(t_{k+1}) = \mathbf{x}(t_k) + \mathbf{l}_k, \quad (2.2)$$

leads to the most efficient search. Thus the definition of efficiency has to be chosen to reflect the desired outcome of the search (i.e. in what sense?). For example, when we consider that a searcher may have limited energy to travel (i.e. maximum travel length), a typical efficiency function could take in the average step length, and the number of targets found as inputs. This comes with the assumption that the terrain is flat and has no constraints. Other cost functions may optimize information gain, entropy, or some other metric. Additionally, these can be combined into one linear cost function, such that, weights can be applied to the individual metrics with the highest priority - called a multi-objective cost function. However, in more complex and nonlinear systems, we know this is not the case, and the associated cost functions would need to incorporate those constraints to achieve the proper outcome.

When the random search is optimized, the optimal parameters found are dependent on the assumed target distribution, initial starting location, and the assumptions related to the searcher (e.g. vehicle dynamics and properties), the environment (e.g. the terrain or energy landscape) and the properties of the targets themselves (e.g. if the targets are destructive or non-destructive). Thus the science of search deals with the how to best choose the search distribution, position update rules, etc.

2.2 Collective Motion and Steering

The term ‘collective motion’ refers to spontaneous emergence of ordered movement within a system of agents (or ‘searchers’). These agents, individually, are thought of as autonomous agents with basic governing rules. The term ‘emergent’ refers to when these individual agents are allowed to interact with one another as a group, bringing new properties or capabilities. Given a domain $\Omega \subseteq \mathbb{R}^n$, the i -th searcher’s goal is to navigate from some initial starting point $\mathbf{x}_i(0) = \mathbf{x}_{i0}$ towards some target point (e.g. food source, etc.), to maximize some global cost function (typically the same for all searchers in the group). Like in Section 2.1 the searchers may be given some information (or a priori) of the physical system, such that, the searcher’s step $\mathbf{l}_i$ (direction and magnitude) depend on a probability distribution. The difference from Section 2.1 and that in the collective motion case, is that the agents now have access to information from the group and they can use that information to interact together as an intelligent and adaptive system. This coupling between purely random search and information sharing is what promotes emergent behavior.

There are examples in nature where groups interact in a way to exhibit emergent behavior. Examples include foraging ants and how they adaptively optimize paths [52] or the flocking behavior of starlings and schools of fish [53]. Additionally, groups of searchers or agents can be used for more formal problems, such as optimization, where meta-heuristic and bio-inspired algorithms have been developed. Algorithms such as the particle swarm optimization (PSO) [54], the grey wolf algorithm [55], or even the cuckoo search algorithm [56].

Collective motion and steering also refers to the deployment policies of interacting searchers or agents. Policies that, for example, can add spatial structure to the searchers (e.g. rules to avoid collision or to promote area coverage) or govern the movement to benefit the minimization of a desired cost function (e.g. gradient descent). Similarly to Section 2.1, the efficiency of the collective motion and steering policies should reflect the complexities of the system of interest.

2.3 A Case Study: On Optimal Tempered Lévy Flight Foraging

From observations in nature, we can see animals searching / foraging for food. Given no a priori information, what is the most optimal way to search? This question has been asked by many researchers over the years.

One common approach to the animal movement patterns is to use the scheme of optimizing random search [57; 58; 59]. In a random search model, single or multiple individuals search a landscape to find targets whose locations are not known *a priori*, which is usually adopted to describe the scenario of animals foraging for food, prey or resources. The locomotion of the individual has a certain degree of freedom which is characterized by a specific search strategy such as a type of random walk and is also subject to other external or internal constraints, such as the environmental context of the landscape or the physical and psychological conditions of the individual. It is assumed that a strategy that optimizes the search efficiency can evolve in response to such constraints on a random search, and the movement is a consequence of the optimization on random search.

Many researchers have concentrated on the study of different animals' foraging movements. It is shown that when the environment contains a high density of food items, foragers tend to adopt Brownian walks, characterized by a great number of short step lengths in random directions that maintain foragers in a small portion of the available space [60; 61]. In contrast, when the density of food items is low, individuals tend to exhibit Lévy flights, where larger step lengths occasionally occur and relocate the foragers in the environment. Due to the fact that the density of food items is often low, many animals behave a Lévy flight when foraging and their movements have been found to fit closely to a Lévy distribution (power law distribution) with an exponent close to 2 [62; 63]. For instance, the foraging behavior of the wandering albatross on the ocean surface was found to obey a power law distribution [64]; the

foraging patterns of a free-ranging spider monkey in the forests was also found to be a power law tailed distribution of steps consistent with Lévy walks [65; 66].

On this basis, researchers mainly consider two issues: one is to model the foraging behavior as a Lévy flight and the other one is to study the searching efficiency theoretically or experimentally. It is assumed that the forager behaves a random walk according to the distribution $p(l) \sim l^{-\mu}$, $1 < \mu < 3$. Then it is proven that the highest searching efficiency can be obtained when μ is close to 2 for the non-destructive case. While the searching efficiency is higher when μ tends to 1 for the destructive case. Later, many more complex situations are considered. Due to the fact that foragers are always searching in a bounded area, [67] and [68] studied the searching efficiency of Lévy flight in a bounded area. [61] took the spatial memory of foragers into consideration and concluded that the spatial information influenced the foraging behavior significantly according to the experimental results. Interestingly, it was claimed that the Lévy flight foraging behavior can also be interpreted by a composite search model [69; 70]. The model consists of an intensive search phase, followed by an extensive phase, if no food is found in the intensive phase. Particularly, [71] considered the waiting time between two successive flights and formulated the master equation for such foraging behavior.

Though many studies have proven that it is usually more efficient to utilize Lévy flight foraging strategy, standard Lévy flight cannot be followed in practice because of many uncertain factors. For instance, the forager may encounter obstacles or natural enemies and extremely large flight distance cannot be reasonable due to the restriction of the forager's flight ability. In this paper, we take these conditions into consideration and temper the Lévy distribution with an exponential decaying function, which results in a tempered Lévy distribution $p(l) \sim e^{-\rho l} l^{-\mu}$. It is then shown that a higher searching efficiency will be derived when a smaller ρ or μ is chosen, both by simulation and theoretical analyses. Further, two different types of master equations are derived: one is the standard diffusion equation and the other one is the tempered fractional diffusion equation. Since the first and second order moments exist, the foraging movement will finally result in a Gaussian motion, which indicates that the tempered fractional diffusion equation is in fact another expression for the standard diffusion.

The remainder of the case study is organized as follows. Section 2.3.1 provides the basic foraging model and some basic results are also given. In Section 2.3.2, we study the searching efficiency when a tempered Lévy distribution is considered.

2.3.1 Basic definitions and model description

In this section, we mainly recall the original model and basic results of Lévy flight optimal random search. Assume that target sites are uniformly distributed and the forager behaves as follows:

- (1) If a target site lies within a “direct vision” distance r_v , then the forager moves on a straight line to the nearest site. A finite value of r_v , no matter how

large, models the constraint that no forager can detect a target site located an arbitrarily large distance away.

- (2) If there is no target site within a distance r_v , then the forager chooses a direction randomly and a distance l_j from a probability distribution. It then incrementally moves to the new point, constantly looking for a target within a radius r_v along its way. If it does not detect a target, it stops after traversing the distance l_j and chooses a new direction and a new distance l_{j+1} ; otherwise, it proceeds to the target as rule (1).

In the case of non-destructive foraging, the forager can visit the same target site many times. In the case of destructive foraging, the target site found by the forager becomes undetectable in subsequent flights. Let λ be the mean free path of the forager between two successive target sites (for two dimensions $\lambda = (2r_v\phi)^{-1}$ where ϕ is the target-site area density).

On the basis of above behaviors, assume that the flight distance is distributed as the Lévy distribution

$$p(l) \sim l^{-\mu}, \quad l \geq r_v, \quad 1 < \mu < 3. \quad (2.3)$$

As shown in Fig. 2.1, researchers find that $\mu \rightarrow 2$ and $\mu \rightarrow 1$ will result in an optimal searching efficiency for the non-destructive case and destructive case, respectively. For more details about the model and existing results, one may refer to the works of [62; 63] and references therein.

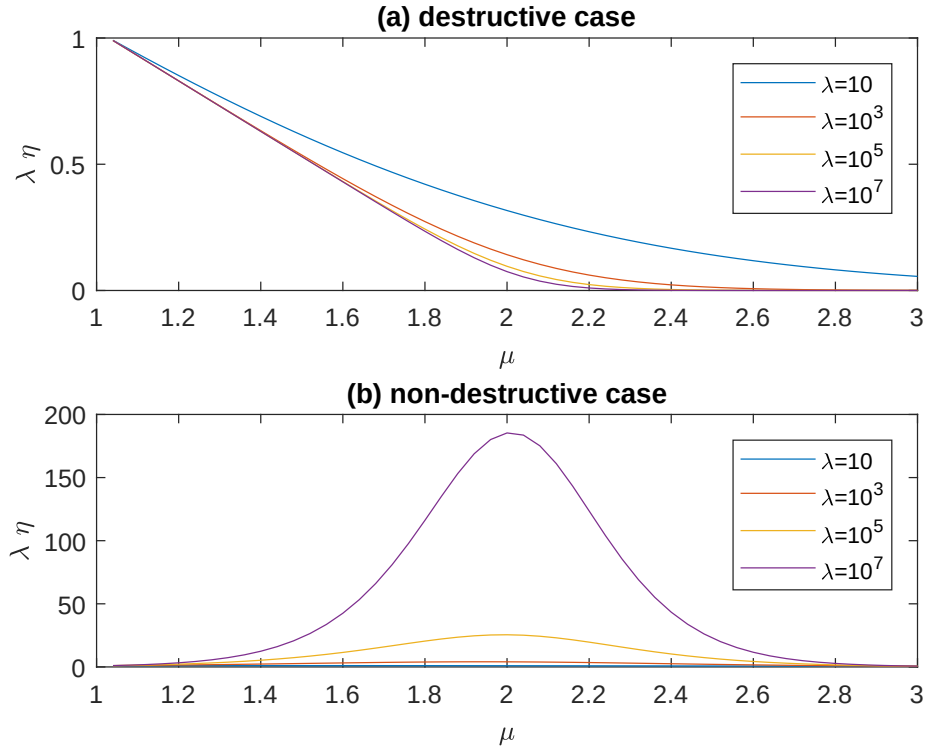
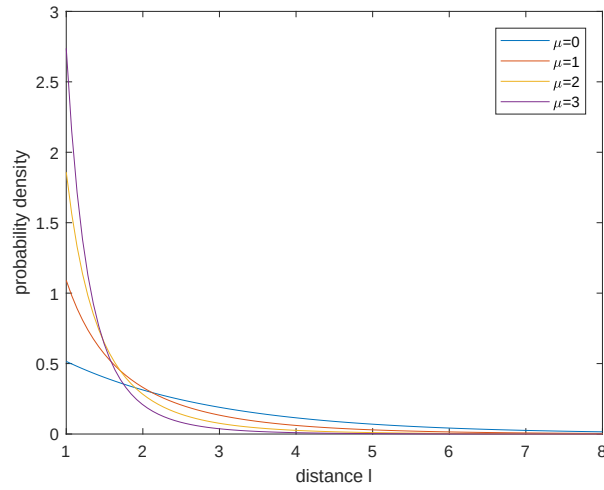
2.3.2 Searching efficiency with a tempered Lévy flight

In almost all the existing literatures about Lévy flight foraging, it is assumed that the flight distance at each step is independently distributed as (2.3). Distribution (2.3) is power-law decaying, which indicates that a large jump length will appear more frequently compared with the traditional Gaussian distribution. In practical foraging, after the forager determines the flight distance at some step, the flight will be interrupted by some unknown reasons, such as obstacles on the flight direction, natural enemies in the vision distance, and restrictions in the energy storage for each flight. Because of these reasons, we can assume that the flight distance is distributed as

$$p(l) \sim e^{-\rho l} l^{-\mu}, \quad r_v \leq l, \quad \rho > 0, \quad \mu \geq 1, \quad (2.4)$$

which indicates that the forager can keep the flight direction with the probability of an exponential distribution. Fig. 2.2 shows the probability density function (pdf) of a tempered Lévy distribution and one can find that the density decreases slower with a smaller μ , which means that a larger jump length is more likely to happen.

Remark The difference between (2.3) and (2.4) is that the power law distribution is tempered by an exponential decaying $e^{-\rho l}$. The exponential part $e^{-\rho l}$ can be viewed

Figure 2.1: Searching efficiency of standard Lévy flight for different mean free path λ Figure 2.2: Probability density for tempered Lévy distribution with $\rho = 0.5$ for different μ

as the probability density that the forager can keep its flight direction before he completes one flight in the existence of some unknown factors and ρ is determined by

the environment. Because Lévy distribution is now tempered by $e^{-\rho l}$, the first and second order moments of distribution (2.4) exist for arbitrary $\mu \in \mathbb{R}$. In the paper, we will discuss the problem in a wider range $\mu \in [1, \infty)$ rather than $(1, 3)$ for the Lévy distribution.

The non-destructive case

In this part, we will borrow the idea from [62] to optimize the searching efficiency. Given the probability density function of the flight distance as (2.4), the mean flight distance can be calculated as

$$\begin{aligned} \langle l \rangle &= \frac{\int_{r_v}^{\lambda} e^{-\rho|x|} |x|^{-\mu+1} dx + \lambda \int_{\lambda}^{\infty} e^{-\rho|x|} |x|^{-\mu} dx}{\int_{r_v}^{\infty} e^{-\rho|x|} |x|^{-\mu} dx} \\ &= \frac{\Gamma_{up}(\rho r_v, 2 - \mu) - \Gamma_{up}(\rho \lambda, 2 - \mu) + \lambda \rho \Gamma_{up}(\rho \lambda, 1 - \mu)}{\rho \Gamma_{up}(\rho r_v, 1 - \mu)} \end{aligned} \quad (2.5)$$

where, the incomplete gamma function Γ_{up} is defined as

$$\Gamma_{up}(x, a) = \int_x^{\infty} t^{a-1} e^{-t} dt. \quad (2.6)$$

Let N be the mean number of flights taken by a Lévy forager while travelling between two successive target sites. Since the first and second order moments of tempered Lévy distribution exist, the trajectory of the forager will result in a Brownian motion. According to the existing results by [62], for the non-destructive case, it follows that the mean flight number between two successive targets can be estimated as

$$N_n \approx \left(\frac{\lambda^2}{2D} \right)^{\frac{1}{2}} \quad (2.7)$$

where, D is the diffusion constant. According to the standard diffusion equation, it is found that the diffusion constant $D = \frac{a}{2b}$, where a is the second order moment of flight distance and b is the mean of the waiting time. Since we do not take the time into consideration, one can conclude that N_n is proportional to $\left(\frac{\lambda^2}{a} \right)^{\frac{1}{2}}$. Here, a can be calculated as

$$a = \int_{r_v}^{\infty} e^{-\rho l} l^{2-\mu} dl = \rho^{\mu-3} \Gamma_{up}(\rho r_v, 3 - \mu). \quad (2.8)$$

Based on the above analyses, we can then calculate the searching efficiency which is defined as

$$\eta = \frac{1}{N \langle l \rangle}. \quad (2.9)$$

Take r_v as 1 when simulating and the results for different mean free path λ are shown in Fig. 2.3. Following observations can be drawn

- (1) For fixed mean free path λ and ρ , a smaller μ will result in a higher searching efficiency.
- (2) For fixed mean free path λ and μ , a smaller ρ will result in a higher searching efficiency.
- (3) The mean free path λ almost has no influence on the choice of μ and ρ to derive the highest searching efficiency.

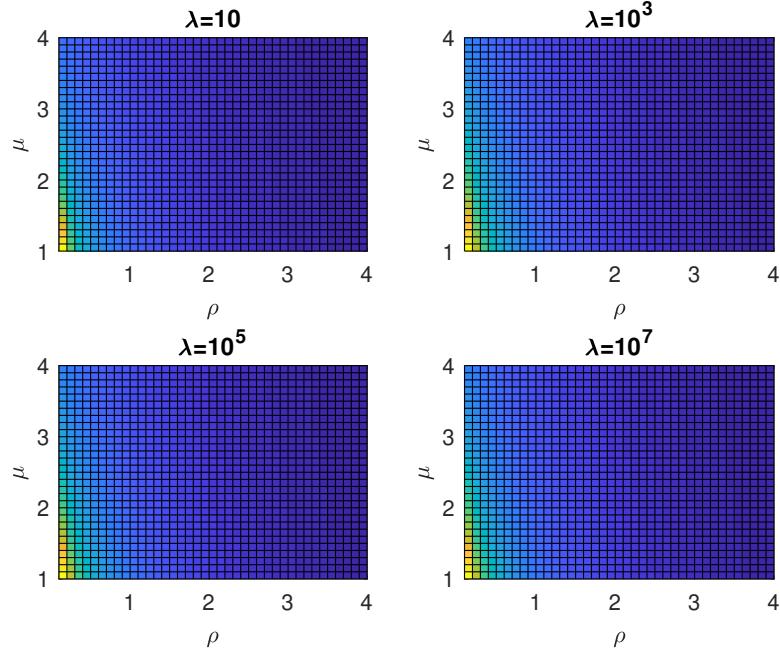


Figure 2.3: Searching efficiency $\eta\lambda$ for different order μ and ρ : the nondestructive case with different λ

As interpreted in the existing papers, the Lévy distribution can lead to a higher efficiency in a sparse area due to the higher probability of large jump lengths. For this issue, a smaller μ or ρ will both decrease the decaying speed of the probability density, which means that the large jump lengths are more likely to appear. Hence, observations (1) and (2) can be explained since frequently large jump lengths can help covering a wider range where it is more likely to find a target in a sparse area. Generally, the density of target site is sparse in practice which means that λ is usually large. Due to the exponential decaying of tempered Lévy distribution, the value of $\lambda \int_{\lambda}^{\infty} e^{-\rho|x|} |x|^{-\mu} dx$ is quite small and almost has no influence on the searching efficiency. It can then explain why the results of Fig. 2.3 with different λ are similar.

One can also interpret the observations from the practical perspective. As discussed before, the tempered item $e^{-\rho l}$ can be viewed as the probability density that the forager can keep its flight direction before he completes one flight in the existence

of some unknown factors. Thus, a smaller λ means that the probability of a forager to encounter some uncertain factors is lower and the foraging efficiency should be higher.

The destructive case

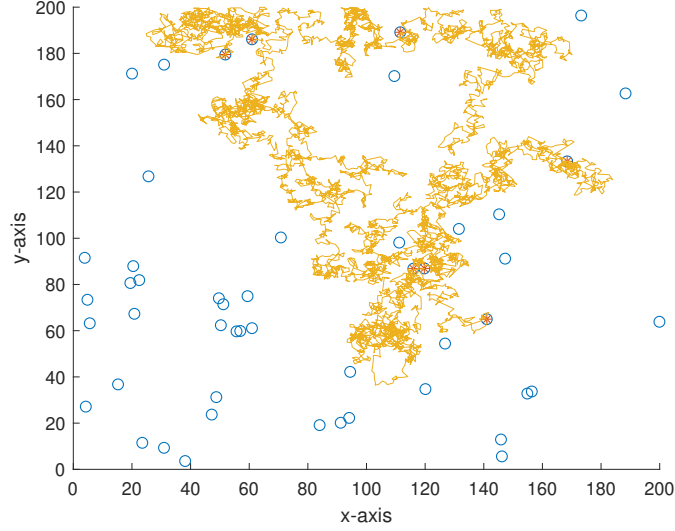


Figure 2.4: A typical example of foraging procedure with $\lambda = 1$ and $\mu = 1$

For the destructive case, the mean number N can be expressed as

$$N_d \approx \frac{\lambda^2}{2D}. \quad (2.10)$$

Similar to the non-destructive case, one can then calculate the searching efficiency using (2.9). The results are shown in Fig. 2.4, which is very similar to the non-destructive case. It is found that a smaller μ or ρ will both result in a higher search efficiency. The mean free path λ almost has no influence on the optimal choice of parameters μ and ρ (see Fig. 2.5). We have shown that for the Lévy distribution, $\mu \rightarrow 1$, where a large jump length appears more likely, will lead to a higher searching efficiency. Thus, a smaller ρ and μ will also result in a larger searching efficiency because large jump lengths are more likely to happen.

Experimental results

We also implement an experiment for validate the theoretical analyses. Consider a 200×200 area and 50 targets are uniformly distributed in this area. The vision distance is $r_v = 1$ and the total flight distance is no longer than 10000 which can be viewed as the flight capability of the forager. The searching efficiency is estimated as $\frac{N_{num}}{L_{total}}$ where N_{num} is the number of found targets and L_{total} is the total flight distance.

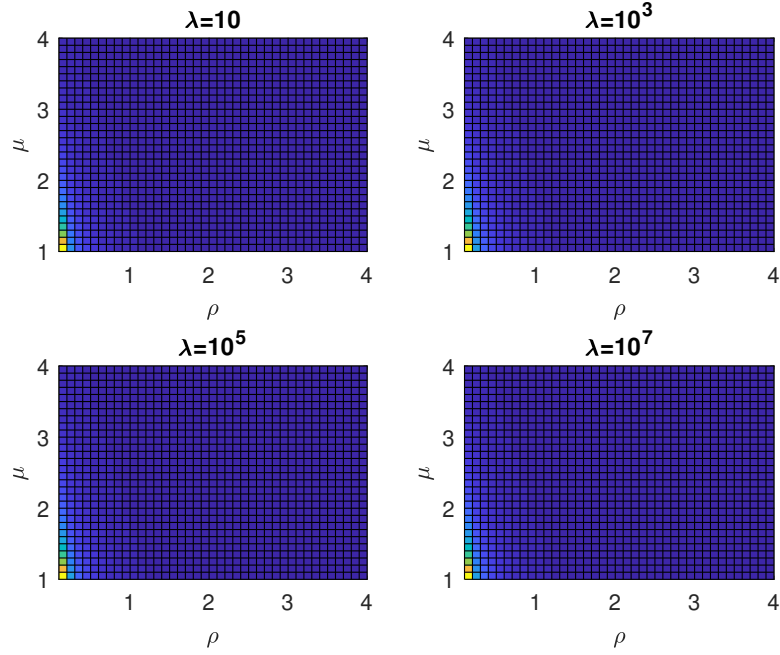


Figure 2.5: Searching efficiency $\eta\lambda$ for different order μ and ρ : the destructive case with different λ

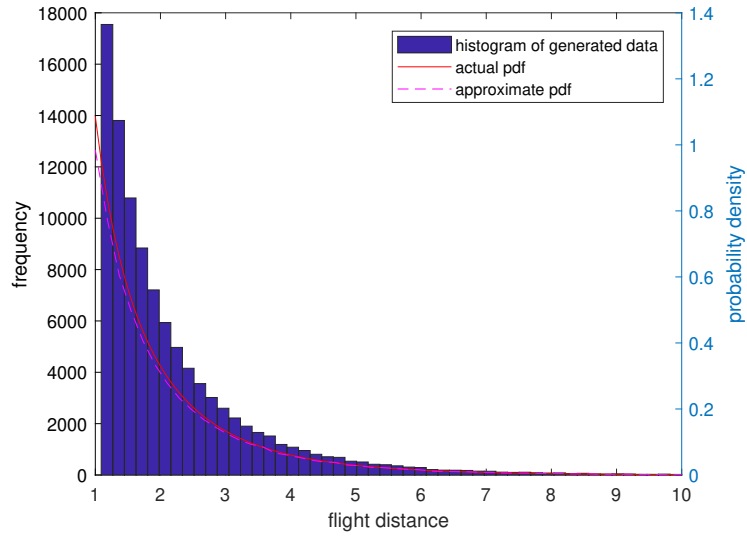
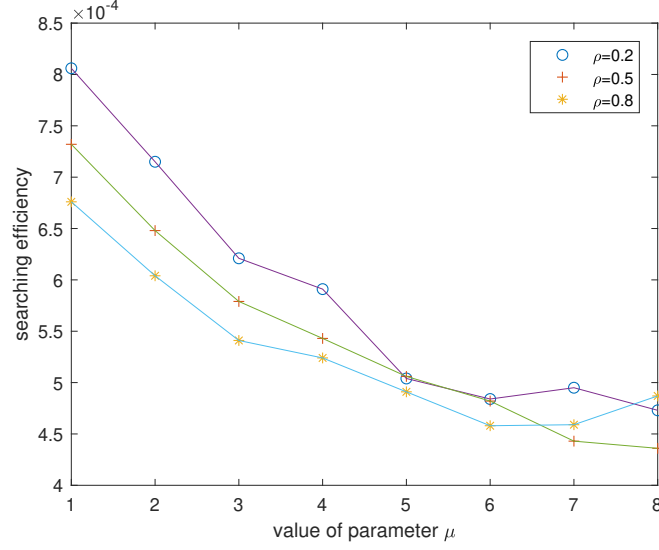
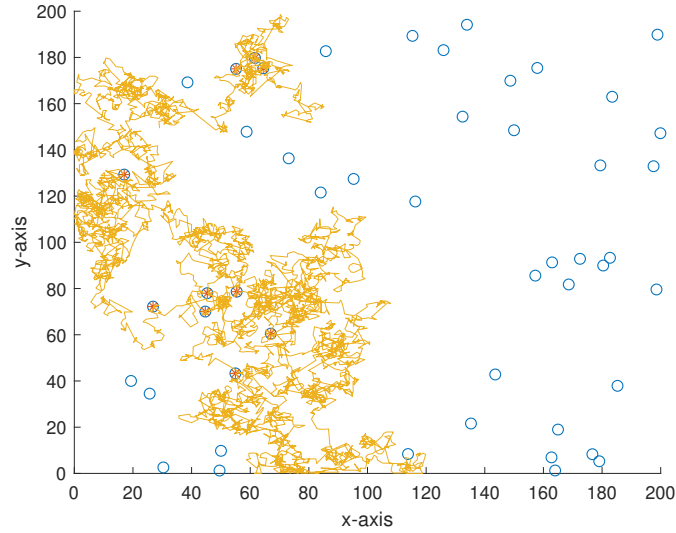


Figure 2.6: The actual tempered Lévy density function with $\rho = 0.5$ and $\mu = 1$ and the statistic results of generated jump lengths

From Fig. 2.7 where the searching efficiency is derived by averaging 100 independent runs, one can find that a smaller ρ and μ will both lead to a higher searching efficiency, which is consistent with the theoretical analyses. Because a larger μ will make the

Figure 2.7: Experimental results of searching efficiency for different λ and μ Figure 2.8: A typical example of foraging procedure with $\lambda = 0.5$ and $\mu = 1$

density function decrease quickly, the range of jump lengths is then very tight. Thus, the searching efficiency is very close for a large μ where the jump lengths are all around the vision distance r_v . Fig. 2.8 - Fig. 2.9 give some typical foraging procedure for different parameters and one can find that all of them perform a Brownian motion. Additionally, larger jump lengths frequently appear in Fig. 2.8 compared with the other two figures, for which the searching efficiency is the highest.

Remark: In this study, we numerically generate the jump lengths distributed as a tempered Lévy distribution and Fig. 2.6 shows the actual density function and the

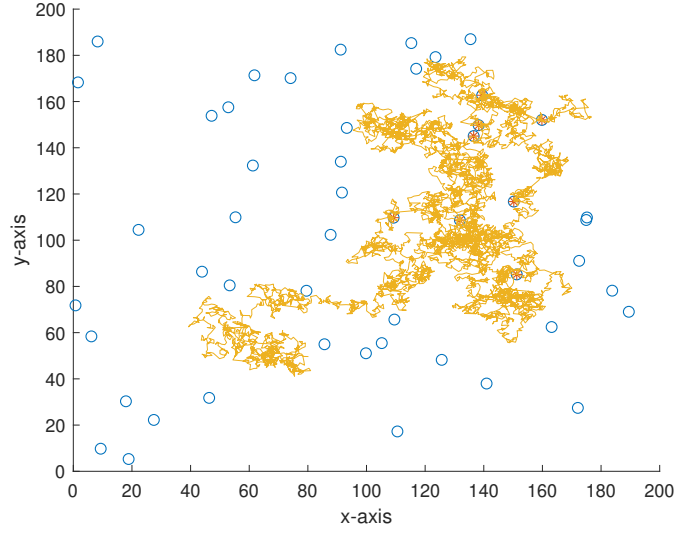


Figure 2.9: A typical example of foraging procedure with $\lambda = 0.5$ and $\mu = 3$

statistic result of generated jump lengths. It is found that the statistic result is very close to the actual density function.

2.3.3 Conclusion

To conclude, we consider the optimal random foraging whose flight distance is distributed according to a tempered Lévy distribution $p(l) \sim e^{-\rho l} l^{-\mu}$. It is found that a higher searching efficiency can be derived when we choose a smaller ρ or μ , which results in a slower decaying speed.

2.4 A Case Study: Individual and Collective Foraging in Autonomous Search Agents with Human Intervention

Foraging is oftentimes studied as an activity produced by individual organisms. A single bird flies over an area in search of food, a tiger roams the jungle, or a person scans their terrain for resources. Other times, foraging is studied as a collective activity that groups of organisms engage in [72; 73; 74; 75]. Collective foraging occurs when groups of organisms interact and move together while searching for resources, and it is more often associated with organisms that have limited cognitive capacities for planning and decision-making. Organisms with greater cognitive capacities may also engage in collective foraging, but they are more likely to exhibit flexibility in switching between different foraging modes depending on various factors. For example, a lion may choose to hunt alone or team up with other lions to find and take

down prey [76]. A person may choose to help others harvest a large patch of berries or head off alone in search of unfound patches. Foragers can also communicate information about resources and conditions to each other, and thereby help individuals make decisions about where and with whom to forage [72; 73; 74; 75; 77; 78].

In the present study, we investigate whether foraging can benefit from the ability to vary between individual versus collective search modes, as a means of improving group foraging performance. A given mode may be more or less advantageous depending on various foraging factors and conditions. For instance, the conditions under which collective foraging is advantageous are often studied through the lens of social foraging theory [79]. Group cooperation can outperform individual or independent foraging strategies by exploiting food resources more quickly [80], and also by providing security from predation among other social benefits [76; 53]. However, these benefits may not always be available or sufficiently salient, and they may be outweighed when foraging becomes sufficiently competitive [81]. This dependence on conditions suggests that the ability to modulate between independent and collective modes of foraging may be advantageous for collective foraging performance.

For example, Harel et al. showed that Eurasian griffon vultures (*Gyps fulvus*) switch between independent and cooperative foraging strategies in response to social cues [82; 83]. They found evidence that vultures can exchange information about a found carcass even when they are far from its location. Specifically, vultures who previously visited a carcass were more likely to be followed when revisiting the same carcass location. Uninformed vultures apparently used visible cues, such as blood stains on the head and body, to flexibly choose when to collectively forage and with whom. Flexibly switching between these two strategies increased group foraging success by helping uninformed vultures find new carcasses more quickly, and together they were able to consume carcasses more quickly, thereby leaving less chance for other scavengers to share in the meal. Similarly, Geoffroy's spider monkeys (*Ateles geoffroyi*) engage in information transference between naïve and informed foragers which helps to improve their success [84].

In general, efficient foraging needs to result in both finding as well as consuming enough food for individuals and groups to survive [83; 85; 86; 87]. For example, Beauchamp (2005) designed a social foraging simulation that demonstrated this principle of efficient foraging indirectly. Social foragers consumed food at a faster rate compared with individual foragers, but faster consumption came at the price of consuming less food per individual. There was also reduced variability in the rate of food consumption among social foragers, indicating that social foraging may confer protection against long periods of famine.

The benefits of social foraging have been shown in other birds of prey as well [82; 88; 89; 90], and the apparent prevalence of collective foraging in nature has led researchers to develop and test formal models of collective foraging, to investigate the underlying principles and processes. For example, Liu and Passino (2004) created a collective foraging model based on balancing forces of attraction and repulsion so that agents tended to position themselves nearby neighbors while keeping some dis-

tance. The goal was for agents to find food resources by following gradients to their locations. Coordinating movements with nearby agents was useful for them to collectively follow otherwise unreliable gradients in the simulated environment that led to resources [91; 92; 93; 94; 95; 96; 97].

In the present study, we developed an agent-based simulation in which both individual and collective foraging strategies have their advantages, so being able to vary between them should be advantageous to the group. We created a social foraging simulation in which we could test collections of autonomous agents with movement rules that varied between more and less collective modes of interaction. For our targeted loose coupling condition, rules were parameterized to balance individual versus collective search behaviors. We were also able to test how humans move and interact with autonomous foraging agents by including a condition in which a human player controlled one of the search agents. Our aims were to 1) understand how coupling strength in the movements of autonomous agents plays a role in foraging performance by diversifying their collective movement patterns, and 2) test whether human agents with high cognitive capabilities while searching engage in loose coupling with autonomous agents in the service of collective foraging.

Our results showed that balancing attraction and repulsion promoted and sustained coordinated movements because random variability inherent to individual foraging behaviors was averaged out as a result of loose coupling. This balance of attraction and repulsion was accomplished using the so-called Lennard-Jones potential [98; 99] applied to govern the degree to which agents are attracted to or repelled from each other as a function of their distance apart.

We manipulated the degree of coupling strength using the Lennard-Jones potential [100] plus a flocking term that correlated the direction of movement among nearby search agents. We also investigated how the addition of a human agent who can switch between independent and cooperative search strategies affects collective movement patterns and foraging efficacy. We expected the intervention to improve group performance, and we also tested whether the degree of loose coupling, in terms of distancing and flocking among autonomous agents, affects human search performance. By comparing agent-based simulations with and without human intervention, we also investigated the benefits of human memory and decision-making in managing the balance of independent versus collective foraging behaviors.

2.4.1 Methods

Simulation

The agent-based foraging model was implemented in NetLogo with a 200×200 grid of pixels with periodic boundary conditions and based on a previous social foraging model [53]. The grid was empty except for one target at time. The target occupied four pixels and was located at random. The goal was for ten agents to search the task space for a gold star target and find as many of these targets as possible within a set amount of time. Ten agents were chosen to provide enough grouping variability while

also allowing human intervention to have a meaningful impact on group coordination and performance. While searching for targets, all agents moved at a constant velocity of 1 pixel per time step. Agents could not “see” the target until they came within a $d_v = 22.5$ pixel radius of their position, which meant the visible area for each agent was 4% of the total game space. With ten agents searching together, it did not take long to find each next target, which helped to find multiple targets within minutes of time.

Search was also facilitated by a chaining effect whereby agents could see when other agents in their view had detected the target. The rationale was that agents change their behavior when detecting a target, e.g. they take a more direct path towards a specific point. Other agents may see this behavioral change before they see the target itself, so the behavioral change becomes a target itself to draw an agent to share in the find. This dynamic was implemented by setting a flag on agents who detect the target and adding a rule to drive unflagged agents to converge on a visible flagged agent as they would converge on the target itself.

When agents arrived at target locations, they immediately started consuming the target, one unit of consumption per time unit, where each target consisted of 500 consumption units. Therefore, it required 500 time steps for one agent to fully consume a target. Less time was required to consume as more agents arrived at the target location and consumed it together, simultaneously. Each target disappeared after it was completely consumed, and a new target appeared at a new random location.

Each simulation session was defined by the rules applied to all ten autonomous agents that generated search movements to find each target or agents who found the target (i.e. visual chaining [88]). The default rule present in all conditions was a correlated random walk (CRW), which made each agent wander through space randomly with some tendency to maintain the current heading. Next was a flocking rule that added a tendency for movement in the average direction of nearby agents. Third and final was a distancing rule based on a generalization of the Lennard-Jones potential [98; 100], that drove agents to maintain a given distance from each other. The governing equations for the CRW, flocking, and distancing forces on the i -th agent (given $i \neq j$) at time t are as follows:

$$\text{CRW}(\vec{d}_N) : \quad \vec{d}_{i,N} = [\cos(\theta_i(t - \delta t) + \theta_N), \sin(\theta_i(t - \delta t) + \theta_N)], \quad (2.11)$$

$$\text{Flocking}(\vec{d}_A) : \quad \vec{d}_{i,A} = \sum_{j=1}^{\Omega} \vec{d}_j(t - \delta t), \quad (2.12)$$

$$\text{Distancing}(\vec{d}_{LJ}) : \quad \vec{d}_{i,LJ} = - \sum_{j=1}^{\Delta} \left[\left(\frac{s}{\|\vec{d}_{i,j}\|} \right)^4 - \left(\frac{s}{\|\vec{d}_{i,j}\|} \right)^3 \right] \frac{\vec{d}_{i,j}}{\|\vec{d}_{i,j}\|}. \quad (2.13)$$

For the CRW, θ_i represents the current heading of the i -th agent and θ_N is a correlated random angle given by $\theta_N = \theta_R - \theta_L$, where $\theta_R \sim U(0, 180)$ and $\theta_L \sim U(0, 180)$ are independent random turning angles between 0 and 180 degrees. The difference

between these random uniform turning angles, produces a symmetric probability distribution from -180 to 180, linearly weighted towards zero. For flocking, the vector $\vec{d}_j$ represents the directional heading of the j -th agent at time $t - \delta t$, where δt represents the time step. The flocking direction $\vec{d}_A$ is calculated as the sum of the set of all the agents within the vision distance d_v denoted as $\omega = \{\vec{d}_j \mid \|\vec{d}_{i,j}\| < d_v\}$. The vector $\vec{d}_{i,j}$ represents the distance in pixels from the i -th and j -th agents. The distancing $\vec{d}_{LJ}$ is calculated from the separation parameter s (the desired distance between agents), and the distance $\vec{d}_{i,j}$ between pairs of agents belonging to the set $\Delta = \{\vec{d}_j \mid \|\vec{d}_{i,j}\| < 1.5d_v\}$. To avoid losing sight of the agents, the separation distance s was set to $s = 15$ pixels, which is inside the vision distance d_v . The exponents 4 and 3 in the distancing rule $\vec{d}_{LJ}$ represent the repulsion and attraction terms, respectively. The original exponents values for the Lennard-Jones potential are 12 and 6, respectively. These values were chosen arbitrarily, through trial and error, to reflect the loose coupling between agents. The result is a mostly repulsive force with a weak attractive component that fades away as the agents separate greater than s . The governing equation for the i -th agent at time t is found by combining additively, $\vec{d}_i = \vec{d}_{i,N} + \vec{d}_{i,A} + \vec{d}_{i,LJ}$. Simulation conditions were defined by turning off or on the flocking and distancing rules such that the CRW rule was always in effect, resulting in four different movement conditions: Random (CRW only), Flocking (and CRW), Distancing (and CRW), and Loose Coupling (all three rules combined). See Fig. 2.10 for visual illustrations of the movement rules and see Fig. 2.11 for general trajectory examples of each movement condition.

Each simulation session lasted 13,500 game time steps long, referred to hereafter as ticks, which corresponded to about eight minutes in real time when simulated through the NetLogo interface. Performance was measured in terms of number of targets found and consumed, which was based on the times needed to find and consume targets. Each movement condition was tested in 60 sessions to match the number of sessions with data collected from participants in the human intervention experiment, described next.

Participants, Materials, and Experimental Setup

The simulated collective foraging sessions described in the previous section consisted of autonomous agents only, with no human intervention. Matching sessions were conducted with human players, where each session including one human controlling one of the ten agents, as follows.

All methods involving human subjects were carried out in accordance with relevant guidelines and regulations. Additionally, all experimental protocols were approved by the University of California, Merced IRB committee and informed consent was obtained from all subjects prior to participation. All subjects were above the age of 18 years old. Sixty participants from the University of California, Merced were recruited for course credit. After being instructed about the game and filling out the consent form, each participant controlled one of the ten search agents and engaged in

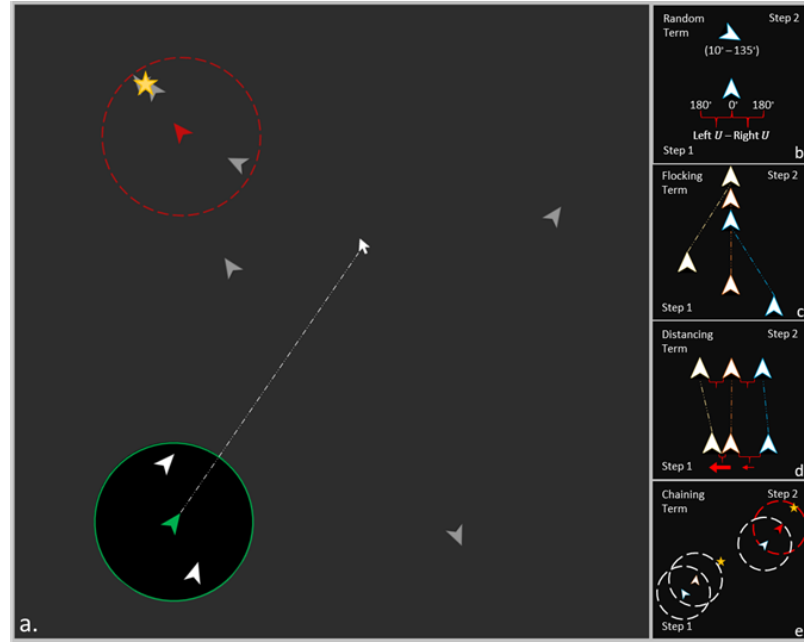


Figure 2.10: (a) Illustrative view of the task space (not shown to players). The green agent and circle represent the human’s agent and their field of view (detailed more below). Area outside their field of view was occluded (greyed out area). The dotted line represents that the human agent moved towards the mouse pointer position, so the human could control movement direction by moving the mouse. (b) Random movement shown to be random angular deviations of movement from each previous heading. (c) Flocking term directed agents to converge towards a similar shared movement trajectory. (d) Distancing term prompted agents to separate from one another when close, and towards each other when further away. (e) Visual Chaining prompted agents to move directly toward an agent flagged as detecting the target.

collective foraging with the goal of finding and consuming as many targets as possible in the allotted time. Each participant controlled their on-screen avatar by using a computer mouse to place a pointer at the desired location, and then their avatar moved in the direction of the pointer. If the avatar reached the desired location, it kept moving in the same direction until the human player moved the mouse to change course. Participants had the same visual radius as autonomous agents, and their avatar affected autonomous search agent dynamics according to the rules as described above. Autonomous agents did not affect movement of the human avatar except for the chaining rule—the human avatar went straight to an autonomous agent if they were tracking towards the target.

The foraging game was designed to give human players the same operational information and latitude as automated agents, so that the only difference with human players was their memory about prior states of the environment and foraging outcomes, and decision-making based on memory and strategies developed through

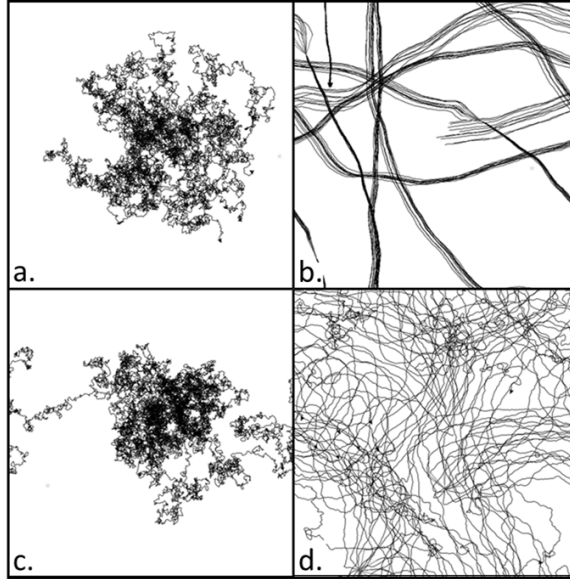


Figure 2.11: Example movement trajectories for 2000 ticks for each movement condition: (a) Random; (b) Flocking; (c) Distancing; and (d) Loose Coupling.

experience. The simulation ran at about 35 msec per update, and each human player ran through one session of each of the four movements conditions, in counterbalanced order across participants. The number of time steps was chosen to be long enough to elicit variability in performance, but short enough to run each session in about eight minutes.

Measures and Analyses

The overall measure of performance for each game session was total targets found (session length was constant), which was determined by the time needed to find and then consume each target. These two components of performance were measured by search time and consumption time. Search time was operationalized as the time elapsed prior to any agent detecting the target, and consumption time was defined as the amount of time it took to fully consume a given target (maximum of 500 ticks). Both measures were computed on a per target basis, and then averaged for each session. This aggregation created equally sized samples across all conditions for each measurement. Overall performance was also measured by the average total trial time, i.e. the sum of search and consumption time, where lower times corresponded with better performance.

Two factors were manipulated to test efficacy of loose coupling and the role of memory and strategy in collective foraging. The efficacy of loose coupling was tested by comparing different movement rules for autonomous search agents across different sessions, where all agents in each session were governed by the same set of rules. The role of memory and strategy was tested by comparing sessions with and without

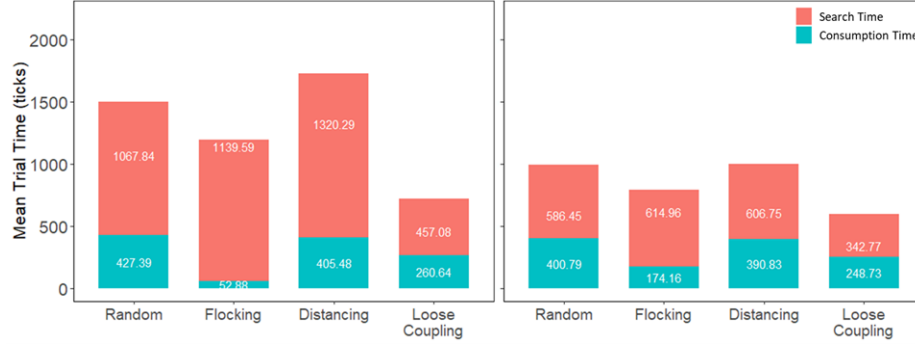


Figure 2.12: Mean trial time per session as a function of movement type without human intervention (left) and with human intervention (right). Mean trial times are divided into their composite search times (red) and consumption times (teal).

human intervention, and because simulations are not statistically comparable to humans in terms of their variability, we ran separate analyses with and without human intervention.

To test the outcomes of these measures independently, we conducted separate within-subject analyses of variance (ANOVA) with each as the dependent measure, and two independent variables, human intervention (present or absent) and movement type (random, flocking, distancing, and loose coupling). Any result lower than $p < 0.05$ are considered as statistically reliable. Estimated effect sizes of ANOVA results are represented by partial eta squared (η_p^2). Lastly, we also conducted Tukey honestly significant difference (HSD) post-hoc analyses to conservatively test for differences between specific pairs of conditions, even though many of these tests were planned a priori.

2.4.2 Results

Loose coupling and human intervention promote collective foraging success

We first determined group search performance by assessing the average search time, consumption time, and total targets found in each movement condition with and without intervention.

Results showed that search performance as measured by mean trial time was better with loose coupling and human intervention, as seen in the lowest average trial times in Fig. 2.12. Movement type had a reliable effect on performance without human intervention, $F(1,59) = 27.65$, $p < 0.001$, $\eta_p^2 = 0.319$, and with human intervention, $F(1,59) = 20.85$, $p < 0.001$, $\eta_p^2 = 0.261$. The specific direction of effect was supported by post-hoc Tukey HSD comparison tests showing that loose coupling was significantly better than other movement types both with and without human intervention (each $p < 0.001$). By necessity, the same pattern of results was found when performance was measured by number of targets found per session: On average, more targets were

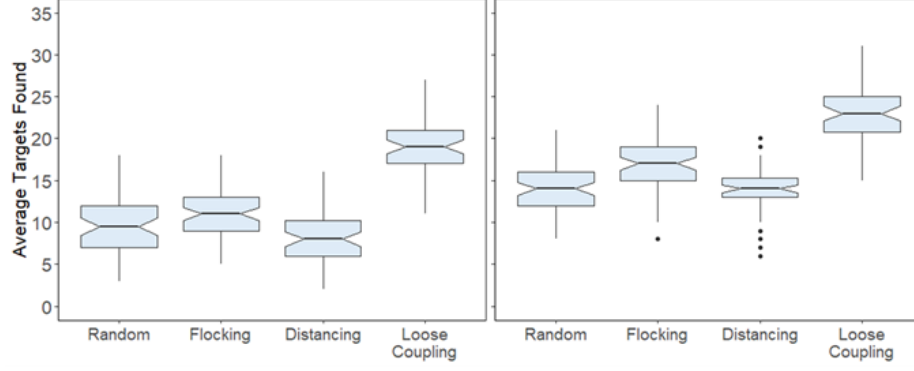


Figure 2.13: Mean targets found per session as a function of movement type without human intervention (left) and with human intervention (right).

found with human intervention ($M = 16.94$) than without ($M = 5.89$), and more targets were found with loose coupling ($M = 16.26$), compared with other movement types ($M = 9.8$). Human intervention did not interact with movement type, $F(1,59) = 2.55$, $p = 0.116$, $\eta_p^2 = 0.141$, which indicates that human intervention resulted in more targets found on average per session for all movement conditions ($M = 11.04$), see Fig. 2.13 below.

Next we broke the performance measure of mean trial time into its two component parts of search time (i.e. the time from start of a trial to when any of the agents detecting the target) and consumption time (the time from the first agent landing on the target to its complete consumption, which went faster as additional agents landed to share in consumption). Mirroring mean trial times, search times were fastest in the loose coupling condition regardless of human intervention: without, $F(1,59) = 48.17$, $p < 0.001$, $\eta_p^2 = 0.449$, and with intervention, $F(1,59) = 41.24$, $p < 0.001$, $\eta_p^2 = 0.411$, refer back to Fig. 2.12. Tukey HSD for both loose coupling conditions were significantly faster than all other respective movement conditions, $p < 0.001$. By contrast, consumption times were fastest in the flocking condition: without, $F(1,59) = 542.2$, $p < 0.001$, $\eta_p^2 = 0.902$, and with intervention, $F(1,59) = 56.79$, $p < 0.001$, $\eta_p^2 = 0.49$. All Tukey HSD flocking condition comparisons were significant, $p < 0.001$. Flocking produced faster consumption times because agents were usually clumped together when the target was found, so they all landed on the target to consume it together. This effect of flocking was predicted to occur, and we also predicted that the distancing condition would produce the fastest search times by means of a divide and conquer strategy. Results were not consistent with this latter prediction because adding flocking to distancing actually improved search times. We return to this unexpected result later when we present analyses of the rate at which agents collectively covered the search area.

Analyses of search times and consumption times as a function of human intervention found that, again mirroring mean trial times, human intervention improved search times substantially across all four movement conditions, albeit less reliably for

loose coupling because of an apparent ceiling effect (loose coupling without human intervention already produced fast search times): $F(1,59) = 22.24$, $p < 0.001$, $\eta_p^2 = 0.086$. By contrast, human intervention improved consumption times in most conditions, $F(1,59) = 80.00$, $p < 0.001$, $\eta_p^2 = 0.253$, but surprisingly, humans caused slower consumption rates in the flocking condition, comparison Tukey HSD $p < 0.001$. The apparent detriment of human intervention on flocking consumption times can be explained by humans finding targets on their own, without the benefit of other agents nearby to join in consumption. This explanation is further addressed in the next section.

Loose coupling diversifies groupings of search agents

Collective foraging performance was best with loose coupling, which was predicted based on the hypothesis that loose coupling balances the benefit of flocking versus distancing. This balance should result in more flexibility in agent groups as they merge and split over time—the agents only partially affect each other’s movements, thereby allowing interactions between search agents to vary as they come in and out of view of each other. To quantify flexibility in grouping, we examined the distribution of numbers of agents in view for each given tick, trial, and agent. If groupings do not change much within each trial, then there should be little variation in the numbers of agents in view, and the distribution should have a sharp peak. By contrast, if groupings vary during a trial, then the numbers of agents in view should vary, and hence their distribution should be more spread out.

We used Shannon entropy to quantify the degree to which the frequency distributions in groupings were more peaked (low entropy) versus more spread out (high entropy). Entropy has been used previously for capturing the fission-fusion dynamics for various groups animal species [101]. Hereafter, we refer to it as Grouping entropy which was calculated as, $-\sum p(x_i) \log p(x_i)$ where x_i is the number of agents viewed by i -th agent over time, and p is the probability associated with the proportion of time that x_i agents were in view.

To focus on grouping entropy from the perspective of autonomous agents, we removed the human player from entropy calculations, and to make analyses comparable, we removed a simulated agent at random in sessions without intervention so that entropy was computed over zero to eight possible agents in view in both conditions. The first 14 ticks at the start of each new trial (when each new target was generated) was removed to avoid initial transients due to agents starting together from the previous target location. Entropy was computed over the subsequent ticks for each trial, up to the tick when the next target was detected by one of the agents. We also computed grouping entropy with respect to the human agent, and again we removed one autonomous agent at random so that entropy was computed over zero to eight possible agents in view.

Figs. 2.14 & 2.15 show the proportion of agents in view aggregated over trials and individuals for each movement condition with respect to autonomous agents in the simulation (Fig. 2.16) and with respect to human agents in the experiment (Fig. 2.17).

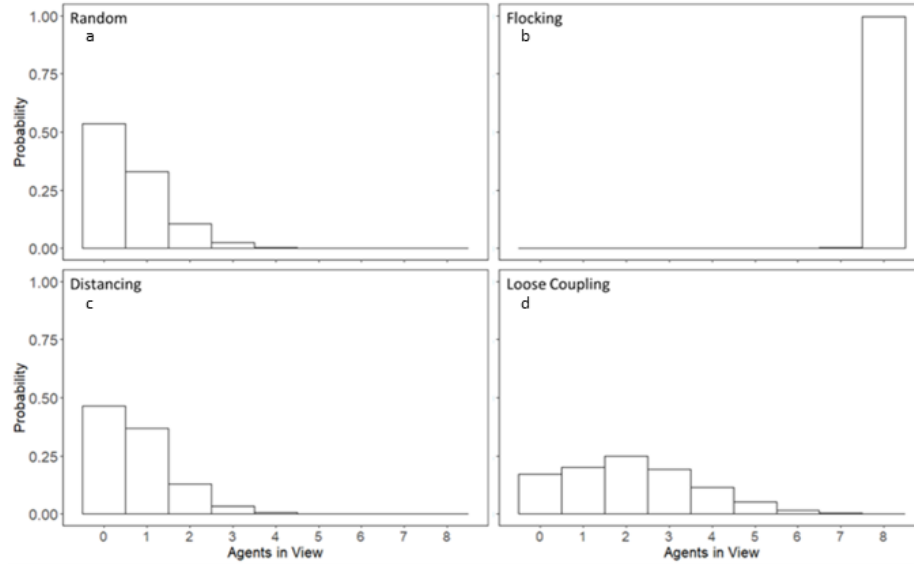


Figure 2.14: Normalized histogram of the number of agents in view during the search time period respective to one autonomous agent without intervention.

These histograms show that the rules governing agent movements and interactions had large effects on agent groupings. The random and distancing conditions were similar in that agents traveled solo much of the time, with another agent in view sometimes, and two or three more on occasion. Adding the flocking rule to each of these two conditions resulted in opposite effects on grouping entropy than the other movement conditions. Flocking agents constantly maintained each other within their respective fields of view, best show in Fig. 2.14b. Flocking plus correlated noise (the random condition) resulted in all agents converging and moving together such that variability caused by noise was not enough to disperse the single grouping of autonomous agents once it was formed. By contrast, adding the distancing term to flocking (along with the correlated noise) was sufficient to counteract flocking and disperse agents such that their flight configurations varied over time. This variation resulted in more varied group sizes and hence more variability and greater entropy in the numbers of agents in view.

We tested the effects of movement type and human intervention on grouping entropy using two different statistical analyses. First, we tested entropy values for individual simulated agents with and without human intervention, as a function of movement type, as shown in Fig. 2.16. Entropy values were minimal in the flocking condition because agents were always in a unified group, so we removed this condition from statistical analyses. Second, we ran the same analyses after removing trials in the human intervention condition when the human player was first to find the target, so that entropy values were not directly affected by human intervention. The second analysis allowed us to ascribe differences to the effect of human intervention on the movements of autonomous agents.

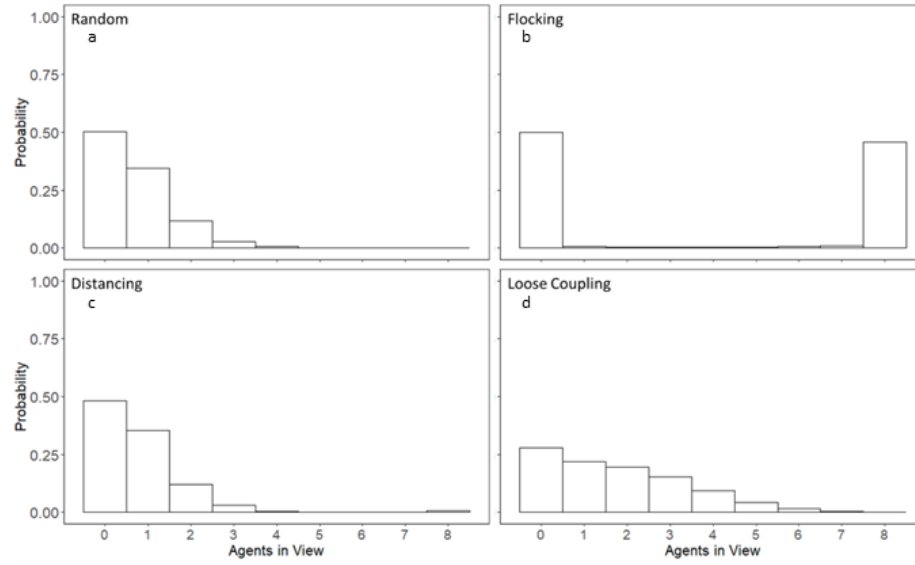


Figure 2.15: Normalized histogram of the number of agents in view during the search time period respective to only human agents.

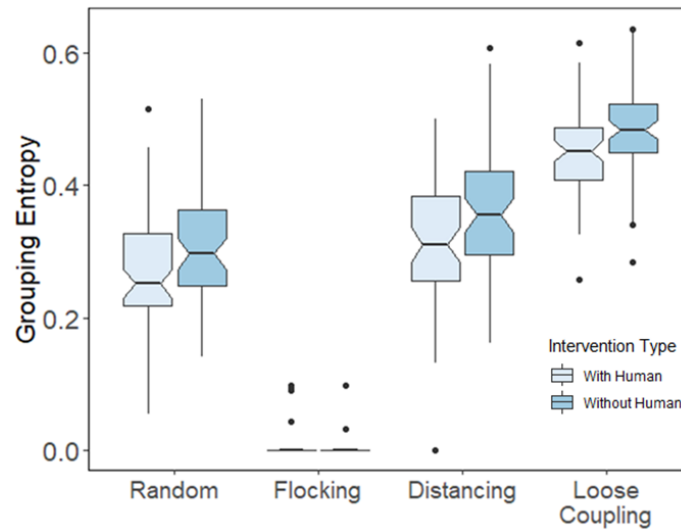


Figure 2.16: Entropy values for simulated agents as a function of movement type (Random, Flocking, Distancing, and Loose Coupling) in the experiment (With Human) versus the simulation (Without Human).

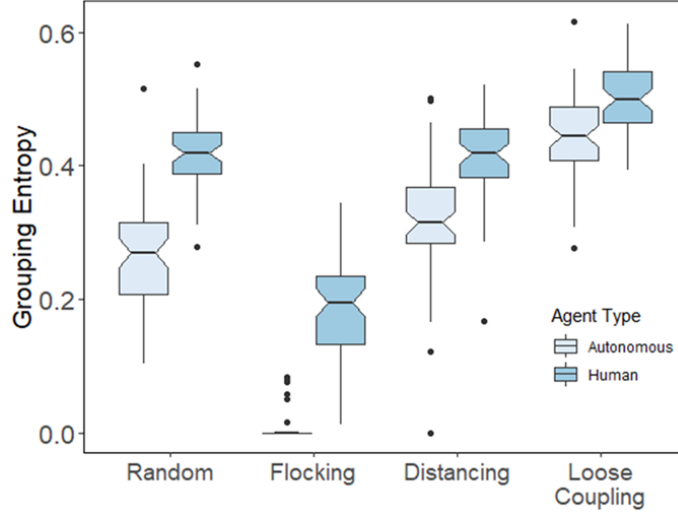


Figure 2.17: Entropy values by movement condition (Random, Flocking, Distancing, and Loose Coupling) for human players (Human) versus simulated agents (Autonomous) in the experiment.

We conducted a mixed-effects ANOVA with movement condition as a within-subjects factor, human intervention as a between-subjects factor, and entropy as the dependent variable. First there was a significant main effect of human intervention whereby human players caused autonomous agents to exhibit less entropy in their distributions over agents in view, $F(1,59) = 21.43$, $p < 0.001$, $\eta_p^2 = 0.058$; and a marginally significant main effect of movement type, $F(2,59) = 2.41$, $p = 0.091$, $\eta_p^2 = 0.014$. The interaction was non-significant, $F(2,59) = 0.49$, $p = 0.613$, $\eta_p^2 = 0.003$. Individual post-hoc tests confirmed that grouping entropy was highest with loose coupling compared with the distancing and random conditions, $p < 0.001$. Human intervention appeared to decrease grouping entropy for autonomous agents by giving them less time to group by means of converging on targets. This decrease in grouping entropy was evident even in the random condition when humans had no direct effect on agent movements—instead, humans had indirect effects because they helped find and consume targets more quickly, thereby decreasing the time available for agents to converge on targets, leaving them less grouped and more disbursed in general.

In our second analysis, we compared the same grouping entropy measure as before but for human players against grouping entropy for individual simulated agents in the experiment with human intervention (Fig. 2.17). We ran another ANOVA like the previous analysis, but with intervention type replaced by agent type (human or autonomous) as a between-subjects factor, again excluding the flocking condition from movement type. We found that grouping entropy was greater for humans compared with autonomous agents, $F(1,59) = 213.85$, $p < 0.001$, $\eta_p^2 = 0.379$, and grouping entropy was again influenced by the movement condition, $F(2,59) = 8.11$, $p < 0.001$, $\eta_p^2 = 0.044$, with post-hoc tests showing that entropy was greatest with loose coupling,

$p < 0.001$. There was also an interaction such that grouping entropy for human players was more like autonomous agents when the latter were loosely coupled compared with other movement types, $F(2,59) = 16.9$, $p < 0.001$, $\eta_p^2 = 0.088$. Moreover, human movements exhibited the most grouping entropy when coordinating with loosely coupled agents based on our post-hoc comparisons, $p < 0.001$.

In summary, grouping entropy was higher, and performance was better, with loose coupling overall, and with human intervention overall (human intervention lowered entropy for simulated agents, but only as a byproduct of shortening time for them to converge on targets). We infer from the main pattern of results that collective foraging in our simulation benefits from loose coupling between autonomous agents as well as between agents and humans.

Human intervention benefits search performance for non-random agents

Entropy analyses in the previous section showed that human intervention decreased the grouping entropy of autonomous agents, even though performance was generally better with human intervention and with increased grouping entropy. Therefore, it is unclear whether human intervention improved the way that autonomous agents searched, or if humans are simply better searchers and therefore find and consume more targets than autonomous agents.

To test the search performance of autonomous agents themselves, we measured how fast they covered the game space when searching for each next target, and we compared their rates of search area coverage with and without human intervention as a function of movement type. Specifically, area search rate was computed as the number of unique pixels searched on each trial, divided by the time spent searching prior to finding the target, and converted into a percentage of total pixels ($200 \times 200 = 40,000$ pixels).

To test more specifically how human intervention affected autonomous agents, we measured area search rate at both the individual and collective levels for autonomous search agents. For the individual level, search rate was computed per agent and then averaged to gauge how fast each agent covered space separately, whereas for the collective level, search rate was computed for all agents simultaneously to gauge how fast the group covered space collectively. Fig. 2.18 shows area search rates for autonomous agents with and without human intervention (the human is always removed from rate calculations, and as before, trials were excluded when search was terminated by the human player finding the target first), for individual search as well as collective search. We used ANOVA models as in the previous results for grouping entropy, but with area search rate as the dependent measure instead, and flocking was included this time.

We found that human intervention improved both individual search rates and collective search rates, $F(3,59) = 43.09$, $p < 0.001$, $\eta_p^2 = 0.044$, but had no effect on random movement conditions because humans had no direct effect on random agent search movements, all post-hoc tests $p > 0.95$. The benefit of intervention was greater for collective versus individual area search rates, $F(3,59) = 17.28$, $p < 0.001$,

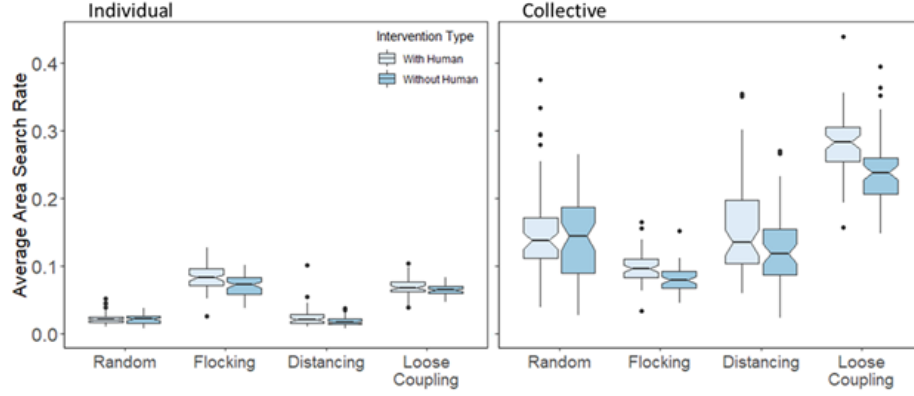


Figure 2.18: Area search rate for autonomous agents, averaged for each session, and plotted as a function of movement type and human intervention for agents individually (left) and collectively (right).

$\eta_p^2 = 0.018$, indicating that human intervention reduced overlap in autonomous agent search areas, as well as reduced the degree to which individual agents returned to areas they already covered. Conducting a post-hoc Tukey HSD analysis found that individual area search rates were not reliably different between the flocking and loose coupling conditions, all $p > 0.7$, but collective search rates were greater with loose coupling compared with flocking, all $p < 0.001$. These results indicate that loose coupling preserved the individual diffusiveness of flocking agents, whose individual grouping entropy was relatively high. By contrast, the addition of distancing helped to reduce agent overlap—collective grouping entropy was highest for loose coupling—and thereby improve collective search performance.

Human search benefits from coordinating with loosely coupled agents

The previous section focused on the beneficial effect of human intervention on the individual and collective search performance of autonomous agents as a function of different movement rules. We can also test whether different movement rules have different effects on human search performance. In theory, human players could search on their own, unresponsive to the movements of other agents. However, to the extent that players try to guide or otherwise coordinate with autonomous agents, the efficacy of human search movements may be affected by the way agents move and coordinate. Results presented earlier showed that human intervention affected autonomous agents via their grouping entropy, and agents affected human players in kind. Given that human players showed the greatest grouping entropy when agents themselves showed the greatest grouping entropy in the loose coupling condition, we can hypothesize that human search performance may benefit from coordination with loosely coupled agents.

To test the effect of movement rules on human search performance, we computed area search rates for the human players individually. Not surprisingly, as shown in

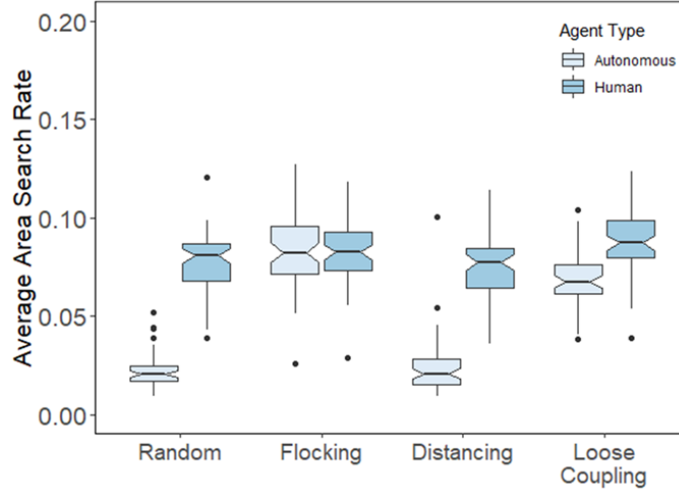


Figure 2.19: Area search rate averaged for each session and plotted as a function of movement type and agent type for human players and individual agents with intervention.

