

Adjoint-based Optimization Scheme for Characterizing Sources of Fugitive Emissions

by

Carol Anne Brereton

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mechanical Engineering

Carleton University

Ottawa, Ontario

© 2018

Carol Anne Brereton

Abstract

Fugitive emissions, such as leaking components, are important sources of methane and other pollutants within the oil and gas sector. These gas emission sources are difficult to mitigate because their existence, location, and magnitude are unknown, especially within facilities which may have thousands of potential source components. Improving the speed at which these sources are characterized (located and quantified) can improve these mitigation efforts.

Using sparse concentration information (from finite sensor locations) and estimated wind fields, a scalar transport adjoint-based optimization approach was developed to characterize both single and multiple simultaneous gas releases representative of the fugitive emissions problem. This approach was tested using data from an open-field gas release to determine a single source location within 5 m and source magnitude within 13%. Simulated simultaneous releases over a complex 3D geometry based on an Alberta gas plant were also characterized using both detailed transient wind and wind fields approximated with a series of steady-state wind simulations more readily implemented in a field application. Magnitudes were predicted within 10% and major regions located.

By extending the method with a database of pre-computed retro-tracers generated on simplified steady-state wind fields, the required computational time for the solution optimization was reduced by a factor of 200-600, making computations feasible on a desktop machine and raising the possibility of near-continuous fugitive emissions quantification in the future. Major sources were successfully located, and magnitudes estimated within -75 to -32% , even with limited wind coverage (60° direction variation).

Turbulent Schmidt number selection (which scales diffusivity) had little effect on estimated source locations, but strongly influenced estimated magnitudes. For a tested turbulent Schmidt number range of 0.33 (diffuse) to 2.0 (highly advective), the predicted emission rates for the open field release varied between -35 to $+128\%$ of actual. Buildings dampened this effect, suggesting that open-field estimates can act as an error bound.

Preface

This thesis contains chapters which have already been published or have been prepared for publication jointly as journal articles.

Chapter 4 has been peer-reviewed and was published in *Atmospheric Environment* and reproduced according to the journal's copyright policy. Only minor clarifications and formatting edits have been made following feedback from the thesis examination committee and for formatting consistency. This paper was co-authored by the thesis author, Carol Brereton, her thesis supervisors Prof. Matthew Johnson and Prof. Lucy Campbell, as well as former student Ian Joynes who was involved in initial discussions and background development work. Ms. Brereton set-up, developed, and tested the relevant procedures and code; generated synthetic test data; performed the test simulations and initial data analysis; and wrote the first draft of the manuscript. Theory development, final data analysis, and editing and revision of the manuscript were conducted jointly. The paper should be cited as:

Brereton, C. A., Joynes, I., Campbell, L. J., Johnson, M. R., 2018. Fugitive Emission Source Characterization Using a Gradient-Based Optimization Scheme and Scalar Transport Adjoint. *Atmos. Environ.*, 181, 106-116 (doi: 10.1016/j.atmosenv.2018.02.014).

Chapter 5 and Chapter 6 have been prepared for future submission to *Atmospheric Environment*. These papers were co-authored by the thesis author, Carol Brereton, and her thesis supervisors Prof. Matthew Johnson and Prof. Lucy Campbell. Ms. Brereton set-up, developed, and tested the relevant procedures and code; generated synthetic test data; performed the relevant test simulations and initial data analysis; and wrote the first draft of the manuscript. Theory development, final data analysis, and editing and revision of the manuscript were conducted jointly. The following paper citations are preferred, and this thesis should be cited only if the appropriate citation for the included articles cannot be found.

Brereton, C. A., Campbell, L. J., Johnson, M. R., 2018. “Computationally Efficient Quantification of Unknown Fugitive Emissions Sources”, *draft manuscript to be submitted to Atmospheric Environment*.

Brereton, C. A., Campbell, L. J., Johnson, M. R., 2018. “Influence of Turbulent Schmidt Number on Fugitive Emissions Source Quantification”, *draft manuscript to be submitted to Atmospheric Environment*.

Statement of Student Involvement

I, Carol Brereton, was fully involved and responsible for developing test cases, procedures, code, obtaining data, analyzing results, and writing first drafts of the material presented in the accompanying thesis.

Carol A Brereton

The student, Carol Brereton, was fully involved and responsible for developing test cases, procedures, code, obtaining data, analyzing results, and writing first drafts of the material presented in the accompanying thesis.

Matthew R Johnson

Table of Contents

Abstract	ii
Preface	iii
List of Tables	ix
List of Figures	x
Nomenclature	xv
Chapter 1 Introduction.....	18
1.1 Motivation	18
1.2 Thesis Objectives	19
1.3 Thesis Overview.....	19
Chapter 2 Fugitive Emissions in the Oil and Gas Industry	21
Chapter 3 Fugitive Source Characterization from Sparse Concentration Sensors	27
3.1 Gas Transport Modelling	27
3.2 Types of Source Characterization Problems	28
3.2.1 Single Source Estimation Problems.....	28
3.2.2 Multi-Source Estimation Problems.....	29
3.2.3 Other Types of Source Characterization Problems.....	29
3.3 Approaches to Solving Source Characterization Problems.....	30
3.3.1 Trajectory Statistical Methods	30
3.3.2 Sequential Data Assimilation Methods.....	30
3.3.3 Variational Data Assimilation.....	34
3.4 Target Application and Proposed Approach	37
3.4.1 Dispersion Model Selection.....	38
3.4.2 Source Definition	39

3.4.3	Optimization Algorithm.....	39
3.4.4	Cost Function.....	40
3.4.5	Cost Function Gradient.....	40
Chapter 4 Fugitive Emission Source Characterization Using a Gradient-Based		
	Optimization Scheme and Scalar Transport Adjoint	41
4.1	Abstract	41
4.2	Introduction	42
4.3	Adjoint-based optimization methodology.....	45
4.3.1	3D Wind Field Data.....	46
4.3.2	Source Guess.....	47
4.3.3	Forward Scalar Transport Model.....	47
4.3.4	Cost Function.....	48
4.3.5	Adjoint Sensitivity and Cost Function Gradient.....	49
4.4	Implementation.....	52
4.5	Results	52
4.5.1	Project Prairie Grass Test Case.....	53
4.5.2	PPG Open Field Reconstruction Results	55
4.5.3	Simulated Gas Plant Test Case	57
4.5.4	Simulated Gas Plant Case Results	62
4.6	Conclusions	67
4.7	Acknowledgements	67
Chapter 5 Computationally Efficient Quantification of Unknown Fugitive Emissions		
	Sources	69
5.1	Abstract	69
5.2	Introduction	70

5.3	Fugitive Emissions	71
5.4	Emission Characterization Methodology	72
5.4.1	Concurrent-Transport-Computation (CTC) Approach	72
5.4.2	Pre-computed Retro-Tracer (PRT) Approach.....	74
5.4.3	Comparison of Approaches and Implementation	76
5.5	Results	79
5.5.1	Open Field Test Case	79
5.5.2	Bluff Body Test Case Based on an Alberta Gas Plant Geometry.....	83
5.6	Conclusions	92
5.7	Acknowledgements	93
Chapter 6 Influence of Turbulent Schmidt Number on Fugitive Emissions Source		
	Quantification	94
6.1	Abstract	94
6.2	Introduction	95
6.3	Turbulent Schmidt Number and Scalar Transport Modelling.....	97
6.4	Emission Characterization.....	100
6.4.1	Optimization procedure to determine unknown source location(s) with unknown emission rate(s)	101
6.4.2	Optimization procedure assuming known source location(s) but unknown emission rate(s)	102
6.5	Transport Solver Implementation.....	103
6.6	Results	103
6.6.1	Open Field Case.....	104
6.6.2	Partial Gas Plant (Bluff Body) Test Case	113
6.7	Conclusions	119
6.8	Acknowledgements	120

Chapter 7 Conclusions and Future Work	122
7.1 Conclusions	122
7.2 Major Contributions	123
7.3 Future Work	125
References.....	126
Appendix A: Plume and Gas Dispersion	139
Appendix B: L-BFGS Optimization Scheme	145
Appendix C: Wind Profiles for Varying Atmospheric Stability.....	149
Appendix D: Project Prairie Grass Mesh Refinement	152
Appendix E: Pre-convergence Location Predictions for Project Prairie Grass	154
Appendix F: Optimization from Different Initial Conditions.....	155

List of Tables

Table 3.1: Summary of major source characterization methods.....	37
Table 4.1: PPG Mesh Statistics.....	54
Table 4.2: Mesh details for the simulated gas plant test case	59
Table 5.3: Number of measurement intervals available for comparison for each gas plant data set	86
Table A.1: Typical Gaussian plume model property representation (Lagzi et al., 2013)	142
Table C.1: Obukhov length and atmospheric stability (Alinot, 2003).....	150
Table C.2: Expressions of ϕm for different atmospheric conditions (Alinot, 2003).....	150
Table C.3: Expressions of Ψm for different atmospheric conditions (Alinot, 2003)	151
Table C.4: Expressions of $\phi \epsilon$ for different atmospheric conditions (Alinot, 2003)	151
Table D.1: Mesh spacing compared with Medium Mesh.....	152
Table F.1: Actual Release Information and Tested Initial Conditions for Project Prairie Grass	155

List of Figures

Figure 2.1: Methane emissions from upstream oil and gas production in Alberta, Canada in 2014 (based on Johnson et al., 2017)..... 22

Figure 4.1: Simplified optimization steps with corresponding paper section numbers and major variable names. Source parameters refer to emission rate per cell for this paper. 46

Figure 4.2: PPG domain with source (star) and sensor locations for 1.5 m height (small dark circles) and towers (unfilled boxes)..... 55

Figure 4.3: Cost function and mass flow rate convergence 56

Figure 4.4: Cell emission rate predictions for various iterations at 0.5 m, 1.5 m, and 3 m heights from the ground (lines represent 50 m horizontal increments) 57

Figure 4.5: Building a) full domain with inner domain cutout marking large (L) and small (S) source locations and sensors (black dots), b) mesh detail..... 59

Figure 4.6: Wind at 10 m input height..... 60

Figure 4.7: Inlet wind at 10 m height for SSWS case..... 62

Figure 4.8: Predicted emissions at 1 m height using 30-sec transient ideal wind data with marked actual release locations (white circles) and 50-m sensor grid (white points) for a) linear scale, b) log scale, c) linear scale with regional filter, and d) log scale with regional filter. Emission centroids are marked with black dots..... 64

Figure 4.9: Predicted emissions at 1 m height using the series of steady wind simulations (SSWS) and averaged sensor data with marked actual release locations (white circles) and 50-m sensor grid (white points) for a) linear scale, b) log scale, c) linear scale with regional filter, and d) log scale with regional filter. Emission centroids are marked with black dots. 65

Figure 4.10: Comparison of normalized emissions per second for predicted and actual major source regions for the actual, ideal, and series of steady wind simulations (SSWS) 66

Figure 5.1: Simplified optimization steps with major variables using a) concurrent transport computations (CTC) (modified from Brereton et al., 2018a), and b) pre-computed database of retro-plumes	78
Figure 5.2: Emission rate predictions at 0.5 m vertical height using sensor concentrations and cost function gradient computed with a) the CTC method (Brereton et al., 2018a), and b) PRT method	81
Figure 5.3: a) Convergence of the predicted emission rate (normalized by the actual release value) and cost function (normalized by the cost function for no sources) for the CTC (black) and PRT (blue) methods. b) Equivalent single processor computational time for the CTC and PRT methods. Potentially reusable computations are shown in blue diagonal hatching and non-reusable computations are shown in red cross-hatching.	82
Figure 5.4: Wind rose at 10 m height ahead of main building cluster for a) original transient wind directions, b) original transient plus additional wind directions, c) steady wind intervals for original directions, and d) steady wind intervals with additional directions.....	84
Figure 5.5: a) Top view of inner 3D building domain marking large (L) and small (S) sources (filled circles), wind towers (T) locations (unfilled circles), and 50 m spaced sensor locations (dots) and b) mesh detail (Brereton et al., 2018a).....	87
Figure 5.6: Emission rate predictions at 1.5 m slice for a) the SSWS case using the CTC method, b) the original wind case using the PRT method, and c) the PRT method with modest increase in wind directional coverage. Source locations (white circles), sensor locations (white dots) and major source prediction region centroids (black dots) are marked.....	89
Figure 5.7: Emission rate predictions for sources L1 and L2 for various methods and data sets. Specific source points P1, P2, P3, P4 defined in previous figure.....	91
Figure 5.8: Computational time using a single processor for CTC method (Brereton et al., 2018a) using a full transient wind field, CTC method using a series of steady state RANS wind simulations (SSWS) consecutively placed in 30 s intervals to approximate the transient flow, and PRT method (original wind directions).	92

Figure 6.1: Simplified optimization steps with major variables (modified from Brereton et al., 2018a).....	101
Figure 6.2: PPG Domain with source (star), sensor (dot) and tower (unfilled square) locations (Brereton et al., 2018a).....	105
Figure 6.3: Isolated Sc_t effects on predicted emission rate locations at 1.5 m height (main sensor level) for extremes of Sc_t mismatch. a) Synthetic data set generated with $Sc_{t,actual} = 2.0$ while using an assumed $Sc_{t,assumed} = 0.33$ during the source characterization optimization, and b) a synthetic data set generated with a $Sc_{t,actual} = 0.33$ while using $Sc_{t,assumed} = 2.0$ during the source optimization. The actual location of the release is marked by white circle.....	107
Figure 6.4: Isolated Sc_t effects on normalized predicted emission rate with unknown and unconstrained location.	108
Figure 6.5: Isolated Sc_t effects on normalized predicted emission rate from simulated annealing procedure with a) known point source location, and b) predicted point source location based on unconstrained location prediction.	110
Figure 6.6: Combined effects of Sc_t mismatch and measurement and wind modelling error on the predicted source location in the Project Prairie Grass data set. The images show the 400 m \times 300 m domain at 0.5 m vertical height for with $Sc_{t,assumed}$ of a) 0.33, b) 0.67 and c) 1.33. Actual location of release is marked by a white circle.	111
Figure 6.7: Normalized prediction of source emission rate in PPG experimental test data versus reciprocal of assumed turbulent Schmidt number. Curves show results assuming either an unknown location, known exact location, or fixed location based on predicted centroid from unknown location method.....	112
Figure 6.8: a) Top view of 3D building domain marking large (L) and small (S) sources (circles) and sensor locations (dots) and b) mesh detail (Brereton et al., 2018b).....	115
Figure 6.9: Predicted emission rate locations at 1.5 m height for $Sc_{t,assumed}$ of 1.0. (a-c) observations were generated on the same wind fields (isolating the effects Sc_t mismatch) with $Sc_{t,actual}$ of a) 0.33, b) 1.0, c) 2.0. (d-f) observations generated detailed transient wind fields (combining effects of Sc_t mismatch and wind and sensor error) with $Sc_{t,actual}$	

of d) 0.33, e) 1.0, and f) 2.0. Actual location of released marked by circles. Horizontal centroid position of important (>5% total) emission regions are marked with black dots. 117

Figure 6.10: a) Isolated effects of Sc_t , and b) combined effects of Sc_t and wind error on source location accuracy (distance from predicted centroid to actual release location) and spread of predicted source (emission-weighted average spread from predicted centroid). 118

Figure 6.11: Normalized emission rate versus turbulent Schmidt number for the unknown (unconstrained), known, and predicted locations using $Sc_{t,assumed} = 1.0$ for a) wind fields matching between model and observations, and b) wind field approximation..... 119

Figure E.1: Project Prairie Grass emission rate prediction at a height of 0.5 m for various convergence iterations. Actual release location is marked by white circle. 154

Figure F.1: Project Prairie Grass emission rate prediction with an initial guess of a 10× larger source, 5 m above the actual source. Actual release location is marked by white circle. Results after 3502 iterations showing (a) predicted sources at 0.5 m elevation (elevation of the actual source) and (b) predicted sources at 5 m elevation (elevation of initial guess source), with enlarged image of region within white circle illustrating a source predicted at initial guess location. 156

Figure F.2: Predicted emission rates and convergence for PPG with an initial guess of a 10× larger source, 5 m above the actual source. 157

Figure F.3: Project Prairie Grass emission rate prediction with an initial guess of a horizontally offset source at same elevation (-25 m 0 m 0.5 m) and similar magnitude to the actual source. Actual source location is marked by white circle. Results showing predicted sources at 0.5 m elevation after 2000 iterations. 158

Figure F.4: Predicted emission rates and convergence for PPG with an initial guess of a horizontally offset source at same elevation (-25 m 0 m 0.5 m) and similar magnitude to the actual source. 158

Figure F.5: Project Prairie Grass emission rate prediction with an initial guess of a 10× larger source, 50 m above the actual source. Actual release location is marked by white circle. Results shown at 0.5 m elevation after 216 iterations..... 159

Figure F.6: Predicted emission rates and convergence for PPG with an initial guess of a 10× larger source, 50 m above the actual source. Black lines and symbols show the zero initial source guess case of Chapter 4 for comparison. Emission rate of the initial guess location (blue triangle) remains effectively unchanged, while predict source near actual iterates toward correct answer. 160

Figure F.7: Project Prairie Grass emission rate prediction with an initial guess of a ~10 m×10 m×10 m volumetric source 15 m north and 4.5 m above the actual source. Results shown at a height of 0.5 m for iteration 191. Actual release location is marked by white circle..... 161

Figure F.8: Predicted emission rates and convergence for PPG with an initial guess of ~10 m×10 m×10 m volumetric source 15 m north and 4.5 m above the actual source.. 161

Nomenclature

<i>Symbol</i>	<i>Description</i>	<i>First Equation Usage</i>
C	Time-averaged concentration [mass/volume]	4.1
C^*	Retro-tracer concentration	5.6
C^{obs}	Concentration observations [mass/volume]	4.2
D	Turbulent diffusion coefficient (matrix) [length ² /time]	4.1
\mathbf{D}	Covariance matrix of Errors in Emission Estimates	3.2
\mathbf{G}	Gain matrix	3.2
g	Misfit	4.2
J	Cost function	4.3
\mathbf{M}	Measurement observations	3.1
\mathbf{n}	Outward normal unit vector	4.7
\mathbf{N}	Noise Matrix	3.2
p	Probability	3.1
\mathbf{P}	Partial Derivative Matrix	3.2
\mathbf{p}	Estimation parameters	4.2
S	Sources/Sinks [mass/volume/time]	4.1
\mathbf{S}	Source state vector	3.1
Sc_t	Turbulent Schmidt Number	6.2
t	Time [time]	4.1
\mathbf{u}	Time-averaged velocity vector [length/time]	4.1
\mathbf{x}	Position	4.2
x, y, z	Cartesian coordinates [length]	3.7
Γ	Control surface boundaries	4.7
λ	Adjoint variable	4.5
Ω	Volumetric domain	4.3
ν_t	Eddy viscosity [length ² /time]	6.2
μ	Measurement distribution	5.6

Acronyms

ABC	Approximate Bayesian Computation
AWP	Alternate Work Practice
BFGS	Broyden-Fletcher-Goldfarb-Shanno
CAPP	Canadian Association of Petroleum Producers
CFD	Computational Fluid Dynamics
CTC	Concurrent Transport Computations
ECCC	Environment and Climate Change Canada
GHG	Greenhouse Gas
GHGRP	Greenhouse Gas Reporting Program
GPS	Global Positioning System
GWP	Global Warming Potential
IR	Infrared
L	Large
L-BFGS	Limited-Memory Broyden-Fletcher-Goldfarb-Shanno
LDAR	Leak Detection and Repair
MCMC	Monte Carlo Markov-Chaining
OGI	Optical Gas Imaging
P	Point
PPG	Project Prairie Grass
PRT	Pre-computed retro-tracer
RANS	Reynolds-Averaged Navier Stokes
S	Small
SSWS	Series of Steady Wind Simulations
T	Tower

TSM	Trajectory Statistical Method
UOG	Upstream Oil and Gas
URANS	Unsteady Reynolds-Averaged Navier Stokes
USEPA	U. S. Environmental Protection Agency
VOC	Volatile Organic Compound

Chapter 1 Introduction

1.1 Motivation

Fugitive emissions in the form of “unmonitored, unintended and/or uncontrolled” gas releases (Brereton & Johnson, 2012a) are important sources of atmospheric greenhouse gases and volatile organic compounds. Official inventory estimates suggest that fugitive methane from the oil and gas sector makes up 25-33% of methane emissions in the United States and Canada (USEPA, 2016a; ECCC, 2017). Given the importance of methane as a short-term climate forcer (with global warming potential 34/96 times worse than CO₂ on a 100/20-year time horizon, Gasser et al., 2017), the reduction of methane emissions is an important part of any climate change mitigation strategy. Additionally, fugitive methane emissions usually represent lost product and can result in safety hazards, giving companies additional incentives to find and repair equipment responsible for such releases. Typical sources for fugitive emissions in the oil and gas sector are leaking valves, seals, fittings; evaporation losses; and process faults (Carras, 2006). These are often economic to mitigate once they are known, however, the largest challenge to mitigating fugitive emission sources is the difficulty in locating and quantifying them. A typical plant may have miles of piping and thousands of potential release locations, some in hazardous locations. Improved methods for fugitive emission location and quantification could unlock significant environmental benefit. Additionally, locating and quantifying sources (or source term estimation) is applicable to a large number of other industries and problems. Characterizing releases from sparse concentration measurements is useful in applications as diverse as agricultural emissions determination (Wilson et al., 2012) to global inversion of atmospheric CO₂ data to estimate emissions (Liu et al., 2016). Further applications include using limited concentration information to determine the source and/or rate of releases of hazardous material from nuclear reactor accidents (Zhang et al., 2014), nuclear weapons tests for nuclear treaty compliance (De Muetter et al., 2017), or terrorism incidents (Norige et al., 2009).

1.2 Thesis Objectives

This thesis focusses on the location and quantification of fugitive emissions within the upstream oil and gas industry. Reducing the time to locate sources of emissions and scaling their release rates to prioritize repair could produce significant greenhouse gas reductions, cost savings, and safety benefits. This is approached as a source term estimation problem using modelled gas releases, assuming that sparse concentration measurements, general wind information (from anemometers), and facility geometry are available. Although fugitive methane emissions are considered the most significant target for this application, other gases could readily be considered if concentration measurements can distinguish them from background concentration levels.

The objectives of this thesis are to:

- research an adjoint-based optimization method capable of quantifying multiple unknown simultaneous sources with complex 3D domains,
- develop and implement the method code using open source tools,
- explore the robustness of the method and ability to use simplified wind fields as input,
- research potential for using pre-computed information to reduce computational cost,
- investigate the effects of turbulent Schmidt number selection for the developed technique, and
- assess the potential for using the developed techniques for a real-world application.

1.3 Thesis Overview

The focus of this thesis is proof of concept and robustness testing of an adjoint-based inverse method of locating and quantifying emission sources from sparse networks of concentration sensors on the scale of an oil and gas facility.

Chapter 2 gives a summary of fugitive emissions from the oil and gas industry. Current estimations of emission importance, emission source distribution, and current detection and measurement technology are overviewed.

Chapter 3 summarizes relevant approaches to characterizing sources from sparse concentration sensor networks before proposing the approach investigated in this thesis specifically aimed at fugitive emission source location and quantification within potentially complex bluff-body dominated oil and gas facilities.

Chapter 4 consists of a paper outlining an adjoint-based optimization method for determining multiple source locations and magnitudes simultaneously from sparse sensor network information, as well as a demonstration on multiple proof-of-concept test cases.

Chapter 5 consists of a paper that outlines and demonstrates an approach to achieve significant computational resource savings using pre-computed inverse scalar transport solutions. Compared with full gas transport solutions, this approach could make the method practical to use in the field.

Chapter 6 consists of a paper investigating the influence of turbulent Schmidt number selection on the ability to accurately locate and quantify unknown fugitive sources.

Chapter 7 summarizes the main conclusions and outlines potential avenues for future work.

Chapter 2 Fugitive Emissions in the Oil and Gas Industry

Methane is a potent greenhouse gas with a global warming potential (GWP) 34/96 times stronger than CO₂ over a 100/20 year period (Gasser et al., 2017). Given methane's potency as a climate forcer in the near-term, methane emission mitigation becomes an important short-term climate change strategy in conjunction with longer-term CO₂ reductions (Shoemaker et al., 2013). Within Canada, the oil and gas sector is the largest contributor to methane emissions, totaling 1728 kt (~40% of the country total) in 2015 (ECCC, 2017). Approximately two thirds of these emissions are from the province of Alberta, one quarter from Saskatchewan, and the remaining ~9% from other provinces.

Across the upstream (or exploration and production) oil and gas sector, there are three main emission source categories of methane emissions (Lyon, 2016a):

1. vented (expected and deliberate),
2. incomplete combustion, and
3. fugitive (unexpected and unintentional) emissions.

Focusing on the upstream oil and gas (UOG) sector of Alberta, fugitive emissions are the largest source of methane emissions, accounting for an estimated 39% of total UOG methane (Johnson et al., 2017). An additional 31% comes from unreported venting sources (Johnson et al., 2017) (e.g. instrument vent gas, compressor start gas, purge gas discharged to atmosphere, dehydrator still column off-gas etc., Clearstone Engineering Ltd., 2014a), which in general are poorly estimated.

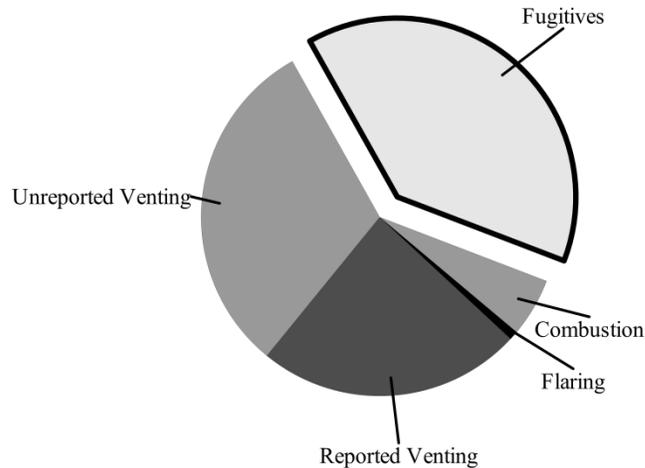


Figure 2.1: Methane emissions from upstream oil and gas production in Alberta, Canada in 2014 (based on Johnson et al., 2017)

Fugitive emissions can generally be thought of as unintentional losses of gas from sources such as leaking components. Within the upstream oil and gas industry, such leaks can arise from (Clearstone Engineering Ltd, 2014b)

- normal wear and tear,
- improper/incomplete component assembly,
- inadequate material specification,
- manufacturing defects,
- damage (from installation or use),
- corrosion,
- fouling, or
- environmental effects (e.g. vibrations, temperature changes).

Individual sources of fugitive emissions are often located within facilities with hundreds or thousands of components. Finding and quantifying leaks for repair prioritization is usually more costly than fixing the leaks themselves (Brandt et al., 2016).

Traditionally, facility surveys are performed with manual component inspection or qualitative optical gas imaging (OGI). The U.S. EPA’s original Method 21 mandates the use of a “portable instrument used to detect VOC [volatile organic compound] leaks from individual sources” (USEPA, 2017) and is intended to locate and classify leaks. A leak threshold of 10,000 ppm is defined for methane. In the original document, some suggested

detector types include catalytic oxidation, flame ionization, infrared absorption, and photo ionization. Leaks may also be found by applying soap solution and checking for bubbles under certain component conditions. This process is very labour intensive, and potentially hazardous, leading to approximately 20% of components not being checked on a regular basis (Chambers, 2006).

An alternative working practice allowing OGI (Method 21 AWP) was proposed in 2006 (USEPA, 2006) and formally added in 2008. This increases the screening speed significantly, though is qualitative. The most common type of OGI used for methane detection is with IR (infrared) cameras as methane and other hydrocarbons absorb at this spectrum (Ravikumar et al., 2017). The cameras themselves are expensive (~100K USD) (Ravikumar et al., 2017) and require trained operators. Thus, surveys are also performed intermittently. Overall, OGI finds an order of magnitude fewer leaking components than the original Method 21 screening (Ravikumar & Brandt, 2017), but is more effective for large emitters (Ravikumar et al., 2017).

IR camera leak detection effectiveness can be influenced by (Ravikumar & Brandt, 2017):

- environmental conditions,
- operator practices,
- underlying leak distribution, and
- gas composition.

High wind adversely affects OGI surveys and, especially important for Canadian climates, effectiveness drops abruptly at freezing temperatures (Ravikumar & Brandt, 2017).

Besides traditional survey methods, in recent years more work has been done on mobile surveys, though these are typically best suited for screening emissions from entire facilities or regions rather than individual leaking components. Vehicles equipped with cavity-ring down spectrometers have been used to map natural gas pipeline leaks (e.g. Jackson et al., 2014; Phillips et al., 2013). Facility-level methane emissions with measurements in a mobile laboratory have been quantified using downwind tracer flux measurements, before being combined with onsite IR cameras to identify emission points in natural gas gathering and processing facilities (e.g. Mitchell et al., 2015). Well pad total emissions have been

quantified using a vehicle equipped with a methane analyzer, anemometer, and GPS (e.g. Rella et al., 2015). Similarly, a vehicle-based campaign in a region with unconventional natural gas development subsequently linked emissions to nearby facilities (Atherton et al., 2017). Airborne measurement studies to locate and quantify emission sources, using both emission mapping techniques and mass balance based on wind and multiple airborne measurements have also been used for regions, and/or facilities (e.g. Thompson et al., 2015; Thorpe et al., 2016; Tratt et al., 2014; Karion et al., 2013; Lavoie et al., 2015; Johnson et al., 2017). While these surveys are useful to estimate total regional emissions or find super-emitting facilities, they are still by necessity sporadic and not used as part of a regular leak survey schedule.

Component surveys at regular intervals are mandated in many jurisdictions. In Alberta, Directive 060 in 2007 required formal leak detection and maintenance programs. A report published by the Canadian Association of Petroleum Producers (CAPP) suggested that following the implementation of directed inspection and maintenance programs by several companies, their estimated fugitive emissions decreased by as much as 75% (Clearstone Engineering Ltd., 2014b). However, currently fugitive emissions are still significant as shown by Figure 2.1. In fact, the significance may be underestimated by inventory reports.

A review of Canadian and American literature by Brandt et al. (2014) suggested that not only is the oil and natural gas sector an important methane contributor, but inventories are consistently low and measured leakage rates are higher than official estimates. Zimmerle et al. (2015) has suggested that total emissions from natural gas transmission and storage are under-reported by a factor of 2.6 under the U.S. EPA's Greenhouse Gas Reporting Program (GHGRP) rules.

In Canada, a recent "top-down" airplane study and comparison with "bottom-up" inventory estimates for two regions in Alberta (Johnson et al., 2017) noted similar discrepancies. While comparisons of the airplane measurements agreed well with the inventory in Red Deer Alberta area, 19% of emissions came from fugitive leaks and 53% from unreported venting, giving a total of ~70% of emissions from unidentified and/or unreported sources.

In the Lloydminster region of Alberta, aircraft measured emissions were ~3.6 times higher than bottom-up inventory estimates.

Brandt et al. (2014) suggested several potential reasons for these underestimations in emission inventories:

1. sampled devices do not represent current practices,
2. sample sizes are too limited (due to expense),
3. emissions have “heavy-tails” rather than normal distributions so small sample sizes are unrepresentative, and
4. activity and device counts are not representative.

A review of ~15,000 measurements from 18 studies was performed by Brandt et al. (2016) and found that natural gas leakage is indeed heavy-tailed with 5% of leaks contributing to the majority of emissions. This heavy-tailed distribution was ubiquitous across all data sets for various facility types, components, and scales. Zimmerle et al. (2015) noted a long tail distribution in most emission models based on field study and partner data and that, for most models, 5% of measurements are responsible for at least 30% of emissions. Allen et al. (2015) found that, for pneumatic controllers, a small subset of devices (19%) was responsible for 95% of emissions. On a site level, 15% of production sites are functionally super-emitters within the Barnett Shale area and are responsible for approximately three-fourths of methane emissions (Zavala-Araiza et al., 2015).

Super-emitters, or the small percent of leaking sources responsible for a large percent of total leakage (Brandt et al., 2016), are made up of three main types (Munnings & Krupnick, 2017):

1. chronic and predictable,
2. episodic with high emissions over brief periods, and
3. stochastic.

The third type (stochastic) are generally created by leaks which are difficult to predict. Work by Brantley et al. (2014) found that only 10% of methane emission variation could be explained by production rate changes, suggesting that unpredictable events are important.

Brandt et al. (2016) recommended formal leak detection and repair (LDAR) programs and frequent checkups as the best way to deal with such sources. However, even with regular LDAR, leaks can emit for months before they are found. Ravikumar et al. (2017) took this further by suggesting that continuous monitoring, though perhaps with reduced sensor sensitivity to reduce costs, as a future direction to reduce leak duration. Johnson et al. (2017), who found that significant emissions are from unmeasured or unreported sources, suggested that focus on these emissions is necessary for future mitigation.

Overall, reduction in leak duration, especially of large leaks, can have a significant impact on total green-house gas emissions. The faster leaks can be identified, and their importance scaled, the sooner they can be mitigated for maximum benefit. Lyon (2016b) investigated the use of single continuous sensors on gas pad sites and concluded that 6 scfh (standard cubic feet per hour) leaks could be noticed within a few days for tested sites (useful for directing maintenance to super-emitting sites), though these had relatively simple geometries. An attractive option is to identify the leaking components as well as leaking sites on this or shorter timescales. While multiple sensors are required, this reduces survey time and labour to find the component, so repairs can be made quickly, especially for larger facilities and/or when facility geometries become more complex. Additionally, continuous monitoring data both before and after maintenance provides proof of effective mitigation, which can improve operations management while opening the door to regulatory strategies based on carbon pricing and/or crediting.

Chapter 3 Fugitive Source Characterization from Sparse Concentration Sensors

Determining emission sources from sparse concentration information is a type of ill-posed problem that has been studied on a variety of scales for decades. Locating and quantifying potentially multiple unknown fugitive emission sources from a network of concentration sensors falls within the larger context of source-term estimation problems. In general, gas transport is modelled and compared with measurements at sensor locations to determine the best representation of emission releases. In the target application, the sensor measurements will be sparse (i.e. the number of available sensors and measurements will be limited compared with the variables representing the emission releases); thus, the problem is ill-posed and a unique solution cannot be guaranteed. Appropriate approaches to the transport model, problem type, and solution procedure depend on the requirements of the specific application. This chapter reviews the main approaches for each and subsequently discusses their relation to the target application.

3.1 Gas Transport Modelling

The first key component to estimating unknown source releases is choosing a method to determine the species transport from estimated emission releases that can be compared with measurement data. Traditional methods of modelling gas dispersion used in source-term estimation (such as Gaussian plume models, Lagrangian particle tracking, puff models, and the Eulerian approach to the advection-diffusion equation) are summarized in Appendix A. Overall, these models march the release of the gas forward in time, using various simplifications of the governing fluid mechanics with the most popular being the Gaussian plume model (Hutchinson et al., 2017).

Since it is computationally intractable to simply guess forward transport solutions for all possible unknown source locations until predicted dispersion matches sensor readings, an alternate approach is necessary. Unlike the traditional forward gas transport that starts with

the release of a gas at a location and marches forward in time to see where this gas is transported, inverse transport modelling flips this paradigm to consider where the gas could have originated upstream of a sensor location. The release is modelled from where and when the sample was measured and marched backward in time against the wind. The intent is to determine the sensitivity of that sensor measurement to locations in the domain. Simplistically, for a high concentration measurement, there is likely some gas source upstream of the measurement location. Conversely, for a low concentration measurement, it is unlikely that there is a source upstream of the sensor. Since the number of sensors is limited and known, while the potential sources are unknown, in some cases inverse transport may reduce computational cost while still providing similar information (such as concentration or sensitivity) to multiple forward transport simulations. It is also possible to perform weight regularization or renormalization of this inverse model (Issartel et al., 2007) to take advantage of geometry statistics (locations of sensors, buildings etc.), which can be used as a priori information in certain data assimilation techniques or for the reduction of artifacts near sensor locations.

3.2 Types of Source Characterization Problems

The allowable source types for a particular problem, the gas transport model, and the methods best suited to determine these sources(s) are inter-related. Sources are generally defined by location and emission rate (which may be constant or time-varying), with the location type potentially varying between point, area, and line sources. A summary of the main problem types is given below.

3.2.1 Single Source Estimation Problems

3.2.1.1 Single source with known location and unknown emission rate

The simplest problem type is determining the emission rate of a known source from downstream measurements. In the most common application, the emission rate is assumed constant, although there are scenarios where determining varying emission rates is desired (e.g. the Fukushima accident) (Hutchinson et al., 2017).

3.2.1.2 *Single source with unknown location and unknown emission rate*

This is the typical scenario under study for most source term estimation research. The variables being determined are minimal, even in cases in which additional plume variables (such as wind speed or direction) are defined.

3.2.2 Multi-Source Estimation Problems

Estimation of multiple simultaneous sources increases the complexity (and computational cost) from single source estimation. If plume models for each source must be performed in various combinations, the challenge grows with every source added. Due to problem complexity, studies looking at multiple sources are notable (e.g. Peneko et al., 2002; Allen et al., 2007; Yee, 2008; Lane et al., 2009; Huang et al., 2010; Konda et al., 2010; Annunzio et al., 2012; Sharan et al., 2012; Wade and Senocak, 2013; Cantelli et al., 2015; Singh and Rani, 2015).

Sources are typically defined by, at minimum, a position and an emission rate. Usually, a limited number of sources (or maximum number) is defined in advance (e.g. Allen et al., 2007; Cantelli et al., 2015; Sharan et al., 2012; Huang et al., 2010; Wade and Senocak, 2013; Singh and Rani, 2015). In some cases, comparisons are made between assumptions of source number (Wade and Senocak, 2013) and these are compared to find the optimal number. In other cases, source strength is neglected as a useful value (Annunzio et al., 2012) to reduce the number of unknown terms. As an alternative to using plume release coordinates, a mesh of potential emission positions can be defined (e.g. Konda et al., 2010). The latter lends itself to working in an Eulerian framework, as it is already discretized into a spatial mesh and plume modelling does not increase in complexity if more sources are defined within the domain.

3.2.3 Other Types of Source Characterization Problems

In some cases, best values for other variables in addition to the source location and strength are also desired. These can be related to selected transport model inputs, allowing them to be treated as part of the optimization rather than specified in advance of the source characterization procedure. Some examples from other researchers include wind speed

(e.g. Kopka et al., 2016; Mons et al., 2017), wind direction (e.g. Kopka et al., 2016; Mons et al., 2017), and dispersion estimates (e.g. Konda et al., 2010).

3.3 Approaches to Solving Source Characterization Problems

There are several possible approaches for quantifying and locating emissions from downstream measurements. In these types of inverse problems, the solution gives the model parameters that generate a desired model response. Potential methods can be categorized in several ways. For this review, they will be separated into direct trajectory-based methods, sequential data assimilation methods, and variational data assimilation methods, and discussed in the context of source term estimation. Advantages and disadvantages of each major type are summarized in Table 3.1 at the end of the section.

3.3.1 Trajectory Statistical Methods

Trajectory statistical methods (TSMs) are some of the simplest direct inverse methods. To determine where contaminated gas could be originating, the wind field is inverted and Lagrangian trajectories are calculated backwards in time from each concentration measurement location. Rather than attempt to determine emission rates, the areas that these trajectories pass through or near are weighted to determine regions in the domain that may have caused high concentration readings.

TSMs were developed for continental scale problems where they have been used extensively (e.g. Hsu et al., 2003; Ashbaugh et al., 1985, Lupu & Maenhaut, 2002; Cheng & Lin, 2001). Previous work (Brereton & Johnson, 2012a; Brereton & Johnson, 2012b) performed proof-of-concept testing of various trajectory-based methods over a synthetic industrial gas-plant domain and focused on assessing algorithm robustness to noise and the possibility of wind simplification to speed up computational processing. The major weakness of such trajectory-based methods is that, while capable of giving location information with very low computational time, release size is only indicated in a relative, qualitative sense.

3.3.2 Sequential Data Assimilation Methods

Sequential data assimilation evolves the estimation of a system state forward in time with corrections that use the difference between the predicted observations and measured

observations as observations become available (Nichols, 2003). Two popular sequential methods are Bayesian inference (Lee, 1989) and Kalman filtering (Kalman, 1960). These approaches determine a probability density function of the desired model parameters instead of a single solution, which has the advantage that uncertainty estimation is inherent to each method.

3.3.2.1 *Bayesian Inference*

Bayes' theorem for source reconstruction is (Monache et al., 2008)

$$p(\mathbf{S}|\mathbf{M}) = \frac{p(\mathbf{M}|\mathbf{S})p(\mathbf{S})}{p(\mathbf{M})} \quad (3.1)$$

where $p()$ is a probability distribution, \mathbf{S} is the state vector of source parameters, and \mathbf{M} is the measured observations. Bayes' theorem relates the probability of a certain emissions state given a measurement state (the posterior), to the probability of certain measurements given a source state (given here as $p(\mathbf{M}|\mathbf{S})$ also called the likelihood function). $p(\mathbf{S})$ is known as the prior distribution and, if the source is permitted to be located anywhere within the domain, is chosen to be uniform (Chow et al., 2008). $p(\mathbf{M})$ is the prior predicative distribution and used to normalize the posterior (Chow et al., 2008). Computing the likelihood function is difficult and cannot be done analytically. Thus, sampling and Monte Carlo Markov-Chain (MCMC) methods (or similar) are used to approximate the posterior distribution (Chow et al., 2008).

Substantial performance loss can be experienced when using experimental data due to poor models of the likelihood function (Hutchinson et al., 2017). In addition to traditional Bayesian inference, Approximate Bayesian Computation (ABC) can be used to replace the likelihood function in the Monte-Carlo algorithm with a measure of differences between predicted and measured concentrations to improve estimates with noisy data or model inaccuracies (Hutchinson et al., 2017).

Bayesian inference combined with MCMC has the advantage of determining a probability distribution and uncertainty directly (Senocak et al., 2008). Probability methods do not require selecting an objective function and result in the probability around critical points (Keats et al., 2007).

Bayesian inference techniques have been used to estimate the source distribution for numerous studies, including for small scale releases. Sohn et al. (2003) determined release locations in an office building. Keats et al. (2007) estimated emission location, rate and start/end times for a single source for the Mock Urban Setting Test and Joint Urban 2003 data sets. Chow et al. (2008) also applied Bayesian inference to the Joint Urban 2003 data set and successfully identified emissions sources. Yee (2008) used a Bayesian framework to determine multiple sources, though this was computationally expensive. Kopka et al. (2016) used an approximate Bayesian computation to examine contamination from mobile sources for seven source parameters. Wang et al. (2017) provided a comparison of Bayesian source estimation method likelihood functions and distance measures for a single open field release with Gaussian plume models. Other examples of Bayesian framework include Ristic et al. (2015), Senocak et al. (2008), Rajaona et al. (2015), Lucas et al., (2016), Huang et al., (2010), and Lane et al. (2009).

A disadvantage of Bayesian inference is that it requires extensive release simulations to estimate the probability distribution, which becomes computationally expensive, especially for complex 3D domains. For example, Chow et al. (2008) were forced to simplify and reduce the potential release locations to make the problem tractable. Senocak et al. (2008) utilized Bayesian inference combined with MCMC to locate the source of chemical and biological agents and use of the Gaussian plume dispersion model allowed quick computation of the forward simulations. However, these authors only considered simple steady flow fields. Guo et al. (2009) performed a probabilistic single source characterization with a finite volume release able to characterize the complex flow typical of urban settings. Zheng and Chen (2011) noted the inefficiency of these probability-based methods due to this computational cost and the extensive literature review by Hutchinson et al. (2017) noted that multiple source releases “posed a significant problem”. The additional computational time to determine the source number and characterize each is a significant increase from single source cases (Hutchinson et al., 2017).

3.3.2.2 *Kalman Filter*

The Kalman filter (Kalman, 1960) is a data assimilation method to estimate system state with a wide array of applications including inverse modelling (e.g. Hartley & Prinn, 1993).

In general, the state vector (e.g. vector of emission rates) is combined with the results of a forward transport model and this model's output is used to update the state vector to improve the match with the observations (Rao, 2007). Hartley and Prinn (1993) demonstrated the use of a linear Kalman filter in inverse sources-estimation modelling to determine surface fluxes of CFCl_3 on a regional scale using:

$$\mathbf{S}_{t+1} = \mathbf{S}_t + \mathbf{G}_t \left(\mathbf{M}_t^{\text{obs}} - \mathbf{M}_t^{\text{model}} \right) \quad (3.2)$$

where the source strengths \mathbf{S} at time t , are updated at time $t+1$ based on a comparison of mixing ratios at each measurement station (\mathbf{M}) and an updated gain matrix:

$$\mathbf{G}_t = \mathbf{D}_t \mathbf{P}_t^T \left[\mathbf{P}_t \mathbf{D}_t \mathbf{P}_t^T + \mathbf{N}_t \right]^{-1} \quad (3.3)$$

For Equation (3.3), \mathbf{D}_t is the covariance matrix of errors in the estimated emissions; \mathbf{P}_t is the partial derivative matrix at each measurement station with respect to the source strengths; and \mathbf{N}_t is the “noise matrix” or covariance matrix of measurement errors at each site (Hartley & Prinn, 1993). Determining the covariance matrix is the difficult portion of this method (Rao, 2007), and this restricts the problem size.

An ensemble Kalman filter (EnKF) is an approximate filtering method that attempts to deal with this complexity by representing the state distribution with an ensemble of draws from the state distribution, effectively reducing the problem dimensionality (Katzfuss et al., 2015). A variation of the EnKF is the Local Ensemble Transform Kalman filter (LETKF) (Hunt et al., 2007), which performs calculations independently for each particular grid point using only “nearby” data (Liu et al., 2016), lending itself to parallelization.

Mulholland and Seinfeld (1995) applied Kalman filtering to improve CO emissions inventory for the California Air Resources Board. Hass-Laursen et al. (1996) extended the work of Hartley & Prinn (1993) to deal with time varying sources on global scale emissions problems. Zhang et al. (2014) applied a modified EnKF for use in emergency releases of a known source location (but unknown buoyancy and wind). Liu et al. (2016) compared LETKF with 4D-Variational data assimilation methods for global CO_2 flux inversion, finding the largest difference in results in sparse data regions.

3.3.3 Variational Data Assimilation

Solving inverse problems is, in general, a two-stage process of estimation and appraisal (Sneider, 1998). In the estimation phase, a set of model parameters is predicted based on the observed data. In the appraisal stage, the correspondence of the generated data and the observations is evaluated (Addepalli et al., 2009). This measure of misfit is known as the cost or objective function and may also include penalties for distance away from an expected emissions source characterization.

Variational data assimilation methods attempt to minimize the objective function by optimizing solution candidates iteratively. Numerous algorithms exist to perform both global and local optimizations. A brief overview of some common methods that have been used for source estimation problems are given below.

3.3.3.1 Global Optimization

Genetic Algorithms

Genetic algorithms (Holland, 1975) represent model inputs on the idea of a chromosome. A population of initial solutions is generated randomly and the fitness of potential solutions is determined. These potential solutions are recombined, using the idea of “mating” to shuffle parameters and add random mutations to prevent getting stuck in local minima (Cartwright & Harris, 1993). In the case of emission releases, fitness can be evaluated by modelling a forward gas release and evaluating the cost function. Allen et al. (2007) used simple Gaussian plume models to evaluate the fitness of solutions to dispersion in an empty field. Haupt et al. (2007) expanded this work with the addition of puff dispersion models and testing additional cost functions. Khalafi et al. (2009) characterized sources on regional scales using Gaussian plume dispersion and genetic algorithms. Annunzio et al. (2012) found multiple release locations but did not estimate source strength to reduce unknown values. Cantelli et al. (2015) looked at emission rates and position of up to three sources with a Gaussian dispersion model. For multiple sources and sensors, a single string chromosome no longer contains all the necessary information and must be replaced by a matrix of dimensions # of sensors \times # of sources (Cartwright & Harris, 1993). When the number of sources is unknown, the size of this matrix becomes a matter of user judgment

given that all possible solutions must be simulated forward in time, which adds to the computational cost.

Simulated Annealing

Simulated annealing uses the annealing principle to introduce random perturbations, which are reduced in a controlled fashion to find a global minimum (Spall, 2003). After simulating neighbouring states to the current state, a probabilistic decision is made whether to move to a new state or remain at the current one. The method avoids local minima by allowing occasional moves to states that increase the cost function (Bertsimas & Tsitsiklis, 1993).

Forward transport simulations are generally used to determine the cost function to allow for a comparison between states. Thomson et al. (2007) used simulated annealing with Gaussian plume models and random search algorithms to estimate ethane sources in a desert environment. The computational time for use with a more complex gas dispersion model is prohibitive. Ma et al. (2013) performed comparison tests including both genetic algorithms, simulated annealing, pattern search, neural networks, Nelder-Mead simplex and several hybrid optimization schemes for single source identification and found that simulated annealing showed poorer than expected robustness and took many iterations if the initial guess was poor.

3.3.3.2 Local Optimization

Direct search

Also called derivative-free or black box methods, direct methods use the cost function directly and do not require that the cost function be differentiable or continuous, expanding the range of possible problems that can be solved and giving more reliable solutions when the function is noisy or has multiple local minima (Pulkkinen, 2008). Some examples are Rosenbrock (Rosenbrock, 1960), pattern search (Hooke & Jeeves, 1961), or hill climbing (Nolle, 2006). However, the convergence rates are typically slower than methods that take advantage of the gradient (Pulkkinen, 2008). Zheng and Chen (2011) used a pattern search method for source determination related to emergency releases.

Gradient-based methods

Gradient based optimization algorithms take advantage of the cost function gradient to move toward a minimum of the cost function. These are especially useful when there are many parameters and the cost function is continuous, such as the current case for multi-source releases with unknown locations. Additionally, inverse transport can be used to estimate gradient values in place of multiple forward transport simulations. Disadvantages of these methods are that the cost function gradient must be known or estimated, and they provide a local optimization so may converge to local rather than global minima depending on starting conditions. Common algorithm types include Newton's method, related quasi-Newton techniques such as Broyden-Fletcher-Goldfarb-Shanno (BFGS), and gradient descent. More detail on the BFGS and related methods can be found in Appendix B. Ruder (2016) provides an overview of methods within the gradient-descent family. Nocedal and Wright (2006) provide an extensive review of other gradient-based optimization methods including Newton and quasi-Newton families. For large numbers of parameters, BFGS is recommended due to its efficiency (Zhang, 2007). Limited memory BFGS (L-BFGS) (Liu & Nocedal, 1989) was employed by Elbern et al. (2000) to quantify synthetic emissions sources. Joynes (2013) also did an extensive study on synthetic releases in two-dimensional gas-plant domain using limited memory BFGS with promising results, but this work required extension into three dimensions and additional robustness tests to be practically useful. Bieringer et al. (2015) used L-BFGS as part of combined method of rapid source identification for hazardous releases.

Table 3.1: Summary of major source characterization methods

Method	Advantages/Disadvantages	Examples
<i>Trajectory-based methods</i>	Very low computational cost No quantitative emission estimate	Brereton & Johnson, 2012a
<i>Sequential Data Assimilation</i>		
Bayesian inference	No objective/cost function required Includes uncertainty estimate High computational cost (large number of samples required)	Keats et al. 2007 Yee, 2008 Kopka et al., 2016 Wang et al., 2017
Kalman filter	High computational cost	Hartley & Prinn, 1993 Zhang et al., 2014
<i>Variational Data Assimilation</i>		
Genetic Algorithms	Global optimization Requires parameter range and number pre-specification Number of parameters may become very large High computational cost (large number of simulations to determine fitness)	Allen et al. 2007 Annunzio et al., 2012 Cantelli et al., 2015
Simulated Annealing	Global optimization Does not require objective function derivative High computational cost (large number of simulations required)	Thomson et al., 2007 Ma et al., 2013
Direct Search	Local optimization Does not require objective/cost function gradient Slower convergence than gradient-based methods	Zheng & Chen, 2011
Gradient-based	Local optimization Improves convergence for large optimization problems Good for optimizing many parameters Requires objective/cost function gradient	Joynes, 2013 Bieringer et al., 2015

3.4 Target Application and Proposed Approach

The target application of this work is locating and quantifying unknown fugitive emission sources within a domain potentially containing multiple buildings, bulk equipment, extensive piping, and many possible leaking components (potentially thousands). Ideally, this could be accomplished in a relatively short timeframe, potentially on a continuous or near-continuous basis. Wind information (from anemometers) and concentration data from

a sparse network of sensors could be gathered over the course of several days during favourable weather conditions. This information could be used to estimate flow fields within the domain and subsequently synthesize the data into an estimate of likely emission sources and magnitudes to aid in their mitigation.

Determining fugitive emissions from a sparse sensor network and some wind knowledge has specific requirements. First, the domain can vary from the comparatively small and simple (e.g. a 30-m wide single well oil or gas production site) to the large and complex (e.g. a gas plant hundreds of meters in size with buildings, equipment, piping, and thousands of potential release regions). Second, the number and location of sources is also unknown. Third, computational cost is a consideration if the method is to be used for practical purposes. Ideally, it would be desirable to detect and quantify sources on a regular basis.

3.4.1 Dispersion Model Selection

The potential domain complexity rules out the use of the Gaussian plume model as plumes are expected to bend and meander and have non-Gaussian shapes. Even industrial models with corrections such as AERMOD have difficulty with certain building configurations and are primarily meant for stacks. Lagrangian-based particle tracking models are computationally costly if used to determine concentration. While puff models can deal with some flow-field complexity, Eulerian-based dispersion modelling using the advection-diffusion equation naturally deals with increasing complexity, bluff bodies and terrain, and even porosity (which may be appropriate for some configurations of piping, for example). The geometric complexity can be defined by the user during the geometry and discretization phase.

The use of the advection-diffusion equation in an Eulerian framework assumes that the gas concentration is small and is unlikely to affect the bulk flow in any significant way. This allows the gas transport modelling to be separated from the wind flow. However, even if the velocity field is completely known statistically, closure assumptions must be made due to the number of unknowns (Tavoularis, 2018). One simple closure method is using the gradient transport concept (Tavoularis, 2018). Turbulent diffusivity relates the

concentration gradient to turbulent diffusion (analogous to the molecular diffusion coefficient and Fick's law, but caused by turbulent flow fluctuations). Detail of the advection-diffusion equation and its derivation from mass conservation can be found in Appendix A.

Turbulent diffusivity is implemented as a scalar field, determined from the eddy viscosity predicted for the wind field and scaled based on the turbulent Schmidt number (ratio of turbulent eddy viscosity to turbulent mass diffusivity) selection. The turbulent Schmidt number selection, range, and effects are discussed in depth in Chapter 6.

However, diffusivity is not necessarily independent of direction, such as for high shear flows, and may be more accurately represented by the second-order turbulent diffusivity tensor (Tavoularis, 2018). Thus, while the presented test cases use a scalar field, the equations and derivations presented in Chapter 4 assume that diffusivity may vary with direction for potential use in other implementations.

3.4.2 Source Definition

The selected approach must be capable of dealing with multiple sources. For dispersion modelling with the advection-diffusion equation in an Eulerian framework, once the domain is discretized, combinations of cells can be used to represent various types and numbers of sources without increasing problem complexity. Treating each cell as a potential source with an unknown emission rate meets the criteria of allowing multiple sources at unknown locations. This is a more complex version of the single-height grid of potential sources used by Konda et al. (2010) with the Gaussian plume model in an open field. The sources are then restricted by the mesh discretization and to volumetric regions, rather than by a small finite number.

3.4.3 Optimization Algorithm

A gradient-based algorithm approach is selected for computational efficiency and ability to deal well with a large number of optimization parameters. Specifically, the bounded version of L-BFGS (L-BFGS-B) was chosen (see Appendix B for details on the algorithm itself). This bounded version naturally allows known sinks, sources, or no-source regions to be specified with no changes to the code if more information about the domain is

available. Similar selections were made by Joynes (2013) for 2D proof-of-concept testing, but a 2D domain is unrealistic for bluff-body flows. This earlier work is extended to three-dimensions and significantly expanded to determine the overall performance and practicality.

3.4.4 Cost Function

The cost function, in addition to reducing as source distribution estimates improve, should be differentiable by the model concentration and source terms (from the advection-diffusion equation used to model the forward gas transport). The cost function was therefore defined based on the least-squared difference between modelled and desired (observed) concentration sensor measurements. Sensor measurements are sparse, with no more than one sensor within a single cell of the discretized domain, and each sensor measurement was given equal weighting in the cost function. A regularization term was not applied (using the difference from an expected source distribution) as earlier tests showed that this resulted in diffuse location predictions with no appreciable benefit for this type of problem (Joynes, 2013). Such a formulation assumes that sensor observations are uncorrelated and have the same measurement error. Other cost functions could be used with the presented source characterization approach. For example, for some setups (especially with mixed sensor types) it could be useful to weight sensor importance or normalize by measurement error (such as in Eq. 3.3).

3.4.5 Cost Function Gradient

The gradient of the cost function is also required to use a gradient-based optimization approach such as L-BFGS. While the gradient can in theory be obtained by running a very large number of forward gas transport simulations, the computational cost is prohibitive. Instead, the gradient of the cost function can also be obtained efficiently using inverse dispersion modelling via the advection-diffusion (or scalar transport) adjoint. The derivation of the cost function gradient based on this method is presented in Chapter 4.

Chapter 4 Fugitive Emission Source

Characterization Using a Gradient-Based

Optimization Scheme and Scalar Transport Adjoint

This chapter has been peer-reviewed and was published in Atmospheric Environment. Only minor clarifications and formatting edits have been made following feedback from the thesis examination committee and for formatting consistency throughout the thesis. It should be cited as:

Brereton, C. A., Joynes, I., Campbell, L. J., Johnson, M. R., 2018. Fugitive Emission Source Characterization Using a Gradient-Based Optimization Scheme and Scalar Transport Adjoint. *Atmos. Environ.*, 181, 106-116 (doi: 10.1016/j.atmosenv.2018.02.014).

This paper was co-authored by the thesis author, Carol Brereton, her thesis supervisors Prof. Matthew Johnson and Prof. Lucy Campbell, as well as former student Ian Joynes. Ms. Brereton set-up, developed, and tested the relevant procedures and code; generated synthetic test cases; performed the relevant test cases and initial data analysis; and wrote the first draft of the manuscript. Theory development, final data analysis, and editing and revision of the manuscript were conducted jointly.

4.1 Abstract

Fugitive emissions are important sources of greenhouse gases and lost product in the energy sector that can be difficult to detect, but are often easily mitigated once they are known, located, and quantified. In this paper, a scalar transport adjoint-based optimization method is presented to locate and quantify unknown emission sources from downstream

measurements. This emission characterization approach correctly predicted locations to within 5 m and magnitudes to within 13% of experimental release data from Project Prairie Grass. The method was further demonstrated on simulated simultaneous releases in a complex 3-D geometry based on an Alberta gas plant. Reconstructions were performed using both the complex 3-D transient wind field used to generate the simulated release data and using a sequential series of steady-state RANS wind simulations (SSWS) representing 30 s intervals of physical time. Both the detailed transient and the simplified wind field series could be used to correctly locate major sources and predict their emission rates within 10%, while predicting total emission rates from all sources within 24%. This SSWS case would be much easier to implement in a real-world application and gives rise to the possibility of developing pre-computed databases of both wind and scalar transport adjoints to reduce computational time.

4.2 Introduction

Fugitive emissions, defined for the purposes of this paper as unintended and/or uncontrolled gas releases, are an important source of greenhouse gas and volatile organic compound (VOC) emissions. This work is primarily motivated by the energy sector, where fugitive releases are often economic to mitigate when their sources are known. However, the approach may also be applicable to other domains where locating and quantifying unknown gas releases is of interest. Complex facilities such as gas processing plants may contain almost innumerable potential sources spread over many thousands of fittings, seals, valves, tanks, etc. and connected by miles of piping. This makes the process of identifying release source(s) difficult. Efficient detection, location, and quantification of individual fugitive sources within this type of environment would allow for prioritization of repairs, lower emissions, and significantly improved emissions reporting.

Traditional fugitive characterization methods include manual component testing or qualitative viewing with IR cameras to visualize escaping gas (USEPA, 2016b). In recent years, progress has been made on emission quantification with IR cameras (Gålfalk et al., 2016), mobile drive-by surveys (e.g. Jackson et al., 2014; Mitchell et al., 2015; Phillips et al., 2013; Rella et al., 2015; Albertson et al., 2016; Atherton et al., 2017; Roscioli et al., 2018; Zavala-Araiza et al., 2018) and airborne fly-overs (e.g. Thompson et al., 2015;

Thorpe et al., 2016; Tratt et al., 2014; Karion et al., 2013; Lavoie et al., 2015; Johnson et al., 2017). These methods, however, by design provide intermittent, qualitative screening of leaking components or snapshots of overall emissions at facilities. This potentially allows leaks to go unnoticed for significant periods due to necessarily long measurement intervals. Reducing this time between surveys would allow maintenance and repair to be directed to release locations faster and decrease overall emissions.

The motivation and target application for the current study is location and quantification of unknown fugitive emissions sources (i.e. leaks) within an oil and gas facility, or industrial plant environment of similar scale and complexity. A fixed position network of concentration sensors installed within the facility would allow for quasi-continuous monitoring. However, because a complex facility could contain thousands of valves, seals, and fittings (each a potential emitter), practicality and cost would necessarily prohibit monitoring all possible sources individually. Instead, data from a comparatively sparse network of sensors could be leveraged in an optimization-based scalar-transport modelling framework to determine source regions and quantify release magnitudes. This combination of a continuous sensor network and quasi-continuous modelling could be used to quickly direct maintenance to new leaks within a small search space, reducing both GHG release and lost product. Such a system could also potentially be configured to permit external data transmission sufficient for third-party verification of pollution mitigation and emissions crediting.

Emission source characterization, i.e. determining emission locations and rates that best describe downstream concentration measurements, has been studied on various domain scales from continental (e.g. Houweling et al., 1999) to urban (e.g. Chow et al., 2008) to agricultural fields (e.g. Wilson et al., 2012). Genetic algorithms (Allen et al., 2007), Bayesian inference (Keats et al., 2007; Rajaona et al., 2015) and direct search methods (Zheng and Chen, 2011) are all data assimilation methods that have been directed at source characterization. Additionally, inverse dispersion modelling has been applied successfully at various scales and applications (Wilson et al., 2012; Wilson et al., 2009). Rao (2007) provides an overview of general source characterization methodologies.

For the target application, in which there may be multiple simultaneous releases at unknown locations within complex domains containing multiple building-scale bluff-bodies, source characterization is especially challenging. Modelling all possible forward releases to compare with measurements is intractable. Gas transport adjoint modelling, also known as inverse plume modelling, is a lower computational cost alternative that seeks to use data from known sensor measurement locations to determine where detected gas has originated. Such adjoint methods have been successfully used for continental scale studies, as in Hakami et al. (2007) who presented the adjoint for various atmospheric processes on a large scale atmospheric model. This approach has a parallel in the backwards-in-time-and space (trajectory) method as exploited for inverse dispersion in urban environments by many authors (e.g. Wilson et al., 2009).

Emergency release event prediction takes advantage of this computational savings but is often unconcerned with detailed quantification. For example, Huang et al. (2015) noted that for short term pollution episodes, potential source area deduction was more practical than precise back-calculation. Rudd et al. (2012) and Singh and Rani (2014) optimized variables for characterizing single sources rather than multiple sources as the current application requires. Crenna et al. (2008) investigated quantifying multiple source releases simultaneously and found that wind direction and sensor-source geometry were the key factors for feasibility. In principle, there is no reason multiple sources cannot be quantified with appropriate sensor placement and a variation in wind directions. Proof-of-concept tests (Brereton & Johnson, 2012a) suggested that multiple sources on an industrial scale could be quickly located with inverse trajectories from sensor locations given sufficient wind direction coverage and/or good sensor placement, though this approach did not quantify emission rates.

The current paper presents an adjoint-based optimization approach for characterizing fugitive emissions, allowing for both complex flow fields and multiple sources. A proof-of-concept numerical analysis is presented using two distinct test cases: (i) real-world, single source releases from Project Prairie Grass (PPG) (Barad, 1958), and (ii) numerical simulations of a 3D geometry based on an actual gas plant facility with multiple

simultaneous sources. Approaches to speeding computations using simplified wind fields are also considered.

4.3 Adjoint-based optimization methodology

The target problem of interest is locating and quantifying an unknown number of gas releases within a physical domain, where concentration data would only be available from a sparse network of sensors. Depending on the facility, the physical domain could range from tens to hundreds of meters in size and contain buildings, pipes, and equipment. This physical domain is represented as a 3-D computational domain discretized in space and time, with sufficient detail to capture major wind effects on observed concentrations, such as wakes of buildings and equipment.

Within this discretized domain, the source characterization is represented by a finite number of potential gas release locations within the spatial domain and their emission rates (allowing for zero emissions). In this paper, which takes an Eulerian approach, these locations correspond to the volumetric mesh cells of the discretized spatial domain (Ω), and the desired emission rates are assumed to constant over the modelled timeframe.

To determine these emission rates, source characterization was treated as an optimization problem utilizing both forward and inverse modelling methods. Before beginning the optimization, the wind fields are estimated. Subsequently, the optimization loop is initialized with a starting source characterization, or initial source guess, such as a zero emission rate field and the forward evolution of the gas release is modelled using the estimated wind fields. While an Eulerian perspective is used for this transport model in this paper, there is no fundamental reason a Lagrangian approach could not be applied. The modelled concentrations are compared with observations from the physical domain to evaluate the current source guess. If the match is poor, an inverse model determines upstream location sensor sensitivity and the optimization algorithm determines how source rates should be updated. A flow chart of the optimization procedure is shown in Figure 4.1, and the major steps are detailed in subsequent subsections.

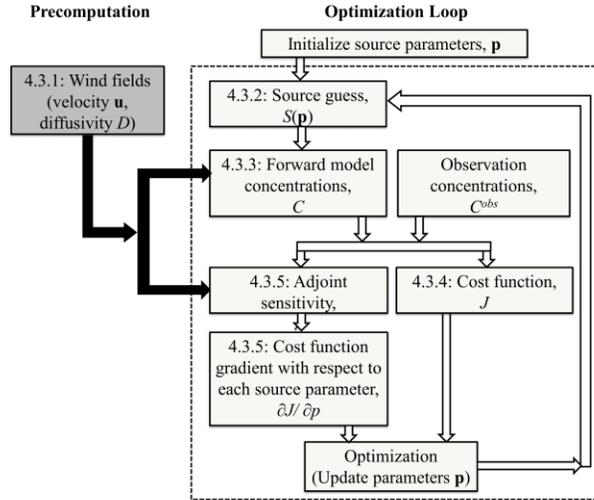


Figure 4.1: Simplified optimization steps with corresponding paper section numbers and major variable names. Source parameters refer to emission rate per cell for this paper.

For the optimization algorithm itself, the quasi-Newton gradient-based L-BFGS-B (Liu & Nocedal, 1989) was chosen for its efficiency with large numbers of parameters, since potentially millions of variables (corresponding to the number of cells) must be optimized.

4.3.1 3D Wind Field Data

Before the optimization loop can be performed, 3D wind fields must be estimated (including a measure of unresolved turbulent wind fluctuations) and used as input for transport models corresponding to the pre-computation step in Figure 4.1. Wind field calculations could be implemented in many ways with varying levels of accuracy. In the simplest case, a horizontally-uniform field of wind statistics may be assumed, parameterized by the friction velocity, Obukov length, roughness height, and mean wind direction. The most complex cases with substantially varied wind statistics in both vertical and horizontal planes may only be simulated in elaborate computations. For the present test cases, wind fields were determined either by directly mapping a vertically varying profile based on measured data (simplest case) or by simulating 3-D flows (either as a transient flow or as a series of varied inlet steady flows) using CFD. For a specified inlet boundary condition and duration, this wind simulation can be performed as a single pre-computation since the bulk wind flow will not be affected by low concentration plume transport. This significantly reduces overall computational time by avoiding the need to

determine the wind field concurrent with the forward model of the tracer gas plume, which may need to be computed 100 s of times.

4.3.2 Source Guess

In the present test cases, the emission source rates (referred to as the Source Guess in Figure 4.1) of the optimization loop were initiated assuming no fugitive sources were present (i.e. the emission rate at each cell in the domain was set to zero). The effects of assuming different initial conditions are shown for the Project Prairie Grass test case in Appendix F. In principle, known sources, sinks, and/or non-emitting areas could be easily considered by specifying values in selected cells as model constants not varied by the optimization scheme. Similarly, area, line, and point sources could be represented as different groupings of cells, with varying source strength, location, and number. Source distribution is then restricted to the chosen mesh resolution used for the forward model and adjoint sensitivity and assumed to maintain the same rate and location for all input data (i.e. constant in time over the course of the simulations).

4.3.3 Forward Scalar Transport Model

Atmospheric dispersion encompasses transport models on a variety of scales ranging from the greatly simplified (e.g. Gaussian plumes, Reible, 1999) to the complex (e.g. large scale reactive weather dependent flows, Peters et al., 1995) depending on application. For the current work, the low computational cost of the simple plume models is desirable, but temporal variation, building wakes, and recirculation zones must be considered as sensors and sources are likely to be within a built area.

Assuming a known time-varying mean, non-divergent velocity field [length/time] (\mathbf{u}) and diffusivity [length²/time] (D) which parameterizes the unresolved turbulent fluctuations, for gas sources (or sinks) [mass/time/length³] (S) of a dilute non-reacting/depositing species, the evolution of concentration [mass/length³] (C) with time t may be predicted with the advection-diffusion equation:

$$\frac{\partial C}{\partial t} + \nabla \cdot (C\mathbf{u} - D\nabla C) - S = 0. \quad (4.1)$$

The advection-diffusion equation is a statistically-based statement of mass conservation using the gradient-diffusion approximation to deal with turbulent motions. As 3-D external bluff-body flows are of interest, the standard three component cartesian directions are assumed throughout this paper. The diffusivity, D , may be a tensor of diffusivity coefficients in cases where directionality has an effect (Kuzmin, 2010), though this is generally taken as a scalar and implemented as such in this paper.

Buoyancy of the target pollutant species (e.g. methane) may be neglected as long as downstream sensors can be expected to be sensitive to the upstream release plumes. Due to entrainment of air, neutrally buoyant source gases are not required as shown by successful oil and gas facility tracer studies with sensors up to several kilometers downwind (Roscioli et al., 2014). When applied to the spatial domain (Ω) with boundaries (Γ), outlets use zero gradient boundary conditions and the inlets must specify flux. In the current cases, no upstream sources were considered. However, even in the presence of upstream sources, these settings can be used with the expectation of additional sources predicted near the upstream domain boundary. Outputs of this forward transport model, in conjunction with the observed sensor concentrations, allow the value of the cost function to be determined (Figure 4.1).

4.3.4 Cost Function

A measure of misfit (here termed g) between the modelled concentrations C (with guessed source strengths in each cell \mathbf{p}) and the field observed concentrations C^{obs} was defined for positions \mathbf{x} and times t as:

$$g(\mathbf{x}, t, \mathbf{p}) = 1/2[F(\mathbf{x}, t)C(\mathbf{x}, t, \mathbf{p}) - C^{obs}(\mathbf{x}, t)]^2 \quad (4.2)$$

where C^{obs} is defined as non-zero only where and when observations exist. These measurements are sparse and no more than one observation is considered per discretized cell volume. Likewise, the function F is set to unity at the positions and times at which observations are measured and zero otherwise such that misfit comparisons are only made when observations exist. Eq. 4.2 defines a dimensional misfit for ease of implementation with the solvers described in Section 4.4, but an appropriate normalization could also be applied if desired and units of the subsequent definitions modified accordingly. The term

g is then integrated from the initial (t_0) to final (t_f) times and over spatial domain (Ω) to determine the cost function, J :

$$J(\mathbf{p}) = \int_{t_0}^{t_f} \int_{\Omega} g d\mathbf{x}dt. \quad (4.3)$$

Optimizing for the minimum of the cost function will then find the best least-squared fit between the measured and modelled concentration fields. While an additional regularization term can be added to the cost function to encourage the solution towards an initial solution estimate, earlier testing (Joynes, 2013) on the current problem found this had little benefit while tending to diffuse the source location predictions.

4.3.5 Adjoint Sensitivity and Cost Function Gradient

To optimize the emission characterization, both the cost function and its gradient are required by the L-BFGS optimization algorithm. The first can be determined by computing the modelled concentration as described above and then using the values at the measurement locations to determine the cost function value defined in above. However, the derivative of Eq. 4.3 with respect to the emission rate in each cell p_i ,

$$\frac{\partial J}{\partial p_i} = \int_{t_0}^{t_f} \int_{\Omega} \frac{\partial g}{\partial p_i} d\mathbf{x}dt = \int_{t_0}^{t_f} \int_{\Omega} \left(\frac{\partial g}{\partial C} \frac{\partial C}{\partial p_i} + \frac{\partial g}{\partial S} \frac{\partial S}{\partial p_i} \right) d\mathbf{x}dt, \quad (4.4)$$

contains $\partial C / \partial p_i$ which is difficult and costly to compute. For the current source and cost function definitions, Eq. 4.4 can be re-derived in a more useful form to reduce computational cost. As the left side of Eq. 4.1 is zero, Eq. 4.1 multiplied by the scalar adjoint variable λ may be subtracted from the misfit within the integrals of Eq. 4.3 without altering the cost function:

$$J = \int_{t_0}^{t_f} \int_{\Omega} g - \lambda \left(\frac{\partial C}{\partial t} + \nabla \cdot (C\mathbf{u} - D\nabla C) - S \right) d\mathbf{x}dt \quad (4.5)$$

The units of the adjoint variable itself depend on the selection of misfit function g . For the current case, if the concentration values are in units of kg/m^3 , the adjoint variable λ should have units of $\text{kg s}/\text{m}^3$. The derivative of Eq. 4.5, which gives the gradient of the cost function required by the optimization scheme, is taken with respect to the model parameters with the last four integrals referring to the evolution, advection, diffusion, and source terms respectively.

$$\begin{aligned}
\frac{\partial J}{\partial p_i} = & \int_{t_0}^{t_f} \int_{\Omega} \frac{\partial g}{\partial C} \frac{\partial C}{\partial p_i} + \frac{\partial g}{\partial S} \frac{\partial S}{\partial p_i} d\mathbf{x}dt \\
& - \int_{t_0}^{t_f} \int_{\Omega} \lambda \frac{\partial}{\partial t} \left(\frac{\partial C}{\partial p_i} \right) d\mathbf{x}dt - \int_{t_0}^{t_f} \int_{\Omega} \lambda \nabla \cdot \left(\frac{\partial C}{\partial p_i} \mathbf{u} \right) d\mathbf{x}dt \\
& + \int_{t_0}^{t_f} \int_{\Omega} \lambda \nabla \cdot \left(D \nabla \left(\frac{\partial C}{\partial p_i} \right) \right) d\mathbf{x}dt + \int_{t_0}^{t_f} \int_{\Omega} \lambda \frac{\partial S}{\partial p_i} d\mathbf{x}dt
\end{aligned} \tag{4.6}$$

If the diffusivity coefficient matrix D is restricted to be diagonal, divergence theorem and integration by parts allow Eq. 4.6 to be rearranged as:

$$\begin{aligned}
\frac{\partial J}{\partial p_i} = & \int_{t_0}^{t_f} \int_{\Omega} \left(\frac{\partial \lambda}{\partial t} + \mathbf{u} \cdot \nabla \lambda + \nabla \cdot (D^T \nabla \lambda) + \frac{\partial g}{\partial C} \right) \frac{\partial C}{\partial p_i} \\
& + \left(\frac{\partial g}{\partial S} + \lambda \right) \frac{\partial S}{\partial p_i} d\mathbf{x}dt - \int_{\Omega} \lambda \frac{\partial C}{\partial p_i} \Big|_{t_0}^{t_f} d\mathbf{x} \\
& + \int_{t_0}^{t_f} \int_{\Gamma} \left(\lambda \left(D \nabla \left(\frac{\partial C}{\partial p_i} \right) - \frac{\partial C}{\partial p_i} \mathbf{u} \right) - \frac{\partial C}{\partial p_i} D^T \nabla \lambda \right) \cdot \mathbf{n} d\mathbf{x}dt
\end{aligned} \tag{4.7}$$

Finally, it is possible to select an adjoint variable that eliminates $\partial C / \partial p_i$ entirely from the cost function gradient computation by reducing terms containing it to zero. The appropriate adjoint variable can be found by solving

$$-\frac{\partial \lambda}{\partial t} + \nabla \cdot (-\lambda \mathbf{u} - D^T \nabla \lambda) - \frac{\partial g}{\partial C} = 0 \quad (4.8)$$

with conditions

$$\lambda(\mathbf{x}, t = t_f) = 0$$

$$(\nabla \lambda) \cdot \mathbf{n} = 0 \text{ where } \mathbf{u} \cdot \mathbf{n} \leq 0 \quad (4.9)$$

$$(\lambda \mathbf{u} + D^T \nabla \lambda) \cdot \mathbf{n} = 0 \text{ where } \mathbf{u} \cdot \mathbf{n} > 0.$$

Eq. 4.8 is similar in form to the forward release (Eq. 4.1) with inverted signs, showing the same form as the analogy driven adjoint operator derivation of Pudykiewicz (1998) (or scalar transport adjoint) with a differing source term appropriate to the current problem determined by the definition of the misfit g . It is worth noting that Eq. 4.8 effectively describes inverse plumes from sensor locations. Substituting in Eq. 4.8 and its boundary conditions into Eq. 4.7 gives the cost function derivative

$$\frac{\partial J}{\partial p_i} = \int_{t_0}^{t_f} \int_{\Omega} \left(\frac{\partial g}{\partial S} + \lambda \right) \frac{\partial S}{\partial p_i} d\mathbf{x} dt \quad (4.10)$$

Since the misfit, g , as defined in the previous section is independent of the source terms, this simplifies to

$$\frac{\partial J}{\partial p_i} = \int_{t_0}^{t_f} \int_{\Omega} \lambda \frac{\partial S}{\partial p_i} d\mathbf{x} dt. \quad (4.11)$$

With the source definition in the previous section, $\partial S / \partial p_i$ has a value of one for cell i and zero elsewhere. Summation over the domain allows the gradient to be computed for each emissions parameter, in this case the source strength in each cell. The gradient is then fed into the optimization scheme (Figure 4.1) and used to improve the source guess.

4.4 Implementation

The optimization loop (see Figure 4.1) employed the limited-memory variation of Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm, implemented starting from the C-wrapped FORTRAN subroutines for bounded L-BFGS of Zhu et al. (2011). As necessary, the optimization loop requests the cost function and cost-function gradients from external solvers for specific optimization parameter arrays (in this case, the field of emission rates for the domain). L-BFGS was previously investigated for source characterization using synthetic test releases (Elbern et al., 1997; Joynes, 2013). The required maximum corrections for the limited memory matrix was set to five. Sinks were precluded by setting a minimum source bound of zero emissions in each cell. No upper bound constraints were applied.

The external solvers to determine the cost function value and its gradient were implemented using the partial differential equation solvers in the OpenFOAM toolkit, which contains extensive existing libraries appropriate for finite-volume CFD and related applications. Both forward (Eq. 4.1) and adjoint (Eq. 4.8) transport solvers could read precomputed velocity and diffusivity information in matching OpenFOAM CFD formats. Sensor reading locations in time and space, likewise, were treated as temporally variable fields. The cost function and gradient were concurrently numerically integrated in time, and subsequently in space, corresponding to Eq. 4.3 and Eq. 4.11.