Fig. 2.19, humans overall covered the search space at a faster rate than the individual agents they foraged with, $F(1,59) = 536.04$, $p < 0.001$, $\eta_p^2 = 0.534$. The exception to this overall effect was in the flocking condition where humans searched at about the same rate as agents they were coordinating with. Individual area search rates were relatively high for flocking agents because maintaining a single, steady bearing is a reasonably good strategy for covering a space with periodic boundary conditions.

We also found marginal differences in human area search rates depending on the movement rules governing autonomous agents, $F(3,59) = 2.57$, $p = 0.055$, $\eta_p^2 = 0.032$. As predicted, a post-hoc analysis showed human search performance was better with loosely coupled agents compared with distancing ($p < 0.001$) and random ($p < 0.01$) movement conditions. Although no other comparisons were reliable, Fig. 2.19 shows that human search rates were also elevated when coordinating with flocking agents, which may be attributable to humans sometimes following the flock as they search at a faster rate than agents in the random and distancing conditions.

Taken together with results from the previous section, we can conclude that human players and loosely coupled agents benefitted from each other to improve search performance by virtue of flexibly coordinated movement patterns, as evidenced by higher area search rates along with higher values of grouping entropy.

Human players adapt their foraging strategies to agent behaviors

To this point, we have interpreted the result that human players improve collective foraging by means of memory and strategy, but the evidence has not been direct. It is difficult to infer specific strategies from game play data alone, but one apparent choice that players can make in collective foraging is the emphasis on finding versus consuming targets. Players may try to find targets with other agents following or

not, or they may instead seek out other agents to collectively consume each target so the next one comes faster. Human players may improve collective foraging in part by adapting their emphasis on finding versus consuming targets based on the rules governing movements of autonomous agents.

To measure the emphasis on finding versus consuming targets, we analyzed the proportion of targets found versus consumed by human players as a function of movement type. As a baseline, if humans are no better than their autonomous counterparts, then they should find targets 10% of the time (0.1 proportion of times) and consume 10% of the target units (recall that each target consisted of 500 consumption units), given that the human player is one of ten foraging agents. The difference between finding and consuming proportions is a measure of the emphasis that human players placed on one versus the other component of collective foraging.

Fig. 2.20 plots the two proportions for human players as a function of movement type. First one can see that both proportions were significantly above 0.1 in all movement conditions, all post-hoc tests $p < 0.001$. Greater-than-chance proportions are evidence that at least some benefit of human intervention for collective foraging comes from the superiority of human players, in that they both find and consume more targets than autonomous agents. Post-hoc analysis showed this benefit was reliably less in when agents were loosely coupled compared with other movement types, all $p < 0.001$, because loose coupling was the most effective movement rule.

Regarding adaptations in strategy, Fig. 2.20 indicates that the differential between proportions varied as a function of movement type, $F(3,59) = 5.74$, $p < 0.001$, $\eta_p^2 = 0.035$. Specifically, players emphasized consuming targets over finding targets when coordinating with loosely coupled agents, $t(59) = 4.96$, $p < 0.001$, whereas players did the opposite when coordinating with flocking agents, $t(59) = -11.43$, $p < 0.001$. This contrast is evidence that players adapted their strategy to the movement rules for agents—it was more beneficial for players to help with search when the collective area search rate was relatively low in the flocking condition, whereas it was more beneficial for players to help with consumption when collective area search rate was relatively high in the loose coupling condition.

Discussion

The main goal of this study was to investigate the benefit of balancing individual and collective modes of interaction to succeed in social foraging. We designed an agent-based model in which finding targets benefitted from individual agents diffusing across the search space, whereas consuming targets benefitted from convergence of agents on target locations. We formulated movement rules that balanced a tendency towards distancing with a tendency towards flocking, and simulations showed that loose coupling among agents was beneficial to group foraging performance. Analyses of grouping entropy showed that loose coupling diversified the range of more individuals to more collective search configurations because flocks of various sizes formed and dissolved as search unfolded.

Our simulation results show that loose coupling is beneficial even when agents

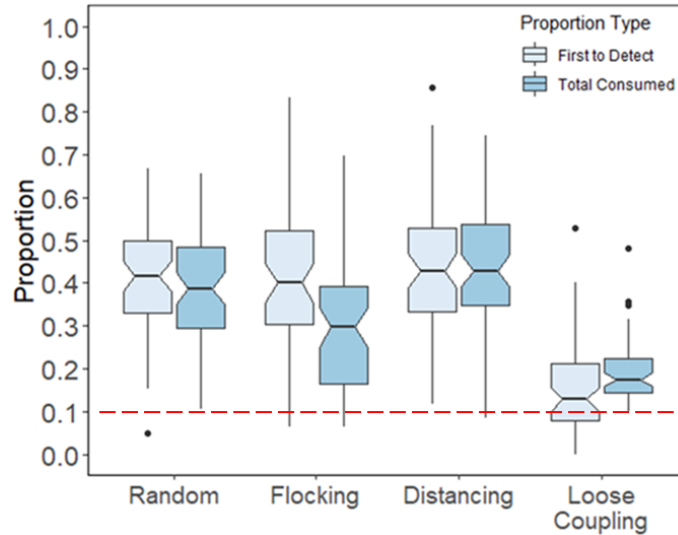


Figure 2.20: Finding and consuming proportions for human players as a function of movement type. The red line indicates the expected proportion to be found and consumed if human performance was no better than that of simulated agents. Proportions above the expected baseline indicate the degree to which humans outperformed simulated agents.

are memoryless and unable to learn, adapt, or develop strategies through experience. The diversity of search patterns did not come from decision-making of any kind—it was instead driven primarily by injecting noise with the CRW movement component, plus additional randomness from positioning of targets. Our simple model of loose coupling is useful in its economy of mechanism and may be appropriate for collections of simple organisms and artificial agents with minimal capacity for computation. However, foraging is necessary for the survival of all mobile species, including for humans and other social animals with extensive capacities of memory, learning, and strategy to drive decision-making as foraging unfolds. To test how decision-making capacities might interact with simple rules of loose coupling, we compared simulations of autonomous agents with and without human intervention.

Human players exhibited an even greater diversity of search configurations than loosely coupled agents, and both humans and autonomous agents covered the search area at faster rates in the loose coupling condition compared with too much distancing or too much flocking. These results provide evidence that abilities like learning and memory may complement simpler rules of loose coupling to support social foraging, rather than supplant them. It is difficult to determine whether movements were controlled based on strategies learned through experience, but we did find evidence that players adapted their strategies to coordinate with agents differently depending on their search behaviors. Human players emphasized finding and consuming targets individually when agents distanced too much or flocked too much, whereas they coordinated more with simulated agents were loosely coupled. This finding suggests

that intelligent agents can learn to leverage other agents for the good of the group, depending on their abilities and performance. Humans varied in performance and behavior from one player to the next (individual differences are universal to human behavior), but more information is needed about the backgrounds and intentions of human players to understand these individual differences.

Our agent-based model proved useful for demonstrating the benefits of loose coupling and adaptive foraging strategies, but future studies could undertake more thorough analyses of the model and its parameters to understand which aspects are most important for loose coupling in social foraging. Also, while the distancing and flocking rules combined to produce loose coupling, they were not quite complementary on their own. Flocking had the desired effect of longer search times offset by shorter consumption times, relative to random search, but distancing did not have the reverse effect—instead, distancing did not have an appreciable effect beyond noise from the CRW rule, although it synergized strongly with the flocking rule. Future studies may consider a different form of distancing that has the opposite effect of flocking, i.e. shorter search times offset by longer consumption times. Future work may also consider implementing measures of grouping and movement coherence [102]. Such measures that help to identify to the effect of movement conditions on the velocity of the whole group.

Finally, it would be informative to study how groups of human foragers coordinate to play our social foraging game. The most salient question is whether players would still exhibit signs of loose coupling in terms of flexibly diverse groupings, and whether group performance would still benefit loose coupling. Theories of self-organization suggest that loose coupling may be generally useful for adapting coordinated behavior to respond to changes in conditions as they unfold [103]. For instance, it may be useful to follow one or more agents when they are first encountered in the hopes of finding targets and consuming them together, but it may become more beneficial to break from the group and seek new search opportunities as time goes by following the group with no success. In this scenario, loose coupling may enable agents to enact a “stay-or-go” decision between exploiting nearby agents or exploring new opportunities, similar to the stay-or-go decision at the heart of optimal foraging theory [104; 105; 106]. The agent-based modeling and experimental paradigm introduced herein may be extended to investigate these and other questions about individual and collective foraging.

2.5 A Case Study: Swarm Robotic Source Seeking with Fractional Fluxotaxis

Source seeking, in general, is a process that involves a searcher that seeks to maximize some unknown scalar valued function $f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^d \rightarrow \mathbb{R}$ that can be accessed via a measurement operator $h(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^d \rightarrow \mathbb{R}$. Typically, the output $y = h(\mathbf{x}) = f(\mathbf{x}) + \eta(t)$ is used to construct a gradient $\nabla f(\mathbf{x})$ or estimate of the gradient $\nabla \hat{f}(\mathbf{x})$, where $\eta(t) \sim \mathcal{N}(\mu, \sigma)$ is a Gaussian white noise process with mean μ and standard deviation

σ . In some cases, such as with extremum seeking control (ESC)[107], the gradient is not explicitly calculated but rather estimated through a high-pass filter and sinusoidal perturbation. For example, this can be seen in nature by foraging insects [108] or in the homing of salmon and turtles [109; 110]. Typically single agent source seeking is employed for the simple cases. In many of these cases, different types of strategies are used that only focus on measurements of one parameter (e.g. concentration with chemotaxis, wind with anemotaxis, etc.). Heuristic methods that rely on some arbitrary yet practical rules for updating agent movement after measurements occur, such as zig-zag (e.g. casting) or spiral-surge [111]. However, for more complex cases, multi-agent approaches have been employed. For example, the source seeking problem has been looked at from an optimization point of view, such as with particle swarm optimization (PSO) [112; 113] or with Optimotaxis [114]. The Optimotaxis algorithm utilizes a random jump-diffusion framework by designing a run and tumble controller that converges to some shaping function [114]. The PSO algorithm utilizes a set of candidate solutions, called particles, whose position and velocity are updated via their respective best position (or solution) as well as the global best position. The performance of PSO depends on the parameter selection and the complexity of the function map or field. In the simplest case, the field f is smooth and continuous. However, in more extreme cases the field f can be turbulent and consequently more complex to solve or be represented in analytical form, restricting access to f . Many bio-inspired approaches were developed to address this issue, such as [115; 116; 117] to name a few. One potential advantage of having access to the spatial-temporal data of the multi-agent group is that it can be used to increase performance of the source seeking method. One such approach is fluxotaxis [100], a physics-based approach that utilizes the gradient divergence of the mass flux (GDMF) by using individual and neighboring observations. This can be described as a type of gradient descent algorithm. More recently, several advancements on fractional order gradient descent (FOGD) algorithms have been developed and compared in [118]. It was shown that FOGD methods have robustness to the extreme value point as well as improvements in convergence speed for $\alpha > 1$. FOGD methods have also been compared in [119] to accelerated gradient descent methods, such as Nesterov accelerated gradient descent (NAGD) and momentum gradient descent (MGD) which was shown to be more robust to the target function or equivalently the step size to the optimal location.

In this work, we introduce a decentralized time fractional order fluxotaxis (FO-fluxotaxis) scheme to approach the source seeking problem using the plume odor simulation (POSIM) framework [120; 1]. In section (2.5.1) we discuss the preliminaries of this work, section (2.5.2) we introduce FO-fluxotaxis, section (2.5.3) we detail the simulations we explored, section (2.5.4) we analyze the results, and in section (2.5.5) we provide some conclusions and discuss future work.

2.5.1 Preliminaries

Fractional Calculus

Fractional Calculus is a type of calculus that deals with differentiation and integration of non-integer orders. This idea dates back as early as 1695 when, in a letter to Leibniz, L'Hôpital asked what if $n = 1/2$, referring to Leibniz's d^n/dt^n notation. Today we have several definitions for fractional integration and differentiation operators, of which, we focus on two. The first, being the Riemann-Liouville definition for fractional order integration given as,

$${}_a^{RL}D_t^{-\alpha}f(t) = \frac{1}{\Gamma(\alpha)} \int_a^t \frac{f(\tau)d\tau}{(t-\tau)^{1-\alpha}} = \frac{t^{\alpha-1}}{\Gamma(\alpha)} * f(t), \quad (\alpha > 0). \quad (2.14)$$

The Riemann-Liouville fractional derivative defined as an integer order derivative of a fractional order integral,

$${}_a^{RL}D_t^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_a^t \frac{f(\tau)d\tau}{(t-\tau)^{\alpha+1-n}}, \quad (n-1 < \alpha < n). \quad (2.15)$$

The second, is the Caputo fractional derivative of order α , written as,

$${}_a^CD_t^\alpha f(t) = \frac{1}{\Gamma(\alpha-n)} \int_a^t \frac{f^{(n)}(\tau)d\tau}{(t-\tau)^{\alpha+1-n}}, \quad (n-1 < \alpha < n), \quad (2.16)$$

where $\Gamma(\cdot)$ is the usual Gamma function. In the limit as $\alpha \rightarrow n$ the Riemann-Liouville and Caputo definitions of the fractional derivative becomes the n -th derivative $f^{(n)}(t)$ [121]. Let us recall the Laplace transform given by,

$$\tilde{f}(s) = \mathcal{L}\{f(t) : s\} = \int_0^\infty e^{-st}f(t)dt. \quad (2.17)$$

The Laplace transform of a fractional derivative or integral can be shown to be,

$$\mathcal{L}\{{}_a^{RL}D_t^\alpha f(t)\} = s^\alpha \tilde{f}(s) - \sum_{k=0}^{n-1} s^k \left[{}_a^{RL}D_t^{\alpha-k-1} f(t) \right]_{t=0}, \quad (2.18)$$

$$\mathcal{L}\{{}_a^CD_t^\alpha f(t)\} = s^\alpha \tilde{f}(s) - \sum_{k=0}^{n-1} s^{\alpha-k-1} f^{(k)}(0), \quad (2.19)$$

where the difference in the two definitions can be seen clearly in the initial conditions. Thus, the Caputo definition enables the use of integer order initial conditions which provides a practical approach for implementation.

Fluxotaxis Strategy

The fluxotaxis strategy is a multi-agent approach that utilizes information from each agent to approximate the direction of the source by estimating the gradient of the

divergence mass flux, **GDMF**. To control the formation of the agents and provide a way for seamless obstacle avoidance, a generalized version of the Lennard-Jones potential is used [100] to generate a formation control vector,

$$F_{formation} = \mathbf{n}_F = c_1 \frac{\mathcal{F}(D)^{c_3}}{r^{c_5}} - c_2 \frac{\mathcal{F}(D)^{c_4}}{r^{c_6}}, \quad (2.20)$$

$$\mathcal{F}(D) = \left(\frac{c_1 D^{c_6 - c_5}}{c_2} \right)^{\frac{1}{c_4 - c_3}}, \quad (2.21)$$

where hyperparameters are chosen for simplicity to be the same as [100] ($c_1 = c_2 = c_3 = 1$, $c_4 = c_5 = 2$, and $c_6 = 3$), such that $F_{formation} = \mathbf{n}_F = \mathcal{F}(D)/r^2 - \mathcal{F}(D)^2/r^3$ and $\mathcal{F}(D) = D^{1/2}$. The variable D represents the desired separation distance between the agents.

Given some closed surface S which encloses a control volume W and some wind vector field $\mathbf{V} = [U, V]^T$, the divergence theorem states,

$$\int_W \nabla \cdot (\rho \mathbf{V}) dW = \oint_S (\rho \mathbf{V}) \cdot d\mathbf{S}, \quad (2.22)$$

where W is the control volume and S is the bounding surface [122]. For a point source the divergence mass flux is given as,

$$\nabla \cdot (\rho \mathbf{V}) = \lim_{W \rightarrow p} \oint_S \frac{(\rho \mathbf{V}) \cdot d\mathbf{S}}{|W|}, \quad (2.23)$$

such that, a positive value results from a source, conversely, a negative value a sink. The gradient can then be given as,

$$\mathbf{GDMF} = \nabla(\nabla \cdot (\rho \mathbf{V})) = \nabla \left(\rho \frac{\partial U}{\partial x} + U \frac{\partial \rho}{\partial x} + \rho \frac{\partial V}{\partial y} + V \frac{\partial \rho}{\partial y} \right). \quad (2.24)$$

Given a centralized virtual swarm structure (i.e. fixed spatially), the **GDMF** can be approximated using finite differences. However, decentralize methods provide more flexibility given obstacles or disturbances in the flow. The decentralized scheme involves calculating the i -th agent's relative flux between the j -th neighboring agents given by [100] is,

$$\mathbf{GDMF}_i \approx F_{ij} = \rho_j |\mathbf{V}_j| \cos \theta_r = \rho_j \frac{\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|} \cdot \mathbf{V}_j. \quad (2.25)$$

The angle between $\mathbf{V}_j$ and separation vector, $\mathbf{r}_{ij}$, is given by θ_r . The update rule for the fluxotaxis algorithm becomes,

$$\frac{d\mathbf{v}_i}{dt} = \mathbf{n}_F + \mathbf{n}_{ij}, \quad (2.26)$$

$$F_{ij}^* = \begin{cases} \arg \min_{\forall j \in \Omega} F_{ij}, & \text{if } \exists (F_{ij} < 0) \\ \arg \max_{\forall j \in \Omega} F_{ij}, & \text{else,} \end{cases} \quad (2.27)$$

where $\mathbf{v}_i$ is the velocity vector control of the i -th agent, and $\mathbf{n}_{ij} = \mathbf{r}_{ij}/|\mathbf{r}_{ij}|$ points in the direction of F_{ij}^* . If no agents are around or no agent detects a concentration, $\mathbf{n}_{ij} \rightarrow 0$.

When no flux is detected, other strategies can be deferred to. The secondary source seeking strategies we consider are: chemotaxis, anemotaxis, and casting [123]. In chemotaxis, the agent is driven in the largest gradient between the i -th and j -th local neighbors. In anemotaxis, the agent will be driven upwind if it detects any chemical concentration. In casting, a back and forth sweeping in 45° paths are executed to help localize the plume. The fluxotaxis strategy uses each of secondary strategies in the order that they appear above depending on the agent's available information. For instance, if no neighboring information is available fluxotaxis will revert to anemotaxis. If no concentration information is available anemotaxis will revert to casting. Additionally, if there is no fluxotaxis information (i.e. $F_{i,j}^* = 0$), fluxotaxis will revert to chemotaxis. If no concentration information is available chemotaxis reverts to casting.

2.5.2 Methods

In this section we introduce the FO-fluxotaxis strategy into the decentralize fluxotaxis framework [100; 122] by modifying the update equation (given $\alpha > 0$),

$${}_a^C D_t^\alpha \mathbf{v}_i = (\mathbf{n}_{ij} + \mathbf{n}_F). \quad (2.28)$$

Under the assumption that the initial conditions are zero, we can rewrite (2.28) using the Laplace transform,

$$s^\alpha \tilde{v}_i = \tilde{\mathbf{n}}_{ij} + \tilde{\mathbf{n}}_F. \quad (2.29)$$

For practical reasons, our update equation should be in terms of acceleration, otherwise it would be unclear how to drive the agents.

$$s \tilde{v}_i = s^{1-\alpha} (\tilde{\mathbf{n}}_{ij} + \tilde{\mathbf{n}}_F) = s^\beta (\tilde{\mathbf{n}}_{ij} + \tilde{\mathbf{n}}_F), \quad (2.30)$$

where $\beta = 1 - \alpha$. Then, the position update for $\beta > 0$ given the α -order is,

$$\frac{d\mathbf{v}_i}{dt} = {}_a^C D_t^\beta (\mathbf{n}_{ij} + \mathbf{n}_F), \quad (2.31)$$

where $\mathbf{n}_F$ is defined as in (2.27). When $\beta < 0$, fractional integral is taken in the Riemann-Liouville sense. The fractional order derivative can be realized as an infinite impulse response (IIR) digital filter based on power series expansion using trapezoidal (Tustin) rule [124]. The accuracy of the estimation depends on the n -th order of the filter, hence the memory of the signal. From [125], $n \geq 20$ is sufficient, however, for the time-scales of the dynamics we are interested in, we extended $n = 100$, which is equivalent to 5 seconds of memory (simulated at 20Hz). Techniques like this are also used in fractional order control [126].

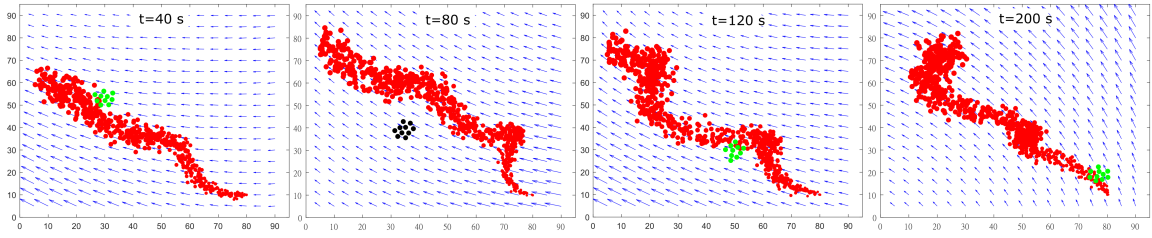


Figure 2.21: An example of the FO-fluxotaxis algorithm given a meandering plume ($a = b = 0.04$ and $G = 5$). The color of the source seeking agent represents the source seeking strategy (green is FO-fluxotaxis). When the plume signal is lost and neighboring information is available, the algorithm reverts back to the secondary source seeking strategy casting (shown in black).

2.5.3 Simulation

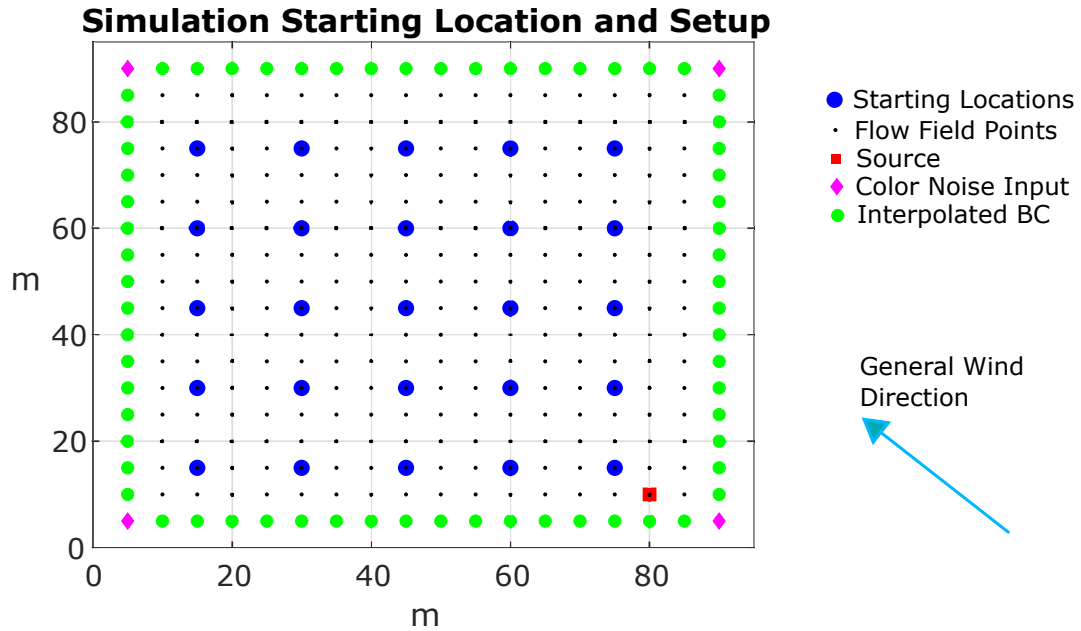


Figure 2.22: Simulation agent starting locations (blue), general wind direction $\bar{\mathbf{V}}$ (light blue), and source location $\mathbf{x}_s$ (red) for both simulation cases.

In this work we conduct simulation experiments to test the performance of the FO-fluxotaxis strategy given different α values using POSIM [120] (a filament-based short-timescale plume model).. The FO-fluxotaxis is compared with the fluxotaxis, anemotaxis, chemotaxis, and casting strategies under different wind conditions driven by colored white noise.

The differences between the simulation experiment in [100] and this work, are

in the starting location of the agents. This work releases the agents from a grid-like structure within the region where the mean wind field is crossing (see Fig. 2.22) rather than 200 random locations, which would require many runs to smooth out the stochasticity. Additionally, we do not consider the expansion or contraction of the agents (as seen in [100]) to try and focus on the performance of the strategy and not the spatial dependence of the individual observations. Given the 90m x 90m search area, each simulation was run for $T = 200$ seconds to provide enough time for most strategies to converge.

In the simulation experiment, 36 plumes were generated based on different wind conditions given a single source starting location $\mathbf{x}_s = [80, 10]^T$ and fixed wind direction at -45° from the vertical. The 36 different plumes are permutations from setting $\bar{\mathbf{V}} = [-1, 1]^T$, $[-2, 2]^T$ and $[-4, 4]^T$, colored noise filter parameters $[a, b, G] = [0.1, 0.1, 1]$ (no meander), $[0.5, 0.1, 3]$ (small meander), $[0.4, 0.4, 5]$ (large meander), and $[0.4, 0.4, 10]$ (larger meander), and the intermediate-scale $\sigma_\psi = 2, 4$, and 6 (see [120] for more implementation details). There are 25 swarm starting locations in a grid pattern inside the plume from $15 \leq x \leq 75$ and $15 \leq y \leq 75$ (see Fig. 2.22). The swarm sizes vary from 7 to 70 agents. The values of α are varied from 0.8, 0.9, 1.1, and 1.2. The total number of trials run are 9000.

To check the performance of each strategy several performance metrics were considered, namely:

1. *Arrival time* - t_a : The time it takes the any one of the swarm agents to reach the source, within a small positive value $\xi > 0$. Here we use $\xi = D$;
2. *Distance travelled* - L_D : Is the time integrated distance traveled during the simulation, normalized by the distance to the source given the initial starting location from time $t = 0$ to $t = t_a$ (or $t = T$);
3. *Closeness* - L_C : A length measure of how close the strategy came to the source, normalized on the initial and final time, T , given by $L_c = 1 - \frac{|\mathbf{x}(T) - \mathbf{x}_s|}{|\mathbf{x}(0) - \mathbf{x}_s|}$.

For the cases to which the strategy converges to the source, the *arrival time* t_a gives good performance information. For the cases to which the strategy does not converge to the source, the *Distance travelled* L_D and *Closeness* L_C , gives information to compare the path traveled and ‘how close’ the strategy came, on average. Additionally, the *Distance travelled* metric can be used to give insight on the optimal path travelled when comparing strategies that converged (given a certain set of wind conditions and initial starting location).

2.5.4 Results

The results for simulation experiments can be seen in Figs. 2.23-2.25, where the performance of fluxotaxis and anemotaxis clearly out rival the other source seeking strategies. It can be seen from Fig. 2.23 that the anemotaxis strategy is more susceptible to meandering under smaller swarm sizes. Also, it is interesting to note, that

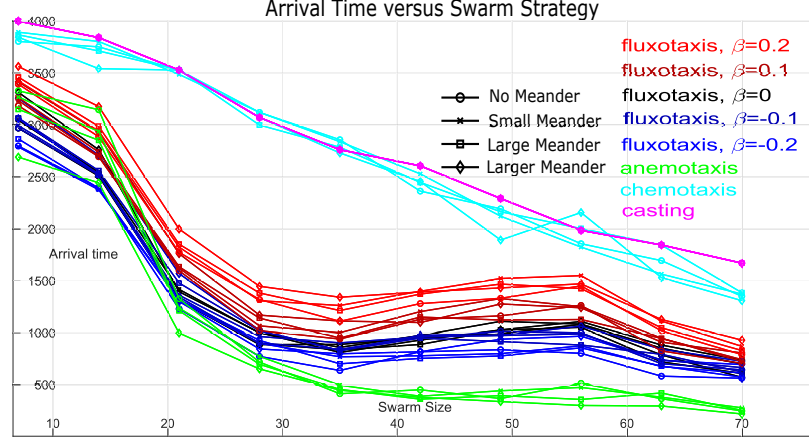


Figure 2.23: Case I: *Arrival time* performance of different swarm strategies given wind conditions driven by colored noise.

the FO-fluxotaxis strategy does improve as $\beta < 0$ and seems to get worse for $\beta > 0$. A possible explanation has to deal with the sub-diffusive behavior of the plume and the sub-diffusive response of the fractional order update scheme. Additionally, the non-local operator may provide a better spatial gradient approximation as opposed to the local integer order one. An interesting phenomenon is shown in Fig. 2.24 where initially the *distance travelled* is small before getting large and then getting small again. This is thought to be related to plume information and probability of detection. For instance, if the agents cover small spatial area as compared to the entire simulation area, there may be some instances where no information is gathered (i.e. no travelling). Furthermore, when some information is gathered but is subject to occasional information losses, the *distance travelled* could be quite large. Given a large enough swarm size the area covered would be sufficient enough to keep the information stable and *distance travelled* low, aside from interactions in between agents. Interestingly enough, as swarm size grows, the *distance travelled* gets worse (aside from the FO-fluxotaxis strategy). Looking at Fig. 2.25 we can see that anemotaxis and FO-fluxotaxis perform quite well after as swarm size increases. For small swarm sizes we can see that when $\beta < 0$, FO-fluxotaxis *closeness* is slightly better.

To understand the performance efficiency of each strategy we devised an efficiency metric, $\mathcal{E} = L_c/t_a L_D$. The *Closeness*, L_c , here represents the success rate like term of the strategy. The *arrival time* and *distance travelled* are things we would like to minimize. The results of this performance efficiency for each strategy can be seen in Fig. 2.26, where it shows the FO-fluxotaxis strategy is more efficient for $\beta < 0$.

2.5.5 Conclusions

In this work we introduced the FO-fluxotaxis scheme to the multi-agent source seeking problem. We conducted a simulation case study based on meandering generated from colored noise. The results from the simulation case study showed that the FO-

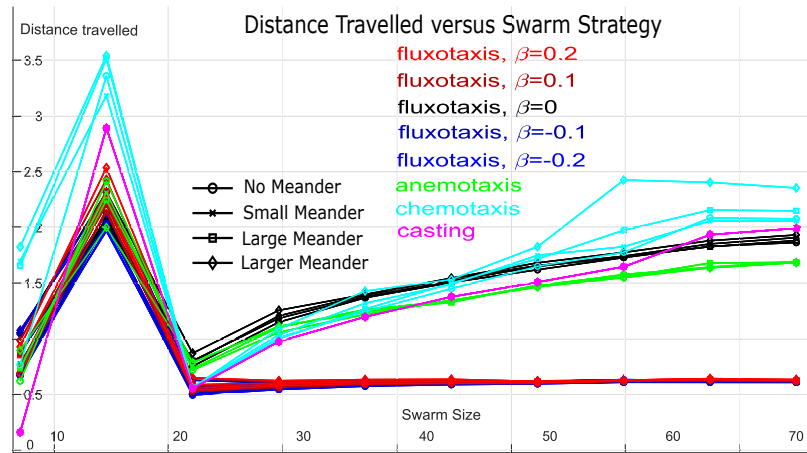


Figure 2.24: Case I: *Distance travelled* performance of different swarm strategies given wind conditions driven by colored noise

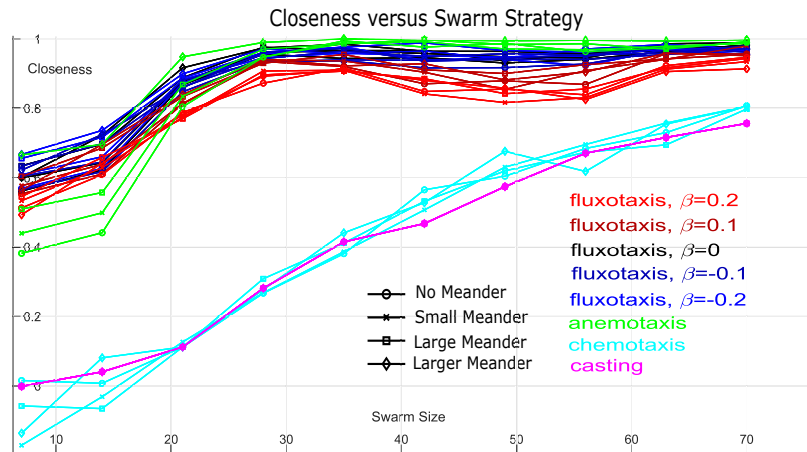


Figure 2.25: Case I: *Closeness* performance of different swarm strategies given wind conditions driven by colored noise.

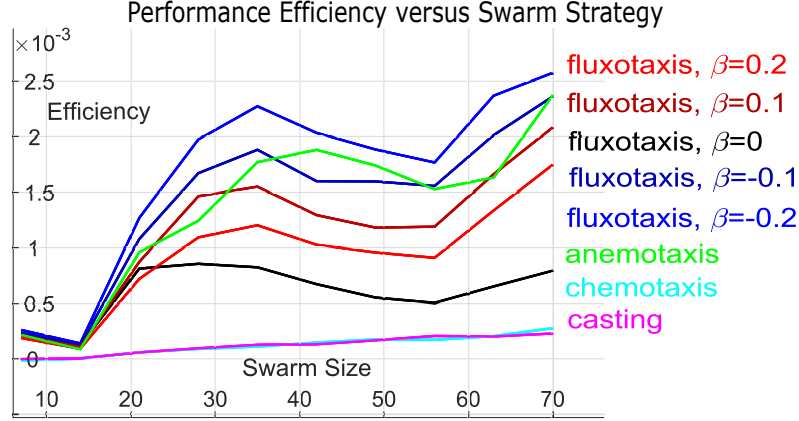


Figure 2.26: The performance efficiency, $\mathcal{E}$, of all the swarm strategies.

fluxotaxis strategy shows good performance at smaller swarm sizes as compared to other strategies. The performance of the anemotaxis strategy outperformed others in arrival time at larger swarm sizes. It was also observed that smaller swarm sizes benefit from larger meandering conditions.

For future research, it is worth exploring the performance response by adding more terms into the approximation into (2.31), for the FO-fluxotaxis update. The inclusion of additional terms is analogous to adding memory within the fluxotaxis strategy, making the agents ultimately smarter. Additionally, exploration into the effect of obstacles on the FO-fluxotaxis strategy as function of memory is needed as well as exploration into the effects of smaller swarm size and desired separation distance. Lastly, the inclusion of entropy based source seeking (such as in entrotaxis [127]) into fluxotaxis strategy should be considered, which could help for searching in regions with more variable change or uncertainty.

2.6 A Case Study: A More Optimal Stochastic Extremum Seeking Control Using Fractional Dithering for A Class of Smooth Convex Function

2.6.1 Introduction

Source seeking is an important topic when it comes to gas and oil industries. Where even small leaks can cause big (and sometimes fatal) problems, such as the 2010 San Bruno incident in southern California. These leaks are sometimes referred to as fugitive leaks and are difficult to resolve. There has been work done to localize leaks over the years. [128] looked at spatially distributed systems and inverted the steady state solution. [114] looked at Optimotaxis using a stochastic switching method

based on *Escherichia Coli* which mimics a chemotactic approach using a run and tumble controller. Other researchers like [129] used an information based approach which uses the a combination of the maximum likelihood principle and the Fischer information matrix to calculate the gradient of the Cramer Rao bound. This points the searching in the direction of the source in a similar way to the flux measurements shown in [130]. Most of the work done using source seeking is approximate or requires heavy computational resources to calculate the analytical solution or numerically simulate. Thus trying source seeking from a control perspective with model free approach seems desirable.

In source seeking the plant information is generally unknown and the measured output of the plant is noisy and sparse. Therefore, these systems are not explicitly convex. Considering that the average system is convex, such as the Gaussian plume model, but contains local extrema due to turbulence or some other perturbation. One can possibly leverage convex optimizers such as the extremum seeking control (ESC) to minimize some cost function J for the system.

In this case study we look at fractional dithering into the ESC framework by introducing fractional Gaussian noise (fGn) and Symmetric alpha stable (S α S) as a stochastic perturbation, organized as follows: section 2.6.2, overviews of ESC framework; section 2.6.3, covers the definition of fractional Gaussian noise; section 2.6.4, class of convex functions; section 2.6.5, goes over averaging principle and the non-robust stability analysis; section 2.6.6, simulation; section 2.6.7, results and performance; section 2.6.8 discusses conclusions and future work.

2.6.2 ESC Overview

The idea of ESC has been shown as early as [131] on electric railways. There are many different types of ESC including; Analog optimizers, Numerical optimization, Sliding mode control, and adaptive control ([132]). These variations can even include state machines. In this paper we are focusing on a specific type of perturbation based ESC called stochastic ESC (SESC). The ideas of SESC can be found in the works of ([133; 134]) and under constraints ([135]). Even under a discrete setting ([136])

The principle of ESC relies on the minimization of a cost function given some dynamical system. The cost function takes the output of the plant as an input and its output exploited using a perturbation strategy. This involves taking the fluctuations of the cost function by way of high pass filter and multiplying the perturbed control signal to it. This high pass filter can be thought of as a taking the derivative of the signal as to explore the slope of the cost function. The multiplication then will give rise to a mainly negative or mainly positive signal. Typically, a deterministic sinusoidal signal is used as the perturbation (such as $\nu = A \sin \omega t$ in Fig. 2.27). Here we do not explicitly use a cost function.

[135; 133] introduced stochastic probing signals into ESC ($\nu = \eta$, in Fig. 2.27) due to the bounded nature of the sine function and the possibility to restrict the algorithms region of attraction. The nature of the sinusoidal perturbations require

the mean to be zero and variance to be positive. Using Gaussian white noise this is preserved. Additionally the stochastic ESC can be bounded as $\nu = A \sin \eta$ in Fig. 2.27.

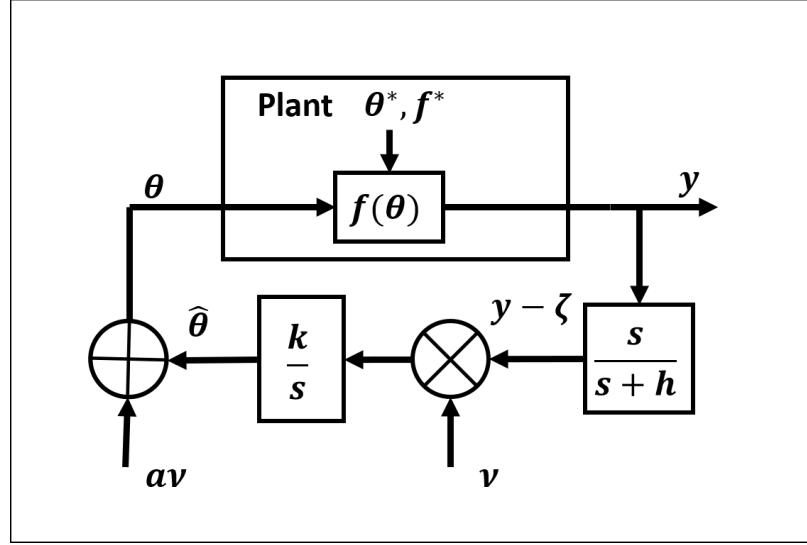


Figure 2.27: Control diagram of ESC.

2.6.3 Dithering Noises

In digital audio, dithering refers to the low level noise added to reduce errors during changing of bit depth. The word dither means “nervous vibration” and can be used to improve reduced images as well. In the context of ESC the stochastic perturbations can be seen as a type of dithering. However, not all noise is the same and thus we introduce fractional dithering to improve performance.

Fractional Gaussian Noise

Fractional Gaussian Noise (fGn) can be derived in a similar manner as Gaussian Noise (Gn). By taking the difference between successive steps in Brownian motion (Bm) one can arrive at Gn. The Riemann-Liouville fractional integral can be used to define fractional Brownian motion (fBm) and is given below,

$$B_H(t) = \frac{1}{\Gamma(H + 1/2)} \int_0^t (t - s)^{H-1/2} dB(s).$$

Here the term $dB(s)$ is the general definition of white noise, the term H is the Hurst parameter and $\Gamma(\cdot)$ is the usual gamma function. We can see that depending on Hurst parameter the motion can be,

- Brownian motion with $H = 1/2$
- Positively correlated $H > 1/2$
- Negatively correlated $H < 1/2$.

For $H > 0.5$ the process exhibits long-range dependence such that,

$$\sum_{n=1}^{\infty} E[B_H(1)(B_H(n+1) - B_H(n))] = \infty.$$

We can denote the fGn based on Hurst parameter as, $fGn_H(k) = B_H(k+1) - B_H(k)$ where k here is used as a discrete time step and not the gain used in the control diagram. In our case we use low pass filtered white noise with high cutoff ratio initially. This process is sometimes referred to as an Ornstein-Uhlenbeck (OU) process. Replacing white noise with fGn creates a fractional OU process with new tuning parameter H . To calculate or compute the fractional integral one can use the Gaussian quadrature, Cholesky decomposition method, or circulant embedding by [137]. Observation of fGn by itself may be difficult to abstract the correlations in-between time steps. These characteristics can be quite easily seen by looking at fBm (see Fig. 2.28). The probability density function (PDF) in Fig. 2.29 shows that the step lengths still obey a Gaussian distribution. However, by examining the power spectrum density (PSD) we can observe the persistent behavior of fGn at values of $H \geq 0.5$. If $H < 0.5$ the slope of the PSD line will become positive, indicating that there is much more high frequency behavior (anti-persistent).

Symmetric α -Stable Noise

Alpha stable noise can be written in terms of four key parameters α , β , γ , and b which can change the distribution uniquely. α is called the characteristic exponent and measures the "thickness" of the tails in the distribution. β , is a symmetry parameter that can skew the distribution to the left or right, here we let $\beta = 0$ which results in a symmetric α -stable or SaS. γ is the scale parameter, also called dispersion, and is similar to the variance in the normal Gaussian case when $\alpha = 2$. The last parameter is b and is referred to as the location parameter. It is the mean when $1 < \alpha \leq 2$ and the median when $0 < \alpha < 1$.

The standard SaS density function is given below where $\alpha > 0$.

$$f_{\alpha}(x) \begin{cases} \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} \Gamma(\alpha k + 1) |x|^{-\alpha k} \sin \frac{k\alpha\pi}{2}, & \alpha < 1 \\ \frac{1}{\pi \alpha} \sum_{k=1}^{\infty} \frac{(-1)^k}{2k!} \Gamma(\frac{2k+1}{\alpha}) x^{2k}, & 1 \leq \alpha \leq 2. \end{cases}$$

The mean is defined and zero for $1 < \alpha \leq 2$ and variance is ∞ for $\alpha < 2$. Fractional lower order moments (FLOM) can be used to further analyze SaS for $\alpha > 1$. The p -th order moment of SaS must be less than α otherwise the moment is ∞ . See [138; 139], for more details.

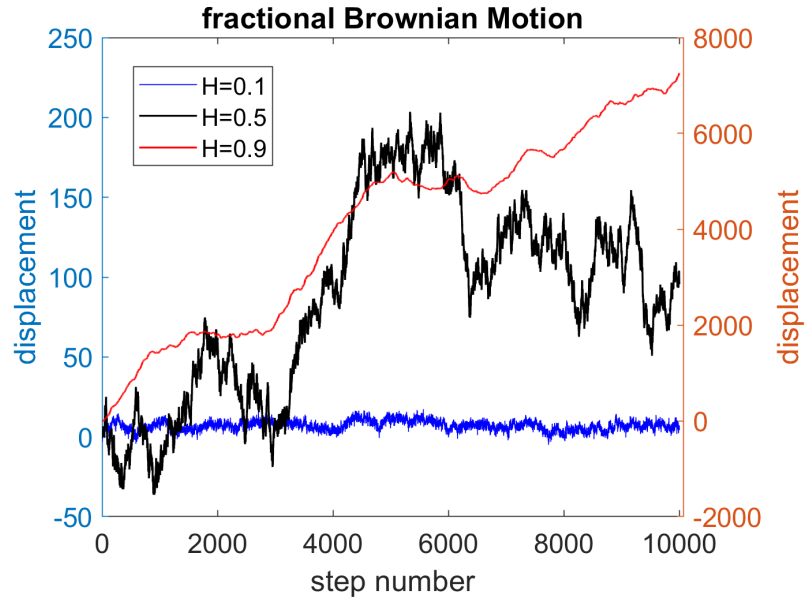


Figure 2.28: Plot of fBm showing (left axis) neutral $H = 0.5$, and anti-persistent behavior $H = 0.1$ and (right axis) persistent behavior $H = 0.9$.

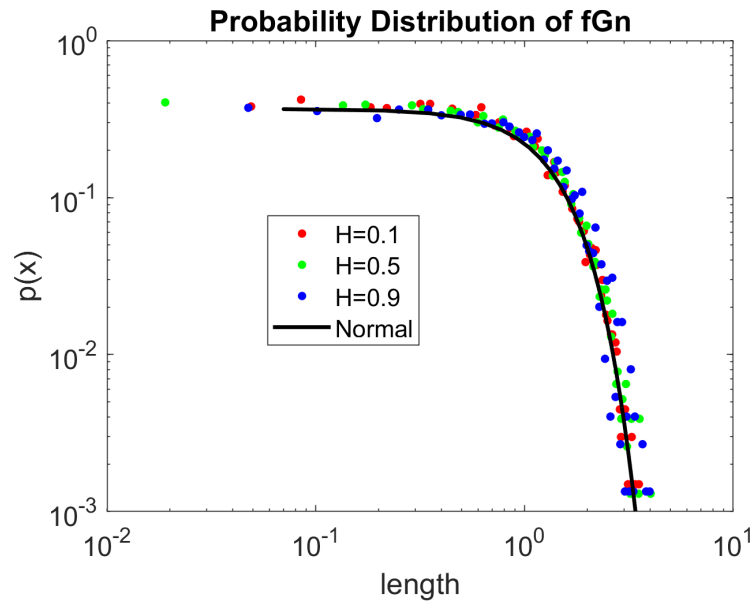


Figure 2.29: Step length distribution for different values of H .

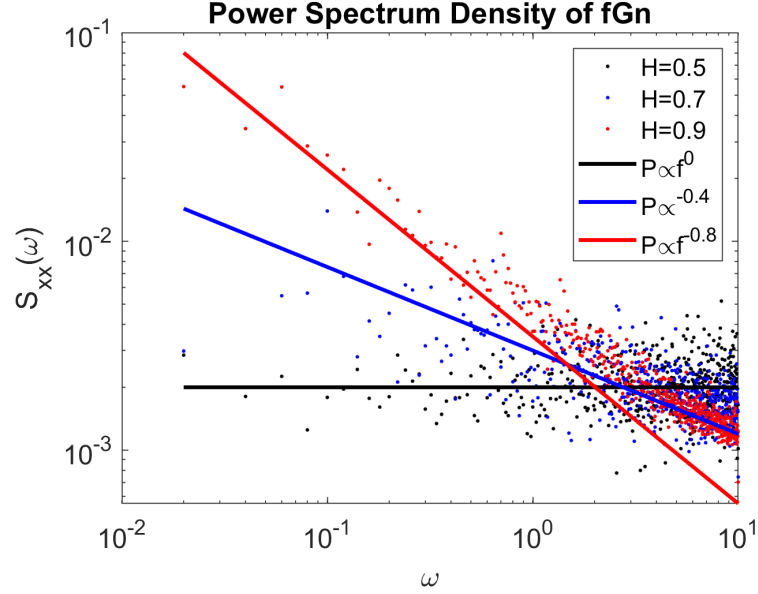


Figure 2.30: The PSD shows the behavior in fGn for $H > 0.5$ which $S_{xx}(\omega) \propto \omega^{-(2H-1)}$.

The PDF of SaS is given in Fig. 2.31. The heavy tail behavior can be observed as they decay at a much slower rate than the traditional normal distribution as α decreases.

2.6.4 Non-smooth Convex Functions

Consider a class of convex functions that can be noisy, contain local minimums and maximums, and is separable into an average function $\bar{f}$ and deviation function $\tilde{f}$.

$$f(\vec{x}, u) = \bar{f}(\vec{x}, u) + \tilde{f}(\vec{x}, u).$$

The function $\bar{f}(\vec{x}, u)$ can be given in the time interval T such that it is convex $\forall \vec{x} \in \Omega$. The expectation of $\tilde{f}(\vec{x}, u)$ is zero and the standard deviation is given as $\sigma_{\vec{x}}$ for some time interval T . The standard deviation σ in this class of functions represents the amplitude of noise or turbulence in the function.

2.6.5 Stability Analysis

Using the averaging principle, as illustrated nicely in [133], the error dynamics of the averaged system can be shown to be exponential stable in the Lyapunov sense.

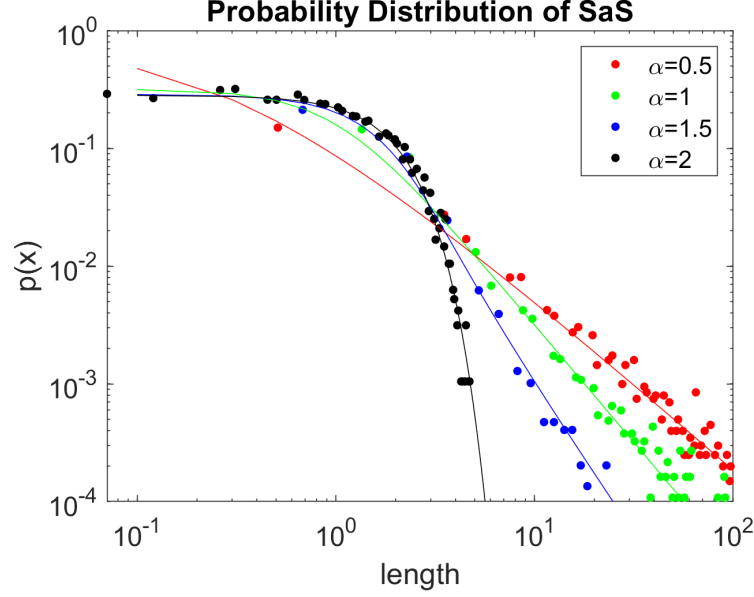


Figure 2.31: The long range dependence can be seen in the heavy tails of the probability distribution function of SaS for various α .

Averaging principle

Given a system with a small positive parameter ϵ and a function $f(x, u)$ that is T -periodic in t .

$$\frac{dZ_t^\epsilon}{dt} = \epsilon f(Z^\epsilon(t+T), \xi_t) = \epsilon f(Z(t), \xi_t), \quad Z_0^\epsilon = x.$$

We make the change of variables $X_t^\epsilon = Z_{t/\epsilon}^\epsilon$ such that,

$$\frac{dX_t^\epsilon}{dt} = f(X_t^\epsilon, \xi_{t/\epsilon}), \quad X_0^\epsilon = x.$$

Considering that the system is bounded, continuous and satisfies global Lipschitz condition the average dynamics can be written as

$$\frac{d\bar{X}_t^\epsilon}{dt} = \bar{f}(\bar{X}_t), \quad \bar{X}_0 = x,$$

given, ξ_t is periodic or the sum of periodic functions. Then by using the Gronwall-Bellman Lemma ([140]) the error dynamics can be written as,

$$|X_t^\epsilon - \bar{X}_t| \leq K \int_0^t |X_s^\epsilon - \bar{X}_s| ds + C(\epsilon)$$

$$\sup_{0 \leq t \leq T} |X_t^\epsilon - \bar{X}_t| \leq C(\epsilon) e^{KT}$$

$$C(\epsilon) \triangleq \sup_{0 \leq t \leq T} \left| \int_0^t [f(\bar{X}_s^\epsilon, \xi_{s/\epsilon}) - \bar{f}(\bar{X}_s)] ds \right|.$$

It can be shown that $C(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$. Thus the averaged system will converge to the original system.

Extension from Gaussian Stability

To show the stability of the stochastic ESC with Gaussian noise, we consider a control law $u = g(x, \theta)$ and smooth function $x = l(\theta)$ s.t. $f(x, u) = f(\theta)$. Defining parameter θ as,

$$\theta(t) = a\eta(t) + \hat{\theta}(t)$$

Here $\eta(t)$ is the stochastic unbounded noise that is added according to the transfer function where h is the high frequency cutoff.

$$\eta(t) = \frac{\sqrt{\epsilon}q}{\epsilon s + h}[Gn].$$

We can define the error of parameter $\tilde{\theta}$ as,

$$\tilde{\theta}(t) = \theta^* - \hat{\theta}(t),$$

where the $\hat{\theta}(t)$ is the estimate of the parameter θ . Thus, the error dynamics then become,

$$\dot{\tilde{\theta}}(t) = -\dot{\hat{\theta}}(t) = k\eta(t)f(\theta(t)).$$

We now need to find the function $f(\theta)$ which can be approximated by using the Taylor series expansion around θ^* .

$$\begin{aligned} f(\theta) &= f(a\eta + \hat{\theta}) \\ &\approx f(\theta^*) + f^{(1)}(\theta^*)(a\eta - \tilde{\theta}) + \frac{1}{2}f^{(2)}(\theta^*)(a\eta - \tilde{\theta})^2. \end{aligned}$$

Given that θ^* is a minimum to $f(\theta)$ then $f^{(1)}(\theta^*) = 0$. Using this finding and expanding the terms in the previous equation we can substitute back into the equation for the error dynamics.

$$\dot{\tilde{\theta}}(t) \approx k\eta(t) \left[f(\theta^*) + \frac{1}{2}f^{(2)}(\theta^*)(a^2\eta^2 - 2a\eta\tilde{\theta} + \tilde{\theta}^2) \right].$$

By regrouping the terms of $\eta(t)$ we can then take the expectation and average the system. The error dynamics can then be broken into three terms: first moment, second moment, and third moment.

$$\begin{aligned} \dot{\tilde{\theta}}(t) \approx k \left[\eta(t)[f(\theta^*) + \frac{1}{2}f^{(2)}(\theta^*)] - \eta^2(t)af^{(2)}(\theta^*)\tilde{\theta}(t) \right. \\ \left. + \eta^3(t)\frac{a^2}{2}f^{(2)}(\theta^*) \right] \end{aligned}$$

$$\lim_{t \rightarrow \infty} E[\eta(t)] = 0, \quad \lim_{t \rightarrow \infty} E[\eta^2(t)] = \frac{q^2}{2}, \quad \lim_{t \rightarrow \infty} E[\eta^3(t)] = 0.$$

We can then define the Lyapunov function $V = \frac{1}{2}\tilde{\theta}^2$ where,

$$\dot{V} = \tilde{\theta}\dot{\tilde{\theta}} \approx -\frac{kaq^2}{2}f^{(2)}(\theta^*)\tilde{\theta}^2. \quad (2.32)$$

Thus in order for the error dynamics to be exponentially stable in the averaged system we need $f(x, u)$ to be convex around the equilibrium point, $f^{(2)}(\theta^*) > 0$. We also need the quantity, $ka > 0$.

Since the fGn has zero mean and variance tends to infinity as $H \in (0.5, 1)$, we can substitute it into η such that $\epsilon d\eta_H = -\eta_H dt + \sqrt{\epsilon} q dB_H$. Consider that for values of $H > 0.5$ the result in (2.32) still holds.

For S α S, the first moment is zero for $1 < \alpha < 2$ and undefined for $0 < \alpha < 1$. Here we assume that it is zero. The third moment is undefined due to the second moment being ∞ . FLOM can be used to characterize S α S but would not be directly applicable to the above analysis. Consider that by truncating the heavy tail of a S α S and normalizing the sample variance we can achieve a computational variance $< \infty$.

2.6.6 Simulation

When given a function map to optimize a minimum or maximum we want the SESC algorithm to be robust to the event where we may get stuck in a local minimum or the map is not sufficiently smooth (i.e. in source seeking problems). Here we explore the use of a static map. However, a dynamic map can be used given the dynamics change very slowly, otherwise, it will invoke internal model principle. Using the Rastrigin function as the function map,

$$f(\vec{x}) = An + \sum_{i=1}^n [x_i^2 - A \cos(2\pi x_i)]$$

gives direct knowledge of the well width, ε , and amplitude, A . In the source seeking framework this can be thought of as the spatial distribution snapshot of a fluid subject to turbulence. We then initialize $\theta(t=0) = \theta_0 = -3$ such that we have multiple extremums before the algorithm can reach the global extremum at $\theta = 0$ (see Fig. 2.32). To understand the characteristics for each type of noise under these conditions we ran 1000 trials per configuration choice of H and α . The values of H and α range from $0.1 < H < 0.9$ and $0.5 < \alpha < 2$.

Parameter Selection

In order to compare the different noises in both the bounded and unbounded case we needed a metric. Using the root mean square (RMS) of a sinusoidal signal ($A_{RMS} = A/\sqrt{2}$) we can scale the amplitude to be proportional to well width, ε .

$$\begin{cases} aA \sin \sigma_\eta / \sqrt{2} \geq \sqrt{2}\varepsilon, & \text{Stochastic Bounded} \\ a\sigma_\eta / \sqrt{2} \geq 2\varepsilon, & \text{Stochastic Unbounded.} \end{cases}$$

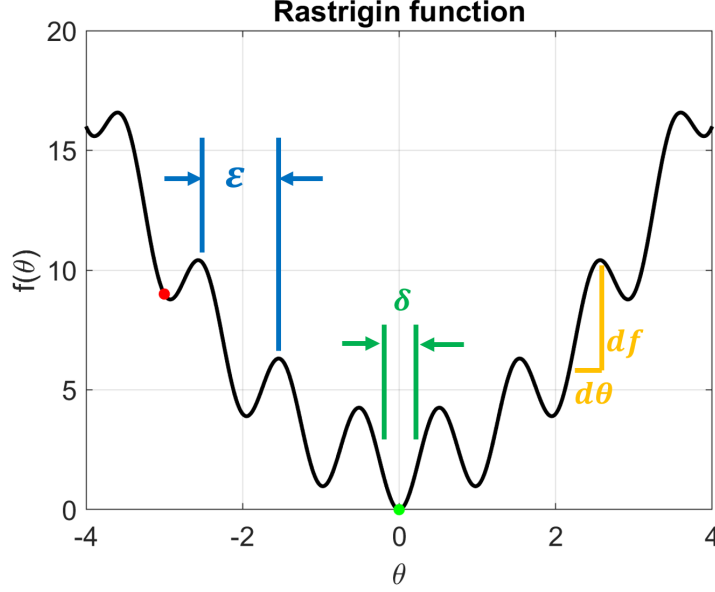


Figure 2.32: Rastrigin function.

Here the term σ_η represents the scaled standard deviation of the generated noise Y to satisfy the above inequality, such that $Y = X\sigma_\eta/\sigma_X$ where X is either fGn or SαS. Due to the periodic behavior of the sine function we chose to have the σ_η for the bounded case to be $\pi/2$ to maximize the output to be $A[-1, 1]$. If the σ_η is chosen too big we see harmonics show up in the results as H or α is changed. Alternatively, if σ_η is chosen too small the output is too small to jump outside of the well width.

The value of k depends proportionally to the inverse of the slope of the function.

$$k \propto \frac{1}{\max_{x_i \in \Omega} \left| \frac{\partial f}{\partial \theta_i} \right|} = \frac{1}{m}, \quad \text{for } i = 1, 2, \dots, n.$$

Here we chose $k = k_m/m$ with $k_m = 20$. The value of a can be used to scale the perturbation on the filtered output and θ . Here we choose it to be 1 for simplicity. However, results can be further improved by tuning a .

The filter parameter ϵ is chosen to be 0.25, where for small ϵ the $\lim_{t \rightarrow \infty} E[\eta^2] = q^2/2$. The gain of the filter becomes $\sqrt{\epsilon}q/h$ and the cutoff frequency is $h/2\pi\epsilon$. For simplicity we choose $h = 5$ and $q = 1$ for all configurations leaving the noise parameters to be optimized.

Performance Metrics

The time convergence t^* is given as,

$$t^* = \min_{t \rightarrow T} t, \quad \text{s.t. } |\theta^* - \hat{\theta}| < \text{tol.}$$

The stopping criteria tolerance is chosen to be 5% of ε . Selection of a small tolerance can mean no convergence. The rigorous proof of [141] in Theorem 5.1 says that there exist a ball of initial conditions and under the assumption that $f(\theta)$ is smooth, $\dot{\theta} = f(\theta)$ is locally exponentially stable for $\theta \in \Omega$ and $f^{(1)}(\theta^*) = 0$ and $f^{(2)}(\theta^*) < 0$ that solution exponentially converges to a neighborhood of order $O(A + \omega + \text{small positive constant})$. Since our A is tuned based on well width ε it seems appropriate to set the tolerance relative to the system not the perturbation.

The δ -neighborhood is defined as the standard deviation of the estimate of θ ,

$$\delta = \sqrt{E[(\hat{\theta} - \theta^*)^2]_{t^* < t < T}}.$$

The reported value of θ estimate is given as,

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^N \hat{\theta}_i, \quad \text{for } i = 1, 2, 3 \dots N,$$

where $\hat{\theta}_i$ is the value of $\hat{\theta}_{t=T}$ and T is the length of the simulation. Here we choose $T = 100$ seconds with time step $\Delta t = 0.01$.

2.6.7 Results

From the performance plots of the SESC with appropriate choices of A and k (see Figs. 2.33-2.38), suggests that the type of noise matters. We can observe that the performance of α decreases around $\alpha = 1.5$ before approaching a similar performance for the $\alpha = 2$ case. Intuitively it seems that this should not be the case. However, when scaling the S α S noise the signal is comprised mainly of low frequency parts with the occasional long jump. Due to the low pass filter the high frequency jumps get attenuated. The result is a more narrow frequency band of jump sizes and because the amplitude of the perturbations affect the SESC performance for jumping from one well to another it makes sense t^* is reduced. For fGn, the anti-persistent behavior results tends to prohibit the well jumps to happen. For $H < 0.4$ and $H > 0.8$ the global convergence is not met (Figs 2.35 and 2.38). For larger values of H (persistent behavior) it can be seen in Figs. 2.34 and 2.37 that the δ -neighborhood grows exponentially. A more optimal selection of noise can be found by minimizing a cost function based on the performance metrics mentioned earlier for $H \in (0.1, 0.9)$ and $\alpha \in (0.5, 2)$,

$$\min J \quad \text{s.t.} \quad J = t^* + \delta/A + \hat{\theta}$$

The optimal value for bounded/unbounded cases of fGn is found to be around $H = 0.63$ in this case. For S α S, the optimal values for bounded/unbounded cases is $\alpha = 2, 0.5$, which for $\alpha = 2 \rightarrow$ Gn. The result for fGn is surprising in that some LRD can improve the convergence behavior. While the result from S α S is not surprising given the effect of filtering. All of the heavy tail behavior is lost due to the type of filter. Also the increase in performance when α approaches 0.5 results in more impulse

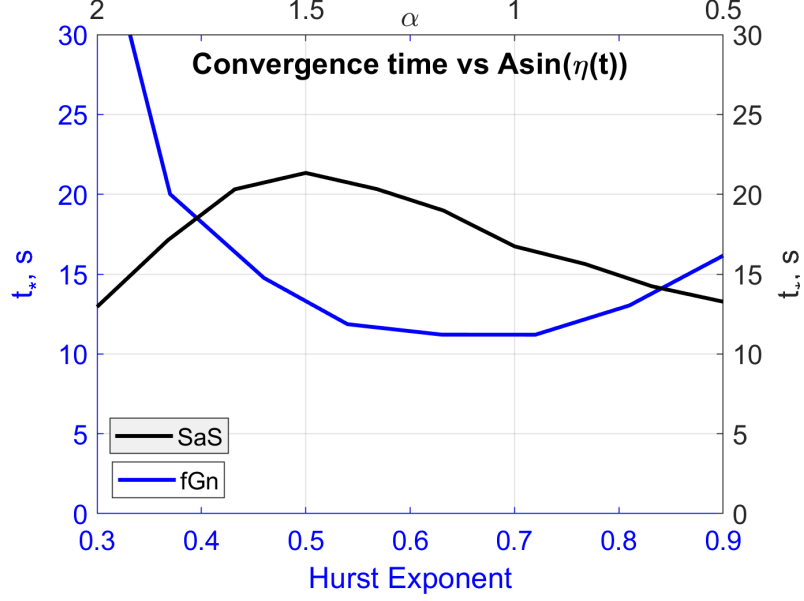


Figure 2.33: The temporal convergence to the global minimum from the initial condition θ_0 , denoted t^* , is shown as a function of fGn and SαS dithering.

behavior. This in turn increases the frequency and in the PESC as ω increases, so does the convergence speed. However, some systems may have bandwidth limitations. This makes the persistent behavior of fGn more desirable.

2.6.8 Conclusion

In this case study we examine SESC with fractional dithering noises from SαS and fGn for a class of smooth convex functions. The stability depends on $f^{(2)}(\theta^*) > 0$, $ka > 0$ and the second moment of η (i.e. choice of H or α). We found that:

- when parameter selection is based on the well width and the inverse of the maximum gradient of $f(\theta)$. The SESC will converge to the global extremum more consistently.
- the presence of fractional noise can improve convergence in the presence of local extrema.
- A more optimal SESC can be achieved with $H = 0.63$, in this example.

Future work will reconsider SαS and fGn under fractional order filtering and explore tempered Lévy distributions ([44]) as well as examine non-static functions.

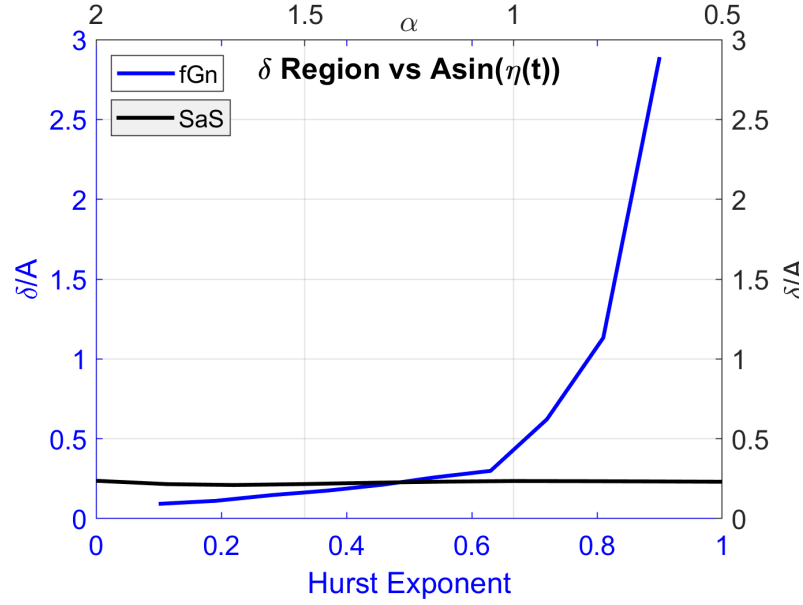


Figure 2.34: The spatial deviation δ around the converged θ is shown as a function of fGn and SaS dithering.