4.5 Results

Two test cases were chosen to investigate the potential of the described characterization method. The benchmark simple open-field release experiment, Project Prairie Grass (Barad, 1958), was selected as the initial test case. This is a well-known and documented set of atmospheric gas dispersion experiments often used to demonstrate the validity of plume models. The emission rate optimization method could then be tested separately from the ability to model wind complexities introduced by bluff bodies. The second test case was a simplified industrial geometry based on a gas plant in Alberta, Canada, containing multiple bluff bodies more representative to the envisioned target application. Multiple, simultaneous releases were simulated in this latter test case. Potential for using simplified

wind fields, consisting of piece-wise assemblies of 30-s duration steady flow simulations, were also considered.

4.5.1 Project Prairie Grass Test Case

Project Prairie Grass (PPG) experiments (Barad, 1958) consisted of 68 individual SO₂ releases in an open field conducted throughout the summer of 1956. Ten-minute average concentrations at an elevation of 1.5 m were measured at sensors 50 m, 100 m, 200 m, 400 m, and 800 m from the release location at 0.46 m elevation. Additional measurements were collected on six 17.5 m wind towers, 100 m from the release location. For the test simulations, the domain was restricted to an area encompassing the first two arcs of sensors (which includes the towers) where the highest readings could be found, forming a 300 m × 200 m × 50 m volume as shown from a top-down view in Figure 4.2. Mesh statistics are shown in Table 4.1 (and mesh convergence information is given in Appendix D). Sensor locations were reduced where necessary to prevent data from multiple sensors within a single cell region. As the elevation of the field in which the releases were performed varied by only ±1 m over the relevant area, the open field was taken as flat.

Eq. 4.9 gives a description of the necessary boundary conditions for the adjoint transport equation. In contrast to the forward case, where the inlet concentration is typically set (Dirichlet condition), the reverse inlet (forward outlet) for the adjoint has a zero-flux condition. However, this is more difficult to implement for many solvers. The effects of using either Dirichlet or flux boundary conditions on the adjoint transport equations were investigated (tested using a mesh coarsened by a factor of ~3 to speed computations). Final emission rates with these two conditions differed by less than 1% with no visible location changes. This is likely to remain true for any outlets where advective transport dominates.

Two release experiments with matching release rates (101.8 and 102.0 g/s) but different wind directions, and without missing data or other experimental problems, were selected (releases 19 and 49) from within the PPG data sets. Two wind directions represent the minimum amount of data that can be expected to successfully find source locations, and thus represents a conservative (more challenging) test case. Preliminary tests incorporating various combinations with additional experimental data sets produced no substantial

changes in results. Wind fields for the PPG test case were not simulated with CFD. Instead, profiles were mapped onto the domain directly based on the experimental friction velocity, Obukhov length, and an assumed roughness height of 0.006 m. A detailed discussion and review of appropriate profiles for various stability classes can be found in Alinot (2003) (and Appendix C). For these particular cases, both data sets correspond to an unstable atmosphere. The appropriate turbulent Schmidt number is often assumed to be near unity, but a review by Gualtieri et al., (2017) shows that for atmospheric dispersion there is no fixed value. A range of 0.6-0.7 is likely more appropriate (Wilson et al., 2009; Wilson, 2013; Wilson, 2015). Thus, a value of 0.67 was used in the present simulations. Future work (Brereton et al., 2018c) is underway to explore the sensitivity of results to turbulent Schmidt number selection.

Forward release and scalar transport adjoint solves were performed when called by the optimization loop (see Figure 4.1) with input from the 10-minute average wind fields. The transport solves used a first-order Euler temporal scheme with a 0.1 s timestep for a total of 60 s physical time. Observations and model concentrations were compared at the end of this physical interval to determine the cost function. Tests showed minimal differences in reconstructed emission fields from simulating the full transient over 600 s with various timesteps and averaging to match the PPG experimental conditions.

Table 4.1: PPG Mesh Statistics

Number of Cells	Domain bounding size [m]	Vertical spacing at ground [m]	Max vertical cell expansion ratio	$\sqrt[3]{\text{Avg Cell Volume}}$ [m]
1.5 Million	$300 \times 200 \times 50$	0.1	1.2	0.8

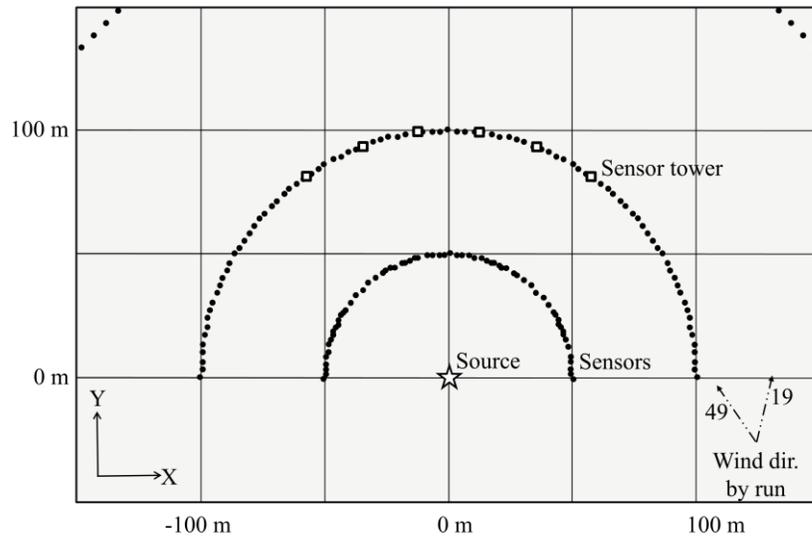


Figure 4.2: PPG domain with source (star) and sensor locations for 1.5 m height (small dark circles) and towers (unfilled boxes)

4.5.2 PPG Open Field Reconstruction Results

Figure 4.3 plots the convergence of the cost function (normalized by initial value and shown on the right axis), predicted emission rate integrated over the entire domain (normalized by actual emission rate), and predicted emission rate within a connected volume near the identified source (also normalized by actual emission rate) for the adjoint optimization of the PPG experiment. The boundary of the connected predicted source volume was determined by applying a minimum threshold of 1% relative to the maximum predicted emission rate per volume, and has the effect of capturing most emissions while avoiding prediction noise near sensor locations. Within fifty iterations, the cost function had decreased by nearly a factor of twenty-five, and the predicted total emission rate was within 20% of the final iteration. After 181 iterations, the predicted mass flow rate of the release converged to a final value that was within 37% of the expected emission rate for the entire domain, and within 13% when only considering the region near the predicted source. Although it is noted that the predicted magnitude is affected by the choice of the turbulent Schmidt number, this agreement demonstrates the utility of the adjoint-optimization approach.

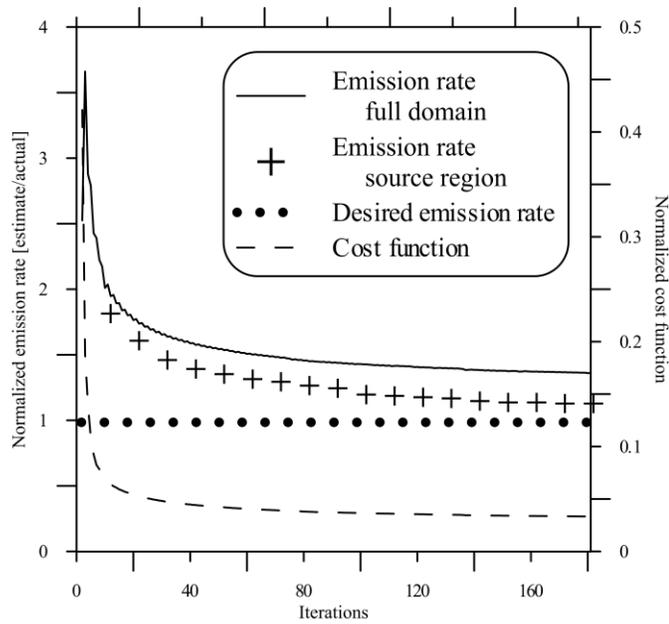


Figure 4.3: Cost function and mass flow rate convergence

The predicted release location for the adjoint-optimization of the PPG experiment is shown in Figure 4.4. By iteration fifty, the emission prediction location is already giving useful information about the location. Intermediate predicted locations during convergence are presented in Appendix E. Observing the final iteration, the source location has tightened in three dimensions and is clearly visible (highlighted by the circle), with a smear around the source location of lower emitting cells. Very little source emission is now predicted at the 3 m level.

In addition to the main source location, there are cells of high emission rate close to the sensors (the majority of which are located at 1.5 m vertical height) for both iterations shown. This prediction noise near sensors acts as a compensating factor for measurements unexplained by mutual upstream locations and are caused by model errors due to wind field and dispersion estimation. The difference between the mass convergence of the entire domain and the inner region (see Figure 4.3) is the result of this effect. The emission rate-weighted average location of the main emission source was within 4.6 m of the actual release location.

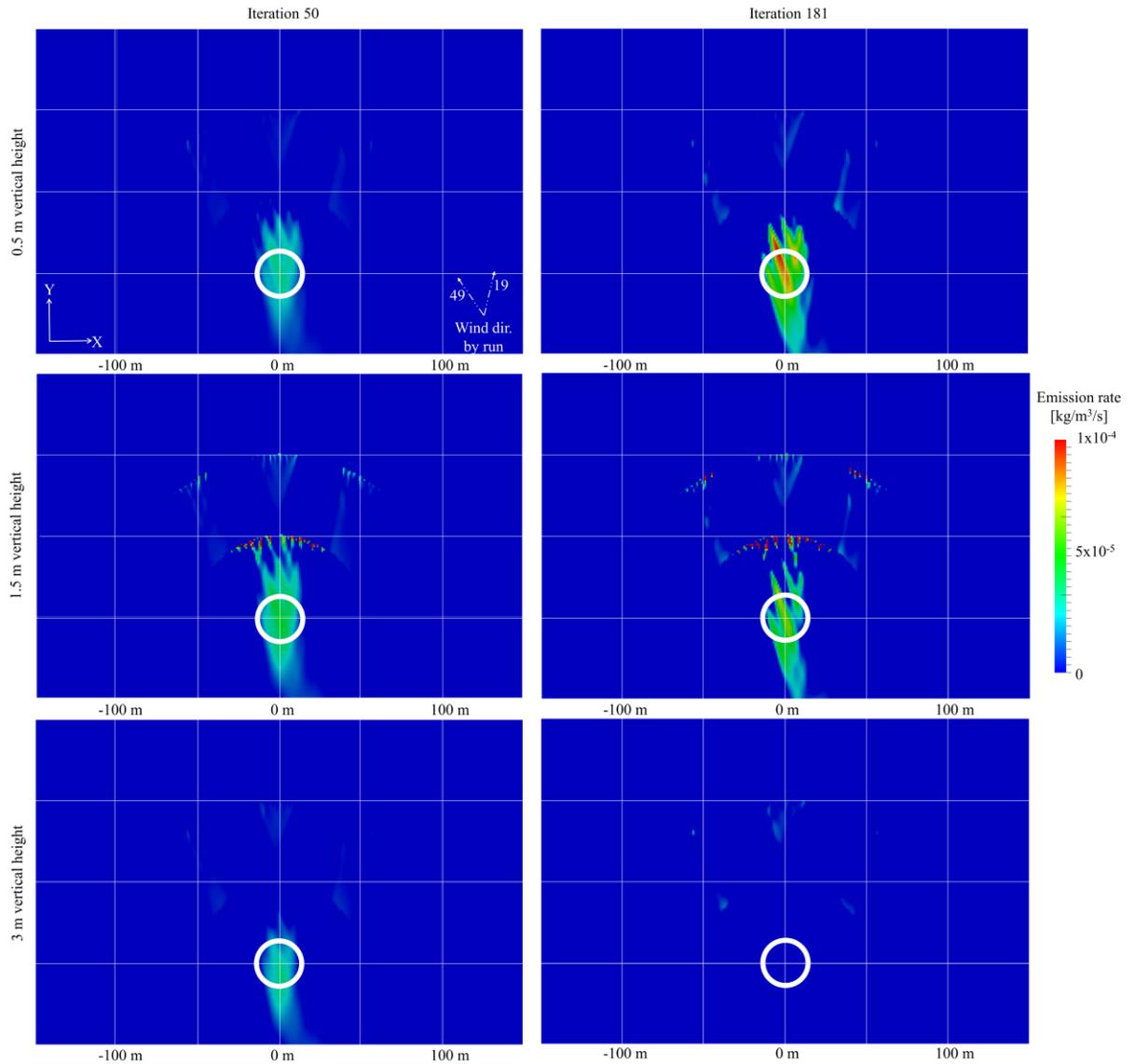


Figure 4.4: Cell emission rate predictions for various iterations at 0.5 m, 1.5 m, and 3 m heights from the ground (lines represent 50 m horizontal increments)

4.5.3 Simulated Gas Plant Test Case

4.5.3.1 Simulation of Measured Concentration Data

In the second test case, a 400 m by 300 m by 100 m subsection of a gas plant in Alberta, Canada containing 15 rectangular bluff bodies ranging in height from 3 m to 10 m was used to create an industrially-relevant bluff-body dominated geometry. These bluff bodies

may be considered buildings or represent large blocks of equipment that block air flow as typically found in such an environment. While not used for this demonstration case, for specific domains it may be more appropriate to treat equipment blocks in more detail or treat them as porous. As shown in top-down view in Figure 4.5a, the rectangular region of interest was situated within a larger triangular domain (3200 m base and height with 400 m vertical thickness). The buildings were placed a minimum of 300 m from the outlet with sufficient vertical clearance to be unaffected by the top boundaries. Duplicates of the buildings were added near the entrance of the triangular domain as a fetch (Figure 4.5) to disturb the flow during the CFD wind simulation. Mesh sizes are summarized in Table 4.2 and a close-up view of the mesh around the building region is shown in Figure 4.5b.

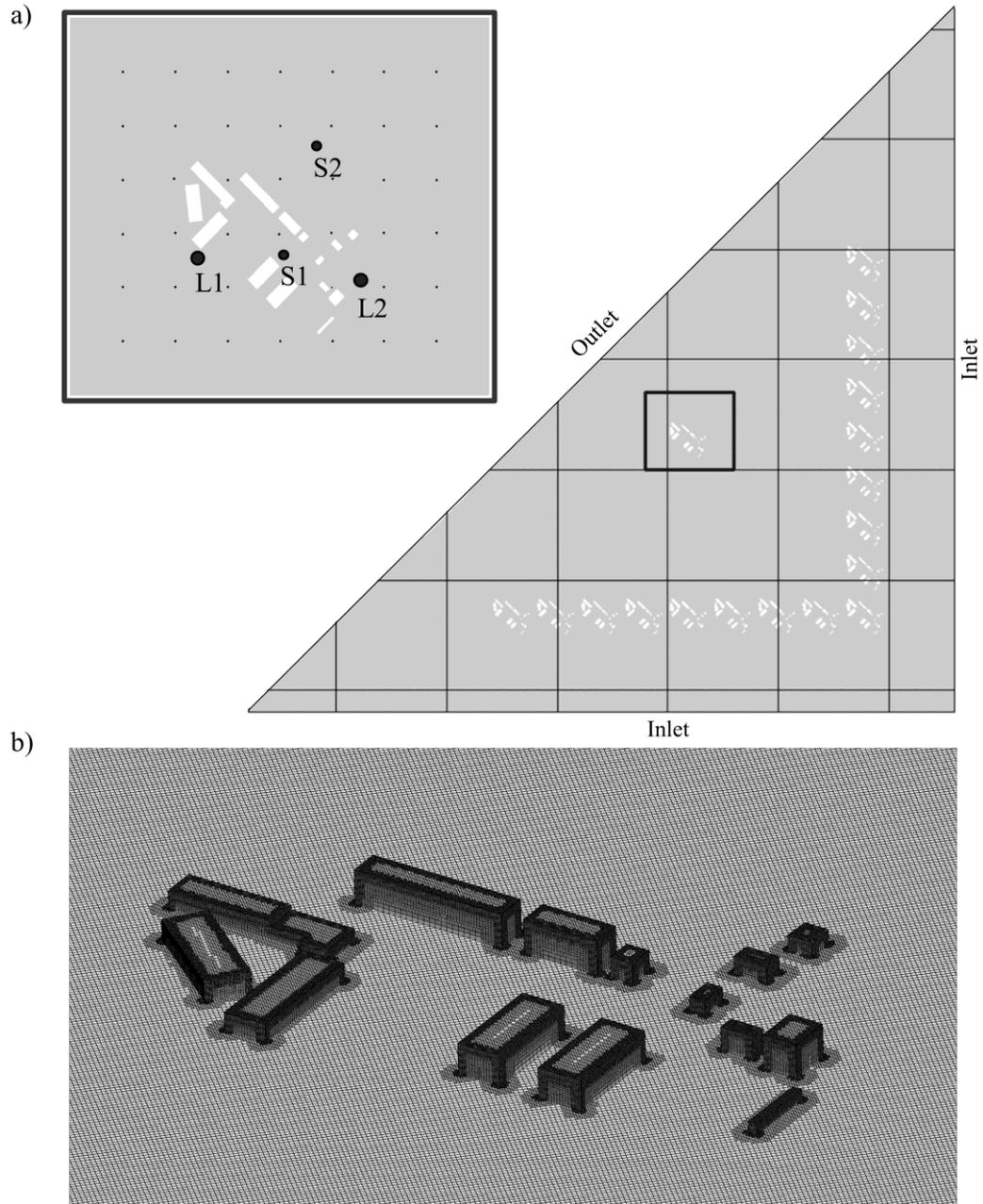


Figure 4.5: Building a) full domain with inner domain cutout marking large (L) and small (S) source locations and sensors (black dots), b) mesh detail

Table 4.2: Mesh details for the simulated gas plant test case

Mesh	Domain bounding size	Number of Cells	$\sqrt[3]{\text{Avg Cell Volume}}$ [m]
Full	3200 m × 3200 m × 400 m	10316272	5.83
Inner	400 m × 350 m × 100 m	1411061	2.15

To generate an appropriate flow field on which to compute virtual sensor concentrations, time-varying inlet wind velocity was taken from a randomly chosen continuous 14 minutes of 10 m-height tower data of the National Wind Technology Center (NREL, 2011) with temporal gusts added by Gaussian randomization based on the measured tower statistics towards the North-West (see Figure 4.6). This information was subsequently fitted to the Richards and Hoxey (1993) profile equations appropriate for neutral stability and specified at the upstream inlet boundaries in the South and East of the domain.

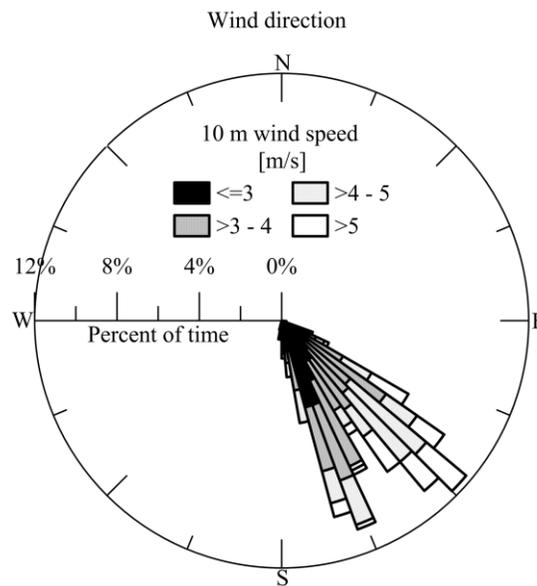


Figure 4.6: Wind at 10 m input height

These inlet conditions were used in a detailed transient incompressible URANS wind simulation was performed using pimpleFoam, with a maximum timestep of 0.03 s (varying to maintain a Courant number less than 1.0). The simulation used the first-order Euler temporal scheme and a k-epsilon RNG turbulence model with a maximum convergence residual of 1×10^{-6} . Initial conditions were set to a simpleFoam steady RANS simulation based on the initial inlet profiles. This detailed bluff-body simulation (with mesh sized to sufficiently resolve influence of building wakes, see Figure 4.5) served as a basis for generating a synthetic gas release to be used in subsequent tests. For the synthetic release generation, only concentrations in the inner domain region surrounding the building were of interest so the subsequent transport modelling was confined to this region and the temporal resolution was reduced to 0.21 s.

To generate simulated sensor measurement data, four simultaneous release locations were chosen (see Figure 4.5) at 1.0 m heights, with large (L) and small (S) release rates set to normalized mass emission rates of 10 and 1 per second respectively. Fugitive emissions survey data suggest that a small number of leaks are typically responsible for the vast majority of fugitive emissions (Picard et al., 1998), represented here by the large sources, while smaller sources may also be present. Forward releases were performed using timesteps of 0.21 s and the first-order Euler temporal scheme and a Gauss limited linear divergence scheme. Virtual concentration sensors were placed every 50 m in a grid from (-50 m, 50 m) to (250 m, 300 m) at a 1.5 m height (working level). Sensors falling within buildings were subsequently removed, leaving a total of forty-one sensors. Concentrations were extracted from these virtual sensors on the detailed wind simulation at these locations every 30 s. Additionally, the 30 s average concentrations were also extracted at the sensors.

4.5.3.2 Evaluation of the Adjoint Optimization Approach

The ability of the adjoint optimization procedure to correctly predict the locations and magnitudes of simulated releases within the complex gas plant geometry was evaluated in two different scenarios. For the first scenario, referred to as the “ideal” case, the simulated instantaneous concentration data was read every 30 s (representing a limited sensor response time) and the detailed wind field (matching that used when generating the test concentration observations) was used in the forward and adjoint transport solves. This allowed for a proof-of-concept trial under idealized conditions with perfect wind data and no noise or sensitivity floor in the extracted sensor observations.

The second case was based on what information could be more practically obtained in the field: the basic geometry, general wind direction and speed, and average concentration values in 30-s intervals. Instead of a transient wind field, a series of steady incompressible RANS simulated winds (with different inflow boundary conditions) were used to approximate the temporal variation. Thirty-second average velocities at a 10 m height were extracted upstream of the central building group (see Figure 4.5). Using neutral boundary profiles based on these velocities (Richards & Hoxey, 1993), steady RANS simulations were performed with simpleFOAM and a k-epsilon RNG turbulence model

until residuals stabilized. A series of these steady RANS wind field simulations (each representing 30 s of physical time) were stitched consecutively to model a 14-minute period (matching the simulated physical time of the comparable fully transient URANS simulation). Figure 4.7 shows the wind rose and the time-varying inlet velocities at an example height of 10 m for the series of steady wind simulations (SSWS). This SSWS was used in all forward and adjoint transport solves during the optimization loop, testing the robustness of the algorithm to imperfect, low quality wind data. In addition, rather than using exact instantaneous sensor concentration data, 30-s-averaged concentrations were extracted from the original detailed transient simulation. These lower-resolution data were used as the observed concentrations (C^{obs}) during the optimization procedure, again testing the robustness of the method. Although additional noise was not added to the sensor readings prior to taking the 30 s averages, separate simulations by Joynes (2013) have shown that the method is highly robust to sensor noise.

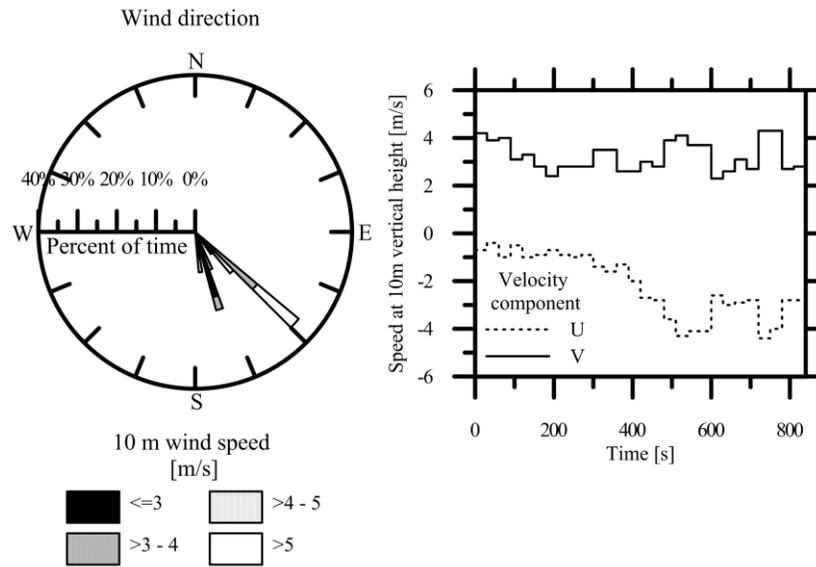


Figure 4.7: Inlet wind at 10 m height for SSWS case

4.5.4 Simulated Gas Plant Case Results

4.5.4.1 Ideal transient matching wind field

For the ideal case with the same wind field used to generate both the observed and forward model concentrations, information loss is primarily due to the scarcity of concentration

sensor locations and temporal readings. Figure 4.8 shows the predicted source locations for this case on both linear and logarithmic intensity scales for a vertical slice of the domain corresponding to the release height. The large sources are accurately located and clearly distinguishable from the background. Although at one-tenth the source strength the smaller sources are not easily seen on the linear scale plot, the log-scale plot reveals predicted emissions in the vicinity of each of these sources. However, there are additional low-level emissions between source L2 and S2 of similar magnitude to S2. This could act as a false positive for an additional small source. However, it is again noted that the present results were achieved using only 14 minutes of simulated wind data, blowing predominantly from the south-east. Additional simulation time, especially from a broader range of wind directions would tighten these predictions, where computations with easterly or westerly winds could be expected to eliminate the false source between L2 and S2. Regardless, even with limited simulation time, the overall predicted emission rates were within 13% of the actual release rate.

4.5.4.2 Identification of Most Important Sources and Magnitudes

A simple algorithmic, post-processing approach was tested to objectively locate and quantify the significant predicted sources. First, all cells with predicted emission rates greater than 1% of the maximum predicted emission rate in the domain were selected. Next regions of connected cells were identified. The emission rates within each of these potential source regions were summed, and regions contributing more than a nominal 5% of the total emission rate of the domain were considered significant. Figure 4.8c,d plots these selected regions on linear and log scales. The computed horizontal positions of the emission-rate-weighted centroids of each region are indicated with a black dot. These predicted locations were both within 4 m of release location of the large sources. For L1, the predicted emission rate in the identified region was 8% higher than the specified release, whereas for L2, the predicted emission rate was 7% higher.

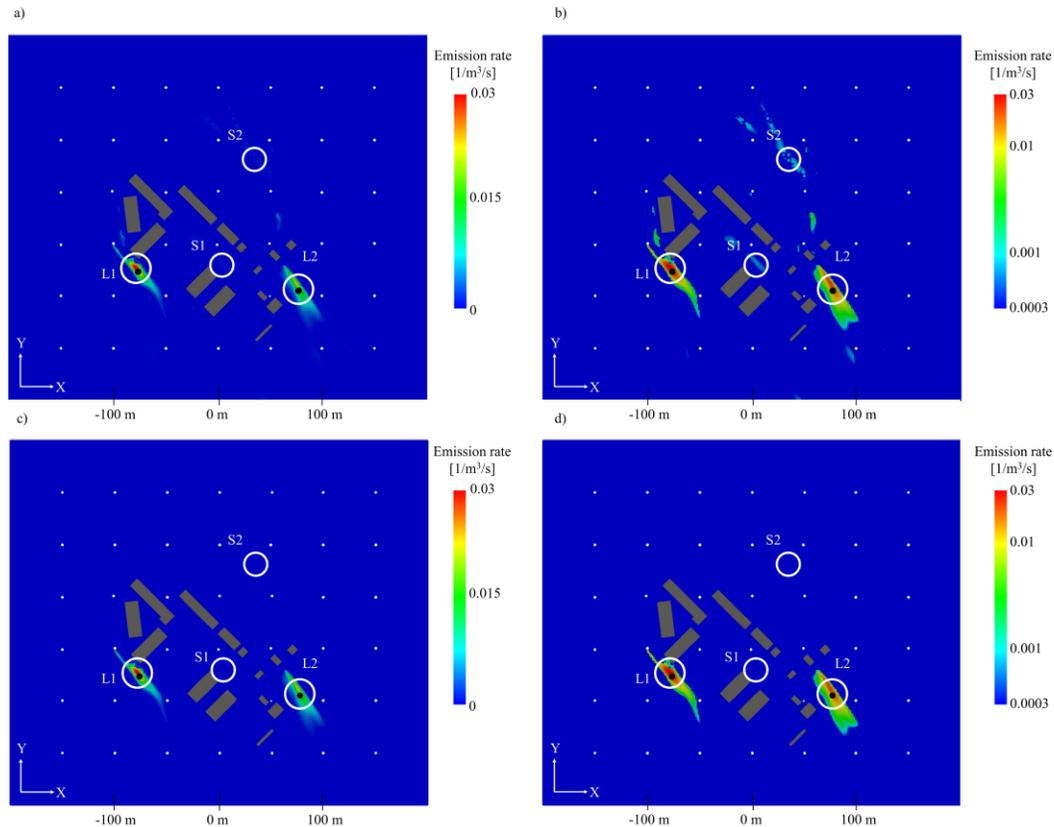


Figure 4.8: Predicted emissions at 1 m height using 30-sec transient ideal wind data with marked actual release locations (white circles) and 50-m sensor grid (white points) for a) linear scale, b) log scale, c) linear scale with regional filter, and d) log scale with regional filter. Emission centroids are marked with black dots.

4.5.4.3 Series of steady wind simulations and averaged concentrations

The much more challenging scenario (but easier to implement in a real-world application) is where the 3D transient wind field is approximated by a series of simplified steady RANS simulated wind fields. Results of this scenario are shown in Figure 4.9. Even with only 14 minutes of physical wind time considered, the large source locations are still readily apparent. However, the predicted sources are also smeared over a longer distance in the predominant south-easterly wind direction and, as shown on a log-scale in Figure 4.9b, the smaller sources are no longer distinguishable from small predictions made upstream of sensors. This is likely because for the 14-minute duration of the wind simulation, both S1 and S2 are downstream of L2 whose plume is complicated in both horizontal and vertical directions by building wakes further obscuring the ten-times smaller S1 and S2. Additional

simulation time with more easterly or westerly winds would reduce the smearing and tighten the predictions, even when using this simplified wind field approximation.

Figure 4.9c shows the identified major source regions using the simple post-processing algorithm described previously. The emission-rate weighted centroid of the region near source L2 was within 12.5 m of the actual release location. Similarly, the location of source L1 was predicted within 4 m. Peak emission rate values are higher in the L1 region than the L2 region, where the closest downstream sensor is behind a set of buildings.

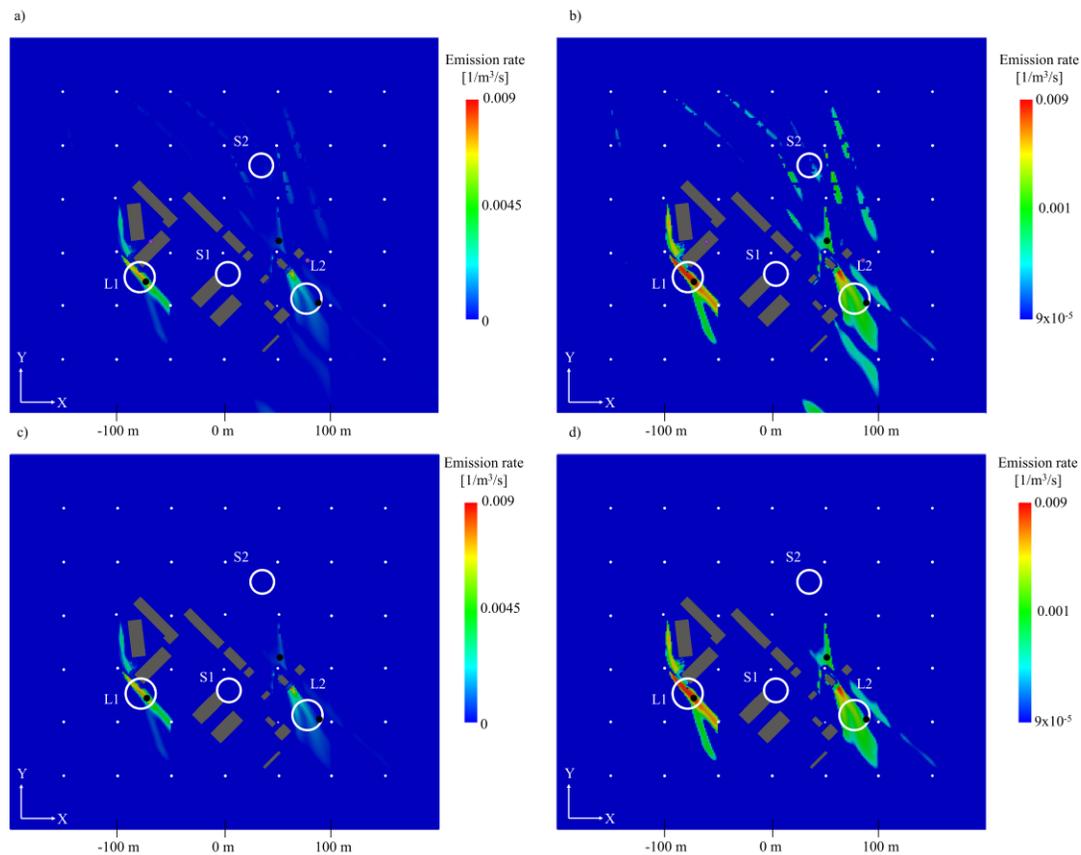


Figure 4.9: Predicted emissions at 1 m height using the series of steady wind simulations (SSWS) and averaged sensor data with marked actual release locations (white circles) and 50-m sensor grid (white points) for a) linear scale, b) log scale, c) linear scale with regional filter, and d) log scale with regional filter. Emission centroids are marked with black dots.

Figure 4.10 plots the predicted emission rates for the identified source regions near L1 and L2 and for the overall domain and compares results for the ideal and the simplified series of steady wind simulation (SSWS) scenarios. Despite the wind field and concentration

data simplifications, the total overall emission rate prediction in this scenario was still within 24% of the actual release rate. Both simulation scenarios were able to predict the magnitudes of the two large sources with similar accuracies. Source L1 was overpredicted by 8% in the ideal scenario, and underpredicted by 2% for the SSWS scenario. Predictions for L2 were 7% high for the ideal scenario and 9% low for the SSWS scenario. The main differences in overall emission rate between the two scenarios are from outside the major predicted source regions. This is consistent with the results of the PPG case. As illustrated in Figure 4.9, the larger overestimation in the SSWS scenario is attributable to the appearance of non-negligible source contributions in the North-East tailing around sensors near S2 and downstream of L2.

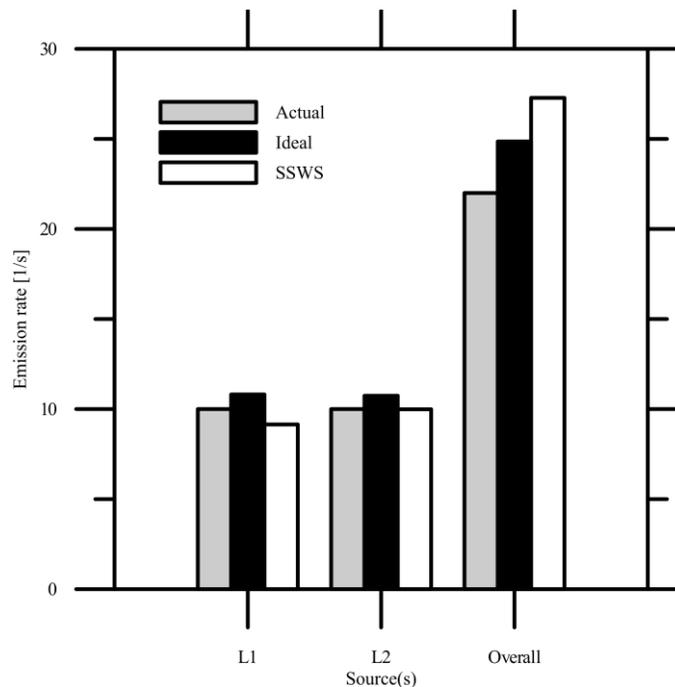


Figure 4.10: Comparison of normalized emissions per second for predicted and actual major source regions for the actual, ideal, and series of steady wind simulations (SSWS)

That these results were achievable based on only 14 minutes of physical wind data with one primary direction is very encouraging. Increasing the amount of data available to the algorithm is likely to improve the estimations. Brereton and Johnson (2012) noted the importance of sensor coverage from varying wind directions in locating emission sources in a complex environment. For a real case, the desired application would allow selection

of multiple appropriate data subsets over the course of a few days, which would likely include more wind variation than the current test case. The applicability of simplified steady wind flows gives rise to the notion of a precomputation database for common conditions that could be used to increase the computational tractability of on-the-fly calculations.

Both the ideal and SSWS tests presented here were limited to relatively in-plane sensor and source release locations, but, as demonstrated by the PPG releases, this is not a requirement. Future work concurrent with extending the potential of pre-computed field database for reducing overall computational cost includes increased variability in source position (including vertical), sensor density, and the robustness to measurement error.

4.6 Conclusions

The adjoint-optimization approach was successfully applied to experimental release data from Project Prairie Grass, predicting the release location within 4.6 m and emission magnitude within 13%. The approach was then further evaluated in numerically simulated releases in a complex, 3-D, bluff-body geometry based on a gas plant in Alberta, Canada. Initial proof of concept tests using only 14-minutes of a transient wind field, demonstrated the ability to locate multiple simultaneous sources, while predicting overall emissions within 13%. More importantly however, adjoint-optimization calculations using a series of steady wind simulations (SSWS) were also found to be capable of correctly locating major sources and predicting overall emission rates within 24%. This latter scenario would be much easier to implement in a real-world application and gives rise to the possibility of developing pre-computed databases of both wind and scalar transport adjoints to reduce computational time. A larger wind data set size would likely improve estimates, but even the present short sample time appears sufficient to meet the objectives of locating and quantifying major sources to direct mitigation actions.

4.7 Acknowledgements

This project was supported by Natural Resources Canada (Project Manager Michael Layer), Petroleum Technology Alliance of Canada (PTAC), and the Natural Sciences and Engineering Research Council of Canada (NSERC, Grant # 261966 and 446199).

Computations were performed on resources and with support provided by the Centre for Advanced Computing (CAC) at Queen's University in Kingston, Ontario. The CAC is funded by: the Canada Foundation for Innovation, the Government of Ontario, and Queen's University.

Chapter 5 Computationally Efficient Quantification of Unknown Fugitive Emissions Sources

This chapter has been written as a paper intended for future submission to Atmospheric Environment.

Brereton, C. A., Campbell, L. J., Johnson, M. R., 2018. “Computationally Efficient Quantification of Unknown Fugitive Emissions Sources”, *draft manuscript to be submitted to Atmospheric Environment.*

This paper was co-authored by the thesis author, Carol Brereton, her thesis supervisors Prof. Matthew Johnson and Prof. Lucy Campbell. Ms. Brereton set-up, developed, and tested the relevant procedures and code; generated synthetic test cases; performed the relevant test cases and initial data analysis; and wrote the first draft of the manuscript. Theory development, final data analysis, and editing and revision of the manuscript were conducted jointly.

5.1 Abstract

Fugitive emissions or unintentional losses of gas (e.g. leaks) are a significant source of greenhouse gas within the oil and gas sector. Previous work has demonstrated the potential of a scalar transport adjoint method for using sparse sensor data to locate and quantify multiple simultaneous unknown fugitive emission sources within a bluff-body dominated facility environment. This paper builds directly on that work and demonstrates the significant computational time reductions that can be achieved by modifying this approach to use a database of pre-computed retro-tracers (PRT). The computational cost, as well as estimated source emission rates and locations, were compared for both an open field release and multiple releases in a bluff-body dominated domain when using the PRT method versus the concurrent gas transport solves used in previous work. For the open-field

release, given the same wind input there were no significant differences in emission rate and location predictions of the two approaches. For the bluff-body dominated multiple source case (a domain representative of an actual gas plant in the upstream energy sector), using simplified wind fields for the PRT database generation allowed major sources to be successfully located and their emission rates predicted within -75 to -32% . When the wind direction coverage was increased to 110° from $\sim 60^\circ$, the emission rate was predicted within $\sim 30\%$ to $\sim 25\%$. The total computational cost for both methods was of a similar order of magnitude when including the initial database generation for the PRT method, but non-reusable computational time was reduced by a factor of 200-600 times making the PRT method feasible on a standard desktop computer once the database is generated. This is a noteworthy achievement as it gives rise to the possibility of continuous or near-continuous characterization of unknown fugitive emissions sources within a complex facility which could allow sources to be identified as they arise.

5.2 Introduction

Fugitive emissions, such as leaks, are an important source of greenhouse gases and can result in safety hazards and lost product. Within the upstream oil and gas industry, these sources are often economic to mitigate, but only once they are found. Facilities can range in size from tens to hundreds of meters and contain buildings, equipment, and piping with numerous valves and fittings that are all potential fugitive release sources. Characterizing fugitive emission sources is essentially a problem of locating and quantifying an unknown number of gas release sources at unknown locations. This type of source term estimation using sparse concentration information and wind estimates is an ill-posed inverse problem that is relevant in a variety of applications over a range of physical scales. An overview of the main solution approaches is available in Rao (2007) and Hutchinson et al. (2017).

Brereton et al. (2018a) recently demonstrated the potential for characterizing unknown sources via a gradient-based optimization procedure using the scalar-transport adjoint, while placing minimal restrictions on source location(s). Results showed that wind field information and concentration data at sparse measurement locations within or downstream of a facility could be used to locate and quantify multiple simultaneous unknown sources within a large, bluff-body dominated 3-D domain. However, the computational resource

requirements of this approach were potentially prohibitive in a practical application. For large and complex domains, the computation time on a desktop workstation could run from days to weeks, or even longer.

Simplified wind fields (which can be computed in advance as part of a database of common flows to reduce computational cost) can be effective input for locating and quantifying sources (Brereton et al., 2018a). However, this idea of a pre-computed database can be extended further. If sensor positions are fixed (i.e. non-mobile or on pre-set paths), inverse (or adjoint) transport modelling can be performed in advance for common wind fields and re-used. Similar inverse transport computations for source term estimation are often used in the context of Bayesian inference on a specific wind field to render Monte-Carlo sampling tractable (e.g. Xue et al., 2017). However, for the present application, wind simplifications will also necessarily be considered, as well as the effects of wind directional coverage when these wind simplifications are present.

This paper demonstrates the significant potential computational saving of inverse sensor-based pre-computations as an extension of the method demonstrated in Brereton et al. (2018a). Computational time savings and results are compared for an open field point release based on a subset of Project Prairie Grass (Barad, 1958) where the wind fields are the same for both methods. Emission characterization predictions are also compared for a synthetic bluff-body case where detailed transient wind fields are approximated by a series of steady RANS simulations used to generate a pre-computed inverse transport database.

5.3 Fugitive Emissions

Leaks in the oil and gas industry have a long tail distribution, with outliers being responsible for the majority of gas released (Brandt et al., 2016). Traditionally, fugitive emissions are located by manual component or qualitative visual IR camera inspection (USEPA, 2016b). Such inspections can be labor intensive, require trained personnel, and may require expensive specialized equipment. As a result, inspections are performed on an intermittent basis. While recent progress has been made in the area of mobile surveys both by motor-vehicle (e.g. Zavala-Araiza et al., 2018; Roscioli et al., 2018; Atherton et al., 2017; Rella et al., 2015; Mitchell et al., 2015; Jackson et al., 2014; Phillips et al., 2013)

and aircraft (e.g. Johnson et al., 2017; Thorpe et al., 2016; Lavoie et al., 2015; Thompson et al., 2015; Tratt et al., 2014; Karion et al., 2013), these methods are also generally limited to providing intermittent snapshots of facility emissions. Fugitive sources can potentially release gas for months or longer before being discovered.

Minimizing the time between the appearance of a release source and its repair is critical in reducing fugitive gas emissions. Ideally, fugitive sources would be discovered as they arise for quick mitigation using a continuous monitoring system. Cost and practicality, however, prohibits installation of sensors at every possible source. Instead, sparse sensor concentration network data can be used in conjunction with sufficiently diverse wind to characterize emission sources. This approach also allows for quantification of source emission rates. As the required time to process acquired concentration and wind data and derive emission source locations and magnitudes is reduced, the potential advantages of this approach become even more significant.

5.4 Emission Characterization Methodology

Details of the scalar transport-adjoint method for characterizing unknown fugitive sources are provided in Brereton et al. (2018a) and are summarized in Figure 5.1a. A 3D computational representation of the physical domain of interest is divided into cells, each with an estimated emission rate. Gas releases are modelled forward in time and compared with concentration sensor network readings. If the mismatch is too large, an inverse approach is used to compute the sensitivity of the mismatch from the sensors, which feeds a gradient-based quasi-Newton optimization scheme (L-BFGS-B, Liu & Nocedal, 1989). Wind field estimates are required as input and performed in advance. Beginning with an initial estimation of no sources, the emission characterization is improved iteratively. Source resolution is limited only by mesh size. Area, line, and point sources can be treated as combinations of cells where emission rates are assumed to be constant over the chosen measurement interval.

5.4.1 Concurrent-Transport-Computation (CTC) Approach

Figure 5.1a outlines the Concurrent-Transport-Computation (CTC) approach, where the transport models are solved within each optimization loop iteration. The forward transport

solve determines model concentrations at sensors and the adjoint transport solve determines sensitivity of upstream locations to the current mismatch at the sensors as described in the following subsections. Pre-computed wind field information is used as input for both transport models.

5.4.1.1 *Forward transport*

Model concentrations at sensors must be determined to compare with sensor measurements for the emission rates estimate. Assuming an incompressible flow field that is not impacted by the gas release, concentration (C) evolution over time (t) may be found with the advection-diffusion equation,

$$\frac{\partial C}{\partial t} + \nabla \cdot (C\mathbf{u} - D\nabla C) - S = 0. \quad (5.1)$$

Eq. 5.1 is an expression of mass conservation using the gradient-diffusion approximation (sometimes called gradient-transfer theory, eddy-diffusivity theory, or K-theory) (Wallace and Hobbs, 2006) for turbulent scale motions. The velocity (\mathbf{u}) [length/time], diffusivity (D) [length²/time], and source term (S) [mass/length³/time] must be specified.

5.4.1.2 *Cost function and its gradient*

The cost function (J) is a measure of how well the emission characterization matches the sensor readings and is defined as

$$J(\mathbf{p}) = \int_{t_0}^{t_f} \int_{\Omega} g d\mathbf{x} dt \quad (5.2)$$

where the misfit g was selected to provide a least-square minimization of the difference between model and measured concentrations where observations exist (Brereton et al., 2018a).

The gradient of the cost function with respect to an individual source parameter p_i (the emission rate in cell i) is (Brereton et al., 2018a)

$$\frac{\partial J}{\partial p_i} = \int_{t_0}^{t_f} \int_{\Omega} \lambda \frac{\partial S}{\partial p_i} d\mathbf{x}dt, \quad (5.3)$$

where $\partial S/\partial p_i$ has a value of one in cell i and zero elsewhere. The appropriate λ is the solution to

$$-\frac{\partial \lambda}{\partial t} + \nabla \cdot (-\lambda \mathbf{u} - D^T \nabla \lambda) - \frac{\partial g}{\partial C} = 0 \quad (5.4)$$

with boundary conditions

$$\begin{aligned} \lambda(\mathbf{x}, t = t_f) &= 0, \\ (\nabla \lambda) \cdot \mathbf{n} &= 0 \text{ where } \mathbf{u} \cdot \mathbf{n} \leq 0, \\ (\lambda \mathbf{u} + D^T \nabla \lambda) \cdot \mathbf{n} &= 0 \text{ where } \mathbf{u} \cdot \mathbf{n} > 0. \end{aligned} \quad (5.5)$$

Once the cost function and its gradient are known, these are used by the optimization loop to improve the source parameters or end the optimization process.

5.4.1.3 *CTC Approach with pre-computed, simplified wind fields*

Modelling the transient wind field, as required to enable subsequent transport modelling, requires significant computation time. However, common flow conditions can be modelled in advance and used and re-used during appropriate reoccurring conditions to reduce this computational expense. Brereton et al. (2018a) successfully replaced detailed transient wind fields with a series of steady wind simulations (SSWS) (with different inlet conditions) stitched consecutively in time. Using this SSWS, concentrations could then be modelled with Eq. 5.1 and used in the optimization procedure outlined in Figure 5.1a. The duration of each steady-flow period must be chosen to allow the wind field to be representative of the actual transient wind but is not otherwise related to measurement intervals or locations, domain size, or velocity.

5.4.2 **Pre-computed Retro-Tracer (PRT) Approach**

As noted in the introduction, a key objective of this paper is to extend the idea of a pre-computed series of steady wind simulations to seek further computational savings that

might allow real-time or near-real-time characterization of fugitive sources within a facility. The premise of this approach is that a retro-tracer concentration representing how much air from a particular location reaches a sensor during its measurement interval can be used as an alternative to computing concentrations with the advection-diffusion equation. A collection of retro-tracers can then be used to determine sensor concentration for any combination of source distributions, as is discussed at length in Hourdin and Talagrand (2006). The critical advantage of this approach is that it makes it possible to pre-compute a database of retro-tracers, extending the benefit of pre-computed wind fields and allowing transport solutions to also be precomputed and reused. This has the potential for significant computational savings. This pre-computed retro-tracer (PRT) approach is outlined in Figure 5.1b.

5.4.2.1 *Retro-tracer transport and sensor concentration*

In an Eulerian reference frame and assuming incompressible flow, the retro-tracer equation (Hourdin and Talagrand, 2006) may be rearranged into its volumetric form, which is equivalent to Eq. 5.1. Using backward integration,

$$-\frac{\partial C^*}{\partial t} + \nabla \cdot (-C^* \mathbf{u} - D^T \nabla C^*) - \mu = 0 \quad (5.6)$$

allows the retro-tracer concentration (C^*) to be found for a measurement distribution μ (where a unit release per volume is made at the measurement cell location over the measurement interval). Integrating over the spatial (Ω) and temporal domain from initial (t_i) to final (t_f) times, the mean concentration at the sensor (C_{sensor}) over the release interval is

$$C_{sensor} = \int_{t_i}^{t_f} \int_{\Omega} S C^* d\mathbf{x} dt \quad (5.7)$$

Unlike the advection-diffusion equation (Eq. 5.1), determining the concentrations with Eq. 5.6 and Eq. 5.7 requires a separate solution for each sensor measurement. However, the transport solution can be found without specifying the source distribution S , which is only required for the final integration in Eq. 5.7. Thus, the retro-tracer transport solution can be re-used to determine sensor concentration for different source distributions. Since Eq. 5.6

is the same form as Eq. 5.4 and the equations are linear, the retro-tracer pre-computations can be re-used to determine the cost function gradient (Eq. 5.4) if the source term is appropriately re-scaled.

5.4.2.2 *Challenge of Initial Conditions when Pre-Computing Gas Transport*

There is one important complication that arises when pre-computing the transport solution for a domain with varying inlet wind conditions. Essentially, the pre-computed transport solution can only account for source influence within the equivalent replaced time interval. Gas already in the domain from the previous time-interval needs to be accounted for as an initial condition. Unfortunately, this is impractical to implement in a precomputed library. For each given inlet wind condition (direction and speed) during a selected time-interval for pre-computation of the transport solution, there could be many different initial conditions (based on the range of possibilities of the wind condition for the previous interval). In a practical application, the permutations of all wind conditions and previous interval wind conditions are likely too numerous to facilitate creation and storage of a pre-computed library of solutions.

The alternative is to allow time for the gas initially in the domain to be transported out prior to comparing model concentrations with observations. Thus, by purposely ignoring sensor data at the start of each new interval (corresponding to a changed inlet condition) the sensors are only influenced by gas emitted within the interval. In practice, the required wait time can be simply specified as the estimated residence time of gas in the domain based on the current inlet wind condition. While this reduces the amount of concentration data available for comparison during the optimization, it is necessary to preclude model error due to unmatched initial conditions. This loss of data can also be compensated for by collecting more data at a broader range of conditions, as further discussed below.

5.4.3 Comparison of Approaches and Implementation

The outlines of both the original (CTC) and pre-computed retro tracer (PRT) procedures for characterizing unknown sources are given in Figure 5.1, which highlights the wind determination and transport solution steps. Not including the precomputed wind field estimation, in the CTC approach, solving the transport equations (Eq. 5.1 and Eq. 5.4) is

most of the computational cost. The overall solution time depends on both the cost of an individual transport solution and how many times these solutions must be computed.

For the PRT approach (Figure 5.1b), retro-plume concentrations are used as pre-computed input, allowing the transport solver calls (Eq. 5.6) to remain outside of the iterative optimization process. Note that an equivalent advection-diffusion (Eq. 5.1) database would require that computations be performed for every cell within the domain rather than the limited number of cells containing sensor positions.

For one-off source characterizations, any computational savings caused by removing the transport solves from the optimization loop may be offset by the need to compute retro-tracers for each sensor and wind combination. However, for a permanent sensor setup on a geometry that is unlikely to change, these pre-computations become re-useable under similar wind conditions. Additionally, if a database of transport solves is performed for common wind conditions in advance, the optimization procedure no longer requires concurrent access to computational resources capable of detailed CFD modelling. Such resources are now only required during the initial setup phase. Scaling and summing the pre-computed transport solutions becomes feasible on a typical desktop machine even for highly detailed meshes.

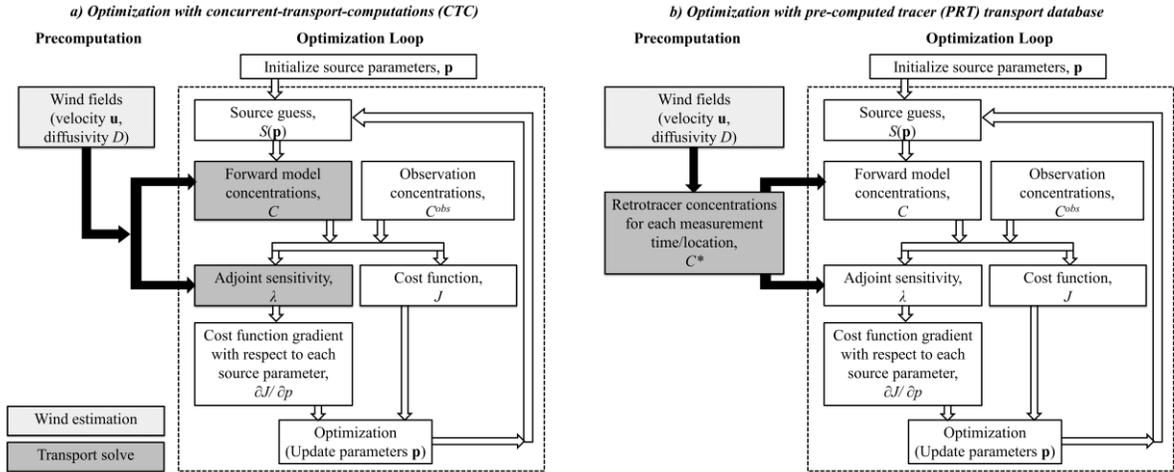


Figure 5.1: Simplified optimization steps with major variables using a) concurrent transport computations (CTC) (modified from Brereton et al., 2018a), and b) pre-computed database of retro-plumes

5.4.3.1 Implementation

Transport solvers were implemented with the OpenFOAM toolkit and the L-BFGS-B algorithm used the code by Zhu et al. (2011) as described in Brereton et al. (2018a). The optimization implementation was the same for both the CTC and pre-computed approaches, but requests for the cost function and its gradient were satisfied differently.

When using the concurrent transport solver solutions, the optimization loop called a remote supercomputer to run the transport model for the current source guess before continuing with the optimization. When instead using the pre-computations, requests for the cost function and its gradient result were completed by the local desktop machine, which could efficiently sum and scale the pre-computed fields. These latter operations were performed in serial but could be parallelized to further increase computational speed. The pre-computation database consisted of steady unit inverse releases on all relevant wind fields for each test case. In the results presented below, computational times have been adjusted to an equivalent single processor for comparison purposes.

5.5 Results

Two main test cases were chosen, which matched those used in Brereton et al. (2018a) allowing results with and without the pre-computation approach to be directly compared. The first test case considered a subset of two releases from the Project Prairie Grass (Barad, 1958) open field experiments. The second test case considered multi-source release simulations in a simplified industrial geometry based on a gas plant in Alberta, Canada. Results for each case are discussed separately below.

5.5.1 Open Field Test Case

Project Prairie Grass (PPG) (Barad, 1958) consisted of a set of SO₂ releases in an approximately flat open field (vertical variation ~1 m) under a variety of conditions. Ten-minute average concentration data were recorded from sensors at 1.5 m vertical heights in five sensor arcs (50 m, 100 m, 200 m, 400 m, and 800 m from the source location) and on six 17.5 m tall towers 100 m from the source.

A 300 m × 200 m × 50 m simulation domain was chosen, which encompassed the first two sensor arcs and tower sensors. This domain was discretized into a 1.5 million cell rectangular mesh (minimum 0.1 m vertical ground spacing with a maximum cell height expansion ratio of 1.2 between adjacent cells). From the PPG data sets, a subset of two releases (Run 19 and Run 49) were chosen. These had similar release rates but different wind directions (where a minimum of two different wind directions is necessary for source location) with no tracer gas entering the domain from outside.

Wind profiles (Alinot, 2003) based on the experimental Obukhov length and friction velocity, with specified roughness height of 0.006 m for each release were mapped onto the domain mesh with an assumed turbulent Schmidt number of 0.67. These wind fields were used as input in the Concurrent-Transport-Computation (CTC) method (see Figure 5.1a) or when generating the pre-computed database of unit retro-tracer (PRT) releases (see Figure 5.1b). Results were compared with the original CTC results presented in Brereton et al. (2018a).

5.5.1.1 *Location prediction*

Figure 5.2 shows the predicted source emission rates per volume at a 0.5 m vertical slice of the domain. Results using the CTC approach are shown in Figure 5.2a and results using the PRT approach are shown in Figure 5.2b. In each figure, the actual source location is marked by the white circle. The results are nearly identical for the two approaches. The same major prediction region shapes are visible, including trails near sensors that compensate for model error. Although the variability of cell emission rates within the source region is slightly higher for the PRT approach, this difference is not likely to be significant. The emission-rate-weighted centroid of the predicted source region (calculated for connected cells around the cell with the highest emission rate per volume and having greater than 1% of this rate) differed by less than 1 m (less than 1 cell width) between the two methods.

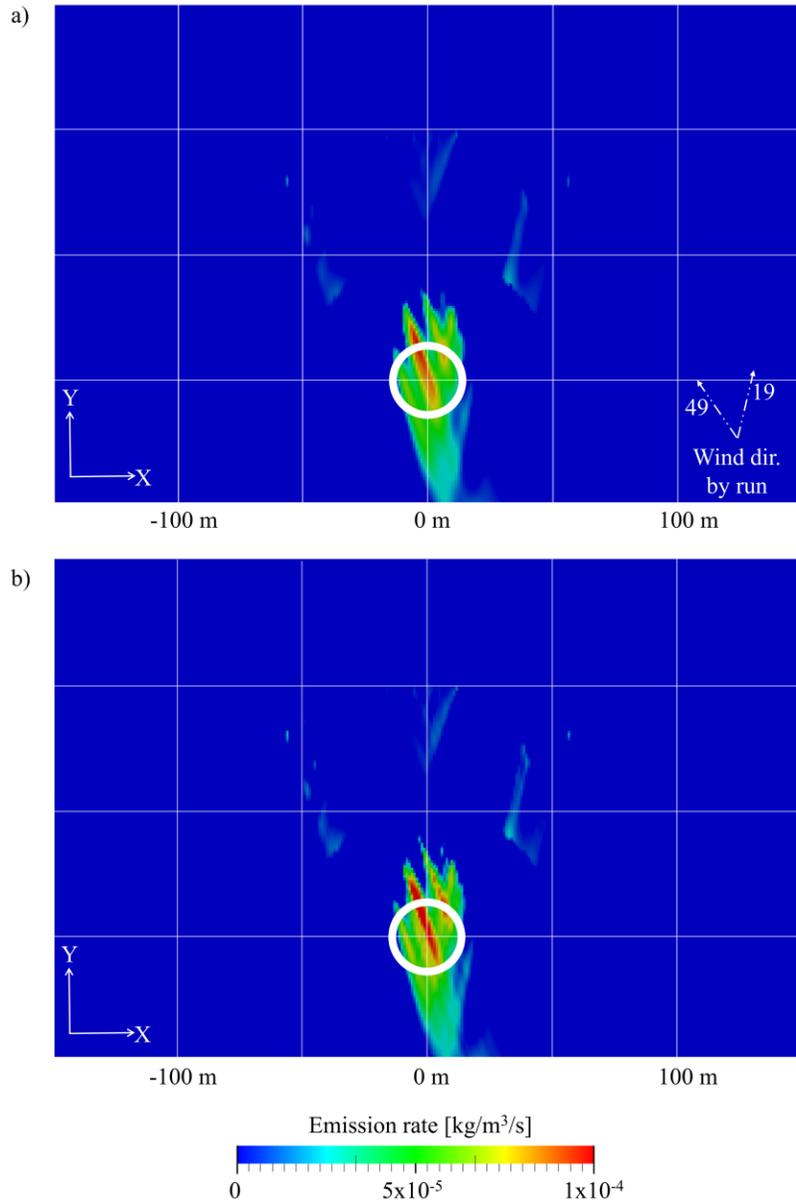


Figure 5.2: Emission rate predictions at 0.5 m vertical height using sensor concentrations and cost function gradient computed with a) the CTC method (Brereton et al., 2018a), and b) PRT method

5.5.1.2 Comparison of Emission Rate Predictions, Cost Function Convergence, and Required Computational Effort

Figure 5.3a compares the convergence of the cost function and predicted emissions for the two approaches. Figure 5.3b compares the required computational time. The convergence characteristics of the two approaches are so similar they are difficult to distinguish. The

overlap of the blue and black lines and symbols in Figure 5.3a suggests the PRT on its own (i.e. without additional wind field simplifications) results in no loss of accuracy. However, as shown in Figure 5.3b, the computational time savings are substantial.

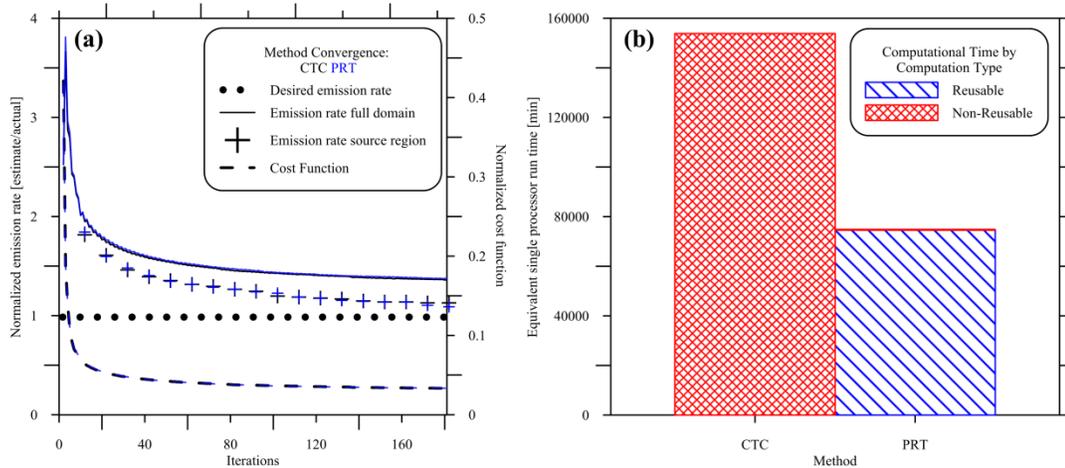


Figure 5.3: a) Convergence of the predicted emission rate (normalized by the actual release value) and cost function (normalized by the cost function for no sources) for the CTC (black) and PRT (blue) methods. b) Equivalent single processor computational time for the CTC and PRT methods. Potentially reusable computations are shown in blue diagonal hatching and non-reusable computations are shown in red cross-hatching.

Converting computational times to a single processor equivalent (actual transport solves were performed in parallel) shows that even the overall computational time is reduced by more than half when using the pre-computed approach. However, more importantly, as indicated by the blue diagonal hatching, almost all of time in the PRT approach is spent generating the reusable transport computations. Stored in a database, results of these computations could be re-used indefinitely so long as the domain and sensor network geometry does not change. Less than 0.5% of the total time for the PRT method was spent on the optimization loop (non-reusable computations). Thus, once the initial database was created, the PRT approach ran approximately 600 times faster than the CTC method for the open-field test case. This computational savings is sufficient to allow the optimization to efficiently run on a modern desktop computer. This is a noteworthy achievement as it

gives rise to the possibility of near-continuous characterization of unknown fugitive emissions sources, which could allow sources to be identified as they arise.

5.5.2 Bluff Body Test Case Based on an Alberta Gas Plant Geometry

The more challenging test case was based on an Alberta gas plant geometry and consisted of a collection of buildings ranging in height from 3 m to 10 m within a $400 \text{ m} \times 350 \text{ m} \times 100 \text{ m}$ domain. As described in Brereton et al. (2018a), when simulating wind fields to generate the initial case, this inner domain was placed within a larger triangular $3200 \text{ m} \times 3200 \text{ m} \times 400 \text{ m}$ domain with buildings copies placed upstream as a fetch.

An initial detailed URANS transient wind simulation using the k-epsilon RNG turbulence model was computed on this building configuration (including the upstream fetch) for slightly less than 14.5 minutes of physical time with the wind blowing primarily from the South-East (total direction variation of ~ 60 degrees, Figure 5.4a). Inlet boundaries were based on neutral conditions with speed and direction taken from wind tower data. Additionally, for this paper, three additional shorter transient wind simulations (each representing 5 minutes physical time) were performed for which the inlet profile speed was identical to the initial five minutes of physical time of the transient simulation, but directions were varied (creating data sets with wind from the South-East with different velocity components, East, and South) modestly increasing the total wind direction variation to ~ 110 degrees (Figure 5.4b). These additional simulations were used to investigate the benefits of greater wind direction coverage in compensating for data loss associated with using the PRT method. Previous work (Brereton & Johnson, 2012a) identified broader wind coverage as a key to better source location. Wind directions Figure 5.4) at 10 m vertical height for the upstream tower location (see Figure 5.5) are shown in a-b.

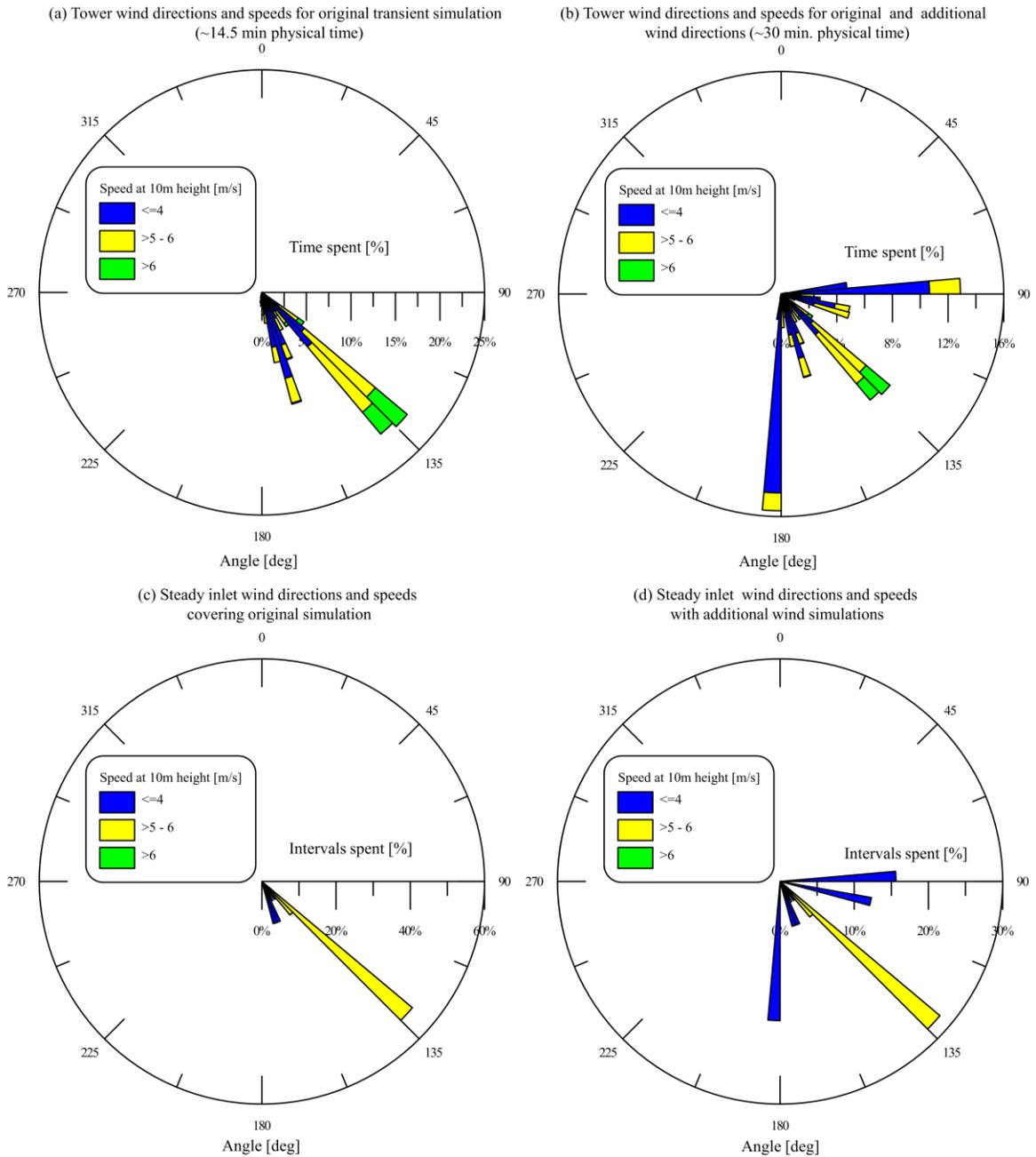


Figure 5.4: Wind rose at 10 m height ahead of main building cluster for a) original transient wind directions, b) original transient plus additional wind directions, c) steady wind intervals for original directions, and d) steady wind intervals with additional directions

For each detailed transient wind field, four simultaneous gas releases were simulated consisting of two large sources (L) ten times the emission rate of two of small sources (S) (see Figure 5.5) at 1.0 m elevation. Concentrations were subsequently extracted at virtual sensors within the domain located on a nominal 50 m grid (omitting sensors that would

overlap with buildings) resulting in 41 sensors at 1.5 m elevation. Data from these virtual sensors were used as inputs in the optimization loops (see Figure 5.1).

Results from two CTC cases from Brereton et al. (2018a) were compared with results from different cases using the PRT method. In the first CTC case, the detailed transient wind (Trans.) was used as input to the optimization loop, with instantaneous concentrations compared every 30 s of physical time. For the second CTC case, a simplified series of steady wind simulations (SSWS) was used in place of the detailed transient wind field, while 30 s average sensor concentrations were extracted from the initial detailed transient simulation. Each steady RANS simulation in the SSWS represented 30 s of physical time, and 28 of these were stitched together in series to create an equivalent 14-minute duration matching the fully transient case. Each RANS simulation was performed on the inner gas plant domain using wind profiles based on the average velocity at a 10 m height ahead of the main building cluster (representing a condition that could be implemented from a single 10-m wind anemometer). Further detail is provided in Brereton et al. (2018a).

For the PRT method, precomputation of the retro-tracers requires a priori specification of the wind fields. Thus, it is only practical to pre-compute retro-tracers for different steady RANS simulations. As noted previously, the retro-plumes can be used to compute sensor concentrations, but within a pre-computed database it is not practical to consider myriad initial conditions corresponding to all possible concentration fields initially within the domain. To overcome this consideration, sufficient time must be given for the gas already within the domain to clear (so initial conditions do not affect the measured concentrations). A multi-criteria algorithm was used to identify quasi-steady wind intervals and select appropriate wait times.

First, wind speeds were extracted from 10 m high virtual wind towers (shown in Figure 5.5) within the detailed URANS simulations. This approach matches what would be easily achieved in a real-world field application. Using the tower most upstream of the building region, average wind speeds at a 2 m height were estimated (from the boundary profile equations). The required time for the flow to clear the domain was estimated using this average wind speed and the longest distance from a sensor to the domain boundary. This

clearance time plus an additional 30 s of usable measurement time, was identified as a candidate PRT interval. Intervals were selected with the criteria that the variation in wind direction during this PRT interval was sufficiently steady – defined as having a standard deviation of wind angle of less than 10 degrees over the entire period (including the sensor measurement time) – and the 30 s of usable sensor measurement intervals did not overlap that of any other chosen PRT interval. This assessment is computationally fast and could easily be used to select PRT intervals on the fly.

From the available 14.5-minutes of physical time, 17 quasi-steady PRT intervals were identified (ranging from 96 s to 171 s clearance time plus an additional 30 s of useable sensor measurement time for each). For cases including the additional wind direction coverage, the additional URANS simulations each allowed for five extra intervals (clearance times of 130 s-150 s plus 30 s of useable sensor measurement time). Figure 5.4c-d show the selected steady-interval PRT wind directions compared with original wind data in Figure 5.4a-b.

Thus, while 30 s physical measurement intervals (instantaneous and average) were used in all cases, the number of actual comparisons between the modelled and observed concentrations differed. The total number of measurement sets compared by the cost function in each case is given in Table 5.3. Convergence of the optimization loop was set to a consistent cost-function drop of less than 0.01% of the initial value between iterations.

Table 5.3: Number of measurement intervals available for comparison for each gas plant data set

Data set	Number of measurement comparisons
CTC with transient wind	29
CTC with SSWS	29
PRT covering original wind period only	17
PRT with additional wind coverage	32

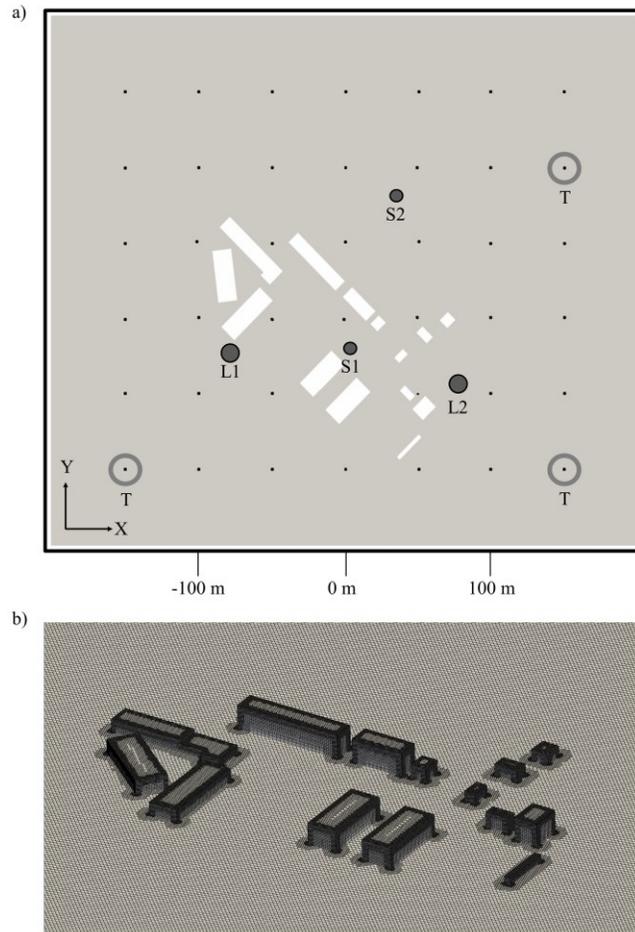


Figure 5.5: a) Top view of inner 3D building domain marking large (L) and small (S) sources (filled circles), wind towers (T) locations (unfilled circles), and 50 m spaced sensor locations (dots) and b) mesh detail (Brereton et al., 2018a).

5.5.2.1 Location Prediction

Figure 5.6 shows the emission rate per volume predictions at 1.5 m elevation (which provides the best visibility of major source regions in all cases). Results using the CTC method with the SSWS are shown in Figure 5.6a, and results using the PRT method in Figure 5.6b and c. The horizontal position of major emission regions (>5% of the total domain emissions) are marked with black circles (although these may be at varying elevations) and actual source locations indicated in white. To determine major emission regions, cells with rates at least 10% of the domain maximum were selected and then aggregated with contiguous bounded cells having 1% of the maximum per region.

The overall maximum cell emission rate was higher for the PRT cases (Figure 5.6b-c) compared with the SSWS case (Figure 5.6a). The major source regions are clearly visible for all cases except the PRT with the original 17 wind directions. With only these wind directions, the source prediction for L2 is downstream of the buildings next to the actual source. Source L1 is predicted with all methods, through the spread differs between the CTC and PRT cases. With even a small amount of additional wind coverage (total wind direction variation of 60 vs. 110 degrees), the L1 source region shrinks by removing some of the predicted tail (reducing the search space) although two separate sources are indicated. For source L2, the location prediction is improved greatly with the inclusion of additional wind, with source P4 becoming a significant emission prediction region (more significant than P3).

If the smaller wind data set is used (Figure 5.6b), two additional spurious sources appear toward the bottom of the image which are 18% and 11% larger than source prediction near S2. Increasing the directional coverage (Figure 5.6c) reduces the influence of this region, eliminating one of the sources as a significant region and reducing the other by 30%. This suggests that added wind coverage can compensate for the simplifications and wind data exclusion (i.e. as required for domain clearance time between changes in averaged wind direction) inherent to the PRT method.

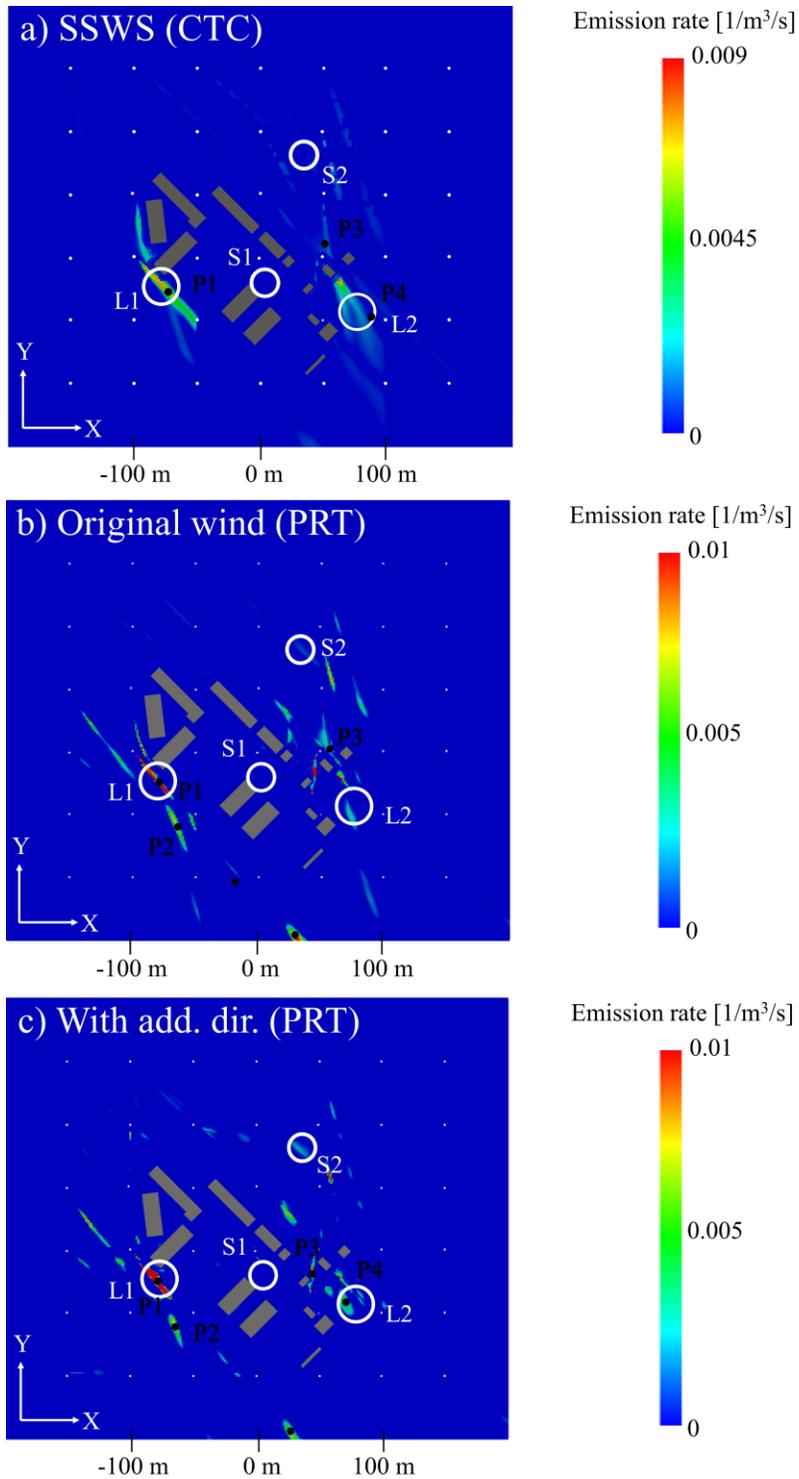


Figure 5.6: Emission rate predictions at 1.5 m slice for a) the SSWS case using the CTC method, b) the original wind case using the PRT method, and c) the PRT method with modest increase in wind directional coverage. Source locations (white circles), sensor locations (white dots) and major source prediction region centroids (black dots) are marked.