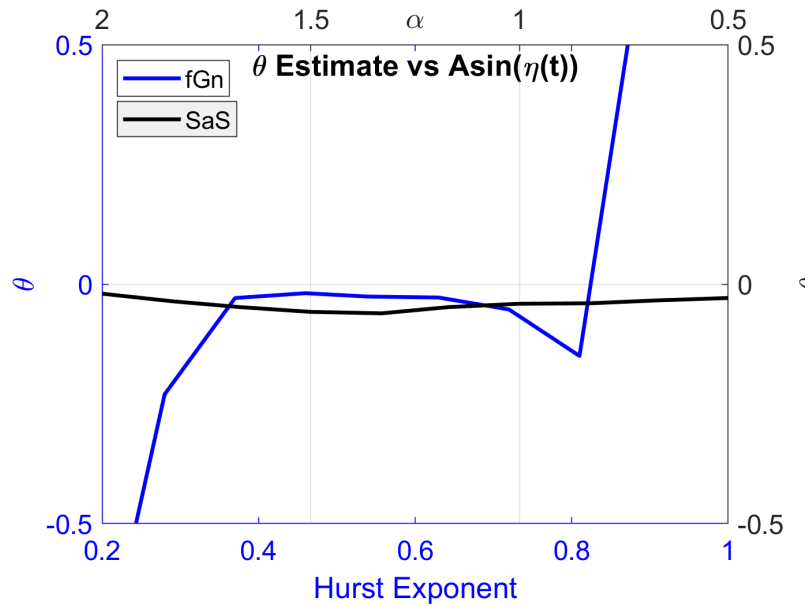


Figure 2.35: The final converged θ is shown as function of fGn and SaS dithering.

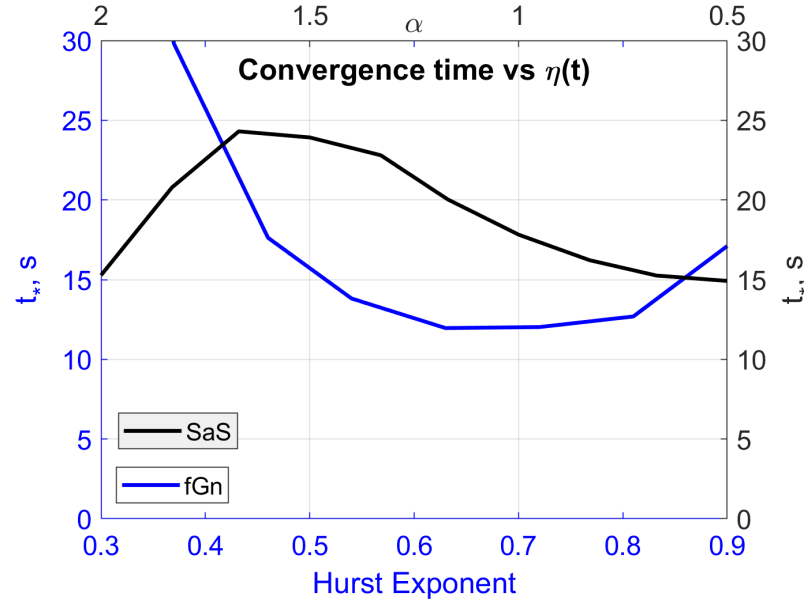


Figure 2.36: The temporal convergence to the global minimum from the initial condition θ_0 , denoted t^* , is shown as a function of fGn and SaS dithering.

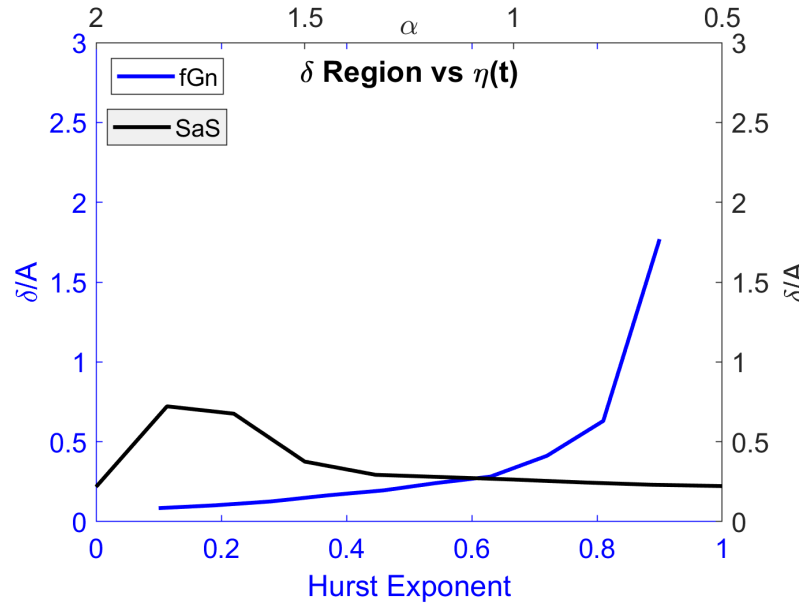


Figure 2.37: The spatial deviation δ around the converged θ is shown as a function of fGn and SaS dithering.

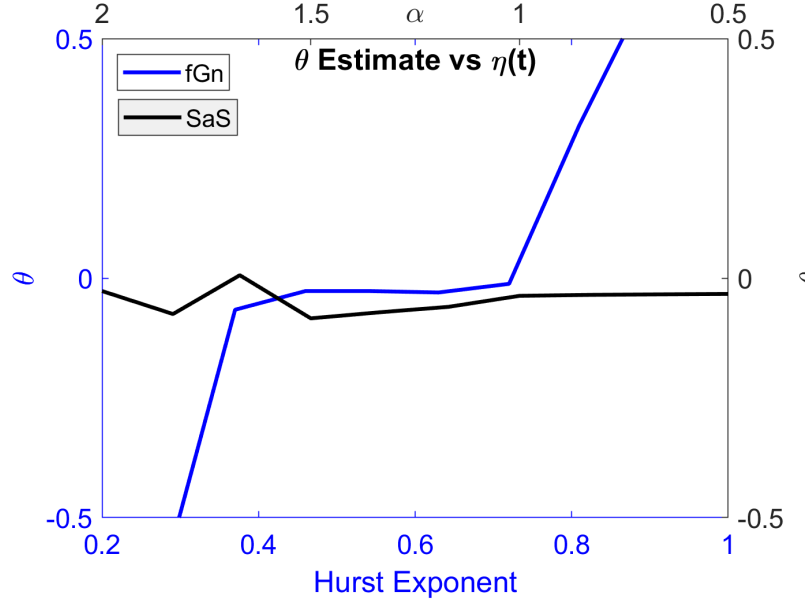


Figure 2.38: The final converged θ is shown as function of fGn and SaS dithering.

2.7 Chapter Summary

In this chapter we explored several case studies related to random search, collective foraging, source seeking, and model-free optimization. It was shown that higher searching efficiency can be obtained when smaller values of ρ and μ are chosen. For collective foraging we seen that the agent based modeling of a loosely coupled system provides benefits and adaptive foraging behavior. In source seeking, we show that fractional order gradients can provide a slight performance increase over that of the original fluxotaxis algorithm. Lastly, for model-free optimization, we have seen that through the use of fractional dithering (e.g. fractional Gaussian noise) a more optimal extremum seeking control algorithm can be executed.

For the fractional order fluxotaxis algorithm, future work would benefit from a simulation benchmark. In the literature, many papers present searching efficiency but differ in simulation setup and design. Therefore, it would be beneficial for comparing source seeking strategies by developing a simulation benchmark with fixed scenarios, starting locations, and plume data.

Chapter 3

Leak Detection and Quantification with sUAS

In this chapter, leak detection and quantification with small unmanned aircraft systems (sUAS) are explored. In the previous chapter, random search, source seeking and model-free optimization were utilized to detect and find the source location explicitly. In practice, a mobile robot can be deployed to carry out such tasks. For practical applications, sUAS are often looked to for carrying out remote sensing techniques. This may be due to their relatively low cost, maneuverability, flexibility, and frequency at which they can be deployed. However, depending on the sensing apparatus and its working principle, the sensor will likely need to be integrated on the sUAS to avoid measurement impact from equipment directly or from the environment (e.g. airflow surrounding the propeller intake and exhaust). This is especially the case for making on-board wind measurements, which is important for emission quantification precision and accuracy. The chapter is organized as follows: Section 3.1 discusses the required sensors and platforms for LDAQ; Section 3.2 will overview quantification methods with and without sUAS; Section 3.3 provides an analysis of quantification methods based on measurement modalities; Section 3.4 will overview detection and localization experiments with sUAS; Section 3.5 will showcase controlled release field test experiments; Section 3.6 will showcase real world field test experiments; and Section 3.7 will conclude and discuss future directions.

3.1 Sensors and Platforms

There are many types of sensors that can be used on-board sUAS, granted they are light enough for a given platform's payload capacity. In this section, we overview a few key sensors used and refer interested readers to a thorough review paper for more on chemical sensing drones [142]. Sensors used in many of the works reviewed here are briefly overviewed.

There are generally two types of sensing modalities, passive and active. Passive sensing encompasses any sensor that receives information from the environment.

A common example of this includes optical cameras, such as visual spectrum cameras (e.g., RGB), thermal cameras (e.g., thermal infrared (TIR) [143; 144], near infrared (NIR), short-wave infrared (SWIR), mid-wave infrared (MWIR), and long-wave infrared (LWIR)).

TIR cameras tend to span a larger bandwidth of wavelengths, whereas hyperspectral cameras can control which wavelengths to focus on. For example, in [145], Telops (Quebec, Canada) used a standoff tripod mounted hyperspectral camera to estimate the flow rate by integrating the mass per unit area and multiplying by the mean velocity of the gas. They utilized a two layer model to calculate the background radiance,

$$L_{tot} = [L_{bkg}\tau_{plume} + L_{plume}(1 - \tau_{plume})]\tau_{atm} + L_{atm}(1 - \tau_{atm}). \quad (3.1)$$

When the hyperspectral camera was optimized for methane detection (as in [146], 7.7 μm band), two controlled release tests showed flow rates calculated (measured) were found to be 25.3 ± 2.8 g/h (23 ± 2.3 g/h) and 102.9 ± 5.8 g/h (100 ± 10 g/h). The authors claimed that the approach is 40 to 100 times more sensitive and can potentially be mounted to an aerial platform, remotely sensing from several hundred meters, deeming it suitable for both natural and anthropogenic sources. Active sensing encompasses any sensor that actively transmits information into the environment, probing a response.

An example of active sensing includes tunable diode laser absorption spectroscopy (TDLAS), which can come in several forms. The working principle relies on the gas species entering the sensor region or laser path, such that some of the power is absorbed by the gas, and a power drop is detected. One form of TDLAS is the closed path TDLAS (e.g., sensing region enclosed in controlled environment), where the emitter and detector are apart of the same device at a fixed distance apart, optimized for a desired detection species.

Other variations of the TDLAS include: the open path TDLAS (e.g., sensing region is open to the environment (see the open path laser spectrometer (OPLS) [9])), the backscatter TDLAS (bs-TDLAS) where the laser is reflected off the natural environment before being received at the detector, and long path TDLAS (e.g., used with retro-reflectors not connected to the physical instrument). Several examples of bs-TDLAS include: LiDAR based (Continuous wave laser absorption LiDAR, Pulsed Differential Absorption LiDAR (DIAL)), Pergam Laser Methane Mini [147], RMLD, Gasfinder2 [148], and Gasfinder3 [149] [150].

Other more sensitive laser based instruments are also used in practice, such as cavity ring-down spectroscopy (CRDS) and off-axis integrated cavity output spectroscopy (ICOS). An example of these types of instruments can be seen from the Los Gatos Research Inc. (LGR) micro greenhouse gas analyzer (MGGA) (also referred to in the literature as the ultra-portable greenhouse gas analyzer (UGGA)) or the Picarro G2301 and G4302. These instruments are typically the gold standard for sensing gas, albeit they are also typically the most heavy as well. Other sensors used consist of non-dispersive infrared (NDIR), ceramic metal oxide sensors

(CMOS) [151; 152; 153; 154], photo-ionization detectors (PID, such as the Honeywell MiniRAE® 3000) and electro-chemical sensors (review of applications [155]).

A recent survey paper in [156] outlined new electronic nose technologies and applications. As the need for low cost sensing solutions increases and becomes commercially available, we are faced with evaluating the accuracy and characteristics of these sensors for practical use, such as in the work by [157] where low cost commercially available sensors were evaluated for precision and accuracy in a gas mixing chamber, providing promise for applications in continuous monitoring applications. Once these sensors can be evaluated and integrated on a platform with suitable sensor characteristics, they can be applied in practice.

For example, in [154], they demonstrated a proof of concept using a chemical multi-sensor payload for gas monitoring based on the DJI S900 platform. Or in [158], a semiconductor type sensor (Testo Gas Detector, Testo SE & Co. KGaA, Titisee-Neustadt, Germany) was used on a DJI M600 to analyze the spatial distribution of methane at a landfill, as well as compared different spatial interpolation techniques. This required calibration and consideration of the vehicle's critical velocity.

For sUAS, lightweight and accurate wind sensors are needed to provide in situ measurements that can be used in the quantification methodology. Some examples of lightweight sensors can range from five hole probes (or multi-hole probes) to ultrasonic anemometers that utilize time of flight (e.g., Anemoment (Longmont, CO, USA) Trisonica used in [159] or the Gill (Lymington, Hampshire, UK) WindMaster used on the OPTOKopter [2]) and resonance based (FT Technologies FT742 and FT205 used in [160]) to more custom micro electrical mechanical systems (MEMS)-based solutions (such as in [161]). These wind sensors can also be applied to general wind profiling and mapping applications. For example, in [162], they used in situ wind measurements on sUAS for understanding the atmospheric boundary layer by developing wind profiling measurements using wind-induced perturbations. Mapping wind distributions over complex terrain was explored in [163], where they utilized a Gill WindMaster 3D ultrasonic anemometer mounted on a octocopter (called the WindLocator).

Wind and temperature profiling from fires were explored in [164] based on wind measurements using the Trisonica. Vertical velocity measurements of aerosol cloud interactions were compared with ground-based radar in [165]. If payload limitations prevent integration of an on-board wind sensor, wind estimation techniques can also be explored, such as in [166; 167]. For a more thorough understanding of different sensors and estimation techniques, we refer the reader to [168].

The choice of a platform and ancillary equipment depends directly on the choice payload system (e.g., the collection of on-board sensors) that needs to be integrated onto the sUAS. For LDAQ, this typically includes lightweight methane and wind sensors. Due to sUAS payload capacity limitations, this often leads to integration problems as the weight of methane sensors can vary greatly depending on the desired sensitivity and response time (e.g., from a couple hundred grams to a couple kilograms). Lightweight and low cost methane sensors (such as CMOS) are slower in response and are less sensitive. For high accuracy and fast response sensors, which

typically weigh more, this ultimately affects the selection process of sUAS, requiring bigger and more costly platforms to maintain safe stability and control.

For this reason, ground-based wind sensors are often used (placed adjacent to operating area) with smaller sUAS platforms gathering the in situ methane measurements. The data from these two systems are collected by a local data acquisition system for live observation and post-processing. This kind of configuration, implicitly assumes that the average wind, or mean wind field, represents the overall sensing region quite well. This is usually only the case in rural areas, where there are little to no obstructions (i.e., trees, hills, buildings, infrastructure, etc.). Most of the scenarios faced in practice, however, deal with obstructions and require on-board, or in situ, wind measurements.

This decision requires platform specific knowledge (e.g., hardware setup, autopilot, co-pilot software, etc.) as well as desired payload knowledge, which will vary depending on the desired application, sensitivity, measurement mode, and so on. For these reasons, and since this manuscript's focus is on detection and quantification methodologies, we omit these details in this manuscript, and suggest that the interested reader see [169] for a guide on sUAS platform selection.

Payload integration strategies for methane sensors include several configurations, such as the boom-mounted, bottom-mounted, or top-mounted (see Figure 3.1). Boom-mounted approaches typically consist of TDLAS based sensors, which are subject to disturbances in the measurement from downwash of the propellers. This is avoided by placing the sensor out front of the aircraft along a boom and sampling when the effective wind speed over the sensor is greater than 2 m/s [170].

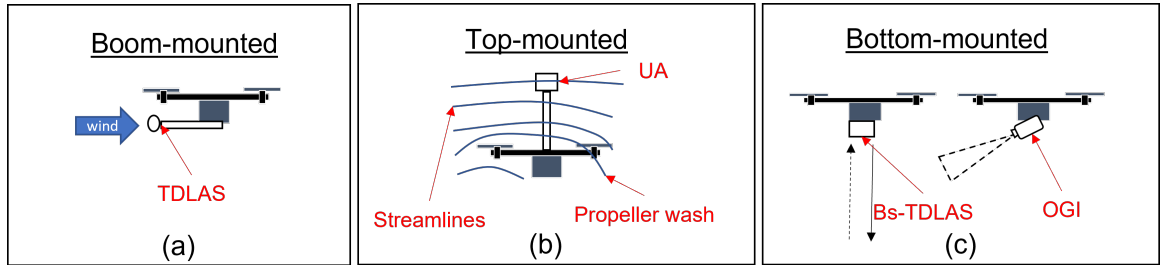


Figure 3.1: Payload configuration examples of (a) boom-mounted TDLAS (e.g. see [1] with an OPLS), (b) top-mounted anemometer showcasing the effect of propellers on the streamlines (e.g. see [2] in wind tunnel tests), and (c) bottom-mounted bs-TDLAS or OGI (e.g. see [3] with a RMLD).

Bottom-mounted approaches are typically used with bs-TDLAS or OGI. The sensor is mounted on a gimbal system or sometimes hard mounted to the aircraft frame (see [142] for more details). On the contrary, top-mounted wind sensors can provide high accuracy if translational and induced wind velocities can be removed [2].

3.2 Leak Quantification Methods

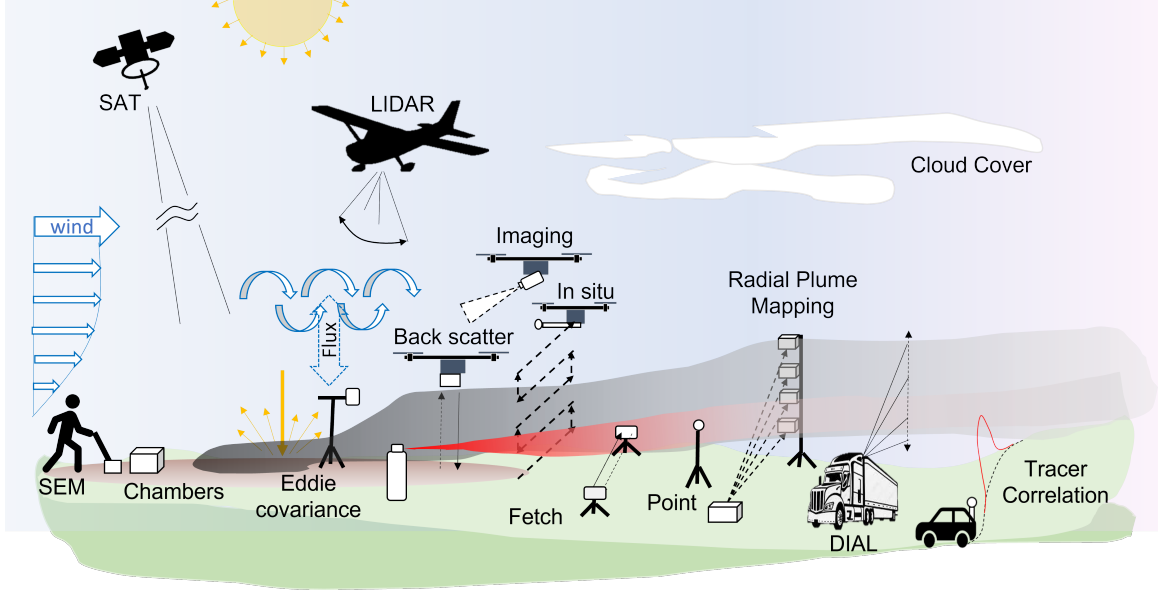


Figure 3.2: Leak detection and quantification methods and modalities.

In this section, we overview the conventional and sUAS-based advanced leak detection and quantification (LDAQ) methodology. The LDAQ methods utilize several concepts and approaches within numerics, control, and optimization as well as approaches based on different available sensing modalities (see Fig. 3.2 for effective length-scales). In this manuscript, we divided these approaches into five general categories, namely: Simulation-based (Section 3.2.1), Optimization-based (Section 3.2.2), Mass-Balance-based (Section 3.2.3), Imaging-based (Section 3.2.4), and Correlation-based (Section 3.2.5).

In the Simulation-based approaches, the methods depend heavily on simulation and computational tools for solving dynamic partial differential equations, which are used to determine the source rate estimation. Sometimes other source parameters are also estimated in the process and this is typically referred to as source term estimation (STE) or the source determination problem (SDP). Two methods that show up in the literature are backwards Lagrangian stochastic (bLS) and mesoscale recursive Bayesian least squares inverse (RB-LSI).

The optimization-based methods showcased in this manuscript depend on some form of a parameterized system model, which undergoes a model fitting or recursive optimization (statistical or information based). Many of these methods include several variations of the point source Gaussian (PSG) solution of the classical Gaussian plume model. This is seen in the PSG approach based on the EPA's other test method (OTM) 33A, where the data is gathered from a single sensor downwind and undergoes model fit of the peak concentration measured.

Next is the conditionally sampled PSG (PSG-CS) approach that utilizes meteorological data in the model fitting process using conditionally sampled concentration data based on the incremental changes in wind direction. Another variation to this is the recursive Bayesian PSG (PSG-RB) that utilizes a moving sensor and meteorological data to condition the models likelihood function and prior for updating the posterior distribution that is used to quantify the source estimate. This approach also considers past knowledge about equipment characteristics if this is known.

A different approach to the Bayesian way of thinking is to solve for the parameters of the model conditioned on the observations. This approach also utilizes a particle filter and Markov Chain Monte Carlo (MCMC) to update the posterior and is referred to as the PSG sequential Bayesian MCMC (PSG-SBM). The last optimization approach mentioned in this manuscript is the Near-Field Gaussian Plume Inversion (NGI) and the modified NGI approach.

The NGI utilizes fitting the Gaussian plume model based on sampling of a perpendicular plane downwind of the source. The vertical and horizontal dispersion relations are used to find the center of the plume within the perpendicular plane and minimize, by least square fit, the difference between the modeled concentration and the observed concentration (integrated over the lateral dispersion direction).

The next category is the Mass-Balance-based approaches, which includes methods that utilize equations based on mass conservation and continuity. The simplest approach is the vertical flux plane (VFP), which takes a control volume approach to estimating the emission rate by measuring the flux entering and leaving the control volume. Traditionally, the plume is sampled using a raster scanning approach in a perpendicular plane upwind and downwind of the source. The sparse set of observations within the plane undergo a spatial interpolation process and are combined with the wind to estimate the source rate.

A direct variation to this approach is the cylindrical flux plane (CFP), which the sensing system measures concentrations on successive loops around the source at different altitudes. The flux going into and out of this cylindrical plane is used to estimate the flux. Using different sensing modalities (such as imaging or backscatter-based sensors), a path integrated vertical flux plane (PI-VFP) method can be formulated. Both aircraft and sUAS-based approaches to PI-VFP have been implemented, which rely on horizontal scanning of the area of interest.

For sUAS-based PI-VFP, concentric circles are flown to confirm that sources are contained inside the path before estimating the source rate. A flux plane approach has also been explored using a series of TDLAS-based laser fetches at different altitudes and utilizes the time-average of the line-integral of the instantaneous product of the wind speed and concentration. This is advantageous to other VFP approaches as it provides very good performance and does not take time to scan the plane. However, it is in ways impractical as it requires setup of the laser fetches and knowledge of the source geometry.

The next method is the Gauss divergence theorem (GDT) approach. It utilizes the CFP approach with mass flux continuity as well as the expected time rate of change

of the mass within the control volume to estimate the source rate. Another VFP approach was included in this review that uses Gaussian plume model optimization with a general linear model (GLM) to help determine the contributions of multiple sources. This approach is referred to as the VFP-GLM.

The last Mass-Balance-based approach is vertical radial plume mapping (VRPM). The VRPM approach uses a ground based laser with retro-reflectors at different altitudes downwind of the source. The path integrated concentrations are measured at different radial angles and used to estimate the flux.

The next category is the Imaging-based approaches that utilize MWIR, hyperspectral cameras, and absorption spectroscopy (such as iterative maximum a posteriori differential optical absorption spectroscopy (IMAP-DOAS)).

The last category covered in this manuscript is the correlation-based approaches, which includes the traditional Eddy covariance (EC) method (in brief) and the tracer correlation method (TCM). The TCM has also been referred to in the literature as the tracer dispersion method (TDM) and atmospheric tracer method (ATM).

3.2.1 Simulation-Based

Forward Modeling

Forward modeling is typically used for projecting or forecasting dispersion. Forward modeling is not directly used in emission quantification by itself, but rather paired with feedback in the optimization sense. This can include numerically solving a governing set of equations, such as the advection diffusion equation (ADE) or applying a parameterized general model (such as the Gaussian plume). It is also common in practice to utilize existing numerical models, such as the WindTrax 2.0, WRF model, FLEXible PARTicle-Weather Research and Forecasting (FLEXPART-WRF), SCIPUFF, QUIC, and others that can be Lagrangian-base, include turbulence e.g., Large eddie simulation (LES), and Reynolds averaged Navier Stokes (RANS). Interested readers can check the review paper from [171] on dispersion models.

Backward Lagrangian Stochastic (bLS)

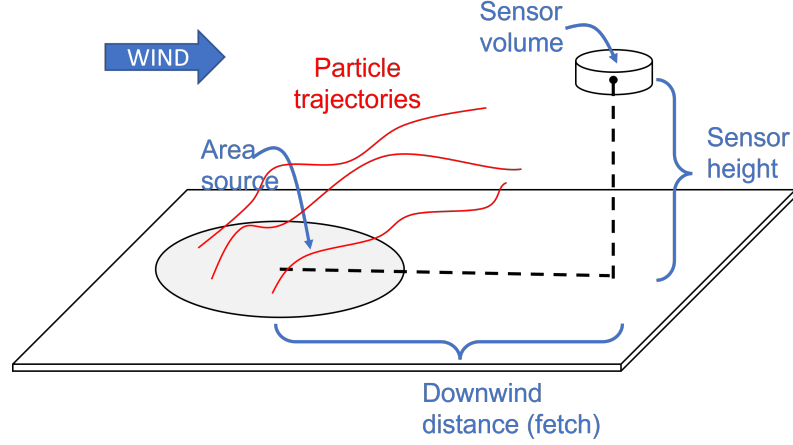


Figure 3.3: A diagram depicting the bLS approach (e.g. see [4]).

The accepted backward modeling approach used in the draft OTM- 33A document [5] and in several applications (e.g., Dairy Farm [172], etc.) is the backwards Lagrangian stochastic (bLS) approach by [4]. The bLS approach aims to answer the general questions: What is the proper form of the LS trajectory model? As well as, how can source estimates be extracted from the particle's backward LS trajectory? The forward model, formulated as a generalized Langevin equation, is evolved jointly as a Markov process,

$$du_i = a_i(\mathbf{x}, \mathbf{u}, t)dt + b_{i,j}(\mathbf{x}, \mathbf{u}, t)d\xi_j, \quad dx_i = u_i dt, \quad (3.2)$$

where the particle position is given by $\mathbf{x} = (x_1, x_2, x_3)$, and $d\xi_j$ is a random increment governed by a Gaussian process. The functions a_i and $b_{i,j}$ have to be specified such that the velocity probability density function, $g_a(\mathbf{x}, \mathbf{u}, t)$, satisfies the Fokker–Planck equation (FPE) [4],

$$\frac{\partial g_a}{\partial t} = \frac{\partial}{\partial x_i}(u_i, g_a) - \frac{\partial}{\partial u_i}[a_i(\mathbf{x}, \mathbf{u}, t)g_a] + \frac{\partial}{\partial x_i}[B_{i,j}(\mathbf{x}, \mathbf{u}, t)g_a]. \quad (3.3)$$

This method provides a source estimation for an area source given the source location (with unknown source rate), assuming horizontally uniform surface source, and that the atmosphere is in horizontal equilibrium (see Fig. 3.3). To make an emission estimate using bLS, the method utilizes the dispersion model relation,

$$\frac{Uy}{Q} = n = f(\mathbf{x}_m, \mathbf{x}_0, L, h, G), \quad (3.4)$$

where U is the wind speed, y is the measured concentrations, Q is the source rate, L is the Monin–Obukhov length, h is the depth of the mixing layer, G describes the set

of parameters characterizing the plume, and x_{3m} represents the measurement height. As the particles from the back trajectories touchdown in the source area, the vertical velocities, w_0 are logged and used to estimate n ,

$$n(x_{3m}) = \frac{y(x_{3m})U(x_{3m})}{Q} = \frac{1}{N} \sum \left| \frac{2}{w_0/U(x_{3m})} \right|. \quad (3.5)$$

Once n is known, an estimate of the source rate can be determined using the measured concentration and wind speed, $Q = n/(yU)$. In this approach, due to the time-averaged ensemble, the accuracy improves over time (nominal averaging period of 15 min [4]). An alternate expression for the emission estimate is given as

$$Q_{bLS} = \frac{y - y_b}{(y/Q)_{sim}}, \quad (3.6)$$

where y_b is the background concentration and $(y/Q)_{sim}$ is calculated using

$$(y/Q)_{sim} = \frac{1}{N} \sum \left| \frac{2}{w_0} \right|. \quad (3.7)$$

This Monin–Obukhov similarity theory (MOST)-based bLS emission estimation methodology was validated against the mass balance approach (given the along-wind distance of the source d),

$$Q_m = \frac{1}{d} \int_0^\infty y(x_3)U(x_3)dx_3, \quad (3.8)$$

and field tested in [148; 173; 174]. A lagoon environmental leak was simulated and explored with the bLS approach by constructing a large 45 m by 45 m emission source on a pond. The accuracy was shown to be lower during the summer period due to more frequent unstable atmospheric conditions [175].

Mesoscale Recursive Bayesian Least Squares Inverse (RB-LSI)

Utilizing the NOAA P-3 aircraft and a wavelength-scanned CRDS, [176] used a mesoscale bayesian least squares approach to solve the inverse problem of estimating emissions. They use the FLEXPART-WRF to model the forward problem, which was compared to physical observations and minimized on an iterative cost function that assumes lognormal distributions,

$$J = \frac{1}{2} (\ln(y_0) - \ln(Hx))^T (\ln(y_0) - \ln(Hx)) + \frac{1}{2} \alpha (\ln(x) - \ln(x_b))^T (\ln(x) - \ln(x_b)), \quad (3.9)$$

where the observed concentration enhancements are given as y_0 , posterior solutions are x , FLEXPART-WRF outputs are H , prior fluxes are x_b , error covariance matrix from observations are R , and error covariance matrix from prior fluxes are B in the lognormal space.

3.2.2 Optimization-Based

In this section, we discuss the emission quantification techniques that utilize some form of optimization in the methodology that fits a model.

Point Source Gaussian (PSG) - OTM33A

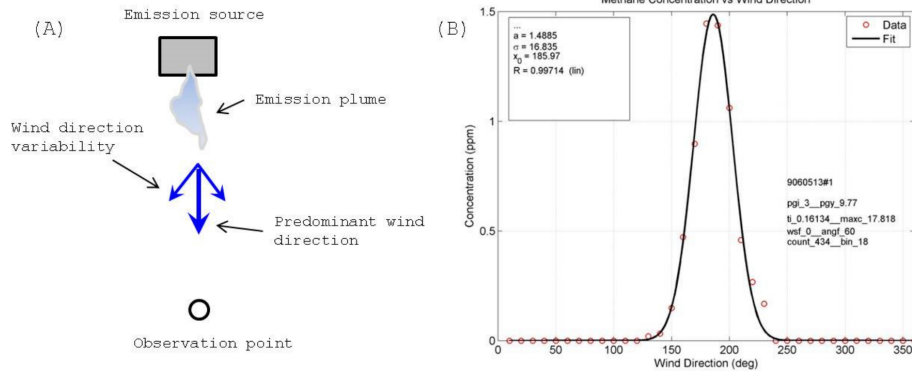


Figure 3.4: (a) Depiction of Gaussian plume dispersion with an observer making a stationary measurement downwind. (b) Resulting time-integrated data with a Gaussian fit applied [5].

In [5], the point source Gaussian (PSG) is discussed. The measurement involves a vehicle with a concentration measurement instrument (CMI) to park downwind of the known source with the vehicle off. The CMI (such as Picarro or LGR UGGA) collects data at roughly 2.5 m above ground at a known distance from the source. The variations in the wind direction are measured using a sonic anemometer (e.g., R.M. Young). The PSG calculations are based off of enhanced emission levels and can be calculated as the fifth percentile of the concentration time series signal. The PSG estimate then becomes a simple 2-D Gaussian integration with no reflection term. The source estimate is given as,

$$Q_E = 2\pi\sigma_{x_2}\sigma_{x_3}U_my_p, \quad (3.10)$$

where y_p is the peak concentration from the Gaussian fit, U_m is the mean wind speed, σ_{x_3} and σ_{x_2} are the vertical and lateral plume dispersion that can be determined from the meteorological conditions, such as the Pasquill–Gifford stability classification curves [177] (see Fig. 3.4). The accuracy of the OTM33A method is explored in [178; 179].

Conditionally Sampled PSG (PSG-CS)

To capture the ensemble mean of the downwind plume behavior, a dispersion model is used in [6], where we denote $\mathbf{x} = [x_1, x_2, x_3]^T$ is the downwind distance, crosswind

distance ($x_2 = 0$ is the plume center) and vertical position, respectfully. The model is a function of downwind distance and dispersion factors $D_{x_2}(x_1, x_2)$ and $D_{x_3}(x_1, x_3)$, given as,

$$y_m(\mathbf{x}) = \frac{Q}{U} D_{x_2}(x_1, x_2) D_{x_3}(x_1, x_3). \quad (3.11)$$

This method essentially aims to determine the source rate, Q , using the conditional mean concentration data, y_m , of the downwind plume. The lateral dispersion downwind of a continuous point source can be shown to have a Gaussian distribution such that it can be represented as,

$$D_{x_2}(x_1, x_2) = \frac{1}{\sqrt{2\pi}\sigma_{x_2}} \left[-\frac{1}{2} \left(\frac{x_2}{\sigma_{x_2}} \right)^2 \right]. \quad (3.12)$$

However, the vertical dispersion (assuming vertical eddy diffusivity and wind speed that scales vertically to a power law) can be formulated as a parameterized stretched exponential (originally expressed in [180]),

$$D_{x_3} = D_{x_3}(x_1, x_3) = \frac{A}{\overline{x_3}} \exp \left[- \left(\frac{Bx_3}{\overline{x_3}} \right)^s \right]. \quad (3.13)$$

The parameters $\overline{x_3}$, s , A , and B are functions of the atmospheric stability and downwind distance, x_1 . A and B can be described using the usual Gamma function, $\Gamma(\cdot)$ as,

$$A = s\Gamma(2/s)[\Gamma(1/s)]^2, \quad (3.14)$$

$$B = s\Gamma(2/s)\Gamma(1/s). \quad (3.15)$$

The conditional averaged concentration can be calculated using,

$$\langle y|\theta \rangle = \frac{1}{n} \sum_{\theta_i \in \Theta} y(\theta_i), \quad (3.16)$$

where the set $\Theta(\theta) = \{\theta_i : |\theta - \theta_i| < \Delta\theta/2, \forall i = 1, 2, \dots, n\}$ and $\Delta\theta = 2^\circ$. The basic idea is to capture the plume geometry in the crosswind direction, which is further used to derive the least squares source estimate,

$$Q = \left[\sum_{i=1}^N \frac{D_{x_2} D_{x_3}}{U} \langle y|\hat{d}_i \rangle \right] / \left[\sum_{i=1}^N \left(\frac{D_{x_2} D_{x_3}}{U} \right)^2 \right]. \quad (3.17)$$

As shown in [6], the lateral dispersion can be determined in two ways: classically, using atmospheric stability (for constants a_{x_2} and p_{x_2}) [181],

$$\sigma_{x_2} = a_{x_2} x_{30} 1.9 (L_x / x_{30})^{p_{x_2}}; \quad (3.18)$$

and by reconstructing the lateral dispersion,

$$\sigma_{x_2} = \sqrt{\frac{1}{N} \sum_{i=1}^N \hat{d}_i}, \quad (3.19)$$

where the N is the number of values in $\langle y|\hat{d} \rangle$, and $\hat{d}$ are $\hat{d}$ values that are greater than the minimum concentration (i.e., background) and $\pm 40^\circ$ off the plume center θ_p . The distance $\hat{d}$ is calculated as,

$$\hat{d}(\theta) = L_x \sin(\theta - \theta_p), \quad (3.20)$$

with $\theta_p = \arg \max_{\theta} \langle y|\theta \rangle$ (see Figure 3.5).

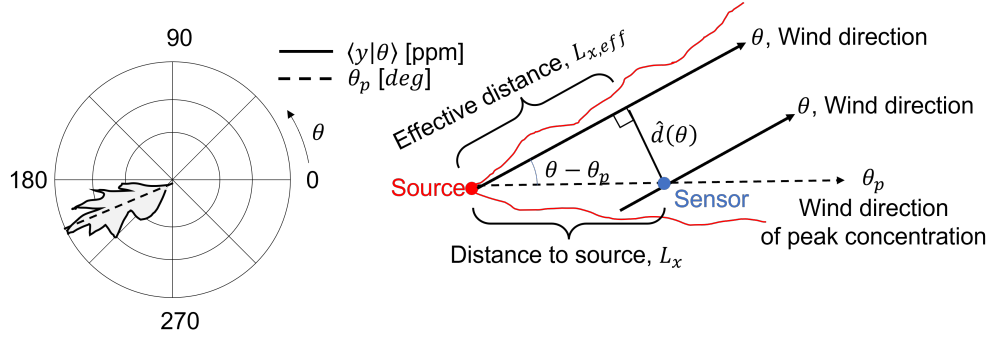


Figure 3.5: (left) Polar plot with the wind direction, θ as the radial axis, and the conditionally averaged concentration, $\langle y|\theta \rangle$ as the distance from the center. θ_p is the peak wind direction located at the maximum conditionally averaged concentration. (right) Illustration of the wind direction geometry for conversion of θ to crosswind position $\hat{d}$ with the source plume represented by the dashed lines [6].

Recursive Bayesian Point Source Gaussian Method (PSG-RB)

In work from [182; 183], a moving sensor measured a point source concentration that can be formulated as,

$$y(\mathbf{x}) = \frac{Q}{\bar{U}} D_{x_2}(x_1, x_2) D_{x_3}(x_1, x_3). \quad (3.21)$$

The source rate is given as Q , the effective wind speed is $\bar{U}$, and the lateral and vertical dispersion is characterized by $D_{x_2}(x_1, x_2)$ and $D_{x_3}(x_1, x_3)$, respectively. The equation is formulated such that the downwind distance, x_1 , is aligned with the predominant wind direction. Since the measurement is taken at closer distances to the source, the lateral dispersion is taken as a random function such that,

$$\int_{-\infty}^{\infty} D_{x_2}(x_1, x_2) dx_2 = 1. \quad (3.22)$$

This can be advantageous for instantaneous plumes. The integrated lateral concentration can be written as,

$$y^{x_2}(x_1, x_3) = \frac{Q}{U} D_{x_3}(x_1, x_3). \quad (3.23)$$

The choice of the vertical dispersion D_{x_3} (originally expressed in [180]) is that of a parameterized stretched exponential function,

$$D_{x_3} = D_{x_3}(x_1, x_3) = \frac{A}{\bar{x}_3} \exp \left[-\left(\frac{Bx_3}{\bar{x}_3} \right)^s \right], \quad (3.24)$$

where $\bar{x}_3$, s , A , and B are functions of atmospheric stability and downwind distance, x_1 . The lateral dispersion is given as

$$D_{x_2} = \frac{1}{\sqrt{2\pi}\sigma_{x_2}} \exp \left[-\frac{1}{2} \left(\frac{x_2}{\sigma_{x_2}} \right)^2 \right]. \quad (3.25)$$

Then, by numerically integrating (3.21) and incorporating the vehicle movement V ,

$$y^{x_2} = \sum_{i=0}^{\infty} y(\mathbf{x}_i) \Delta t V = Q \sum_{i=0}^{\infty} \frac{\Delta t V}{U_i} D_{x_3}(x_{1_i}, x_{3_i}) D_{x_2}(x_{1_i}, x_{2_i}). \quad (3.26)$$

The recursive Bayesian approach described here is based on well pads and oil and gas production, which are used to help inform the path planning of the mobile sensor. For brevity, we will only cover the formulation of the quantification only. Starting with the definition of the posterior distribution,

$$p(Q|M, W, \Lambda) = \frac{p(Q|W)p(M|Q, \Lambda)}{p(M|\Lambda)}, \quad (3.27)$$

where M is the concentration data, W is the ancillary information (e.g., well pad characteristics), Λ is the meteorological conditions, $p(Q|W)$ is the prior, $p(M|Q, \Lambda)$ is the likelihood, and $p(M|\Lambda)$ is the evidence (which can be thought of as a normalization constant for the likelihood [184]). The prior is given as

$$p(Q|W) = \frac{1}{\beta} \exp \left[-\left(1 + \gamma \frac{Q - \mu}{\beta} \right)^{-\frac{1}{\gamma}} \right] \left(1 + \gamma \frac{Q - \mu}{\beta} \right)^{-1-1/\gamma}, \quad (3.28)$$

where the hyperparameters need to be fit to the application (for well-pad source, $\gamma = 1$, $\mu = 0.19$, $\beta = 0.23$ based on [185]). The likelihood function is chosen to be a Gaussian,

$$p(M|Q, \Lambda) = \frac{1}{\sqrt{2\pi}\sigma_e} \exp \left[-\frac{1}{2} \left(\frac{y^{x_2} - y^{x_2, M}}{\sigma_e} \right)^2 \right], \quad (3.29)$$

where $y^{x_2, M}$ is the modeled concentration for a given source rate, and σ_e is the combined model and measurement error (outlined in [186]). The recursive approach involves replacing the prior with the previous posterior distribution found using the likelihood function,

$$p(Q|W)_i = \begin{cases} p(Q|W), & i = 1, \\ p(Q|M, W, \Lambda)_{i-1}, & i > 1. \end{cases} \quad (3.30)$$

As the number of passes increases, the posterior distribution improves and can be used to estimate the source rate,

$$\hat{Q} = \int_{Q_{min}}^{Q_{max}} Q p(Q|M, W, \Lambda) dQ. \quad (3.31)$$

Variations of this method were seen in [183], where the measurement noise was assumed to be Gaussian and also included a UAV with sensor noise and utilized the flux plane mass balance method to estimate the source rate, which was used in the calculation of the posterior distribution. Further field tests of this method were carried out in [187].

Point Source Gaussian Sequential Bayesian Markov Chain Monte Carlo (PSG-SBM)

Utilizing the Gaussian plume model for the likelihood of a sequential Bayesian Markov Chain Monte Carlo (MCMC) method, a UAV scans horizontally to update the estimated posterior distribution in [188]. The parameters are given as Θ_k where $\Theta_k = [\mathbf{x}_s^T, Q_s, u_s, \phi_s, \zeta_s]^T$, the position is $\mathbf{x}_s$, source rate Q_s , wind speed and direction u_s and ϕ_s , and the model diffusion coefficients $\zeta_s = [\zeta_{s1}, \zeta_{s2}]^T$. The point source observations, $\mathbf{y}_{1:k} = \{y_1, y_2, \dots, y_k\}$ are used within Bayes rule to update the posterior,

$$p(\Theta_{k+1}|\mathbf{y}_{1:k+1}) = \frac{p(\mathbf{y}_{k+1}|\Theta_{k+1})p(\Theta_{k+1}|\mathbf{y}_{1:k})}{p(y_{k+1}|\mathbf{y}_{1:k+1})}. \quad (3.32)$$

The likelihood model, $\mathcal{M}(\mathbf{x}_k, \zeta_k)$, in [188], based on observational data, $\bar{y}_k = \mathcal{M}(\mathbf{x}_k, \zeta_k) + \bar{v}_k$, was taken to be detection event, $p(\bar{y}_k|\Theta_k)$, if $y_k > y_{thr}$,

$$p(\bar{y}_k|\Theta_k) = \frac{1}{\sigma_k \sqrt{2\pi}} \exp \left[-\frac{(\bar{y}_k - \mathcal{M}(\mathbf{x}_k, \zeta_k))^2}{2\sigma_k^2} \right], \quad (3.33)$$

and a non-detection event otherwise,

$$p(\underline{y}_k|\Theta_k) = \left(\frac{p_b}{2} \left[1 + \operatorname{erf} \left(\frac{y_{thr} - \mu_b}{\sigma_b \sqrt{2}} \right) \right] \right) + p_m + \left(\frac{p_s}{2} \left[1 + \operatorname{erf} \left(\frac{y_{thr} - (\mu_b + \mathcal{M}(\mathbf{x}_k, \zeta_k))}{\sigma_b \sqrt{2}} \right) \right] \right). \quad (3.34)$$

The three terms in the non-detection event account for instrument noise, turbulence, and observing concentrations above the threshold, where $p_b + p_m + p_s = 1$, and μ_b and σ_b are mean background noise and standard deviation, respectively. Using a particle filter, the posterior can be approximated by a set of n weighted random samples $\{\Theta_k^{(i)}, w_k^{(i)}\}_{i=1}^n$,

$$p(\Theta_k | \bar{y}_{1:K}) \approx \sum_{i=1}^n w_k^{(i)} \delta(\Theta_k - \Theta_k^{(i)}), \quad (3.35)$$

where δ is the Dirac delta function. The un-normalized weights are then updated using

$$\bar{w}_{k+1}^{(i)} = w_k^{(i)} \cdot p(y_{k+1} | \Theta_k^{(i)}). \quad (3.36)$$

Once the weights are determined they can be normalized by dividing by the summation of all the weights. Additionally, an effective sample size must be considered to avoid the degeneracy problem. The new samples undergo a MCMC step that is accepted with the likelihood probability distribution described earlier (see more details in [188]).

Near-Field Gaussian Plume Inversion (NGI)

The near-field Gaussian plume inversion (NGI) method [189; 190] is a mass continuity model in principle, where the upwind and downwind concentration measurements, combined with wind measurements, of an emission source are differenced to quantify emission flux. The NGI method is typically sampled around 100 m from the source. The sampling aims to capture the time-invariant behavior of the plume, which, under turbulent conditions, may not map out the characteristic Gaussian plume shape.

This is because it is assumed that spatial variability in the time-averaged plume is Gaussian. This method was initially carried out with a DJI S900 equipped with an ultra portable greenhouse gas analyzer (UGGA) by Los Gatos Research Inc. (LGR). The flux estimate is derived by fitting the experimentally measured flux values, q_{me} , to the modeled flux values, q_{m0} given as,

$$q_{me} = (y - y_b)U(x_3)\rho, \quad (3.37)$$

where the modeled flux is given by the Gaussian model,

$$q_{m0} = \frac{Q_e}{2\pi\sigma_{x_2}(x_1)\sigma_{x_3}(x_1)} \exp\left(\frac{-(x_2 - x_{2c})^2}{2\sigma_{x_2}(x_1)^2}\right) \left(\exp\left(\frac{-(x_3 - h)^2}{2\sigma_{x_3}(x_1)^2}\right) + \exp\left(\frac{-(x_3 + h)^2}{2\sigma_{x_3}(x_1)^2}\right) \right). \quad (3.38)$$

The lateral and vertical dispersion relations are typically looked up in the PGT stability tables, however, in this method, they are assumed to be linearly proportional to downwind distance,

$$\tau_{x_2} = \sigma_{x_2}(x_1)/x_1, \quad \tau_{x_3} = \sigma_{x_3}(x_1)/x_1. \quad (3.39)$$

Trying to solve (3.38) is not always well constrained, and thus the method proposes to separate (3.38) and fit the model along the x_3 -direction,

$$q_{me,x_2} = q_{me} \frac{\tau_{x_3} x_1 \sqrt{2\pi}}{\left(\exp\left(\frac{-(x_3 - h)^2}{2(\tau_{x_3} x_1)^2}\right) + \exp\left(\frac{-(x_3 + h)^2}{2(\tau_{x_3} x_1)^2}\right) \right)}. \quad (3.40)$$

The spatial variability in the x_3 -direction has to be sampled to determine τ_{x_3} . The lateral spatial variability τ_{x_2} and plume center x_{2c} are determined simultaneously,

$$x_{2c} = \frac{\sum_j (q_{me,x_2} x_{2j})}{\sum_j (q_{me,x_2})}, \quad (3.41)$$

$$\tau_{x_2} = \sqrt{\frac{\sum_j (q_{me,x_2j} (\frac{x_{2j}-x_{2c}}{x_{1j}})^2)}{\sum_j (q_{me,x_2j})}}. \quad (3.42)$$

Once the unknowns, τ_{x_3} , τ_{x_2} , and x_{2c} are determined, the source emission rate, $\hat{Q}$, can be estimated by minimizing the least square fit between q_{me} and q_{mo} , parameterized by Q_e and τ_{x_3} . The uncertainty in $\hat{Q}$ and the impact of limiting τ_{x_3} are given in [189].

Modified NGI

To improve the optimization routine of the NGI method, the modified NGI (mod-NGI) method looks at conditioning the parameter estimates by using the following flux-based likelihood function, represented by,

$$P(q|\theta) = D_{x_2}(x_1, x_2; \sigma_{x_2}, \mu_{x_2}) D_{x_3}(x_1, x_3; \sigma_{x_3}, \mu_{x_3}), \quad (3.43)$$

where $\theta = [Q, \mu_{x_2}, \mu_{x_3}, \sigma_{x_2}, \sigma_{x_3}]^T$ and the dispersion functions are,

$$D_{x_2}(x_1, x_2; \sigma_{x_2}, \mu_{x_2}) = \frac{1}{\sqrt{2\pi}\sigma_{x_2}(x_1)} \exp\left(\frac{-(x_2 - \mu_{x_2})^2}{2\sigma_{x_2}(x_1)^2}\right), \quad (3.44)$$

$$D_{x_3}(x_1, x_3; \sigma_{x_3}, \mu_{x_3}) = \frac{\exp\left(\frac{-(x_3 - \mu_{x_3})^2}{2\sigma_{x_3}(x_1)^2}\right) + \exp\left(\frac{-(x_3 + \mu_{x_3})^2}{2\sigma_{x_3}(x_1)^2}\right)}{\sqrt{2\pi}\sigma_{x_3}(x_1)}. \quad (3.45)$$

When the source rate is multiplied we have the modelled flux equation,

$$q_{mo} = \mathcal{M}(\mathbf{x}, \theta) = Q D_{x_2}(x_1, x_2; \sigma_{x_2}, \mu_{x_2}) D_{x_3}(x_1, x_3; \sigma_{x_3}, \mu_{x_3}). \quad (3.46)$$

Consider that experimental measurements do not give access to the above likelihood function but can rather give an estimate of it and let's assume an initial source rate estimate $\hat{Q}$ such that,

$$\hat{P}(q_{me}|\hat{\theta}) \approx \frac{Q}{\hat{Q}} D_{x_2}(x_1, x_2; \hat{\sigma}_{x_2}, \hat{\mu}_{x_2}) D_{x_3}(x_1, x_3; \hat{\sigma}_{x_3}, \hat{\mu}_{x_3}). \quad (3.47)$$

The likelihood of observing the entire dataset then becomes,

$$\hat{P}(D|\hat{\theta}) = \prod_{i=1}^N \hat{P}(q_{me}^i|\hat{\theta}). \quad (3.48)$$

Given that analytically solving for an optimizer using (3.45) can be complex, we make the assumption that the plume is far from the ground only for the parameter initialization steps. This allows for the approximation of (3.45) by,

$$D_{x_3}(x_1, x_3; \hat{\sigma}_{x_3}, \hat{\mu}_{x_3}) \approx \frac{1}{\sqrt{2\pi}\sigma_{x_3}(x_1)} \exp\left(\frac{-(x_3 - \mu_{x_3})^2}{2\sigma_{x_3}(x_1)^2}\right). \quad (3.49)$$

Solving for the maximum log likelihood estimate (MLE) yields,

$$\hat{\tau}_{x_2} = \sqrt{\frac{1}{Nx_1^2} \sum_{i=1}^N (x_{2i} - \mu_{x_2})^2}, \quad \hat{\tau}_{x_3} = \sqrt{\frac{1}{Nx_1^2} \sum_{i=1}^N (x_{3i} - \mu_{x_3})^2}. \quad (3.50)$$

It can be readily apparent that the MLE is the definition of the standard deviation conditioned only on the spatial coordinates. To make this MLE conditioned on the flux measurements one can substitute for the weighted standard deviation. For example, the horizontal scale factor is given by,

$$\hat{\tau}_{x_2} = \sqrt{\frac{N}{(N-1)x_1^2} \frac{\sum_{i=1}^N q_{me}^i (x_{2i} - \mu_{x_2})^2}{\sum_{i=1}^N q_{me}^i}}. \quad (3.51)$$

In order to estimate τ_{x_2} and τ_{x_3} , the plume centers need to be computed. This is done using the measured flux as weights, and computing the center of mass,

$$\hat{\mu}_{x_2} = \frac{\sum_{i=1}^N q_{me}^i x_{2i}}{\sum_{i=1}^N q_{me}^i}, \quad \hat{\mu}_{x_3} = \frac{\sum_{i=1}^N q_{me}^i x_{3i}}{\sum_{i=1}^N q_{me}^i}. \quad (3.52)$$

Alternatively, the plume widths can be estimated directly without knowledge of plume location,

$$\hat{\sigma}_{x_2} = \sqrt{\frac{N}{(N-1)} \frac{\sum_{i=1}^N q_{me}^i (x_{2i} - \mu_{x_2})^2}{\sum_{i=1}^N q_{me}^i}}, \quad (3.53)$$

and similarly for σ_{x_3} . Once the dispersion and plume center parameters are estimated and a initial estimate of the source rate is established for $\hat{Q}$, the optimization of parameters θ can be undertaken, such that,

$$\hat{\theta} = \min_{\theta} J(\mathbf{X}, \theta), \quad J(\mathbf{X}, \theta) = \frac{1}{N} \sum_{i=1}^N (\mathcal{M}(\mathbf{x}_i, \theta) - q_{me}^i)^2, \quad (3.54)$$

where $\mathbf{X}$ is the observation data set $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_1, \dots, \mathbf{x}_N]$. This optimization can be carried out using the MATLAB *fminsearchbnd* function [191].

3.2.3 Mass Balance Based

The mass balance approach aims to estimate an emission source by balancing the mass flux leaving or entering a control volume. Generally, there are two path planning approaches to the mass balance method: (1) rectangular vertical flux plane (or

curtain) downwind of the source and (2) a cylindrical flux plane enclosing the source. For a well behaved plume under stable atmospheric conditions, the downwind plume contains all the flux. The sampling distance from the source may vary based on each submethod. The measured flux plane data can be sparse and is typically subject to spatial interpolation.

Vertical Flux Plane (VFP)

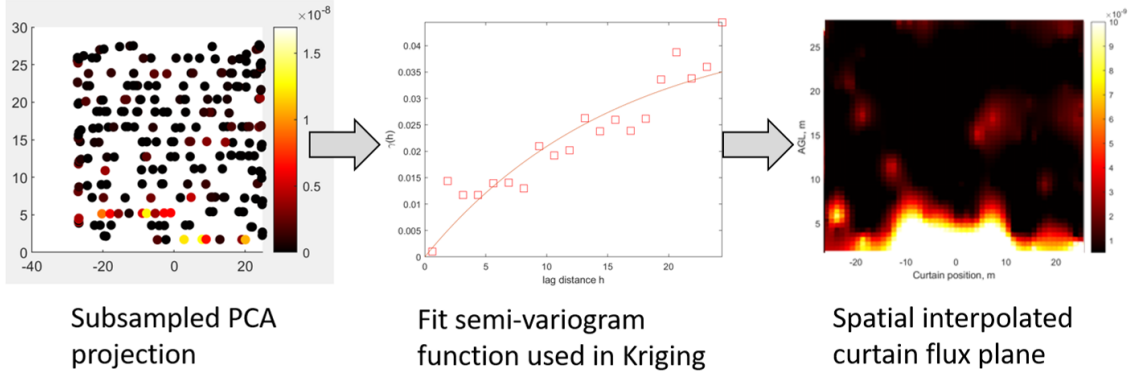


Figure 3.6: Demonstration of using sampled flux plane data and applying kriging to it for spatial interpolation [7].

The flux plane method generally involves sampling within a plane, vertically or horizontally, upwind and downwind, of an emission source. It has been applied in several works [192; 32; 29; 193; 194; 1; 195; 196; 197; 198; 199; 200; 201; 202; 203]. The plane is typically sampled using a raster-scanning approach, capturing the plume within the width and height of the plane. The emission rate (in moles s^{-1}) can be estimated as,

$$Q_c = \iint_{\Omega_p} n_{ij}(y - y_b) \mathbf{u} \cdot \hat{\mathbf{n}}_f dx_2 dx_3, \quad (3.55)$$

where n_{ij} is the mole density of air (given standard temperature and pressure), $(y - y_b)$ is the enhanced mole fraction (referenced to air), y_b is the background mole fraction, $\mathbf{u}$ is the wind speed vector, and $\hat{\mathbf{n}}_f$ is the flux plane normal vector (see Fig. 3.6). Since the measurements are sparse, the integral irregularly spaced. To combat this, the sparsely sampled points are spatially interpolated using techniques, such as inverse distance weighting (IDW) [204] or kriging [205]. This is a common problem in geostatistics to interpret unknown data, $y(\mathbf{x}_0)$, from desired spatial locations $\mathbf{x}_0$ in domain $\Omega \in \mathbb{R}^2$ (in our case is the domain of the plane $\Omega_p \in \mathbb{R}^2$), only using N sparse sampling points, $y(\mathbf{x}_i)$, based on some optimal weights, λ_i ,

$$\hat{y}(\mathbf{x}_0) = \sum_{i=1}^N \lambda_i y(\mathbf{x}_i). \quad (3.56)$$

For example, in ordinary kriging [205], a semivariogram is used to model the spatial variability and, given a spatial distance, h , is defined as,

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (y(\mathbf{x}_i) - y(\mathbf{x}_i + h))^2. \quad (3.57)$$

This experimental semivariogram can be fitted to the model semivariogram with one of several common functions: circular, spherical, exponential, Gaussian, or linear. The weights are determined by solving

$$\sum_{j=1}^N \lambda_j C(\mathbf{x}_i - \mathbf{x}_j) + \mu(\mathbf{x}_0) = C(\mathbf{x}_i - \mathbf{x}_0), \text{ for } i = 1, 2, \dots, N, \quad (3.58)$$

where $C(\cdot)$, in this context, represents the point support covariance matrix. This matrix is related to the semivariogram, $\gamma(h) = C(0) - C(h)$ [205], and the mean square prediction error is $\sigma_e^2 = \text{Var}(y(\mathbf{x}_0) - \hat{y}(\mathbf{x}_0))$, which, for ordinary kriging, is minimized to make the estimated values $\hat{y}(\mathbf{x}_0)$ optimal. Furthermore, the estimator should be unbiased (e.g., $E[\hat{y}(\mathbf{x}_0)] = E[y(\mathbf{x}_0)]$), which requires $\sum \lambda_i = 1$ and the spatial mean to be stationary $E[y(\mathbf{x})] = \mu, \forall \mathbf{x} \in \Omega$.

If the kriging process is not stationary, it is considered, at best, an approximate solution to the spatial interpolation problem and incorrect at worst. A better approach could be to apply a spectral method that takes into consideration non-stationarity and higher frequencies, namely, the high frequency kriging method [206]. Consideration of temporal observations could be included as well, see quantile kriging in [207].

An enhanced version of the IDW was proposed in [208] to include an adaptive distance-decay parameter based on the density characteristics of the sampled points. Available tools, such as Kriging Assistant (KA) [209], Golden Software Surfer, or ESRI Geostatistical Analyst for ArcMap have been used in the literature before. For irregular geographical units with different sizes and shapes, the interested reader should consult [210].

A variation of the VFP technique is illustrated in [211], where a path-averaged long open path dual-comb spectroscopy is operated from a ground vehicle to a sUAS with a retro-reflector. A vertical profile is flown downwind of the source to conduct the VFP. This technique is also very similar to VRPM.

Cylindrical Flux Plane (CFP)

A variation to the VFP is the Cylindrical Flux Plane (CFP). This method has been used with manned aircraft as it is not as easy to raster-scan a rectangular flux plane. The methodology is essentially very similar to the VFP and can be found in the work by [212], omitted here for brevity.

Path Integrated Vertical Flux Plane (PI-VFP)

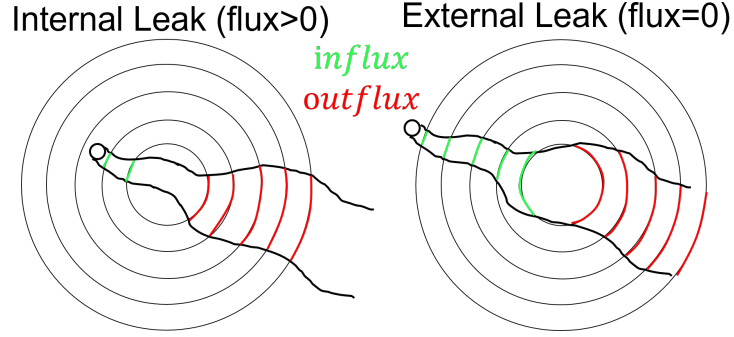


Figure 3.7: Example of the VFP-PI strategy via a UAV sensing in circular trajectories with (a) being an internal leak producing a net positive flux and (b) being an external leak producing a net zero flux. The color of the arcs are indicative of methane flux strength with green being more negative and red being more positive (see [3] for more details).

A variation of the VFP is the path integrated vertical flux plane (PI-VFP). This method utilizes a bs-TDLAS approach in that the instrument points straight down and scans or circles the emission source (see Fig. 3.7). In [213], the AVIRIS-NG manned aircraft used IMAP-DOAS technique to retrieve methane concentrations and estimated fluxes using a PI-VFP type calculation. This approach was compared with the GDT and Gaussian inverse approaches during a joint-flight campaign.

The emission rates were estimated by, $Q \approx \mathbf{u} \cdot \mathbf{n} \sum_i V_i \Delta s_i$, where V_i represents the vertically integrated concentration, and Δs_i is a path segment along the boundary. The individual measurements are integrated together (referred to as integrated methane enhancement (IME)) such that $\text{IME} = k \sum X_{CH_4}(i) \cdot S(i)$. The value X_{CH_4} is the methane plumes that exceed the minimum threshold of 200 ppm/m and k is a conversion factor.

Using an RMLD sensor fitted to a small quadrotor UAV, a circular scanning approach can be applied to sample horizontally a site of interest. The sensor uses a bsTDLAS to measure integrated methane emissions from a known height. The resulting measurements are then combined with wind measurements to estimate the flux [214; 3],

$$Q = \oint_{\mathcal{S}} \mathbf{u} \cdot \mathbf{n}_s (y - y_b) ds, \quad (3.59)$$

where $\mathbf{n}_s$ is the normal with respect to the path of travel such that $s \in \mathcal{S}$, y is the column measured concentration and y_b is the column background concentration. This calculation encompasses a single circular loop and if the source is encapsulated, multiple passes can be used to estimate the source. For instance, given n passes,

$$\hat{Q} = \frac{1}{n} \sum Q_i. \quad (3.60)$$

In practice, the circular flight path is actually made up of line segments that are box-like. The source location was also identified by course raster scanning over the area of interest followed by a more fine flight pattern free approach combined with triangular natural neighbor interpolation. The maximum observed concentration was used for the source location.

Micrometeorological Mass Difference (MMD)

Utilizing the technique from [215], sampling the plume far enough downwind of the source, the averaged MMD can be calculated as,

$$Q = \overline{\iint U_{(x_2, x_3)} (y_{(x_2, x_3)} - y_b) dx_2 dx_3} = \int \chi(x_3) dx_3, \quad (3.61)$$

where $U_{(x_2, x_3)}$ is the normal wind speed relative to the plane. The work in [216] utilized the time-average of the line-integral of the instantaneous product of U and y in the x_2 -direction. Alternatively, while using a laser fetch, an instantaneous product of a single wind measurement U and line-averaged laser concentration was used,

$$\chi \approx \Delta x_2 \overline{U_{(x_3)}} (\overline{\rho_{L(x_3)}} - y_b). \quad (3.62)$$

This method can also be used to calculate the turbulent fluxes,

$$\frac{Q_{tur}}{Q} = \frac{(Q_{\overline{Uy}} - Q_{\overline{Uy}})}{Q_{\overline{Uy}}}, \quad (3.63)$$

where $Q_{\overline{Uy}}$ is calculated from the flux term in (3.62) and $Q_{\overline{Uy}}$ in (3.64),

$$\chi \approx \Delta x_2 \overline{U_{(x_3)}} (\overline{y_{L(x_3)}} - y_b). \quad (3.64)$$

This prescription of the flux does not capture the turbulent component of the horizontal flux (albeit wrong), is often necessary due to the short time-scale behavior of the wind (e.g., limitations in wind measurement devices).

Gauss Divergence Theorem (GDT)

In the paper by [217], Conley et al. they focused on the continuity equation,

$$Q_c = \left\langle \frac{\partial m}{\partial t} \right\rangle + \iiint \nabla \cdot y \mathbf{u} dV, \quad (3.65)$$

where m is the mass of the aerosol, $\langle \cdot \rangle$ is the expectation or average, $y = Y + y'$ is the concentration (comprised of an average term and a deviation term), $\mathbf{u}$ is the wind speed, and V is the volume of the area of interest. The flux divergence can be expanded as,

$$\nabla \cdot y\mathbf{u} = \mathbf{u} \cdot \nabla y + y \nabla \cdot \mathbf{u}. \quad (3.66)$$

The surface integral is taken to be a cylinder, which can be broken into several parts: the floor, the walls of the cylinder, and the top. The height of the cylinder is taken such that the emission is encapsulated with the minimum and maximum height. The resulting emission rate can be calculated as

$$Q_c = \left\langle \frac{\partial m}{\partial t} \right\rangle + \int_0^{z_{max}} \oint y' \mathbf{u}_h \cdot \hat{\mathbf{n}} dl dx_3, \quad (3.67)$$

where x_3 represents the altitude, and l the flight path. The temporal trend of the total mass ($\frac{\partial m}{\partial t}$) within the volume can be estimated from the measurements. The cylinder passes can be vertically binned and discretely summed up,

$$Q_c = \frac{\Delta m}{\Delta t} + \sum_{x_3=0}^{x_3=x_{3,max}} \left(\sum_0^L (\rho \cdot u_n) \cdot \Delta s \right) \cdot \Delta x_3, \quad (3.68)$$

and ρ represents the scalar air density.

Vertical Flux Planes with GLM (GLM-VFP)

In [30], a 3D grid of airborne measurements are collected across multiple landfill sites. The resulting downwind observational points are then spatially interpolated with IDW and used to calculate the total mass flux. The multiple steady state Gaussian dispersion models,

$$y(\mathbf{x}) = \frac{Q}{2\pi\sigma_{x_2}\sigma_{x_3}U} \exp\left(\frac{-x_2^2}{2\sigma_{x_2}^2}\right) \left(\frac{1}{\sqrt{2\pi\sigma_{x_3}}}\right) \exp\left[\frac{-(x_3 - L)^2}{2\sigma_{x_3}^2}\right], \quad (3.69)$$

are applied to a fixed grid (50 m by 50 m), where the mixing ratios found over each individual landfill was used to calculate a model mass flux (for each site, integrated along the x_1 , x_2 , and x_3 directions). The experimental measurements are then used with simulation measurements and a general linear model,

$$\min_{\alpha} |MF - \sum_{i=1}^{max} (MMF_i \cdot \alpha_i)|, \quad (3.70)$$

to approximate the emission coefficient, α_i , from multiple landfill sources. The emission findings are further corroborated with a local eddy covariance tower measurement.

Vertical Radial Plume Mapping (VRPM)

The vertical radial plume mapping approach (compared with other methods in [218]), utilizes a long path TDLAS instrument from the ground. The laser is aimed at retro-reflectors, situated perpendicular and downwind of the source. The height of the retro-reflector constitutes the different radial angles where the path-integrated

concentrations are combined with the normal wind component to estimate the flux (similar to VFP or MMD). An illustration of this is seen in Fig. 3.2 (see [218] for more details).

3.2.4 Imaging-Based

In this section, we overview the imaging-based methodology for quantifying methane emissions. This typically includes techniques that sample images passively, such as TIR, MWIR, or other OGI-based instrumentation. The methods mentioned here that can quantify methane emissions are considered as quantitative optical gas imaging (QOGI).

Mid-Wave Infrared (MWIR) and Hyperspectral

In the work by [219], the detection limits of MWIR band of a hyperspectral data was explored using the Spatially-Enhanced Broadband Array Spectrograph System (SEBASS) airborne instrument. They also provided a comparison between LWIR and MWIR (see [219] for more details) using the radiative transfer model,

$$R_s = (R_T^\uparrow + R_S^\uparrow) + t\{\epsilon_s B(T_s) + (1 - \epsilon_s)\left[\frac{R_T^\downarrow + R_S^\downarrow}{1 - S(1 - \epsilon_s)}\right]\}, \quad (3.71)$$

where R_s is the total radiance at the sensor, $R_T^\uparrow$ is the upwelling emitted atmospheric path radiance, $R_T^\downarrow$ is the downwelling emitted atmospheric path radiance, $R_S^\uparrow$ is the scatter path radiance at the sensor, $R_S^\downarrow$ total solar radiance that reaches the surface, t is the atmospheric transmittance, ϵ_s is the surface emissivity, and $B(T_s)$ is the blackbody radiation at the surface temperature.

Other works, such as [150], have used MWIR cameras combined with two Pergam Methane Mini G lasers in pipeline leak detection. In [220], a FLIR GF320 and a RMLD were used together to make volumetric flow rate calculations in the laboratory using a data fusion approach. In [221], they utilized a thermal camera and steady state energy balance approach to estimate methane emissions from thermal anomalies in urban landfills.

Iterative Maximum a Posteriori Differential Optical Absorption Spectroscopy (IMAP-DOAS)

The IMAP-DOAS method was applied to the AVIRIS-NG [27; 26] aircraft and measures reflected solar radiation between 0.35 μm and 2.5 μm with 5 nm spectral resolution and sampling. Using a nonlinear iterative minimization of the differences between modeled and measured radiance. The measured concentrations can be applied to the PI-VFP method to calculate fluxes [213]. Variations in this approach for retrieving methane concentrations has been seen in [222] for albedo correction and [34] anomaly-based mass balance.

3.2.5 Correlation-Based

Tracer Correlation (TCM)

The tracer correlation method, or isolated source tracer ratio method, initially proposed and implemented in works by Lamb et al [223] and Czepiel et al [224], aims to quantify the emission rate of an unknown gas species by releasing a tracer gas at a known flow rate while measuring both the tracer and the unknown signals collocated downwind. This method assumes that the location of the source is known and, at the measurement location, the plume is well mixed. The elevated signal downwind also needs to typically be greater than 50 ppb. The authors report uncertainty estimates of $\pm 15\%$. The general equation is given as

$$Q_m = Q_t \frac{y_m}{y_t}, \quad (3.72)$$

where Q_t is the tracer release rate, and y_m and y_t are the elevated mixing ratios of the unknown source gas and tracer gas, respectively. A comparison study between TCM and other fugitive emission quantification methods are studied in [218]. The effect of wind on accuracy of the TCM was explored for landfills using WRF model [225]. An in situ method was used to evaluate the collection efficiency of gas extraction wells based on tracer gas [226].

Variations of the quantification of TCM were explored in [24], which quantified emission rates based on the plume integration of a transect, peak height of the transect using a scatter plot to calculate the ratio (best fit line), and comparison with fitted Gaussian plume model. A landfill field comparison of methane emission models were compared to measured emissions using TCM [227]. The TCM method was also applied to quantifying emissions from dairy farms in [172].

A dual tracer method was explored in [228]. The second tracer provides for closer downwind measurements that can be refined by assessment of plume position as well as in the far-field measurements the second tracer becomes an internal standard to the measurement. A mobile version of the TCM approach was proposed in [229].

Eddy Covariance (EC)

The Eddy covariance method aims to estimate the emission flux from a footprint area given the boundary layer meteorology. Historical developments and current implementations of this method are summarized in [230]. This method generally assumes stationarity of the measured data and fully developed turbulent conditions [231]. One way it can be expressed is,

$$Q = \frac{1}{t_f - t_i} \int_{t_i}^{t_f} (y(t) - \bar{y})(w(t) - \bar{w}) dt, \quad (3.73)$$

where the time-averaged concentration and vertical wind speed is $\bar{y}$ and $\bar{w}$, respectively. There are several assumptions required to make this flux calculation (for more details see [232]).

Table 3.1: Figures of merit for defining complexity of an estimation method

FOM	(%)	Low (2.5)	Medium (5)	Med-High (7.5)	High (10)
Operator skill	30	Little	Moderate	Professional	Expert
Number of operators	25	1	2	3	3+
Equipment cost	15	<\$10,000	<\$50,000	<\$100,000	>\$100,000
Setup Time	20	<1hr	<4hr	<8hrs	8+hrs
Survey Time	10	<0.5hr	<1hr	<2hrs	2+hrs

3.3 Analysis of Methods

In an attempt to analyze the methods covered in the previous section we decided to use the following Figure of Merit metrics, such as: required assumptions, sample distance, survey time, complexity, average precision, average accuracy, and average cost. The required assumptions are meant to inform the practitioner so the best method can be applied to a given problem. For example, if the source location is unknown, the PSG method may not be directly applicable unless a source location estimate is supplied. The sample distance is defined as the distance from the source at which the required method is able to take measurements from. The survey time consists of the time required to make a single flux estimate. Understandably, some methods may require multiple flux estimates in order to approximate the emission source to within an acceptable error. Complexity is the measure of how difficult it is to implement any given method. In order to determine a value for complexity, a scheme was developed using figures of merit (FOM) that assigns factors and weights to the metrics (detailed in Table 3.1). Determining the values for these factors were based on loose estimates, inferred from papers found in the literature. Ranges were assigned to the metrics, as to capture variations in the factors due to either the operators or the equipment being used, and are given in Table 3.2. For example, some setups may use more expensive equipment or more people for the same method, and as a result are reflected in the complexity metric.

Evaluating methane quantification techniques is important and much work has already gone into this topic through controlled release experiments and evaluation frameworks. Examples from controlled release facilities (CRF) consist of but are not limited to the following: In the Joint Urban 2003 study [233; 234], static sensors were distributed in an urban setting to measure the dispersion of tracer particulates; In [235], area-averaged velocity and turbulent kinetic energy profiles are derived from data collected at the Mock Urban Setting Test (MUST); Mock Urban Setting test (MUST) was also evaluated with photo-ionization detectors (PID) [236; 237]; MUST was further simulated using MISKAM 6 [238]; In [239], the WRF model was used to model wind and turbulence inside the Quick Urban & Industrial Complex (QUIC) model for comparing simulated and observed plume transport; A test plan for Jack Rabbit II was developed in [240] which aimed to improve chemical hazard modeling, produce better planning for release incidents, improve emergency response, and

improve mitigation measures.

More recently, single-blind tests at the Methane Emission Technology Evaluation Center (METEC) in Fort Collins, Colorado evaluated several types of LDAQ sensing modalities as apart of the Stanford/EDF Mobile Monitoring Challenge (MMC) and the Advanced Research Projects Agency-Energy (ARPA-E) MONITOR program (such as by vehicle, plane, and drone - see [196; 241] for more details). In the Stanford EDF MMC it was observed that the drone based technologies performed quite well (e.g. SeekOps) with an $R^2 = 0.42$ [196]. While, the results in [196; 241] seem quite promising, there is still exists some improvements in precision that can be done. In the ARPA-E MONITOR program, 6 of the 11 participants tested their technologies at the METEC facility in [242] against 6 other industry-based participants. Due to confidentiality agreements at the time of testing, the data gathered from the 12 participants were aggregated to compare the methodologies based on measurement type (handheld, mobile and continuous monitoring). However, to the best of authors knowledge only four of the MONITOR program participants have published data regarding the METEC tests (see [243; 244; 245; 241]). In the white paper by Bridger Photonics [241], a sUAS-based approach using LiDAR-based sensor has also shown promising results even though the uncertainty is not given. In [3], a RMLD is used on a sUAS with the PI-VFP method. In contrast, [243] utilized a portable TDLAS-based instrument and the PSG method to quantify emissions. Lastly, [244] uses a dual frequency comb spectrometer (from over one kilometer away) with the non-zero minimum bootstrap method (see [245]) and the Gaussian plume model to estimate the source rate. Examples from active operations with comparison to conventional OGI-based methods are conducted in the Alberta Methane Field Challenge (AMFC) [200; 246; 13] which aim to answer the questions: Are Leak detection and repair (LDAR) programs effective at reducing methane emissions and can new technologies provide more cost-effective leak detection compared to existing approaches?

In order to compare the performances of the each of the methods to one another, their performance metrics were garnered from different studies where the method was utilized in either a field study or a controlled release scenarios and recorded in Table 3.2. Performance values were gathered from the standard deviations of consecutive flux estimates of a singular source leak scenario. Accuracy pertains to error of the flux estimate to the known source rate. This information was limited primarily to controlled release scenarios. For each method, performances and details were separated into the broad types of sampling strategies: fixed/static, on foot, mounted on a vehicle, mounted on an aircraft, and mounted on a sUAS. This prevents convolution of performance values between, for example, long aircraft sampling flights at far distances and short sampling flights near the source via sUAS.

3.3.1 Summary of Methods

After analyzing the quantification methods we can separate methods based on whether they have used sUAS or not. In this manuscript we observed that the sUAS-based

Table 3.2: Summary of methods and their assumptions, operational details, complexity, cost, average precision, and average accuracy are generalized over implementations found in the literature. The average precision and accuracy are given as unitless values, normalized by the true source rate (estimated source rate). (measured using ¹fixed, ²foot, ³vehicle, ⁴manned aircraft, or ⁵UAV; cost with \$ $\leq$ \$10,000, \$\$ $\leq$ \$50,000, \$\$\$ $\leq$ \$100,000, \$\$\$\$ $\geq$ \$100,000; (.) represents precision normalized on estimated source rate average complexity.

Method	Assumptions	Sample tance	Dis-	Survey Time	Complexity (1-10)	Avg sion	Preci- racy	Accu- racy	Avg Cost
bLS	x_s , horizontally uniform surface source atmosphere in horizontal equilibrium	20-441m ¹		15min-4days ¹	3.8-8.5 ¹	$\pm 0.16-0.36^1$ ($\pm 0.07-0.85$) ¹	$\pm 0.02-0.30^1$		\$-\$\$\$ ¹
PSG	x_s , steady state source rate, point source, plume evolution via ground-level Gaussian dispersion with no obstructions	441m ¹ 500m ³ 50m ⁵	18-	4days ¹ 58min ³ 20min ⁵	4.5-8 ¹ 3.6-6.1 ³ 3.3-5.5 ⁵	$\pm 0.0022-0.43^3$ $\pm 0.50^5$			\$\$-\$\$\$ ³
PSG-RB	PSG assumptions, vertical eddy diffusivity and wind speed approximated by power law scheme	20-200m ³		6min ³	3.6-6.4 ³	—	$\pm 0^3$		\$\$-\$\$\$
PSG-CS	PSG assumptions, continuous source emission, constant wind speed, vertical eddy diffusivity and wind speed approximated by power law	18-106m ³		20min ³	3.6-5.9 ³	$\pm 0.20-0.67^3$ ($\pm 0.19-0.47$) ³	$\pm 0.02-0.26^3$		\$\$-\$\$\$ ³
NGI	constant source rate, σ_x and σ_y linear functions of distance to source	50-82.25m ⁵		7.35- 29.62min ⁵	3.3-5.5 ⁵	$\pm 0.21-0.58^5$ ($\pm 0.06-0.53$) ⁵	$\pm 0.11-0.13^5$		\$-\$\$\$ ⁵
MMMD	x_s	12-27m ¹		15min ¹	4.5-8.1 ¹	$\pm 0.06^1$	$\pm 0.10^1$		\$\$\$-\$\$\$\$ ¹
GDT	near to no meandering, steady state source rate	3-8km ⁴		1hr ⁴	5.5-7.8 ⁴	$\pm 0.07^4$ (± 0.08) ⁴	$\pm 0.13^4$		\$\$\$-\$\$\$\$ ⁴
VFP	GDT assumptions	4.875-10km ⁴ 19.08-510m ⁵		1.5-4.5hr ⁴ 30min ⁵	5.8-8.3 ⁴ 5.5 ⁵	$\pm 0.30-0.53^4$ $\pm 0.17-0.37^5$ ($\pm 0.013-0.62$) ⁵	$\pm 0.10-0.50^4$ $\pm 0.03-0.50^5$		\$\$\$-\$\$\$\$ ⁴ \$- \$\$ ⁵
PI-VFP	GDT assumptions	3km ⁴ 0-6.77m ⁵		20-30min ⁴ 15- 20min ⁵	5.3-7.5 ⁴ 3.3- 5.5 ⁵	$\pm 0.34-0.58^4$ $\pm 0.82^5$	$\pm 0.27-52^5$		\$\$\$-\$\$\$\$ ⁴ \$- \$\$ ⁵
CFP	GDT assumptions	3-17.84km ⁴		1hr ⁴	5.5-7.8 ⁴	—	—		\$\$\$-\$\$\$\$ ⁴
GLM-VFP	GDT assumptions	0.4-2.2km ⁴		2.5hr ⁴	6-8.3 ⁴	(± 0.21) ⁴	—		\$\$\$-\$\$\$\$ ⁴
VRPM	GDT assumptions	10-100m ¹		1hr ¹	7.1-8.6 ¹	$\pm 0.18-0.21^1$ ($\pm 0.21-0.33$) ¹	$\pm 0.05-0.43^1$		\$-\$\$\$ ¹
QOGI	temperature and pressure of gas at leak location are the same, gas plume length in direction of optical path is small	30m ¹		1min ¹	3-6.3 ¹	$\pm 0.01-0.02^1$ ($\pm 0.02-0.03$) ¹	$\pm 0.20-0.24^1$		\$-\$\$\$ ¹
TCM	leak plume and tracer plume are well mixed	100-3546m ³		0.5-2hr ³	4.6-7.6 ³	$\pm 0.06-0.24^3$ ($\pm 0.06-0.74$) ³	$\pm 0.0056-0.17^3$		\$\$\$-\$\$\$\$ ³
EC	stationarity, fully developed turbulent conditions	25-228m ¹		—	4.5-7.8 ¹	(± 0.08) ¹	—		\$-\$\$\$ ¹

methods consist of: Near-field Gaussian plume inverse (NGI), vertical flux plane (VFP), and the path-integrated vertical flux plane (PI-VFP). The non-sUAS-based methods consist of: backwards Lagrangian Stochastic (bLS), point source Gaussian (PSG), recursive Bayesian point source Gaussian (PSG-RB), conditionally sampled point source Gaussian (PSG-CS), micrometeorological mass difference (MMD), Gauss divergence theorem (GDT), VFP, PI-VFP, cylindrical flux plane (CFP), general linear model vertical flux plane (GLM-VFP), vertical radial plume mapping (VRPM), quantitative optical gas imaging (QOGI), tracer correlation method (TCM), and Eddy covariance (EC). When comparing their performances in Table 3.2, it can be seen that, when categorizing by means of mobility (i.e. fixed, on-foot, etc.), methods using static sensors show a trend of having higher complexity values while UAV-based methods display generally lower complexity values. For a subset of the methods, the survey times, sample distances, and average accuracies can be seen in Fig. 3.8. This subset was specifically displayed for these methods had both upper and lower bounds for survey times and sample distances along with accuracy data which allowed for the plotting of these quantities for each method in the form of ellipses on a log-log plot. When analyzing this plot, it can be seen that the sUAS-based methods are generally lower in sample distances and survey times as opposed to manned aircraft-based methods (being the one of the highest in both). The bLS and TCM methods are shown to have the best average accuracy with several sUAS and mobile methods close in accuracy. The long sample times of bLS method are due to the values reported in [247], and it is possible that these values don't reflect typical bLS sample times. The advantages and disadvantages of each of the methods can be seen in Table 3.3 along with what typical application fields that they were applied in.

The final ranking of the methods depends heavily on the desired application, which also depends on factors such as sample distance, sample time, and desired accuracy. For that reason, it is difficult to rank the methods in general. Thus, we provide a ranking of the methods in terms of complexity (outlined in Table 3.1) with highlights from the precision and cost in Fig. 3.9. The results indicate that the simplest methods, in terms of complexity, are the sUAS-based NGI⁵ and VFP⁵ as well as fixed QOGI¹. The most complex methods include bLS¹ and manned aircraft-based approaches. In terms of precision, bLS¹, NGI⁵, GDT⁴, VFP⁵, QOGI¹, TCM³, and EC¹ tend to be the best. Thus, for sUAS-based methods, NGI⁵ and VFP⁵, are the most promising approaches. Additionally, the GDT⁴, TCM³, and EC¹ approaches can be treated as candidate methods for future implementation using sUAS.

3.4 Case Studies: Leak Detection and Localization

In this section we overview leak detection and localization experiments carried out using sUAS, verifying detection capabilities and back trajectory heuristics. Additionally we overview a simulation based approach at autonomous leak detection/localization and test the algorithm's localization capability using experimental data obtained using a hand-held sensor on foot.

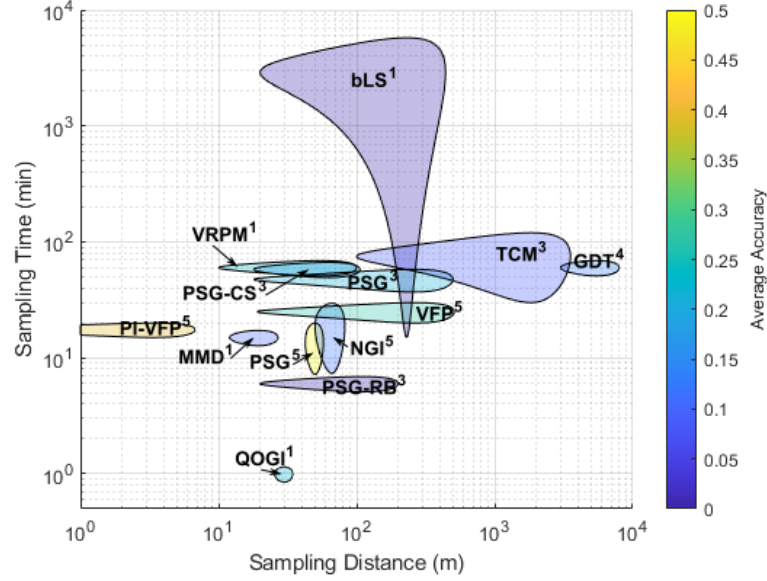


Figure 3.8: Diagram of summary of methods (based on Table 3.2) showing relationship between typical survey time versus sample distance and there associated normalized accuracy, where lower values represent more accurate measurements. (measured using ¹fixed, ²foot, ³vehicle, ⁴manned aircraft, or ⁵UAV)

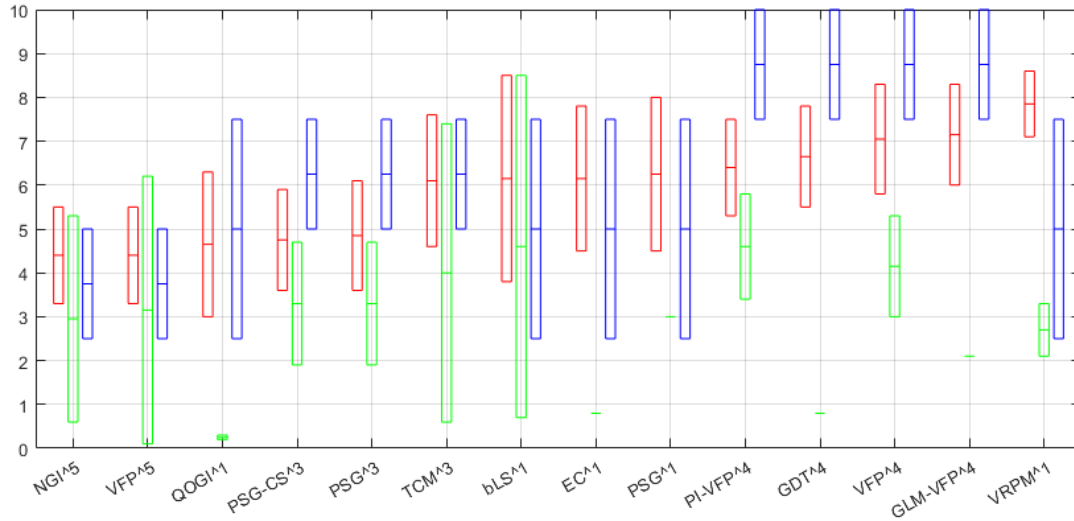


Figure 3.9: Diagram of the complexity ranking of the methods (based on Table 3.1), showing the relationship between the method complexity (red), precision (green), and cost (blue). The precision is normalized on the source estimate multiplied by 10 and the cost is ranked from 0 to 10. (measured using ¹fixed, ²foot, ³vehicle, ⁴manned aircraft, or ⁵UAV)

Table 3.3: Summary of method advantages and disadvantages along with fields of application

Method	Application	Advantages	Disadvantages
bLS	General, Oil and Gas, Agriculture	able to quantify area source and point source emissions	sensor is fixed, multiple measurements, negatively impacted by obstacles
PSG	Biogas, Oil and Gas	No tracer required	Negatively impacted by obstacles and low wind speeds due to plume advection, mobile sensors limited to roads
PSG-RB	Oil and Gas	No tracer required, accurate source quantification in open environments	<i>PSG limitations</i>
PSG-CS	Oil and Gas	No tracer required	<i>PSG limitations</i>
NGI	General, Oil and Gas	assumes near-field plume turbulence and wind meandering	near-field (100m), prop-wash interference if unaccounted for
MMD	Agriculture	able to give instantaneous flux estimates	fixed sensors
GDT	Regional	capable of giving emission quantifications of large areas	Sample around closed volume typically large areas, unable to capture instantaneous methane flux
VFP	General, Oil and Gas, Landfill	does not require exact source location, ease of mobility	stable atmospheric conditions, unable to capture instantaneous methane flux
PI-VFP	General, Oil and Gas	<i>VFP advantages</i>	<i>VFP limitations</i>
CFP	General, Urban	capable of giving emission quantifications of large areas	<i>GDT limitations</i>
GLM-VFP	Landfill	<i>VFP advantages</i>	<i>VFP limitations</i>
VRPM	Landfill	able to give instantaneous flux estimates	fixed sensors
QOGI	Biogas	able to give instantaneous flux estimates	gas velocity determined via gas camera - velocity component parallel to image plane, can be difficult to process images
TCM	General, Oil and Gas, Landfill, Urban	does not rely on meteorological measurements	mobile sensors limited to roads, application difficulty due to outside methane source interference
EC	General, Oil and Gas	can capture emission variations due to long time series	fixed sensor(s), stable atmospheric conditions, sensitive to time of day, typically requires long sampling times

The methane sensor utilized by this project is the Open Path Laser Spectrometer (OPLS) instrument. It is a tunable diode laser spectrometer entirely designed, built, and tested at JPL. The diode laser, which only became commercially available in the last few years, emits mid-infrared (IR) light near $3.27 \mu\text{m}$ (3057 cm^{-1}). This allows the OPLS sensor to probe the strongest available absorption signals of methane. Prior tunable diode laser spectrometers utilized light in the near-IR at $1.6 \mu\text{m}$ for which the methane signal is nearly $50\times$ weaker as shown in Fig. 3.10.

In addition to the new type of laser, OPLS utilizes miniaturization of electronics and mechanical structure developed by NASA for the Mars Curiosity Rover Tunable Laser Spectrometer instrument currently measuring methane (and many other chemicals) on Mars. A key design implementation of OPLS is that it is open-path as shown in Fig. 3.10. This enables OPLS to be lightweight (no pump nor inlet controllers) and make measurements nearly as quick as the time it takes air to flush out the optical head. Under typical wind speeds and flight operations, this results in a 0.1 second instrument response time. This makes OPLS comparable to the LICOR 7700 but with OPLS being $20\times$ less mass with nearly identical sensitivity. The platform utilized for these experiments was the 3DR Solo, as a part of work done with NYSEARCH. There are several reasons for using this platform. It has proven to be a reliable platform with easy-to-use, vetted software and little down time for maintenance. Second, the platform and spare parts are inexpensive. Third, the platform is relatively small at 1.8 kg. Fourth, the 3DR Solo uses an open-source Pixhawk autopilot that enables the research team to implement ancillary sensors like the ground-ranging LIDAR. The LIDAR has proven invaluable for accurate height information from which to better

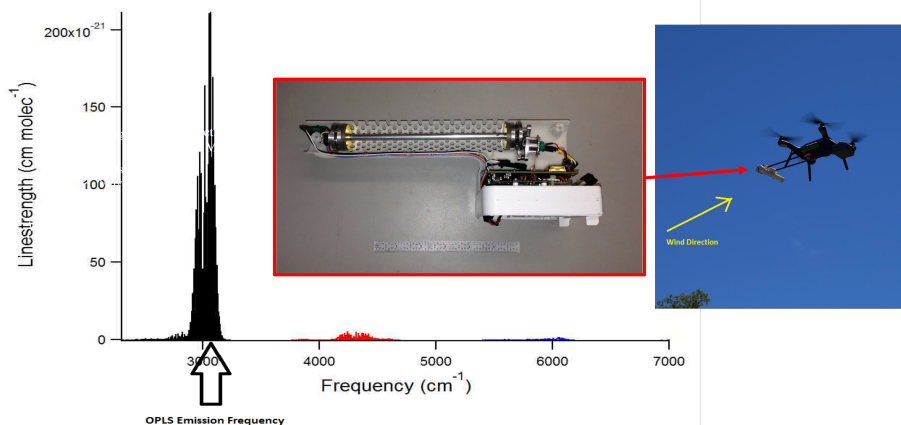


Figure 3.10: Image showing the OPLS emission frequency ($3057 \text{ cm}^{-1} = 3.27 \mu\text{m}$), the OPLS sensor showing the open path optical head, and the integrated OPLS on the 3DR Solo quadcopter hovering and oriented into the wind

understand the data. A key flight implementation aspect is that the OPLS sensor is placed outside the propeller wash of the quadcopter (see [170]). Tests in 2015 with smoke generators and wind tunnels demonstrated that placing the quadcopter just a few centimeters in front of a line between the centers of the forward facing propellers (with respect to the wind or flight direction) [170]. Typical flight orientations are as shown in Fig. 3.10 where the OPLS sensor is pointed into the wind, if the platform is not in motion, or in the direction of flight if the platform is in motion, such that the wind data is modified by the 5-min moving average.

3.4.1 Merced Vernal Pools Grassland Reserve

During February 14-15, 2017, a field experiment with a controlled release was conducted at the Merced Vernal pools and Grassland Reserve (MVPGR) in Merced, CA, directly northeast of the UC Merced Campus (see Fig. 3.11). UC Merced has gained special permission from the reserve director to perform test flights at this location, and we are free to move about the 6,500 acre reserve to find suitable locations for testing. The current test location chosen is south of an unused barn, and approximately 0.5 mi from the UC Merced campus.

Temperatures during testing ranged from 13°C to 18°C . Wind speeds ranged from 0 to 3 m s^{-1} , generally from all directions. The combination of fairly low wind speeds and variable wind direction necessitated longer time in the field to adjust for this variability. There was no precipitation. Cloud coverage was 20% during the entire experiment. A cylinder of 100% methane was placed on location as shown in Fig. 3.11. A regulator was used to bring the pressure down to 15 PSI and the methane gas was flowed into a perforated Teflon tube embedded in coarse gravel as shown in Fig. 3.11. Flow rates ranged from 0.1 to 25 SCFH. A sonic anemometer (R.M. Young 81000) was placed near the methane source. It transmitted wind data

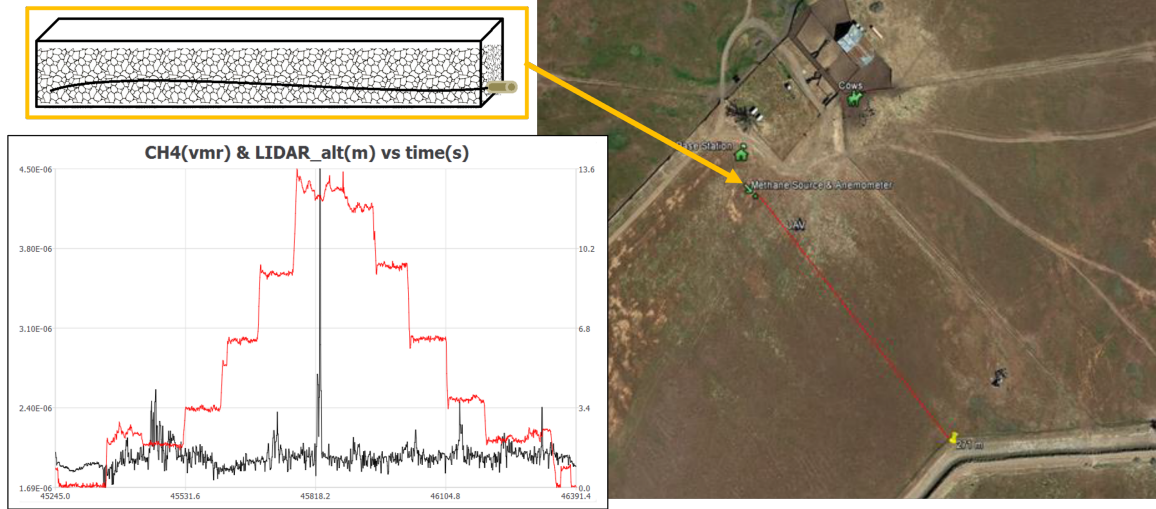


Figure 3.11: The right shows a satellite view of MVPGR test site. The buildings to the north are an abandoned farm house. The red line is 270 m long. On the left, shows the gas diffuser setup and data from a 30 m downwind flight with a 5 SCFH methane release.

(direction and speed) wirelessly to the ground computer control station. This wind data was interpolated into the data acquired by the quadcopter and methane sniffer.

The controlled release campaign consisted of 27 flights at several different leak rates (shown in Table. 3.4). Typically, flights were 30 meters away from the leak source at altitudes ranging from 3 to 25 meters above ground. Two types of flight profiles were flown. Downwind flights consisted of flying downwind of the leak source at 30 meters distance along a 50-m long line whose center point was parallel to the average wind direction for the preceding 5 min. For the downwind flights, several altitudes were flown in order to understand the vertical extent of the methane plume. For those cases where a discernable upper altitude could be determined, mass-balance analysis can be performed. Flight speeds were between 2 and 4 m s⁻¹. The sensor was always oriented facing the emission source. Average flight times were 9 minutes.

The flights conducted here occurred with wind speeds 1.3 ± 0.9 m s⁻¹ with wind direction variance ($\sigma\theta$) of 65°. A typical flight pattern where different altitudes were investigated as shown in Fig. 3.11.

It is helpful to contrast the Feb. 14-15, 2017 measurements with measurements taken in August 2016 for another project under similar conditions except that the wind was more stable (2.8 ± 0.4 m s⁻¹, $\sigma\theta = 25^\circ$). The August 2016 measurements were taken at the same location, at the same methane flowrate but under more laminar wind conditions. In the case of August 2016, the occurrence of methane signal above 10 m altitude was nearly zero whereas in February 2017, 20% of the passes downwind of the emission resulted in a measureable methane signal like that observed in Fig. 3.11.

Data from both experiments were analyzed for the fraction of downwind passes

Table 3.4: Testing matrix for February 14-15 Merced Vernal Pools controlled release test experiments. The flow rate is given in standard cubic feet per hour (SCFH).

Flt #	Date	Type	Description	Flow Rate	Mass Balance?
1	Feb 14	Downwind	SE of leak	5	Yes
2	Feb 14	Downwind	SE of leak	25	Yes
3	Feb 14	Downwind	SE of leak	25	Yes
4	Feb 14	Downwind	SE of leak	25	Yes
5	Feb 14	Downwind	NW of leak	25	Yes
6	Feb 14	Downwind	SE of leak	25	Yes
7	Feb 14	Downwind	SW of leak	5	Yes
8	Feb 14	Downwind	W of leak	5	Yes
9	Feb 14	Downwind	SW of leak	5	Yes
10	Feb 14	Downwind	NW of leak	5	Yes
11	Feb 14	Downwind	NW of leak	1	Yes
12	Feb 14	Downwind	NW of leak	1	Yes
13	Feb 14	Survey	near leak	1	No
14	Feb 15	Flt survey	near leak and cows	5	No
15	Feb 15	Flt survey	near leak	5	No
16	Feb 15	Downwind	NW of leak	0.2	No
17	Feb 15	Downwind	NW of leak	0.2	No
18	Feb 15	Flt survey	near leak	0.2	No
19	Feb 15	Downwind	NW & SW of leak	5	Yes
20	Feb 15	Flt survey	near leak	5	Yes
21	Feb 15	Flt survey	near leak	5	No
22	Feb 15	Flt survey	near leak	1	No
23	Feb 15	Flt survey	near leak	7	No
24	Feb 15	Flt survey	near leak	0.4	No
25	Feb 15	Flt survey	near leak	12	No
26	Feb 15	Flt survey	near farm house	NA	No
27	Feb 15	Flt survey	near farm house	NA	No

Table 3.5: Number of passes needed at Pasquill-Gifford atmospheric stability class and above ground level (AGL) altitude to gain 95% due diligence confidence that a 5 SCFH leak was discovered.

Date	Class	Ave wind (m/s)	Distance(m)	AGL(m)	Detection/ Pass	#Passes → 95%
Feb. 2017	B	1.3	30	5	0.55	4
Feb. 2017	B	1.3	30	10	0.32	8
Aug. 2016	C	2.8	30	5	0.85	2
Aug. 2016	C	2.8	30	10	0.08	36

that had clear leak indications. This meant that when flying 30 m downwind of the 5 SCFH leak, along a 50-m long line perpendicular to wind direction, we measured how often the methane signal > 100 ppb was observed (above background). An examination of these detection fractions as a function of altitude from both August 2016 and February 2017 are shown in Fig 4.4. During the laminar conditions of August 2016, at 5 m altitude, nearly 85% of passes indicated the presence of a leak. However, in the more turbulent, less atmospheric stable conditions of Feb. 2017, at 5 m altitude, only 55% of passes indicated a leak. Thus to obtain the number of passes required to obtain 95% confidence that the upwind area was appropriately sampled for leaks (N95), more passes are needed for the unstable conditions, which have lower fraction of passes with detections (P), as shown in Table. 3.5.

However, while unstable atmospheric conditions are not conducive to direct downwind monitoring, such conditions greatly increase the chance of detection at higher altitudes. The unstable atmospheric conditions also affected the efficacy of leak localization. This is shown in Fig. 3.12 for 4 out of 15 downwind localization test flights. These four examples exemplify the range of results for localization. The perpendicular (shortest) distance from the leak source to the average back-trajectory line varied widely from 2 to 30 meters. In contrast, for the more laminar conditions of Aug. 2016, the average distances between the back-trajectory line and the leak location varied between 0 and 5 m. In summary, there is a huge effect on the efficacy of back-trajectory leak localization and due diligence leak detection between an average wind speed of 1.3 m s^{-1} versus 2.8 m s^{-1} .

3.4.2 PG&E Livemore Test Facility

During March 27-29, 2017, flights were conducted at a simulated neighborhood at the Pacific Gas & Electric (PG&E) Livermore training facility. The Livermore training facility allows for single and multiple controlled release points of natural gas from specific locations inside the simulated neighborhood (see Fig. 3.13).

There are 16 houses, 4 in each row. Each house is within 5 meters, of the actual leak location. They are numbered 1-16 for readability of the test matrix. Fig. 3.13 bottom show actual leak locations. On March 27 (Day One), tests were done where the leak locations were shared between the test operators (NYSEARCH + PG&E)

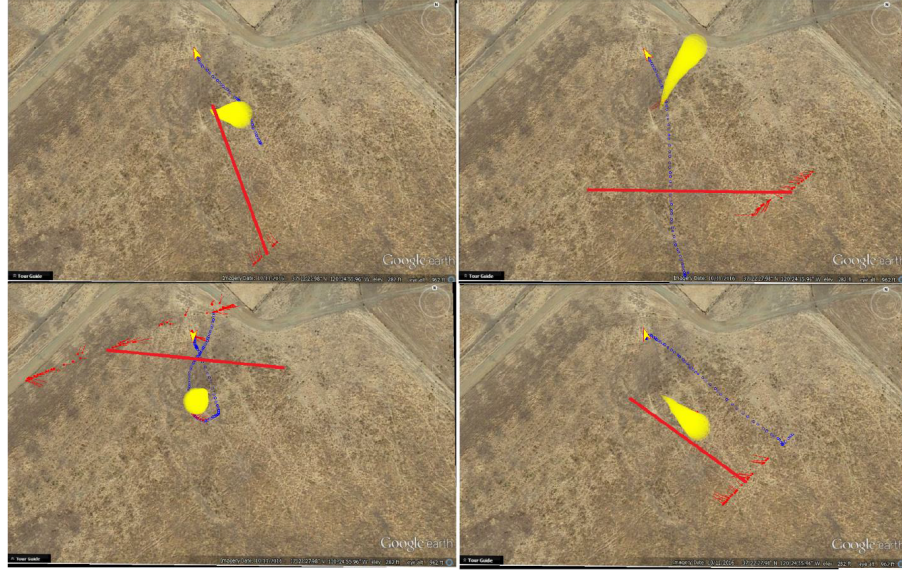


Figure 3.12: Localization back-trajectories from Feb. 2017 Merced Field test. Shown are 4 examples. Small red lines are back-trajectories from individual methane signals. Fat red lines are the averaged back-trajectories for that flight. The yellow is the predicted plume where the leak location is at the sharp end of the yellow plume.

and the sUAS test team (JPL and UC Merced). One or two leak locations were emitting simultaneously during these tests. On March 28 (Day Two), blind tests were conducted in which the test operators opened up one or more leaks and the sUAS test team had to identify the leak location and how many leaks were present. One through four simultaneous leak locations were emitting during these blind tests. On March 29 (Day Three), leak locations were shared between test operators and the sUAS team. One through four leak locations were simultaneously emitting during these tests.

The test matrix is shown in Table. 3.6. Large leaks refer to leaks estimated to be > 10 SCFH. Medium leaks refer to leaks 5-10 SCFH. Small leaks are < 5 SCFH. It should be noted that actual leak rates were not determined for these tests and leak rates estimated from past experience by PG&E personnel.

For blind tests, the whole 16 house area was surveyed. Then, the areas where there were methane signals were investigated. During tests, the issue of leak localization versus leak pinpointing arose. On average, leaks were discovered and localized to ± 2.5 meters within 6 minutes, using a whole neighborhood scanning approach. However, to pinpoint the leak location to < 1 m typically resulted in another 10+ minutes of flight time. There were several instances during tests where the wind direction shifted nearly 120° over the course of minutes. This turned out to be advantageous for leak detection because it enables the leak surveyor to triangulate the position of the leak. A good example of this is test #5 shown in Fig. 3.14. Early in the test (Fig. 3.14a), back trajectories of high methane signal suggested a leak location

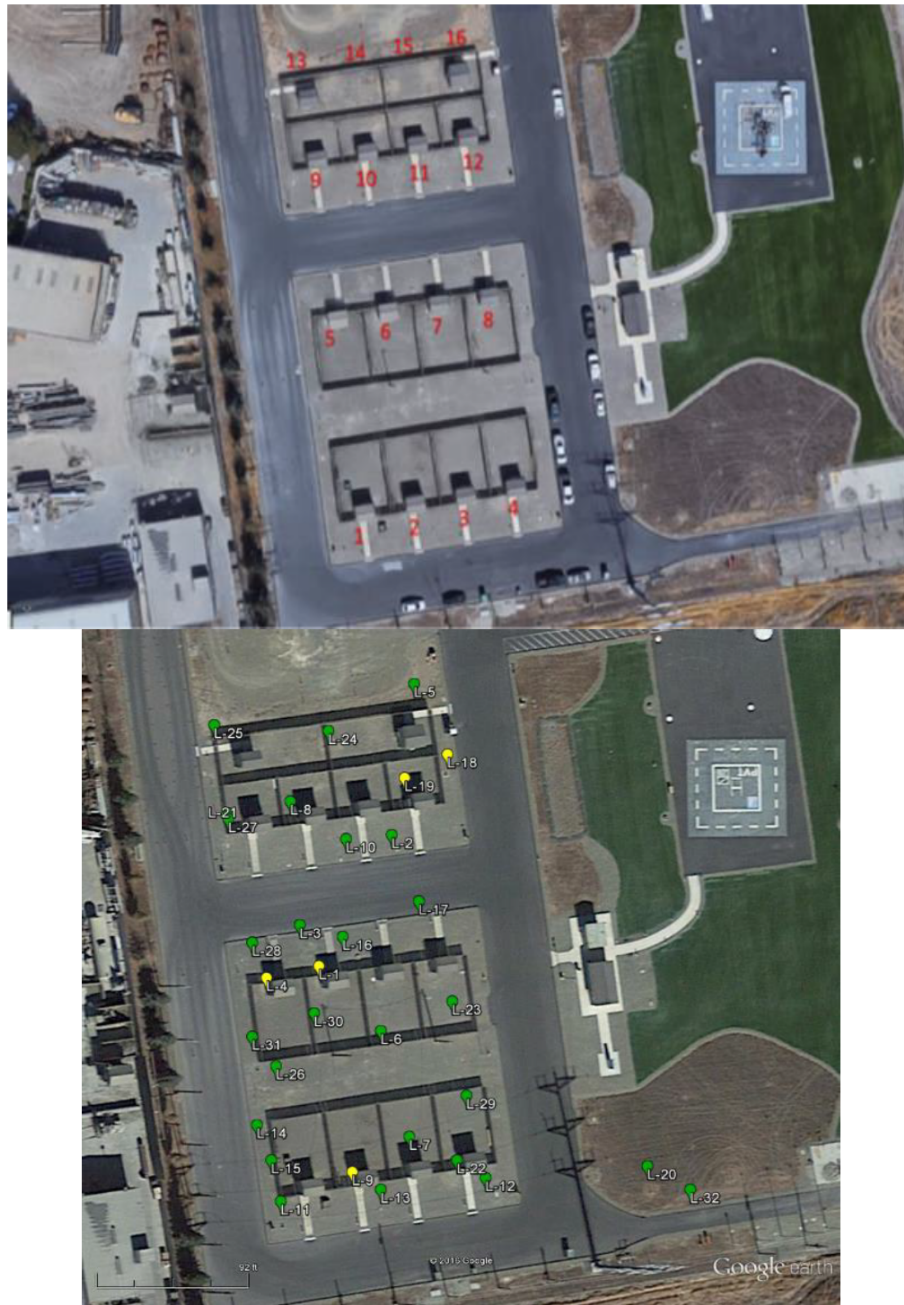


Figure 3.13: The PG&E Livermore training center. Top: The numbers in red refer to release points that map to the test matrix provided above. Bottom: Actual release points where green are below-ground leaks and yellow are above-ground leaks.

Table 3.6: Test Matrix for PG&E Livermore Testing Facility Field Experiments.

Test #	Date	# Leaks	Leak Locations	Leak size	Blind
1	Mar 27	1	1	Large	No
2	Mar 27	1	9	Large	No
3	Mar 27	1	16	Large	No
4	Mar 27	1	4	Med	No
5	Mar 27	1	5	Med	No
6	Mar 27	1	13	Med	No
7	Mar 27	1	6	Small	No
8	Mar 27	1	7	Large	No
9	Mar 27	1	15	Medium	No
10	Mar 27	2	1, 8	Large	No
11	Mar 27	2	9,16	Large	No
12	Mar 27	2	4,8	Large	No
13	Mar 27	1	1	Med	No
14	Mar 28	1	6	Large	Yes
15	Mar 28	1	11	Med	Yes
16	Mar 28	2	4,16	Small	Yes
17	Mar 28	3	6,8,11	Small	Yes
18	Mar 28	3	5,10,13	Small	Yes
19	Mar 28	2	3,9	Small	Yes
20	Mar 28	4	4,6,10,15	Small	Yes
21	Mar 29	1	5	Large	No
22	Mar 29	4	5,6,7,8	Large	No
23	Mar 29	3	4,8,12	Large	No

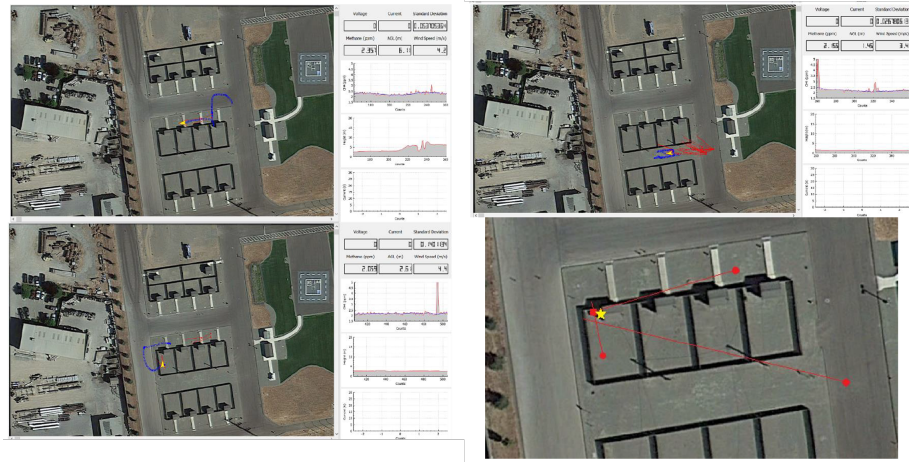


Figure 3.14: (a) top: early back trajectories (red lines); (b) 2nd from top: mid-flight back trajectories; (c) 3rd from top: late back trajectories; (d) averaged back trajectories and predicted leak location (red square) vs. actual leak location (yellow star).

Table 3.7: Metrics of blind test leaks. The time represents the time to localize all the leaks and the leak size is given in SCFH.

Test #	Leak Size	Time(min)	# leaks	# Leaks found
14	Large	16	1	1
15	Med	41	1	1
16	Small	28	2	2
17	Small	23	3	3
18	Small	24	3	2
19	Small	29	2	2
20	Small	31	4	3

somewhere along a line near houses 5, 6, 7. Midway through the test, the wind shifted and the back trajectory indicated a leak location somewhere along the line near houses 5, 9 (Fig. 3.14b). Where these lines intersect was the west side of house 5. Later during the test, back trajectories confirmed the intersection point near the west side of house 5 (Fig. 3.14c). Fig. 3.14d shows the average back trajectory vector for all these cases and compares it with the true leak location (yellow star). The predicted leak location is within 1 meter of the actual leak location.

Seven blind tests were conducted. The sUAS team was not told where or how many leaks were induced. Surveys were usually conducted in the following manner by flying in a grid pattern. Whenever a leak was encountered, about 10-30 seconds were used to better determine where the leak may be coming from before moving on. Metrics of the blind surveys are tabulated in Table. 3.7. In summary, 14 of 16 leaks were found and the leaks that were not found, were masked by other leaks.

3.4.3 Autonomous Leak Detection and Localization

Following the work by Matthes et al (2005) [128], Carslaw (1959), and Roberts (1923) [248] the solution to a single point source advection diffusion equation (ADE) can be solved for a dynamic system approximately by making a quasi-steady state assumption if the variance and transient behavior of the wind small. W_0 is the Lambert function.

$$\begin{aligned} \frac{\partial y}{\partial t} - D \frac{\partial^2 y}{\partial x_i^2} + v \frac{\partial y}{\partial x_i} &= 2q_0 \delta(t - t_0) \delta(x_i - x_{i,0}) \\ \bar{y}(\bar{x}_2, x_0, q_0)_i &= \frac{q_0 \exp(\frac{\bar{v}(\bar{x}_2 - x_{2,0})}{2D})}{\pi^{\frac{2}{3}} D d} \\ d_i(y_i, x_0, q_0) &\approx \frac{2D}{\bar{v}} W_0\left(\frac{\bar{v}q_0}{4\pi D^2 y_i} \exp\left(\frac{v}{2D}(\bar{x}_2 - x_{2,0})\right)\right) \\ \min_{q_0, x_{2,0}} : \sum_{i,j=1}^m &(x_{1,i}(x_{2,0}, q_0) - x_{1,j}(x_{2,0}, q_0))^2 \end{aligned}$$

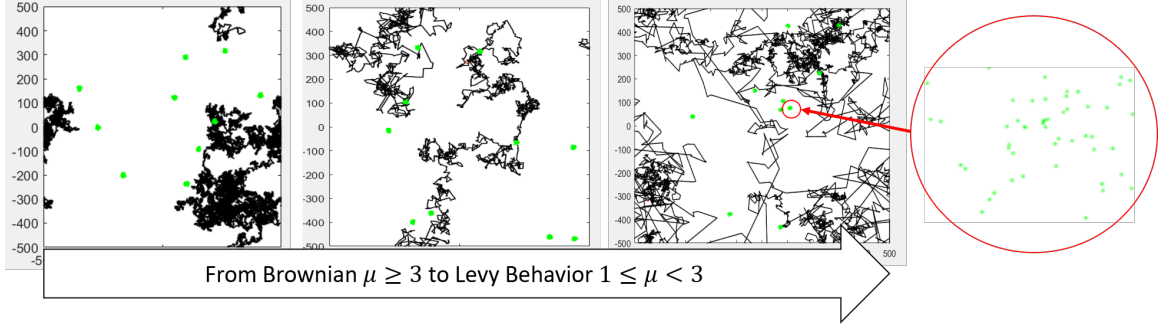


Figure 3.15: This diagram depicts the change in searching strategy from Brownian motion to Lévy following the work in [8]

In the foraging literature the Levy walk has been shown to be effective at searching sparse environments. However, Brownian motion is more efficient in dense areas. This adaptive search model [8] can switch dynamically from Levy to Brownian (see Fig. 3.15) based on finding targets using tumble probability $P(z(t))$, $z(t)$ is governed by the stochastic differential equation (SDE) below,

$$P(z(t)) = e^{-z(t)}, \quad 0 \leq z \leq 5$$

$$\dot{z} = -\frac{\partial U}{\partial z} A + \epsilon, \quad \begin{cases} U = (z - h)^2, \epsilon : \begin{cases} H = \frac{1}{2}, N(0, \sigma) \\ H \neq \frac{1}{2}, \text{fGn} \end{cases} \\ A = \max(A_{min}, \alpha(t)) \end{cases}$$

$$\alpha_k = C_\alpha \alpha_{k-1} + k_t F \begin{cases} F = 1, \text{found target} \\ F = 0, \text{otherwise} \end{cases},$$

where $k_t > 0$ is the gain, and U is the potential function. We extend [8] by adding, fGn to the SDE, defined as $Y_j = B_H(j+1) - B_H(j)$ where fractional Brownian motion $B_H(j)$ is given below,

$$B_H(x) \approx \sum \phi(x-y) B(\Delta y) \quad \phi(x) = \frac{\Gamma(H+1-d/2+||x||)}{\Gamma(||x||+1)\Gamma(H+1-d/2)} \approx \frac{||x||^{H-d/2}}{\Gamma(H+1-d/2)}$$

The adaptive search model has shown to adjust from Brownian motion to Levy walks in a 2D random search. By reducing the problem to a 1D path problem (i.e. survey route) adding decision trees and modeling fugitive gas with a small time scale filament model [120] we have the opportunity to optimize random search for application (see Fig. 3.16). To do this we need to gather enough information to form a set of sample data sets to use in the inversion method for a zeroth order approximation of source localization $(x_{1,0}, y_{2,0})$ and quantification (q_0) .

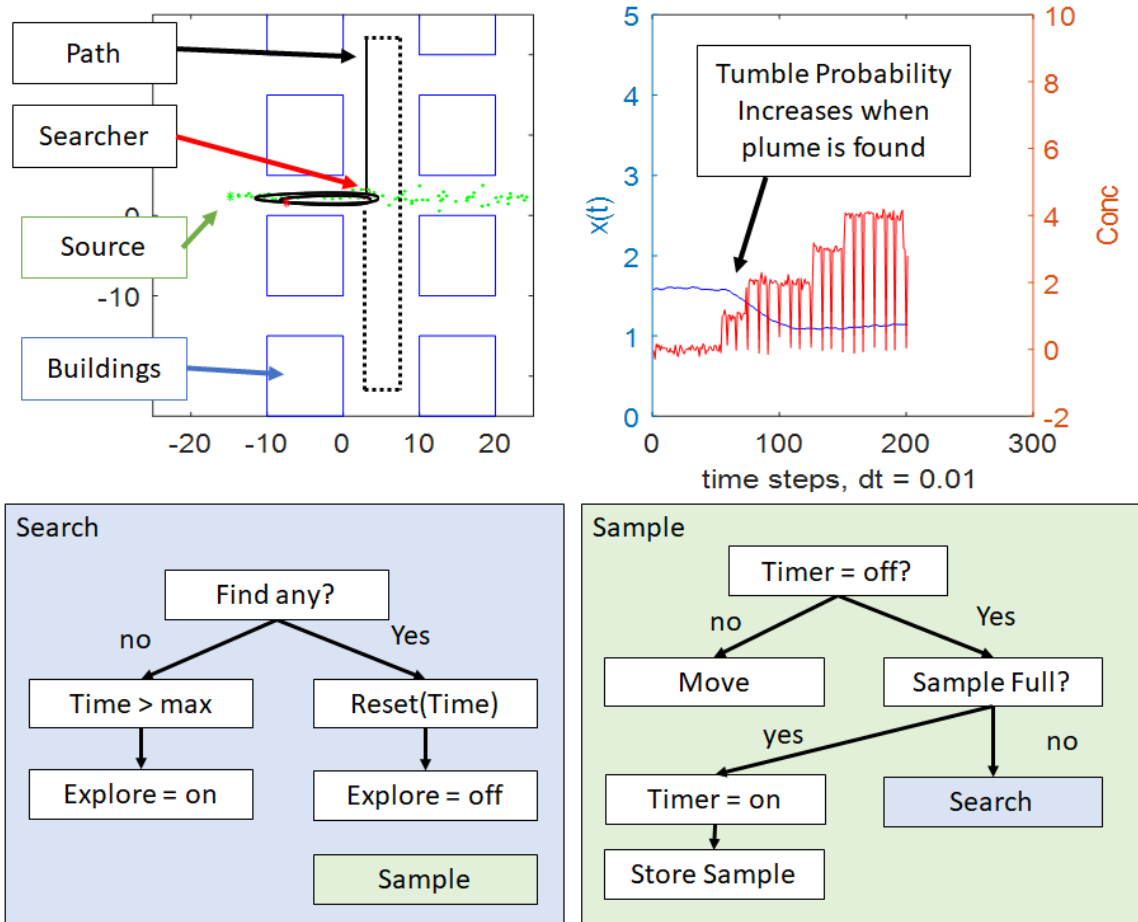


Figure 3.16: (top) Depicts the autonomous searching path where the searcher detects the emission and switches the searching strategy to a Brownian-based random walk, slowing transitioning to a Lévy walk before continuing on the path. (bottom) Schematic of the switching algorithm.

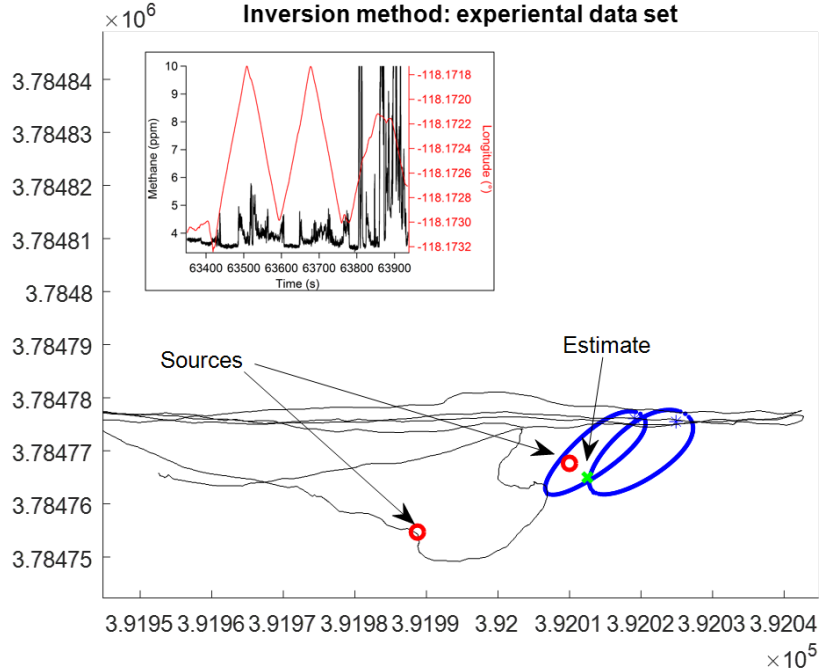


Figure 3.17: An experimental example applying the quadi-steady state inversion method to localize a real-world leak. In this case, two sources are present but based on the measured data (shown in the reduced image) only one of the sources can be located.

Using the quasi-steady inversion method on experimental data we can see the results from just two samples (blue) in the presence of two sources (red) in Fig. 3.17, reveal only one solution. Only taking a small section of raw data from each longitudinal pass we can approximate the source location (green) from our measurement with the OPLS [9]. It should be noted, that in the presence of multiple leaks, this methodology still has limitations and requires further investigation.

3.5 Case Studies: Controlled Release Testing and Emission Quantification

The next step in leak detection and localization is quantification. In this section three case studies are showcased and insight is given in a post-campaign fashion. This includes early work doing controlled release experiments at the MVPGR (in 2017 and 2018) as well as some lessons learned at METEC (in 2021).

3.5.1 Feb 2017 MVPGR

During the flights in the MVPGR (outlined in 3.4), the mass balance vertical flux plane method (see 3.2.3) was carried out to quantify a controlled point source emission

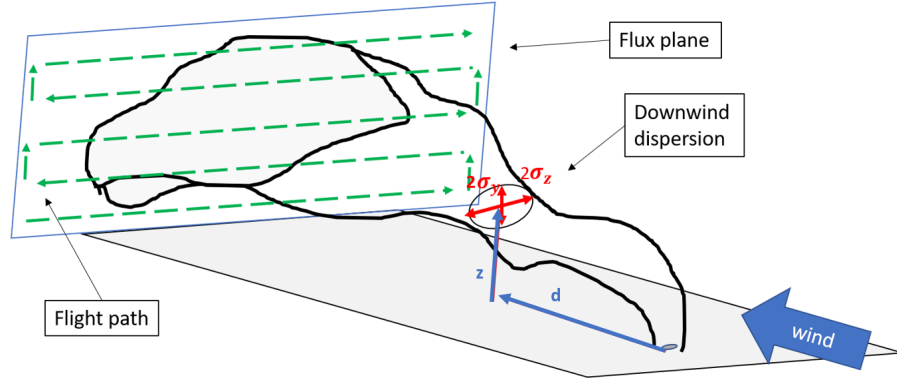


Figure 3.18: An example curtain flight path for mass balance calculations.

released from ground level. The emission rate was varied from 1-25 SCFH. The total number of flights quantified consisted of 14, of which, some of the flights contained an ascending and descending flight path (counting as two individual estimations). The results can be seen in Fig. 3.19, which show that for most of the estimates, as turbulence intensity gets smaller, the estimates become better. This suggests that as the atmosphere becomes more stable the estimates become more accurate. From a theoretical perspective, this makes sense as the time averaged plume going through the measurement plane, would be governed by the mass balance equation. However, the atmospheric conditions tend to continuously change and provide disturbances to the estimates.

It is worth noting that the sequence of releases (shown in Fig 3.6) goes from 5 SCFH all the way to 25 SCFH and eventually down to 1 SCFH. Therefore, it may be plausible that residual emissions from previous higher emission rates, resulted in over-estimations of the emission rate in the 1 SCFH case. In fact, from Fig. 3.19 it can be seen that if the 1 SCFH case was processed at a 5 SCFH emission rate, three of the estimates would be below 200% error - consistent with the rest of the estimates.

3.5.2 Aug 2018 MVPGR

In the same testing location as the February 2017 flights (in the MVPGR), the source setup included a pure methane bottle connected to a pressure regulator, followed by several meters of PFTE (Polytetrafluoroethylene) tubing, where a rotameter flowmeter gauge was placed for controlling the release rate. Additional PFTE tubing was used to connect the output of the flowmeter to a 5 Gallon bucket filled with coarse rocks (≈ 1 cm in diameter). The rocks served to slow the stack velocity down and let the wind carry the methane, simulating a surface point source. The source rate was initially set to 10 SCFH, whereas after laboratory testing the source rate was calibrated to be $Q_* = 12.9$ SCFH. The landscape of the site was relatively flat with little topology and no surface obstructions (see Fig. 3.20). A sUAS, namely the DJI M210 (see Fig. 3.21), was equipped with the OPLS sensor. The atmospheric stability

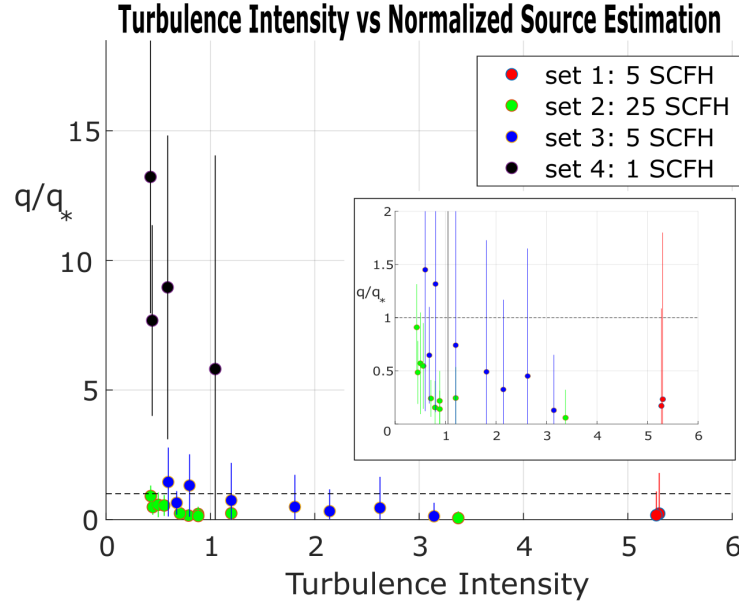


Figure 3.19: Emission quantification from controlled release tests in the MVPGR February 2017.

during the experiment was determined to be C and B (slightly to moderately unstable) using the PG curves [177]. The wind speed and direction was measured with a RMYoung 3D ultrasonic anemometer. The experiment consisted of 16 downwind flight paths of horizontal transects ascending from ground level till a clean air fetch was reached, referred to as a curtain (see Fig. 3.18).

The results can be seen in Fig. 3.22. It can be observed that as the turbulence intensity (TI) is decreasing, the accuracy of the mass balance method improves, which suggests that the atmospheric stability (more stable $\rightarrow$ smaller TI) is important for accuracy in the emission estimate as well as for improving the estimate variability. This result echos the result found in the previous Section.

3.5.3 Apr 2021 METEC

In April 2021, a series of controlled release experiments were undertaken with Aero-metrix Services Inc. (Victoria BC, Canada) at the Methane Emission Technology Evaluation Center (METEC) over a five day period. The controlled release tests were conducted concurrently with other teams on site simultaneously at each of the sub-divided regions within METEC (see Fig. 4.6). The testing configurations were changed every hour (called rounds) and the teams rotated to another sub-division or were given an hour rest. The 20 measurement rounds had mean measured methane concentrations ranging from 2.4 to 5.6 ppm - all significantly above the natural background. The emission detections at the drone, resulted in maximum methane concentrations from 3.1 to 152.8 ppm.

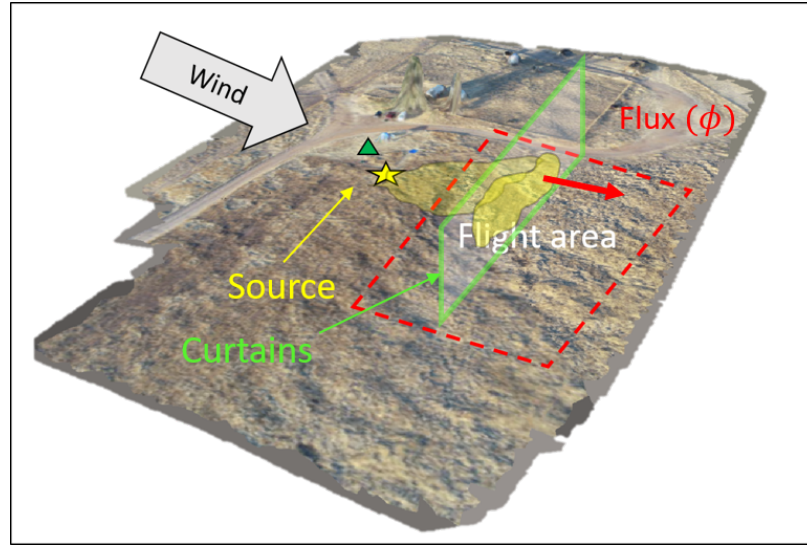


Figure 3.20: Experiment location at Merced Vernal Pools and Grassland Reserve. The yellow star indicates the source, the green triangle is the RMYoung anemometer, and the flight area is shown red dotted line.



Figure 3.21: DJI Matrice 210 (M210) equipped with a sensitive methane sniffer developed by [9] and attached to the DJI Skyport.

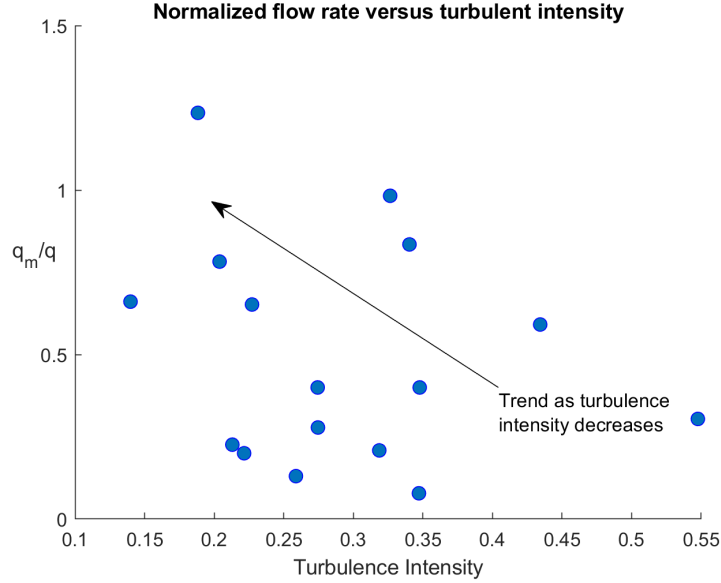


Figure 3.22: Normalized controlled release estimation using the mass balance method of the August 2018 MVPGR field test.