5.5.2.2 Emission rate prediction

Emission rates of predicted sources using the PRT and CTC methods are compared in Figure 5.7. For the PRT case with the smallest data set (see Figure 5.6b), the main emission region (P1) is ~50% lower than the actual rate of L1. If the nearby source P2 (overlapping the region predicted by the CTC method) is included, the combined rate increases to within ~30%. Adding a modest amount of additional wind directions as shown in (Figure 5.6c) improves the overall emission rate estimate slightly, but increases the proportion correctly attributed at the actual location of L1.

The predictions for source L2 are also underestimated in both PRT cases. For this source, sensor concentrations are heavily influenced by the presence of nearby buildings. For the PRT method with the smaller dataset, the nearest predicted source at P3 underestimated L2 by ~75%. The added wind coverage improved this to an underestimation of ~25% with the majority of the emissions now coming from P4 (which is closer to the actual source location than P3). It is expected that any additional wind data from the east or north would further improve the results.

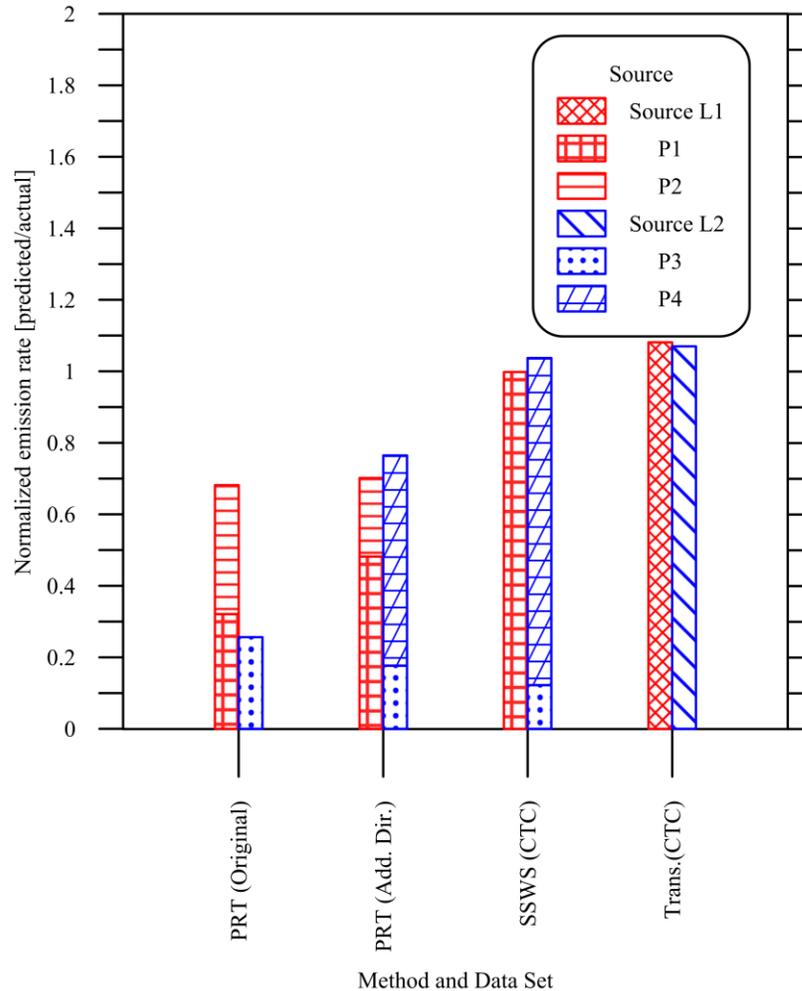


Figure 5.7: Emission rate predictions for sources L1 and L2 for various methods and data sets. Specific source points P1, P2, P3, P4 defined in previous figure.

Overall, emission rates of source L1 and L2 could each be predicted within ~30% even with under 30 minutes of wind data and limited directional coverage. While the predicted emission rates with the PTC method were less accurate than those using the CTC method in Brereton et al. (2018a), the CTC cases also benefit from greater temporal wind variation. Extrapolating the improvements seen with the increase in wind directions suggest that the PRT could potentially match the accuracy of the CTC method.

5.5.2.3 Computational Cost

Figure 5.8 compares the computational time for full CTC cases with detailed wind information to the computational time to develop and use the pre-computed transport

database (PRT method). The time to generate the PRT database is comparable (25 to 85% longer) to that required to the complete CTC case runs. However, the PRT database has the significant advantage that it can be reused in the future during similar wind conditions. Additionally, it is not necessary to perform the computations in series (allowing for greater parallelization). Less than 0.5% of the computational time in the PRT approach was dedicated to the non-reusable optimization loop iterations as compared with 100% of time for the CTC method. This represents a factor of ~ 200 increase in on the fly solution speed. Database generation computational time approximately scales by the number of sensors and wind directions. If the database is already generated, the optimization loop can be performed on a regular desktop machine in hours, which can potentially be further improved through parallelization.

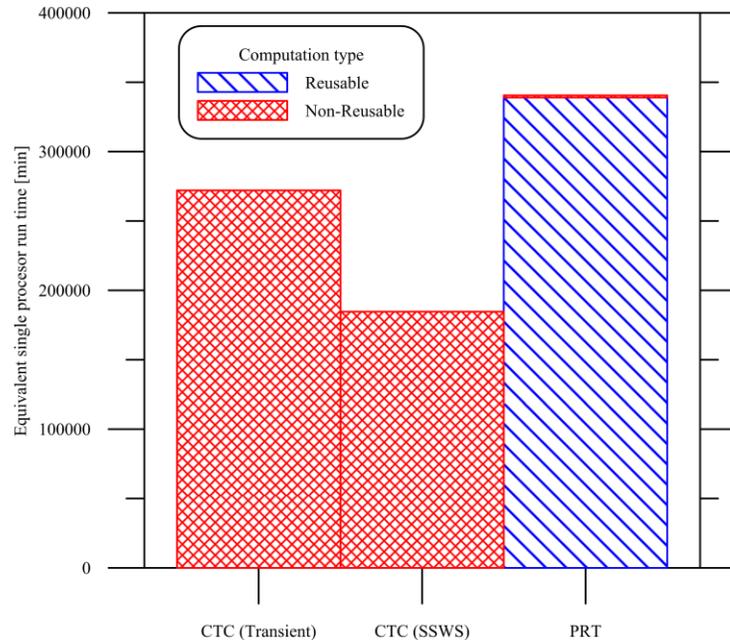


Figure 5.8: Computational time using a single processor for CTC method (Brereton et al., 2018a) using a full transient wind field, CTC method using a series of steady state RANS wind simulations (SSWS) consecutively placed in 30 s intervals to approximate the transient flow, and PRT method (original wind directions).

5.6 Conclusions

A pre-computed retro-tracer approach to solving the scalar transport adjoint was able to predict major source locations for both the open field and synthetic bluff body releases

with equivalent results to CTC method. Using the PRT method (which also requires a limited number of steady-state wind flows) gave comparable emission location predictions to the CTC method when steady-state wind information was used. The accuracy of the emission rate quantification was reduced when the wind was further simplified as part of the database generation process for the bluff body cases, although this could be at least partially and possibly fully compensated for with increased wind direction coverage. However, even with the limited (<30 minutes) of wind data used in the test cases, the emission rates could be usefully estimated with the PRT method and large source locations could be distinguished while reducing computational time. Most significantly, non-reusable computational cost was reduced between 200-600 times in the cases tested. For practical applications, this enables useful location and source magnitude calculations on a typical desktop computer in hours, rather than days, weeks, or longer. The PRT method is a promising development in the goal of near-continuous fugitive emissions source location and quantification within industrial facilities.

5.7 Acknowledgements

This project was supported by Natural Resources Canada (Project Manager Michael Layer), Petroleum Technology Alliance of Canada (PTAC), and the Natural Sciences and Engineering Research Council of Canada (NSERC, Grant # 261966 and 446199). Computations were performed on resources and with support provided by the Centre for Advanced Computing (CAC) at Queen's University in Kingston, Ontario. The CAC is funded by: the Canada Foundation for Innovation, the Government of Ontario, and Queen's University.

Chapter 6 Influence of Turbulent Schmidt Number on Fugitive Emissions Source Quantification

This chapter has been written as a paper intended for future submission to Atmospheric Environment.

Brereton, C. A., Campbell, L. J., Johnson, M. R., 2018. “Influence of Turbulent Schmidt Number on Fugitive Emissions Source Quantification”, *draft manuscript to be submitted to Atmospheric Environment.*

This paper was co-authored by the thesis author, Carol Brereton, her thesis supervisors Prof. Matthew Johnson and Prof. Lucy Campbell. Ms. Brereton set-up, developed, and tested the relevant procedures and code; generated synthetic test cases; performed the relevant test cases and initial data analysis; and wrote the first draft of the manuscript. Theory development, final data analysis, and editing and revision of the manuscript were conducted jointly.

6.1 Abstract

Finding and quantifying unknown (fugitive) gas releases from downstream concentration data is a critical, challenging problem with important environmental, safety, economic, and health implications. Many proposed solutions involve wind and gas dispersion modelling for which a turbulent Schmidt number (Sc_t), the ratio of the eddy kinematic viscosity to the turbulent diffusivity, is usually chosen to scale the estimated diffusivity within an optimization process. This model constant has a range of physically reasonable values. Numerical simulations were performed on multiple test cases to quantify the impact of Sc_t uncertainty on the ability to locate and quantify fugitive emissions sources. Both a gradient-based optimization scheme with unconstrained location and an emission rate only optimization with prescribed point locations were investigated. An initial control case

considered synthetic open field point releases generated with different prescribed $Sc_{t,actual}$ values (ranging from 0.33 to 2.0) after which sources were characterized using a different range of $Sc_{t,assumed}$. Additional test cases used published Project Prairie Grass field data. A second case examined synthetic releases within a more complex bluff-body geometry based on an actual gas-plant and investigated source characterization with various Sc_t with and without additional wind and sensor response error. Although major source location predictions were relatively independent of turbulent Schmidt number, it strongly influenced emission rate quantification. For the controlled open field cases, assuming a midrange $Sc_{t,assumed}$ value of 0.67 resulted in a predicted emission rate within of -33% to $+48\%$ of actual for the unconstrained location method over all $Sc_{t,actual}$ values considered. This range increased for fixed location predictions. With the same $Sc_{t,assumed} = 0.67$, the field experiment release was predicted with a $\sim 10\%$ overestimation using the unconstrained location approach, though results could vary from -35% to $+128\%$ for $0.33 \leq Sc_{t,assumed} \leq 2.0$. The presence of bluff bodies and additional wind error tended to dampen the Sc_t effects. This suggests that if computational restrictions do not allow for multiple Sc_t values to be tested during source characterization procedures, open field results could be used to provide a conservative error estimate.

6.2 Introduction

Emission source characterization or determining unknown emission locations and rates from downstream measurements, is useful for a wide variety of applications and has been investigated on spatial scales ranging from continental (Houweling et al., 1999), to urban environments (Chow et al., 2008; Wilson et al., 2009), to industrial plants (Brereton et al., 2018a), to agricultural fields (Wilson et al., 2012). The turbulent Schmidt number can be an influential parameter in these techniques, but it is generally not known with high precision. This paper investigates the effects of the turbulent Schmidt number selection on source characterization. The research is motivated by the problem of finding, quantifying, and fixing sources of fugitive gas emissions within complex industrial facilities typical of the energy sector, though it is relevant to other domains where unknown sources must be quantified from sparse concentration data.

Fugitive emissions, such as leaking valves and fittings, or other uncontrolled and/or unintended gas releases, are important pollutant sources and significant contributors of greenhouse gas (UNFCCC, 2015) and volatile organic compound (VOC) emissions (Brantley et al., 2015). They can also result in safety concerns and induce economic losses due to lost product. Common approaches to locating unknown fugitive sources include intermittent surveys or snapshots using manual component testing and qualitative IR cameras (USEPA, 2016b). Recent work has been done on mobile surveys, both drive-by (e.g. Jackson et al., 2014; Mitchell et al., 2015; Phillips et al., 2013; Rella et al., 2015; Atherton et al., 2017; Roscioli et al., 2018; Zavala-Araiza et al., 2018) and airborne (e.g. Thompson et al., 2015; Thorpe et al., 2016; Tratt et al., 2014; Karion et al., 2013; Lavoie et al., 2015; Johnson et al., 2017), as well as attempts at IR camera quantification (Gålfalk et al., 2016). These methods, however, still require manual labour by trained personal and are, by necessity, infrequent.

An alternative to intermittent measurement surveys is continuous monitoring of a sparse network of sensors distributed throughout a facility (Brereton et al., 2018a, Brereton and Johnson, 2012a), where inverse analysis leveraging knowledge of the domain geometry and wind conditions can be used to locate and quantify unknown sources of gas contributing to readings at downstream sensors. Related source characterization methodologies are overviewed in Rao (2007) at various problem scales. Some examples include genetic algorithms (Allen et al., 2007), Bayesian inference techniques (Keats et al., 2007; Rajaona et al., 2015), direct search methods (Zheng and Chen, 2011), and inverse dispersion modelling (Wilson et al., 2009; Wilson et al., 2012).

While wind fields (and associated turbulent diffusivity information) used in these approaches can be modelled in a variety of ways, practical computational requirements generally restrict wind simulations to RANS CFD models. To determine the diffusivity from these RANS simulated wind fields, most standard methods require the specification of the turbulent Schmidt number (the ratio of the eddy kinematic viscosity to the turbulent diffusivity, Sc_t), a model constant which has no universal value (Gualtieri et al., 2017). For known emission source locations and rates, the selection of the turbulent Schmidt number has potentially strong effects on model results (Tominaga and Stathopoulos, 2007;

Gualtieri et al., 2017). This is exacerbated in the current application, in which the source location(s) and release rate(s) may also vary.

This paper investigates the effects of incorrectly estimating the turbulent Schmidt number when determining emission rates from both unknown and known release locations. For the first scenario, tests were performed using the scalar transport adjoint sensor-based methodology outlined in Brereton et al., (2018a). This source characterization approach allows for multiple sources at unknown locations which are located and quantified through optimization of emission rates within cells of a computational model of the physical domain. For the second scenario, source rate optimization was performed assuming known point source locations using a global simulated annealing optimization scheme. Simulations were completed using both an open-field geometry, drawing on experimental data from the Project Prairie Grass experiments (Barad, 1958), and a complex bluff body geometry based on an actual gas plant in Alberta, Canada. For each geometry, the influence of different assumed Schmidt number values on the ability to correctly locate and quantify unknown source(s) or to quantify sources with known locations was evaluated. Results of these simulations help bound the potential range of turbulent Schmidt number induced uncertainty in source characterization problems as specifically relevant to fugitive emissions detection.

6.3 Turbulent Schmidt Number and Scalar Transport Modelling

In Brereton et al. (2018a), forward gas releases for a guessed or estimated source characterization were modelled using the advection-diffusion (or scalar transport) equation in an Eulerian framework. The source characterization was then updated via an optimization procedure. This approach was tractable using inverse (or adjoint) modelling to determine the sensitivity of potential sources to downstream sensors, significantly reducing the amount of transport modelling required (Brereton et al., 2018a). Both the forward scalar transport equation, Eq. 6.1 below, and its adjoint (which determines upstream sensitivity to sensors) require diffusivity and velocity fields as inputs.

Using mass conservation and assuming that the time-varying mean velocity field \mathbf{u} [length/time] is unaffected by input of a dilute gas species, for gas sources (or sinks) S

[mass/time/length³], the evolution of concentration C [mass/length³] of this dilute species with time t may be predicted with the advection-diffusion (or scalar transport) equation:

$$\frac{\partial C}{\partial t} + \nabla \cdot (C\mathbf{u} - D\nabla C) - S = 0. \quad (6.1)$$

The gradient-diffusion approximation, with diffusivity D [length²/time], is used to capture the effects of turbulent motions not dealt with by the bulk velocity. While the velocity field may be determined through separately performed wind modelling, the diffusivity still requires assumptions. Generally, for wind flows, diffusion is dominated by turbulent effects and molecular diffusion can be neglected (Gualtieri et al., 2017). The relation between momentum and mass diffusivity is usually determined by assuming a turbulent Schmidt number, the non-dimensional ratio of the eddy kinematic viscosity (ν_t) to the turbulent diffusivity (D_t):

$$Sc_t = \frac{\nu_t}{D_t} \quad (6.2)$$

As the eddy kinematic viscosity is estimated by the wind model, specifying the turbulent Schmidt number and neglecting molecular diffusion effects gives a scalar diffusivity of:

$$D = \frac{\nu_t}{Sc_t} \quad (6.3)$$

Unlike the analogous Schmidt number for molecular scales, Sc_t has no universal value (Gualtieri et al., 2017). Spalding (1971) and Launder (1978) suggested values of 0.7 and 0.9, respectively, based on various experimental observations and these are often used as default values in commercial codes (Tominaga and Stathopoulos, 2007). A wider range, however, has been determined from various atmospheric measurement studies, and numerical studies have found that modifying this value may improve plume dispersion results.

Tominaga and Stathopoulos (2007) and Gualtieri et al. (2017) provide extensive reviews of Sc_t values for atmospheric flows in various measurement or numerical studies. Tominaga and Stathopoulos (2007) noted that the optimal Sc_t values for atmospheric flows could range from 0.2 to 1.3, while having significant impacts on the predicted dispersion results. They recommended that Sc_t be selected only after considering the dominant

turbulent mass transport effects for each individual case. They also gave a tentative recommendation of $Sc_t = 0.3$ for open country but noted that this may be compensation for RANS model errors for this type of flow. The more recent review by Gualtieri et al. (2017) of values selected for atmospheric flows showed an even larger range of 0.1 to 2.5.

Flesch (2002) estimated turbulent Schmidt number for various stability and wind conditions using pesticide emissions by comparing a turbulent Schmidt number dependent flux-gradient emission rate computation with a second independent flux estimate. A range of 0.18 to 1.34 was found with an average value of 0.6. Additionally, there was no clear relationship between stability, time of day, or wind velocity. Simulation work by Riddle et al. (2004) found that, for isolated stacks in flat terrain, a value of 0.3 improved over the “standard value” of 0.7 when comparing ground level concentrations. Blocken et al. (2008) compared numerical results for pollutant dispersion for isolated stacks and buildings, as well as high momentum exhaust when surrounded by several rooftop structures and found concentration profiles were highly sensitive to turbulent Schmidt number selection.

Wilson et al., (2009) calibrated forward models based on Project Prairie Grass release experiments. Wilson et al. (2013) measured different vertical fluxes to determine turbulent Schmidt number values. Wilson (2015) compared flux footprints, or zone upwind of an instrument that contributes to flux measurements. These works indicated that an Sc_t range of 0.6-0.7 may be more appropriate than selecting unity for atmospheric flows.

Numerical experiments by Yu and Thé (2016), while comparing both turbulence models and turbulent Schmidt number selections for numerical releases around rectangular buildings, noted better performance in various concentration profile metrics at $Sc_t < 0.7$, with a final selection of 0.4.

To the authors’ knowledge, only the recent paper by Xue et al. (2017) concentrated on turbulent Schmidt number effects on unknown location source term estimation. Working with a Bayesian inference framework on a single-source, single wind direction case in a symmetric geometry well-covered by downstream sensors, this study found that different values could move the predicted location (higher values tended to move the location upstream) and change the emission rate prediction. This further raises important questions

for the desired fugitive emissions application, where multiple sources are expected within a complex, asymmetric domain and computational cost restrictions most likely prohibit a Bayesian inference framework in a practical application. Additionally, data from multiple wind directions would be important since site logistics and cost restrictions would likely result in poor sensor coverage of source regions (since fugitive emissions can occur at any one of thousands of possible emission source locations).

Overall, while loose guidelines exist for different flow types, there is no universal value of turbulent Schmidt number for atmospheric flows. Additionally, most numerical studies have been limited to known locations and/or emission rates. In the current paper, the influence of Sc_t effect on location estimation as well as emission rate prediction is investigated for both single and multiple simultaneous sources using an optimization framework with multiple wind directions and two different domains of varying complexity.

6.4 Emission Characterization

The overall goal is to use concentration readings from a sparse sensor network combined with wind field estimates to locate and quantify (i.e. characterize) an unknown number of gas release sources in an outdoor domain potentially containing buildings, pipes and equipment. The physical domain, which might range from tens to hundreds of meters in size depending on the target facility, is represented by a 3D computational domain discretized in space and time. This discretization must be sufficiently detailed to capture important wind effects on plume dispersion such as wakes and recirculation zones of buildings or other bluff bodies. Although low gas concentrations allow the wind field to be treated separately from the gas transport, practical limits on computational time typically necessitate the use of RANS (Reynolds-averaged Navier-Stokes) for environmental CFD simulations on this scale (Gualtieri et al., 2017). This paper utilizes an Eulerian framework in which the computational domain is divided into discrete cell volumes. For unknown location optimization, all cells are considered as potential emitters, limiting source prediction resolution to mesh resolution. In specific test cases when locations are assumed to be known, emissions were allowed to vary from zero only in the cells identified as emission sources.

6.4.1 Optimization procedure to determine unknown source location(s) with unknown emission rate(s)

Figure 6.1 shows an overview of the simulation and optimization procedure to identify unknown emission sources and quantify emissions rates. This approach is fully detailed in Brereton et al. (2018a) and only briefly reviewed here. Using either pre-computed wind field simulations or mapped experimental wind profiles with pre-computed retro-tracers (Brereton et al., 2018b), initially guessed source distributions are iteratively improved via a gradient-based optimization approach, where the source gradients are determined via adjoint transport modelling.

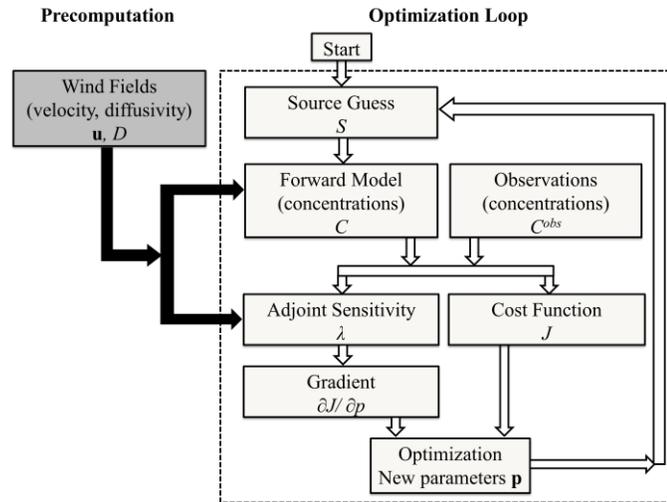


Figure 6.1: Simplified optimization steps with major variables (modified from Brereton et al., 2018a)

6.4.1.1 Forward Transport Model and Cost Function

To determine the fitness of the source guess, gas releases are modelled forward in time for comparison with the observed sensor concentration network readings (C^{obs}). The advection-diffusion (or scalar transport) equation presented in Eq. 6.1 was used. Inlets to the computational domain (based on wind direction) had set concentration values, while outlet and wall boundaries utilized zero gradient boundaries. Input model parameters (\mathbf{p}) are the emission rate in each cell of the domain. The misfit measure (g) was defined as half the squared difference between model concentrations and observations (C^{obs}) where measurements existed, and zero elsewhere. Integrating the misfit over both the spatial and temporal domain gives the cost function, J , which the optimization procedure minimizes.

6.4.1.2 Cost Function Gradient and Inverse Model

The optimization scheme requires the cost function gradient ($\partial J/\partial p_i$) with respect to the optimization parameters (i.e. the emission rate in each cell) to determine where to increase or decrease emission rate estimates. The gradient of the cost function with respect to the source parameters can be determined using the scalar transport adjoint with a forcing term based on the model misfit from observations (details may be found in Brereton et al., 2018a).

The adjoint equation has a similar form to Eq. 6.1 (including the diffusivity term), though is now marched backwards in time against the flow. This can be thought of as an inverse plume run from the sensor locations backwards in time. Also note that to obtain the necessary diffusivity for both the forward and adjoint model calculations, a value for the turbulent Schmidt number must be specified. Once the gradient is known, this can be fed into the optimization scheme to make an iterative improvement to the source guess (Figure 6.1) as part of a least-squares minimization of the cost function.

6.4.1.3 Optimization Scheme

Optimizing the source guess, or emission rate in each cell, could require that millions of variables be considered. The quasi-Newton, gradient-based optimization scheme L-BFGS-B (Liu & Nocedal, 1989) as implemented by Zhu et al., (2011) was selected for its speed and ability to handle large numbers of variables. Emission rates of each cell were optimized from a starting guess of no sources and assumed to be constant throughout the modelled simulation.

6.4.2 Optimization procedure assuming known source location(s) but unknown emission rate(s)

The simpler problem of known source locations but unknown emission rates could arise in cases where additional domain knowledge is available (e.g. detected leaks of unknown magnitude from IR camera investigation) or as a follow-on analysis using location results of the unconstrained characterization described above in Section 6.4.1. For example, in this latter case one might assign a point source location to the centroid of an identified potential source region, or an operator might recognize a particular piece of equipment as

the likely source within an area. This simpler scenario is also relevant to closely related applications in which inverse methods might be used with fence-line or mobile survey data to estimate magnitudes from known sites.

Relative to the unconstrained case, required variables for optimization are significantly reduced from every cell in the domain to only those few cells containing identified sources. Thus, for the known source location case, the global optimization simulated annealing algorithm (Spall, 2003) was thus adopted since it is more appropriate than the gradient-based L-BFGS-B scheme. This simulated annealing method requires only the cost function magnitude. Taking advantage of the linearity of Eq. 6.1, forward unit releases were made from each known source location and sensor concentrations from any combination of source strengths could be found by scaling and superposition. A key question to be answered was whether the influence of unknown turbulent Schmidt number might be reduced if source locations were known or could be assumed a priori.

6.5 Transport Solver Implementation

Transport solvers for the forward source release, Eq. 6.1, and the adjoint sensitivity release, were implemented with the OpenFOAM toolkit, which contains libraries appropriate for finite volume CFD modelling. These transport solvers or the pre-computed data they generated were called at appropriate intervals during the optimization loop (see Figure 6.1). The same forward transport solver was used for the forward unit release required by the simulated annealing optimization (Section 6.4.2). Wind field information, sensor location, and concentration observations were read from separate data files, with the turbulent Schmidt number treated as a constant scalar value used by both forward and adjoint solvers.

6.6 Results

Two main test cases representing both open and built-up areas were chosen to test the effects of turbulent Schmidt number selection. Chavez et al. (2011), studying the effects of adjacent buildings on pollutant releases using both CFD and wind tunnel experiments, noted that adjacent buildings reduced the impact of turbulent Schmidt number selection on numerical simulation of pollutant dispersion, while isolated stack releases and buildings showed the strongest dependence. An open field domain where the highest turbulent

Schmidt number dependence is expected was thus taken as the initial test scenario so that there would be no bluff-body influence on the gas release. However, as built environments are the domains of interest for most fugitive emissions source location applications, a second test case using building geometries based on an Alberta gas plant was created as previously used in Brereton et al. (2018a).

6.6.1 Open Field Case

6.6.1.1 Test Case Details

The open field domain was based on the classic emission release experiment Project Prairie Grass (PPG) (Barad, 1958), where 68 individual SO₂ releases were performed under various conditions in a flat field (vertical variation ± 1 m). Ten-minute average concentrations were measured using sensors at 1.5 m heights arranged in arcs 50 m, 100 m, 200 m, 400 m, and 800 m from the source and mounted on six 17.5 m tall sensor towers 100 m from the source. The source release height was 0.46 m.

A 300 m x 200 m x 50 m model domain was chosen which encompassed the first two sensor arcs (see Figure 6.2). A minimum of two wind directions is required to determine a distinct source location. Two release data sets with no missing data or experimental problems, corresponding to PPG experimental runs 19 and 49, were chosen as a base case (corresponding to Brereton et al., 2018a). The releases rates were similar (within 1%), but wind directions differed. The domain itself was divided into a mesh of 1500000 cells with 0.1 m vertical cell spacing at the ground and increasing with a maximum ratio of 1.2 between adjacent cells.

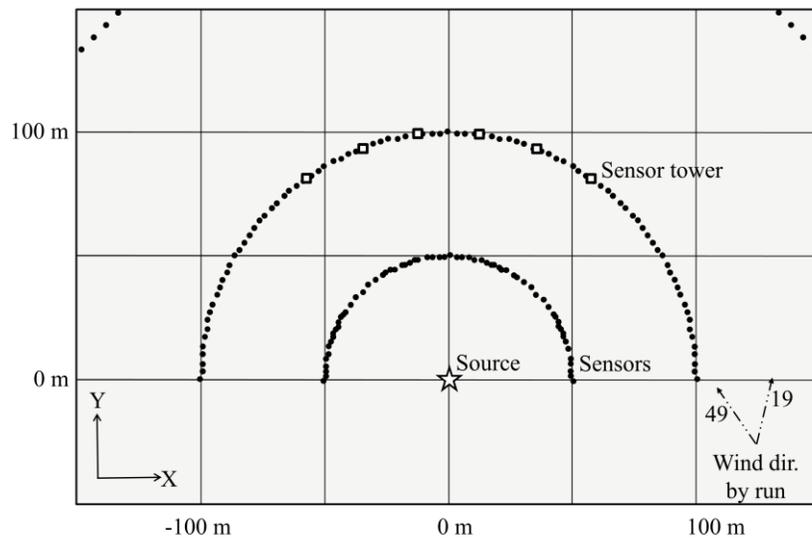


Figure 6.2: PPG Domain with source (star), sensor (dot) and tower (unfilled square) locations (Brereton et al., 2018a)

Wind profiles were mapped directly onto the mesh with an assumed roughness height of 0.006 m and measured experimental friction velocity and Obukhov length. For these particular releases, atmospheric conditions were unstable. Additional details can be found in Brereton et al. (2018a).

Pre-computed inverse releases from the sensors were performed using adjoint transport. These inverse releases ran backward in time from sensors using each wind field and selected turbulent Schmidt numbers as described in Brereton et al. (2018b). Sensor concentrations for specific emission distributions were then computed by appropriate scaling and integration of these inverse releases (see Brereton et al., 2018b for details). The cost function gradient was also computed by scaling and summing these pre-computed inverse releases. As detailed in Brereton et al. (2018b), optimization results using pre-computations are equivalent to those obtained using the standard advection-diffusion equation with gradient computations (Brereton et al., 2018a), but are orders of magnitude more computationally efficient. Timesteps of 0.2 s were used for a physical duration of 60 s for all transport solves. The criteria for convergence was set to a cost function decrease of less than 0.1% of the initial cost function value over 20 iterations.

Forward gas releases were numerically computed for turbulent Schmidt numbers from 0.33 to 2.0 based on the range of possible values suggested in the literature reviewed in Section 6.3. The source release rate matched that of the chosen PPG sets. Extracting the computed concentrations at sensor locations, these synthetic data sets were treated as the observation input (C^{obs}) to the optimization scheme where the ‘actual’ turbulent Schmidt number ($Sc_{t,actual}$ for the purposes of the discussion) is known. Optimizations were performed using various ‘assumed’ Sc_t values from the same full range (i.e. $Sc_{t,assumed} = 0.33$ to 2.0), but not necessarily matching the specified value used to generate the test data. This allowed the potential errors induced by incorrect $Sc_{t,assumed}$ used during the optimization procedure to be isolated and explored over the extreme ranges of a potential mismatch with $Sc_{t,actual}$ consistent with the observations. Subsequent simulations used the actual measured concentration values from PPG experiments while attempting to quantify the source for various $Sc_{t,assumed}$. This allowed the combined influence of Sc_t error and all other complexities of a real-world scenario (e.g. wind model and sensor measurement error) to be investigated.

6.6.1.2 Sc_t influence on unknown release location with unknown emission rate

Using the simulated test data described in the previous section, source characterizations were performed assuming various turbulent Schmidt numbers from 0.33 to 2.0 during the optimization procedure, starting from an initial guess of no sources in the domain. For all combinations of $Sc_{t,actual}$ vs. $Sc_{t,assumed}$, the central location of the estimated release region remained within 1.6 m of the actual release location (i.e. less than 2 horizontal cell widths). Figure 6.3 shows the extremes of the predicted source location variation for the cases where $Sc_{t,assumed} = 2.0$ or 0.33 while $Sc_{t,actual} = 0.33$ or 2.0. When the diffusivity is over-estimated, the predicted location is tight. By contrast, the predicted location is spread when the diffusivity is significantly under-estimated (Figure 6.3b). The general height of the predicted source region also increases when Sc_t is underestimated (note that the sensors are higher than the actual source location). In addition, spurious source estimations near the sensor locations increase, presumably compensating for the increased model input error. For extreme underestimations, spurious sources near sensors are also

visible, though at fewer locations. The overall prediction spread is smaller when $Sc_{t,assumed}$ under predicts $Sc_{t,actual}$.

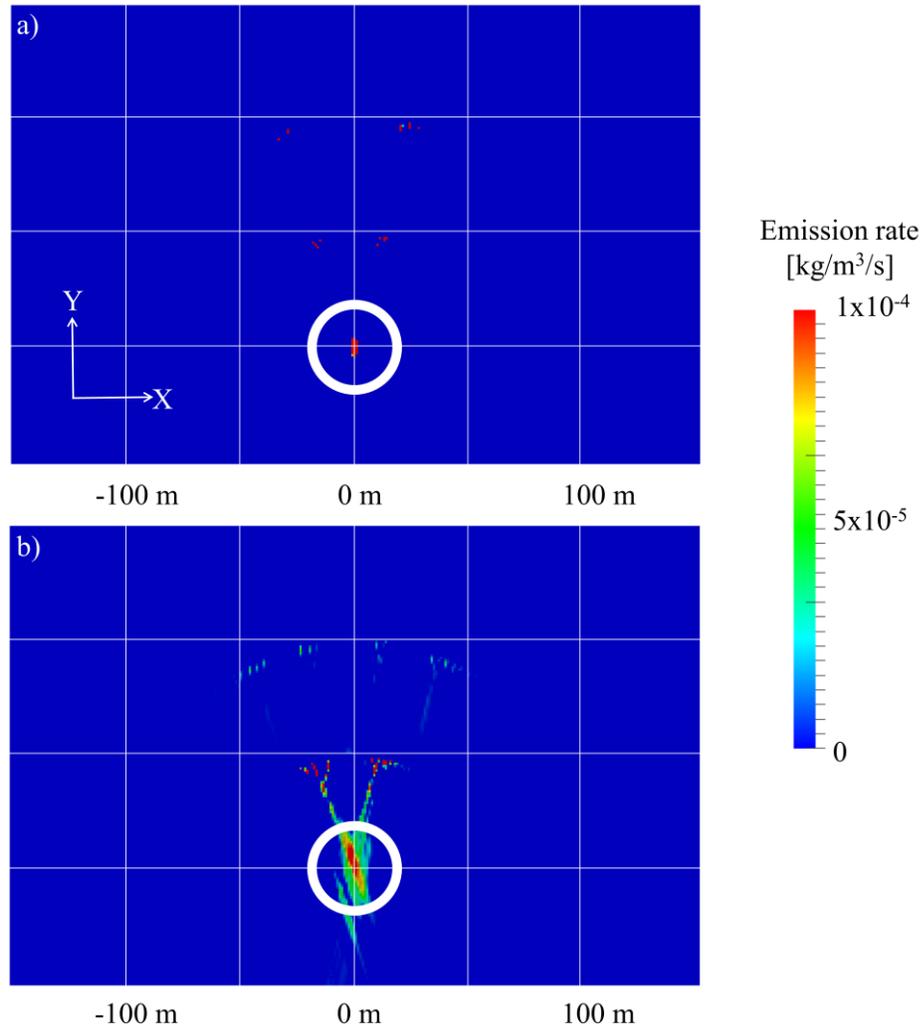


Figure 6.3: Isolated Sc_t effects on predicted emission rate locations at 1.5 m height (main sensor level) for extremes of Sc_t mismatch. a) Synthetic data set generated with $Sc_{t,actual} = 2.0$ while using an assumed $Sc_{t,assumed} = 0.33$ during the source characterization optimization, and b) a synthetic data set generated with a $Sc_{t,actual} = 0.33$ while using $Sc_{t,assumed} = 2.0$ during the source optimization. The actual location of the release is marked by white circle.

The influence of Sc_t mismatch on the accuracy of the predicted source magnitude is plotted in Figure 6.4. The total emission rate within the contiguous area of the main emission region (region boundaries determined using a boundary threshold of 1% of the maximum emission rate per volume) was used as the predicted emission rate. This avoided influence

of spurious predictions that sometimes occur at sensor locations in compensation for model errors. The predicted emission rate was normalized by the actual emission rate used to generate the data set and plotted as a function of $Sc_{t,assumed}$.

Even with $Sc_{t,actual} = Sc_{t,assumed}$ there is a slight over prediction in the source magnitude. This is anticipated given that even with the excellent sensor coverage the predicted source is diffused over a larger area than the actual source. In general, the predicted emission rates show similar trends. For a given $Sc_{t,actual}$, higher $Sc_{t,assumed}$ values (underestimating the diffusivity) reduce the predicted emission rates. Highly advective flows (higher $Sc_{t,actual}$) tend to be over-predicted and highly diffusive flow (lower $Sc_{t,actual}$) underpredicted for the ranges considered.

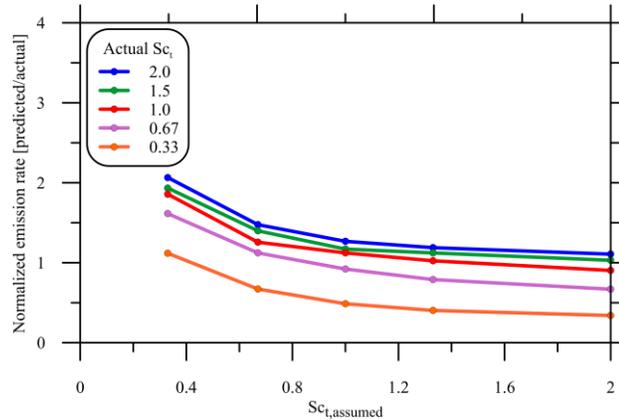


Figure 6.4: Isolated Sc_t effects on normalized predicted emission rate with unknown and unconstrained location.

The emission rates when the location is unconstrained range from -66% to $+109\%$ of the actual rate for combinations of Sc_t between 0.33 and 2.0. However, considering a single assumed turbulent Schmidt number of $Sc_{t,assumed} = 0.67$ (i.e. a midrange value frequently cited as appropriate for an open-field geometry) with actual values $Sc_{t,actual}$ still varying between 0.33 to 2.0, the predicted emission rates lie within a narrower range of -33% to $+48\%$ of the actual emission rate. A similar total range is found at each $Sc_{t,assumed}$ considered. The turbulent Schmidt number selection is clearly an important factor in determining the emission rate of an upstream source at unknown location.

6.6.1.3 Sc_t influence on known location but unknown emission rate

Two scenarios were considered in which the source location could be known or constrained prior to attempting to estimate its magnitude. In Figure 6.5a, the influence of incorrect $Sc_{t,assumed}$ on predicted source magnitude is considered in a case where the actual location is known (e.g. from site knowledge, camera observation, etc.). Figure 6.5b considers the case where the predicted location from the unconstrained adjoint case is assumed prior to seeking the source magnitude via simulated annealing. In both cases, the trends in the predicted emission rate error are similar to those of Figure 6.4. However, now the absolute range of the spread increases as $Sc_{t,assumed}$ is decreased and magnitudes of the prediction errors are generally larger.

For the exactly known location case shown in Figure 6.5a, the predicted emission rate has a broader range of uncertainty – between -80% and $+267\%$ of the actual release rate for the tested range of Sc_t mismatch. Selecting a mid-range turbulent Schmidt number ($Sc_{t,assumed}=0.67$) results in a predicted emission rate between -55% to $+90\%$ of the actual emission release rate for $Sc_{t,actual} = 0.33$ to 2.0 . Notably, these error ranges are greater than that for the unspecified location optimization.

In Figure 6.5b, the locations were assumed to be at the centroid of the prediction region from the unknown location procedure after which the emission rates were predicted via simulated annealing. For low $Sc_{t,assumed}$ values (high assumed diffusivity), the results are nearly identical to the known location results of Figure 6.5a. However, at higher $Sc_{t,assumed}$ values, the slight changes in position have measurable effects on the predicted emission rate. For cases where $Sc_{t,assumed} = Sc_{t,actual}$, emissions are consistently overestimated by $\sim 20\%$ when the predicted location is used. This is a larger error than the unknown location case in Figure 6.4. The emissions now range between -74% to $+261\%$ the actual rate (slightly less than using the actual locations in Figure 6.5a, but still more than the unconstrained case of Figure 6.4). With $Sc_{t,actual} = 0.33$ to 2.0 , selecting a mid-range turbulent Schmidt number ($Sc_{t,assumed}=0.67$) results in a predicted emission rate between -33% to $+109\%$ times the actual emission release rate (still worse than the -33% to $+48\%$ times error obtained in the unconstrained location case).

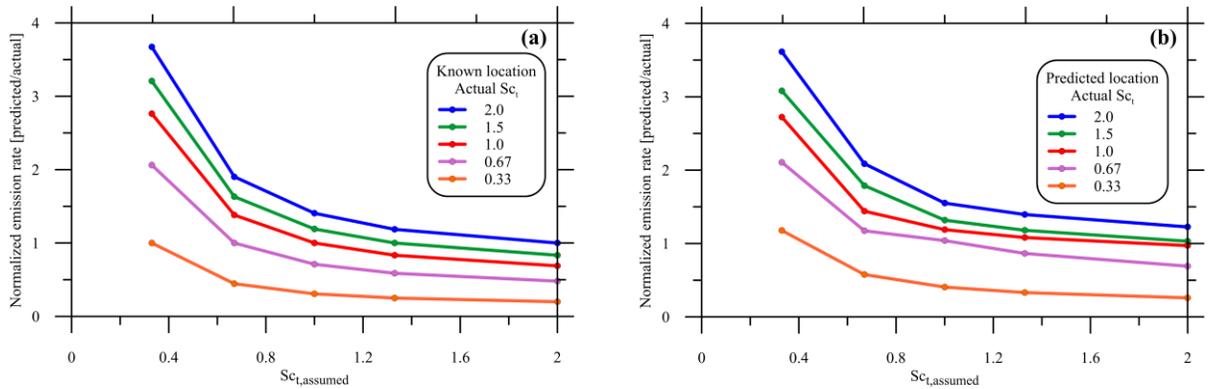


Figure 6.5: Isolated Sc_t effects on normalized predicted emission rate from simulated annealing procedure with a) known point source location, and b) predicted point source location based on unconstrained location prediction.

In the unknown location cases, model error tends to result in spurious sinks/sources in regions with minimal mutual information, or overlap between regions upstream of sensors, usually directly at sensors. However, for the specified location method, encompassing everything into a single variable (the single emission rate) appears to be less robust to model error. This may be exacerbated by the imperfections in point source implementation as volumes found in Eulerian framework dispersion. Overall, fixing the location (either to a known or prediction point) did not improve the error ranges over the unknown location (unconstrained) method in the open field test case.

6.6.1.4 Influence of Sc_t error in combination with all other error sources using Project Prairie Grass field measurement data

The effects of Sc_t variation were also investigated using field measured concentration data from the PPG measurements (Barad, 1958). This allowed Sc_t error effects to be considered in combination with other factors such as measurement uncertainty, wind model error, etc. that are relevant in a field application. Simulations evaluated the accuracy of the predicted source location and magnitude for a range of $Sc_{t,assumed}$ from 0.33 to 2.0. As illustrated in Figure 6.6a-c the central location of the emissions prediction region remained within a 5.3 m distance of the actual release location for all turbulent Schmidt numbers tested. Similar to the synthetic data test cases of Figure 6.3, the strongly advective plumes seen at high $Sc_{t,assumed}$ (Figure 6.6c) produce irregular, less smoothly varying source prediction regions. Compared with Figure 6.3 where Schmidt number effects were considered in isolation, the predicted source area using the field observations is spread over a larger area.

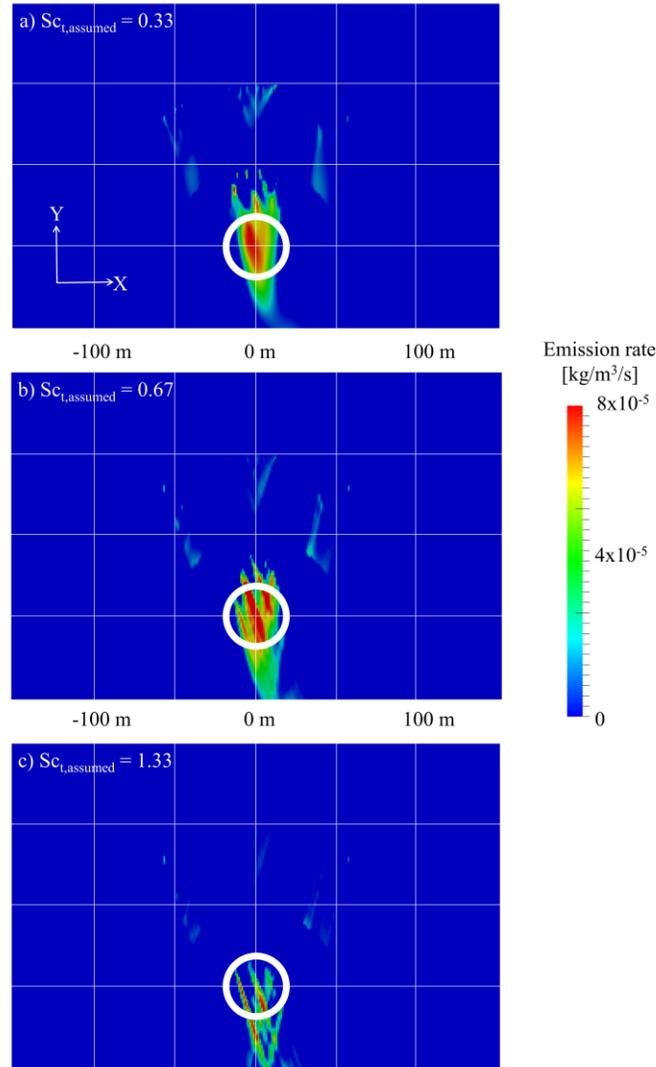


Figure 6.6: Combined effects of Sc_t mismatch and measurement and wind modelling error on the predicted source location in the Project Prairie Grass data set. The images show the $400\text{ m} \times 300\text{ m}$ domain at 0.5 m vertical height for with $Sc_{t,assumed}$ of a) 0.33, b) 0.67 and c) 1.33. Actual location of release is marked by a white circle.

The normalized emission rate was plotted versus $Sc_{t,assumed}$ in Figure 6.7. The unknown source location, known source location, and predicted location emission rate estimations are shown together. For the unknown source location case, the emission rate is calculated from the sum of predicted rates in connecting cells composing the main source region. The centroid of this region was used as the source position for the assumed source location case.

Although the actual source emission rate is known from the experimental data, $Sc_{t,actual}$ is unknown. In general, the unknown source location case predicts a stronger source, the

known location a weaker source, and the assumed location case falls in between. Overall, for an assumed turbulent Schmidt number range of 0.33-2.0, the predicted emission rates were within -35% to $+128\%$ (range of $\sim 170\%$), -75% to $+47\%$ (range of $\sim 120\%$), and -65% to $+99\%$ (range of $\sim 160\%$) of the actual emission rate for the unknown, known, and assumed locations respectively.

Although the error range is slightly reduced using the known or predicted location methods, there is a greater tendency toward underprediction. When assuming a mid-range turbulent Schmidt number appropriate for an open field of $Sc_{t,assumed}=0.67$, the predicted emission rates are $+9\%$, -41% , and -29% of the actual value for the unknown, known, and predicted locations. This suggests that the unknown source location method (scalar-adjoint approach) is still likely to give better results in similar situations matching the isolated Sc_t results presented in previous sections. However, the field data results of Figure 6.7 show that over an extreme range of $Sc_{t,assumed}$, each method can have a case where it provides the closest estimate of the actual source magnitude. Interpolating on Figure 6.7, the implied $Sc_{t,actual}$ to give a normalized emission rate prediction of unity is 0.76 for the unknown location, 0.45 for the exactly known location, and 0.6 for the predicted location cases.

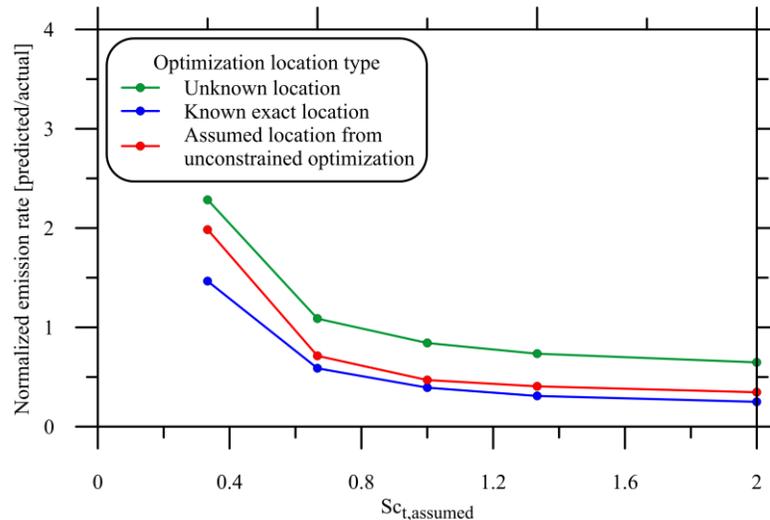


Figure 6.7: Normalized prediction of source emission rate in PPG experimental test data versus reciprocal of assumed turbulent Schmidt number. Curves show results assuming either an unknown location, known exact location, or fixed location based on predicted centroid from unknown location method.

6.6.2 Partial Gas Plant (Bluff Body) Test Case

6.6.2.1 Test Case Details

Turbulent Schmidt number sensitivity was also investigated in a more complicated 3D domain based on an Alberta gas plant relevant to the target application of detecting and quantifying fugitive emissions within industrial facilities. The test geometry (see Figure 6.8) features irregularly placed buildings ranging in height from 3 m to 10 m within a 400 m × 350 m × 100 m domain discretized into 1411061 cells. Four release locations (see Figure 6.8) were chosen, each at a vertical height of 1.0 m. Non-dimensional mass emission rates of 10/s and 1/s were selected, creating two large (L) and two small (S) sources (representing insignificant sources). Further detail can be found in Brereton et al. (2018a).

Brereton et al. (2018a) found that it was possible to use as sequence of steady RANS simulations (with different inlet boundary conditions) to reasonably approximate transient URANS wind fields during the optimization process. Brereton et al. (2018b) showed that this approach could also be used with a pre-computed flow database based on the steady flow wind data approximations. In this paper, due to practical computational restrictions in simulating multiple cases, these approaches were adopted. For the transient URANS simulations, inlets were based on varying 10 m vertical velocities of wind tower data representing approximately 30 min. of physical time. Inlet wind profiles were represented by the recommended neutral stability profiles of Richards and Hoxey (1993) for a roughness height of 0.04 m and an additional fetch of buildings were added upstream of the region of interest. Detail of these transient simulations can be found in Brereton et al. (2018a, 2018b). Using average velocities at 10 m height ahead of the main built region, this transient wind field was approximated with thirty-two steady-state RANS simulations with wind primarily from the South-East. Each interval consisted of 30 s of useable measurement time preceded by 96 s to 171 s of domain clearance time, which allowed time for the plume within the domain to clear after each wind direction change. This is discussed in more detail in Brereton et al. (2018b). Measurement intervals were fully independent and did not overlap. The convergence criterion was a consistent cost function decrease of less than 0.01% of the first iteration value. Known emission location rates were also found

with the simulated annealing optimization assuming both the actual locations and locations estimated from the initial unknown location procedure.

This series of steady RANS simulations was initially treated as a known wind field so the effects of turbulent Schmidt number selection could be isolated. The four virtual releases were again simulated with $Sc_{t,actual}$ ranging from 0.33–2.0 on each steady RANS simulated wind field. Concentrations were extracted at sparse virtual sensor locations (see Figure 6.8). This virtual sensor network was set to a rectangular grid of sensors at 50 m spacing and a vertical height of 1.5 m. These synthetic concentration data sets (each with a different $Sc_{t,actual}$) were treated as the concentration observations. Inverse releases were subsequently performed with $Sc_{t,assumed}=1.0$ from each sensor backwards in time to create a database of pre-computed transport flows. These results were used to calculate model concentrations during the optimization loop. Thus, in these tests the wind input used within the optimization procedure can, barring the turbulent Schmidt number selection, match the observation generation. The general wind directions were chosen so that sources L2 and S2 interfere to investigate the effects of poor source discernment.

In addition to these known wind field cases, additional test cases considered the combined effects of Sc_t error, wind error, and sensor response limitations. Virtual releases with $Sc_{t,actual}$ ranging from 0.33–2.0 were performed on the transient URANS wind fields to generate sets of raw concentration data at the virtual sensors (see Brereton et al., 2018a, 2018b). The source characterization was then performed using the same steady-state RANS simulations as the previous cases (Brereton et al., 2018b). Thirty-second average concentrations at sensor locations were extracted from the transient case (representing limited sensor response) at the end of each selected period represented by the steady-state RANS simulation. Thus, the wind fields used in the optimization only approximated the transient URANS fields used to generate the test sensor data. As noted above, during the source optimization procedure $Sc_{t,assumed}$ was fixed at a nominal value of 1.0. Thus, with $Sc_{t,actual}$ varied in the original transient URANS simulations, the predicted source locations and magnitudes were subject to the combined effects of Sc_t error, wind error, and sensor response.

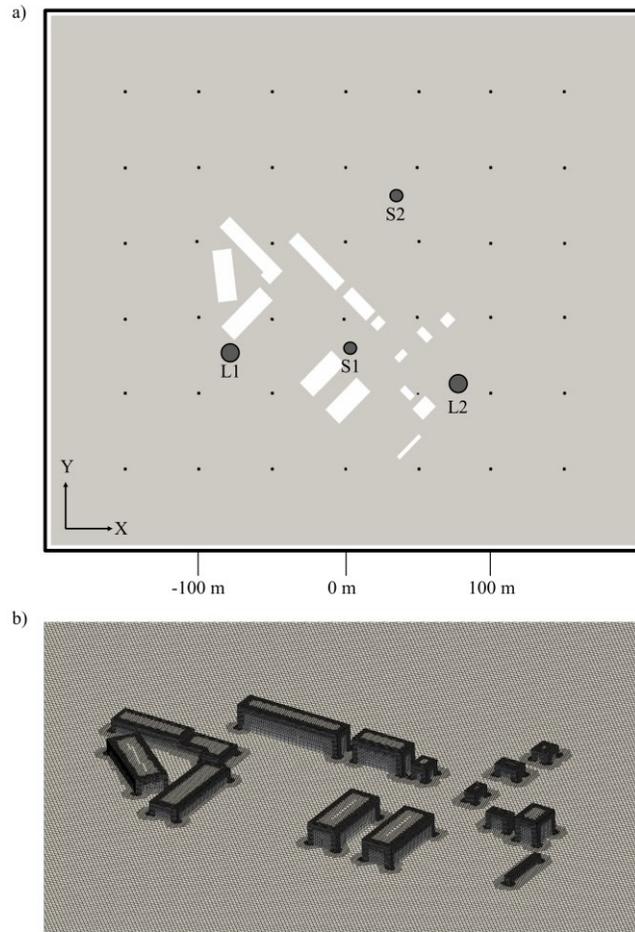


Figure 6.8: a) Top view of 3D building domain marking large (L) and small (S) sources (circles) and sensor locations (dots) and b) mesh detail (Brereton et al., 2018b)

6.6.2.2 *Unknown locations with unknown emission rates*

For $Sc_{t,assumed}=1.0$ and starting from an initial guess of no sources while allowing the scalar adjoint optimization scheme to predict source locations without constraint, Figure 6.9 shows predicted source regions for different $Sc_{t,actual}$. In Figure 6.9a-c, the wind fields used to generate the test sensor data exactly match those used in the source characterization, so that only the effects of Sc_t error are apparent. In Figure 6.9d-f, the sensor observations were generated with detailed transient wind fields that differed from the simplified data used in the source characterization. Thus, Sc_t mismatch errors are no longer isolated and instead are compounded with errors in the wind modelling and sensor response as would be expected in a practical application. In all cases, the centroids of the significant predicted source regions (i.e. regions representing $>5\%$ of total predicted emissions) are shown in

black and emission rates, location, and location prediction spread are compared for the closest significant emission source to the actual major release locations.

Results correctly predict the large sources regardless of the actual turbulent Schmidt number, though as with previous work (Brereton et al., 2018a), smaller sources cannot be discerned. However, tails and recirculation zones are also estimated as potential emission regions. These regions are expected given the limited wind data set (Brereton et al., 2018a).

When isolating for the effects of mismatch between the actual and assumed turbulent Schmidt numbers (Figure 6.9a-c), additional spurious sources become more important as the actual source plumes become more advective, though no additional regions become significant. However, this is not necessarily the case when including additional wind model errors during source characterization (Figure 6.9d-f).

Figure 6.10 plots location error (i.e. distance between the predicted source centroid and the actual source location) and spread of the predicted source (i.e. the emission-rate weighted average distance of each cell in the prediction region to the centroid – a convenient single-value measure of the spread of the source location prediction). Figure 6.10a shows the results when isolating for turbulent Schmidt number mismatch. The predicted locations are always within 4 m of the actual source. However, there is no consistent trend in the location error or spread with turbulent Schmidt number.

Figure 6.10b shows corresponding location error and predicted source spread results when sensor observations were generated with detailed transient, varying wind fields such that the errors of incorrect $Sc_{t,assumed}$ are combined with wind modelling and sensor response errors as would be expected in a practical application. The distance between the predicted source centroid and the actual source location again shows little dependence on $Sc_{t,assumed}$, but has increased to a maximum of 10 m from the actual source for L2. The spread of the predicted source regions is slightly larger than the matching wind cases but show little dependence on turbulent Schmidt number mismatch.

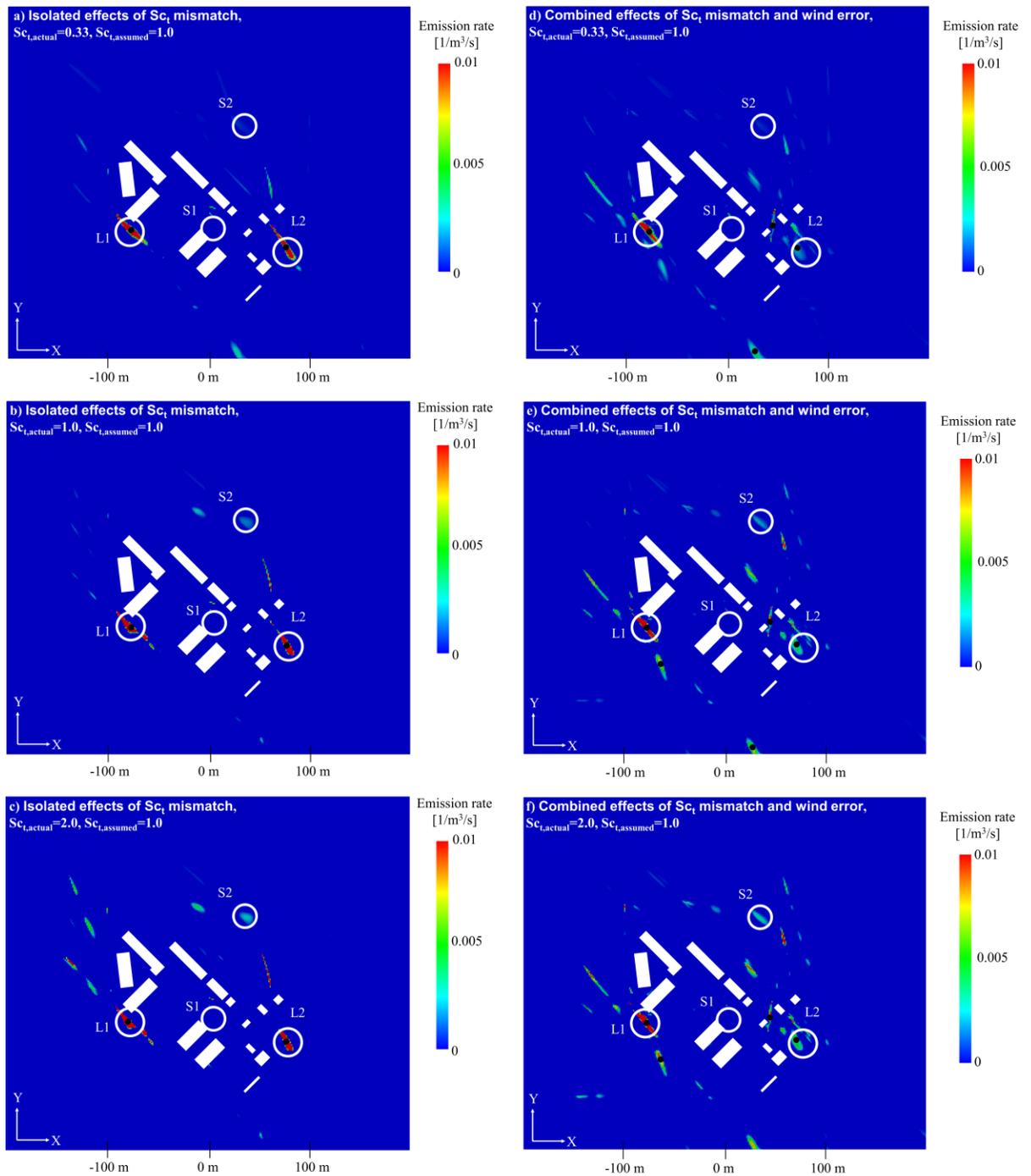


Figure 6.9: Predicted emission rate locations at 1.5 m height for $Sc_{t,assumed}$ of 1.0. (a-c) observations were generated on the same wind fields (isolating the effects Sc_t mismatch) with $Sc_{t,actual}$ of a) 0.33, b) 1.0, c) 2.0. (d-f) observations generated detailed transient wind fields (combining effects of Sc_t mismatch and wind and sensor error) with $Sc_{t,actual}$ of d) 0.33, e) 1.0, and f) 2.0. Actual location of released marked by circles. Horizontal centroid position of important (>5% total) emission regions are marked with black dots.

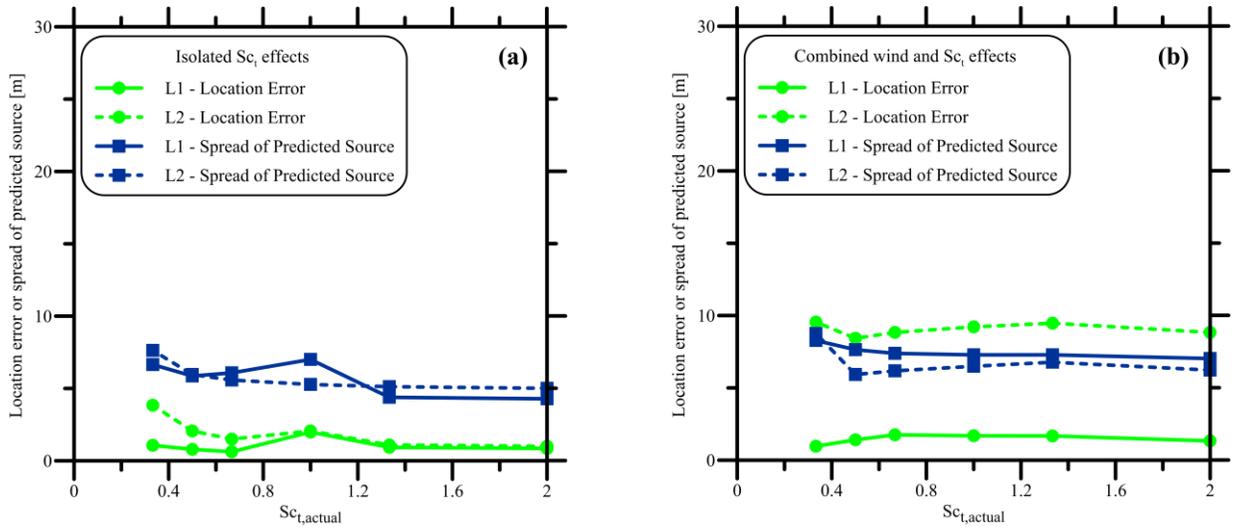


Figure 6.10: a) Isolated effects of Sc_t , and b) combined effects of Sc_t and wind error on source location accuracy (distance from predicted centroid to actual release location) and spread of predicted source (emission-weighted average spread from predicted centroid).

6.6.2.3 Emission rate predictions for unknown, known, and assumed source locations

Figure 6.11a compares predicted and actual emission rates for different $Sc_{t,actual}$ values using observations and optimization model concentrations generated on the same wind fields, which isolates the effects of Sc_t mismatch. The three plots show results using the unknown (unconstrained) location method, the known location with simulated annealing, and the predicted location with simulated annealing. Figure 6.11b compares the same prediction types, but with observations generated using the transient wind fields (i.e. incorporating both $Sc_{t,assumed}$ and wind modelling error). Predicted emission rates are plotted for each of the contiguous high emission regions corresponding to L1 and L2 (with boundaries defined by a threshold of 1% of the highest emission rate per volume).

Considering the effects of Sc_t in isolation (i.e. with matching wind fields), for source L1 the emission rate prediction increases from -33% to $+31\%$ of the actual source as $Sc_{t,actual}$ increases. Source L2, which compared to L1 is poorly covered by the sensor network for these wind directions, has a similar trend (rising from -41% to $+36\%$ of the actual emission rate). When wind model error is also considered (Figure 6.11b), both sources are under-

predicted for all $Sc_{t,actual}$, however the range of predicted values is notably reduced (–61% to –48% for source L1, and –65% to –44% for source L2).

Specifying source locations as known values, such as a case where an IR camera is used to verify location, does not reduce the general sensitivity to $Sc_{t,actual}$. With known wind fields (Figure 6.11a) knowing the location does seem to improve quantification accuracy, but this does not hold when wind error is also considered (Figure 6.11b). If the predicted source locations are specified prior to estimating source magnitude via simulated annealing, the sensitivity to $Sc_{t,actual}$ is potentially worse in both cases. In general, the presence of bluff-bodies tends to reduce the sensitivity to Sc_t error in comparison to the open-field cases, although this remains an important limitation in all cases. The bounds of the turbulent Schmidt number effects on emission rate predictions from the open field cases could be useful as a conservative error estimate in lieu of other information.

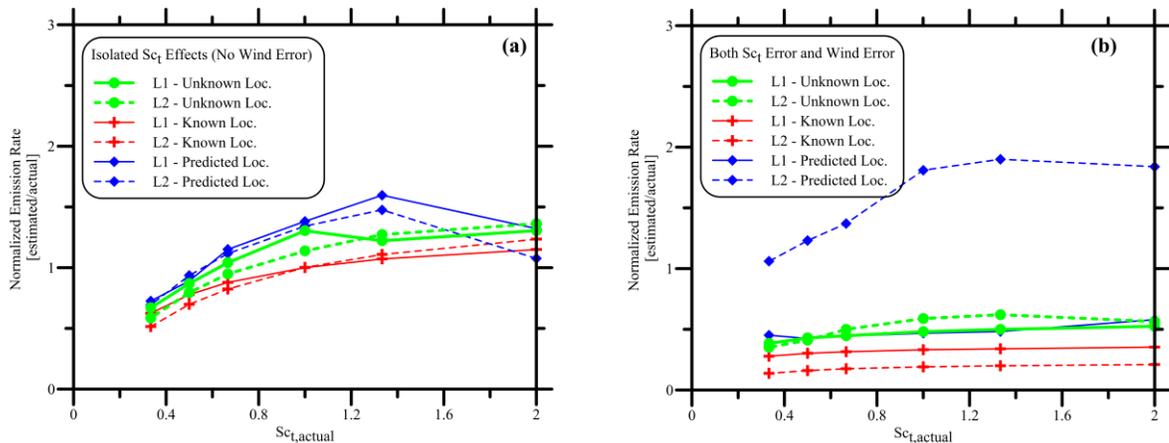


Figure 6.11: Normalized emission rate versus turbulent Schmidt number for the unknown (unconstrained), known, and predicted locations using $Sc_{t,assumed} = 1.0$ for a) wind fields matching between model and observations, and b) wind field approximation.

6.7 Conclusions

Clearly the selection of turbulent Schmidt number is an important criterion for quantifying emission sources using a modelling approach. However, source location accuracy is not particularly sensitive to turbulent Schmidt number. For the open field cases with good sensor coverage, increasing the assumed diffusivity (decreasing $Sc_{t,assumed}$) increased the emission rate prediction. For combinations of actual and assumed turbulent Schmidt

numbers spanning 0.33-2.0, the predicted emission rates could vary from -66% to +109% of the actual rate if locations were unspecified and were made worse (-80% to +267%) if the source location was prescribed. However, by assuming a single nominal turbulent Schmidt number of 0.67, the unconstrained location method could predict emission rates within -33% to +48% over the same range of actual turbulent Schmidt numbers.

For tests using the PPG experiment field data, if a mid-range $Sc_{t,assumed}$ value of 0.67 frequently cited as appropriate for an open-field geometry was selected, the unconstrained location approach predicted the actual source within ~10%. However, results could still vary from -35 to +128% of the actual emission rate if $Sc_{t,assumed}$ values between 2.0 and 0.33 were selected. Higher errors observed when prescribing the source location (even when specifying the actual source locations) were attributed to the restricted model being less able to compensate for model errors, though all methods could reproduce the experimental emission rate by assuming specific turbulent Schmidt numbers within the physically reasonable range.

The presence of bluff bodies and additional wind modelling error tended to reduce the turbulent Schmidt number dependence, consistent with Chavez et al. (2011), however spurious sources could arise especially with strongly advective plumes. Emission rates prediction did not improve by prescribing the predicted location as a point source prior to estimating its magnitude via simulated annealing.

Given the observed results, an ideal approach to placing bounds on turbulent Schmidt number dependence in source characterization applications would be to perform multiple characterizations using a range of different turbulent Schmidt number assumptions. However, in the limits of computational time and cost, a more practical alternative would be to use the open field case emission range as a conservative error bound for bluff-body cases. Simulations show that the main predicted source locations are not expected to change even if actual turbulent Schmidt numbers differ.

6.8 Acknowledgements

This project was supported by Natural Resources Canada (Project Manager Michael Layer), Petroleum Technology Alliance of Canada (PTAC), and the Natural Sciences and

Engineering Research Council of Canada (NSERC, Grant # 261966 and 446199). Computations were performed on resources and with support provided by the Centre for Advanced Computing (CAC) at Queen's University in Kingston, Ontario. The CAC is funded by: the Canada Foundation for Innovation, the Government of Ontario, and Queen's University.

Chapter 7 **Conclusions and Future Work**

7.1 Conclusions

The objective of this thesis was to investigate a scalar transport adjoint-based optimization approach that could efficiently locate and quantify unknown sources of fugitive emissions within potentially complex facilities. Working toward a target application of fugitive emissions mitigation in oil and gas facilities, the approach uses sparse sensor data and simplified wind fields to meet the main requirement of being able to direct maintenance to key areas in a timely manner. The technique was implemented using the partial differential equation solvers in the OpenFOAM toolkit and tested in multiple simulations.

Results showed that single-source open-field releases, based on field release data from Project Prairie Grass, could be located within 5 m of the actual source location and quantified within 13% using measured data and estimated wind fields. Tests of simulated multi-source releases on a complex 3D domain based on a partial Alberta gas plant were also successful in determining major source regions using only 14 min. of data. This was true even if detailed transient winds were approximated by stitching together a series of steady-state wind estimates to cover the physical period and using 30 s average sensor concentrations. This series of steady wind simulations would be much easier to implement for real-world applications and its success suggested that a pre-computed database of common flow conditions could be used to reduce computational time.

Further work successfully demonstrated the utility of a database of plume retro-tracers (PRTs) developed as an extension of the use of steady-state wind series approximations. These retro-tracers could be performed using common flow conditions in advance, front-loading computational cost to an initial system setup phase. Compared to the original method which uses concurrent transport computations (CTC) during the optimization process, using the PRT to estimate transport solves improved computational time by a factor of 200 to 600. The PRT method was shown to give equivalent results to the CTC

method on the open-field test case when using identical wind estimates. However, there was some loss of accuracy when using the PRT method on the bluff-body dominated test case because of the need to use simplified wind fields when developing the database of retro-tracers. In addition, with the PRT method, following each change in wind conditions, it is necessary to allow sufficient time for gas within the domain to clear. Nevertheless, results of test simulations with the PRT method showed that major sources could be successfully located even with minimal data and wind directional coverage (i.e. less than 30 min. of physical time with 110° of wind direction variation). Emission rates of major sources were predicted within -30% , which is sufficient to distinguish them from insignificant sources and prioritize maintenance. Results were also shown to improve with more data and increased wind coverage.

A final phase of the thesis considered the importance of the turbulent Schmidt number value used for source characterization, which has not previously been investigated in multi-source estimation problems. Uncertainty in the value of the turbulent Schmidt number (required to estimate diffusivity) was found to have little impact on source location estimates but could strongly influence predicted emission rates. For the open-field test case, if actual turbulent Schmidt number values ranged from 0.33 to 2.0, assuming a mid-range value of 0.67 resulted in a predicted emission rate uncertainty of -33% to $+48\%$. However, the presence of bluff bodies reduced the impact of the turbulent Schmidt number selection, suggesting that the open field case could be used as an error bound assuming computational costs preclude multiple tests with varying turbulent Schmidt number selections.

7.2 Major Contributions

The research presented in this thesis focused on an adjoint-optimization based approach for characterizing an unknown number of emission sources within potentially complex 3D domains. This is appropriate for locating and quantifying sources of fugitive emissions within oil and gas industry facilities but could be broadly applicable to emission sources in other applications such as hazardous material releases, terrorism incidents, or the location and/or quantification of other industrial pollutants.

The method code was developed and implemented with OpenFOAM, an open-source set of partial differential equation solver libraries rather than proprietary software. This use of open-source tools makes it much easier to adapt and extend the presented results into practical applications. The method and code were investigated with both field release data and developed synthetic releases within a three-dimensional domain representative of the target application.

In addition to demonstrating the utility of the approach in locating and quantifying unknown emission sources within a complex facility geometry, further work revealed the potential of using simplified wind input data to reduce computational overhead. Compared with complex, transient wind inputs, simplified 30-s average wind data were shown to be sufficient to provide useful location and quantification information for unknown sources.

A major extension of this result then demonstrated how pre-computed adjoint releases based on simplified average wind fields could be used to vastly reduce computational cost while still allowing major source regions to be located and emission rates scaled. Following an initial setup phase, this reduced computational time by 200 to 600 times, making near-continuous source determination feasible on a simple desktop computer.

Tests on both single and multi-source releases revealed that turbulent Schmidt number selection did not significantly affect predicted source locations, making it reasonable to use a single turbulent Schmidt number for pre-computation database generation. Although the turbulent Schmidt number was shown to affect quantification accuracy, bluff-bodies in the domain dampened this effect. A further contribution of this work is the suggestion that results for the open-field test case can act as an upper bound to Schmidt number uncertainties in the bluff-body dominated domains, which are expected in the target application.

Overall, this thesis has demonstrated the potential of the scalar transport adjoint-based optimization approach to locate and quantify unknown sources of fugitive emissions within complex facilities. In a field application, this technique has the potential to greatly increase fugitive emission detection speed, allowing emitting components to be dealt with quickly by giving both approximate magnitude and location to prioritize maintenance and repair.

Further, the PRT method, once setup, is sufficiently fast that it gives rise to the possibility of continuous, or near-continuous, emission source identification. Such a system would easily allow for external verification of emissions and reductions, directly enabling regulatory strategies based on carbon pricing and/or crediting.

7.3 Future Work

Several areas remain for future investigations. Wind directional coverage, as well as sensor position relative to source locations, have been shown to have large impacts on the ability to estimate source magnitudes and narrow predicted source regions. Additional test data sets, specifically for varying complexities and size of domain, sensor positioning, and wider range of wind direction values are aspects that need to be explored in the future to develop best practices for use. This includes developing test data sets for field releases over representative bluff bodies and data sets with varying vertical source positioning with the goal of developing minimum data collection limits to facilitate automatic source characterization.

In related work, the importance of sensor positioning also needs to be further explored to develop best practice placement guidelines and minimum levels of sensor coverage. The current work used a regularly placed grid of sensors for the bluff-body tests, but it is possible that placement could be optimized for facility geometries and common wind conditions. Determining the limits of the method with fence-line monitoring, and sensor density requirements for such a sensor setup, also may be useful in cases where monitoring cannot be allowed within facilities.