During this field campaign we came across two notable challenges for measuring emissions in the open atmosphere with the OPLS-based sUAS approach. The first challenge has to do with off-site emissions, meaning, the emissions from adjacent subdivisions of the METEC site that would meander into the area we are measuring. This is a challenge when the task is to detect, localize, and quantify within the subdivision itself. Therefore, the off-site emission needs to be tore out (or accounted for) of the emission estimate of the sub-divided site. This problem adds to the difficulty of measuring accurate emissions, due to the presence of higher emission rates, which leads to higher variability in the magnitude of the measurement signal. For example, given a large emission rate from an adjacent site and a relatively small emission from the target site, the variability between the difference in the upwind and downwind measurements (i.e. flux in and flux out of the control volume) could mask the presence of the smaller leak (see Fig. 3.23). The other challenge we found with the sUAS was with underground and surface based leaks. For example, in Fig. 3.24, the underground leak is first detected from the initial survey but when trying to quantify the leak, the detection became difficult when flying low to the ground as well as slow ($< 1\text{m/s}$). One of the reasons for this is due to the propeller exhaust of the sUAS interacting with the ground, causing the emission to be advected in an outward fashion from the aircraft (relative to the ground). Flying faster delays this affect, resulting in a better detection rate. This ultimately makes quantification difficult as the farther back we fly from the source, the more sensitive the measurement equipment requires to detect the emission. One advantage to measuring farther from the source is with the well mixed condition. This provides a smooth concentration signal that is easier

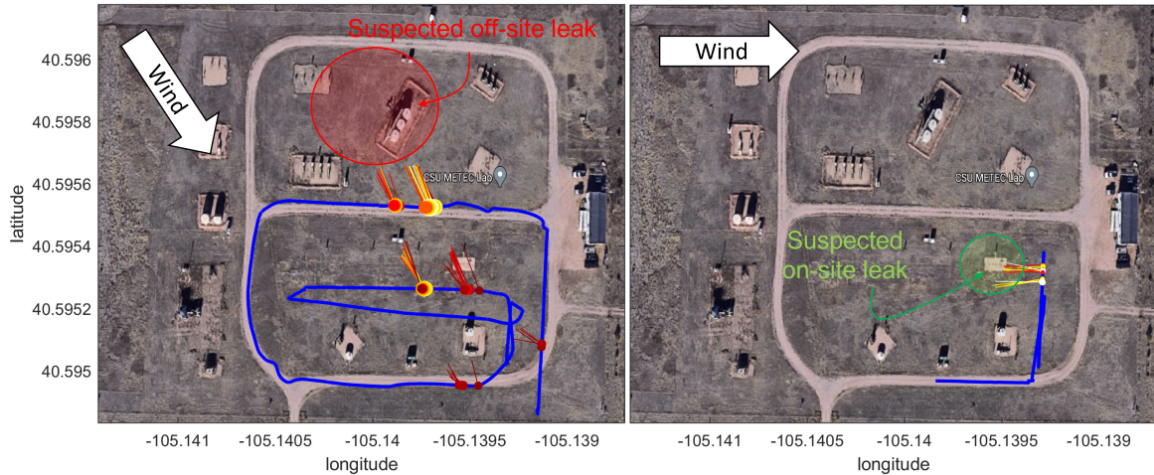


Figure 3.23: Example of an off-site emission source affecting the survey and emission quantification.

to estimate the source with (given the sensor is sensitive enough). Unfortunately, one of the constraints for the experiment was to only measure the source from inside the site fence line and sub-divided region. Which, this was done to mitigate the potential flight safety risk from sUAS and other personnel operating in an adjacent region.

3.6 Case Studies: Real World Emission Quantification

Once a quantification method has been developed the overarching goals would be to subject them to real world applications. In this section we give examples of campaigns undertaken in real world environment scenarios.

3.6.1 Aug 2018: Merced County Regional Waste Management Landfill

The California Air Resources Board (CARB) introduced an implementation document to help municipal solid waste landfills be in compliance with California Code of Regulations title 17, subchapter 10, article 4, subarticle 6 sections 95460 to 95476, Methane Emissions from Municipal Solid Waste Landfills. In short, this requires periodic regulatory surface emissions testing that may vary depending on the historical compliance of the site.

Surface emissions tests can take a considerable amount of time to complete (days to weeks) depending on the spatial scale of a site and how in-compliance the site has been previously. This is due to the fact that surveyors must traverse the site on foot. This requires the site to be broken into grids, which, can pose a resolution

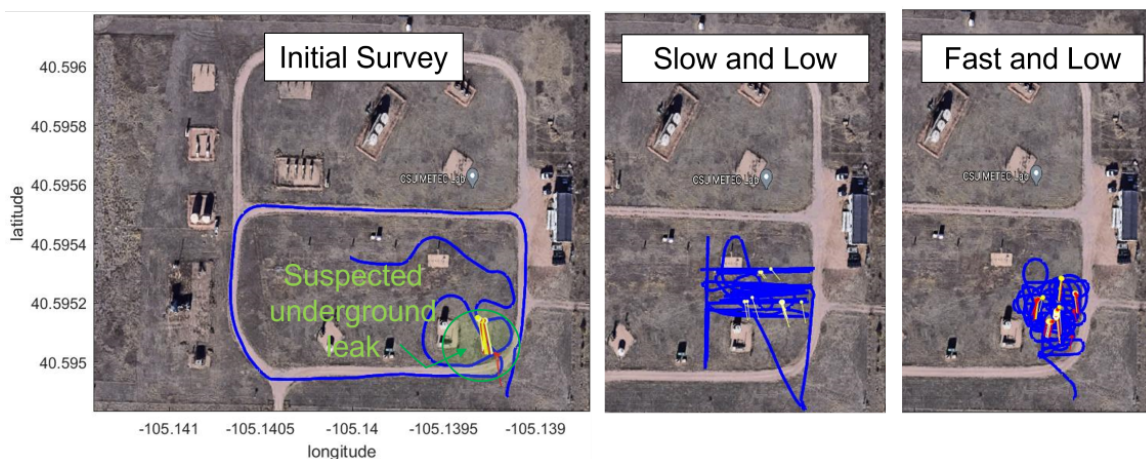


Figure 3.24: Example of a survey with an underground source. The initial survey detected a leak near the south-east corner of the site. The successive flights had difficulty detecting the emission source when flying low and slow ($\approx 1\text{m/s}$ or less), whereas, when flying faster the rate of detection increases.

problem on how well the site is surveyed. This can potentially lead to emissions going undetected as well. Using alternative methods that rely the plume dynamics of the surface emissions can allow for measurements to be taken remotely, such as measurements on-board sUAS equipped with an OPLS and using the mass balance approach.

Since sUAS cost is relatively small and can be deployed rapidly and more frequent than the traditional SEM surveys, leaves a potential to for sUAS-based surveying to provide a faster alternative to leak detection as well as quantification.

In August of 2018, the Merced County Regional Waste Management (MCRWM) site was flown (see Fig. 3.25) with a DJI M210 equipped with the OPLS [9]. Due to the limited amount of surveys conducted and the variability in the wind, yielding poor confidence in the accuracy of the estimates, the emission estimates are likely off. The emission estimates for both flights shown here above 200 SCFH (see Fig. 3.26), which is equivalent to the emissions of about 400 cows (given an average cow can typically produce 0.3-0.5 SCFH). One major challenge behind the use of sUAS, in this case, is that when measuring large areas using the mass balance approach requires the atmosphere to be quite stable during the flight (which is on the order of the endurance of the measurement systems ≈ 20 minutes). Additionally, the size of the curtain makes the flight time requirement substantially larger than point source type measurements.

3.6.2 Jun 2019: Alberta Methane Field Challenge

In June of 2019, six teams with different technologies (drone-based, truck-based, hand-held, and fixed) were subject to surveying around 50 active oil and gas sites. The sur-



Figure 3.25: At the MCRWM site a new cell is being developed, such that, the bank shown in the image has no membrane and only a dirt cover.

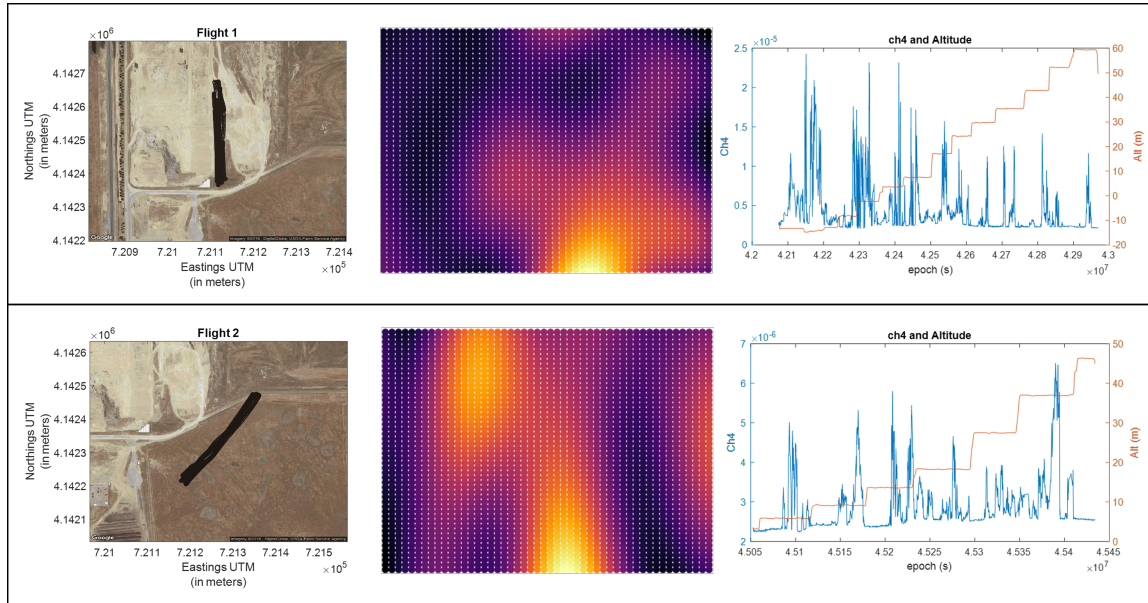


Figure 3.26: Two surveys of an active landfill cell without a membrane are conducted where measurements are taken from: (top) near the source; (bottom) further back. The measurements near the source show the emission emanating from the ground whereas when the measurement is taken from further back the emission is subject to more turbulence.

Table 3.8: Summary of the performance of the participating technologies in the first phase of the Alberta Methane Field Challenge (AMFC)[13]. Days refers to the number of participation days, rate refers to the number of sites per day, and time is the survey in minutes per site.

Team	Method	Days	Rate	Time	Sites	Detection (site-level, %)
Univ. of Calgary	Truck	12	8	10	90	83
Altus Geomatics	Truck	10	15	9	138	87
SeekOps	Drone	11	5	36	54	92
Aerometrix	Drone	10	5	20	42	80
Heath (hybrid)	Truck + handheld	12	5	41	53	91

vey would typically consist of three tasks: detection, localization, and quantification (granted the technology supported that). The degree to which each technology was able to achieve each task varied. The results for each of the technologies were showcased first in the American Geophysical Union in 2019 [246], and then in a preprint in [13]. During the AMFC, the author was collaborating with Aerometrix Services Inc. and showcased the results in a Geoscience British Columbia report in [200], of which, has not been publically available as of the completion date of this manuscript.

It is important to note that not all participants were able to produce quantification of the emission sources and the survey times and effectiveness varied. A summary of the first phase of the challenge can be seen in Table. 3.8.

The site level emissions by all the participating teams were compared and averaged. It was mentioned that not all technologies were able to detect and/or quantify all the emissions on the sites, and therefore is likely that the estimates represent a lower-bound for the total emission estimates of the site. Another important aspect with the underestimation between QOGI and other technologies, was the wind. The wind condition, both instantaneous and or time-averaged, provided direct impact to the emission estimate accuracy. This is further compounded by the presence of intermittent sources, making the estimation a difficult process using sUAS-based approaches.

The AMFC project had a second phase later in the year, when the winter conditions set in. Not all the participants from the first phase were willing and able to be at phase two. It is interesting to note, that the total site-level emissions, on average, were approximately 50% of those in phase one. This finding suggests that the summer time (increased atmospheric temperature) leads to more emission leaks.

3.6.3 Nov 2019: Natural Emissions from Alaskan Permafrost Bog

Methane plays an important role in determining the atmosphere’s climate and chemistry. Fluxes of methane from an ecosystem are often measured using eddy covariance



Figure 3.27: The mobile sUAS measurement platform above an eddy covariance tower in an Alaskan peatland bog.

Table 3.9: The quantification results are compared to tower estimates, which show that the sUAS-based estimates are lower than the tower measurements during the time of the flights. However, the estimates are within a reason considering the variability of the wind and the difficulty in measuring small emission rates.

Date	Avg Tower Flux	Avg UAV Flux	UAV Flux-1	UAV Flux-2	UAV Flux-3
9/25	0.0338 ± 0.0054	0.0220 ± 0.0278	0.0610 ± 0.0162	-0.0180 ± 0.0116	-
9/27	0.0411 ± 0.0063	0.0097 ± 0.0254	0.0078 ± 0.0094	0.0170 ± 0.0035	0.0045 ± 0.0126

flux towers (see Fig. 3.27), however, there are disadvantages with this method. Flux towers are expensive to purchase and have high demands with respect to maintenance and cost of operation, especially in remote locations, making replication across the landscape a challenge. In this case study we applied the mass balance method to quantify the emissions from a permafrost bog using upwind and downwind quantification measurements [7].

During a field campaign in September of 2019, the Alaskan Peatland Experiment (APEX) site was conducted to explore the ability to measure natural ecosystem flux from the Alaskan peatland bogs. This work looked at using bootstrapping to understand the uncertainty better and experimented with different spatial interpolation methods (e.g. IDW, Kriging, see [1]) on the mass balance method. Additionally, this work explored the comparison between sUAS-derived wind (using an on-board anemometer) and the eddy covariance tower. An example of the flux plane measurements can be seen in Fig. 3.28 and the emission estimates can be seen in Fig. 3.29 and Table. 3.9.

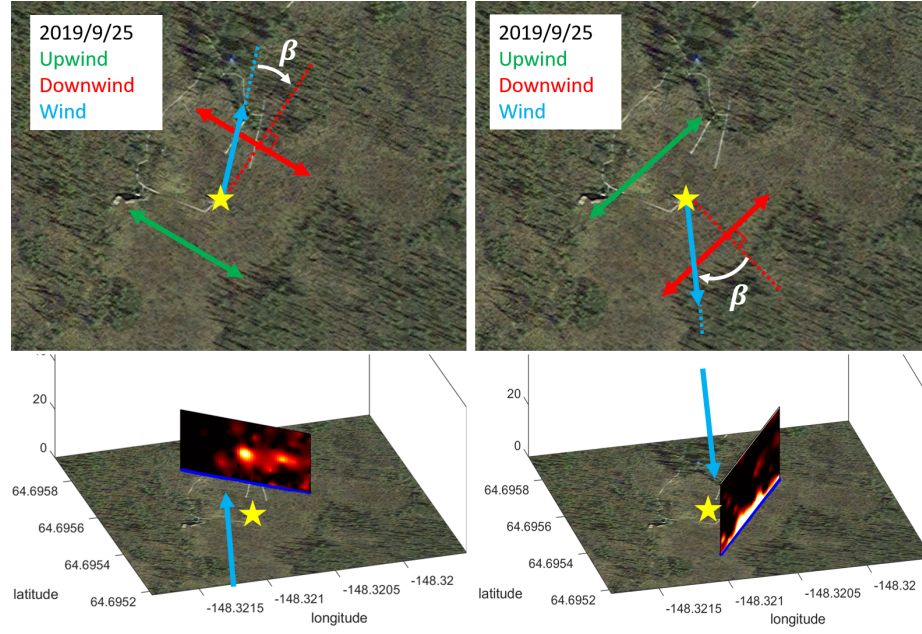


Figure 3.28: A top down view of the Alaskan Peatland Experiment (APEX) site, showing the upwind and downwind measurements for two different wind conditions.

3.7 Chapter Summary

In this chapter, we discussed leak detection and quantification (LDAQ) with sUAS by: examining the different types of sensors and platforms, investigating sUAS and non-sUAS based LDAQ methods (simulation-, optimization-, mass-balance-, correlation-, and imaged-based), providing an analysis of the investigated methods along with a method summary, and providing an overview of several case studies regarding leak detection and localization experiments / simulations, controlled release quantification experiments, and real world emission quantification.

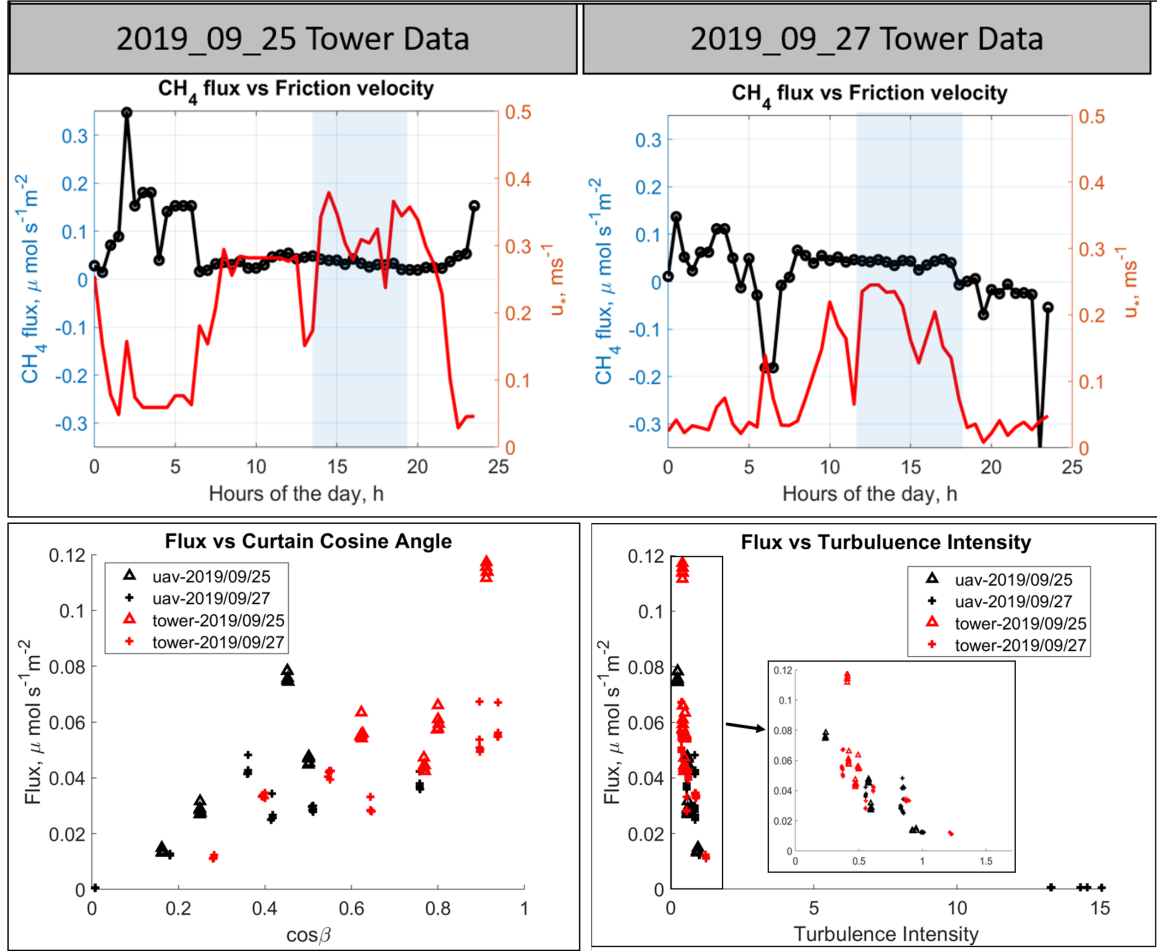


Figure 3.29: (top) Tower-based flux measurement of the APEX site for the days and times of the 2019 September field campaign (bottom-left) Average flux given the cosine angle between the apparent wind direction and the normal vector of the flux plane (bottom-right) The emission quantification values are shown to be higher when the turbulence intensity is small.

Chapter 4

Towards Smarter Environmental Sensing with Digital Twins

Environmental sensing and monitoring includes a variety of applications such as measuring local air quality, soil remediation, oil spills, ground water quality-tracing, etc. These applications can include both fixed and mobile sensors. For example, sensing ecosystem flux can be done with fixed sensors in the soil, above ground sensors, with mobile sensors (e.g. ground vehicles or sUAS [249]) or remote sensing techniques (such as with satellite imagery or manned aircraft [26]). Typically fixed sensor approaches work well to provide precise ecosystem flux measurements (e.g. chambers, autochambers, and eddy covariance towers [250]). In terms of site level emissions, fixed open path lasers have been used to evaluate and compare with accepted Environmental Protection Agency (EPA) methods (e.g. dairy farms) [172]. Within the past few years, several works in the area of environmental sensing with sUAS have been done (see the comprehensive review by [142]). Typically the sensors consist of low cost chemical sensors (e.g. amperometric, metal-oxide, non-dispersive infrared, or photo-ionization detectors) or higher cost tunable diode laser absorption spectroscopy (traditional open/closed cell - TDLAS or backscatter based - sTDLAS). Other techniques like optical gas imaging, which uses thermal camera based sensing have been used with sUAS. For instance, the sTDLAS (e.g. Pergam Methane mini-G (SA3C50A)) has been used for determining spatial methane in permafrost [251]. In the oil and gas industry, the TDLAS (e.g. open path laser spectrometer - OPLS [9], or path integrated) has been used for detection [201; 202] and for quantification [252], such as in quantifying the Alberta Methane Field Challenge (AMFC) site emissions [200]. The use of fixed-wing sUAS have been used as well in both detection [253] and quantification [254].

Due to the large variety of approaches and applications of environmental sensing, the focus of this chapter will be on ground-to-air emissions with an application to gas plumes (e.g. fugitive methane). The motivation of studying fugitive methane (CH_4) gas, stems from the fact that methane has a large green house gas (GHG) warming potential relative to CO_2 and occurs from both natural and anthropogenic sources.

Modeling gas plumes in general can be done using a variety of tools (e.g. deterministic and stochastic approaches) which provide information about how the plume may develop given a set of initial and boundary conditions. However, these visualization require a certain degree of model fidelity to capture realistic behavior (i.e. plume meandering and trapping). Computational fluid dynamics approaches can be applied but suffer from high computational cost. Stochastic methods have shown better computational efficiency but for a given initial condition the output is an ensemble (due to the inherent randomness) and presents difficulty in implementing continuous time. On the other hand, hybrid modeling (which encompasses a combination of stochastic and deterministic methods) can provide computational efficiency while maintaining highest possible fidelity. In practice, computational efficiency of these models would need to be suitable for deployment on a laptop. Thus, there exists a trade-off between model fidelity and run-time. This is important for applying concepts such as Digital Twins (DT) to further improve on environmental sensing (ES) work.

What is a DT? Some of the earliest work of using physical twins (or hardware twins) can be seen as early as National Aeronautics and Space Administration's (NASA) Apollo program [36]. In this case there existed at least two space vehicles or rovers at once. This allowed for the continuous testing of code, maneuvers, and conditions before sending them to the deployed version. Software can also be introduced to the hardware twin, as in the 'iron bird' which combined a physical interface of a plane cockpit with an aerodynamics simulation for training purposes. The simulation was used for a visual aide as well as to compute the meteorological conditions and resulting forces on the plane [36]. Some of the early reports of the current concept of digital twin (as digital equivalence of a physical product) showed up in a product lifecycle management course at the University of Michigan in 2003 taught by Michael Grieves. Since then a white paper was written in reference to virtual factory replication [37]. It is explained that the DT in this scenario contains three parts: (1) physical products in real space (2) virtual products in virtual space and (3) and connections that tie together the real and virtual products. The use cases for DT are: (1) conceptualization - visualization of the physical and virtual products, (2) comparison - compare physical product behavior with the virtual product to improve results and (3) collaboration - solutions found can be applied immediately to other factories (or systems). For environmental sensing, the environment can be full of complexity and having a DT can provide a deeper understanding between the sensors and physical parameters, such as source rate, location, and atmospheric conditions. Additionally, DT's have the ability to update the system parameters, over time, as well as give insight on performance metrics.

This chapter proposes a DT framework for ES from a point source emission as an extended version of the conference paper in [1]. A controlled release case study is used to solve an optimization-based behavior matching problem of the DT framework, using the estimated concentration signal, $\hat{c}_m = f(\mathbf{x}, \mathbf{v}|\theta)$ (where $\mathbf{x}$ is the position of the sUAS, $\mathbf{v}$ represents the wind field and θ is the DT parameters), and the measured concentration signal, c_m , with a single sUAS measurement system. The mass balance

method is applied to experimentally quantify the source emission rate. This includes: defining DT levels and capabilities; adding additional details and figures of the DT model, approach, sensitivity of parameters, and how the observation problem is affected by sensor location (i.e. measurement operator); exploring additional spatial and spatial-temporal interpolation methods to quantify the source and compare data between the DT simulation and physical experiment; proposing the use of Mittag-Leffler function into the spatial-temporal gas distribution modeling framework.

4.1 The Digital Twin Framework and POSIM Digital Twin Platform

In order to discuss a framework there needs to be some agreed upon definition of DT. Over the years there have been several variants of the DT definition such as those shown in Table 4.1. The Mechatronics Embedded Systems and Automation Lab has since added a more general definition which is as follows, “A Digital Twin is the combination of multiple, individual, and detailed simulation models (continuous, discrete, hybrid), where its interconnection represents the dynamics of a complex system, which is updated periodically (windowed or real time) with the system information in order to reflect the system current status as well as predict its future behavior and possible faults”. For ES, the model needs to be selected such that there is a trade off between fidelity and computational speed. Considering the quote from statistician George Box, “all models are wrong, but some are useful”, we need the desirables of the DT model to be capable of: (D1) exhibiting relevant digital artifacts with suitable level of fidelity, (D2) short or long time-scale characteristics evolves along with the real system given the same conditions, and (D3) can be used to solve relevant problems for the system or improving methodology [1].

The DT concept can also be broken into four different stages or levels based on their current implementation: (L1) Pre-Digital Twin, (L2) Digital Twin, (L3) Adaptive Digital Twin, and (L4) Intelligent Digital Twin. Each stage has increasing levels of model sophistication. Aside from L1, all of the levels have a physical twin. As the DT levels go from L2 to L4 the data acquisition from the physical twin increases. Different system health and performance parameters can be extracted for predictive maintenance (from batch updates to real-time updates). Additionally, these stages can start to incorporate machine learning (ML) in operator preferences as well as for the system/environment [38].

The proposed DT framework for a point source first starts on L1 with a proposed model for representing the complex system. In the case of gas concentration, the system is governed by a partial differential equation such as the advection diffusion equation,

$$\partial_t C = \nabla \cdot (D \nabla C) - \nabla \cdot (\mathbf{v} C) + \text{source}, \quad (4.1)$$

where the C represents the concentration of the gas, D is the diffusion coefficient, and $\mathbf{v} = [u, v]^T$ is the wind vector field. The source can be represented using Kronecker

Table 4.1: (top) Definitions of Digital Twin [14] and (bottom) recent Digital Twin platforms and applications.

Definition	Authors	Year
DT is a virtual, digital equivalent to a physical product	Grieves [37]	2003
Up-to-date representation of an actual physical asset in operation	Mathworks [255]	2019
Dynamic virtual representation of a physical object or system, usually across multiple stages of its lifecycle. It uses real-world data, simulation or machine learning models, combined with data analysis, to enable understanding, learning, and reasoning.	IBM [256]	2020
Software representations of assets and processes that are used to understand, predict, and optimize performance in order to achieve improved business outcomes.	GE Digital [257]	2019
DT is a perfect digital copy of the physical world: a digital twin. This twin would enable you to collaborate virtually, intake sensor data and simulate conditions quickly, understand what-if scenarios clearly, predict results more accurately, and output instructions to manipulate the physical world..	Deloitte [258]	2020
An integrated Multiphysics, multiscale, probabilistic simulation of an as-built vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its corresponding flying twin	NASA[259]	2012
A digital twin is a multi-faceted dynamic set of smart digital models of a system or a subsystem along with all its constituents, which accurately represent the design of a product, production process or the performance of a product or production system in operation.	Dufour et al. [260]	2018
DT are precise, virtual copies of machines or systems driven by data collected from sensors in real time, these sophisticated computer models mirror almost every facet of a product, process or service.	Tao et al. [261]	2019
Platforms / Application	Authors	Year
Environmental sensing using sUAS with POSIM.	Hollenbeck et el. [1]	2020
Twin Control approach for Machine Tool Design, Process design, Process control and Maintenance.	Armedia et el. [262]	2020
Digital Twin in remote labs with Amazon Web Services (AWS).	Wuttke et el. [263]	2019
Model-based Digital Twin of UAV in product lifecycle management.	Nikolaev et el. [264]	2018
Generic Digital Twin Architecture for packed-bed thermal storage (aligned with Reference Architecture Model Industry 4.0 (RAMI4.0)).	Steindl et el. [265]	2020
Digital Twins in smart cities, engineering, and construction.	Sepasgozar [266]	2021
Warehouse Inventory and Traceability with Blockchain-based system.	Fernandez et al. [267]	2019
Automatic guided vehicles in smart manufacturing.	Martínez-Gutiérrez et al. [268]	2021
Digital twin of robot Baxter for practicing spatial manipulation tasks.	Verner et al. [41]	2019
Cooperation of UAV swarms for intelligent network reconstruction.	Lei et al. [269]	2020
Helicopter dynamics using multibody systems.	Guivarch et al. [270]	2019
Multisource model for robotic assembly systems.	Li et al. [271]	2020
Digital twin enabled smart control engineering framework.	Viola et al. [272]	2020
Information-based model discrimination using digital twin behavior matching.	Viola et al. [273]	2020
Smart control engineering of multiple input multiple output (MIMO) system using PHELP.	Viola et al. [274]	2020

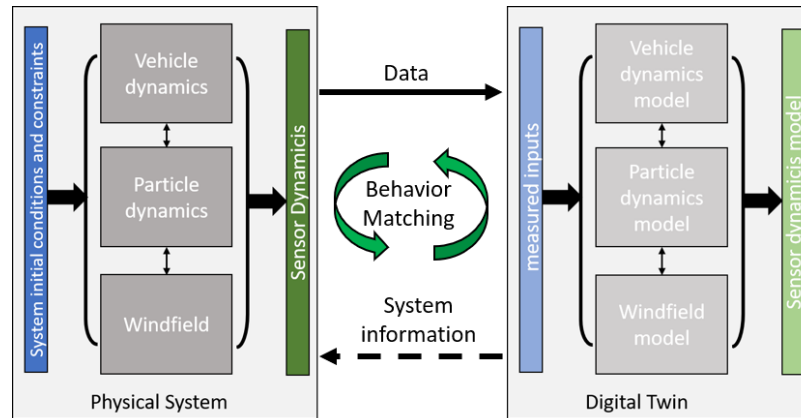


Figure 4.1: Digital twin framework diagram.

delta function $\delta_s = \delta(x - x_s)\delta(y - y_s)\delta(z - z_s)$, where $\mathbf{x}_s = [x_s, y_s, z_s]^T$ is the location of the point source. The wind field parameters can be solved generally by using the something like the incompressible Navier-Stokes equation,

$$\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} - \nu \nabla^2 \mathbf{v} = -\nabla w + \mathbf{g}. \quad (4.2)$$

Here ν is the kinematic viscosity, w is the thermodynamic work, and $\mathbf{g}$ is the gravitational acceleration. Solving these systems consecutively can be expensive and thus, some assumptions need to be applied to simplify the models. One such approach to solve this problem was proposed by Farrell et al [120] and is summarized here for completeness. The first assumption is to break the concentration dynamics into discrete filaments (or packets of molecules) based on different length scales (i.e. large scale advection (a), intermediate scale turbulent mixing and stirring (m), and local small scale diffusion (d)). The transport of the k -th concentration molecule can then be described by, $\dot{\mathbf{x}}_k = \mathbf{v}(\mathbf{x}) = \mathbf{v}_a + \mathbf{v}_m + \mathbf{v}_d$. However, the small scale diffusion term can be absorbed into the shape of the filament, such that the filament's i -th location can be described as,

$$\dot{\mathbf{x}}_i = \mathbf{v}(\mathbf{x}) = \mathbf{v}_a + \mathbf{v}_{m_i}. \quad (4.3)$$

Now the only thing left to resolve is the wind vector field. One approach is to look at average wind vector such that $\mathbf{v} = \bar{\mathbf{v}} + \mathbf{v}'$, where the overbar ($\bar{\cdot}$) represents the average and the prime ($'$) denotes the deviation from the average. Substituting the decomposition of the wind vector into (4.2) and using the following assumptions the Navier-Stokes equation can be drastically simplified:

1. Coriolis forces, geostrophic winds, and molecular viscosity can be deemed small and neglected;
2. Measurements are conducted close to the ground where winds are relatively constant w.r.t. the altitude.

The wind field can then be represented by the following equations that resembles the viscous Burgers equation,

$$\begin{aligned} \frac{\partial \bar{u}}{\partial t} &= -\bar{u} \frac{\partial \bar{u}}{\partial x} - \bar{v} \frac{\partial \bar{u}}{\partial y} + K_x \frac{\partial^2 \bar{u}}{\partial x^2} + K_y \frac{\partial^2 \bar{u}}{\partial y^2} \\ \frac{\partial \bar{v}}{\partial t} &= -\bar{u} \frac{\partial \bar{v}}{\partial x} - \bar{v} \frac{\partial \bar{v}}{\partial y} + K_x \frac{\partial^2 \bar{v}}{\partial x^2} + K_y \frac{\partial^2 \bar{v}}{\partial y^2}. \end{aligned} \quad (4.4)$$

This is realized by using the simplest K -closure method, which relates the diffusivity ($K_{x,y}$) to, $\overline{u'u'} = -K_x \frac{\partial \bar{u}}{\partial x}$ and $\overline{u'v'} = \overline{v'u'} = -\frac{1}{2}(K_x \frac{\partial \bar{v}}{\partial x} + K_y \frac{\partial \bar{u}}{\partial y})$ and similarly for the $\overline{v'v'}$, and $\overline{v'u'}$. For the simulation aspect, the wind field meandering can be produced by perturbing the flow with colored noise using a second order transfer function approach, $H(s) = Ga/(s^2 + bs + a)$ driven by white noise [120] where G is the gain, and a and b are parameters that affect the natural frequency and damping ratio. This colored noise is updated on the corners of the finite grid and adjacent

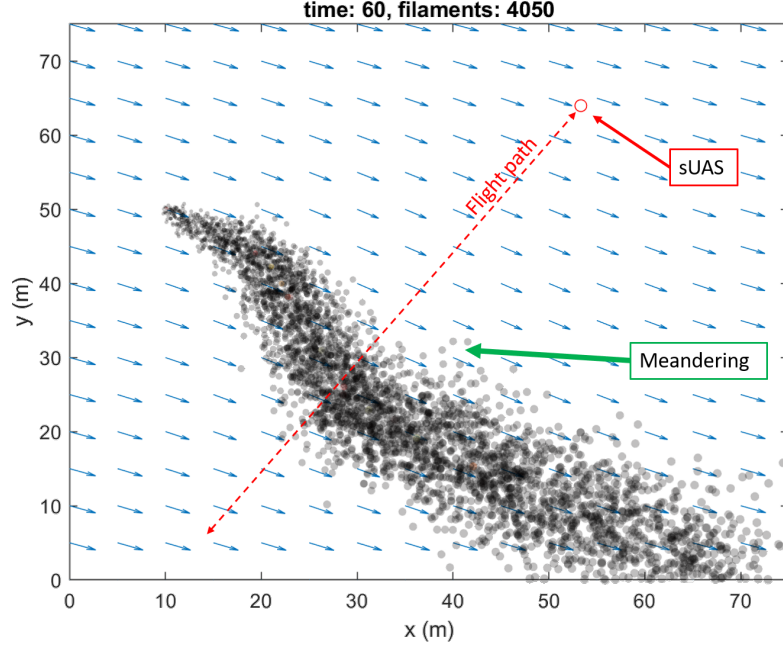


Figure 4.2: MATLAB example of the digital twin with meandering

boundary conditions are linearly interpolated (see Fig. 4.3). The meandering behavior (see Fig. 4.2) can be adjusted by scaling a , b , G and $K_{x,y}$. The interior nodes can be solved implicitly [275] or explicitly by finite differences [276].

The large scale advective term, $\mathbf{v}_a$, can be calculated by using a bilinear interpolation between the nearby wind vectors points or estimated using the nearest neighboring point relative to the filament position, $\mathbf{x}_i$. The grid size used to calculate the the wind field can remain coarse (large Δx and Δy) as it reflects the larger length scales of the model. Additionally the large spatial separation between points helps to increase simulation runtime. To maintain efficiency we used the nearest neighboring point for this work. To calculate the intermediate scale term, $\mathbf{v}_m$, we use a random process satisfying, $\dot{\omega} = A'\omega + B'\eta$ and $\mathbf{v}_m = C'\omega + D'\eta$, where $\omega \in R^n$ and η is a white noise random process with spectral density σ_η^2 . The matrices A' , B' , C' , and D' are appropriately sized. The resulting standard deviation grows as $G\sigma_\eta\sqrt{t}$. In the simplest 2D case $D' = \mathbf{0}$, $(B')^T = C' = [1, 1]$, and $A' = I$, where I is the identity matrix.

To calculate the small scale diffusion, $\mathbf{v}_d$ the use of a growth term, $R(t)$, is needed. This growth term represents the standard deviation of the packet of molecules inside the discrete filament,

$$\frac{dR_i}{dt} = \frac{\gamma}{2R_i} \quad \rightarrow \quad R_i(t) = (R_{i0}^2 + \gamma t)^{1/2},$$

where $r_i(t) = \|\mathbf{x} - \mathbf{x}_i\|$ is the Euclidean distance from the sensor and the i -th filament, $R_{i0} = R_i(0)$ and γ is a positive scalar. This allows for the concentration of individual

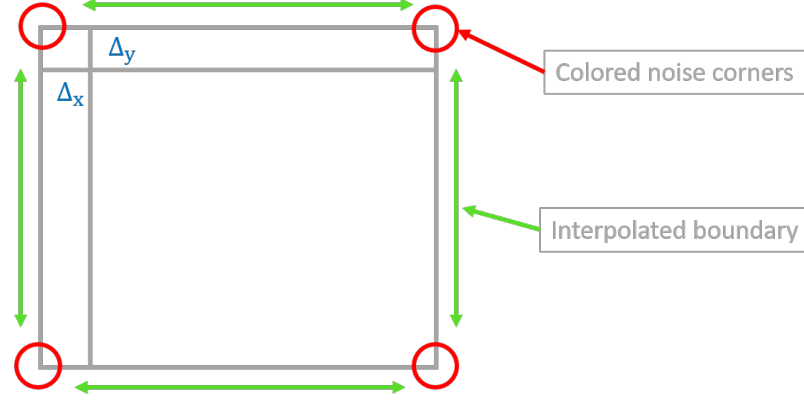


Figure 4.3: Scheme for updating the colored noise in the corners interpolating the adjacent boundaries.

molecules in a filament to be calculated by,

$$C_i(\mathbf{x}, t) = \frac{Q}{\sqrt{8\pi^3} R_i^3(t)} \exp\left(-\frac{r_i^2(t)}{R_i^2(t)}\right), \quad (4.5)$$

where Q is the filament source rate in molecules per filament. The total concentration measured, $C(\mathbf{x}, t)$, can be calculated given an effective sensor area, A_e , where N represents the number of filaments inside $A_e = \pi r_e^2$.

$$C(\mathbf{x}, t) = \sum_{i=1}^N C_i(\mathbf{x}, t) \quad (4.6)$$

The effective area can be approximated as a circle with radius r_e . The sensor can be modeled as a low pass filtered system with thresholding, $\dot{c} = f_{BW}(C(\mathbf{x}, t) - c)$. The sensor bandwidth is f_{BW} and threshold is λ . The measured concentration, c_m would then be,

$$c_m(t) = \begin{cases} c(t), & \text{if } c(t) > \lambda \\ 0, & \text{otherwise} \end{cases} \quad (4.7)$$

This current model can represent the gas plume as 2D flow. However, in most scenarios the gas needs to be represented as a 3D flow usually with building resolving capabilities. To start we will examine just rural case with no buildings. Furthermore solving the wind field for 3D case results in an increase in computation because the number of discrete points grows as $1/\Delta^3$ if $\Delta = \Delta x = \Delta y = \Delta z$.

In [1], the model was extended to 3D by utilizing a detection probability based on experimental observations in [10; 277]. The experiments showed that the detection probability varies as a function of altitude and wind speed. An example of this can be seen in Fig. 4.4, where the sigmoid function can be fit to the probability of detection given different wind conditions or atmospheric stability classes.

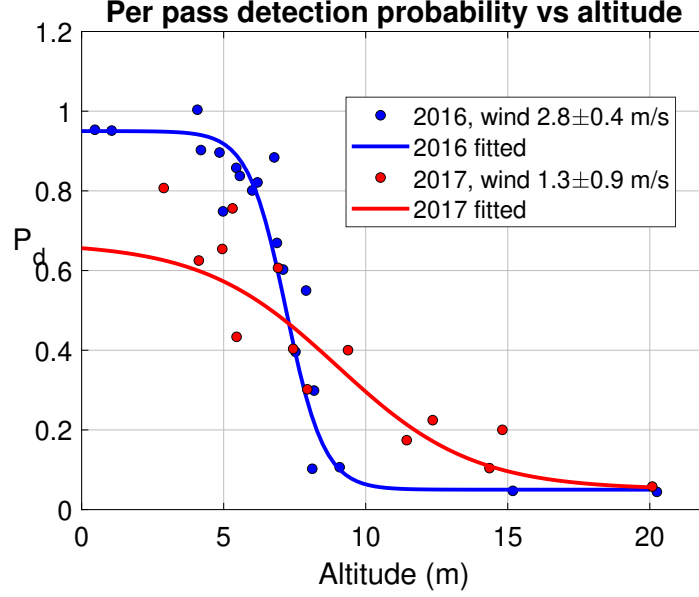


Figure 4.4: Contrast in measurements of two different wind conditions stable (2016) vs unstable (2017) under controlled methane release of 5 standard cubic feet per hour (SCFH) at 30m downwind [10].

By using this knowledge one can essentially extend the 2D model to 3D. Given that sensors do not always detect the target gas, we can introduce a conditional detection probability P_d . Letting the P_d be a function of downwind distance d , altitude z , and wind velocity $|\mathbf{v}_a|$. Making the assumption that the lateral or cross wind position does not affect the probability with respect to the simulation, the relation becomes, $P_d(\delta_d|z) = f(d, z, |\mathbf{v}_a|)$. Given the nature of the mass balance method, the downwind distance d does not readily change, and therefore can be approximated as

$$P_d(\delta_d|z) \approx \frac{P_d(\delta)}{1 + e^{M(z-z_{bias})}}. \quad (4.8)$$

Here M is the slope at the half probability point, and z_{bias} is the altitude at which half probability occurs. From the sigmoid curves in [10] (see Fig. 4.4), $P_d(\delta) = [0.62, 0.9]$, $M = [0.42, 1.5]$, and $z_{bias} = [7.2, 9]$, where the average can be chosen to mirror the conditions of the experiment. However, if the gas is not in contact of the ground the distribution does not capture the vertical behavior. By using a Gaussian normal distribution,

$$P_d(\delta_d|z) \approx P_d(\delta) \exp\left(\frac{-(z - z_{bias})^2}{2\sigma_z^2}\right), \quad (4.9)$$

we can alleviate this limitation. Although this only works if the source location and downwind distance are known ahead of time. Additionally, there is a relationship between the value of M , z_{bias} , and the atmospheric stability classes that needs to

be further investigated. The atmospheric stability classes (A-F), first introduced by Pasquill in 1961 and later reformulated by Gifford, are typically referred to as the Pasquill-Gifford (PG) curves [177]. The PG curves relate wind speed and solar irradiance to the dispersion coefficients that govern the spread of the plume. This is represented in the Gaussian plume model, or sometimes referred to as the infamous ‘bell curve’,

$$C(d, d_l, z, H) = \frac{Q_p}{\pi \sigma_{d_l} \sigma_z U} \exp\left(\frac{-d_l^2}{2\sigma_{d_l}^2}\right) \times \left[\exp\left(\frac{-(z-H)^2}{2\sigma_z^2}\right) + \exp\left(\frac{-(z+H)^2}{2\sigma_z^2}\right) \right], \quad (4.10)$$

where d_l is the lateral (or crosswind) distance from the plume centerline, $H = h + \Delta H$ is the effective height of the plume, h is the stack height, Q_p is the source rate in g/s and ΔH is the plume rise [278]. The plume rise term can be calculated using models such as Holland’s formula which is dependant on temperature and wind speed. However, Holland’s formula was designed for exhaust stacks and also takes into consideration the pressure, stack diameter, and temperature of the ambient air and stack. As a result if the stack velocity is small, such as with sub-surface leaks that diffuse through the ground, the plume rise term tends toward zero. Therefore other factors should be considered, such as lapse-rate, which takes into consideration how the temperature changes as a function of altitude. Furthermore, for small leaks, even surface temperature may come into effect by introducing vertical momentum through natural convective forces. Holland found that there is a small correlation between the plume rise and the temperature gradient near the ground. Briggs mentioned that the plume rise is only dependent on the temperature gradient of the air to which the plume is rising. It is also mentioned that the gradient near the ground is not a good representation of the gradient at higher altitudes [279]. However, temperature dependence within the DT framework will be reserved for future implementations (see section 4.2).

The simulation source rate, $\hat{Q}_*$, given in molecules per second, and is determined by the number of filaments that are discretely released from the point source location. The particular way the filaments are released depends on two parameters, the number of filaments per puff n_f and the number of puffs per filament n_p . The filament source rate can then be defined by $Q = \hat{Q}_*/(n_f n_p)$, which is used in (4.5). The different combinations of these parameters can result in qualitative changes in the measured signal from the sensor. This qualitative behavior can be matched to the measured signal of the physical system and sensor (i.e. behavior matching). The behavior matching can be undertaken given an observable data set that captures these qualitative changes within the model. The key point here is observable data and thus the question becomes, how to get it? The answer depends generally on two things: sensing in the right place and sensing at the right time. Given a set of observations $\mathbf{s}$ we want to infer the states or plume field, $\mathbf{x}_i$. This observation problem can be formulated as a state estimation problem using orthogonal decomposition methods

based on singular value decomposition (SVD) [280]. Essentially the problem can be represented as a $\mathcal{L}_2$ minimization,

$$\mathbf{s} = \mathbf{H}\mathbf{x} \approx \mathbf{H}\Phi\boldsymbol{\nu}, \quad (4.11)$$

$$\boldsymbol{\nu} \in \arg \min_{\tilde{\boldsymbol{\nu}}} \|\mathbf{s} - \mathbf{H}\Phi\tilde{\boldsymbol{\nu}}\|_2^2, \quad (4.12)$$

where Φ represents the orthogonal modes and $\boldsymbol{\nu}$ the associated coefficients. This approach can yield solutions that are often ill-posed and not necessarily unique. However, it can be shown that the estimation error depends on the approximation basis Φ (orthogonal modes) as well as the measurement operator, $\mathbf{H}$,

$$\|\mathbf{x}_i - \hat{\mathbf{x}}_i\| = \|(\mathbf{I} - \Phi(\mathbf{H}\Phi)^+\mathbf{H})\mathbf{x}_i\|. \quad (4.13)$$

The symbol $^+$ represents the Moore-Penrose pseudo-inverse. The measurement operator depends entirely on the sensor location and not the orthogonal modes. Therefore, the choice of sensor location is an important task. For the single sUAS measurement system, this comes in the form of path planning and is subject to disturbances from changes in the stability of the wind. When applied to the mass balance method, path planning involves scanning back and forth at different altitudes inside a region of interest. The size of this region can incorporate twice the expected plume dispersion (deviation from the plume centerline) σ_{d_l} and σ_z , such as in [10] to ensure a likely encapsulation of the plume (see Fig. 3.18). This approach aims to capture rich signal data for conducting the mass balance method. However, this process takes time to complete with a single sensor system, and thus can present spatial-temporal issues when reconstructing the mass flux.

To advance the pre-Digital Twin model (L1) to a Digital Twin (L2) we need to have the physical twin exist and have the virtual system model of the physical twin. This step requires a system identification like optimization of the pre-Digital Twin model, given a dataset with observable properties and tunable parameters to minimize on. Once this minimization is complete the Digital Twin will allow for testing different path planning strategies to improve mass balance method results. The tunable parameters (see Fig. 4.5) in this DT model consist of: the effective radius r_e , the number of filaments per puff n_f , the number of puffs per second n_p , the intermediate scale mixing and stirring dispersion σ_η , and a correction factor c_f where the corrected simulation source rate is $(\hat{Q}_*)_c = c_f\hat{Q}_*$. The value of $\hat{Q}_*$ is chosen to be the same as the physical source rate, such that $Q = (\hat{Q}_*)_c/(n_fn_p)$, and c_f is determined through behavior matching optimization.

4.2 Methane Odor Abatement Simulator with Digital Twins (MOABS/DT)

In this section we outline a DT for use in reducing fugitive gas emissions by early detection and quantification, namely, methane odor abatement simulator (MOABS/DT), which, is an advancement to the POSIM-based DT in the previous Section.

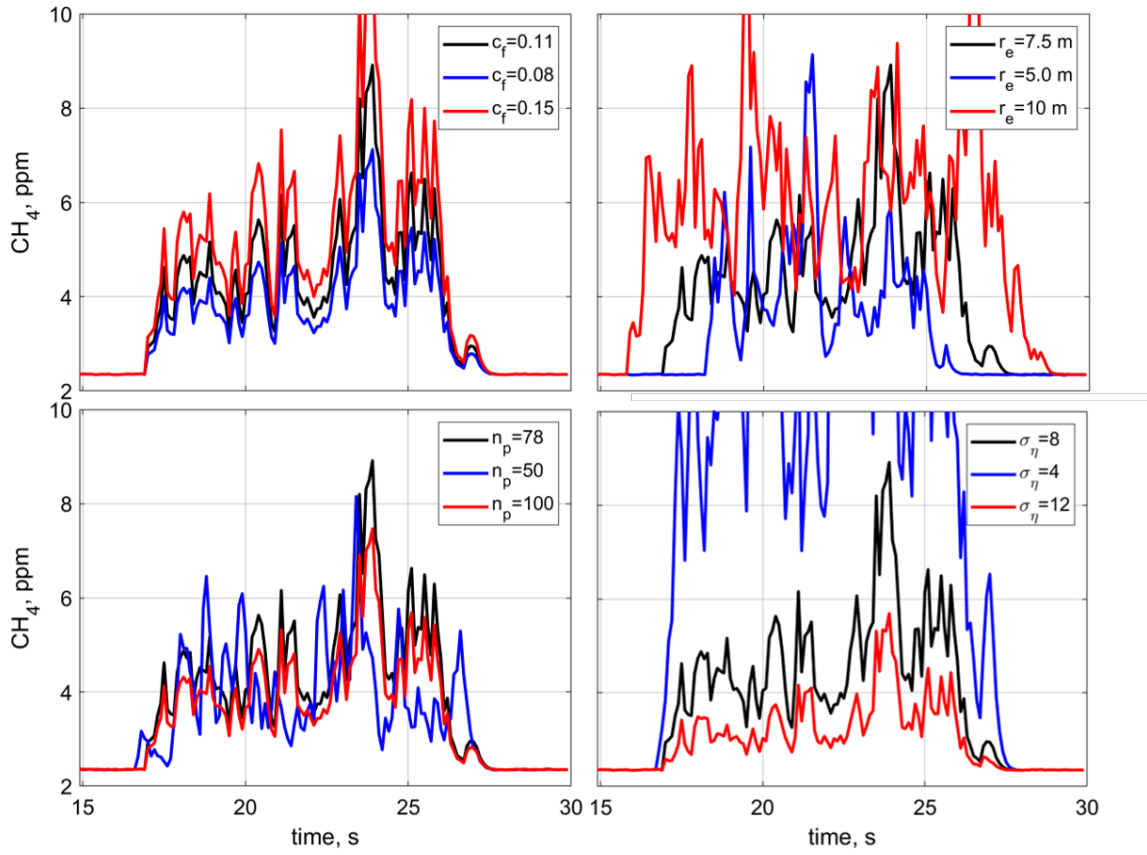


Figure 4.5: The timeseries behavior and sensitivity with tuning parameters θ given one pass through the plume.



Figure 4.6: A satellite image of the Colorado State University methane emission technology evaluation center (METEC), sub-divided into sites and equipment pads are numbered for simplicity in topology modeling.

The real world site we choose to focus on is METEC (see Fig. 4.6), as it was designed for testing and evaluating new technologies as well as provides a benchmark for future research efforts.

The first step is the target system definition. We want the simulation requirements to be: create a digital representation of METEC that can simulate short time-scale behavior of emission dynamics and provide real-time or near real-time operation; create a L2 DT for analyzing performance of detection and quantification strategies; and conduct offline control of DT for studying how environmental variables impact behavior of the DT (such as source location, atmospheric stability, and surface topology). The overall purpose of the system is to connect a DT representation of a physical asset (i.e. METEC) for running virtual controlled release experiments.

The second step is to gather all the system documentation and known problems. Since direct feedback from the L2 DT is not applicable in this case study, control algorithms seek to perform analysis, in an offline sense. For example, the source location using measured data from sparse sensors throughout the field [281] and perform behavior matching using synthetic and real timeseries data [1]. The sensor used in this application is a tunable diode laser absorption spectrometer (TDLAS). This sensor makes point based measurements at 5Hz with sensitivities in the 10 ppb s^{-1} [9]. It is mounted out front of a multi-rotor aircraft (or drone) that actuates the location in space and time. The full state of the drone can be realized using the vehicle dynamics

and motor inputs or treated as a subsystem in guided control mode with single or double integrator dynamics. One of the common problems with methane emissions is detection itself. Having good measurements of the local wind field is not generally possible, so in situ wind or static wind measurements are used to inform the operator of where to search (e.g. downwind and around equipment). Flying upwind and downwind of the source can be used to isolate leaking equipment or sections of a site. Multiple point source emissions are often difficult to distinguish using point based measurements, which include sources from adjacent sites that may drift into the measurement region through variable wind conditions. Utilizing on-board computers and WiFi based communication systems, data can be streamed to a local computer for feedback of the real system. Upon completion of physical data gathering process, the DT can be used to behavior match the system and exploit new methods of detection and quantification performance.

The third step in the construction is the multi-domain simulation. For this case study, there is a need to model diffusion and transport of the methane gas, as well as the model the dynamics of the wind field. Since these domains are governed by partial differential equations, namely, the advection diffusion equation and viscous Burger's equation, the computational cost can be quite high depending on numerical strategy. For these reasons, we adopted the 2D small time-scale filament model by [120], which stochastically treats the chemical transport at different length scales. The wind field is solved using an implicit finite differences approach [275]. The dynamics of the chemical sensor can be modeled with an effective sensor area and a low pass filter. The dynamics of the actuator (e.g. multi-rotor drone) can be modeled using single or double integrator dynamics as well. To extend the model to 3D power law scaling of the wind is applied [282]. The site topology can be modeled using simple geometric shapes defined at locations specified by the equipment pads (see Fig. 4.6). The interaction between topology and gas filaments can be implemented with a filament collision model.



Figure 4.7: (a) Site level emission survey and quantification from METEC using TDLAS equipped drone. (b) The survey begins with a perimeter flight to identify leaking equipment. (c) Once a leak is identified the quantification can be done using downwind and upwind mass balance flights.

The fourth step is the DT assembly and behavior matching. The DT is assembled

by first connecting wind field model with the chemical filament dynamics and the collision model defined by the topology. Next, the interaction with the sensor and actuator models for measuring observations in space and time. Once the DT is assembled, the behavior matching to the physical system of interest can be done by calibrating simulation leak rates with observed timeseries concentration signals [1] (see Fig. 4.1). This is done with a training data set (a subset of the overall data gathered) that captures rich measurement data and knowledge of meteorological conditions (e.g. atmospheric stability [177]). These data sets have been gathered, in this example, through a series of survey flights that include: initial perimeter sweeps for detection of leaks, and mass balance flights for quantification of emissions (see Fig. 4.7).

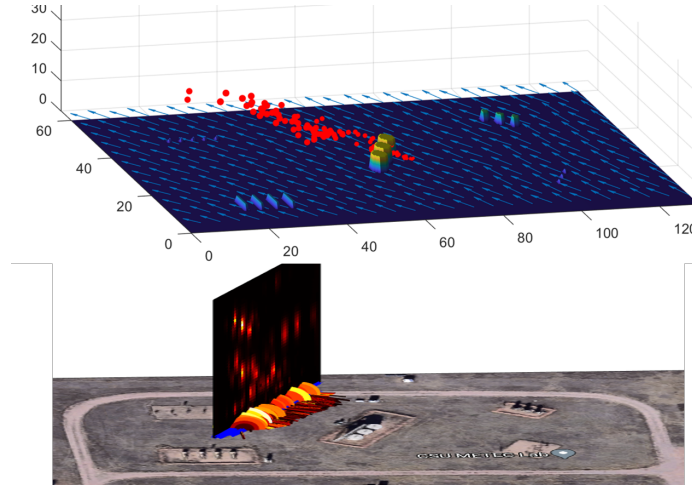


Figure 4.8: Running the DT in parallel with the physical system, the dynamics of the gas filaments are being lofted into the air due to the cylinders on equipment pad 3 at site 1 (top). The physical observations seen from experiment and calculated with the mass balance (bottom) reflect those in the DT.

The fifth and last step in the construction process, is the DT validation and deployment. The validation can be done by comparing the quantification results (previously unseen data) to the behavior matched DT (from step 4) in an offline setting. A well defined/trained DT will produce results suitable for prediction defined within the purpose of the DT. The DT can then be deployed to run in parallel with the physical system (see Fig. 4.8) gathering system wide data in the process (from static sensors or other mobile systems deployed). This opens up the possibility to perform online behavior matching of parameters, ultimately, providing estimates to key parameters (such as source location [281]) for smarter path planning as well as improving quantification efforts [51].

As we continue to use energy sources such as natural gas in our homes and methane emission reduction continues to be an focus in society, there will always be a need to detect and quantify emissions for repair and validation. This work proposes the use of DT's for improving and testing the limitations of advance leak detection and quantification techniques. Improving accuracy and frequency of surveys that can

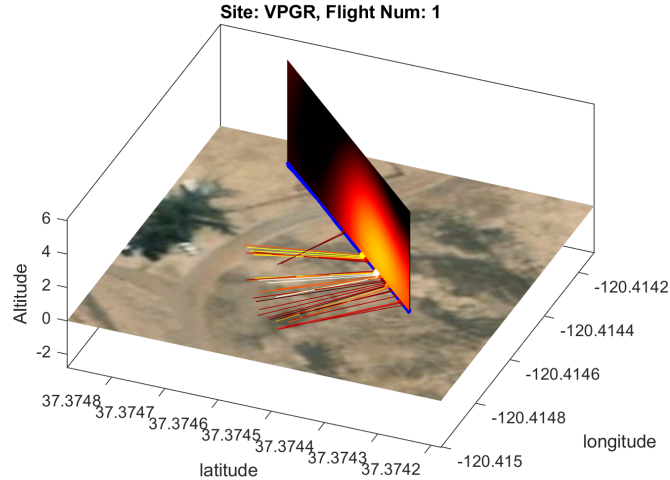


Figure 4.9: Example mass balance measurement for one curtain flight.

lead to reductions in overall emissions. The DT provides a way to cost effectively test techniques and optimize approaches. The DT can also provide an added layer of intelligence for site performance and smarter path planning. Additionally, if component level DT's are incorporated within the site equipment (valves, tanks, etc.), prognosis features such as remaining useful life (RUL) can be achieved. Using analytics such as RUL, the frequency of surveys can be prescriptively increased for areas near the failure limit.

4.3 A Case Study: Single-sUAS-based Characterization of Ground-To-Air Emissions

The experimental data used to convert the L1 Digital Twin to L2 was based off a controlled release single source field experiment in the Merced Vernal Pools and Grassland Reserve just north of the University of California, Merced campus (See experimental setup in Section 3.5.2).

4.3.1 Behavior Matching

Once the experimental data is collected, the most data information rich flights are to be selected to perform behavior matching with the L1 digital twin. In this experiment, some of the curtain flights (such as flights 13 and 14) resulted in no detected emissions and cannot be used to behavior match the model (see Table 4.2). After analysing the 16 flights using the mass balance method (see Fig. 4.3), flight 1 showed stable wind conditions, consistent methane detection throughout the ascending transects, and good estimation of the true source rate. Thus, flight 1 was chosen for behavior matching the sensor signal of the L1 digital twin to the experimental sensor signal (see

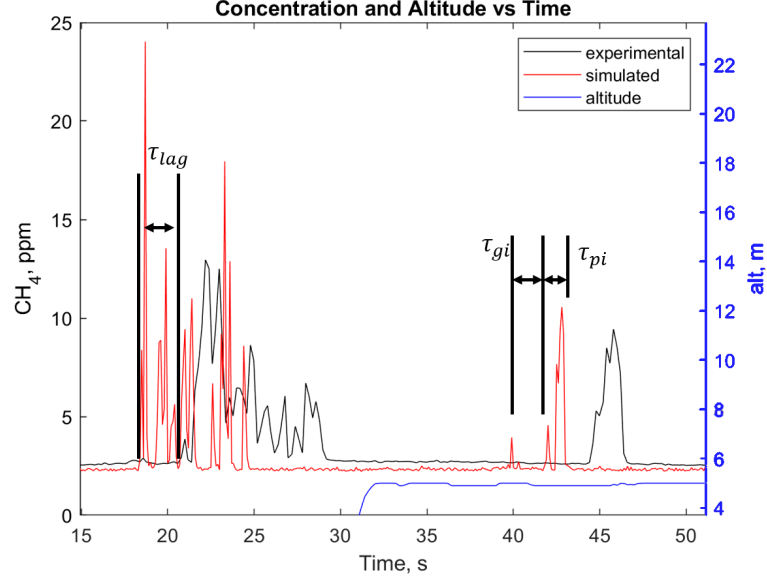


Figure 4.10: The concentration time series in ppm during flight curtain 1, where $n_p = 40$ puffs/sec and the symbols τ_{gi} , τ_{pi} , and τ_{lag} represent the i -th gap length, i -th pulse length, and the time lag between the simulation and experimental data [1].

Fig. 4.10). The simulation source rate $\hat{Q}_*$ was set to ($\hat{Q}_* = 2.4 \times 10^{24}$ molecules/s). Using the PG curves given the current meteorological conditions (i.e. atmospheric stability) the initial condition of σ_η was chosen to be 8 m. The wind speed and direction measured from the experiment was used as a direct input into the L1 digital twin, using the mean wind field approximation. Using the behavioral characteristics of the hyperparameters shown in section 4.1 and in an attempt to reduce complexity, let $n_f = 1$, such that $n_f n_p$ has units as filaments per second. Heuristically tuning the hyperparameters, to give a best guess of $\theta = [c_f, r_e, n_p, \sigma_\eta]$, the following initial conditions were chosen: $c_f = 0.2$, $r_e = 5$ m, $n_p = 40$ puffs/sec, and $\sigma_\eta = 8$ m. It can be also seen that there is a lag present in the system (see Fig. 4.10) which can be attributed to model inconsistencies as well as initial conditions of the plume. Since in this case study there is currently no implementation of plume rise (see MOABS/DT in Section 4.2), the average altitude where concentration detection occurred was used to determine the z_{bias} term and the vertical dispersion was set to $\sigma_z = 5$ m.

If we represent the L1 digital twin given hyperparameters θ as a function at time t ,

$$\{\hat{c}_m, \mathbf{x}_i\}_t = \{f(\mathbf{x}, \mathbf{v}|\theta)\}_t, \quad (4.14)$$

with the sUAS position (i.e. sensor position) $\mathbf{x}$ and $\mathbf{v}$ describes the wind speed and direction measurement using a model 81000 3D ultrasonic anemometer from the R. M. Young company, USA. The output consists of a concentration measurement $\hat{c}_m(t)$ and filament positions $\mathbf{x}_i$. In the physical experiment the locations of the plume filaments are not observable given that methane gas is not visible to the human

eye. The plume can potentially be observed using a large number of high quality sensors, of which, is expensive, difficult to set up, and only be inferred using spatial interpolation techniques. Thus, the only output available to optimize over is the concentration measurement, $c_m(t)$, in the least squares sense. In order to compensate for the stochastic behavior of the plume dynamics, the cost function was designed to average $n_T = 20$ trials. Unfortunately, this leads to multiple possible solutions to the minimization. Therefore, additional constraints on the hyperparameter space, Ω , and regularization can be added to constrain the magnitude of the concentration measurement such that,

$$\theta = \arg \min_{\theta \in \Omega} J(\theta), \quad \Omega = \{\theta | \theta_{min} \leq \theta \leq \theta_{max}\}, \quad (4.15)$$

$$J(\theta) = \frac{1}{n_T} \sum_{k=1}^{n_T} \left[0.2 \left(\sum_{i=1}^{n_t} \{c_m\}_i - \sum_{j=1}^{n_t} \{\hat{c}_m\}_j \right)^2 + 0.8 (c_{max} - \hat{c}_{max})^2 \right], \quad (4.16)$$

where $c_{max} = \max(\mathbf{c}_m)$. The cost function weights were chosen to balance the effects of each term and were determined from trial and error.

The solution to (4.15) can be found using a variety of optimization methods and techniques, such as Gradient Descent or Stochastic Gradient Descent. Faster convergence can be achieved using Hessian based approaches but due to the cost of the forward model calculation and considering complexity of implementation these methods were not considered. Other gradient free approaches are more desirable for this problem such as Genetic Algorithm, Extremum Seeking Control, or Pattern (Direct) Search can be applied. With simplicity in mind, the Pattern Search Optimization was used here to behavior match the digital twin to the experimental data. After 65 iterations and over 390 cost function evaluations (7880 forward model runs) the gradient of the cost function $dJ/dt \approx 0$ (see Fig. 4.11). The model parameters were found to be: $c_f = 0.11$, $r_e = 7.5$ m, $n_p = 78$ puffs/s, and $\sigma_\eta = 8$ m.