References

- Addepalli, B., Sikorski, C., Pardyjak, E.R. (2009). Source Characterization of Atmospheric Releases using Quasi-Random Sampling and Regularized Gradient Optimization. University of Utah Computer Science, p. 19 Technical Report, UUCS-09-001. Available at: <https://www.cs.utah.edu/docs/techreports/2009/pdf/UUCS-09-001.pdf>
- Albertson, J. D., Harvey, T., Foderaro, G., Zhu, P., Zhou, X., Ferrari, S., Thoma, E. D. (2016). A Mobile Sensing Approach for Regional Surveillance of Fugitive Methane Emissions in Oil and Gas Production. *Environmental Science & Technology*, 50(5), 2487–2497.
- Alinot, C. (2003). Analyse Aérodynamique Des Parcs Éoliens Immergés Dans Une Couche Limite Terrestre Ayant Différentes Conditions De Stabilité Thermique. (PhD. thesis). University of Quebec, Montreal, Canada.
- Allen, C. T., Young, G. S., Haupt, S. E. (2007). Improving pollutant source characterization by better estimating wind direction with a genetic algorithm. *Atmospheric Environment*, 41, 2283-2289.
- Allen, D.T., Pasci, A.P., Sullivan, D.W., Zavala-Araiza, D., Harison, M., Keen, K., Fraser, M.P. Hill, A.D., Sawyer, R.F., Seinfeld, J.H. (2015). Methane emissions from process equipment at natural gas production sites in the United States: Pneumatic Controllers. *Environmental Science & Technology*, 49(1), 633-640.
- Andrew, G. (2008). *Overview of Quasi-Newton optimization methods*. Retrieved Sept. 5, 2013, from <http://homes.cs.washington.edu/~galen/files/quasi-newton-notes.pdf>
- Annunzio, A.J., Young, G.S., Haupt, S.E. (2012). A multi-entity field approximation to determine the source location of multiple atmospheric contaminant releases. *Atmospheric Environment*, 62, 593-604.
- Ashbaugh, L., Malm, W., Sadeh, W. (1985). A residence time probability analysis of sulphur concentrations at Grand Canyon National Park. *Atmospheric Environment*, 19, 1263-1270.
- Atherton, E., Risk, D., Fougere, C., Lavoie, M., Marshall, A., Werring, J., Williams, J.P., Minions, C. (2017). Mobile measurement of methane emissions from natural gas developments in Northeastern British Columbia, Canada. *Atmospheric Chemistry & Physics*, 17, 12405-12420.
- Barad, M.L. (1958). Project Prairie Grass, a field program in diffusion, vols. 1 and 2. *Geophysical Research Papers U.S.*, 59.
- Bertsimas, D., Tsitsiklis, J. (1993). Simulated Annealing. *Statistical Science*, 8(1), 10–15. doi:10.1214/ss/1177011077
- Bieringer, B.E, Rodriguez, L.M., Vandenberghe, F., Hurst, J.G., Bieberbach, G., Sykes, I., Hannan, J.R., Zaragoza, J., Fry, R.N. (2015). Automated source term and wind

- parameter estimation for atmospheric transport and dispersion applications. *Atmospheric Environment*, 122, 206-219.
- Blocken, B., Stathopoulos, T., Saathoff, P., Wang, X. (2008). Numerical evaluation of pollutant dispersion in the built environment: Comparisons between models and experiments. *Journal of Wind Engineering & Industrial Aerodynamics*, 96, 1817-1831.
- Blomgren, P. (2013). *Numerical Optimization Lecture Notes 18: Quasi-Newton Methods –The BFGS method*. San Diego: San Diego State University. Retrieved Jan. 6, 2014, from http://terminus.sdsu.edu/SDSU/Math693a_f2013/Lectures/18/lecture.pdf
- Brandt, A. R., Heath, G.A., Cooley, D. (2016). Methane leaks from natural gas systems follow extreme distributions. *Environmental Science & Technology*, 50. 12512-12520.
- Brandt, A.R., Heath, G.A., Kort, E.A., O’Sullivan, F., Petron, G., Jordaan, S.M., Tans, P., Wilcox, J., Gopstein, A.M., Arent, D., Wofsy, S., Brown, N.J., Bradley, R., Stucky, G.D., Eardley, D., Harriss, R. (2014). Methane leaks from North American natural gas systems. *Science*, 343(6172), 733-735.
- Brantley, H.L, Thomas, E.D., Eisele, A.P. (2015). Assessment of volatile organic compound and hazardous air pollutant emissions from oil and natural gas well pads using mobile remote and on-site direct measurements. *Journal of the Air & Waste Management Association*, 65(9), 1072-1082.
- Brantley, H.L., Thoma, E.D., Squier, W.C., Guven, B.B., Lyon, D. (2014). Assessment of methane emission for oil and gas production pads using mobile measurements. *Environmental Science & Technology*, 48(24), 14508-14515.
- Brereton, C. A, Johnson, M.R. (2012a). Identifying Sources of Fugitive Emissions in Industrial Facilities using Trajectory Statistical Methods. *Atmospheric Environment*, 51, 46–55.
- Brereton, C. A., Johnson, M.R. (2012b). Steady-state Wind Field Reconstruction for Pollutant Emission Location. In Air & Waste Management Association (AWMA) 105th Annual Conference. San Antonio, TX.
- Brereton, C.A., Joynes, I., Campbell, L.J., Johnson, M.R. (2018a). Fugitive emission source characterization using a gradient-based optimization scheme and scalar transport adjoint. *Atmospheric Environment*, 181, 106-116.
- Brereton, C.A., Campbell, L.J., Johnson, M.R. (2018b). Computationally Efficient Quantification of Unknown Fugitive Emissions Sources, *manuscript prepared for submissions to Atmospheric Environment included in this thesis as Chapter 5*.
- Brereton, C.A., Campbell, L.J., Johnson, M.R. (2018c). Influence of Turbulent Schmidt Number on Fugitive Emissions Source Quantification *manuscript prepared for submissions to Atmospheric Environment and included in this thesis as Chapter 6*.

- Byrd, R. H.; Nocedal, J.; Schnabel, R. B. (1994). Representations of Quasi-Newton Matrices and their use in Limited Memory Methods. *Mathematical Programming*, 63 (4): 129–156.
- Cantelli, A., D’Orta, F., Cattini, A., Sebastianelli, F., Cedola, L. (2015). Application of genetic algorithm for the simultaneous identification of atmospheric pollution sources. *Atmospheric Environment*, 115, 36-46.
- Carras, J. N., Franklin, P. M., Hu, Y., Singh, A. K., Tailakov, O., Picard, D., et al (2006). Chapter 4: Fugitive emissions. *2006 IPCC guidelines for national greenhouse gas inventories*. Hayama, Japan: IPCC.
- Cartwright, H. M., Harris, S. P. (1993). Analysis of the distribution of airborne pollution using genetic algorithms. *Atmospheric Environment*, 27A(12), 1783-1791.
- Chambers, A.K. (2006). Refinery demonstration of optical technologies for measurement of fugitive emission and for leak detection. (ARC Contract Report No CEM 9643). Edmonton, AB: Carbon and Energy Management, Alberta Research Council Inc.
- Chavez, M., Hajra, B., Stathopoulos, T., Bahloul, A. (2011). Near-field pollutant dispersion in the built environment by CFD and wind tunnel simulations. *Journal of Wind Engineering and Industrial Aerodynamics*, 99, 330-339.
- Cheng, C. (2013). Gradient Descent Methods for Large-Scale Linear Inverse Problems. Emory University.
- Cheng, M., Lin, C. (2001). Receptor modeling for smoke of 1998 biomass burning in Central America. *Journal of Geophysical Research*, 106, 22871-22886.
- Chow, F.K., Kosović, B., Chan, S. (2008). Source Inversion for Contaminant Plume Dispersion in Urban Environments Using Building-Resolving Simulations. *Journal of Applied Meteorology & Climatology*, 47, 1553–1572.
- Clearstone Engineering Ltd. (2014a). UOG Emissions Inventory Vol. 3: UOG Emissions Inventory Methodology Manual. Gatineau, QC, Canada. Environment Canada.
- Clearstone Engineering Ltd. (2014b). Update of Fugitive Equipment Leak Emission Factors. Calgary, AB, Canada. Canadian Association of Petroleum Producers (CAPP). Retrieved Feb 21, 2018 from <https://www.capp.ca/~media/capp/customer-portal/publications/238773.pdf>
- Crenna, B.P., Flesch, T.K., Wilson, J.D. (2008). Influence of Source-Sensor Geometry on Multi-Source Emission Rate Estimates. *Atmospheric Environment*, 42, 7373-7383.
- Davidson, W.C. (1959). Variable metric method for minimisation. *A.E.C. Res. Develop. Rept.* (No. ANL-5596).
- de Haan, P. (1999). On the use of density kernels for concentration estimations within particle and puff dispersion models. *Atmospheric Environment*, 33, 2007-2021.

- De Meutter, P., Camps, J., Delcloo, A., Termonia, P. (2017). Assessment of the announced North Korean nuclear test using long-range atmospheric transport and dispersion modelling. *Scientific Reports*, 7(8762), 1-9.
- Environment and Climate Change Canada (ECCC). (2017). Canadian Environmental Sustainability Indicators: Greenhouse Gas Emissions. Retrieved on Feb. 21, 2018 from: http://www.canada.ca/content/dam/eccc/migration/main/indicateurs-indicators/f60db708-6243-4a71-896b-6c7fb5cc7d01/ghgemissions_en.pdf
- Elbern, H., Schmidt, H., Talagrand, O., Ebel, A. (2000). 4D-variational data assimilation with an adjoint air quality model for emission analysis. *Environmental Modelling & Software*, 15, 539–548.
- Flesch, T.K. (2002). Turbulent Schmidt number from a tracer experiment. *Agricultural and Forest Meteorology*, 111, 299-307.
- Flesch, T.K., Harper, L.A., Powell, J.M., Wilson, J.D. (2009). Inverse dispersion calculation of ammonia emissions from Wisconsin dairy farms. *Transactions of the American Society of Agricultural and Biological Engineers*, 52, 253-265.
- Foken T. (2006). 50 years of the Monin–Obukhov similarity theory. *Boundary-Layer Meteorology*. 119. 431-447.
- Gålfalk, M., Olofsson, G., Crill, P., Bastviken, D. (2016). Making methane visible. *Nature Climate Change*, 6, 426-430.
- Gasser, T., Peters, G.P., Fuglestedt, J.S., Collins, W.J., Shindell, D.T., Ciais, P. (2017). Accounting for the climate-carbon feedback in emission metrics. *Earth System Dynamics*, 8(2), 235-253.
- Goyal, A., Small, M. J., Von Stackelberg, K., Burmistrov, D., & Jones, N. (2005). Estimation of fugitive lead emission rates from secondary lead facilities using hierarchical Bayesian models. *Environmental Science & Technology*, 39(13), 4929–4937.
- Gualtieri, C., Angeloudis, A., Bombardelli, F., Jha, S., Stoesser, T. (2017). On the values for the Turbulent Schmidt Number in environmental flows. *Fluids*, 2, 17.
- Guo, S., Yang, R., Zhang, H., Weng, W., Fan, W. (2009). Source identification for unsteady atmospheric dispersion of hazardous materials using Markov Chain Monte Carlo method. *International Journal of Heat and Mass Transfer*, 52, 3955–3962.
- Hakami, A., Henze, D. K., Seinfeld, J. H., Singh, K., Sandu, A., Kim, S., et al. (2007). The adjoint of CMAQ. *Environmental Science & Technology*, 41, 7807–7817.
- Hartley, D., Prinn, R. (1993). Feasibility of determining surface emissions of trace gases using an inverse method in a three-dimensional chemical transport model. *Journal of Geophysical Research*, 98(D3), 5183-5197.
- Haas-Laursen, D. E., Hartley, D. E., Prinn, R. G. (1996). Optimizing an inverse method to deduce time-varying emissions of trace gases. *Journal of Geophysical Research*, 101(D17), 22823-22831.

- Haupt, S. E., Young, G. S., Allen, C. T. (2007). A genetic algorithm method to assimilate sensor data for a toxic contaminant release. *Journal of Computers*, 2(6), 85-93.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. Ann Arbor, MI: The MIT Press.
- Hooke, R., Jeeves, T.A. (1961). "Direct Search" Solution of Numerical and Statistical Problems. *Journal of the ACM.*, 8, 212–229.
- Hourdin, F., Talagrand, O. (2006). Eulerian backtracking of atmospheric tracers. I: Adjoint derivation and parameterization of subgrid-scale transport. *Quarterly Journal of the Royal Meteorological Society*, 132, 567-586.
- Houweling, S., Kamsink, T., Dentener, F., Lelieveld, J., Heimann, M. (1999). Inverse modeling of methane sources and sinks using the adjoint of a global transport model. *Journal of Geophysical Research*, 104, 26137–26160.
- Hsu, Y., Holsen, T. M., Hopke, P. K. (2003). Comparison of hybrid receptor models to locate PCB sources in Chicago. *Atmospheric Environment*, 37, 545-562.
- Huang, C., Hsing, T., Cressie, N., Ganguly, A.R., Protopopescu, V.A., Rao, N.S. (2010). Bayesian source detection and parameter estimation of a plume model based on sensor network measurements. *Applied Stochastic Models in Business and Industry*, 26(4), 331-348.
- Huang, Z., Wang, Y., Yu, Q., Ma, W., Zhang, Y., Chen, L. (2015). Source area identification with observation from limited monitor sites for air pollution episodes in industrial parks. *Atmospheric Environment*, 122, 1-9.
- Hunt, B.R., Kostelich, E.J., Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230(1-2), 112-126.
- Hutchinson, M., Oh, H., Chen, W-H. (2017). A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors. *Information Fusion*, 36, 130-148.
- Issartel, J., Baverel, J. (2003). Inverse transport for the verification of the Comprehensive Nuclear Test Ban Treaty. *Atmospheric Chemistry and Physics*, 3, 475-486.
- Issartel, J.P., Sharan, M., Modani, M. (2007). An inversion technique to retrieve the source of a tracer with an application to synthetic satellite measurements. *Proceedings of Royal Society A.*, 463(2087), 2863-2886.
- Jackson, R.B., Down, A., Phillips, N.G., Ackley, R.C., Cook, C.W., Plata, D.L., Zhao, K. (2014). Natural Gas Pipeline Leaks Across Washington, DC. *Environmental Science & Technology*, 48, 2051–2058.
- Johnson, M. R., Tyner, D. R., Conley, S., Schwietzke, S. Zavala-Araiza, D. (2017). Comparisons of Airborne Measurements and Inventory Estimates of Methane Emissions in the Alberta Upstream Oil and Gas Sector. *Environmental Science & Technology*, 51(21), 13008–13017.

- Joynes, I. M. (2013). Proof-of-Concept Inverse Micro-Scale Dispersion Modelling for Fugitive Emissions Quantification in Industrial Facilities. (M.A.Sc. Thesis), Carleton University.
- Jung, Y., Park, W., Park, O. (2003). Pollution dispersion analysis using the puff model with numerical flow field data. *Mechanics Research Communications*, 30, 277-286.
- Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME-Journal of Basic Engineering*, 82, 35–45.
- Karion, A., Sweeney, C., Pétron, G., Frost, G., Hardesty, R.M., Kofler, J., et al. (2013). Methane emissions estimate from airborne measurements over a western United States natural gas field. *Geophysical Research Letters*, 40, 4393–4397.
- Katzfuss, M., Stroud, J.R., Wilke, C.K. (2015). Understanding the Ensemble Kalman Filter. *The American Statistician*, 70(4), 351-357.
- Keats, A., Yee, E., Lien, F.-S. (2007). Bayesian inference for source determination with applications to a complex urban environment. *Atmospheric Environment*, 41, 465–479.
- Khlaifi, A., Ionescu, A., Candau, Y. (2009). Pollution source identification using a coupled diffusion model with a genetic algorithm. *Mathematics and Computers in Simulation*, 79(12), 3500–3510.
- Konda, U., Cheng, Y., Singh, T., Scott, P.D. (2010). Source identification of puff-based dispersion models using convex optimization. In IEEE: 13th Conference on Information Fusion (FUSION), Edinburgh (2010), 1-6.
- Kopka, P., Wawrzynczak, A., Boryxiewicz, M. (2016). Application of the Approximate Bayesian Computation methods in the stochastic estimation of atmospheric contamination parameters for mobile sources. *Atmospheric Environment*, 145, 201-212.
- Kuzmin, D. (2010). A Guide to Numerical Methods for Transport Equations. University of Erlangen–Nuremberg, Germany.
<http://www.mathematik.uni-dortmund.de/~kuzmin/Transport.pdf>
- Lagzi, I., Meszaros, R., Gelybo, G., Leelossy, A. (2013). Atmospheric Chemistry. Eötvös Loránd University, Budapest, Hungary. Retrieved Feb 2 2018 from: <http://elte.prompt.hu/sites/default/files/tananyagok/AtmosphericChemistry/index.html>
- Lane, R., Briers, M., Copsey, K. (2009). Approximate Bayesian computation for source term estimation. In IMA Conference on Mathematics in Defence 2009. Farnborough, Hampshire, UK.
- Launder, B.E. (1978). Heat and Mass Transport. In: Bradshaw, P. (Ed.), Topics in Applied Physics, Turbulence, vol. 12. Springer, Berlin.
- Lavoie, T.N., Shepson, P.B., Cambaliza, M.O.L., Stirm, B.H., Karion, A., Sweeney, C., et al. (2015). Aircraft-Based Measurements of Point Source Methane Emissions in the Barnett Shale Basin. *Environmental Science & Technology*, 49, 7904–7913.

- Lee, P.M. (1989). *Bayesian statistics: an introduction*. New York, Oxford Press.
- Liu, J., Bowman, K.W., Lee, M. (2016). Comparison between the Local Ensemble Transform Kalman Filter (LETKF) and 4D-Var in atmospheric CO₂ flux inversion with the Goddard Earth Observing System-Chem model and the observation impact diagnostics from the LETKF. *Journal of Geophysical Research: Atmospheres*, 141(21), 13066-13087.
- Liu, D.C., Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45, 503–528.
- Lucas, D.D., Gowardhan, A., Cameron-Smith, P., Baskett, R.L. (2016), Impact of meteorological inflow uncertainty on tracer transport and source estimation in urban atmospheres. *Atmospheric Environment*, 143, 120-132.
- Lupu, A., Maenhaut, W. (2002). Application and comparison of two statistical trajectory techniques for identification of source regions of atmospheric aerosol species. *Atmospheric Environment*, 36, 5607-5618.
- Lyon, D. R. (2016a). “Chapter 3: Methane Emissions from the Natural Gas Supply Chain.” In *Environmental and Health Issues in Unconventional Oil and Gas Development*. Ed. Debra A. Kaden. Ed. Tracie L. Rose. Amsterdam, Netherlands: Elsevier, 33-48.
- Lyon, D.R. (2016b). Quantifying, assessing, and mitigating methane emissions from super-emitters in the oil and gas supply chain. (PhD thesis). University of Arkansas, Fayetteville, Arkansas, USA.
- Ma, D., Deng, J., Zhang, Z. (2013). Comparison and improvements of optimization methods for gas emission source identification. *Atmospheric Environment*, 81, 188-198.
- Mitchell, A.L, Tkacik, D.S., Roscioli, J.R., Herndon, S.C., Yacovitch, T.I., Martinez, D.M. et al. (2015). Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results. *Environmental Science & Technology*, 49, 3219–3227.
- Monache, L.D. Lundquist, J. K., Kosovic, B., Johannesson, G. Dyer, K. M., Aines, R. D., Belles, R. D., Hanley, W. G., Larsen, S. C., Loosmore, G. A., Nitao, J. J. Sugyama, G. A. and Vogt, P.J. (2008). Bayesian inference and Markov Chain Monte Carlo sampling to reconstruct a contaminant source on a continental scale. *Journal of Applied Meteorology and Climatology*, 47, 2600–2613.
- Mons, V., Margheri, L., Chassaing, J.C., Sagaut, P. (2017). Data assimilation-based reconstruction of urban pollutant release characteristics. *Journal of Wind Engineering & Industrial Aerodynamics*, 169. 232-250.
- Mulholland, M. and Seinfeld, J.H. (1995). Inverse air pollution modelling of urban-scale carbon monoxide emissions. *Atmospheric Environment*, 29, 497–516.
- Munnings, C., Krupnick, A. (2017). Comparing policies to reduce methane emissions in the natural gas sector. Washington, DC: Resources for the Future. Retrieved Feb. 21, 2018 from <http://www.rff.org/files/document/file/RFF-Rpt-Methane.pdf>

- Nichols, N.K. (2003). Data Assimilation: Aims and Basic Concepts. In *Data Assimilation for the Earth System*. Swinbank, G. Shutyaev, V., Lahoz, W.A. Ed., Kluwer Academic, 2003.
- Nocedal, J. (1980). Updating Quasi-Newton Matrices with Limited Storage, *Mathematics of Computation*, 35, 773-782.
- Nocedal, J., Wright, S.J. (2006). Numerical Optimization. Second Edition. Springer, New York, USA.
- Nolle, L. (2006). On a Hill-Climbing Algorithm with Adaptive Step Size: Towards a Control Parameter-Less Black-Box Optimization Algorithm. *Computational Intelligence Theory & Applications*, 38, 587-595.
- Norige, A., Thornton, J., Shiefelbein, C., Rudzinski, C. (2009). High density distributed sensing for chemical and biological defense. *Lincoln Laboratory Journal*, 18(1), 25-40.
- NREL. (2011). National Wind Technology Center M2 Tower. Retrieved Feb. 10, 2011 from http://www.nrel.gov/midc/nwtc_m2/
- Peneko, V., Baklanov, A., Tsvetova, E. (2002). Methods of sensitivity theory and inverse modeling for estimation of source parameters. *Future Generation Computer Systems*, 18(5), 661-671.
- Peters, L.K., Berkowitz, C.W., Saylor, R.D., Tsang, T.T., Carmichaels, G.R., Easter, R.C. (1995). The current state and future direction of Eulerian models in simulating the tropospheric chemistry and transport of trace species: a review. *Atmospheric Environment*, 29(2), 189-222.
- Peterson, R.L., Guerra, S.A., Bova, A.S. (2017). Critical review of the building downwash algorithms in AERMOD. *Journal of the Air and Waste Management Association*, 67(8), 826-835.
- Phillips, N.G., Ackley, R., Crosson, E.R., Down, A., Hutyra, L.R., Brondfield, M., Karr, J.D., Zhao, K., Jackson, R.B. (2013). Mapping urban pipeline leaks: Methane leaks across Boston. *Environmental Pollution*, 173, 1-4.
- Picard, D., Stribny, M., Harrison, M.R. (1998). Handbook for estimating methane emissions from Canadian natural gas systems. Clearstone Engineering Ltd., Enerco Engineering Ltd. And Radian International, Don Mills, ON, Canada.
- Pudykiewicz, J. (1998). Applications of Adjoint Tracer Transport Equations for Evaluating Source Parameters. *Atmospheric Environment*, 32, 3039-3050.
- Pulkkinen, S. (2008). A Review of Methods for Unconstrained Optimization: Theory, Implementation and Testing. University of Helsinki.
- Rajaona, H., Septier, F., Armand, P., Delignon, Y., Olry, C., Albergel, A., Moussafir, J. (2015). An adaptive Bayesian inference algorithm to estimate the parameters of a hazardous atmospheric release. *Atmospheric Environment*, 122, 748-762.
- Rao, K. S. (2007). Source estimation methods for atmospheric dispersion. *Atmospheric Environment*, 41, 6964-6973.

- Ravikumar, A.P., Brandt, A.R. (2017). Designing better methane mitigation policies: the challenge of distributed small sources in the natural gas sector. *Environmental Research Letters*, 12, 044023.
- Ravikumar, A.P., Wang, J., Brandt, A.R. (2017). Are optical gas imaging technologies effective for methane leak detection? *Environmental Science & Technology*, 51, 718-724.
- Reible, D. D. (1999). Fundamentals of environmental engineering - 1999 publication. Boca Raton, FL: Lewis Publishers.
- Rella, C.W., Tsai, T.R., Botkin, C.G., Crosson, E.R., Stelle, D. (2015). Measuring Emissions from Oil and Natural Gas Well Pads Using the Mobile Flux Plane Technique. *Environmental Science & Technology*, 49, 4742–4748.
- Richards, P.J., Hoxey, R.P. (1993). Appropriate boundary conditions for computational wind engineering models using the k-epsilon turbulence model. *Journal of Wind Engineering & Industrial Aerodynamics*, 46&47, 145-153.
- Riddle, A., Carruthers, D., Sharpe, A., McHugh, C., Stocker, J. (2004). Comparisons between FLUENT and ADMS for atmospheric dispersion modelling. *Atmospheric Environment*, 38, 1029-1038.
- Ristic, B., Gunatilaka, A., Gailis, R., Skvortsov, A. (2015). Bayesian likelihood-free localization of a biochemical source using multiple dispersion models. *Signal Processing*, 108, 13-24.
- Roberts, P.J.W., Webster, D.R. (2002). Turbulent diffusion, in *Environmental Fluid Mechanics: Theories and Application* [H., Shen, A., Cheng, K.-H. Wang, M.H. Teng, and C. Liu (eds.)]. American Society of Civil Engineers, Reston, VA, USA.
- Roscioli, J. R., Herndon, S.C., Yacovitch, T.I., Knighton, W.B., Zavala-Araiza, D., Johnson, M.R., and Tyner, D.R. (2018). Characterization of Methane Emissions from Five Cold Heavy Oil Production with Sands (CHOPS) Facilities. *Journal of the Air and Waste Management Association*, in press (doi: 10.1080/10962247.2018.1436096).
- Roscioli, J. R., Yacovitch, T. I., Floerchinger, C., Mitchell, A. L., Tkacik, D. S., Subramanian, R., Martinez, D. M., et al. (2014). Measurements of methane emissions from natural gas gathering facilities and processing plants: measurement methods. *Atmospheric Measurement Techniques*, 7, 12357–12406.
- Rosenbrock, H.H. (1960). An Automatic Method for Finding the Greatest or Least Value of a Function. *The Computer Journal*, 3, 175–184.
- Rudd, A.C., Robins, A.G., Lepley, J.J., Belcher, S.E. (2012). An inverse method for determining source characteristics for emergency response applications. *Boundary Layer Meteorology*, 144 (1), 1–20.
- Ruder, S. (2016). An overview of gradient descent optimization algorithms. Retrieved Mar. 10, 2018 from <http://arxiv.org/abs/1609.04747>

- Senocak, I., Hengartner, N. W., Short, M. B., Daniel, W. B. (2008). Stochastic event reconstruction of atmospheric contaminant dispersion using Bayesian inference. *Atmospheric Environment*, 42(33), 7718–7727.
- Sharan, M. Singh, S.K., Issartel, J. (2012). Least square data assimilation for identification of the point source emissions. *Pure and Applied Geophysics*, 169(3), 483-497.
- Schulman, L.L., Strimaitis, D.G., Scire, J.S. (2000). Development and evaluation of the PRIME plume rise and building downwash model. *Journal of the Air and Waste Management Association*, 50(3), 378-390.
- Shoemaker, J. K., Schrag, D.P., Molina, M.J., Ramanathan, V. (2013). What Role for Short-Lived Climate Pollutants in Mitigation Policy? *Science*, 342 (December), 1323–1324.
- Singh S., Rani, R. (2014). A least-squares inversion technique for identification of a point release: Applications to Fusion Field Trials 2007. *Atmospheric Environment*, 92, 104-117.
- Singh, S.K., Rani, R. (2015). Assimilation of concentration measurements for retrieving multiple point releases in atmosphere: a least squares approach to inverse modelling. *Atmospheric Environment*, 119, 402-414.
- Singh, S.K., Sharan, M., Issartel, J-P. (2015). Inverse modeling methods for identifying unknown releases in emergency scenarios: an overview. *International Journal of Environment and Pollution*, 57, 68-91.
- Skajaa, A. (2010). Limited memory BFGS for nonsmooth optimization. New York University.
- Slater, J.W. (2008). NPARC Alliance CFD Verification and Validation Web Site: Examining Spatial (Grid) Convergence. Retrieved Mar 2, 2018 from: <https://www.grc.nasa.gov/WWW/wind/valid/tutorial/spatconv.html>
- Snieder, R. (1998). The role of nonlinearity in inverse problems. *Inverse Problems*, 14, 387–404.
- Sohn, M.D., Sextro, R.G., Gadgil, A.J., Daisey, J.M. (2003). Responding to sudden pollutant releases in office buildings: 1. Framework and analysis tools. *Indoor Air*, 13, 267–276.
- Spalding, D.B. (1971). Concentration fluctuations in a round turbulent free jet. *Chemical Engineering Science*, 26, 95–107.
- Spall, J. C. (2003). *Introduction to stochastic search and optimization: Estimation, simulation, and control*. Hoboken, NJ: John Wiley and Sons.
- Stohl, A., Eckhardt, S., Forster, C., James, P., Spichtinger, N., and Seibert, P. (2002). A replacement for simple back trajectory calculations in the interpretation of atmospheric trace substance measurements. *Atmospheric Environment*, 36, 4635–4648.

- Stohl, A., Forster, C., Frank, A., Seibert, P., Wotawa, G. (2005). Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmospheric Chemistry & Physics*, 5, 4739–4799.
- Tavoularis, S. (2018). MCG5152 Theory of Turbulence Lecture Notes. Department of Mechanical Engineering, University of Ottawa. Ottawa, Canada.
- Thompson, D.R., Leifer, I., Bovensmann, H., Eastwood, M., Fladland, M., Frankenberg, C., et al. (2015). Real-time remote detection and measurement for airborne imaging spectroscopy: a case study with methane. *Atmospheric Measurement Techniques*, 8, 4383–4397.
- Thomson, L. C., Hirst, B., Gibson, G., Gillespie, S., Jonathan, P., Skeldon, K. D., et al. (2007). An improved algorithm for locating a gas source using inverse methods. *Atmospheric Environment*, 41, 1128–1134.
- Thorpe, A.K., Frankenberg, C., Aubrey, A.D., Roberts, D.A., Nottrott, A.A., Rahn, T.A., et al. (2016). Mapping methane concentrations from a controlled release experiment using the next generation airborne visible/infrared imaging spectrometer (AVIRIS-NG). *Remote Sensing of Environment*, 179, 104–115.
- Tominaga Y., Stathopoulos, T. (2007). Turbulent Schmidt numbers for CFD analysis with various types of flowfield. *Atmospheric Environment*, 41, 8091–8099.
- Tratt, D.M, Buckland, K.N., Hall, J.L, Johnson, P.D., Keim, E.R., Leifer, I., et al. (2014). Airborne visualization and quantification of discrete methane sources in the environment. *Remote Sensing of Environment*, 154, 74–88.
- UNFCCC. (2015). Greenhouse Gas Inventory Data. Retrieved Feb. 13, 2018, from <http://di.unfccc.int>
- US Environmental Protection Agency (USEPA). 2006. Alternative Work Practice To Detect Leaks From Equipment, Retrieved Feb 21, 2018 from: <https://www.federalregister.gov/documents/2006/04/06/E6-5005/alternative-work-practice-to-detect-leaks-from-equipment>
- US Environmental Protection Agency (USEPA). (2016a). Inventory of US Greenhouse Gas Emission and Sinks: 1990-2014. EPA 430-R-16-002. Retrieved on Feb 21, 2018 from: <https://www.epa.gov/sites/production/files/2016-04/documents/us-ghg-inventory-2016-main-text.pdf>
- U. S. Environmental Protection Agency, (2016b). *Control Techniques Guidelines for the Oil and Natural Gas Industry* (EPA-453/B-16-001), U. S. Environmental Protection Agency, Triangle Park, NC.
- US Environmental Protection Agency (USEPA). (2017). Method 21 – Determination of Volatile Organic Compound Leaks. Retrieved on Feb 21, 2018 from: https://www.epa.gov/sites/production/files/2017-08/documents/method_21.pdf
- Wade, D., Senocak, I. (2013). Stochastic reconstruction of multiple source atmospheric contaminant dispersion events. *Atmospheric Environment*, 74, 45–51.

- Wallace, J. M., Hobbs, P.V. (2006). *Atmospheric science: An introductory survey*. Amsterdam: Elsevier Academic Press.
- Wang, Y., Huang, H., Huang, L., Ristic, B. (2017). Evaluation of Bayesian source estimation methods with Prairie Grass observations and Gaussian plume model: A comparison of likelihood functions and distance measures. *Atmospheric Environment*, 152, 519-530.
- Wilson, J.D. (2015). Computing the Flux Footprint. *Boundary-Layer Meteorology*, 156(1), 1-14.
- Wilson, J.D. (2013). Turbulent Schmidt numbers above a wheat crop. *Boundary-Layer Meteorology*, 148, 255-268.
- Wilson, J.D., Flesch, T.K., Creena, B.P. (2012). Estimating surface-air gas fluxes by inverse dispersion using a backward Lagrangian stochastic trajectory model. Pages 149-161 of: Lin, J., et al. (Eds), *Lagrangian Modeling of the Atmosphere*. Geophysical Monograph, no. 200. American Geophysical Union.
- Wilson, J.D., Yee, E., Ek, N., d'Amours, R. (2009). Lagrangian Simulation of Wind Transport in the Urban Environment. *Quarterly Journal of the Royal Meteorological Society*, 135, 1586-1602.
- Windtrax. (2018). Lagrangian stochastic dispersion modeling. Retrieved Mar. 2, 2018 from: <http://www.thunderbeachscientific.com/>
- Yee, E. (2008). Theory for reconstruction of an unknown number of contaminant sources using probabilistic inference. *Boundary Layer Meteorology*, 127(3), 359-394.
- Yu, H., Thé, J. (2016). Simulation of gaseous pollutant dispersion around an isolated building using the $k-\omega$ SST (shear stress transport) turbulence model. *Journal of the Air and Waste Management Association*, 67(5), 517-536.
- Xue, F., Li, X., Ooka, R., Kikumoto, H., Zhang, W. (2017). Turbulent Schmidt number for source term estimation using Bayesian inference. *Building and Environment*, 125, 414-422.
- Zavala-Araiza, D., Lyon, D., Alvarez, R.A., Palacios, V., Harriss, R., Lan, X., Talbot, R., Hamburg, S.P. (2015). Towards a functional definition of methane super-emitters: application to natural gas production sites. *Environmental Science & Technology*, 49(13), 8167-8174.
- Zavala-Araiza, D., Herndon, S.C., Roscioli, J.R., Yacovitch, T.I., Johnson, M.R., Tyner, D.R., Omara, M., and Knighton, B. (2018). Methane emissions from oil and gas production sites in Alberta, Canada. *Elementa: Science of the Anthropocene*, 6(1):27.
- Zhang, L. (2007). *Large-scale simulations using first and second order adjoints with applications in data assimilation*. (Master's thesis). Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Zhang, X.L., Su, G.F., Yuan, H.Y., Chen, J.G., Huang, Q.Y. (2014). Modified ensemble Kalman filter for nuclear accident atmospheric dispersion: Prediction improved and source estimated. *Journal of Hazardous Materials*, 280, 143-155.

- Zheng, X., Chen, Z. (2011). Inverse calculation approaches for source determination in hazardous chemical releases. *Journal of Loss Prevention in the Process Industries*, 24, 293–301.
- Zhu, C., Byrd, R., Nocedal, J., Morales, J.L. (2011). L-BFGS-B. Retrieved Feb 18, 2012 from <http://users.iems.northwestern.edu/~nocedal/lbfgsb.html>
- Zimmerle, D.J., Williams, L.L., Vaughn, T.L., Quinn, C., Subramanian, R., Duggan, G.P., Willson, B., Opsomer, J.D., Marchese, A.J., Martinez, D.M., Robinson, A.L. (2015). Methane emissions from the natural gas transmission and storage system in the United States. *Environmental Science & Technology*, 49(15), 9374-9383.

Appendix A: Plume and Gas Dispersion Modelling

Modelling of gas transport, or plume dispersion modelling, is important for many atmospheric chemistry applications. In addition to source characterization, modelling plumes can determine impacts of existing or proposed structures and facilities. Various paradigms and simplifications exist depending on available resources or application, but are typically based on mass conservation.

A.1 Mass Conservation and the Advection-Diffusion Equation

For a fixed volume Ω with control surface boundaries Γ and a concentration c (per unit volume) of a tracer species (Kuzmin, 2010),

$$\frac{\partial}{\partial t} \int_{\Omega} c \, d\mathbf{x} + \int_{\Gamma} \mathbf{f} \cdot \mathbf{n} \, d\Gamma = \int_{\Omega} S \, d\mathbf{x} \quad (\text{A.1})$$

must be satisfied where

t is time,

S is the term for sources or sinks of mass within the control volume,

\mathbf{f} is the flux vector which specifies the rate tracer species enters or leaves the volume through the boundaries, and

\mathbf{n} is the outward normal unit vector at the surface.

Applying divergence theorem, the evolution of c in time (t) is governed by (Kuzmin, 2010)

$$\frac{\partial c}{\partial t} + \nabla \cdot \mathbf{f} = S. \quad (\text{A.2})$$

One form of mass transport for a tracer species is molecular diffusion. Fick's law relates this diffusion driven mass transport to concentration gradient and, in three dimensions, can be generalized to (Roberts and Webster, 2002)

$$\mathbf{q} = -D_c \nabla c \quad (\text{A.3})$$

where

\mathbf{q} is the flux vector (mass transport rate per unit area per unit time), and

D_c is the molecular diffusion coefficient.

The second major transport mechanism is advective transport (also sometimes called convective transport) and denoted $\mathbf{U}c$ in three dimensions (Roberts and Webster, 2002) where \mathbf{U} is the velocity vector. Putting these transport mechanisms into the Eq. (A.2) gives

$$\frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{U}c + \mathbf{q}) = \frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{U}c - D_c \nabla c) = S. \quad (\text{A.4})$$

The source term S may contain terms for sources and sinks, including deposition, chemical decay, and radioactive decay. Gravitational settling may be added to advection component if necessary. (Lagzi et al., 2013)

Using Reynolds decomposition, both the wind and concentration terms can be split into mean (ensemble-average) and perturbed components,

$$\mathbf{U} = \mathbf{u} + \mathbf{u}', \quad (\text{A.5})$$

and

$$c = C + c'. \quad (\text{A.6})$$

Substituting these split terms into Equation (A.4) and averaging results in

$$\frac{\partial C}{\partial t} + \nabla \cdot (\mathbf{u}C) + \nabla \cdot (-D_c \nabla C) + \frac{\partial(\overline{u'c'})}{\partial x} + \frac{\partial(\overline{v'c'})}{\partial y} + \frac{\partial(\overline{w'c'})}{\partial z} = S \quad (\text{A.7})$$

where u' , v' , and w' are the perturbed velocity components. The three new terms on the left hand side are the turbulent flux (Roberts and Webster, 2002). The x -component of the molecular dispersion and turbulent flux can be written as (Roberts and Webster, 2002)

$$\frac{\partial \left(-D_c \frac{\partial C}{\partial x} + \overline{u'c'} \right)}{\partial x}. \quad (\text{A.8})$$

D_c is usually very small so $\overline{u'c'} \gg D_c \frac{\partial c}{\partial x}$ and the molecular transport is generally neglected unless considering small scale mixing (Roberts and Webster, 2002), retaining the turbulent transport only. Other components directions are similar. A common assumption to treat the turbulent transport is based on gradient theory and treats the directional turbulent flux as proportional to the concentration gradient (Lagzi et al., 2013). Using this assumption, Equation (A.7) can be expressed as

$$\frac{\partial C}{\partial t} + \nabla \cdot (\mathbf{u}C - D\nabla C) = S \quad (\text{A.9})$$

where D is the matrix turbulent diffusion coefficients and is a property of the flow. Equation (A.9), often referred to as the advection-diffusion equation or the convection-diffusion equation, may be solved in a number of ways, but the three main approaches are (Lagzi et al., 2013):

1. Gaussian, where the equation is solved analytically with simplifying assumptions,
2. Lagrangian, a stochastic approach where the concentration field is given by particle superposition, and
3. Eulerian, where a spatially discretized mesh is used.

Each of these approaches and common industry implementations are discussed below.

A.2 Gaussian Plume Dispersion

Assuming a homogenous steady-state flow and source, Equation (A.9) can be solved analytically. Assuming coordinate directions where the x-axis is aligned with the wind direction, the basic form of the standard Gaussian plume model is (Jung et al., 2003)

$$C = \frac{\dot{m}}{2\pi U \sigma_y \sigma_z} \exp\left[-\frac{y^2}{2\sigma_y^2}\right] \exp\left[-\frac{(z-H)^2}{2\sigma_z^2}\right] \quad (\text{A.10})$$

where C is the steady-state concentration, \dot{m} is the emission rate, U is the wind speed at the release height, y the horizontal distance from the plume center, z the distance from the ground and H the effective release height (which may include compensation for plume buoyancy). σ_y and σ_z are the standard deviations in spread for their respective coordinate

directions and describe the crosswind and vertical plume mixing respectively. They are typically assumed to be proportional to time and/or distance (Reible, 1999) and depend on atmospheric stability. Additional corrections for ground and inversion layer reflection can also be added. The model itself is very fast to compute even by hand, a trade-off with its simplifications (see Table A.1) and is typically used for source characterization research (Hutchinson et al., 2017).