4.3.2 Quantification Methods

To calculate and perform the mass balance method, sometimes called the box method, requires a way to measure the flux entering or leaving a control volume (see Section 3.2.3). Typically, upwind and downwind curtain flights are used to measure the flux. However, in the case of a controlled release the downwind curtain flight is sufficient to calculate the flux, granted the plume is encapsulated in the flight path (see Fig. 3.18). Using the timeseries vector $\mathbf{c}_m = [c_m(\Delta t), c_m(2\Delta t), \dots, c_m(n_t\Delta t)]^T$ with the corresponding measurement locations, $\{\mathbf{x}\}_1^{n_t}$, a concentration matrix, $\mathbf{C}_m$ can be estimated on an n by m spatial grid through spatial interpolation methods. The flux calculation can be then given as the integration over curtain area with the average wind vector multiplied by the cosine angle of the normal vector to the curtain with

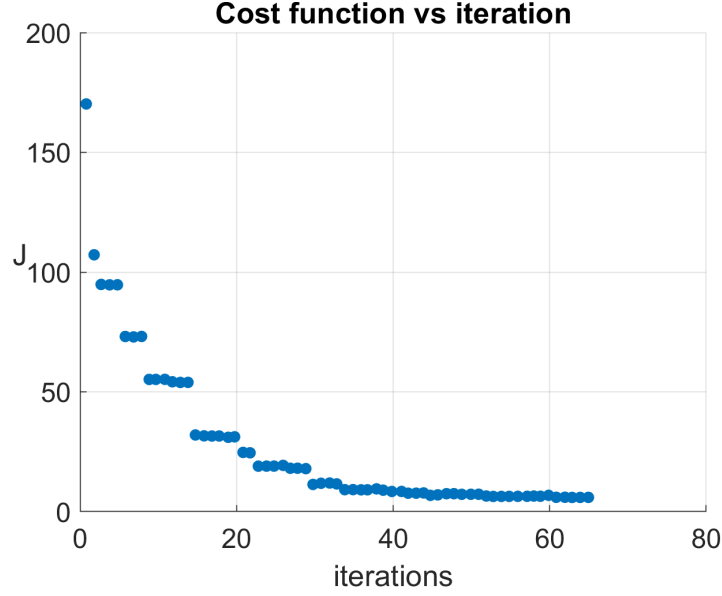


Figure 4.11: Pattern search method convergence plot for behavior matching of flight 1 given the cost function, $J(\theta)$. The algorithm converged as mesh size tolerance was reached to ≈ 5.813 .

the enhanced concentration,

$$\phi = \int_A \bar{\mathbf{v}} \cdot \hat{\mathbf{n}}[(\mathbf{C}_m - C_b)]dA. \quad (4.17)$$

Here C_b represents the background concentration measurement. An example of a flux plane calculation can be shown in Fig. 4.3. Several kinds of techniques can be applied to the timeseries data to generate $\mathbf{C}_m$, such as Ordinary Kriging, Inverse Distance Weighting (IDW) or statistical gas distribution modeling (e.g. Kernel DM+V/ W). For the case of Kriging, a linear unbiased estimator, can be solved using a variogram (or semi-variogram),

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (s(x) - s(x+h))^2. \quad (4.18)$$

Generally, the redundancy and closeness covariance matrix are calculated from fitting the experimental variogram to a function with respect to the spatial distance or lag, such as the exponential or Gaussian, and knowing the covariance of the measured points (referred to as the sill). This method requires second order stationarity in the spatial observations [205]. Unfortunately this is only approximately the case during stable atmospheric conditions. Methods such as IDW look at the inverse distance when determining weights and does not consider other spatial statistics as in Kriging. There has been some efforts to improve this by looking at minimum error variance within IDW [283]. Wind information has been incorporated into Kernel

DM+V/W approach in 3D [284]. Another adaptation to statistical gas distribution mapping approach was implemented in [285] by using the Gaussian plume kernel, which incorporates the wind information into the covariance function in determining the weights. The downside to this approach is that it assumes the source location. This is not ideal for determining the flux in this case. The spatial-temporal effects on modeling have been looked at as a separation between the spatial covariance and the exponential decay between temporal measurements [286; 287],

$$w_i^{(k)} = \mathcal{N}(|\{\mathbf{x}\}_i - \mathbf{x}^{(k)}|, \sigma) \varphi(t_*, t_i), \quad (4.19)$$

where $\mathcal{N}$ is a Gaussian weight function, $(\cdot)^{(k)}$ represents the k -th grid point, $\{\mathbf{x}\}_i$ are the sensor measurements at time t_i and φ is given as,

$$\varphi(t_*, t_i) = e^{-a_\varphi(t_* - t_i)}. \quad (4.20)$$

However, when dealing with complex systems such as in turbulent environments there may be several factors that determine what value of a_φ is. If we look at the weight summation of exponentials, with elements increasing towards infinity, it can be described by fractional order dynamic behavior [288]. The Mittag-Leffler function [289] can be used to describe this behavior,

$$\mathcal{E}_{\alpha, \beta}(t) = \sum_{k=1}^{\infty} \frac{t^k}{\Gamma(\alpha k + \beta)}, \quad (4.21)$$

such that when $\alpha = \beta = 1$, $\mathcal{E}_{1,1}(t) = e^t$. Given that $\beta = 1$ and letting α decrease we can see in Fig. 4.12 that the tail behavior decays much more slowly than the exponential function. By combining the approach in (4.19) with (4.21) we can apply a new weighting scheme with tuning parameter, α ,

$$w_i^{(k)} = \mathcal{N}(|x_i - x^{(k)}|, \sigma) \mathcal{E}_{\alpha, 1}(-a_\varphi |t_* - t_i|^\alpha). \quad (4.22)$$

The value of t_* can be chosen for each grid point based on the time t_i of the closest measured point that minimizes $|\{\mathbf{x}\}_i - \mathbf{x}^{(k)}|$. For the single sensor mapping problem we test the change in sensitivity for $\alpha = 1$, 0.5, and 0.1.

4.3.3 Discussion

After optimization of the L2 digital twin, using the experimental position and wind measurements, each flight was simulated 10 times. The spatial interpolation was carried out using Kriging, IDW, and Kernel DM/ V to see a comparison in quantitative results across the different approaches. It can be observed that all the spatial interpolation methods tend to perform similarly. Also, for the spatial-temporal method using (4.22) does not show noticeable performance improvement. Future work can be done to address the optimal configurations of these methods and their hyperparameters. It can be observed that the source rate estimations fluctuate in value depending

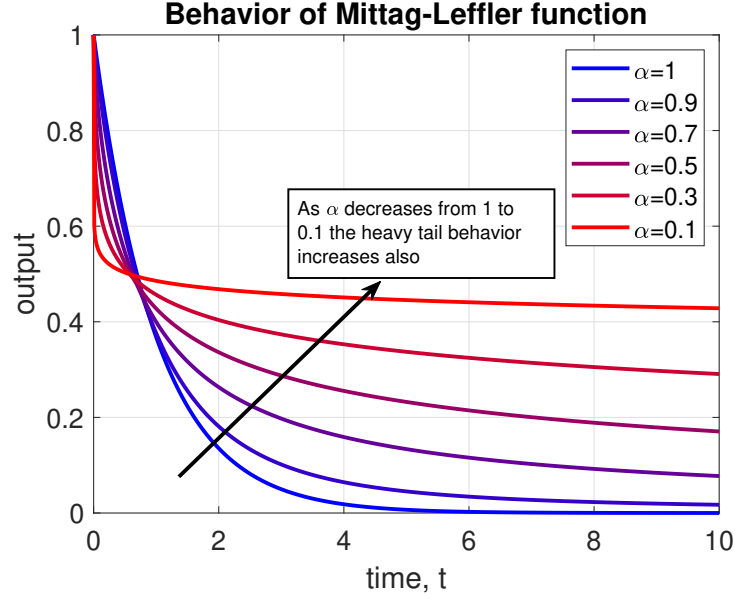


Figure 4.12: Response of the single parameter Mittag-Leffler function for $\mathcal{E}_\alpha(-t^\alpha)$ as $\alpha \in [0, 1]$.

on current weather conditions and turbulence (see Table 4.2). Given small deviations in the wind, the sUAS curtain flight path can appropriately capture the spatial data through the mass balance plane. When the wind deviates slightly, the data we capture can sometimes be spread out over the flux plane (i.e. like a blurred image). As a result, this can lead to higher source rate estimations or to poor detection and characterization of the plume. Since the wind direction fluctuations correlate to different stability classes it makes sense that there should be some correlation with the turbulence intensity as well (see Fig. 4.13). Thus, mindfulness of the current atmospheric conditions are important when planning and conducting these flights. Furthermore, the wind behavior, to some extent, is always changing in magnitude and shifting in direction. It seems this behavior is very rarely calm and stable, which may put physical limitations on the sensitivity of accurate spatial sampling with single mobile sensor systems (at least in the case for air). Comparing the experimental data to that of the digital twin one can observe qualitative similarities in the distribution of sensor detections, as seen in Fig. 4.14. Additionally, larger less frequent spikes at lower altitudes were observed that are consistent with what was experimentally observed (see Fig. 4.15).

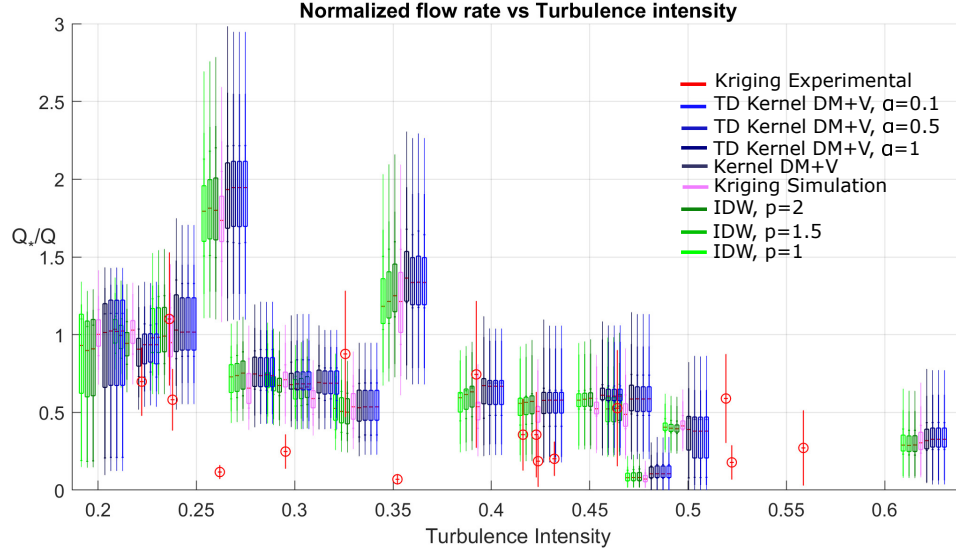


Figure 4.13: Here the turbulence intensity, $T_i = \sigma_u/\bar{u}$, of the wind is plotted against the normalized flow rate, which trends toward $\hat{Q}_*/Q = 1$ as $T_i \rightarrow 0$. A boxplot of the source estimation for each flight (set of 10 simulations), given a spatial interpolation method, is shown for each T_i condition observed from the experimental data. A smaller boxplot (shown within each boxplot) contains tails that represent the uncertainty. We can observe 1) several flights agree with experimental data 2) majority of flights underestimate the flow rate in short due to increased T_i and 3) the remaining flights that disagree are likely not capturing vertical plume behavior.

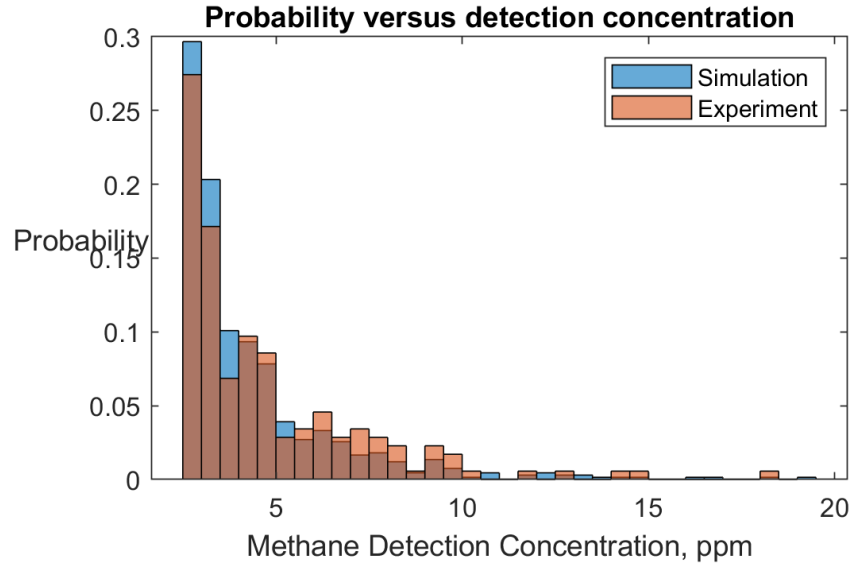


Figure 4.14: The detection concentrations from the simulation and experimental sensor are compared in a probabilistic histogram. While the distributions show that the majority of indications are near background levels, they also agree with the less likely larger concentrations detected in the experiment.

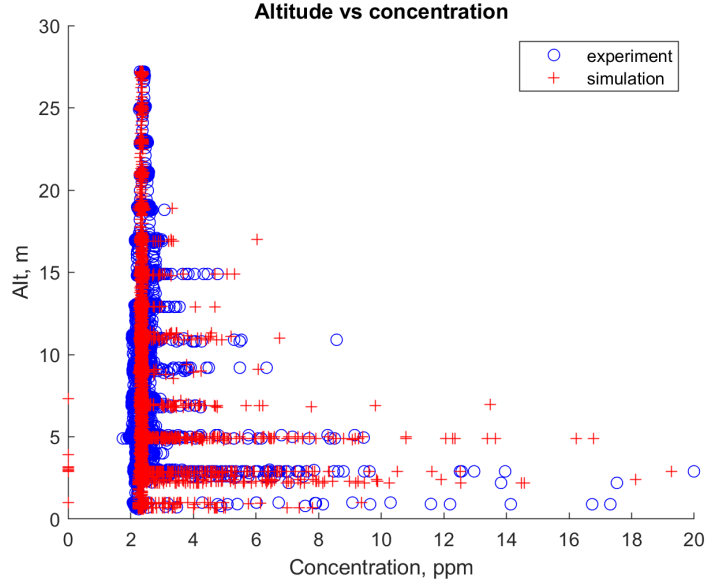


Figure 4.15: The detection altitude versus measured concentrations from the simulation and experimental sensor are compared. The distributions match well with experiment and the probability of detection declines rapidly as the altitude increases from 10-15 m.

Table 4.2: Results from mass balance method for 16 curtain flights are shown below along with the average experimental wind speed, $\bar{u}_e$ and the $\cos \beta_e$ between the normal vector of the flux plane. The flow rates (experimental Q_* and simulation $\hat{Q}_*$) are in SCFH, the subscripts denote the spatial interpolation method, and the superscripts are the associated parameters, where K is Kriging, IDW is inverse distance weighting, and KDM is Kernel DM+V.

Flt	Q_*	$(\hat{Q}_*)_K$	$(\hat{Q}_*)_{IDW}^{p=2}$	$(\hat{Q}_*)_{IDW}^{p=1}$	$(\hat{Q}_*)_{KDM}$	$(\hat{Q}_*)_{KDM}^{\alpha=1}$	$(\hat{Q}_*)_{KDM}^{\alpha=0.1}$	$\bar{u}_e$	$\cos \beta_e$
1	14.2±3.9	13.0±2.8	12.1±2.6	12.4±2.6	12.0±3.2	12.2±3.2	12.2±3.2	1.82±0.43	0.83
2	11.3±3.7	7.0±2.3	6.5±2.2	6.9±2.3	7.0±2.7	7.0±2.6	7.0±2.6	1.75±0.57	0.96
3	2.4±1.5	1.0±0.7	1.1±0.8	1.1±0.8	1.5±1.3	1.5±1.4	1.5±1.4	1.18±0.50	0.67
4	9.6±4.3	6.6±3.0	7.8±3.5	7.3±3.3	8.2±4.2	8.1±4.1	8.1±4.1	1.30±0.51	0.87
5	6.8±3.4	6.2±3.1	6.9±3.4	6.7±3.3	7.6±4.2	7.6±4.1	7.6±4.1	1.25±0.58	0.91
6	7.5±1.8	12.8±3.1	13.3±3.2	13.2±3.2	13.8±4.5	13.8±4.2	13.8±4.3	2.27±0.54	0.97
7	4.6±2.1	6.7±3.3	7.5±3.7	7.5±3.6	8.0±4.4	7.9±4.3	7.9±4.3	1.37±0.57	0.67
8	3.2±1.0	9.2±3.0	8.7±2.8	9.1±2.9	9.0±3.2	9.1±3.2	9.0±3.2	1.93±0.57	0.93
9	4.6±2.5	6.3±3.5	7.1±3.9	6.9±3.8	7.2±4.3	7.3±4.3	7.3±4.3	1.49±0.63	0.83
10	3.5±2.2	4.1±2.7	3.9±2.5	3.9±2.6	4.2±3.3	4.4±3.3	4.3±3.3	1.45±0.81	0.86
11	2.6±1.0	8.4±2.2	9.5±2.5	9.3±2.4	9.8±3.3	9.8±3.2	9.8±3.2	1.25±0.54	0.52
12	2.3±1.0	7.5±2.0	8.5±2.3	8.4±2.3	9.0±2.8	8.9±2.7	8.9±2.7	0.90±0.47	0.43
13*	0.9±0.3	15.6±5.4	16.8±5.8	15.7±5.4	17.7±6.9	17.3±6.8	17.5±6.8	2.47±0.87	0.98
14*	1.5±0.4	22.5±6.0	23.3±6.2	22.9±6.1	24.4±7.8	24.6±7.6	24.6±7.6	2.52±0.66	0.94
15	9.0±2.0	13.2±2.9	12.1±2.7	13.2±2.9	11.4±3.4	11.7±3.2	11.8±3.3	2.07±0.46	0.96
16	7.6±2.6	5.4±1.9	5.1±1.7	5.3±1.8	4.7±3.1	4.4±3.2	4.5±3.2	0.52±0.27	0.27

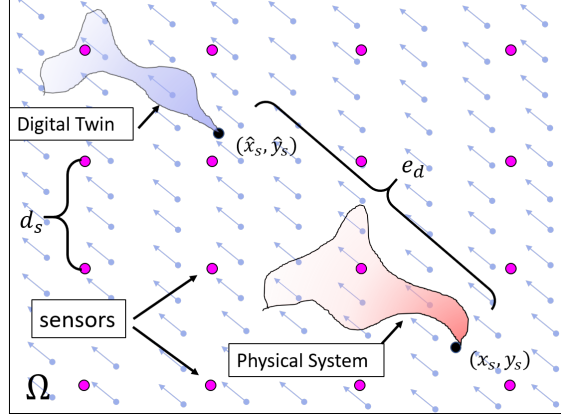


Figure 4.16: Given an unknown source location (x_s, y_s) the DT is behavior matched to minimize the error e_d in the source location estimation $(\hat{x}_s, \hat{y}_s)$ given a sensor distance described by d_s .

4.4 A Case Study: Behavior Matching and Localization Using Fixed Sensor Mesh

Suppose that within an environment of domain Ω , an unknown point source gas release exists at (x_s, y_s) . Other than doing manual surveys (e.g. mobile sensing), the alternative solution is to use fixed sensors, spread throughout the landscape. The placement of these sensors would then depend on the number of sensors considered and where the emission source is likely to be. For a uniform landscape with uniform likelihood of the source position and no obstructions, the sensors take the form of a grid. The spacing of the grid would then depend on the number of sensors (see Fig. 4.16). The goal of this method is to determine the source location based on sparse sensor measurements.

By creating a DT of the system we can analyse the physical and simulated measurements to estimate parameters such as source rate $\hat{q}$ and source location $(\hat{x}_s, \hat{y}_s)$. However, the DT needs to first be behavior matched. Let the DT and physical system be modeled using POSIM [120; 1]. For simplicity, let's consider only the source location problem given N sensors fixed in space. At time k , the i -th sensor location is at $\mathbf{x}_i$ and measures $\rho_{i,k}$. Since the measurements are sparse we can utilize spatial interpolation (e.g. IDW) to improve some of the resolution, from N point source measurement vector ρ_k to a $n \times m$ matrix $\hat{\rho}_k \in \Omega$. Since the wind field is changing and the plume can meander, the observed measurements tend to be noisy. In order to perform behavior matching we need the low frequency behavior of the system to be slower than the control input. In this case, the control is the source location estimate $(\hat{x}_s, \hat{y}_s)$. Utilizing DMD, the low frequency modes can be retrieved from $M + 1$ snapshots of the measurement data directly,

$$Y_1 = \begin{bmatrix} | & | & & | \\ \hat{\rho}_1 & \hat{\rho}_2 & \cdots & \hat{\rho}_M \\ | & | & & | \end{bmatrix}, \quad Y_2 = \begin{bmatrix} | & | & & | \\ \hat{\rho}_2 & \hat{\rho}_3 & \cdots & \hat{\rho}_{M+1} \\ | & | & & | \end{bmatrix}, \quad (4.23)$$

$$\Psi = \mathbf{Y}_2 \mathbf{V}_Y \Sigma_Y^{-1} \mathbf{W}_Y = \begin{bmatrix} | & | & & | \\ \psi_1 & \psi_2 & \cdots & \psi_r \\ | & | & & | \end{bmatrix}. \quad (4.24)$$

The recovered reduced order DMD modes Ψ show the low frequency behavior of the spatially interpolated estimates $\hat{\rho}$ (see Fig. 4.17) such that the r -th mode ψ_r moves slow compared to the control. By performing DMD on both the physical system and the DT, the source location distance error can be defined,

$$e_d = |\mathbf{x}_s - \hat{\mathbf{x}}_s|, \quad (4.25)$$

$$\mathbf{x}_s = \arg \max_{\mathbf{x} \in \Omega} |\psi_r|, \quad \hat{\mathbf{x}}_s = \arg \max_{\mathbf{x} \in \Omega} |\psi_r^{DT}|, \quad (4.26)$$

where ψ_r^{DT} is the DT's r -th DMD mode. The DMD modes can be selected depending on the frequency of behavior of interest. For this case the first mode is used. Given the error, a controller can be designed to provide self optimizing control (such as extremum seeking control [47; 290]) or something as simple as a proportional control law. In this case we chose the latter,

$$\dot{\hat{\mathbf{x}}}_s = k_p(\mathbf{x}_s - \hat{\mathbf{x}}_s). \quad (4.27)$$

The potential limitations to this method are with the persistent of excitation (PE) of the input signal (e.g. the plume measurements) and the physical sensor's properties. Without PE the SVD of the time snapshots can result in the eigenvalues being zero and the DMD modes cannot be computed. Additionally, the sensor's response and relaxation time can potentially limit the performance if they are too slow for the environmental conditions (or meandering wind).

4.4.1 Simulation

To illustrate the behavior matching method, we set up a simulation experiment where the source location is at $\mathbf{x}_s = (80, 10)$, and the initial guess of the DT source location to be $\hat{\mathbf{x}}_s(0) = (45, 45)$. The simulation domain spans a 90×90 unit area. The initial wind conditions are stable with no meandering ($a = 0.01$, $b = 0.01$, and $G = 1$) and unstable with meandering ($a = 0.04$, $b = 0.04$, and $G = 5$). The average wind direction is $\bar{\mathbf{u}} = [-1, 1]$. The DMD modes are computed using $M + 1 = 201$ snapshots using (50×50) IDW point estimates from 100 sparse sensors. The source location distance error is calculated and used to update the DT source location estimate in (4.27). The evolution of this can be seen in Fig. 4.18. The results from the no meandering case and from the meandering case are shown in Fig. 4.19. It can be seen that the meandering wind, in slightly neutral/unstable atmospheric conditions, can

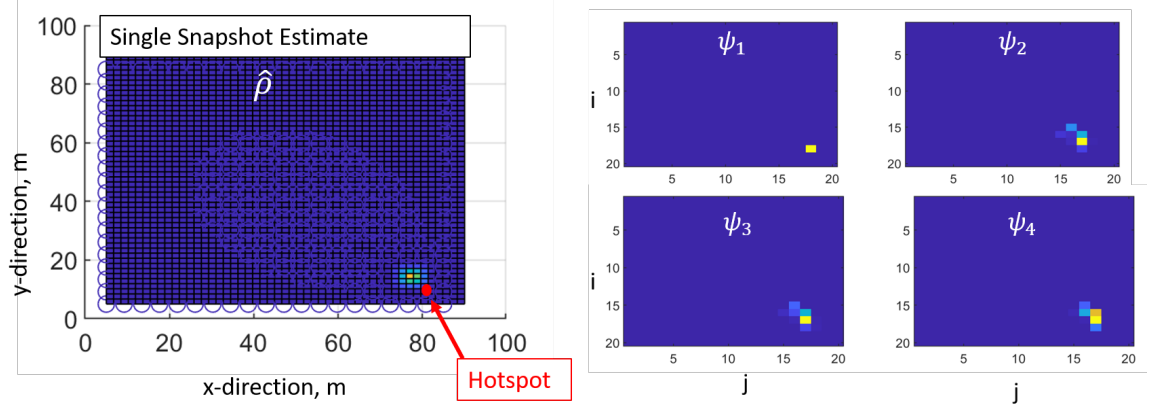


Figure 4.17: The results from a single snapshot of data using IDW is shown on the left and the first four spatially interpolated DMD modes are shown to the right (starting with the top row). The low frequency behavior creates a hotspot that varies less with time, allowing for behavior matching optimization.

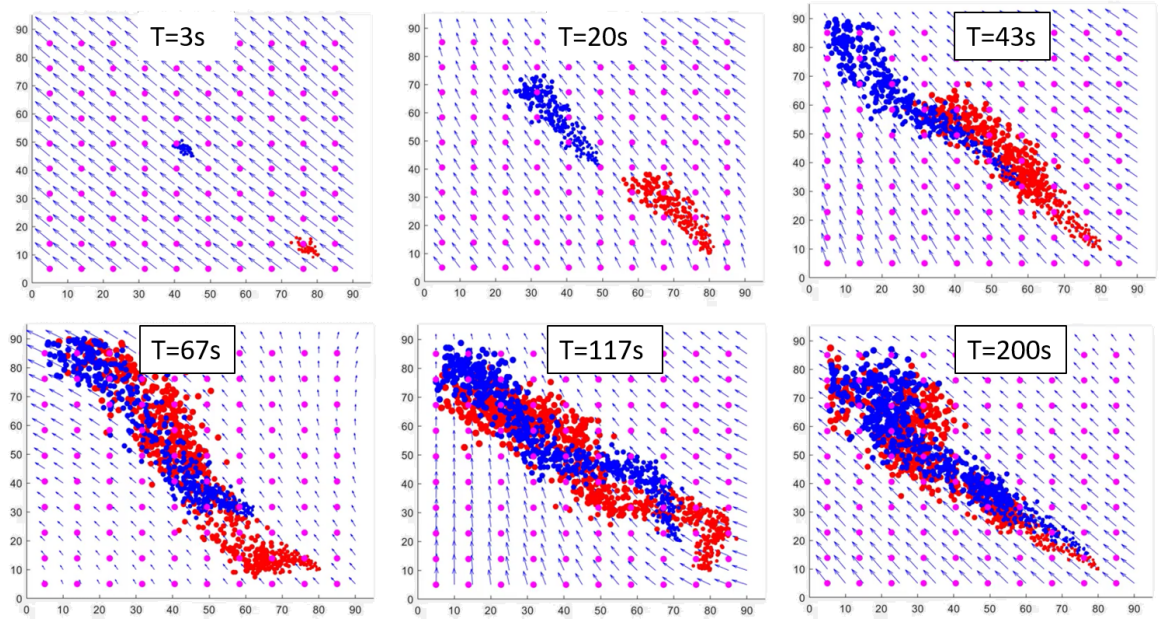


Figure 4.18: The time evolution of the DT behavior matching is shown for various time steps. The DT is shown in blue and the unknown physical system is shown in red.

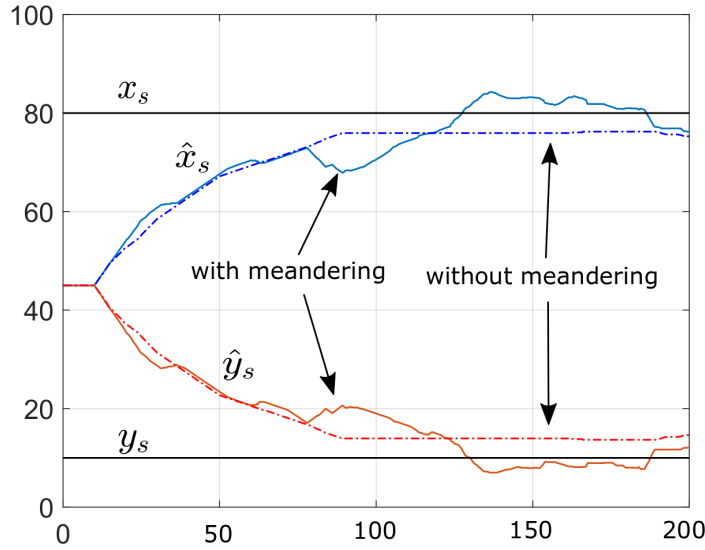


Figure 4.19: Given both meandering and no meandering wind conditions, the behavior matching method converges around the true source location. Even though the additional turbulence from the meandering conditions makes the convergence noisy, the estimation of the source location seems to improve.

push the plume to different locations in the space. For example, a source may be placed within the sparse grid of sensors, in such a way that certain wind directions may provide no observations to the sensors. However, if wind variability is introduced, the plume may redirect onto other nearby sensors, where observations can be made. Depending on the source location and sensor density this can be seen as favorable, however, further investigation is required. Preliminary results on the impact of sensor density and meandering can be seen in Fig. 4.20.

4.5 A Case Study: Multi-sUAS-based Emission Quantification with DE-CVT

In this section we discuss the multi-sUAS continuous source rate estimation method based on CVT coverage control, DMD, and the mass balance flux plane with spatial interpolation (outlined in Section 3.2.3).

4.5.1 CVT Coverage Control

The coverage control problem can involve applications such as path planning, obstacle avoidance or deployment of mobile sensors to name a few. Given a domain of interest Ω , the local positions of the sensors $P = (p_1, p_2, \dots, p_N)$ can be divided into N

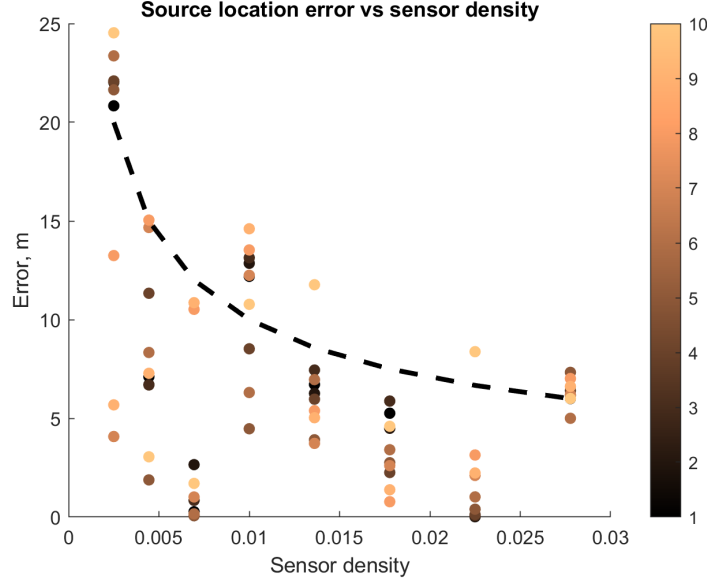


Figure 4.20: Preliminary results in source localization error are shown as a function of sensor density. A single source location, initial guess, and mean wind direction are used to vary meandering amplitude $1 \leq G \leq 10$ (where 10 represents larger meandering) for a fixed natural frequency $\omega_n = 0.2$ and damping ratio $\zeta = 0.1$ ($a = \omega_n^2$ and $b = 2\omega_n\zeta$). The dashed line is the distance between adjacent sensors and shows the method provides some improvements in resolution.

polytopes $\mathcal{V} = (\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_N)$. The polytopes are determined using centroidal Voronoi tessellations (CVT)[291]. The cost function that controls the CVT desired positions is,

$$J(P) = \sum_{i=1}^N \int_{\mathcal{V}_i} \rho(q) |q - p_i|^2 dq, \text{ for } q \in \Omega. \quad (4.28)$$

The partial derivative can be used with the mass, m_i , and center of mass, c_i , to identify the critical point,

$$m_i = \int_{\mathcal{V}_i} \rho(q, t) dq, \quad c_i = \frac{1}{m_i} \int_{\mathcal{V}_i} q \rho(q, t) dq, \quad (4.29)$$

$$\frac{\partial J(P)}{\partial p_i} = 2m_i(p_i - c_i)^T. \quad (4.30)$$

This means that $c_i = p_i$ for all $i = 1, 2, \dots, N$ is a minimizer of the CVT. Treating the relation as a gradient, Lloyd's algorithm can be used,

$$\dot{p}_i = -k(p_i - c_i). \quad (4.31)$$

The general algorithm to compute the locations of actuators by computing the CVT [292; 293; 294] using Lloyd's method. The general approach is described in [281].

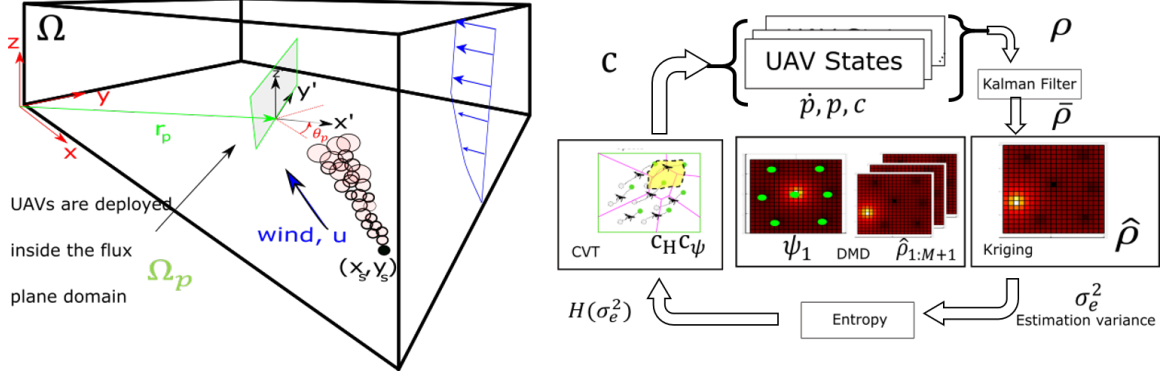


Figure 4.21: This diagram depicts the simulation environment setup and framework implementation flowchart.

4.5.2 Method

In this framework, we seek to make continuous source rate estimation by controlling the UAV positions inside a flux plane Ω_p using CVT coverage control, utilize spatial interpolation with N sparse measurements to build a concentration map estimation, and then apply the mass balance method [215] to quantify the source rate (see Fig 4.21). However, due to the turbulent behavior the measurements made, they tend to be noisy and intermittent. The severity of the disturbances can be attributed to the current atmospheric stability, which is a function of wind speed and solar irradiance [177]. During more stable conditions (e.g. turbulence intensity is near 0) the data can be considered quasi-stationary (in the flux plane frame of reference). To help smooth this incoming data, a Kalman filter [295; 296] can be applied to the i -th UAV measurement at time k to obtain $\bar{\rho}_{k,i}$.

Once the N UAV flux plane measurements $\bar{\rho}_k$ are made, an estimation of the $n \times m$ point spatial concentration map $\hat{\rho}_k \in \Omega_p \subseteq \Omega \in \mathbb{R}^3$ can be computed using ordinary kriging, $\bar{\rho}_k \rightarrow \hat{\rho}_k$ in (3.58). The semivariogram used to find $\hat{\rho}_k$ is the exponential [205],

$$\hat{\gamma}(h) = a + (\sigma^2 - a)(1 - \exp \frac{-3|h|}{r}). \quad (4.32)$$

In (4.32), a is the nugget, σ^1 is the sill, and r is the range. Other potential fits for the semivariogram function include the Gaussian as well as α -stable distribution but were not pursued in this work.

Since the concentration map is assumed to be quasi-stationary, the results from ordinary kriging can be noisy. Therefore we chose to explore the slow moving spatial temporal dynamics of the estimated concentration map inside the domain Ω_p using DMD. Given $M + 1$ snapshots of $\hat{\rho}_k$ the DMD modes can be calculated from (4.23)-(4.24) and used to represent the density in the CVT cost, $J(P)$,

$$J(P) = \sum_{i=1}^n \int_{V_i} [\psi_1(q) + H(\sigma_e^2(q))] |q - p_i|^2 dq, \quad (4.33)$$

for $q \in \Omega_p$.

In addition to computing the DMD modes, a byproduct of the ordinary kriging is the estimation error variance, σ_e^2 . This can in turn be used to calculate the entropy of the error variance as a weight for the CVT algorithm. In adaptive sampling the entropy of information has shown to be better performing in reconstruction efforts [297]. An advantage of using CVT in this case is the automatic distribution of measurement points (e.g. UAVs). This can prevent uncontrolled clustering, which was an issue encountered in [297]. The entropy is defined in the information theory sense as,

$$H(\sigma_e^2) = -\sigma_e^2 \ln \sigma_e^2, \quad (4.34)$$

where σ_e^2 is a normalized quantity. The introduction of the entropy produces a trade-off between exploration (entropy) and exploitation (density). The density based CVT on one hand, seems to exploit more often the high concentration regions depending on how the UAVs are deployed. For instance, if deployed in a region with multiple constant hotspots, the UAVs take time to expand and cover the total area (see Fig. 4.22). When entropy is included, the UAVs recognize there are areas of uncertainty not explored and therefore add weight to the unexplored area. This entropy weight helps the CVT algorithm explore in a much quicker way, which is ideal for plumes that may meander back and forth. The estimation error as a result of including entropy dropped significantly given this scenario (see Fig. 4.23). The UAV position update law was a modification to Lloyds by adding entropy. The update is defined as,

$$\dot{p}_i = -k_p(p_i - c_{\psi,i}) - k_e(p_i - c_{H,i}), \quad (4.35)$$

where, k_p , and k_e , are positive constants used to adjust the weights of density and entropy. The centroids $c_{\psi,i}$ and $c_{H,i}$ for the i -th UAV are solved using,

$$m_{\psi,i} = \int_{v_i} \psi_1(q, t) dq, \quad c_{\psi,i} = \frac{1}{m_{\psi,i}} \int_{v_i} q \psi_1(q, t) dq, \quad (4.36)$$

$$m_{H,i} = \int_{v_i} H(\sigma_e)(q, t) dq, \quad c_{H,i} = \frac{1}{m_{H,i}} \int_{v_i} q H(\sigma_e)(q, t) dq. \quad (4.37)$$

To get the continuous mass flux measurement, a n_a point moving window average is used on the calculated source rate with the first DMD mode as an input,

$$\dot{\mathbb{M}}_k = \frac{k_{cf}}{n_a} \sum_{j=k-n_a}^k \int_{\Omega} \bar{\mathbf{u}} \cdot \hat{\mathbf{n}} \psi_{1,j} dA. \quad (4.38)$$

The variable k_{cf} is the conversion factor that converts to desirable units (e.g. SCFH or kg/s), $\bar{\mathbf{u}}$ is the average wind vector, and $\hat{\mathbf{n}}$ is the normal vector of the flux plane.

The moving average was implemented to help smooth out the spikey instantaneous flux measurements (see Fig. 4.25).

In some cases the meandering can shift the plume outside of Ω_p resulting in loss of plume data. To combat this, a simple adaptive algorithm is constructed that relates to the size and position of the domain Ω_p . The domain is defined by the plane position vector $\mathbf{r}_p$, the plane angle θ_p , the maximum height z'_{max} , and maximum width y'_{max} . Given the center of the domain c_Ω and the center of mass of density c_ψ , the flux plane location and domain size can be adjusted adaptively,

$$\begin{bmatrix} \dot{y}'_{max} \\ \dot{z}'_{max} \end{bmatrix} = \begin{bmatrix} -k_y(c_{\Omega,y} - c_{\psi,y}) \\ -k_z(c_{\Omega,z} - c_{\psi,z}) \end{bmatrix}, \quad (4.39)$$

$$\begin{bmatrix} \dot{r}_{p,x} \\ \dot{r}_{p,y} \end{bmatrix} = -k_y(c_{\Omega,y} - c_{\psi,y})R_{\theta_p} \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (4.40)$$

Under the assumption that the domain Ω_p encapsulates the plume, the perpendicular direction of the flux plane (w.r.t. the mean wind direction) is then used to center the flux plane in the plume.

4.5.3 Simulation

In this section we give two illustrative examples of using this multi-UAV method. The first example looks at exploration versus exploitation behavior of the CVT coverage control given a static map in 2D. The second, looks at implementing the CVT algorithm on a more complex 3D case with turbulence and meandering plumes using an extension of the small time-scale plume odor simulation (POSIM) [120].

4.5.4 2D static case

The first example is in the form of a static concentration map,

$$f(x, y) = \sin \frac{x}{2} \sin \frac{y}{2} \left(\exp \frac{-(x-5)^2}{25} + \exp \frac{-(y-5)^2}{25} \right), \quad (4.41)$$

where the UAVs are deployed within a uniform ball of radius 2 located at the origin. In Fig. 4.22, we show the exploration/exploitation behavior of the density based CVT (D-CVT) and the density/entropy based CVT (DE-CVT), where the position update gains are set to: $k_p = k_e = 2$ and $k_v = 1$. The benefit of adding entropy also improves estimation error over time (see Fig. 4.23).

4.5.5 3D flux plane case

The 3D flux plane case, represents a fugitive gas emission scenario. The scenario consists of a single source rate and $N = 10$ UAVs. The wind speed is relatively constant, such that $\bar{\mathbf{u}} = [-1, 1]^T$. The state of the plume, the wind field, and the

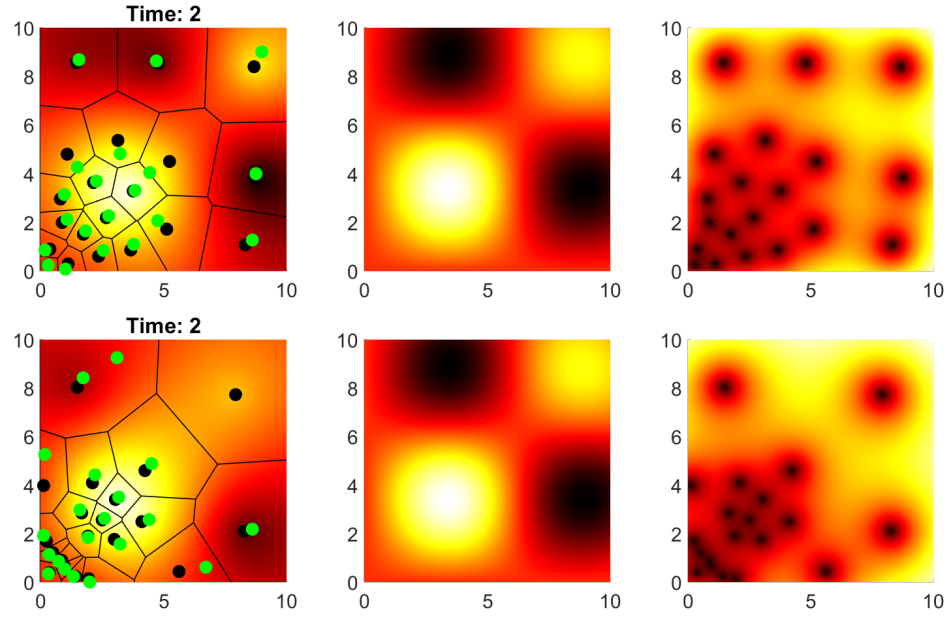


Figure 4.22: These plots show the current position / density reconstruction (left), the ground truth (middle), and the entropy (right) of the DE-CVT (top row) and D-CVT (bottom row) strategies for flux reconstruction.

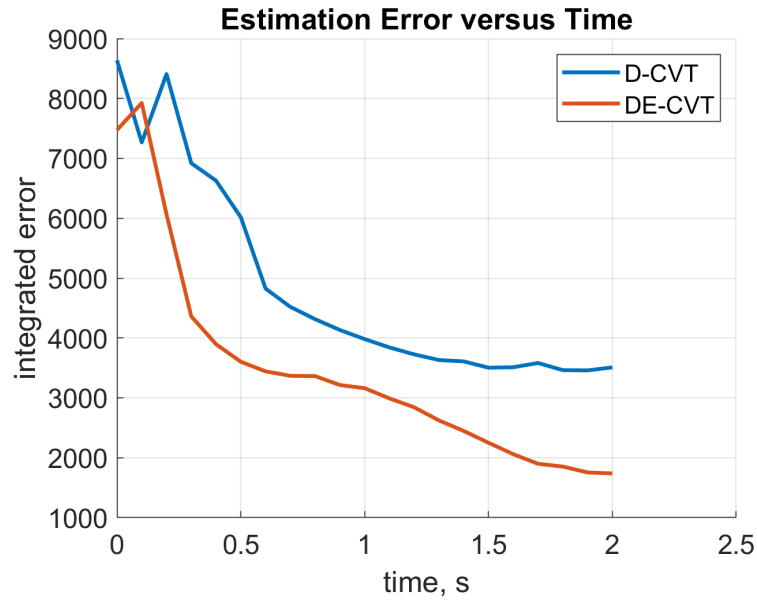


Figure 4.23: Given the 2D static case, the estimation error is reduced faster with the DE-CVT. This is desirable behavior since the field may change over time in practical applications.

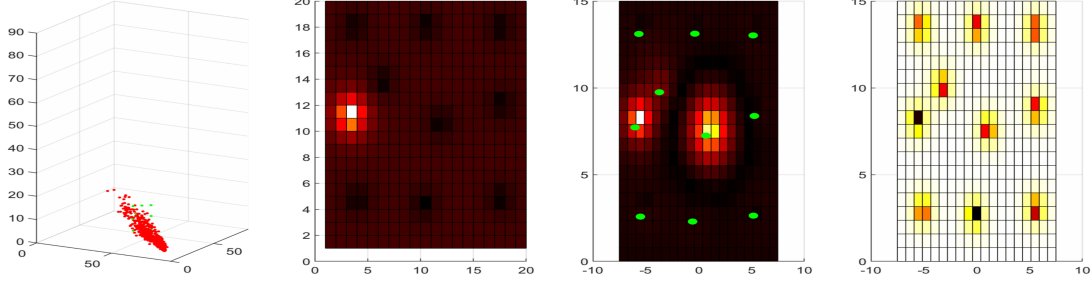


Figure 4.24: (left) 3D fugitive gas emission with deployed UAVs performing DE-CVT algorithm (middle-left) Ordinary kriging estimation given current concentration measurements for the UAV's at timestep k (middle-right) The first DMD mode capturing the low frequency behavior of the kriging results and (right) the entropy map computed from the kriging estimation variance.

UAV sensor is computed using POSIM. In this case we extend the 2D simulation POSIM DT to include a 3D windfield by powerlaw scaling [282] in the z -direction,

$$\mathbf{u} = \mathbf{u}_0 \left(\frac{z}{z_0} \right)^\alpha, \quad (4.42)$$

where $\alpha = 0.7$ for stable atmospheric conditions, and z_0 is the altitude the original 2D wind field is calculated (see Fig. 4.21). The 3D method is shown in Fig. 4.24, where the difference in the instantaneous emission estimate from kriging $\hat{\rho}$ and the DMD mode ψ_1 can be seen. The continuous source rate estimation performance shows good accuracy when measurements are more consistent and varies with the observations. This behavior can be seen in Fig. 4.25.

4.5.6 Conclusion

This case study introduced a new multi-UAV method to make continuous source rate estimations using an optimal coverage control CVT based on DMD mode based density and kriging variance entropy. The 2D simulation shows that adding entropy can provide better exploration behavior. The 3D simulation shows that the continuous source rate estimation using DE-CVT provides a good estimate of the source rate given that the UAV's are inside the plume making good measurements.

For future work, we will not only look at exploring DE-CVT's limitations on source rate estimation (as a function of meteorological conditions, number of UAV's, the source rate, location, etc.) but will seek to compare existing methodologies (single-UAV/sensor approaches such as mass balance [215], tracer correlation method [223; 229], near-field Gaussian inverse [189], OTM 33A [6]) against DE-CVT. Additionally, an evaluation metric for computing the efficiency of each method (on a per UAV basis) will be developed. The efficiency should contain variables such as system cost, survey speed, accuracy, precision, and overall complexity. The next following step is to extend to physical experiments.

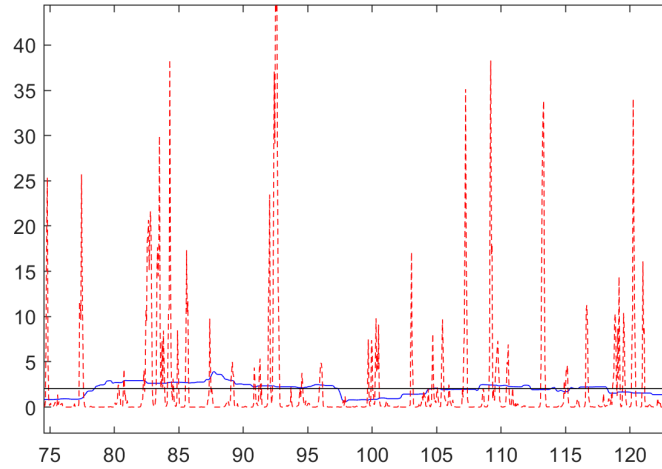


Figure 4.25: A snapshot of the instantaneous source rate (red-dotted), the continuous mass rate using 10 s moving average (blue), and true source rate (black).

4.6 A Case Study: Multi-sUAS-based Emission Quantification Using Mod-NGI

To illustrate the usefulness of the proposed method, a series of simulations are conducted comparing the NGI method and the mod-NGI method. It is assumed that the domain of interest contains the plume and the task is solely for quantification. It is also assumed that from the surface to 2 m is not accessible by the sUAS for safety reasons (see Fig. 4.26). Furthermore, let's consider that the sUAS is controllable at a high level and spatial position can be represented with single or double integrator dynamics.

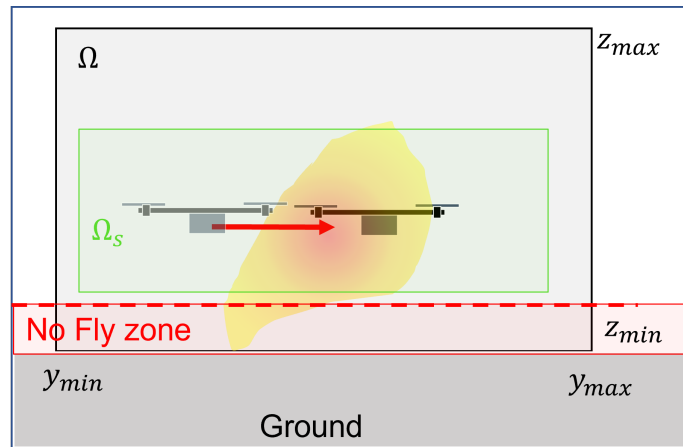


Figure 4.26: The simulation sample domain.

There are two general types of path planning approaches explored here: simple fixed path planning (no vehicle dynamics), and adaptive path planning (single integrator dynamics). In the fixed path planning approaches, a lawnmower pattern (also referred to as raster scanning), and an Archimedes spiral was explored. For the adaptive path planning we look at extremum seeking control (ESC) and centroidal Voronoi tessellations (CVT) (see Fig. 4.27).

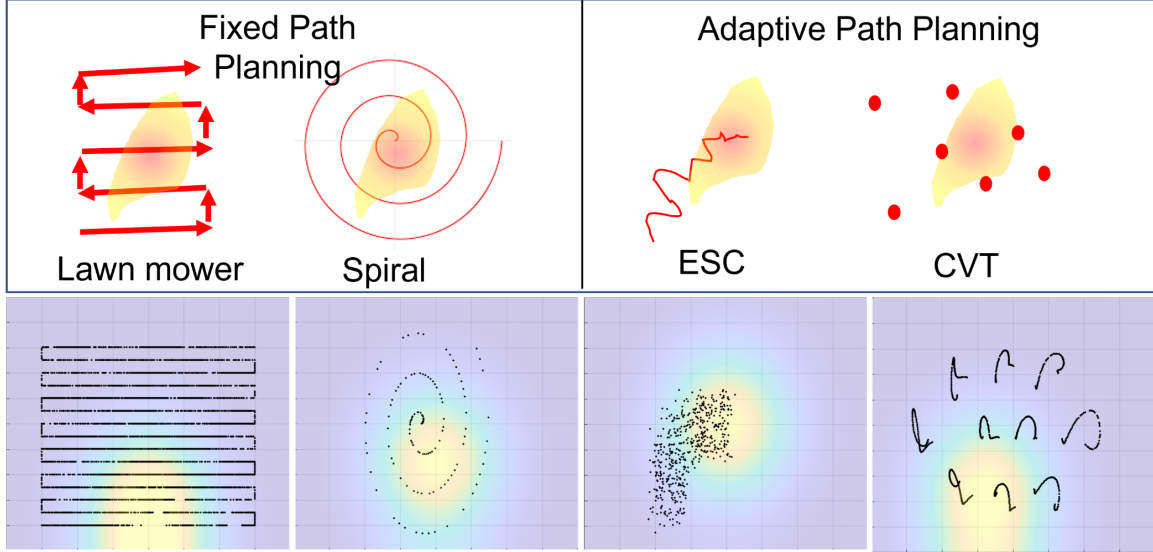


Figure 4.27: This image illustrates the different path planning strategies to sample the plume.

4.6.1 Fixed Path Planning

Path Planning

To test the fixed path planning strategies a series of simulations were conducted, varying sample time, spacing between transects (or spirals) and the plume width (σ). It was observed that sample time and sample spacing (between transects) played an important role in reducing the variability of the estimations. For instance, as the sample time increased (i.e. more sparse), the performance (accuracy and precision of the θ estimates) degraded. Likewise, the sample spacing reflected the ability to capture the center of the plume, often leading to poor performance. The width of the plume, in conjunction with sample time, reflected the ability to even detect the plume. Thus, the performance given different plume widths, σ , was explored for lawnmower and spiral paths (see Fig. 4.28). Conversely, it can be observed that larger plume widths will lead to better estimations and less of a dependence on which path planning approach to choose. For fugitive and small emissions this is not often the case, since at larger plume widths (further distances downwind) the concentration becomes quite small and perhaps even unmeasurable.

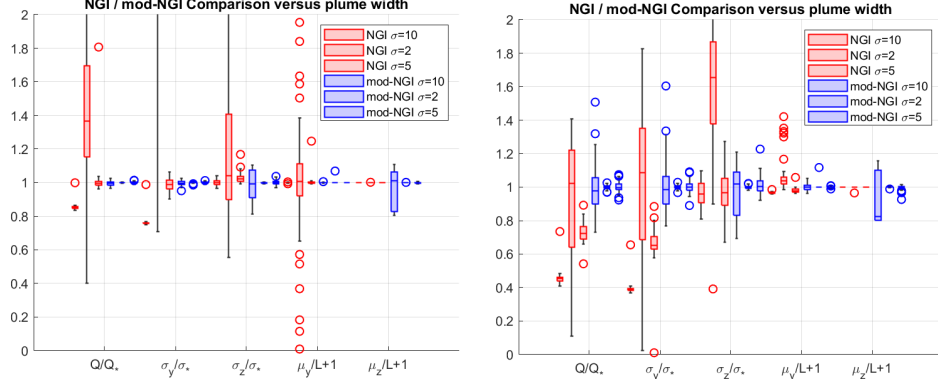


Figure 4.28: In these boxplots, the estimation performance results for static plumes different plume widths are explored for (a) lawn mower & (b) spiral paths

Given a dynamic plume setting, where the plume location has an oscillation e.g. meandering wind that can infringe on the measurement process. To test the effectiveness, the amplitude of the plume meander is increased from 0 m (no meander) to 10 m for several frequencies given the same lawnmower flight path (1m transect spacing, 2 Hz sampling rate, 1m/s horizontal velocity and 10% relative noise). The results are shown in Fig. 4.29, where in some cases the source rate is not as affected and in other cases the source rate is under- and even over-estimated.

Comparison with experimental flight data

The mod-NGI method was compared with the results in [11] using their processed survey data. There were two DJI Spreading Wings S1000+ octocopter sUAS. The first was equipped with a perfluoroalkoxy tubing that tethers a local inlet on the UAV to a ABB Micro-portable Greenhouse Gas Analyzer (MGGA) on the ground. It was also equipped with a wind sensor. The second was equipped with a lighter prototype MGGA (pMGGA). A two-dimensional stationary sonic anemometer was deployed near the boundary of the site for all surveys. The total data consisted of 22 flight surveys (7 for the first, and 15 for the second sUAS) was used to compare NGI and mod-NGI results (see Fig. 4.31). Though the mod-NGI method did not ‘outperform’ NGI, the results were still comparable.

4.6.2 Adaptive Path Planning

Single-sUAS Extremum Seeking Control (ESC)

In nonlinear control theory, adaptive control is a very desirable topic, particularly if the model of the system is unknown or not accessible. Here we consider a model-free strategy called extremum seeking control (ESC), which aims to optimize a set of parameters using a perturbation signal and filtering. One of the first stability analysis works on ESC came from [107]. The ESC has been used to solve many

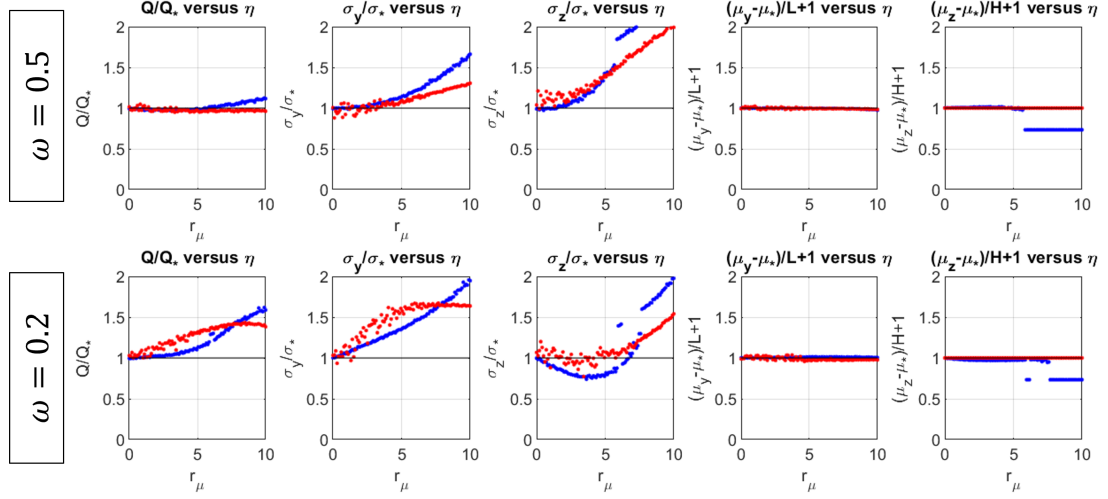


Figure 4.29: In these sets of plots, the performance of the NGI (red) and the mod-NGI (blue) at different amplitudes of meandering are shown and the different frequency responses of the algorithms are shown in each row. It can be observed that both algorithms do not perform well given dynamic (e.g. meandering) conditions. The most variability can be seen in the plume width estimates.

problems, such as electric railway [131], maximum power point tracking [298], and plasma impedance matching [299]. The principle of ESC relies on the minimization of a cost function given some dynamical system. Here we treat the cost function as the concentration of the plume and seek to find the center (see Fig. 4.27). Given several values around the center of the plume, a reasonable estimate of the source could be made. However, the slope of the measurement field may not be smooth. This may lead to slow convergence times and trapping in local minima. It can be observed in Fig. 4.30d that given a dynamic plume and slow convergence, mod-NGI estimations perform poorly with ESC given the frequency of meandering.

Multi-sUAS CVT

Taking into consideration all of the simulations thus far, they all utilize a single mobile sensor for estimation of the parameters. In this section we further explore adaptive path planning by introducing multi-sUAS coverage control strategy, namely, the centroidal Voronoi tessellations (CVT). The CVT strategy is brought into the mod-NGI approach to improve the robustness given a dynamic source (as seen in the previous sections). The CVT strategy looks to optimize the following cost function,

$$J(X) = \sum_{i=1}^N \int_{\mathcal{V}_i} \rho(\xi) |\xi - X_i|^2 d\xi, \text{ for } \xi \in \Omega. \quad (4.43)$$

Given the local positions of the sensors $X = (X_1, X_2, \dots, X_N)$, a domain of interest Ω can be divided into N polytopes $\mathcal{V} = (\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_N)$. The polytopes are determined

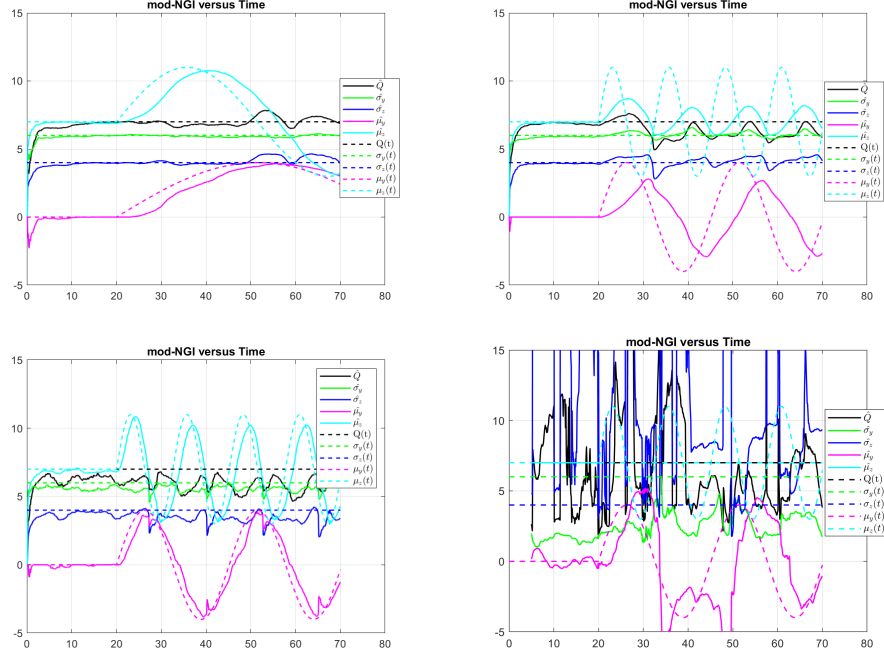


Figure 4.30: Given a moving source these six plots show the effect of the multi-UAS CVT-mod-NGI and single ESC-mod-NGI based algorithms. It can be observed in plots (a)-(b) that as the movement frequency of the source increases the ability of the CVT-mod-NGI algorithm to estimate and track the source diminishes. This behavior is much like the response of a time series signal given a low pass filter. Conversely, as the window of points used to compute the mod-NGI approach is reduced, much like we would expect, the ability to track and estimate the parameters increases. Although the performance in (c) is improved, the estimations become more noisy, and thus there is a trade-off between tracking performance and stable estimations. To illustrate why multi-sUAS systems may be better, an ESC-mod-NGI algorithm is shown in (d), where the ability to estimate parameters is not as robust to meandering and exhibits more difficulty with hyperparameter tuning.

using centroidal Voronoi tessellations [291]. The centroids, c_i , are computed using the density function $\rho(\xi)$ and positions are updated with Llyods algorithm [292; 293; 294], $\dot{X}_i = -k(X_i - c_i)$, where k is a proportional gain to be tuned. In this work we substitute the density function for, $\rho(X) = q_{mo}(X, \Theta)$. Here, $\Theta = [\hat{Q}, \sigma_y^c, \sigma_z^c, \mu_y^c, \mu_z^c]^T$, and the control parameters are related to the mod-NGI parameter estimations by,

$$\sigma^c = (1 + A_\sigma \sin(\omega_\sigma t))\hat{\sigma}, \quad (4.44)$$

$$\mu_y^c = A_\mu \cos(\omega_\mu t) + \hat{\mu}_y, \quad \mu_z^c = A_\mu \sin(\omega_\mu t) + \hat{\mu}_z. \quad (4.45)$$

The results of this approach are shown in Fig. 4.30 given $A_\mu = \omega_\mu = 1$ and $A_\sigma = \omega_\sigma = 0.5$. The control parameters are time varying to introduce a persistent excitation signal into the sampling for improving exploration and parameter estimation

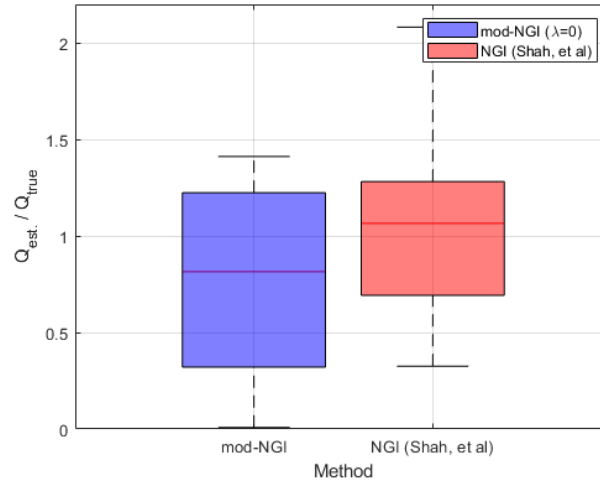


Figure 4.31: Comparisons between the results from Shah et al.’s NGI, our NGI, the mod-NGI, and the true controlled release flux using the controlled release survey data from [11].

convergence.

4.6.3 Conclusions

In this case study we introduced the mod-NGI method for estimating emissions from mobile sensing data. The mod-NGI performance is compared with NGI in static and dynamic plume settings. The methods utilize fixed (lawnmower and spiral) and adaptive (ESC and CVT) path planning based measurements. For single mobile sensor approaches, the proposed method shows improved estimation accuracy and precision given static plume settings, however, when exposed to a dynamic plumes the performance particularly degrades in the source rate, and plume width estimations in both approaches. Regularization based optimization was also explored and offered little improvements in certain instances and degradation in others. The mod-NGI method seen comparable estimation results to previously published NGI data using controlled release survey data. Lastly, the results of the multi-sUAS adaptive path planning approach, CVT-mod-NGI, show good tracking and estimation performance. Furthermore, the use of multi-sensor strategies in simulation, outperform single sensor ones given a dynamic plume. This provides some motivation to utilize multi-sensor strategies in physical experimentation. Future work will look utilizing MOABS/DT [49] simulation (based on the ARPA-E METEC facility) to test the short time-scale behavior of the mod-NGI and CVT-mod-NGI methods. Additionally, comparison against controlled release mass balance based approaches in [1].

4.7 A Case Study: Single and Multi-sUAS Emission Quantification Performance Assessment using MOABS/DT

Understanding performance limitations of a new or existing method, typically, requires many field experiments under different scenarios and weather conditions. The methods are then validated through multiple trial runs (given each configuration), ultimately, making it costly to pursue and difficult to compare against competing methods. This is also true, going from an early method design stage to something more practical and deployable. In this case study MOABS/DT is used to test single and multi-sUAS emission quantification methods and their performance under different meandering and terrain conditions.

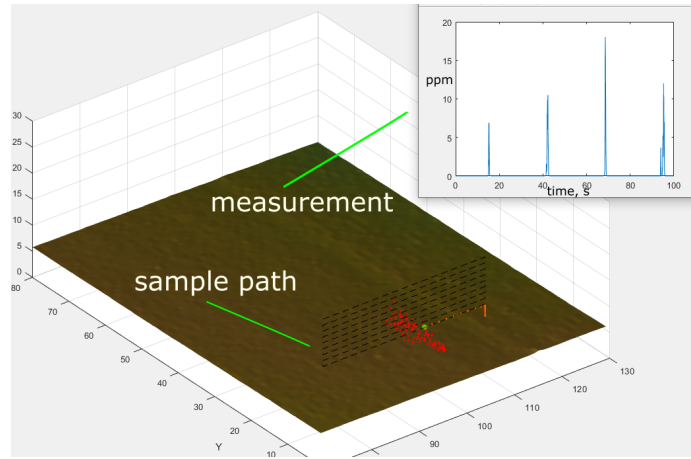


Figure 4.32: To validate the model the sUAS samples along a fixed path downwind of the source. The measured signal is shown in the top right.

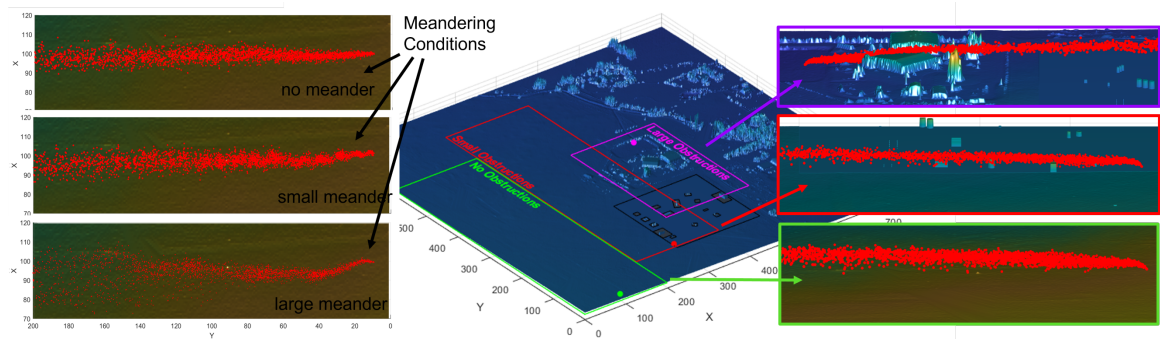


Figure 4.33: Subsections of the METEC location used for the case studies: No Obstacles (green), Small Obstacles (red), and Large Obstacles (Magenta). The colored points correspond to the plume starting locations.

Table 4.3: An overview of the simulation testing scenarios.

Test	Meandering	Wind Speed (m/s)	# Plumes Per Config.	Distances (From x=0)	# Plumes
Near-field & mid-field, No Obstacles	None, Small, Large	1	5	25, 45, 75, 100, 150, 200, 275, 310, 360, 415, 500	15
Near-field, Small Obstacles	None, Small, Large	1	5	25, 45, 75, 100, 150, 200	15
Mid-field, Small Obstacles	None, Small, Large	1	1	275, 310, 360, 415, 500	3
Near-field, Large Obstacles	None, Small, Large	1	5	20, 50, 70, 90, 110, 120	15

4.7.1 Simulation Setup

To understand the performance assessment of the four quantification methods, the simulation setup was broken down into three categories: No Obstacles, Small Obstacles, and Large Obstacles. The emission locations were chosen from the METEC site and surrounding area, which can be seen in Fig. 4.33. The terrain used for the METEC site is derived from LiDAR data. The terrain data is coupled with the MOABS/DT's collision algorithm, which allows for the exploration of the effects of plume-structure interaction with regard to methane source quantification. For each category, the same methane plumes were simulated, ranging from no meandering, small meandering, and large meandering. This determined how dynamic the plume was through time laterally and each of these plumes were independently simulated 5 times due to the stochasticity of the wind field. The wind conditions were chosen to be neutral with the mean wind speed near the ground being 1 m/s. The plumes were modeled with a constant small leak rate of 0.05225 g/s (≈ 10 SCFH).

After the plumes were fully simulated, the four methane quantification methods were tested on plumes at multiple downwind distances for each meandering type. Further details can be seen in Table 4.7 and in Fig. 4.33.

For the single-sUAS case, the sUAS moved in a fixed path consisting of way-points constrained to a flux plane that was 50 m wide and 10 m tall. An example of this path can be seen in Fig. 4.32. The sUAS was modeled as a point mass with double integrator dynamics. The position error between the sUAS and the target position (or waypoint) was passed through a PD-controller. The sUAS was given a maximum velocity of 2 m/s and maximum acceleration of 1 m/s². The sUAS took 225 s to collect the methane data along its path across the fully developed plume.

For the multi-sUAS strategies, the swarms were given 120 seconds to sample the fully developed plume where the swarms calculated the source rate in real time. The last 96 seconds of the flux estimates were averaged to retrieve a single flux estimate for the trial.

4.7.2 Results and Observations

From the plots shown in Fig. 4.34, both of the single sUAS strategies performed well. The NGI method was observed to have performed better than the MB method overall but had larger variability 15 m from the source for the No Obstacles category. A general trend that can be seen from Figs. 4.34, 4.35, and 4.36 is that, for the No Obstacle and Small Obstacle categories, the estimates get worse closer to the source and better downwind. The Single-sUAS methods are generally best in the ranges of ≈ 100 m - 400 m. The Multi-sUAS strategies stay accurate past 400 m with the exception of DE-CVT with 3 UAVs.

When comparing all of the strategies, the NGI method performed better than the MB method overall with both performing similarly in the 100 m - 400 m range. DE-CVT performed well for only a swarm size of 9 past 200 m. The Mod-NGI method performed well for both swarm sizes downwind but performed the best within 150

m of the source with a swarm size of 9. All methods displayed increased variability progressing from the no obstacles to large obstacles within the ranges tested.

The observed effect of swarm sizes in the multi-sUAS strategies indicate that as the number of sUAS increases, the accuracy of the emission estimate gets better. In the DE-CVT method, the emission is marginally overestimated near the source with a swarm size of 9, whereas, for a swarm size of 3, it performed poorly. In the mod-NGI case, the swarm size of 3 shows increased variability in the estimate, which indicates that there must exist an optimal swarm size for estimating the emission.

4.7.3 Conclusions

In this case study we utilize the MOABS/DT platform to assess the performance of the MB, NGI, DE-CVT, and Mod-NGI methane source quantification methods. The plume parameters were gauged via comparing methane signals in simulation to methane signals seen in practice. Results from this study show that all of the methods performed well with the single-UAS strategies and the DE-CVT method decreasing in accuracy and increasing in variability close to the source. The mod-NGI method performance had high variability near the source for a swarm size of 3 but performed better than all of the other strategies with a swarm size of 9 at all ranges.

From this study, we can infer, as a rule of thumb, that the NGI and MB methods work best ≈ 100 m - 400 m from the source. The DE-CVT works best for all distances beyond 200 m for higher swarm sizes. The variability of the quantification increased for all methods (given obstacle categories) except the mod-NGI method with a swarm size of nine.

An advantage of simulating these quantification methods is that it allows for the performances to be compared under the exact same environmental conditions. This eliminates any quantification variations due to changes in atmospheric conditions, allowing for a more accurate means of comparison. Furthermore, utilizing DT's for benchmarking methodology is a very low-cost and efficient way to explore performance compared to field experimentation.

Based on the results observed in this study, future work will: explore further the effects of swarm size on emission quantification, add more quantification methods to compare, and test strategies over various simulation scenarios.

4.8 Chapter Summary

In this Chapter we introduced the Digital Twin framework in the context of environmental sensing with sUAS. Two DT platforms were introduced, namely, the Plume Odor Simulation (POSIM) and the Methane Odor Abatement Simulator with Digital Twins (MOABS/DT). The case studies presented in this Chapter focused on: characterized of a controlled release ground-to-air emission using POSIM under different spatial interpolation strategies; investigating a localization method using behavior matching between physical and DT plume simulations; development of a continuous

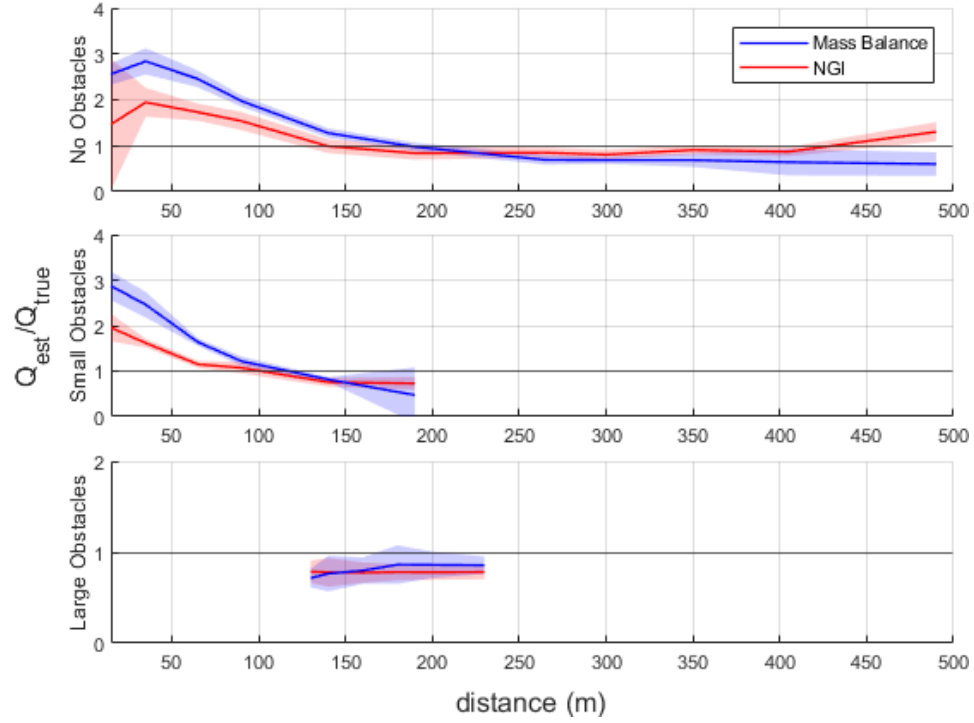


Figure 4.34: Results for the Mass Balance and NGI methods averaged over all of the meandering conditions and plume repetitions for each case study and distance. The solid lines indicate the means at the points while the vertical distance from the solid line to the edge of the shaded region represents the standard deviation at that point.

source rate estimation technique using a CVT-based multi-sUAS strategy with DMD and estimation entropy; development of the modified NGI method (mod-NGI); and a MOABS/DT-based method performance analysis between single-sUAS and multi-sUAS-based quantification strategies.

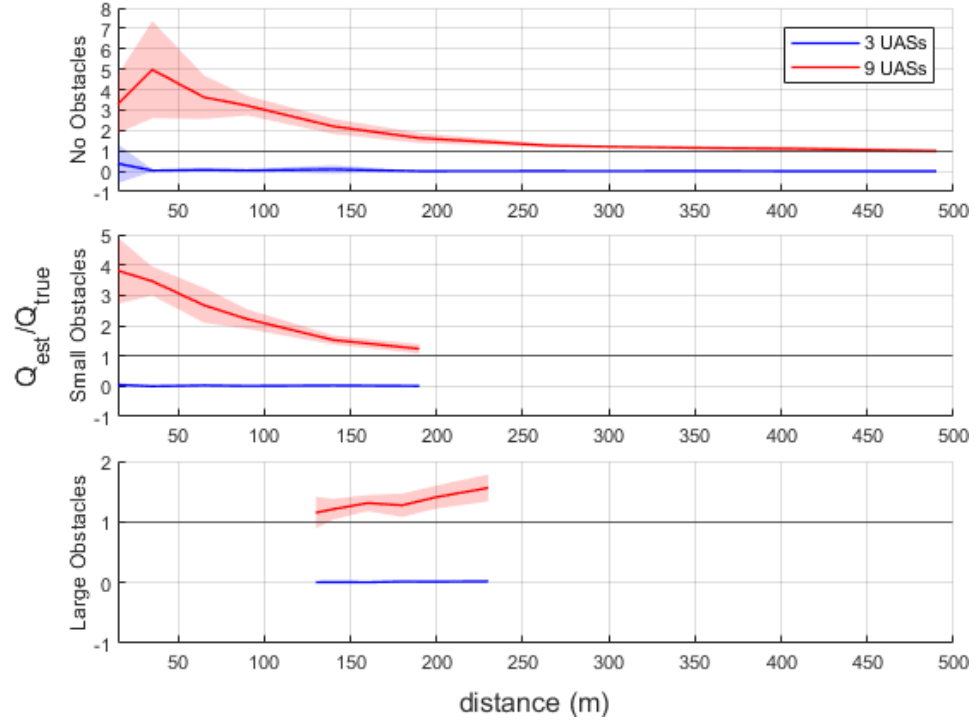


Figure 4.35: Results for the DE-CVT method for a swarm size of 3 UASs and 9 UASs averaged over all of the meandering conditions and plume repetitions for each case study and distance. The solid lines indicate the means at the points while the vertical distance from the solid line to the edge of the shaded region represents the standard deviation at that point.

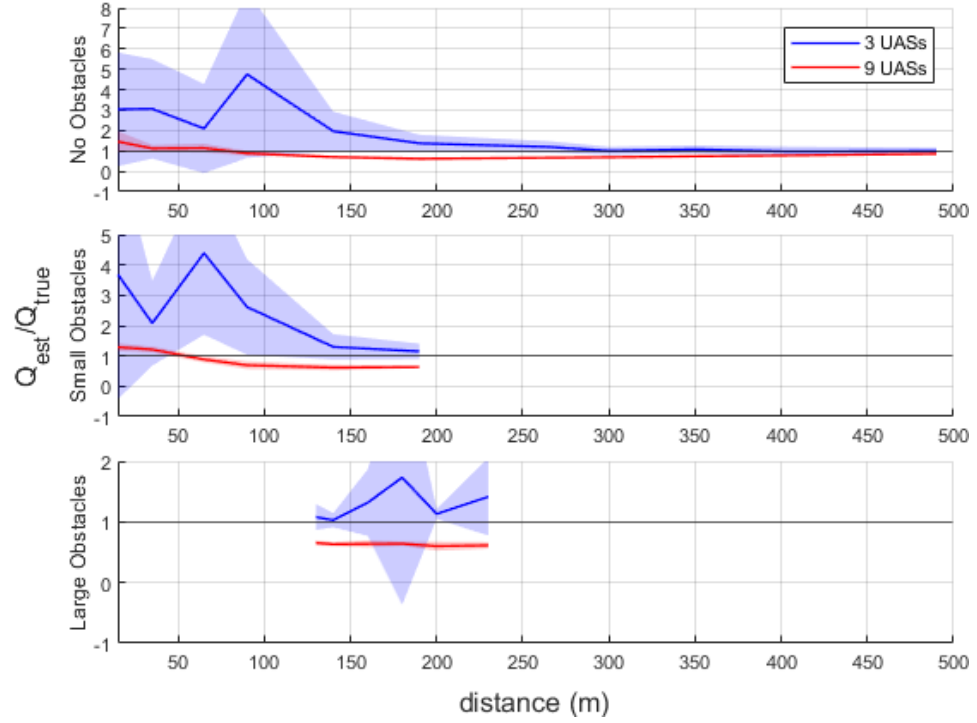


Figure 4.36: Results for the Mod-NGI method for a swarm size of 3 UASs and 9 UASs averaged over all of the meandering conditions and plume repetitions for each case study and distance. The solid lines indicate the means at the points while the vertical distance from the solid line to the edge of the shaded region represents the standard deviation at that point.

Chapter 5

Digital Twin Enabled Smart Sensor Placement and Steering

In the previous chapters, we discussed random search, quantification of methane emissions with sUAS and Digital Twins (DT). One of the major outstanding questions is how to best deploy mobile sensing within the field to extract the most information? i.e. increase the accuracy and precision of the emission estimates. Therefore the purpose of this chapter is to explore the idea of smart sensor placement and steering for improving emission quantification. One of the major challenges of estimating emissions, (e.g. inverse dispersion methodology) using methods outlined in Chapter 3, are such that the problem is often ill-posed. Lets consider the system,

$$\mathbf{u} = \mathbf{A}\mathbf{z} \equiv \int_a^b K(x, s)z(s)ds, \quad (5.1)$$

where $K(x, s)$ is a continuous kernel with partial derivative $\partial K/\partial x$. According to Hadamard, for a problem to be well-posed it must meet three conditions [300; 301]: (1) A solution, $\mathbf{u}$, must exist. (2) There exists a unique solution, $\mathbf{u}$, for a given input $\mathbf{z}$. and (3) The inverse solution, $\mathbf{z} = \mathbf{A}^{-1}\mathbf{u}$, is stable under small perturbations.

For the inverse dispersion problem, inherently apart of emission quantification, this is often not the case and requires optimization and regularization to solve. Therefore, in this chapter we will focus on the sensor placement and steering problem, overview empirical Gramians, and make the connection to Digital Twins in an optimization sense.

5.1 Sensor Placement and Steering Problem

The sensor placement problem suggests that there is a coherent structure within the data such that it can be represented with a sparse vector of measurement points. This problem is very similar to the compressed sensing problem where a transform basis $\Phi \in \mathbb{R}^{n \times n}$ (typically a Fourier or Wavelet basis), and a sparse vector $\mathbf{s} \in \mathbb{R}^n$

that can reconstruct the states or signal $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x} = \Phi \mathbf{s}. \quad (5.2)$$

This method is typically used for image and audio signal reconstruction where they are compressible. Despite having considerable success in real-world applications it still relies on access to the full high-dimensional measurements [302]. In practice, a single mobile sensor system is deployed to make measurements of the system and the field changes spatially and temporally. Thus, the question of where best to steer the mobile sensor to achieve the best performance arises. In the work by Ucinski et al. [303], optimal sensor placement was explored for estimation of distributed parameter systems. Given a possibly nonlinear system,

$$\frac{\partial y}{\partial t} = \mathcal{A}(x, t, y) + f(x, t), \quad (5.3)$$

where $\mathcal{A}$ is a spatial differential operator and f is the source term. The authors utilize time intervals (called stages) where they compute the parameters via maximum likelihood estimator. The forward problem is then computed (using numerical methods and the assumption the wind field is known). Given the forward problem the optimal trajectory can be computed given the Fisher Information Matrix (FIM), F , and D-optimality condition, $\Psi[F] = -\log \det(F)$,

$$F = \sum_{j=1}^N \int_0^{t_f} g(x^j(t), t) g^T(x^j(t), t) dt, \quad (5.4)$$

where

$$g(x, t) = \nabla_{\theta} y(x, t; \theta)|_{\theta=\theta^0}. \quad (5.5)$$

The concept is that the D-optimality condition-based trajectory will lead to better maximum likelihood estimates of the system parameters and thus reduce overall error. Other works have looked into this FIM way of thinking before, such as with [304], or even exploiting the gradient of the Cramer Rao Lower bound in [305].

In essence, the optimal trajectory depends on what the optimality metric or condition is. For instance, there are several other metrics using the FIM. In particular, there is the E-optimality $\Psi[F] = \lambda_{\max}(F^{-1})$ (or smallest eigenvalue), A-optimality $\Psi[F] = \text{tr}(F^{-1})$, or the sensitivity criterion $\Psi[F] = -\text{tr}(F)$. Other types of metrics have been explored in an effort to find better performance. For example, entropy has been used as a coverage control metric for deployment with the continuous source rate estimation method [281] described in Chapter 4, or in source seeking context with the Entrotaxis method [127]. Instead of looking for the observability of a system is another promising deployment metric utilizing control-theoretic techniques [281]. An example in the literature by Hinson et al. [306], where they used the empirical observability Gramian as a cost function to select the best sensor placement for gyroscopic sensing in a hawkmoth wing. The results reflected quite well to what was seen in nature.

However, in order to update the sensor position or decide on a trajectory the forward model needs to be computable, at least on a timescale that is relevant to the application. To compute (5.5), the output needs to be generated for the derivative with respect to each parameter. For high dimensional systems, this can be challenging.

Another method for solving such an issue, is the simultaneous perturbation and stochastic approximation (SPSA) algorithm [307]. When given a p -dimensional system, which requires $2p$ noisy measurements using finite differences with traditional stochastic approximation algorithms, the SPSA algorithm only requires $2q$, such that $q \ll p$. There have been some advancements using digital-twin-enabled smart control engineering with parallel SPSA [308].

5.2 Empirical Observability Gramians

In control theory, the observability Gramian is an important concept for designing state estimators, or performing model order reduction. However, this is typically limited to linear systems theory and not applicable to nonlinear systems. An extension to this concept is the empirical Gramian, which was developed for the use with nonlinear and parametric systems. Given the nonlinear system,

$$\dot{x} = f(t, x(t), u(t), \theta), \quad y(t) = h(t, x(t), u(t), \theta), \quad (5.6)$$

the empirical observability Gramian can be defined as,

$$\hat{W}_O = \frac{1}{|S_x|} \sum_{l=1}^{|S_x|} \frac{1}{d_l^2} \int_0^\infty \Psi^l(t) dt, \quad (5.7)$$

where $\Psi^l = (y^{li}(t) - \bar{y}^{li})^T (y^{li}(t) - \bar{y}^{li})$. The initial state configuration is given as, $x_0^{li} = d_l \epsilon^i + \bar{x}$, $u(t) = \bar{u}$, and $\bar{y} = 1/T \int_0^T y(t) dt$. The term $S_x = \{d_l \in \mathbb{R} : l = 1 \dots L, d_l \neq 0\}$. There are existing tools in MATLAB to compute W_o , such as *emgr* function (outlined in [309]).

If the system has process noise the trajectories can be computed using the forward Euler-Maruyama method for the Itô calculus, such that,

$$X_{t+1} = X_t + f(X_t, u_t) \Delta t + \sigma(X_t, u_t) Z_t \sqrt{\Delta t},$$

where $Z_t \sim \mathcal{N}(0, I)$ and X_0 is sampled from the initial condition distribution [12]. The expectation for the empirical observability gramian can be computed using Theorem 7 in [12],

$$E[W_O^\epsilon(t_1, x_0, u)] = \bar{W}_O^\epsilon + \hat{W}_O^\epsilon, \quad (5.8)$$

$$(\bar{W}_O^\epsilon)_{ij} = \frac{1}{4\epsilon^2} \int_0^{t_1} (E[y^{+i}] - E[y^{-i}])^T (E[y^{+j}] - E[y^{-j}]) dt, \quad (5.9)$$

$$(\hat{W}_O^\epsilon)_{ii} = \frac{1}{4\epsilon^2} \int_0^{t_1} \text{tr}(\text{Cov}[y^{+i}] - \text{Cov}[y^{-i}]) dt. \quad (5.10)$$

The stochastic empirical gramian shows a connection with the condition number of the Fisher information matrix, $F(t)$, and the covariance matrix, R , of the measurement system [12] (see Fig. 5.1). This relation is shown as,

$$F(t) \preceq \bar{\sigma}(R^{-1}) \frac{d}{dt} \lim_{\epsilon \rightarrow 0} W_O^\epsilon(t, x_0, u), \quad (5.11)$$

where $\bar{\sigma}(R^{-1})$ represents the maximum singular value of R^{-1} .

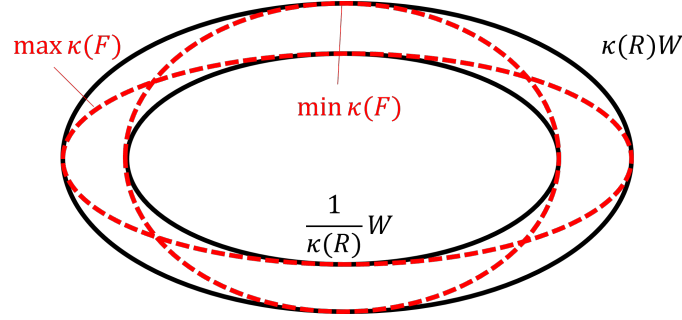


Figure 5.1: The minimum and maximum condition numbers of the Fisher information matrix is bounded by the empirical observability Gramian scaled with the condition number of the measurement noise covariance matrix (see [12] for more details).

Once the empirical observability gramian is obtained, the optimization can be undergone. Optimizations such as those with the objective functions, $J(\cdot)$ using the measures of unobservability [310], or with the $-tr(W^{-1})$, $\logdet(W)$, and $rank(W)$ [311]. Because the observability Gramian is closely related to the estimation uncertainty and the FIM, the eigenvalues of the observability Gramian directly control the Fisher information and inversely control the estimation covariance. The minimum eigenvalue, $\lambda_{min}(W)$, is a measure of the output energy for the least observable mode, while the maximum eigenvalue, $\lambda_{max}(W)$, is the measure of the output energy for the most observable mode. The potential cost functions and their significance are listed in Table 5.1.

5.3 Empirical Identifiability Gramian

With the concept of observability and empirical observability gramians, the output of the system is used to give information on how well the states can be recovered. If we are interested in understanding how the system parameters can be estimated from the output, then we would inquire about identifiability. This is constructed by augmenting the parameter state dynamics into the general states of the system,

$$\begin{pmatrix} \dot{x}(t) \\ \dot{\theta}(t) \end{pmatrix} = \begin{pmatrix} f(t, x(t), u(t), \theta(t)) \\ 0 \end{pmatrix}, \quad y(t) = g(t, x(t), u(t), \theta(t)), \quad (5.12)$$

Table 5.1: Observability metrics that can be used as a minimizing cost function for evaluating a system.

Metric	Significance
$\kappa(W)$	shape of estimation uncertainty ellipsoid [310]
$\sigma_{min}(W)^{-1}$	unobservability index [310]
$tr(W^{-1})$	minimize average estimation uncertainty [312]
$-\log(\det(W))$	minimize estimation uncertainty volume [312]
$-rank(W)$	max observable states [312]
$\lambda_{min}(W)^{-1}$	maximum estimation uncertainty [311]
$\lambda_{max}(W)^{-1}$	minimum estimation uncertainty [311]
$\det(W^{-1})$	volume of estimation uncertainty [311]

where the parameter-states initial value is used for testing parameter perturbations. The resulting augmented observability gramian has the block structure,

$$\hat{W}_O = \begin{pmatrix} W_O & W_M \\ W_M^T & W_P \end{pmatrix}, \quad (5.13)$$

where the mixed state and parameter block is $W_M = W_M^T$. The empirical Identifiability Gramian is then defined using the Schur-complement,

$$W_I = W_P - W_M^T W_O^{-1} W_M. \quad (5.14)$$

In some cases, the indentifiability can be approximated as $W_I \approx W_P$. However, similar to the Fisher information and SPSA case, to compute the empirical identifiability matrix the parameters need to be perturbed. Depending on how the forward model is computed, this may lead to a higher computational cost compared to the empirical observability Gramian.

5.4 Persistence of Excitation and Sufficient Richness

The term, persistence of excitation (PE), commonly refers to a exogenous signal that excites the states of the system. This is important to consider in mobile sensor actuator systems as sensor position is reflected in the output only, and is needed to observe the excited states, where they are observable. The formal definition for PE is given here [313],

$$\sum_{t=t_0+1}^{t_0+N} x(t)x(t)^T \geq \alpha I, \quad (5.15)$$

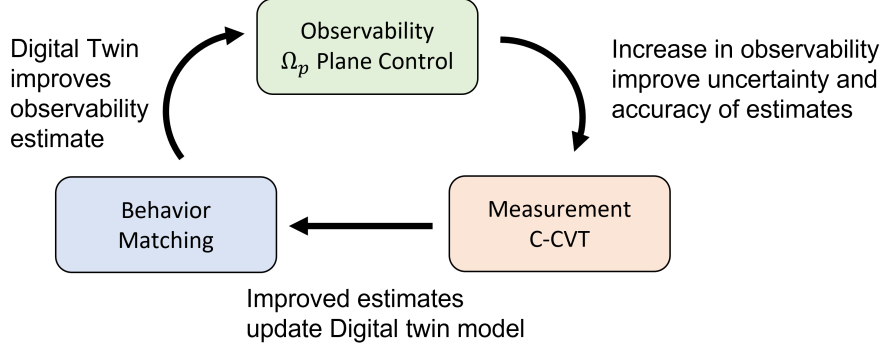


Figure 5.2: A diagram depicting the interconnection between the three major tasks of digital twin enabled smart sensor placement framework.

where $x(t) \in R^N$ is a sequence and $\alpha > 0$. The term, sufficient richness (SR), is used to describe the PE of a control signal in the system that is order m , given N steps.

$$\sum_{t=t_0+1}^{t_0+N} \begin{bmatrix} x(t+1) \\ x(t+2) \\ \vdots \\ x(t+m) \end{bmatrix} [x(t+1)^T, x(t+2)^T, \dots, x(t+m)^T] \geq \alpha I. \quad (5.16)$$

In the work by [314], the definition is given in integral form,

$$\int_{t_0} x(t)x^H(t)dt \geq \alpha I, \quad \forall t_0 \in R. \quad (5.17)$$

The definition for persistently exciting of order l of x is if $\underline{x}_l$ is persistently exciting,

$$\underline{x}_l = [x^{(0)T}, \dots, x^{(l-1)T}]^T. \quad (5.18)$$

The importance of defining PE and SR can be shown in the works like [315], such that, even a single sensor with motion can induce the PE condition and ensure the convergence of the spatial field estimation. Therefore, given multiple mobile sensors, the PE condition requirement should still hold, even though the spatial fields in [315] are linearly expandable with respect to the parameter to be estimated.

5.5 A Case Study: DT-enabled Sensor Placement and Steering with the Empirical Observability Gramian

The introduction of DTs into the sensor placement and steering problem provides a pathway for computing the forward problem efficiently and quite possibly in near real-time. One way to approach the sensor placement problem would be to follow the

work in [303], where the sensor deployment would be updated via one of the Fisher information matrix based metrics (e.g. D-optimality). While this seems very enticing to pursue, the Fisher information matrix requires the computation of the partial derivative of the output with respect to the parameters, leading to more forward model solves to achieve. In the context of DTs, this would require many active DT assets simulating perturbations within the parameter space. However, as we seen in previous sections the empirical observability Gramian can be computed with just the output of the system.

One of the advantages of having an actively behavior matching DT of the system is the ability to converge overtime, assuming the physical system is observable and the information is sufficiently rich, the sensor placement would have need to have persistence of excitation condition to promote parameter estimation convergence [315]. To first expand on this concept, let us consider that the empirical observability Gramian can be applied to find an optimal trajectory, reducing uncertainty and improving parameter estimations.

5.5.1 Problem definition

Let us define a domain $\Omega \subseteq \mathbb{R}^3$ that contains a single concentration source and is perturbed by an unknown wind field $\mathbf{u}$. Let there be N point measurements using mobile sensors (i.e. sUAS) deployed in Ω_p using CVT-based coverage control. The CVT plane, $\Omega_p \subseteq \Omega$ is defined by the position vector $r(t)$, the width w , the height h , and the normal vector heading φ , such that $\Omega_p \perp \underline{\Omega}$, where $\underline{\Omega}$ is the floor of Ω . The position of the i -th sensor is given as $\mathbf{z}_i(t)$ and is used to measure the concentration and wind fields. Let the concentration field is described by,

$$\partial_t \rho = \nabla \cdot (D \nabla \rho) - \underbrace{\mathbf{u} \cdot (\nabla \rho)}_{\text{coupled disturbance}} + \underbrace{q \delta(\mathbf{x} - \mathbf{x}_s)}_{\text{control}}, \quad (5.19)$$

where the advection component is thought of as the coupled disturbance and the source component the control signal. Then, the concentration states can be written as,

$$\dot{\rho} = f(\rho, \mu, \mathbf{u}|D), \quad (5.20)$$

where $\mu = q \delta(\mathbf{x} - \mathbf{x}_s)$. Let us assume that the wind field disturbance can be described by the viscous Burgers equation,

$$\partial_t \bar{u}_i = -\bar{\mathbf{u}} \cdot (\nabla \bar{u}_i) + \mathbf{K} \cdot \nabla^2 \bar{u}_i, \quad (5.21)$$

where $(\bar{u}_i)$ is the average wind in the i -th direction of $\mathbf{u} = [u_1, u_2, u_3]^T$, and $\mathbf{K} = [K_1, K_2, K_3]^T$ are the viscosity terms (see more details in Ch. 4).

Lets assume the mobile sensors can be modelled as single integrator systems and have negligible affect on the physical measurement (i.e. no sensor noise). If the domain Ω is spatially discretized to support a high dimensional number of states, the individual persistence of excitation of any $\rho_{ijk}(t)$ states can be explored. Considering

the definition in (5.18), it is easy to argue that the areas with persistence of excitation are around areas with concentrations that change or are non-zero (relative to background level). This because the control (i.e. source) is acting as the sufficiently rich signal that interacts with its neighbors according to (5.19) and (5.21). Therefore, sampling around areas where the states are excited (given the diffusive behavior of the physical system), are of utmost importance for parameter estimation (i.e. the sensor placement problem). Thus, the problem is (1) how to choose the best place to measure, and (2) how to deploy the sUAS to measure, reduce uncertainty and increase parameter estimation accuracy.

5.5.2 Framework

To solve the problems listed above, we have to define several smaller tasks. The first task (T1), would be to find the general area where the plume is likely to be (i.e. sufficiently rich). The approach taken here, is to control the location of the domain Ω_p with $r(t)$ and φ , (referred to as a plane control task) using the empirical observability Gramian and digital twins (MOABS/DT). By controlling the domain Ω_p , the sUAS can be controlled in a stable and bounded way. It is assumed that the initial deployment of the sUAS into the plane will be undergone before the task is started. Additionally, the plane controller will actuate the plane at slower rate than the dynamics of the sUAS will allow. This is to ensure that the plane does not outpace the sUAS and potentially result in a collision. It is further assumed that the plane control will be subject to constraints, such as restricted areas (e.g. buildings or equipment) within the domain $\underline{\Omega} \subseteq \Omega$, which indicate the points on the floor of Ω , that are not restricted.

The second task (T2) would aim to solve the deployment of the sUAS inside the control plane Ω_p . In previous works, the CVT method has been used to determine the desired location of the sensors, this includes the DE-CVT and mod-NGI-CVT methods. However, the sensor locations are relatively slow moving within the domain. This slow moving (sometimes fixed) behavior may limit the sensor from reaching the states that are excited (or have concentration at some time t). Therefore, there needs to be continuous movement within the plane to increase the chances of measuring an excited state. Furthermore, in [316], the reconstruction of spatial temporal field was explored using radial basis functions (RBF). One of the major assumptions to the Theorems proposed (leading to parameter convergence), is that the sensor passes through the center of the RBFs used to estimate the spatial-temporal field (i.e. the point of maximum concentration for Gaussian). As a result, the concept of phase portraits is used to construct a velocity field with the stable points centered inside the control plane Ω_p to induce the PE condition. The mod-NGI-CVT based desired positions can be included in the computation to result in new objective function for obtaining a continuously moving coverage control method.

The third task (T3) is estimate the parameters of the system. This comes in three subtasks: parameter estimation, localization, and behavior matching. For parameter

Ω_p Plane Controller

Given the current states and parameters of the digital twin, ρ_{DT} and θ_{DT} , a function mapping $h_{DT}(\cdot)$ is needed to map the Lagrangian states to that of an Eulerian ones $\rho_{DT} \mapsto \hat{\rho}_{DT} \in \underline{\Omega}$. This is achieved by uniformly discretizing the domain $\underline{\Omega}$ and integrating along the e_3 direction (see Fig. 5.3). The resulting map of $\hat{\rho}_{DT}$ is then used to compute the empirical observability Gramian, $\hat{W}_O$. Given the nonlinear system,

$$\begin{aligned}\dot{\rho}_{DT} &= f(\rho_{DT}, \mathbf{x}, \mu, \mathbf{u}, \hat{\theta}), \\ \mathbf{y} &= h_{DT}(\rho_{DT}, \mathbf{r}), \\ \dot{\mathbf{r}} &= -k_r(\mathbf{r} - \mathbf{r}_d), \\ \dot{\varphi} &= -k_\varphi \Delta \varphi_d,\end{aligned}\tag{5.22}$$

where $\mathbf{r}_d$ and φ_d are the desired plane location and heading, and $k_r > 0$ and $k_\varphi > 0$ are proportional constants. To avoid computing $\hat{W}_O$ across the whole domain, only points near the current plane location are considered. Furthermore, since $\hat{\rho}_{DT} \subseteq \underline{\Omega}$ The empirical observability Gramian can be computed at some nominal point $\mathbf{r}_0$ using sample trajectories in the e_1 and e_2 directions, defined as,

$$\hat{W}_O = \frac{1}{|S_x|} \sum_{l=1}^{|S_x|} \frac{1}{d_l^2} \int_0^\infty \Psi^l(t) dt,\tag{5.23}$$

where $\Psi^l = (y^{li}(t) - \bar{y}^{li})^T (y^{li}(t) - \bar{y}^{li})$. The initial state configuration is given as, $r_0^{li} = d_l e_i + \bar{r}$, $u(t) = \bar{u}$, and $\bar{y} = 1/T \int_0^T y(t) dt$. The term $S_x = \{d_l \in \mathbb{R} : l = 1 \dots L, d_l \neq 0\}$. There are existing tools in MATLAB to compute W_o , such as *emgr* function (outlined in [309]). Given the estimate of the empirical observability Gramian in the perturbation directions, an objective function $J(W_o)$ can be chosen to evaluate each (see Table 5.1). For each direction and for each candidate objective function, the step will be evaluated and used to compute a plane control vector.

CVT Continuous Coverage Controller

Given local coordinates of the domain Ω_p , the positions of the sensors $\mathbf{z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N]$ can be divided into N polytopes $\mathcal{V} = (\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_N)$. The polytopes are determined using centroidal voronoi tessellations (CVT)[291]. The cost function that controls the CVT desired positions is,

$$\mathcal{H}(\mathbf{z}, t) = \sum_{i=1}^N \int_{\mathcal{V}_i} \mathcal{M}(q, t) |\mathbf{z}_i - q|^2 dq, \quad \text{for } q \in \Omega_p,\tag{5.24}$$

where $\mathcal{M}(\cdot)$ is a... The partial derivative can be used with the mass, m_i , and center of mass, c_i , to identify the critical point,

$$m_i = \int_{\mathcal{V}_i} \mathcal{M}(q, t) dq, \quad c_{m_i} = \frac{1}{m_i} \int_{\mathcal{V}_i} q \mathcal{M}(q, t) dq,\tag{5.25}$$

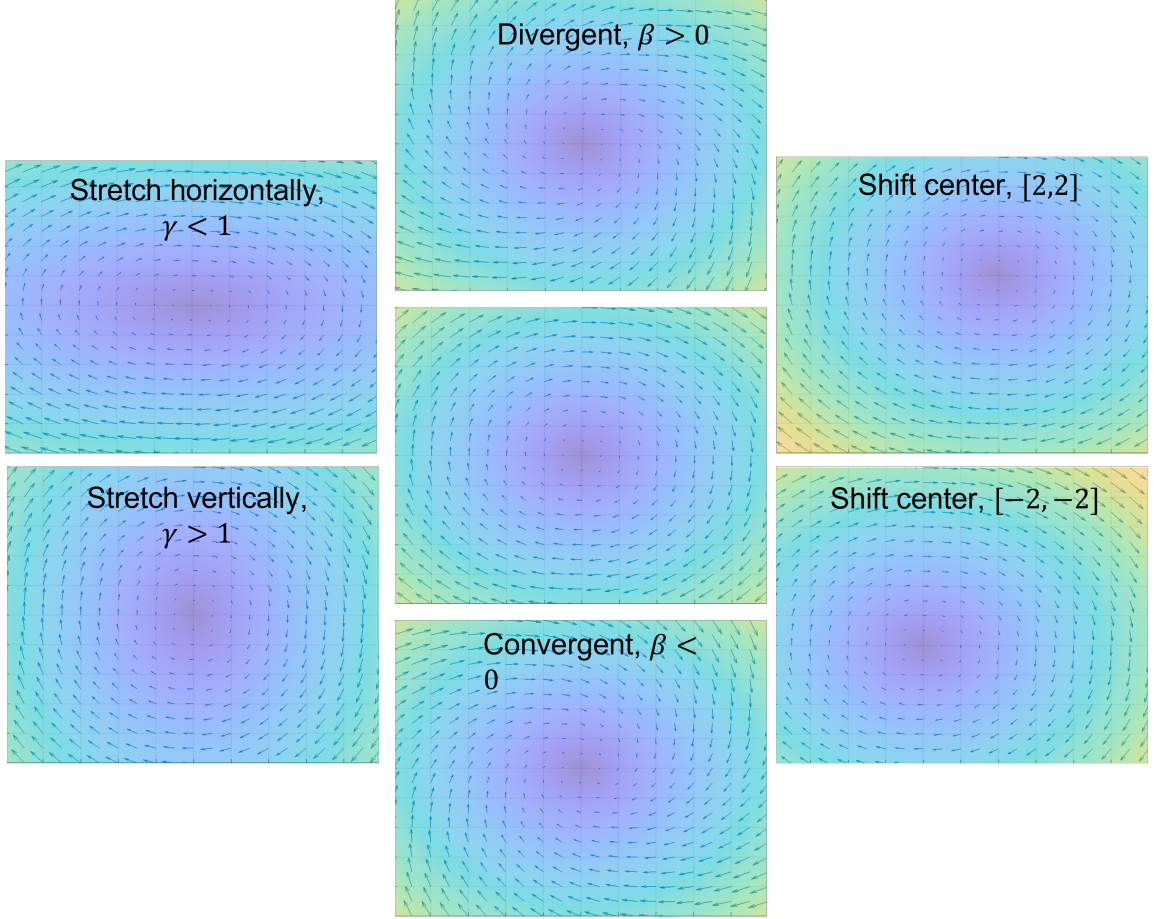


Figure 5.4: The phase portrait for different parameters (outlined in 5.28) can be seen here, where the velocity field is shown to divergent/convergent, stretched, and shifted.

$$\frac{\partial \mathcal{H}}{\partial \mathbf{z}_i} = 2m_i(\mathbf{z}_i - c_{m_i})^T. \quad (5.26)$$

This means that $c_i = \mathbf{z}_i$ for all $i = 1, 2, \dots, N$ is a minimizer of the CVT. Treating the relation as a gradient, Llyod's algorithm can be used,

$$(\dot{\mathbf{z}}_i)_c = -k_c(\mathbf{z}_i - c_{m_i}). \quad (5.27)$$

The general algorithm to compute the locations of actuators by computing the CVT [292; 293; 294] using Lloyd's method. To make the coverage control algorithm continuous, a velocity field can be applied using the following,

$$(\dot{\mathbf{z}}_i)_v = \lambda \frac{A}{|A|} \mathbf{z}_i, \quad A = \begin{bmatrix} \beta & 1 \\ \gamma & \beta \end{bmatrix}, \quad (5.28)$$

where β , γ , and λ are constants. When $\beta > 0$, the system is diverging, $(\mathbf{z}_i)_v \rightarrow \infty$, when $\beta < 0$, the system is converging $(\mathbf{z}_i)_v \rightarrow 0$, and when $\beta = 0$, the system is

critically stable. When $\gamma < -1$ the velocity field is stretched vertically, and when $0 \leq \gamma < -1$ the velocity field is stretched horizontally. Because the system is linear, the dynamics can be easily shifted (see Fig. 5.4). Additionally, by normalizing the matrix A and having a gain λ the resulting velocity field can be scaled to appropriate magnitudes. If the distance gets too far from the stable point, the magnitude of $(\dot{\mathbf{z}}_i)_v$ can be limited, such that,

$$(\dot{\mathbf{z}}_i)_v = v_{max} \frac{(\dot{\mathbf{z}}_i)_v}{|(\dot{\mathbf{z}}_i)_v|}, \quad \text{if } |(\dot{\mathbf{z}}_i)_v| > v_{max}. \quad (5.29)$$

To combine the two forms, an time step approximation of the trajectory (in the direction of the desired velocity field) is used as a desired location with weight k_v , resulting in a new dynamic law,

$$\dot{\mathbf{z}}_i = (\dot{\mathbf{z}}_i)_c - k_v[\mathbf{z}_i - ((\dot{\mathbf{z}}_i)_v \Delta t + \mathbf{z}_i)], \quad k_v > 0. \quad (5.30)$$

If we consider that $k = k_c = k_v$,

$$\dot{\mathbf{z}}_i = -k(\mathbf{z}_i - c_{m_i} - c_{v_i}), \quad k > 0, \quad (5.31)$$

where $c_{v_i} = (\dot{\mathbf{z}}_i)_v \Delta t$. If we consider that the density function $\mathcal{M}(q, t)$ be a radial basis function (such as the Gaussian Plume Model used in the NGI and mod-NGI) centered at the stable point of $(\dot{\mathbf{z}})_v$ dynamics, the orbiting trajectory path would be circular as well.

5.5.4 Simulation Setup

The simulation was initialized with the true source and DT located at $\mathbf{x}_s = [120, 140, 0]$. The wind field was initialized with the mean wind vector as $\bar{\mathbf{u}} = [-1, 0]$. Utilizing the MATLAB random number generator to control the random seed, each simulation was initialized over two stages. The first stage, lets the plumes develop past the Ω_p control plane as well as stabilize the sensor positions within Ω_p . Once this stage has been past, the CVT continuous coverage controller is enabled. At the second stage, the desired plane location is set to 40 meters above the current location, which is designed to cross the path of the plume 30 meters downwind (see the diagram in Fig. 5.5). When any of the sensors detects the presence of the plume (e.g. measurements significantly above background), the observability-based sensor placement and steering approach is enabled. For this experiment, the initial perturbation step size $d_l = 1$ was chosen based on the resolution of the projected DT plume $\hat{\rho}_{DT}$ onto $\underline{\Omega}$. To keep the computational cost to a minimum, we only considered a one step trajectory for each of the basis direction (e_1, e_2) . For each cost function evaluation, the simulation was run for 300 seconds. During the simulation, the emission is being quantified using the mod-NGI method. In the event that the empirical observability Gramian is singular, a chemotaxis-like approach is undertaken, such that, the plane control vector $\mathbf{r}(t)$ is perturbed in the direction of the neighboring cell with the highest concentration (with respect to $\hat{\rho}_{DT}$).

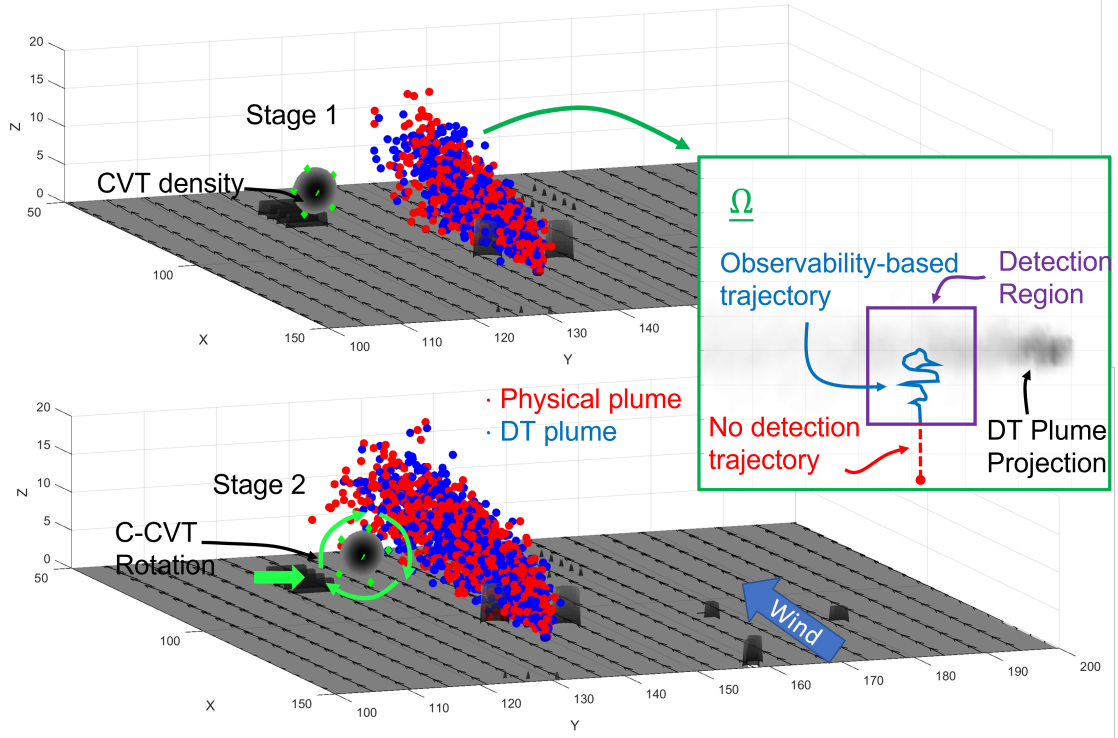


Figure 5.5: The MOABS/DT simulation is first initialized in (Stage 1) by letting the physical and DT plumes develop past the point of measurements, then (Stage 2) initializes the continuous CVT and commands the Ω_p plane to cross the plume. Once a detection is made the sensor placement and steering is activated and begins using the DT projection onto $\underline{\Omega}$ for the empirical observability Gramian computation.

5.5.5 Results

Given the first task of DT-enabled sensor placement and steering framework, the empirical observability Gramian-based cost functions were utilized to create trajectories for the Ω_p plane to estimate the emission rate. The individual trajectories are shown in Fig. 5.6 and it can be observed that several pairs of cost functions have almost identical trajectories. This is likely because each of the cost functions are trying to reduce similar forms of the estimation uncertainty and there are only two basis directions that the trajectory can take. Furthermore, there are only two pairs of trajectories that turned out to be identical, which consist of $\sigma_{min}(W)^{-1}$ and $\lambda_{min}(W)^{-1}$, and $-\log(\det(W))$ and $\det(W^{-1})$, respectfully. It turns out that in this case, $\det(W)^{-1} = \det(W^{-1})$, which means they only differ from a logarithmic scaling. Furthermore, the minimum eigenvalues and minimum singular values turn out to be identical. It can also be observed that the trajectory behavior of $\sigma_{min}(W)^{-1}$, $\lambda_{min}(W)^{-1}$, and $tr(W^{-1})$ tend to travel in the cross wind direction, whereas, $\kappa(W)$ and $\lambda_{max}(W)^{-1}$ tend to travel in the direction of the wind. If we take a look at the linear combination of the condition number and the minimum eigenvalue inverse, $\kappa(W) + \lambda_{min}(W)^{-1}$, we can see both behaviors in the trajectory, which is unique to the weighting chosen (here it is 25). Depending on the desired behavior (given the task at hand) the weight on $\lambda_{min}(W)^{-1}$ can be increased or decreased to see an effect in the trajectory behavior. This type of steering deserves some attention and will be a topic of future research.

Analyzing the continuous emission quantification estimates (see Fig. 5.7), shows that the cost functions with the trajectory that yielded the best results were the unobservability index $\sigma_{min}(W)^{-1}$, $\log(\det(W))$, $\lambda_{min}(W)^{-1}$ and $\det(W^{-1})$. Conversely, we see that the worst trajectory was the $rank(W)$ cost function. In fact, during the simulation if there was no general direction that would lead to an increase in the rank (i.e. the rank condition stayed the same, regardless of the trajectory direction) the sensor placement and steering strategy would default to chemotaxis.

5.6 Chapter Summary

In this Chapter we introduced the concepts of the sensor placement and steering problem, the empirical observability / identifiability Gramians, and persistence of excitation. We then outlined the DT-enabled sensor placement and steering problem framework and proposed a simulation case study around the first task (e.g observability-based sensor placement). Several different observability-based cost functions were explored in the simulation case study and it was found that given a sufficiently matched DT (capable of representing the physical system well), the empirical observability Gramian can be used to steer the Ω_p plane to reduce the estimation error. Additionally, it was observed that some cost functions promoted crosswind trajectory behavior while others promote trajectories along the direction of the wind. It was also observed that the combination of these two different types of metrics can be used to generate a

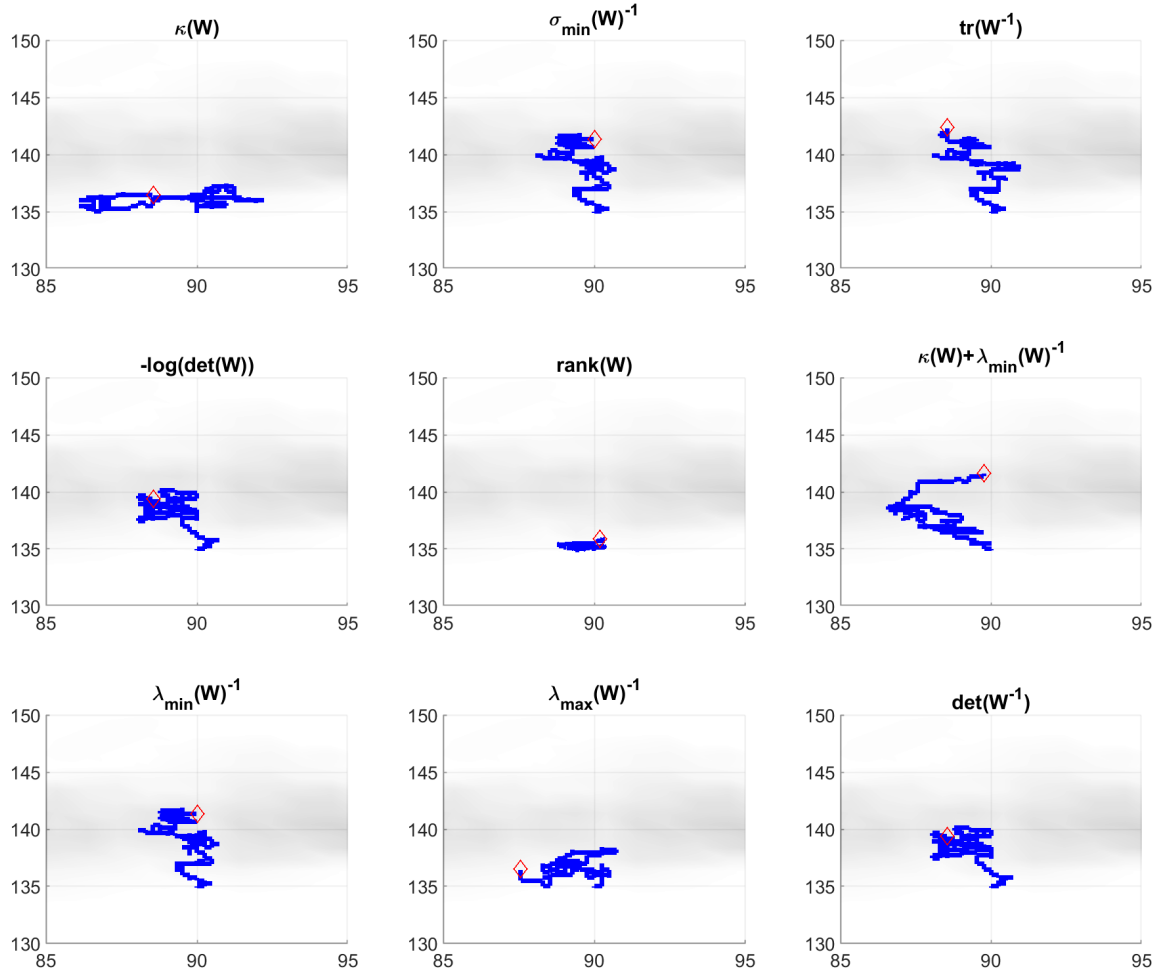


Figure 5.6: The sensor placement and steering trajectory for each cost function are shown here.

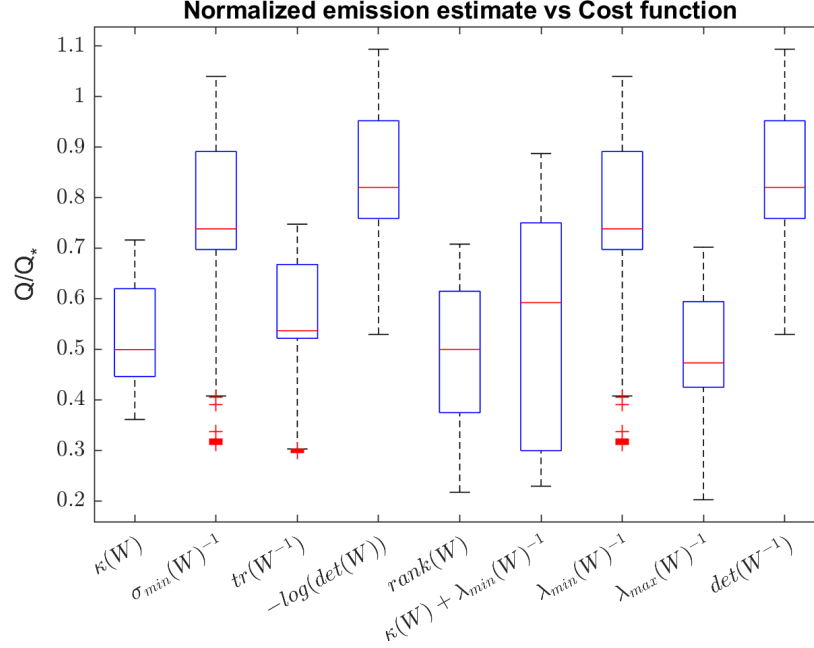


Figure 5.7: The emission quantification using the mod-NGI method is shown against different cost functions. The best performing metrics have smaller uncertainty and better accuracy, such as with $\det(W^{-1})$.

different trajectory. The estimation accuracy and variability was observed to be better when utilizing the determinant-based and the inverse of the minimum eigenvalue (or singular value - also referred to as the unobservability index). These metrics tend to focus on reducing the volume of the ellipsoid and maximum estimation uncertainties, respectfully.

Future shall look into testing task T2 and T3 within the DT-enabled sensor placement and steering framework, through solving the behavior matching problem for localization. Another interesting direction to exploring the identifiability Gramian, which the plume location parameters can be augmented into the states of interest, and may be applied as a behavior matching strategy (for parameters not obtained through other methods, such as with mod-NGI).

Chapter 6

Conclusions and Future Work

In this dissertation, several chapters were presented to build the dissertation topic, “DT-enabled Collective Sensing and Steering for Source Determination Problems”. The chapters contributed key pieces and developments that make steps towards improving emission quantification strategies and methane mitigation goals. In Chapter 2, the concepts of collective motion and the science of search were explored, which includes random search with single searcher undergoing Lévy flights, random search with loose coupled multi-agent searchers (including human intervention), swarm-based search using physics, as well as model-free optimization. We seen in the collective foraging with human intervention case, how loose coupling impacts foraging performance. We shown that source seeking and model-free optimization can be improved through fractional order derivatives and fractional Gaussian noise, respectively. In Chapter 3, the leak detection and quantification (LDAQ) sensors and methodology with sUAS was explored as well as several case studies focused on leak detection and localization, controlled release tests and real world emission quantification. In Chapter 4, environmental sensing with DTs were developed and explored with several case studies utilizing POSIM and MOABS/DT. The case studies focused on single-sUAS/multi-sUAS-based quantification, behavior matching localization with mesh sensors, as well as method performance assessment. The developments and studies are aimed towards making smarter environmental sensing possible. In Chapter 5, the DT enabled smart sensor placement and steering framework was explored, utilizing the empirical observability Gramian. The simulation explores the use of DTs within the sensor placement problem with an application towards the source determination problem. The implications of developing smarter emission quantification systems can help reduce overall emissions by early detection and repair as well as reduce the global warming and climate changing effects of methane emissions in the near-term.

The future challenges and opportunities for the DT-enabled collective sensing and steering for source determination problems are multi-faceted. First and foremost the behavior matching task T3 needs to be solved to fully realize the DT-enabled sensor placement and steering framework. From a practical standpoint, multi-sUAS systems are not readily available and have not been developed on a hardware and software

level in the lab. Additionally, to deploy these systems in practice, a Federal Aviation Administration (FAA) part 107 waiver is required. From a financial standpoint, the cost of highly sensitive lightweight sensors that are capable of carrying out the emission quantification tasks are still rather expensive - making it difficult to experiment with. Therefore the continued development of emission quantification methodology with DTs are still desirable. From a DT standpoint, development of real-time data acquisition and behavior matching systems are needed in order to deploy them in practice. Following the results and findings of many of the case studies explored in this body of work and in the literature, it seems imperative that concepts such as localizability and quantifiability need to be defined and explored. This would mean that for a system to be localizable, the set properties and constraints of the physical system (e.g. weather, terrain, leak type, etc.) and measurement system (e.g. location, sampling sequence, sensor type, etc.) would make it possible to localize the leak(s). Similarly, for a system to be quantifiable, the physical and measurement systems need to contain the properties necessary to quantify the leak(s). Probable candidates could include metrics from Monin-Obukhov Similarity Theory (e.g. Monin-Obukhov length, or friction velocity) regarding atmospheric stability, sensor characteristics (e.g. sensitivity, rise time, relaxation time, selectivity, saturation limit, stability, etc.), sensor placement (e.g. location or trajectory), or even the information about the terrain (such as obstacles or potential leak locations) in conjunction with local wind information. With these concepts emission measurements can become smarter and more reliable, making it more useful for practical applications.

Appendices

List of sUAS-based Field Campaign Missions

Within this dissertation work, several field missions were carried out to improve understanding and sensor capabilities for detection, localization, and quantification of a methane emission source. The field experiments consist of two general types: controlled release tests, and real-world emission scenarios.

Controlled Release Tests

The controlled release tests conducted are outlined below in brief.

- In February 2017, a two day single point source controlled release was conducted in the MVPGR. This includes some detection probability / localization, and quantification results are included in Section 3.4 & 3.5.
- In March 2017, a three day campaign was conducted at the simulated neighborhood in PG&E's Livermore training facility. The facility has the capability of single/multi sources at small, medium, and large leak rates. Some localization results are included in section 3.2. (Further work on localization can be done within this dataset, including Digital Twin representation - see Chapter 4).
- July 2017, as a part of NYSEARCH funded work, a series of flights were conducted at the SoCalGas facility located in Pico Rivera, CA, called Situation City. Multiple leak sources were released at one time and the team was tasked with locating all the leaks.
- In August 2018, a full day of single point source controlled release tests were conducted at the MVPGR. A total of 16 Mass balance flights were flown. (Results are shown in Section 3.5)
- In April 2021, a week long series of tests (3 actual testing days due to weather) were conducted at the Methane Emission Technology Evaluation Center (METEC) as a part of a Shell Global Solutions study (working under Aerometrix Inc.). (Results not yet shown)

Real-world Emission Scenario

The real-world emissions measured using the above techniques are described in brief.

- May 2017, as a part of NYSEARCH funded work, a series of flights were conducted in conjunction with Baltimore Gas and Electric at an undisclosed neighborhood in Baltimore, Maryland. The task was to detect and localize active and known leak sources.
- July 2017, as a part of NYSEARCH funded work, a series of flights were conducted at an undisclosed SoCalGas compressor station. The task was to detect and localized leak sources that may be on site, as well as quantify the total site emissions.
- In August 2018, Two landfills were surveyed (Merced County Landfill, and Yolo County Landfill). The Merced County Landfill has preliminary results for an active cell with 2ft dirt cap shown in Section 3.6. Results for Yolo are not shown here).
- In June 2019, close to 50 active oil and gas sites were visited and measured as a part of the Alberta Methane Field Challenge (AMFC). The results were published in Geoscience BC yearly report [200].
- In Fall 2020, an Alaskan peatland bog was flown at the Alaskan Peatland Experiment (APEX) in an attempt to quantify methane emissions from a natural methane emitter. The results are compared with an on-site Eddie covariance (EC) tower and presented in the American Geophysical Union [7].

Bibliography

- [1] D. Hollenbeck and Y. Chen, “Characterization of ground-to-air emissions with sUAS using a digital twin framework,” in *Proc. of the 2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1162–1166, IEEE, 2020.
- [2] W. Thielicke, W. Hübner, U. Müller, M. Eggert, and P. Wilhelm, “Towards accurate and practical drone-based wind measurements with an ultrasonic anemometer,” *Atmospheric Measurement Techniques*, vol. 14, no. 2, pp. 1303–1318, 2021.
- [3] S. Yang, R. W. Talbot, M. B. Frish, L. M. Golston, N. F. Aubut, M. A. Zondlo, C. Gretencord, and J. McSpirt, “Natural gas fugitive leak detection using an unmanned aerial vehicle: Measurement system description and mass balance approach,” *Atmosphere*, vol. 9, no. 10, p. 383, 2018.
- [4] T. K. Flesch, J. D. Wilson, and E. Yee, “Backward-time lagrangian stochastic dispersion models and their application to estimate gaseous emissions,” *Journal of Applied Meteorology and Climatology*, vol. 34, no. 6, pp. 1320–1332, 1995.
- [5] EPA, “Draft Other Test Method 33A: Geospatial measurement of air pollution, remote emissions quantification - direct assessment (GMAP-REQ-DA),” *Environmental Protection Agency (EPA)*, 2014. <https://www3.epa.gov/ttnemc01/prelim/otm33a.pdf>.
- [6] T. A. Foster-Wittig, E. D. Thoma, and J. D. Albertson, “Estimation of point source fugitive emission rates from a single sensor time series: A conditionally-sampled gaussian plume reconstruction,” *Atmospheric Environment*, vol. 115, pp. 101–109, 2015.
- [7] D. Hollenbeck, K. Manies, Y. Chen, D. Baldocchi, E. Euskirchen, and L. Christensen, “Evaluating a uav-based mobile sensing system designed to quantify ecosystem-based methane,” *Earth and Space Science Open Archive*, p. 15, 2021.
- [8] S. G. Nurzaman, Y. Matsumoto, Y. Nakamura, K. Shirai, S. Koizumi, and H. Ishiguro, “From lévy to brownian: a computational model based on biological fluctuation,” *PLOS One*, vol. 6, no. 2, p. e16168, 2011.

- [9] L. E. Christensen, “Miniature tunable laser spectrometer for detection of a trace gas,” June 2017. US Patent 9,671,332.
- [10] D. Hollenbeck, M. Dahabra, L. E. Christensen, and Y. Chen, “Data quality aware flight mission design for fugitive methane sniffing using fixed wing sUAS,” in *Proc. of the 2019 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 813–818, IEEE, 2019.
- [11] A. Shah, J. R. Pitt, H. Ricketts, J. B. Leen, P. I. Williams, K. Kabbabe, M. W. Gallagher, and G. Allen, “Testing the near-field gaussian plume inversion flux quantification technique using unmanned aerial vehicle sampling,” *Atmospheric Measurement Techniques*, vol. 13, no. 3, pp. 1467–1484, 2020.
- [12] N. Powel and K. A. Morgansen, “Empirical observability gramian for stochastic observability of nonlinear systems,” *arXiv preprint arXiv:2006.07451*, 2020.
- [13] D. Singh, B. Barlow, C. Hugenholtz, W. Funk, C. Robinson, and A. P. Ravikumar, “Field performance of new methane detection technologies: Results from the Alberta Methane Field Challenge,” 2021.
- [14] “Smart control engineering via digital twins (SCE via DT),” Aug 2020.
- [15] I. B. Ocko, T. Sun, D. Shindell, M. Oppenheimer, A. N. Hristov, S. W. Pacala, D. L. Mauzerall, Y. Xu, and S. P. Hamburg, “Acting rapidly to deploy readily available methane mitigation measures by sector can immediately slow global warming,” *Environmental Research Letters*, vol. 16, no. 5, p. 054042, 2021.
- [16] C. E. Kemp, A. P. Ravikumar, and A. R. Brandt, “Comparing natural gas leakage detection technologies using an open-source “virtual gas field” simulator,” *Environmental Science & Technology*, vol. 50, no. 8, pp. 4546–4553, 2016.
- [17] C. E. Kemp and A. P. Ravikumar, “New technologies can cost effectively reduce oil and gas methane emissions, but policies will require careful design to establish mitigation equivalence,” *Environmental Science & Technology*, vol. 55, no. 13, pp. 9140–9149, 2021.
- [18] GASMET, “How to measure greenhouse gas soil fluxes,” 2020.
- [19] D. Pickering, “New solutions for landfill surface emissions monitoring,” 2021.
- [20] E. Luetschwager, J. C. von Fischer, and Z. D. Weller, “Characterizing detection probabilities of advanced mobile leak surveys: Implications for sampling effort and leak size estimation in natural gas distribution systems,” *Elem Sci Anth*, vol. 9, no. 1, p. 00143, 2021.
- [21] H. Maazallahi, J. M. Fernandez, M. Menoud, D. Zavala-Araiza, Z. D. Weller, S. Schwietzke, J. C. von Fischer, H. Denier van der Gon, and T. Röckmann,

- “Methane mapping, emission quantification, and attribution in two european cities: Utrecht (nl) and hamburg (de),” *Atmospheric Chemistry and Physics*, vol. 20, no. 23, pp. 14717–14740, 2020.
- [22] Z. D. Weller, D. K. Yang, and J. C. von Fischer, “An open source algorithm to detect natural gas leaks from mobile methane survey data,” *PLoS One*, vol. 14, no. 2, p. e0212287, 2019.
- [23] K. MacKay, M. Lavoie, E. Bourlon, E. Atherton, E. O’Connell, J. Baillie, C. Fougère, and D. Risk, “Methane emissions from upstream oil and gas production in canada are underestimated,” *Scientific Reports*, vol. 11, no. 1, pp. 1–8, 2021.
- [24] J. G. Mønster, J. Samuelsson, P. Kjeldsen, C. W. Rella, and C. Scheutz, “Quantifying methane emission from fugitive sources by combining tracer release and downwind measurements—a sensitivity analysis based on multiple field surveys,” *Waste Management*, vol. 34, no. 8, pp. 1416–1428, 2014.
- [25] P. V. Riquetti, J. I. Fletcher, and C. D. Minty, “Aerial surveillance for gas and liquid hydrocarbon pipelines using a flame ionization detector (fid),” in *Proc. of the International Pipeline Conference*, vol. 40214, pp. 681–687, American Society of Mechanical Engineers, 1996.
- [26] A. K. Thorpe, C. Frankenberg, D. R. Thompson, R. M. Duren, A. D. Aubrey, B. D. Bue, R. O. Green, K. Gerilowski, T. Krings, J. Borchardt, *et al.*, “Airborne doas retrievals of methane, carbon dioxide, and water vapor concentrations at high spatial resolution: application to aviris-ng,” *Atmospheric Measurement Techniques*, vol. 10, no. 10, 2017.
- [27] A. Thorpe, C. Frankenberg, and D. Roberts, “Retrieval techniques for airborne imaging of methane concentrations using high spatial and moderate spectral resolution: Application to aviris,” *Atmospheric Measurement Techniques*, vol. 7, no. 2, pp. 491–506, 2014.
- [28] T. Rafiq, R. M. Duren, A. K. Thorpe, K. Foster, R. Patarsuk, C. E. Miller, and F. M. Hopkins, “Attribution of methane point source emissions using airborne imaging spectroscopy and the vista-california methane infrastructure dataset,” *Environmental Research Letters*, vol. 15, no. 12, p. 124001, 2020.
- [29] M. Cambaliza, P. Shepson, D. Caulton, B. Stirm, D. Samarov, K. Gurney, J. Turnbull, K. Davis, A. Possolo, A. Karion, *et al.*, “Assessment of uncertainties of an aircraft-based mass balance approach for quantifying urban greenhouse gas emissions,” *Atmospheric Chemistry and Physics*, vol. 14, no. 17, pp. 9029–9050, 2014.

- [30] D. Gasbarra, P. Toscano, D. Famulari, S. Finardi, P. Di Tommasi, A. Zaldei, P. Carlucci, E. Magliulo, and B. Gioli, “Locating and quantifying multiple landfills methane emissions using aircraft data,” *Environmental Pollution*, vol. 254, p. 112987, 2019.
- [31] M. R. Johnson, D. R. Tyner, and A. J. Szekeres, “Blinded evaluation of airborne methane source detection using bridger photonics lidar,” *Remote Sensing of Environment*, vol. 259, p. 112418, 2021.
- [32] D. Allen, S. Stokes, E. Tullos, B. Smith, S. Herndon, and B. Flowers, “Field trial of methane emission quantification technologies,” in *Proc. of the SPE Annual Technical Conference and Exhibition*, OnePetro, 2020.
- [33] A. P. Ravikumar, J. Wang, and A. R. Brandt, “Are optical gas imaging technologies effective for methane leak detection?,” *Environmental Science & Technology*, vol. 51, no. 1, pp. 718–724, 2017.
- [34] I. Leifer, C. Melton, M. L. Fischer, M. Fladeland, J. Frash, W. Gore, L. T. Iraci, J. E. Marrero, J.-M. Ryoo, T. Tanaka, *et al.*, “Atmospheric characterization through fused mobile airborne and surface in situ surveys: methane emissions quantification from a producing oil field,” *Atmospheric Measurement Techniques*, vol. 11, no. 3, pp. 1689–1705, 2018.
- [35] M. Mommert, M. Sigel, M. Neuhausler, L. Scheibenreif, and D. Borth, “Characterization of industrial smoke plumes from remote sensing data,” *arXiv preprint arXiv:2011.11344*, 2020.
- [36] S. Boschert and R. Rosen, “Digital twin—the simulation aspect,” in *Mechatronic futures*, pp. 59–74, Springer, 2016.
- [37] M. Grieves, “Digital twin: manufacturing excellence through virtual factory replication,” *White paper*, vol. 1, pp. 1–7, 2014.
- [38] A. M. Madni, C. C. Madni, and S. D. Lucero, “Leveraging digital twin technology in model-based systems engineering,” *Systems*, vol. 7, no. 1, p. 7, 2019.
- [39] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, “Digital twin-driven product design, manufacturing and service with big data,” *The International Journal of Advanced Manufacturing Technology*, vol. 94, no. 9-12, pp. 3563–3576, 2018.
- [40] R. Ganguli and S. Adhikari, “The digital twin of discrete dynamic systems: Initial approaches and future challenges,” *Applied Mathematical Modelling*, vol. 77, pp. 1110–1128, 2020.
- [41] I. Verner, D. Cuperman, S. Gamer, and A. Polishuk, “Digital twin of the robot baxter for learning practice in spatial manipulation tasks,” in *International*

- Conference on Remote Engineering and Virtual Instrumentation*, pp. 81–92, Springer, 2019.
- [42] M. Farsi, A. Daneshkhah, A. Hosseinian-Far, and H. Jahankhani, *Digital Twin Technologies and Smart Cities*. Springer, 2020.
- [43] V. Isakov, *Inverse problems for partial differential equations*, vol. 127. Springer, 2006.
- [44] Y. Chen, D. Hollenbeck, Y. Wang, and Y. Chen, “On optimal tempered Lévy flight foraging,” *Frontiers in Physics*, vol. 6, p. 111, 2018.
- [45] D. S. Schloesser, D. Hollenbeck, and C. T. Kello, “Individual and collective foraging in autonomous search agents with human intervention,” *Scientific Reports*, vol. 11, no. 1, pp. 1–13, 2021.
- [46] D. Hollenbeck, K. Zheng, D. Zulevic, and Y. Chen, “Swarm robotic source seeking with fractional fluxotaxis,” in *Proc. of the 2023 International Conference on Fractional Differentiation and Its Applications (ICFDA)*, 2023.
- [47] D. Hollenbeck and Y. Chen, “A more optimal stochastic extremum seeking control using fractional dithering for a class of smooth convex functions,” *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 3737–3742, 2020.
- [48] D. Hollenbeck, D. Zulevic, and Y. Chen, “Advanced leak detection and quantification of methane emissions using suass,” *Drones*, vol. 5, no. 4, p. 117, 2021.
- [49] D. Hollenbeck, D. Zulevic, and Y. Chen, “MOABS/DT: Methane Odor Abatement Simulator with Digital Twins,” in *Proceedings of the 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)*, pp. 378–381, IEEE, 2021.
- [50] D. Hollenbeck, D. Zulevic, and Y. Chen, “Single and multi-suass based emission quantification performance assessment using moabs/dt: A simulation case study,” in *Proc. of the 2022 18th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, pp. 1–5, IEEE, 2022.
- [51] D. Hollenbeck and Y. Chen, “Digital twin behavior matching of gas plumes using a fixed sensor mesh and dynamic mode decomposition,” in *Proc. of the 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 85437, p. V007T07A005, American Society of Mechanical Engineers, 2021.
- [52] D. M. Gordon, “Ant encounters,” in *Ant Encounters*, Princeton University Press, 2010.

- [53] T. Vicsek and A. Zafeiris, “Collective motion,” *Physics reports*, vol. 517, no. 3-4, pp. 71–140, 2012.
- [54] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proceedings of ICNN’95-international conference on neural networks*, vol. 4, pp. 1942–1948, IEEE, 1995.
- [55] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey wolf optimizer,” *Advances in engineering software*, vol. 69, pp. 46–61, 2014.
- [56] X.-S. Yang and S. Deb, “Cuckoo search via lévy flights,” in *2009 World congress on nature & biologically inspired computing (NaBIC)*, pp. 210–214, Ieee, 2009.
- [57] F. Bartumeus, M. E. Da Luz, G. Viswanathan, and J. Catalan, “Animal search strategies: a quantitative random-walk analysis,” *Ecology*, vol. 86, no. 11, pp. 3078–3087, 2005.
- [58] G. Viswanathan, E. Raposo, and M. Da Luz, “Lévy flights and superdiffusion in the context of biological encounters and random searches,” *Physics of Life Reviews*, vol. 5, no. 3, pp. 133–150, 2008.
- [59] A. Reynolds, “Optimal random lévy-loop searching: new insights into the searching behaviours of central-place foragers,” *EPL (Europhysics Letters)*, vol. 82, no. 2, p. 20001, 2008.
- [60] F. Bartumeus, J. Catalan, U. Fulco, M. Lyra, and G. Viswanathan, “Optimizing the encounter rate in biological interactions: Lévy versus brownian strategies,” *Physical Review Letters*, vol. 88, no. 9, p. 097901, 2002.
- [61] B. E. Kerster, T. Rhodes, and C. T. Kello, “Spatial memory in foraging games,” *Cognition*, vol. 148, pp. 85–96, 2016.
- [62] G. M. Viswanathan, S. V. Buldyrev, S. Havlin, M. Da Luz, E. Raposo, and H. E. Stanley, “Optimizing the success of random searches,” *nature*, vol. 401, no. 6756, p. 911, 1999.
- [63] G. Viswanathan, V. Afanasyev, S. V. Buldyrev, S. Havlin, M. Da Luz, E. Raposo, and H. E. Stanley, “Lévy flights in random searches,” *Physica A: Statistical Mechanics and its Applications*, vol. 282, no. 1-2, pp. 1–12, 2000.
- [64] G. M. Viswanathan, V. Afanasyev, S. Buldyrev, E. Murphy, P. Prince, and H. E. Stanley, “Lévy flight search patterns of wandering albatrosses,” *Nature*, vol. 381, no. 6581, p. 413, 1996.
- [65] D. Boyer, O. Miramontes, G. Ramos-Fernandez, J. Mateos, and G. Cocho, “Modeling the searching behavior of social monkeys,” *Physica A: Statistical Mechanics and its Applications*, vol. 342, no. 1-2, pp. 329–335, 2004.

- [66] G. Ramos-Fernández, J. L. Mateos, O. Miramontes, G. Cocho, H. Larralde, and B. Ayala-Orozco, “Lévy walk patterns in the foraging movements of spider monkeys (*ateles geoffroyi*),” *Behavioral ecology and Sociobiology*, vol. 55, no. 3, pp. 223–230, 2004.
- [67] B. Dybiec, E. Gudowska-Nowak, E. Barkai, and A. A. Dubkov, “Lévy flights versus lévy walks in bounded domains,” *Physical Review E*, vol. 95, no. 5, p. 052102, 2017.
- [68] K. Zhao, R. Jurdak, J. Liu, D. Westcott, B. Kusy, H. Parry, P. Sommer, and A. McKeown, “Optimal lévy-flight foraging in a finite landscape,” *Journal of The Royal Society Interface*, vol. 12, no. 104, p. 20141158, 2015.
- [69] M. Plank and A. James, “Optimal foraging: Levy pattern or process?,” *Journal of The Royal Society Interface*, vol. 5, no. 26, pp. 1077–1086, 2008.
- [70] F. Bartumeus, E. P. Raposo, G. M. Viswanathan, and M. G. da Luz, “Stochastic optimal foraging: tuning intensive and extensive dynamics in random searches,” *PloS one*, vol. 9, no. 9, p. e106373, 2014.
- [71] C. B. Zeng and Y. Q. Chen, “Optimal random search, fractional dynamics and fractional calculus,” *Fractional Calculus and Applied Analysis*, vol. 17, no. 2, pp. 321–332, 2014.
- [72] M. D. Hauser and P. Marler, “Food-associated calls in rhesus macaques (*macaca mulatta*): I. socioecological factors,” *Behavioral Ecology*, vol. 4, no. 3, pp. 194–205, 1993.
- [73] K. McComb, D. Reby, L. Baker, C. Moss, and S. Sayialel, “Long-distance communication of acoustic cues to social identity in african elephants,” *Animal Behaviour*, vol. 65, no. 2, pp. 317–329, 2003.
- [74] C. R. Brown, M. B. Brown, and M. L. Shaffer, “Food-sharing signals among socially foraging cliff swallows,” *Animal behaviour*, vol. 42, no. 4, pp. 551–564, 1991.
- [75] R. Frey, A. Gebler, K. A. Olson, D. Odonkhuu, G. Fritsch, N. Batsaikhan, and I. W. Stuermer, “Mobile larynx in mongolian gazelle: retraction of the larynx during rutting barks in male mongolian gazelle (*procapra gutturosa pallas*, 1777),” *Journal of Morphology*, vol. 269, no. 10, pp. 1223–1237, 2008.
- [76] J. Lamprecht, “The function of social hunting in larger terrestrial carnivores,” *Mammal Review*, vol. 11, no. 4, pp. 169–179, 1981.
- [77] D. E. Jackson and F. L. Ratnieks, “Communication in ants,” *Current biology*, vol. 16, no. 15, pp. R570–R574, 2006.

- [78] V. Tereshko and A. Loengarov, “Collective decision making in honey-bee foraging dynamics,” *Computing and Information Systems*, vol. 9, no. 3, p. 1, 2005.
- [79] L.-A. Giraldeau and T. Caraco, *Social foraging theory*. Princeton University Press, 2018.
- [80] G. Beauchamp, “Does group foraging promote efficient exploitation of resources?,” *Oikos*, vol. 111, no. 2, pp. 403–407, 2005.
- [81] C. Packer and L. Ruttan, “The evolution of cooperative hunting,” *The American Naturalist*, vol. 132, no. 2, pp. 159–198, 1988.
- [82] R. Harel, O. Spiegel, W. M. Getz, and R. Nathan, “Social foraging and individual consistency in following behaviour: testing the information centre hypothesis in free-ranging vultures,” *Proceedings of the Royal Society B: Biological Sciences*, vol. 284, no. 1852, p. 20162654, 2017.
- [83] F.-X. Dechaume-Moncharmont, A. Dornhaus, A. I. Houston, J. M. McNamara, E. J. Collins, and N. R. Franks, “The hidden cost of information in collective foraging,” *Proceedings of the Royal Society B: Biological Sciences*, vol. 272, no. 1573, pp. 1689–1695, 2005.
- [84] T. Palacios-Romo, F. Castellanos, and G. Ramos-Fernandez, “Uncovering the decision rules behind collective foraging in spider monkeys,” *Animal Behaviour*, vol. 149, pp. 121–133, 2019.
- [85] T. D. Seeley, “Division of labor between scouts and recruits in honeybee foraging,” *Behavioral ecology and sociobiology*, vol. 12, no. 3, pp. 253–259, 1983.
- [86] C. Dreller, “Division of labor between scouts and recruits: genetic influence and mechanisms,” *Behavioral Ecology and Sociobiology*, vol. 43, no. 3, pp. 191–196, 1998.
- [87] M. Lihoreau, M. A. Charleston, A. M. Senior, F. J. Clissold, D. Raubenheimer, S. J. Simpson, and J. Buhl, “Collective foraging in spatially complex nutritional environments,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 372, no. 1727, p. 20160238, 2017.
- [88] A. Cortés-Avizanda, R. Jovani, J. A. Donázar, and V. Grimm, “Bird sky networks: How do avian scavengers use social information to find carrion?,” *Ecology*, vol. 95, no. 7, pp. 1799–1808, 2014.
- [89] P. López-López, J. Benavent-Corai, C. García-Ripollés, and V. Urios, “Scavengers on the move: behavioural changes in foraging search patterns during the annual cycle,” *PloS one*, vol. 8, no. 1, p. e54352, 2013.

- [90] A. L. Jackson, G. D. Ruxton, and D. C. Houston, “The effect of social facilitation on foraging success in vultures: a modelling study,” *Biology Letters*, vol. 4, no. 3, pp. 311–313, 2008.
- [91] Y. Liu and K. M. Passino, “Stable social foraging swarms in a noisy environment,” *IEEE Transactions on automatic control*, vol. 49, no. 1, pp. 30–44, 2004.
- [92] A. Falcón-Cortés, D. Boyer, and G. Ramos-Fernández, “Collective learning from individual experiences and information transfer during group foraging,” *Journal of the Royal Society Interface*, vol. 16, no. 151, p. 20180803, 2019.
- [93] W. I. Sellers, R. A. Hill, and B. Logan, “An agent-based model of group decision making in baboons,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 362, no. 1485, pp. 1699–1710, 2007.
- [94] C. J. Torney, A. Berdahl, and I. D. Couzin, “Signalling and the evolution of cooperative foraging in dynamic environments,” *PLoS computational biology*, vol. 7, no. 9, p. e1002194, 2011.
- [95] B. A. Camley, J. Zimmermann, H. Levine, and W.-J. Rappel, “Emergent collective chemotaxis without single-cell gradient sensing,” *Physical review letters*, vol. 116, no. 9, p. 098101, 2016.
- [96] K. Bhattacharya and T. Vicsek, “Collective foraging in heterogeneous landscapes,” *Journal of The Royal Society Interface*, vol. 11, no. 100, p. 20140674, 2014.
- [97] R. Martínez-García, J. M. Calabrese, T. Mueller, K. A. Olson, and C. López, “Optimizing the search for resources by sharing information: Mongolian gazelles as a case study,” *Physical Review Letters*, vol. 110, no. 24, p. 248106, 2013.
- [98] K. Copenhagen, D. A. Quint, and A. Gopinathan, “Self-organized sorting limits behavioral variability in swarms,” *Scientific reports*, vol. 6, no. 1, pp. 1–11, 2016.
- [99] D. F. Spears, R. Anderson-Sprecher, A. Kletsov, and A. Rebguns, “A statistical framework for estimating the success rate of liquid-mimetic swarms,” in *Physicomimetics*, pp. 475–503, Springer, 2011.
- [100] W. M. Spears and D. F. Spears, *Physicomimetics: Physics-based swarm intelligence*. Springer Science & Business Media, 2012.
- [101] G. Ramos-Fernandez, A. J. King, J. C. Beehner, T. J. Bergman, M. C. Crofoot, A. Di Fiore, J. Lehmann, C. M. Schaffner, N. Snyder-Mackler, K. Zuberbühler, *et al.*, “Quantifying uncertainty due to fission–fusion dynamics as a component of social complexity,” *Proceedings of the Royal Society B: Biological Sciences*, vol. 285, no. 1879, p. 20180532, 2018.

- [102] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, “Novel type of phase transition in a system of self-driven particles,” *Physical review letters*, vol. 75, no. 6, p. 1226, 1995.
- [103] G. C. Van Orden, J. G. Holden, and M. T. Turvey, “Self-organization of cognitive performance.,” *Journal of experimental psychology: General*, vol. 132, no. 3, p. 331, 2003.
- [104] E. L. Charnov, “Optimal foraging, the marginal value theorem,” *Theoretical population biology*, vol. 9, no. 2, pp. 129–136, 1976.
- [105] J. M. Pleasants, “Optimal foraging by nectarivores: a test of the marginal-value theorem,” *The American Naturalist*, vol. 134, no. 1, pp. 51–71, 1989.
- [106] K. A. Ehinger and J. M. Wolfe, “When is it time to move to the next map? optimal foraging in guided visual search,” *Attention, Perception, & Psychophysics*, vol. 78, no. 7, pp. 2135–2151, 2016.
- [107] M. Krstić and H.-H. Wang, “Stability of extremum seeking feedback for general nonlinear dynamic systems,” *Automatica*, vol. 36, no. 4, pp. 595–601, 2000.
- [108] R. T. Cardé, “Odour plumes and odour-mediated flight in insects,” in *Ciba Foundation Symposium*, pp. 54–66, Wiley Online Library, 1996.
- [109] K. J. Lohmann, “How sea turtles navigate,” *Scientific American*, vol. 266, no. 1, pp. 100–107, 1992.
- [110] A. D. Hasler and A. T. Scholz, *Olfactory imprinting and homing in salmon: Investigations into the mechanism of the imprinting process*, vol. 14. Springer Science & Business Media, 2012.
- [111] X. Chen and J. Huang, “Odor source localization algorithms on mobile robots: A review and future outlook,” *Robotics and Autonomous Systems*, vol. 112, pp. 123–136, 2019.
- [112] P. Civicioglu and E. Besdok, “A conceptual comparison of the cuckoo-search, particle swarm optimization, differential evolution and artificial bee colony algorithms,” *Artificial Intelligence Review*, vol. 39, no. 4, pp. 315–346, 2013.
- [113] J. Wei, Y. Chen, Y. Yu, and Y. Chen, “Optimal randomness in swarm-based search,” *Mathematics*, vol. 7, no. 9, p. 828, 2019.
- [114] A. R. Mesquita, J. P. Hespanha, and K. Åström, “Optimotaxis: A stochastic multi-agent optimization procedure with point measurements,” in *International Workshop on Hybrid Systems: Computation and Control*, pp. 358–371, Springer, 2008.

- [115] D. R. Webster, K. Y. Volyanskyy, and M. J. Weissburg, “Bioinspired algorithm for autonomous sensor-driven guidance in turbulent chemical plumes,” *Bioinspiration & Biomimetics*, vol. 7, no. 3, p. 036023, 2012.
- [116] W. Wu, I. D. Couzin, and F. Zhang, “Bio-inspired source seeking with no explicit gradient estimation,” *IFAC Proceedings Volumes*, vol. 45, no. 26, pp. 240–245, 2012.
- [117] N. Saxena, D. Natesan, and S. P. Sane, “Odor source localization in complex visual environments by fruit flies,” *Journal of Experimental Biology*, vol. 221, no. 2, p. jeb172023, 2018.
- [118] Y. Chen, Q. Gao, Y. Wei, and Y. Wang, “Study on fractional order gradient methods,” *Applied Mathematics and Computation*, vol. 314, pp. 310–321, 2017.
- [119] Y. Chen, Y. Wei, Y. Wang, and Y. Chen, “On the unified design of accelerated gradient descent,” in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 59292, p. V009T12A015, American Society of Mechanical Engineers, 2019.
- [120] J. A. Farrell, J. Murlis, X. Long, W. Li, and R. T. Cardé, “Filament-based atmospheric dispersion model to achieve short time-scale structure of odor plumes,” *Environmental Fluid Mechanics*, vol. 2, no. 1-2, pp. 143–169, 2002.
- [121] I. Podlubny, *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*. Academic Press, 1998.
- [122] D. Zarzhitsky, D. F. Spears, and W. M. Spears, “Distributed robotics approach to chemical plume tracing,” in *Proc. of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4034–4039, IEEE, 2005.
- [123] G. Kowadlo and R. A. Russell, “Robot odor localization: a taxonomy and survey,” *The International Journal of Robotics Research*, vol. 27, no. 8, pp. 869–894, 2008.
- [124] I. Petráš, “An effective numerical method and its utilization to solution of fractional models used in bioengineering applications,” *Advances in Difference Equations*, vol. 2011, no. 1, pp. 1687–1847, 2011.
- [125] I. Petráš, “Digital fractional-order differentiator and integrator - new iir type,” Jan 2021. <https://www.mathworks.com/matlabcentral/fileexchange/31358-digital-fractional-order-differentiator-and-integrator-new-iir-type>.
- [126] Y. Chen, I. Petras, and D. Xue, “Fractional order control-a tutorial,” in *2009 American Control Conference*, pp. 1397–1411, IEEE, 2009.

- [127] M. Hutchinson, H. Oh, and W.-H. Chen, “Entrotaxis as a strategy for autonomous search and source reconstruction in turbulent conditions,” *Information Fusion*, vol. 42, pp. 179–189, 2018.
- [128] J. Matthes, L. Groll, and H. B. Keller, “Source localization by spatially distributed electronic noses for advection and diffusion,” *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1711–1719, 2005.
- [129] A. Nehorai, B. Porat, and E. Paldi, “Detection and localization of vapor-emitting sources,” *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp. 243–253, 1995.
- [130] D. Zarzhitsky, D. F. Spears, W. M. Spears, and D. R. Thayer, “A fluid dynamics approach to multi-robot chemical plume tracing,” in *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pp. 1476–1477, IEEE Computer Society, 2004.
- [131] M. Leblanc, “Sur l’électrification des chemins de fer au moyen de courants alternatifs de fréquence élevée,” *Revue générale de l’électricité*, vol. 12, no. 8, pp. 275–277, 1922.
- [132] C. Zhang and R. Ordóñez, *Extremum-seeking control and applications: a numerical optimization-based approach*. Springer Science & Business Media, 2011.
- [133] S.-J. Liu and M. Krstic, *Stochastic averaging and stochastic extremum seeking*. Springer Science & Business Media, 2012.
- [134] C. Manzie and M. Krstic, “Extremum seeking with stochastic perturbations,” *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 580–585, 2009.
- [135] F. Coito, J. Lemos, and S. Alves, “Stochastic extremum seeking in the presence of constraints,” *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 276–281, 2005.
- [136] S.-J. Liu and M. Krstić, “Discrete-time stochastic extremum seeking,” *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 3274–3279, 2014.
- [137] C. R. Dietrich and G. N. Newsam, “Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix,” *SIAM Journal on Scientific Computing*, vol. 18, no. 4, pp. 1088–1107, 1997.
- [138] M. Shao and C. L. Nikias, “Signal processing with fractional lower order moments: stable processes and their applications,” *Proceedings of the IEEE*, vol. 81, no. 7, pp. 986–1010, 1993.
- [139] C. L. Nikias and M. Shao, *Signal processing with alpha-stable distributions and applications*. Wiley-Interscience, 1995.
- [140] H. K. Khalil, *Nonlinear Control*. Pearson New York, 2015.

- [141] M. Krstic and H.-H. Wang, "Design and stability analysis of extremum seeking feedback for general nonlinear systems," in *Proceedings of the 36th IEEE Conference on Decision and Control*, vol. 2, pp. 1743–1748, IEEE, 1997.
- [142] J. Burgués and S. Marco, "Environmental chemical sensing using small drones: A review," *Science of The Total Environment*, p. 141172, 2020.
- [143] L. Fjelsted, T. Thomasen, I. Valbjørn, C. Scheutz, A. Christensen, and P. Kjeldsen, "Development of an innovative UAV-mounted screening tool for landfill gas emissions," in *Proc. of the Sardinia 2015-15th International Waste Management and Landfill Symposium*, CISA Publisher, 2015.
- [144] L. Fjelsted, A. Christensen, J. Larsen, P. Kjeldsen, and C. Scheutz, "Assessment of a landfill methane emission screening method using an unmanned aerial vehicle mounted thermal infrared camera—a field study," *Waste Management*, vol. 87, pp. 893–904, 2019.
- [145] J.-p. Gagnon, "Quantification of VOC Emissions Using Remote Thermal Infrared Hyperspectral Imaging," 2020. Telops Inc.
- [146] M. Gålfalk, G. Olofsson, P. Crill, and D. Bastviken, "Making methane visible," *Nature Climate Change*, vol. 6, no. 4, pp. 426–430, 2016.
- [147] B. J. Emran, D. D. Tannant, and H. Najjaran, "Low-altitude aerial methane concentration mapping," *Remote Sensing*, vol. 9, no. 8, p. 823, 2017.
- [148] T. K. Flesch, S. M. McGinn, D. Chen, J. D. Wilson, and R. L. Desjardins, "Data filtering for inverse dispersion emission calculations," *Agricultural and Forest Meteorology*, vol. 198, pp. 1–6, 2014.
- [149] Z. J. DeBruyn, C. Wagner-Riddle, and A. VanderZaag, "Assessment of open-path spectrometer accuracy at low path-integrated methane concentrations," *Atmosphere*, vol. 11, no. 2, p. 184, 2020.
- [150] T. R. Bretschneider and K. Shetti, "UAV-based gas pipeline leak detection," *35th Asian Conference on Remote Sensing 2014, ACRS 2014: Sensing for Reintegration of Societies*, no. April, 2014.
- [151] D. Bastviken, "Supplement of Technical Note: Cost-efficient approaches to measure carbon dioxide (CO₂) fluxes and concentrations in terrestrial and aquatic environments using mini loggers The copyright of individual parts of the supplement might differ from the CC-BY ," vol. 12, pp. 3849–3859, 2015.
- [152] D. Bastviken, I. Sundgren, S. Natchimuthu, H. Reyier, and M. Gålfalk, "Cost-efficient approaches to measure carbon dioxide (CO₂) fluxes and concentrations in terrestrial and aquatic environments using mini loggers," *Biogeosciences*, vol. 12, no. 12, pp. 3849–3859, 2015.

- [153] D. Bastviken, J. Nygren, J. Schenk, R. Parellada Massana, and N. Thanh Duc, “Technical note: Facilitating the use of low-cost methane (CH_4) sensors in flux chambers-calibration, data processing, and an open-source make-it-yourself logger,” *Biogeosciences*, vol. 17, no. 13, pp. 3659–3667, 2020.
- [154] M. Carrozzo, S. De Vito, E. Esposito, M. Salvato, F. Formisano, E. Massera, G. Di Francia, P. D. Veneri, M. Iadaresta, and A. Mennella, “UAV intelligent chemical multisensor payload for networked and impromptu gas monitoring tasks,” *Proc. of the 5th IEEE International Workshop on Metrology for AeroSpace, MetroAeroSpace*, pp. 112–116, 2018.
- [155] F. Arduini, S. Cinti, V. Scognamiglio, D. Moscone, and G. Palleschi, “How cutting-edge technologies impact the design of electrochemical (bio) sensors for environmental analysis. a review,” *Analytica Chimica Acta*, vol. 959, pp. 15–42, 2017.
- [156] D. Karakaya, O. Ulucan, and M. Turkan, “Electronic nose and its applications: a survey,” *International Journal of Automation and Computing*, vol. 17, no. 2, pp. 179–209, 2020.
- [157] W. T. Honeycutt, M. T. Ley, and N. F. Materer, “Precision and limits of detection for selected commercially available, low-cost carbon dioxide and methane gas sensors,” *Sensors*, vol. 19, no. 14, p. 3157, 2019.
- [158] Y. M. Kim, M. H. Park, S. Jeong, K. H. Lee, and J. Y. Kim, “Evaluation of error inducing factors in unmanned aerial vehicle mounted detector to measure fugitive methane from solid waste landfill,” *Waste Management*, vol. 124, pp. 368–376, 2021.
- [159] D. Hollenbeck, G. Nunez, L. E. Christensen, and Y. Chen, “Wind measurement and estimation with small unmanned aerial systems (sUAS) using on-board mini ultrasonic anemometers,” in *Proc. of the 2018 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 285–292, IEEE, 2018.
- [160] D. Hollenbeck, M. Oyama, A. Garcia, and Y. Chen, “Pitch and roll effects of on-board wind measurements using sUAS,” in *Proc. of the 2019 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1249–1254, IEEE, 2019.
- [161] P. Bruschi, M. Piotto, F. Dell’Agnello, J. Ware, and N. Roy, “Wind speed and direction detection by means of solid-state anemometers embedded on small quadcopters,” *Procedia Engineering*, vol. 168, pp. 802–805, 2016.
- [162] J. González-Rocha, S. F. De Wekker, S. D. Ross, and C. A. Woolsey, “Wind profiling in the lower atmosphere from wind-induced perturbations to multirotor uas,” *arXiv preprint arXiv:2001.02740*, 2020.

- [163] C. Ingenhorst, G. Jacobs, L. Stöbel, R. Schelenz, and B. Juretzki, “Method for airborne measurement of the spatial wind speed distribution above complex terrain,” *Wind Energy Science*, vol. 6, no. 2, pp. 427–440, 2021.
- [164] M. J. Brewer and C. B. Clements, “Meteorological profiling in the fire environment using UAS,” *Fire*, vol. 3, no. 3, p. 36, 2020.
- [165] R. Calmer, G. C. Roberts, J. Preissler, K. J. Sanchez, S. Derrien, and C. O’Dowd, “Vertical wind velocity measurements using a five-hole probe with remotely piloted aircraft to study aerosol–cloud interactions,” *Atmospheric Measurement Techniques*, vol. 11, no. 5, pp. 2583–2599, 2018.
- [166] C. Brosy, K. Krampf, M. Zeeman, B. Wolf, W. Junkermann, K. Schäfer, S. Emeis, and H. Kunstmann, “Simultaneous multicopter-based air sampling and sensing of meteorological variables,” *Atmospheric Measurement Techniques*, vol. 10, no. 8, pp. 2773–2784, 2017.
- [167] C. Brosy, “Hexacopter-based three-dimensional measurements of hydrometeorological variables and methane,” *Diss., Universität Augsburg*, 2019.
- [168] P. Tian, H. Chao, M. Rhudy, J. Gross, and H. Wu, “Wind sensing and estimation using small fixed-wing unmanned aerial vehicles: A survey,” *Journal of Aerospace Information Systems*, vol. 18, no. 3, pp. 132–143, 2021.
- [169] B. Stark, B. Smith, and Y. Chen, “A guide for selecting small unmanned aerial systems for research-centric applications,” *IFAC Proceedings Volumes*, vol. 46, no. 30, pp. 38–45, 2013.
- [170] B. Smith, G. John, B. Stark, L. E. Christensen, and Y. Chen, “Applicability of unmanned aerial systems for leak detection,” in *Proc. of the 2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1220–1227, IEEE, 2016.
- [171] N. S. Holmes and L. Morawska, “A review of dispersion modelling and its application to the dispersion of particles: an overview of different dispersion models available,” *Atmospheric Environment*, vol. 40, no. 30, pp. 5902–5928, 2006.
- [172] C. Arndt, A. Leytem, A. Hristov, D. Zavala-Araiza, J. Cativiela, S. Conley, C. Daube, I. Faloon, and S. Herndon, “Short-term methane emissions from 2 dairy farms in california estimated by different measurement techniques and us environmental protection agency inventory methodology: A case study,” *Journal of Dairy Science*, vol. 101, no. 12, pp. 11461–11479, 2018.
- [173] T. Flesch, J. Wilson, L. Harper, B. Crenna, and R. Sharpe, “Deducing ground-to-air emissions from observed trace gas concentrations: a field trial,” *Journal of Applied Meteorology*, vol. 43, no. 3, pp. 487–502, 2004.

- [174] T. Flesch, J. Wilson, and L. Harper, “Deducing ground-to-air emissions from observed trace gas concentrations: a field trial with wind disturbance,” *Journal of Applied Meteorology*, vol. 44, no. 4, pp. 475–484, 2005.
- [175] K. S. Ro, M. H. Johnson, K. C. Stone, P. G. Hunt, T. Flesch, and R. W. Todd, “Measuring gas emissions from animal waste lagoons with an inverse-dispersion technique,” *Atmospheric Environment*, vol. 66, pp. 101–106, 2013.
- [176] Y. Y. Cui, J. Brioude, S. A. McKeen, W. M. Angevine, S.-W. Kim, G. J. Frost, R. Ahmadov, J. Peischl, N. Bousserez, Z. Liu, *et al.*, “Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The South Coast Air Basin,” *Journal of Geophysical Research: Atmospheres*, vol. 120, no. 13, pp. 6698–6711, 2015.
- [177] C. Hunter, “A recommended Pasquill-Gifford stability classification method for safety basis atmospheric dispersion modeling at SRS,” tech. rep., Savannah River Site (SRS), 2012.
- [178] R. Edie, A. M. Robertson, R. A. Field, J. Soltis, D. A. Snare, D. Zimmerle, C. S. Bell, T. L. Vaughn, and S. M. Murphy, “Constraining the accuracy of flux estimates using OTM 33A,” *Atmospheric Measurement Techniques*, vol. 13, no. 1, pp. 341–353, 2020.
- [179] R. Edie, A. M. Robertson, J. Soltis, R. A. Field, D. Snare, M. D. Burkhardt, and S. M. Murphy, “Off-site flux estimates of volatile organic compounds from oil and gas production facilities using fast-response instrumentation,” *Environmental Science & Technology*, vol. 54, no. 3, pp. 1385–1394, 2019.
- [180] A. Van Ulden, “Simple estimates for vertical diffusion from sources near the ground,” *Atmospheric Environment*, vol. 12, no. 11, pp. 2125–2129, 1978.
- [181] R. M. Eckman, “Re-examination of empirically derived formulas for horizontal diffusion from surface sources,” *Atmospheric Environment*, vol. 28, no. 2, pp. 265–272, 1994.
- [182] J. D. Albertson, T. Harvey, G. Foderaro, P. Zhu, X. Zhou, S. Ferrari, M. S. Amin, M. Modrak, H. Brantley, and E. D. Thoma, “A mobile sensing approach for regional surveillance of fugitive methane emissions in oil and gas production,” *Environmental Science & Technology*, vol. 50, no. 5, pp. 2487–2497, 2016.
- [183] J. R. Gemerek, S. Ferrari, and J. D. Albertson, “Fugitive gas emission rate estimation using multiple heterogeneous mobile sensors,” in *Proc. of the 2017 ISOCS/IEEE International Symposium on Olfaction and Electronic Nose (ISOEN)*, pp. 1–3, IEEE, 2017.
- [184] E. Yee and T. K. Flesch, “Inference of emission rates from multiple sources using Bayesian probability theory,” *Journal of Environmental Monitoring*, vol. 12, no. 3, pp. 622–634, 2010.

- [185] H. L. Brantley, E. D. Thoma, W. C. Squier, B. B. Guven, and D. Lyon, "Assessment of methane emissions from oil and gas production pads using mobile measurements," *Environmental Science & Technology*, vol. 48, no. 24, pp. 14508–14515, 2014.
- [186] E. Yee, "Probability theory as logic: data assimilation for multiple source reconstruction," *Pure and Applied Geophysics*, vol. 169, no. 3, pp. 499–517, 2012.
- [187] X. Zhou, S. Yoon, S. Mara, M. Falk, T. Kuwayama, T. Tran, L. Cheadle, J. Nyarady, B. Croes, E. Scheehle, *et al.*, "Mobile sampling of methane emissions from natural gas well pads in California," *Atmospheric Environment*, vol. 244, p. 117930, 2021.
- [188] M. Hutchinson, C. Liu, and W.-H. Chen, "Source term estimation of a hazardous airborne release using an unmanned aerial vehicle," *Journal of Field Robotics*, vol. 36, no. 4, pp. 797–817, 2019.
- [189] A. Shah, G. Allen, J. R. Pitt, H. Ricketts, P. I. Williams, J. Helmore, A. Finlayson, R. Robinson, K. Kabbabe, P. Hollingsworth, *et al.*, "A near-field gaussian plume inversion flux quantification method, applied to unmanned aerial vehicle sampling," *Atmosphere*, vol. 10, no. 7, p. 396, 2019.
- [190] Adil Shah, "Supplement to A Near-Field Gaussian Plume Inversion Flux Quantification Method, Suitable For Unmanned Aerial Vehicle Sampling," *Atmosphere*, 2020.
- [191] J. D'Errico, "fminsearchbnd, fminsearchcon," Feb 2022.
- [192] G. Allen, P. Hollingsworth, K. Kabbabe, J. R. Pitt, M. I. Mead, S. Illingworth, G. Roberts, M. Bourn, D. E. Shallcross, and C. J. Percival, "The development and trial of an unmanned aerial system for the measurement of methane flux from landfill and greenhouse gas emission hotspots," *Waste Management*, vol. 87, pp. 883–892, 2019.
- [193] M. O. L. Cambaliza, J. E. Bogner, R. B. Green, P. B. Shepson, T. A. Harvey, K. A. Spokas, B. H. Stirm, M. Corcoran, D. Helmig, and A. Wisthaler, "Field measurements and modeling to resolve m^2 to km^2 CH_4 emissions for a complex urban source: An indiana landfill study," *Elementa: Science of the Anthropocene*, vol. 5, 2017.
- [194] J. L. France, P. Bateson, P. Dominutti, G. Allen, S. Andrews, S. Bauguitte, M. Coleman, T. Lachlan-Cope, R. E. Fisher, L. Huang, *et al.*, "Facility level measurement of offshore oil and gas installations from a medium-sized airborne platform: method development for quantification and source identification of methane emissions," *Atmospheric Measurement Techniques*, vol. 14, no. 1, pp. 71–88, 2021.

- [195] R. P. Morales, J. Ravelid, K. P. Brennan, B. Tuzson, L. Emmenegger, and D. Brunner, “Estimating local methane sources from drone-based laser spectrometer measurements by mass-balance method,” in *EGU General Assembly Conference Abstracts*, p. 14778, 2020.
- [196] A. P. Ravikumar, S. Sreedhara, J. Wang, J. Englander, D. Roda-Stuart, C. Bell, D. Zimmerle, D. Lyon, I. Mogstad, B. Ratner, *et al.*, “Single-blind inter-comparison of methane detection technologies—results from the stanford/edf mobile monitoring challenge,” *Elementa: Science of the Anthropocene*, vol. 7, 2019.
- [197] M. Reuter, H. Bovensmann, M. Buchwitz, J. Borchardt, S. Krautwurst, K. Gerilowski, M. Lindauer, D. Kubistin, and J. P. Burrows, “Development of a small unmanned aircraft system to derive CO₂ emissions of anthropogenic point sources,” *Atmospheric Measurement Techniques*, vol. 14, no. 1, pp. 153–172, 2021.
- [198] A. Shah, G. Allen, H. Ricketts, J. Pitt, and P. Williams, “Methane flux quantification from lactating cattle using unmanned aerial vehicles,” *EGU General Assembly*, vol. 20, p. 7655, 2018.
- [199] A. A. Shah, “Methane Flux Quantification Using Unmanned Aerial Vehicles,” *Diss., University of Manchester*, 2020.
- [200] M. Whitarcar, D. Hollenbeck, B. Billwiller, C. Salas, and L. Christensen, “Application of the bc GHGMapper™ platform for the Alberta Methane Field Challenge (AMFC),” *Geoscience BC Summary of Activities 2019: Energy and Water, Geoscience BC, Report 2020-02*, pp. 87–102, 2020.
- [201] M. Whitarcar, L. Christensen, C. Salas, and P. Reece, “Ghgmap: Detection of fugitive methane leaks from natural gas pipelines British Columbia and Alberta,” *Geoscience BC Summary of Activities 2018: Energy and Water, Geoscience BC, Report 2019-2*, pp. 67–76, 2019.
- [202] M. Whitarcar, L. Christensen, C. Salas, and P. Reece, “GHGMap: novel approach for aerial measurements of greenhouse gas emissions British Columbia,” *Geoscience BC Summary of Activities 2017: Energy, Geoscience BC, Report 2018-4*, pp. 1–10, 2018.
- [203] M. Bourn, G. Allen, P. Hollingsworth, K. Kababbe, P. I. Williams, H. Ricketts, J. R. Pitt, and A. Shah, “The development of an unmanned aerial system for the measurement of methane emissions from landfill,” *Sixteenth International Waste Management and Landfill Symposium*, no. October 2017, 2018.
- [204] N. Cressie, “Kriging nonstationary data,” *Journal of the American Statistical Association*, vol. 81, no. 395, pp. 625–634, 1986.

- [205] H. Wackernagel, “Ordinary Kriging,” in *Multivariate Geostatistics*, pp. 79–88, Springer, 2003.
- [206] M. Fuentes, “A high frequency kriging approach for non-stationary environmental processes,” *Environmetrics: The official journal of the International Environmetrics Society*, vol. 12, no. 5, pp. 469–483, 2001.
- [207] H. Lebreznz and A. Bárdossy, “Geostatistical interpolation by quantile kriging,” *Hydrology and Earth System Sciences*, vol. 23, no. 3, pp. 1633–1648, 2019.
- [208] G. Y. Lu and D. W. Wong, “An adaptive inverse-distance weighting spatial interpolation technique,” *Computers & Geosciences*, vol. 34, no. 9, pp. 1044–1055, 2008.
- [209] A. Mazzella and A. Mazzella, “The importance of the model choice for experimental semivariogram modeling and its consequence in evaluation process,” *Journal of Engineering*, vol. 2013, 2013.
- [210] P. Goovaerts, “Kriging and semivariogram deconvolution in the presence of irregular geographical units,” *Mathematical Geosciences*, vol. 40, no. 1, pp. 101–128, 2008.
- [211] K. C. Cossel, E. M. Waxman, E. Hoenig, M. Cermak, C. Choate, D. Hesselius, I. Coddington, and N. R. Newbury, “Micrometeorological flux measurements using spatially-scanned open-path dual-comb spectroscopy,” *Optical Sensors and Sensing Congress*, vol. 2020, no. c, p. EM3C.2, 2021.
- [212] J. M. Ryoo, L. T. Iraci, T. Tanaka, J. E. Marrero, E. L. Yates, I. Fung, A. M. Michalak, J. Tadić, W. Gore, T. P. Bui, J. M. Dean-Day, and C. S. Chang, “Quantification of CO₂ and CH₄ emissions over Sacramento, California, based on divergence theorem using aircraft measurements,” *Atmospheric Measurement Techniques*, vol. 12, no. 5, pp. 2949–2966, 2019.
- [213] C. Frankenberg, A. K. Thorpe, D. R. Thompson, G. Hulley, E. A. Kort, N. Vance, J. Borchardt, T. Krings, K. Gerilowski, C. Sweeney, *et al.*, “Airborne methane remote measurements reveal heavy-tail flux distribution in four corners region,” *Proc. of the National Academy of Sciences*, vol. 113, no. 35, pp. 9734–9739, 2016.
- [214] M. B. Frish, “Monitoring fugitive methane emissions utilizing advanced small unmanned aerial sensor technology,” 2016. <http://www.psicorp.com/sites/psicorp.com/files/articles/SR-2018-3.pdf>.
- [215] O. Denmead, L. Harper, J. Freney, D. Griffith, R. Leuning, and R. Sharpe, “A mass balance method for non-intrusive measurements of surface-air trace gas exchange,” *Atmospheric Environment*, vol. 32, no. 21, pp. 3679–3688, 1998.

- [216] Z. Gao, R. L. Desjardins, and T. K. Flesch, "Comparison of a simplified micrometeorological mass difference technique and an inverse dispersion technique for estimating methane emissions from small area sources," *Agricultural and Forest Meteorology*, vol. 149, no. 5, pp. 891–898, 2009.
- [217] S. Conley, I. Faloona, S. Mehrotra, M. Suard, D. H. Lenschow, C. Sweeney, S. Herndon, S. Schwietzke, G. Pétron, J. Pifer, *et al.*, "Application of gauss's theorem to quantify localized surface emissions from airborne measurements of wind and trace gases," *Atmospheric Measurement Techniques*, vol. 10, no. 9, pp. 3345–3358, 2017.
- [218] A. Babilotte, "Field comparison of methods for assessment of methane fugitive emissions from landfills," *Environmental Research and Education Foundation (EREF)*, 2011.
- [219] R. D. P. M. Scafutto and C. R. d. S. Filho, "Detection of methane plumes using airborne midwave infrared (3–5 μm) hyperspectral data," *Remote Sensing*, vol. 10, no. 8, pp. 1–16, 2018.
- [220] S. Dierks and A. Kroll, "Quantification of methane gas leakages using remote sensing and sensor data fusion," in *Proc. of the 2017 IEEE Sensors Applications Symposium (SAS)*, pp. 1–6, IEEE, 2017.
- [221] G. Tanda, M. Balsi, P. Fallavollita, and V. Chiarabini, "A UAV-based thermal-imaging approach for the monitoring of urban landfills," *Inventions*, vol. 5, no. 4, pp. 1–13, 2020.
- [222] M. D. Foote, P. E. Dennison, A. K. Thorpe, D. R. Thompson, S. Jongaramrungruang, C. Frankenberg, and S. C. Joshi, "Fast and accurate retrieval of methane concentration from imaging spectrometer data using sparsity prior," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 9, pp. 6480–6492, 2020.
- [223] B. K. Lamb, J. B. McManus, J. H. Shorter, C. E. Kolb, B. Mosher, R. C. Harriss, E. Allwine, D. Blaha, T. Howard, A. Guenther, *et al.*, "Development of atmospheric tracer methods to measure methane emissions from natural gas facilities and urban areas," *Environmental Science & Technology*, vol. 29, no. 6, pp. 1468–1479, 1995.
- [224] P. Czepiel, B. Mosher, R. Harriss, J. Shorter, J. McManus, C. Kolb, E. Allwine, and B. Lamb, "Landfill methane emissions measured by enclosure and atmospheric tracer methods," *Journal of Geophysical Research: Atmospheres*, vol. 101, no. D11, pp. 16711–16719, 1996.
- [225] P. T. Imhoff and F. K. Chow, "Assessing Accuracy of Tracer Dilution Measurements of Methane Emissions from Landfills with Wind Modeling Research Pro-

- positional Assessing Accuracy of Tracer Dilution Measurements of Methane Emissions from Landfills with Wind Modeling,” 2014.
- [226] P. Imhoff, R. Yazdani, B. Han, C. Mei, and D. Augenstein, “Quantifying capture efficiency of gas collection wells with gas tracers,” *Waste Management*, vol. 43, pp. 319–327, 2015.
- [227] F. B. De la Cruz, R. B. Green, G. R. Hater, J. P. Chanton, E. D. Thoma, T. A. Harvey, and M. A. Barlaz, “Comparison of field measurements to methane emissions models at a new landfill,” *Environmental Science & Technology*, vol. 50, no. 17, pp. 9432–9441, 2016.
- [228] J. Roscioli, T. Yacovitch, C. Floerchinger, A. Mitchell, D. Tkacik, R. Subramanian, D. Martinez, T. Vaughn, L. Williams, D. Zimmerle, *et al.*, “Measurements of methane emissions from natural gas gathering facilities and processing plants: measurement methods,” *Atmospheric Measurement Techniques*, vol. 8, no. 5, pp. 2017–2035, 2015.
- [229] T. A. Foster-Wittig, E. D. Thoma, R. B. Green, G. R. Hater, N. D. Swan, and J. P. Chanton, “Development of a mobile tracer correlation method for assessment of air emissions from landfills and other area sources,” *Atmospheric Environment*, vol. 102, pp. 323–330, 2015.
- [230] B. B. Hicks and D. D. Baldocchi, “Measurement of fluxes over land: Capabilities, origins, and remaining challenges,” *Boundary-Layer Meteorology*, vol. 177, pp. 365–394, 2020.
- [231] M. Göckede, C. Rebmann, and T. Foken, “A combination of quality assessment tools for eddy covariance measurements with footprint modelling for the characterisation of complex sites,” *Agricultural and Forest Meteorology*, vol. 127, no. 3-4, pp. 175–188, 2004.
- [232] G. Burba, *Eddy covariance method for scientific, industrial, agricultural and regulatory applications: A field book on measuring ecosystem gas exchange and areal emission rates*. LI-Cor Biosciences, 2013.
- [233] K. J. Allwine and J. E. Flaherty, “Joint urban 2003: Study overview and instrument locations,” tech. rep., Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2006.
- [234] M. Leach, “Final report for the joint urban 2003 atmospheric dispersion study in Oklahoma City: Lawrence livermore national laboratory participation,” tech. rep., Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States), 2005.
- [235] M. A. Nelson, M. Brown, E. Pardyjak, and J. Klewicki, “Area-averaged profiles over the mock urban setting test array,” tech. rep., Los Alamos National Laboratory, 2004.

- [236] C. A. Biltoft, “Customer report for mock urban setting test,” *DPG Document*, no. 8-CO, pp. 160–000, 2001.
- [237] C. A. Biltoft and E. Yee, “Overview of the Mock Urban Setting Test (MUST) C.A.,” *Engineering*, 2001.
- [238] J. Eichhorn and M. Balczó, “Flow and dispersal simulations of the mock urban setting test,” *Hrvatski meteorološki časopis*, vol. 43, no. 43/1, pp. 67–72, 2008.
- [239] M. A. Nelson, M. J. Brown, S. A. Halverson, P. E. Bieringer, A. Annunzio, G. Bieberbach, and S. Meech, “A case study of the weather research and forecasting model applied to the joint urban 2003 tracer field experiment. part 2: Gas tracer dispersion,” *Boundary-Layer Meteorology*, vol. 161, no. 3, pp. 461–490, 2016.
- [240] D. K. Nicholson, A. Hedrick, P. Serguievski, and A. A. Martinez, “Detailed test plan for Jack Rabbit (JR) II,” tech. rep., West Desert Test Center Dugway Proving Ground UT, 2015.
- [241] Bridger Photonics, “Gas mapping lidar™ METEC round 1 results.”
- [242] C. S. Bell, T. Vaughn, D. Zimmerle, D. Helmig, and B. Lamb, “Evaluation of next generation emission measurement technologies under repeatable test protocols,” *Elementa: Science of the Anthropocene*, vol. 8, 2020.
- [243] E. J. Zhang, C. C. Teng, T. G. van Kessel, L. Klein, R. Muralidhar, G. Wysocki, and W. M. Green, “Field deployment of a portable optical spectrometer for methane fugitive emissions monitoring on oil and gas well pads,” *Sensors*, vol. 19, no. 12, p. 2707, 2019.
- [244] C. B. Alden, S. C. Coburn, R. J. Wright, E. Baumann, K. Cossel, E. Perez, E. Hoenig, K. Prasad, I. Coddington, and G. B. Rieker, “Single-blind quantification of natural gas leaks from 1 km distance using frequency combs,” *Environmental science & technology*, vol. 53, no. 5, pp. 2908–2917, 2019.
- [245] C. B. Alden, S. Ghosh, S. Coburn, C. Sweeney, A. Karion, R. Wright, I. Coddington, G. B. Rieker, and K. Prasad, “Bootstrap inversion technique for atmospheric trace gas source detection and quantification using long open-path laser measurements,” *Atmospheric Measurement Techniques*, vol. 11, no. 3, pp. 1565–1582, 2018.
- [246] A. P. Ravikumar, B. Barlow, J. Wang, and D. Singh, “Results from the Alberta methane measurement campaigns: New insights into oil and gas methane mitigation policy,” in *AGU Fall Meeting Abstracts*, vol. 2019, pp. A41D–08, 2019.

- [247] J. T. Shaw, G. Allen, J. Pitt, A. Shah, S. Wilde, L. Stamford, Z. Fan, H. Ricketts, P. I. Williams, P. Bateson, *et al.*, “Methane flux from flowback operations at a shale gas site,” *Journal of the Air & Waste Management Association*, vol. 70, no. 12, pp. 1324–1339, 2020.
- [248] O. Roberts, “The theoretical scattering of smoke in a turbulent atmosphere,” *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 104, no. 728, pp. 640–654, 1923.
- [249] K. Manies, E. Yates, L. Christensen, M. Fladeland, R. Kolyer, E. Euskirchen, and M. Waldrop, “Can a drone equipped with a miniature methane sensor determine methane fluxes from an Alaskan wetland?,” *Earth and Space Science Open Archive*, 2018.
- [250] D. Baldocchi, M. Detto, O. Sonnentag, J. Verfaillie, Y. A. Teh, W. Silver, and N. M. Kelly, “The challenges of measuring methane fluxes and concentrations over a peatland pasture,” *Agricultural and Forest Meteorology*, vol. 153, pp. 177–187, 2012.
- [251] F. K. Oberle, A. E. Gibbs, B. M. Richmond, L. H. Erikson, M. P. Waldrop, and P. W. Swarzenski, “Towards determining spatial methane distribution on Arctic permafrost bluffs with an unmanned aerial system,” *SN Applied Sciences*, vol. 1, no. 3, pp. 1–9, 2019.
- [252] L. Golston, N. Aubut, M. Frish, S. Yang, R. Talbot, C. Gretencord, J. McSpirt, and M. Zondlo, “Natural gas fugitive leak detection using an unmanned aerial vehicle: Localization and quantification of emission rate,” *Atmosphere*, vol. 9, no. 9, p. 333, 2018.
- [253] T. E. Barchyn, C. H. Hugenholtz, S. Myshak, and J. Bauer, “A UAV-based system for detecting natural gas leaks,” *Journal of Unmanned Vehicle Systems*, vol. 6, no. 1, pp. 18–30, 2017.
- [254] G. Allen, P. Hollingsworth, K. Kabbabe, J. R. Pitt, M. I. Mead, S. Illingworth, G. Roberts, M. Bourn, D. E. Shallcross, and C. J. Percival, “The development and trial of an unmanned aerial system for the measurement of methane flux from landfill and greenhouse gas emission hotspots,” *Waste Management*, 2018.
- [255] “What is a digital twin?,” 2020. <https://www.mathworks.com/discovery/digital-twin.html>.
- [256] “Digital twin: Helping machines tell their story,” 2020. <https://www.ibm.com/topics/what-is-a-digital-twin>.
- [257] “What is a digital twin?,” 2020. <https://www.ge.com/digital/blog/what-digital-twin>.

- [258] “Digitals twins,” 2020. <https://www2.deloitte.com/us/en/insights/focus/tech-trends/2020/digital-twin-applications-bridging-the-physical-and-digital.html#endnote-10>.
- [259] E. Glaessgen and D. Stargel, “The digital twin paradigm for future NASA and US Air Force vehicles,” in *Proc. of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, p. 1818, 2012.
- [260] C. Dufour, Z. Soghomonian, and W. Li, “Hardware-in-the-loop testing of modern on-board power systems using digital twins,” in *Proc. of the 2018 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)*, pp. 118–123, IEEE, 2018.
- [261] F. Tao and Q. Qi, “Make more digital twins,” *Nature*, vol. 573, pp. 490–491, 2019.
- [262] M. Armendia, M. Ghassempouri, E. Ozturk, and F. Peysson, *Twin-Control: A Digital Twin Approach to Improve Machine Tools Lifecycle*. Springer, 2019.
- [263] H.-D. Wuttke, K. Henke, and R. Hutschenreuter, “Digital twins in remote labs,” in *International Conference on Remote Engineering and Virtual Instrumentation*, pp. 289–297, Springer, 2019.
- [264] S. Nikolaev, M. Gusev, D. Padalitsa, E. Mozhenkov, S. Mishin, and I. Uzhinsky, “Implementation of ‘Digital Twin’ concept for modern project-based engineering education,” in *IFIP International Conference on Product Lifecycle Management*, pp. 193–203, Springer, 2018.
- [265] G. Steindl, M. Stagl, L. Kasper, W. Kastner, and R. Hofmann, “Generic digital twin architecture for industrial energy systems,” *Applied Sciences*, vol. 10, no. 24, p. 8903, 2020.
- [266] S. M. Sepasgozar, “Differentiating digital twin from digital shadow: Elucidating a paradigm shift to expedite a smart, sustainable built environment,” *Buildings*, vol. 11, no. 4, p. 151, 2021.
- [267] T. M. Fernández-Caramés, O. Blanco-Novoa, I. Froiz-Míguez, and P. Fraga-Lamas, “Towards an autonomous industry 4.0 warehouse: A UAV and blockchain-based system for inventory and traceability applications in big data-driven supply chain management,” *Sensors*, vol. 19, no. 10, p. 2394, 2019.
- [268] A. Martínez-Gutiérrez, J. Díez-González, R. Ferrero-Guillén, P. Verde, R. Álvarez, and H. Perez, “Digital twin for automatic transportation in industry 4.0,” *Sensors*, vol. 21, no. 10, p. 3344, 2021.

- [269] L. Lei, G. Shen, L. Zhang, and Z. Li, "Toward intelligent cooperation of uav swarms: When machine learning meets digital twin," *IEEE Network*, vol. 35, no. 1, pp. 386–392, 2020.
- [270] D. Guivarch, E. Mermoz, Y. Marino, and M. Sartor, "Creation of helicopter dynamic systems digital twin using multibody simulations," *CIRP Annals*, vol. 68, no. 1, pp. 133–136, 2019.
- [271] X. Li, B. He, Y. Zhou, and G. Li, "Multisource model-driven digital twin system of robotic assembly," *IEEE Systems Journal*, vol. 15, no. 1, pp. 114–123, 2020.
- [272] J. Viola and Y. Chen, "Digital twin enabled smart control engineering as an industrial AI: A new framework and a case study," *arXiv preprint arXiv:2007.03677*, 2020.
- [273] J. Viola, Y. Chen, and J. Wang, "Information-based model discrimination for digital twin behavioral matching," in *2020 2nd International Conference on Industrial Artificial Intelligence (IAI)*, pp. 1–6, IEEE, 2020.
- [274] J. Viola, C. Rodriguez, and Y. Chen, "Phelp: Pixel heating experiment learning platform for education and research on iai-based smart control engineering," in *2020 2nd International Conference on Industrial Artificial Intelligence (IAI)*, pp. 1–6, IEEE, 2020.
- [275] N. MohameD, "Fully implicit scheme for solving Burgers' equation based on finite difference method," *Egyptian Journal for Engineering Sciences and Technology*, vol. 26, pp. 38–44, 2018.
- [276] R. H. Pletcher, J. C. Tannehill, and D. Anderson, *Computational Fluid Mechanics and Heat Transfer*. CRC Press, 2012.
- [277] B. J. Smith, G. John, L. E. Christensen, and Y. Chen, "Fugitive methane leak detection using sUAS and miniature laser spectrometer payload: system, application and groundtruthing tests," in *Proc. of the 2017 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 369–374, IEEE, 2017.
- [278] A. A. Abdel-Rahman, "On the atmospheric dispersion and Gaussian plume model," in *WWAI'08: Proc. of the 2nd International Conference on Waste Management, Water Pollution, Air Pollution, Indoor Climate*, pp. 31–39, 2008.
- [279] G. A. Briggs, "Plume rise: A critical survey.," tech. rep., Air Resources Atmospheric Turbulence and Diffusion Lab., Oak Ridge, Tenn., 1969.
- [280] N. B. Erichson, L. Mathelin, Z. Yao, S. L. Brunton, M. W. Mahoney, and J. N. Kutz, "Shallow neural networks for fluid flow reconstruction with limited sensors," *Proc. of the Royal Society A*, vol. 476, no. 2238, p. 20200097, 2020.

- [281] D. Hollenbeck and Y. Chen, “Mutli-uav method for continuous source rate estimation of fugitive gas emissions from a point source,” in *Proc. of the 2021 International Conference on Unmanned Aircraft Systems (ICUAS)*, IEEE, 2021.
- [282] J. F. Newman and P. M. Klein, “The impacts of atmospheric stability on the accuracy of wind speed extrapolation methods,” *Resources*, vol. 3, no. 1, pp. 81–105, 2014.
- [283] O. Babak and C. V. Deutsch, “Statistical approach to inverse distance interpolation,” *Stochastic Environmental Research and Risk Assessment*, vol. 23, no. 5, pp. 543–553, 2009.
- [284] M. Reggente and A. J. Lilienthal, “The 3D-kernel DM+V/W algorithm: Using wind information in three dimensional gas distribution modelling with a mobile robot,” in *SENSORS, 2010 IEEE*, pp. 999–1004, IEEE, 2010.
- [285] X. He, J. A. Steiner, J. R. Bourne, and K. K. Leang, “Gaussian-based kernel for multi-agent aerial chemical-plume mapping,” in *Proc. of the Dynamic Systems and Control Conference*, vol. 59162, p. V003T21A004, American Society of Mechanical Engineers, 2019.
- [286] S. Asadi, H. Fan, V. H. Bennetts, and A. J. Lilienthal, “Time-dependent gas distribution modelling,” *Robotics and Autonomous Systems*, vol. 96, pp. 157–170, 2017.
- [287] Y. Xu, J. Choi, S. Dass, and T. Maiti, *Bayesian prediction and adaptive sampling algorithms for mobile sensor networks: Online environmental field reconstruction in space and time*. Springer, 2015.
- [288] R. L. Magin, “Fractional calculus models of complex dynamics in biological tissues,” *Computers & Mathematics with Applications*, vol. 59, no. 5, pp. 1586–1593, 2010.
- [289] A. Shukla and J. Prajapati, “On a generalization of Mittag-Leffler function and its properties,” *Journal of Mathematical Analysis and Applications*, vol. 336, no. 2, pp. 797–811, 2007.
- [290] J. Viola, D. Hollenbeck, C. Rodriguez, and Y. Chen, “Fractional-order stochastic extremum seeking control with dithering noise for plasma impedance matching,” in *Proceedings of the 2021 Conference on Control Technology and Applications*, IEEE, 2021.
- [291] J. Cao, Y. Chen, and C. Li, “Multi-uav-based optimal crop-dusting of anomalously diffusing infestation of crops,” in *Proc. of the 2015 American Control Conference (ACC)*, pp. 1278–1283, IEEE, 2015.

- [292] Y. Chen, Z. Wang, and J. Liang, "Optimal dynamic actuator location in distributed feedback control of a diffusion process," *International Journal of Sensor Networks*, vol. 2, no. 3-4, pp. 169–178, 2007.
- [293] Y. Chen, Z. Wang, and K. Moore, "Optimal spraying control of a diffusion process using mobile actuator networks with fractional potential field based dynamic obstacle avoidance," in *Proc. of the 2006 IEEE International Conference on Networking, Sensing and Control*, pp. 107–112, 2006.
- [294] Y. Chen, Z. Wang, and J. Liang, "Actuation scheduling in mobile actuator networks for spatial-temporal feedback control of a diffusion process with dynamic obstacle avoidance," in *IEEE International Conference Mechatronics and Automation, 2005*, vol. 2, pp. 752–757 Vol. 2, 2005.
- [295] G. A. Terejanu, "Discrete kalman filter tutorial," *University at Buffalo, Department of Computer Science and Engineering, NY*, vol. 14260, 2013.
- [296] G. Welch, G. Bishop, *et al.*, "An introduction to the kalman filter," 1995.
- [297] R. Andrade-Pacheco, F. Rerolle, J. Lemoine, L. Hernandez, A. Meité, L. Juziwelo, A. F. Bibaut, M. J. van der Laan, B. F. Arnold, and H. J. Sturrock, "Finding hotspots: development of an adaptive spatial sampling approach," *Scientific Reports*, vol. 10, no. 1, pp. 1–12, 2020.
- [298] S. L. Brunton, C. W. Rowley, S. R. Kulkarni, and C. Clarkson, "Maximum power point tracking for photovoltaic optimization using ripple-based extremum seeking control," *IEEE transactions on power electronics*, vol. 25, no. 10, pp. 2531–2540, 2010.
- [299] J. Viola, D. Hollenbeck, C. Rodriguez, and Y. Chen, "Fractional-order stochastic extremum seeking control with dithering noise for plasma impedance matching," in *Proceedings of the 2021 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 247–252, IEEE, 2021.
- [300] A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*. V.H. Winston & Sons, 1977.
- [301] A. N. Tikhonov, A. Goncharsky, V. Stepanov, and A. G. Yagola, *Numerical methods for the solution of ill-posed problems*, vol. 328. Springer Science & Business Media, 1995.
- [302] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2019.
- [303] D. Uciński and M. Patan, "Sensor network design for the estimation of spatially distributed processes," 2010.

- [304] C. Tricaud and Y. Chen, *Optimal mobile sensing and actuation policies in cyber-physical systems*. Springer Science & Business Media, 2011.
- [305] B. Porat and A. Nehorai, “Localizing vapor-emitting sources by moving sensors,” *IEEE Transactions on Signal processing*, vol. 44, no. 4, pp. 1018–1021, 1996.
- [306] B. T. Hinson and K. A. Morgansen, “Gyroscopic sensing in the wings of the hawkmoth *manduca sexta*: the role of sensor location and directional sensitivity,” *Bioinspiration & Biomimetics*, vol. 10, no. 5, p. 056013, 2015.
- [307] J. C. Spall, “Multivariate stochastic approximation using a simultaneous perturbation gradient approximation,” *IEEE transactions on automatic control*, vol. 37, no. 3, pp. 332–341, 1992.
- [308] J. Viola and Y. Chen, *Digital-Twin-Enabled Smart Control Engineering: A Framework and Case Studies*. Springer Nature, 2023.
- [309] C. Himpe, “emgr—the empirical gramian framework,” *Algorithms*, vol. 11, no. 7, p. 91, 2018.
- [310] A. J. Krener and K. Ide, “Measures of unobservability,” in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pp. 6401–6406, IEEE, 2009.
- [311] B. T. Hinson, *Observability-based guidance and sensor placement*. PhD thesis, 2014.
- [312] F. L. Cortesi, T. H. Summers, and J. Lygeros, “Submodularity of energy related controllability metrics,” in *53rd IEEE conference on decision and control*, pp. 2883–2888, IEEE, 2014.
- [313] E.-W. Bai and S. S. Sastry, “Persistency of excitation, sufficient richness and parameter convergence in discrete time adaptive control,” *Systems & control letters*, vol. 6, no. 3, pp. 153–163, 1985.
- [314] N. Nordström and S. S. Sastry, “Persistency of excitation in possibly unstable continuous time systems and parameter convergence in adaptive identification,” in *Adaptive Systems in Control and Signal Processing 1986*, pp. 347–352, Elsevier, 1987.
- [315] M. A. Demetriou, “Inducing persistence of excitation through sensor motion in the adaptive estimation of spatial fields,” in *2022 American Control Conference (ACC)*, pp. 1673–1678, IEEE, 2022.
- [316] R. Abdul Razak, S. Sukumar, and H. Chung, “Scalar field estimation with mobile sensor networks,” *International Journal of Robust and Nonlinear Control*, vol. 31, no. 9, pp. 4287–4305, 2021.

-
- [317] D. Hollenbeck, D. Zulevic, and Y. Chen, “A modified near-field gaussian plume inversion method using multi-suas for emission quantification,” in *2022 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1620–1625, IEEE, 2022.