Table A.1: Typical Gaussian plume model property representation (Lagzi et al., 2013)

Represented in most Gaussian models	Neglected in most Gaussian models
Advection	Wind shear
Horizontal and vertical diffusion	Temporal wind change
Ground and inversion layer reflection	Source rate/location change over time
Elevated sources	Deposition
Buoyancy (effective stack height correction)	Gravitational settling
Multiple point sources	Chemical reactions and decay
	Complex terrain
	3D diffusion

A.2.1 AERMOD

One of the most utilized Gaussian-based models for industrial settings (including upstream oil and gas) is AERMOD. AERMOD was developed collaboratively by AERMIC (American Meteorological Society/U.S. Environmental Protection Agency Regulatory Model Improvement Committee) and is an EPA preferred model for regulatory purposes.

The model itself includes turbulence parameterization (based on Monin-Obukhov theory) and models for complex terrain and urban boundary layers. Downwash is handled by PRIME (Plume RISE Model Enhancements) (Schulman et al., 2000), but has difficulty with some types of building configurations (Peterson et al., 2017).

A.3 Lagrangian Dispersion Models

In Lagrangian-based dispersion models, pollutant particles move along trajectories determined by velocity, buoyancy and turbulent effects (Lagzi et al., 2013). Particle tracking on a wind field is computationally simple and stochastic estimations of

concentration can be found from the final distribution if a large number of particles are used (Lagzi et al., 2013). The particle tracking itself is grid free (Lagzi et al., 2013), but a typical method of determining concentration is to use a volume average over a grid cell, though determining appropriate cell sizes can pose difficulties with random fluctuations at small volumes or detail loss at large volumes (de Haan, 1999). For sufficient particle densities at high detail levels, computational time can become quite large compared with Gaussian or Eulerian counterparts (Windtrax, 2018). Alternatively, a density distribution can be used to distribute particles in space rather than using large numbers of particles (de Haan, 1999). This leads to the two main types of Lagrangian-based dispersion models (Lagzi et al., 2013):

1. Trajectory models:

Trajectory models estimate each particle as a point and use the distribution of many particles to determine concentration.

2. Puff models

Puff models generally assume Gaussian dispersion inside each particle and use superposition of multiple particles to determine concentration.

For a typical puff model, puff mass maintains equivalence with the mass released during the initial prescribed time interval and the puff evolves through time. For a single puff, the central location moves with the computed trajectory and the concentration distribution takes a Gaussian form such as (Jung et al., 2003),

$$C = \frac{m}{(2\pi)^{3/2}\sigma_y^2\sigma_z} \exp\left[-\frac{x^2}{2\sigma_y^2}\right] \exp\left[-\frac{y^2}{2\sigma_y^2}\right] \exp\left[-\frac{(z-H)^2}{2\sigma_z^2}\right] \quad (\text{A.11})$$

where m is the mass of the puff released. Concentrations are summed for multiple puffs.

A.3.1 Windtrax

Windtrax (available at <http://www.thunderbeachscientific.com/>) is a short-range atmospheric dispersion modelling software based on Lagrangian stochastic trajectory particle models. In particular, trajectories can be computed in both forward (from sources) and backward (from sensors backwards in time) modes and the software can calculate emission rates for known source locations and estimate sensitivity of source/receptor

relationships for multiple source cases. An example application for emission calculation for a dairy farm can be found in Flesch et al. (2009).

A.3.2 CALPUFF

Another preferred U.S. EPA model is CALPUFF (available at <http://www.src.com/>) recommended for long range transport at distances where winds cannot be considered steady and homogenous throughout. CALPUFF consists of a meteorological model, an air dispersion model based on non-steady-state puffs and a post processing package. The use of puffs allows for multiple species to be represented and has algorithms for subgrid scale effects such as terrain impingement.

A.3.3 SCIPUFF

SCIPUFF, or Second-order closure model, is a 3D Gaussian puff model with diffusion parameters based on turbulence closure theory. Continuous, instantaneous, moving, and stack sources can be modelled for both buoyant and non-buoyant releases for short and long range transport. SCIPUFF has been used in several source characterization studies (e.g. Bieringer et al., 2015; Kopka et al., 2016; Allen et al., 2007)

A.4 Eulerian Approach to Dispersion Modelling

Eulerian-based models numerically solve Equation (A.9) on a spatially discretized mesh. This becomes an initial value problem that is marched forward in time in a manner similar to more familiar CFD codes. Depending on the domain and problem, this can be simple or complex and Eulerian-based models naturally scale between both. For complex domains and near-source regions, CFD and chemical reactions can be performed in conjunction (Lagzi et al., 2013).

Appendix B: L-BFGS Optimization Scheme

Gradient-based optimization schemes are often-used methods for their computational efficiency. The BFGS (Broyden, Fletcher, Goldfarb, Shanno) scheme falls within this category of optimization methods, as does its limited memory counterpart. Overall, the goal is to find the minimum cost function value (a fitness measure), with the cost function definition depending on the problem itself.

B.1 Gradient Descent

Gradient descent (GD) methods are frequently used in function optimization (Cheng, 2013). To find the minimum of a function, GD methods choose a step direction that is the opposite of the function gradient since this is the direction the function decreases fastest (Cheng, 2013). This can be expressed as (Cheng, 2013)

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \alpha_k \mathbf{d}_k \quad (\text{B.1})$$

where α_k is the step length and \mathbf{d}_k is the step direction at step k at point \mathbf{p} .

B.1.1 Newton's Method

For Newton's method, the second order Taylor expansion of the cost function J around \mathbf{p}_k at point \mathbf{p} is (Andrew, 2008)

$$m_k(\mathbf{p}) = J_k(\mathbf{p}) + (\mathbf{p} - \mathbf{p}_k)^T (\nabla J_k) + \frac{1}{2} (\mathbf{p} - \mathbf{p}_k)^T \nabla^2 (J_k) (\mathbf{p} - \mathbf{p}_k) \quad (\text{B.2})$$

where $\nabla^2 (J_k)$ is the Hessian matrix (symmetric matrix of second-order partial derivatives), which represents the curvature of J . This quadratic approximation is minimized when the gradient with respect to \mathbf{p} is the zero vector. Thus, if the gradient is

$$\nabla m_k(\mathbf{p}) = \nabla J_k + \nabla^2 (J_k) (\mathbf{p} - \mathbf{p}_k), \quad (\text{B.3})$$

\mathbf{p} may be found with the local quadratic approximation of J ,

$$(\mathbf{p} - \mathbf{p}_k) = -\nabla^2 (J_k)^{-1} \nabla J_k. \quad (\text{B.4})$$

The inverse of the Hessian matrix, however, requires considerable computational resources that may not be feasible for a large number of model parameters (Andrew, 2008), as is the case for the present source characterization problem.

B.1.2 Quasi-Newton Methods

Quasi-Newton methods reduce the expense of evaluating the Hessian by updating a running approximation of the Hessian (or its inverse). DFP (Davidon, 1959) was the first quasi-Newton method, but this was superseded by BFGS (Broyden, Fletcher, Goldfarb, Shanno) discovered in 1970. While DFP approximates the Hessian, BFGS approximates its inverse.

B.1.3 BFGS

Using an approximate Hessian, \mathbf{B}_k , updated each iteration, and \mathbf{d}_k as the search direction, Equation (B.2) can be rewritten as (Blomgren, 2012)

$$m_k(\mathbf{d}) = J(\mathbf{p}_k) + \nabla J(\mathbf{p}_k)^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T \mathbf{B}_k \mathbf{d} \quad (\text{B.5})$$

and Equation (B.4) as

$$\mathbf{d}_k = -\mathbf{B}_k^{-1} \nabla J(\mathbf{p}_k) \quad (\text{B.6})$$

Once the search direction \mathbf{d}_k and the step length α_k are known, Equation (B.1) (see Gradient Descent) can be used to find the next point \mathbf{p}_{k+1} . The step length must satisfy the Wolfe conditions (Blomgren, 2012)

$$\begin{aligned} J(\mathbf{p}_k + \alpha_k \mathbf{d}_k) &\leq J(\mathbf{p}_k) + c_1 \alpha_k \mathbf{d}_k^T \nabla J(\mathbf{p}_k) \\ \mathbf{d}_k^T \nabla J(\mathbf{p}_k + \alpha_k \mathbf{d}_k) &\geq c_2 \mathbf{d}_k^T \nabla J(\mathbf{p}_k) \\ 0 &< c_1 < c_2 < 1 \end{aligned} \quad (\text{B.7})$$

The approximate Hessian \mathbf{B}_k , unlike in the actual Newton's method, is not actually the matrix of second partial derivatives. Using the curvature information of the current step, this matrix is updated to \mathbf{B}_{k+1} (Blomgren, 2012) with

$$m_{k+1}(\mathbf{d}) = J(\mathbf{p}_{k+1}) + \nabla J(\mathbf{p}_{k+1})^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T \mathbf{B}_{k+1} \mathbf{d} \quad (\text{B.8})$$

Since the Hessian should not change significantly between iterations, the Hessian approximation \mathbf{B}_k should have a gradient equal to the function gradient for steps k and $k+1$ (Andrew, 2008). Thus (Blomgren, 2012),

$$\begin{aligned} \nabla m_{k+1}(-\alpha_k \mathbf{d}_k) &= \nabla J(\mathbf{p}_{k+1}) - \alpha_k \mathbf{B}_{k+1} \mathbf{d}_k = \nabla J(\mathbf{p}_k) \\ \alpha_k \mathbf{B}_{k+1} \mathbf{d}_k &= \nabla J(\mathbf{p}_{k+1}) - \nabla J(\mathbf{p}_k) \end{aligned} \quad (\text{B.9})$$

Defining

$$\begin{aligned} \mathbf{s}_k &= \mathbf{p}_{k+1} - \mathbf{p}_k = \alpha_k \mathbf{d}_k \\ \mathbf{y}_k &= \nabla J(\mathbf{p}_{k+1}) - \nabla J(\mathbf{p}_k) \end{aligned} \quad (\text{B.10})$$

this condition, the Secant Equation, is (Blomgren, 2012)

$$\mathbf{B}_{k+1} \mathbf{s}_k = \mathbf{y}_k \quad (\text{B.11})$$

Rather than determine \mathbf{B}_{k+1} , for BFGS the form of this equation is updated to determine its inverse.

$$\mathbf{H}_{k+1} \mathbf{y}_k = \mathbf{B}_{k+1}^{-1} \mathbf{y}_k = \mathbf{s}_k \quad (\text{B.12})$$

There are an infinite number of symmetric positive definite matrices that satisfy this condition. To determine \mathbf{H}_{k+1} uniquely, it is chosen so that it is closest to the previous iteration using the weighted Frobenius norm (Blomgren, 2012).

$$\begin{aligned} \mathbf{H}_{k+1} &= (\mathbf{I} - \rho_k \mathbf{s}_k \mathbf{y}_k^T) \mathbf{H}_k (\mathbf{I} - \rho_k \mathbf{y}_k \mathbf{s}_k^T) + \rho_k \mathbf{s}_k \mathbf{s}_k^T \rho_k = 1 / (\mathbf{y}_k^T \mathbf{s}_k) \\ \text{where } \mathbf{H} &= \mathbf{H}^T \end{aligned} \quad (\text{B.13})$$

Thus, the BFGS algorithm, assuming a starting point of \mathbf{p}_0 , a convergence tolerance of ϵ , and an initial Hessian approximation of \mathbf{H}_0 (this can be selected in various ways, including setting it to the identity matrix) can be summed up as (Blomgren, 2012)

```

k=0
while (  $\|\nabla J(\mathbf{p}_k)\| > \varepsilon$  )

     $\mathbf{d}_k = -\mathbf{H}_k \nabla J(\mathbf{p}_k)$ 

     $\mathbf{p}_{k+1} = \text{linesearch}(\mathbf{d}_k, \dots)$ 

     $\mathbf{s}_k = \mathbf{p}_{k+1} - \mathbf{p}_k$ 

     $\mathbf{y}_k = \nabla J(\mathbf{p}_{k+1}) - \nabla J(\mathbf{p}_k)$ 

     $\rho_k = 1 / (\mathbf{y}_k^T \mathbf{s}_k)$ 

     $\mathbf{H}_{k+1} = (I - \rho_k \mathbf{s}_k \mathbf{y}_k^T) \mathbf{H}_k (I - \rho_k \mathbf{y}_k \mathbf{s}_k^T) + \rho_k \mathbf{s}_k \mathbf{s}_k^T$ 

    k = k + 1

end-while

```

BFGS is considered the most popular and effective of the quasi-Newton methods (Skaaja, 2010). It is strongly self-correcting of bad matrix information after only a few updates and is robust (Skaaja, 2010).

B.1.4 Limited Memory BFGS

For a large number of variables, storing and manipulating the approximate inverse Hessian becomes difficult. The limited memory version of the BFGS algorithm (Nocedal, 1980) stores a compressed version of the approximate inverse Hessian (an $n \times n$ dense matrix) by using a limited history of vectors pairs \mathbf{s}_k and \mathbf{y}_k (each of size n) instead for matrix operations (Blomgren, 2012). Rather than containing updates from the initial step, the limited memory only uses recent updates. The equivalent starting matrix (at the beginning of the stored updates) is one that is easy to store, for example, a scalar multiple of the identity matrix based on the most recent gradient change (Andrew, 2008). If n is large, the limited memory BFGS can have significant increases in speed over the BFGS algorithm (Skaaja, 2010). Byrd et al. (1994) developed a bounded version of the algorithm (L-BFGS-B) that restricts the values of the predictions, which is useful in the target application where emission sources are positive emissions and chemical sinks are negligible.

Appendix C: Wind Profiles for Varying Atmospheric Stability

C.1 Atmospheric Turbulence, Stability, and the Obukhov Length

Atmospheric turbulence is made up of two main effects: wind shear causing mechanical turbulence, and thermal turbulence from buoyancy (Lagzi et al., 2013). The first of these is highly influenced by terrain, while the second by atmospheric stability and surface radiation (Lagzi et al., 2013).

Atmospheric stability, which describes the tendency to encourage or discourage vertical air motion, can be described in a number of ways. An early, and still popular, method is the Pasquill's stability classes which accounts for wind speed, sun elevation, and cloud cover to define mixing efficiency (Lagzi et al., 2013). However, as Pasquill's classes define only six discrete turbulence intensity values, other methods have increased in usage over time.

One such method is the use of the Obukhov length (also called the Monin-Obukhov length), which combined vertical turbulent momentum and heat fluxes into a dimensioned length value (Lagzi et al., 2013)

$$L = - \frac{(-\overline{u'w'})^{\frac{3}{2}}}{\kappa \frac{g}{T} \overline{T'w'}} \quad (\text{C.1})$$

where T' , u' , w' are temperature, horizontal, and vertical wind fluctuations, T is the temperature and κ is the von Karman constant (~ 0.4). Using the turbulent heat flux (q) and friction velocity (u^*) (Alinot, 2003),

$$u^* = \sqrt{\frac{\tau_w}{\rho}} \quad (\text{C.2})$$

this can be expressed as (Foken, 2006)

$$L = - \frac{u^{*3}}{\kappa \frac{g}{T} \frac{q}{c_p \rho}} \quad (\text{C.3})$$

where c_p is the specific heat and ρ the density of air. A dimensionless parameter can be obtained by scaling the height from the ground (z) by the Obukhov length. Table C.1 summarizes the implications of this non-dimensional parameter.

Table C.1: Obukhov length and atmospheric stability (Alinot, 2003)

z/L	Atmospheric stratification	Implication
$z/L > 0$	Stable	Mechanical turbulence reduced by thermal stratification
$z/L = 0$	Neutral	Purely mechanical turbulence
$z/L < 0$	Unstable	Turbulence increased by thermal stratification

C.2 Wind Profiles and Obukhov Length

Turbulent viscosity, μ_T , may also be expressed as a function of height (z), friction velocity (u^*), and the Obukhov length (L) (Alinot, 2003) as

$$\mu_T = \frac{\rho \kappa u^* z}{\phi_m(z/L)} \quad (\text{C.4})$$

where the function ϕ_m varies with atmospheric stability and the most common expressions are given in Table C.2.

Table C.2: Expressions of ϕ_m for different atmospheric conditions (Alinot, 2003)

Stability	$\phi_m\left(\frac{z}{L}\right)$
Neutral	1
Stable	$1+5(z/L)$
Unstable	$(1-16z/L)^{-1/4}$

The velocity profile (u), for various stratifications with roughness height z_0 , is (Alinot, 2003)

$$u = \frac{u^*}{\kappa} \left[\ln \left(\frac{z}{z_0} \right) - \Psi_m \left(\frac{z}{L} \right) \right] \quad (\text{C.5})$$

where $\Psi_m \left(\frac{z}{L} \right)$ for various conditions may be found in Table C.3.

Table C.3: Expressions of Ψ_m for different atmospheric conditions (Alinot, 2003)

Stability	$\Psi_m \left(\frac{z}{L} \right)$
Neutral	0
Stable	$-5(z/L)$
Unstable	$\ln \left[\left(\frac{1+(1/\phi_m)^2}{2} \right) \left(\frac{1+(1/\phi_m)^2}{2} \right)^2 \right] - 2 \arctan \left(\frac{1}{\phi_m} \right) + \frac{\pi}{2}$

Turbulent energy dissipation can also be expressed in the form (Alinot, 2003)

$$\varepsilon = \frac{u^{*3}}{\kappa z} \phi_\varepsilon(z/L) \quad (\text{C.6})$$

where the function ϕ_ε varies with atmospheric stability and the most common expressions are given in Table C.4.

Table C.4: Expressions of ϕ_ε for different atmospheric conditions (Alinot, 2003)

Stability	$\phi_\varepsilon \left(\frac{z}{L} \right)$
Neutral	1
Stable	$\phi_m \left(\frac{z}{L} \right) - \left(\frac{z}{L} \right)$
Unstable	$1 - \left(\frac{z}{L} \right)$

Appendix D: Project Prairie Grass Mesh Refinement

Forward transport releases were performed on three grids of varying refinements, using the same source definitions, based on the Project Prairie Grass Release 19 (Barad, 1958). The region considered and source location are outline in Chapter 4. The medium mesh (M) corresponds to the mesh used throughout this work. Details are found in Chapter 4.

Table D.1: Mesh spacing compared with Medium Mesh

Mesh	Number of Cells	Cell side length compared with Medium Mesh
Coarser (C)	190 thousand	2
Finer (F)	12 million	0.5

For comparison purposes, the source definition (which is volumetric) was maintained for each simulation. Similarly, concentrations were compared for the same measurement volumes corresponding at the 50 m arc positions. Details of Richardson extrapolation and grid convergence index (GCI) can be found at Slater (2008).

D.1 Grid convergence index: maximum concentration at 50 m

Order of convergence: 4.3

Coarser to Medium Mesh: $GCI_{CM} = 1.968\%$

Medium to Fine Mesh: $GCI_{MF} = 0.079\%$

Asymptotic convergence range check (should be approximately 1.0): 1.015

D.2 Grid convergence index: concentration spread at 50 m

The plume spread at 50 m was compared with the angle of 1% of the maximum concentration (using the Richardson extrapolated definition) on both sides of the plume.

Left side:

Order of convergence: 1.8

Coarser to Medium Mesh: $GCI_{CM} = 0.51\%$

Medium to Fine Mesh: $GCI_{MF}=0.14\%$

Asymptotic convergence range check (should be approximately 1.0): 0.996

Right side:

Order of convergence: 0.9

Coarser to Medium Mesh: $GCI_{CM}=3.2\%$

Medium to Fine Mesh: $GCI_{MF}=1.7\%$

Asymptotic convergence range check (should be approximately 1.0): 1.014

For all comparisons, the GCI was less than 5%, making the medium mesh sufficiently representative of the expected plume dispersion.

Appendix E: Pre-convergence Location Predictions for Project Prairie Grass

Useful information about the location prediction can be obtained even before convergence is achieved. Intermediate location predictions for the Project Prairie Grass case (described in Section 4.5.1) are presented here at a vertical height of 0.5 m (approximately equal to the release height).

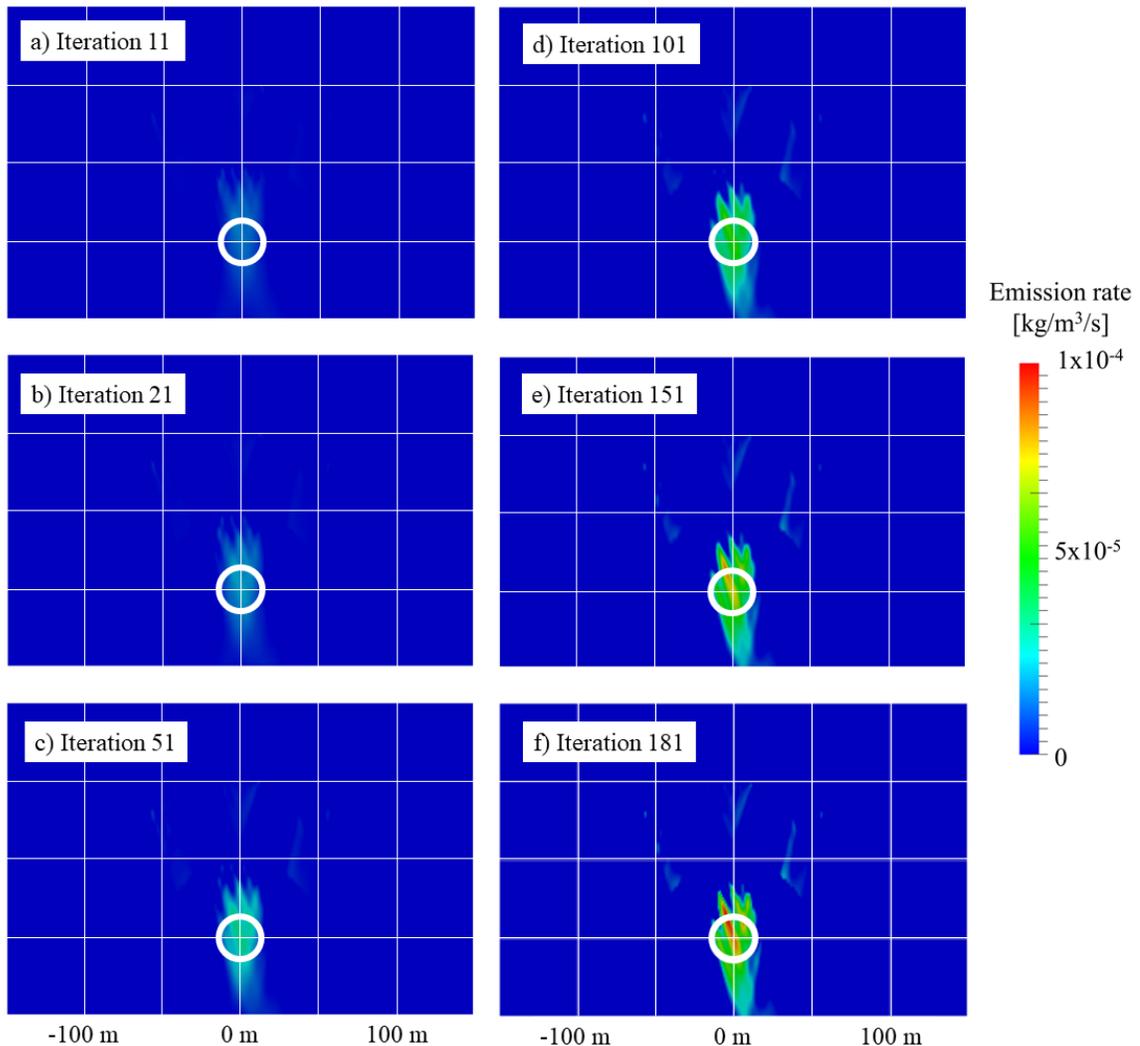


Figure E.1: Project Prairie Grass emission rate prediction at a height of 0.5 m for various convergence iterations. Actual release location is marked by white circle.

Appendix F: Optimization from Different Initial Conditions

Throughout the results presented in the thesis, initial conditions were started by assuming no emissions throughout the domain. This provided the fast convergence and did not bias the result towards a particular initial location. However, as the optimization scheme chosen is a local optimization, beginning at a different initial condition can tend towards a different local minimum point. Using the Project Prairie Grass test case, convergence from several different initial conditions (summarized with the actual release information in Table F.1) are shown here. Cold restarting (treating the previous guess as a new initial condition and removing the time-history from the Hessian calculation) was also attempted in some cases to test convergence as further noted below.

Table F.1: Actual Release Information and Tested Initial Conditions for Project Prairie Grass

Actual / Initial Condition	Release Description
Actual release from PPG (desired result)	Point source (0.102 kg/s) at (0 m 0 m 0.46 m)
No release	No release within domain (results presented in thesis)
10× larger point release near actual source location	1 kg/s at (0 m 0 m 5 m)
Horizontally offset point release of same magnitude that affects sensor readings	0.1 kg/s at (-25 m, 0, 0.5 m)
10× larger point at high elevation (minimal effect on sensor readings)	1 kg/s at (0 m 0 m 50 m)
Volumetric release of correct magnitude dispersed in cube offset from actual source	0.0001 kg/m ³ /s (0.1 kg/s) in cube volume from (-5 m 10 m 0 m) to (5 m 20 m 10 m)

For the results shown below, emission regions were computed (using a 1% of the maximum threshold to define region boundaries). However, for cases where an excessively large (10×) initial guess source could obscure actual sources, comparisons are also shown including other regions where necessary. The presented normalized cost functions are all normalized by the cost function of a zero-emission release throughout for consistency and are presented at the same absolute convergence criteria (a consistent change of less than 0.01% of the initial value per iteration) unless otherwise stated.

F.1 Initial Condition of a Large Point Release Near Actual Source Location

A large point release (an order of magnitude greater than the actual release), 4.5 m above the actual release location, was investigated as a second initial condition case. Results are summarized in Figure F.1 and Figure F.2. A cold restart was performed at iteration 2021. Compared to a zero-source initial guess (Figure 4.3), the convergence is more than an order of magnitude slower. Nevertheless, after 3502 iterations, the algorithm predicts a source with close to the correct magnitude (within 12%) but still at the initial guess location. However, this location error of 4.5 m is comparable to the 4.6 m error for the zero initial source guess case considered in Chapter 4.

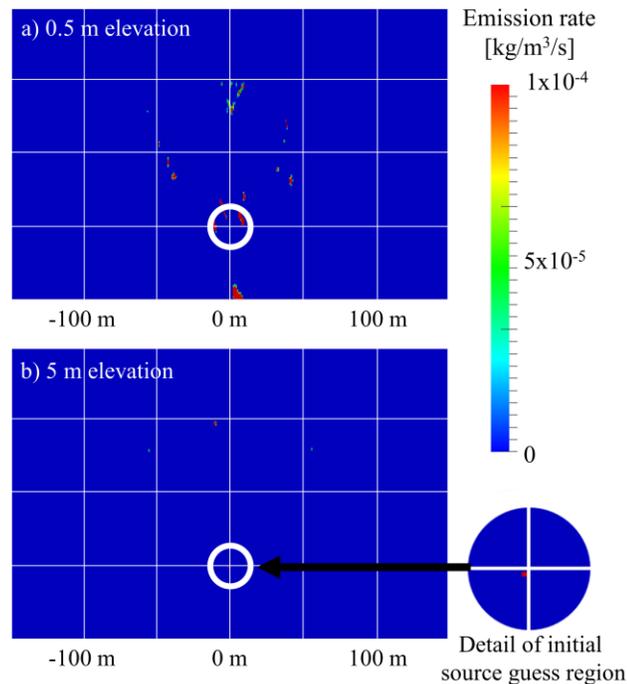


Figure F.1: Project Prairie Grass emission rate prediction with an initial guess of a $10\times$ larger source, 5 m above the actual source. Actual release location is marked by white circle. Results after 3502 iterations showing (a) predicted sources at 0.5 m elevation (elevation of the actual source) and (b) predicted sources at 5 m elevation (elevation of initial guess source), with enlarged image of region within white circle illustrating a source predicted at initial guess location.

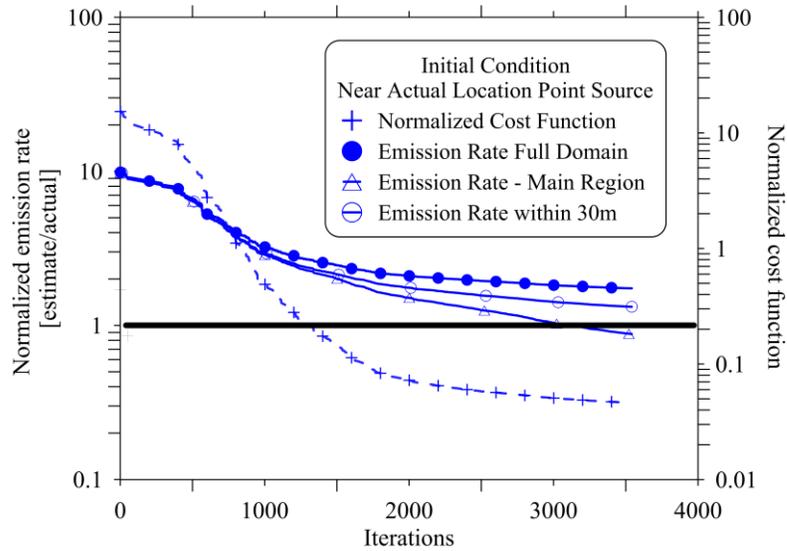


Figure F.2: Predicted emission rates and convergence for PPG with an initial guess of a $10\times$ larger source, 5 m above the actual source.

F.2 Initial Condition of a Horizontally Offset Point Release of Similar Magnitude

Results for the initial guess case of a horizontally offset source at same elevation (-25 m 0 m 0.5 m) and similar magnitude to the actual source are shown in Figure F.3 and Figure F.4. A cold-restart was performed at iteration 1700 and the convergence criteria was met at approximately 2000 iterations. As shown in Figure F.3, the centroid of the predicted source region is within 3 m of the actual source, as good or better than the zero initial source guess case considered in Chapter 4. Figure F.4 shows the convergence of the cost function, as well as the evolution of predicted emission rates in the entire domain, at the initial guess release location, and in dominant predicted source region (which is near the actual source location). The algorithm reduces the initial guess source in a slow, nearly linear fashion with the number of iterations. The main predicted source region (near the actual source) converges within 13% of the actual source emission rate (which is the same as the zero initial source guess case) within 175 iterations. This is comparable to the 191 iterations for convergence of the zero initial source guess case.

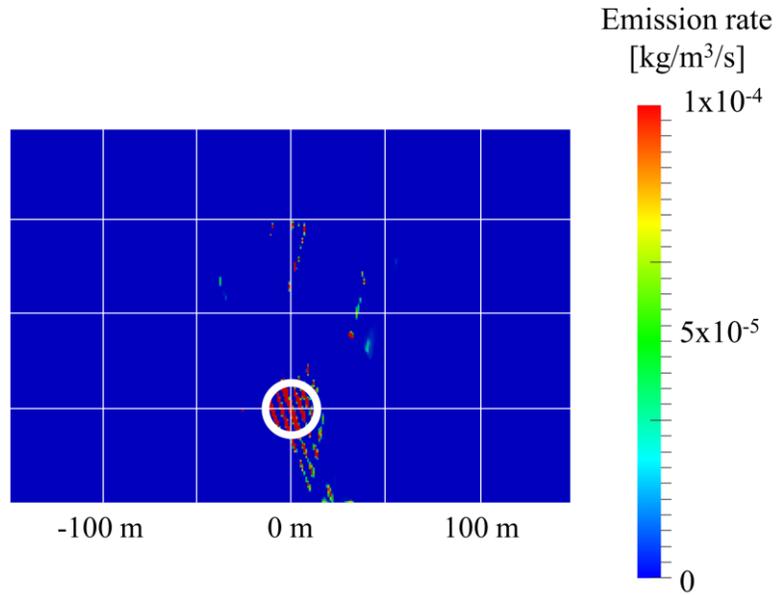


Figure F.3: Project Prairie Grass emission rate prediction with an initial guess of a horizontally offset source at same elevation (-25 m 0 m 0.5 m) and similar magnitude to the actual source. Actual source location is marked by white circle. Results showing predicted sources at 0.5 m elevation after 2000 iterations.

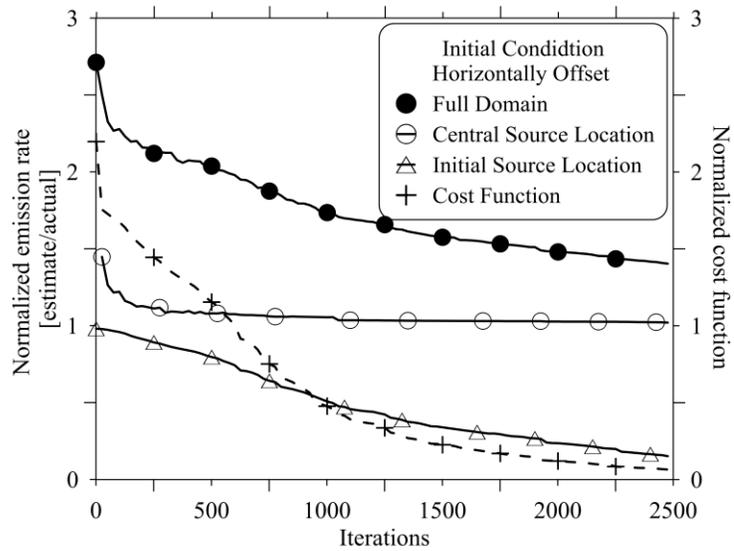


Figure F.4: Predicted emission rates and convergence for PPG with an initial guess of a horizontally offset source at same elevation (-25 m 0 m 0.5 m) and similar magnitude to the actual source.

F.3 Initial Condition of a Large Point Release at High Elevation

A large point release (an order of magnitude greater than the actual release rate), 50 m above the actual release location, was also tested as an initial condition. As shown in Figure F.5 and Figure F.6, the algorithm essentially ignored this initial source, which was too high in elevation to significantly affect the sensors in the domain. A separate source was correctly predicted near the actual source location (within 4.4 m) at the correct magnitude (within 1%), which was of similar accuracy as the zero initial source guess case considered in Chapter 4. Convergence was only slightly slower than the for no-release case for the same convergence criteria.

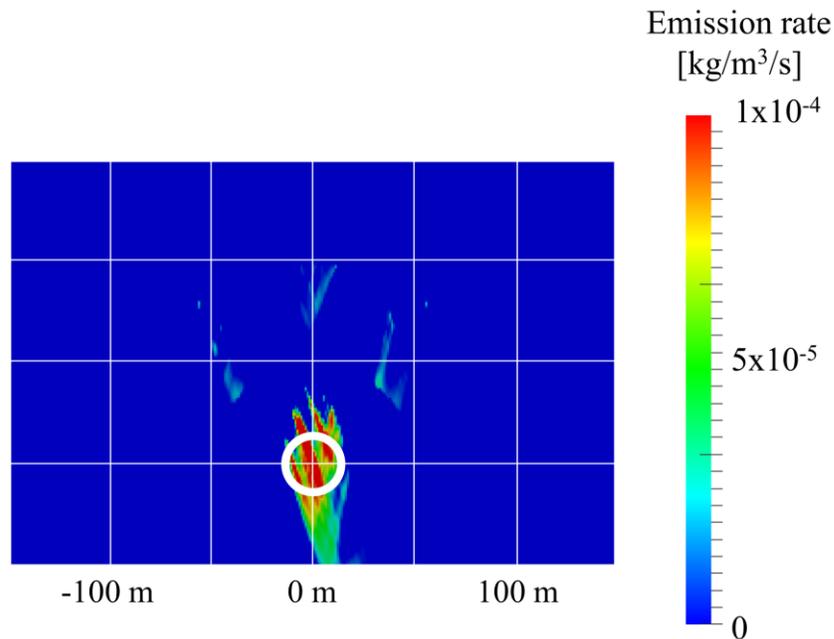


Figure F.5: Project Prairie Grass emission rate prediction with an initial guess of a $10\times$ larger source, 50 m above the actual source. Actual release location is marked by white circle. Results shown at 0.5 m elevation after 216 iterations.

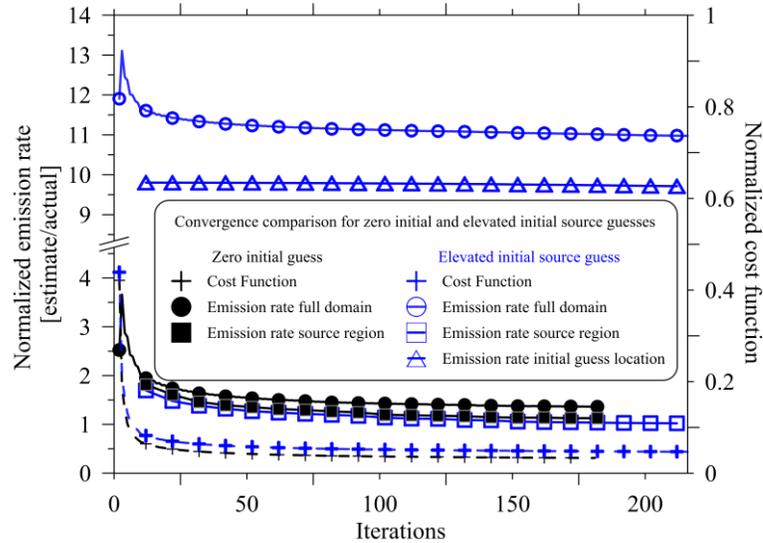


Figure F.6: Predicted emission rates and convergence for PPG with an initial guess of a $10\times$ larger source, 50 m above the actual source. Black lines and symbols show the zero initial source guess case of Chapter 4 for comparison. Emission rate of the initial guess location (blue triangle) remains effectively unchanged, while predict source near actual iterates toward correct answer.

F.4 Initial Condition of a Volume Release

Figure F.7 and Figure F.8 show the predicted source location and convergence characteristics for an initial condition of a volumetric release of the correct magnitude in a $\sim 10\text{ m}\times 10\text{ m}\times 10\text{ m}$ cube located at (0 m, +15 m, +5 m), which is 15 m north and 4.5 m above the actual source. The algorithm predicts the correct source location and magnitude with similar accuracy as the zero initial source guess case considered in Chapter 4.

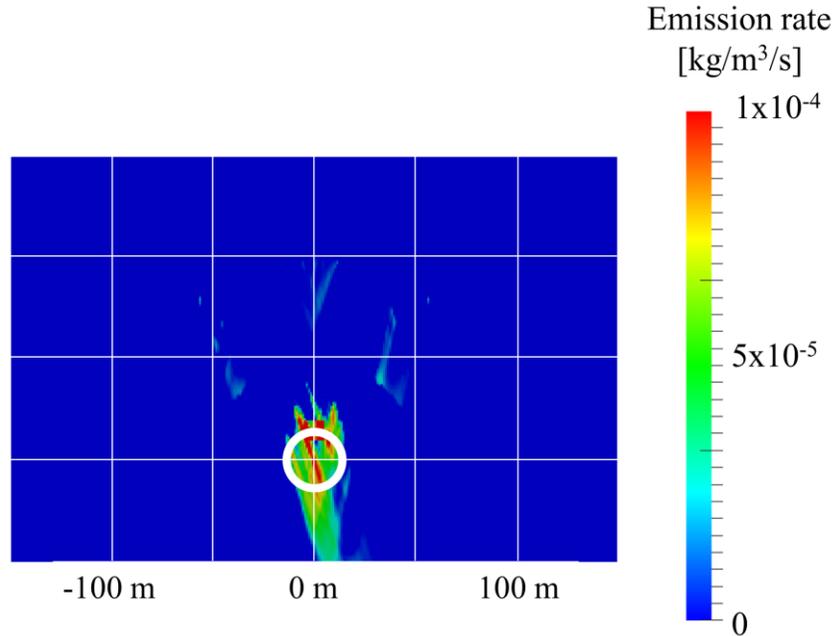


Figure F.7: Project Prairie Grass emission rate prediction with an initial guess of a $\sim 10 \text{ m} \times 10 \text{ m} \times 10 \text{ m}$ volumetric source 15 m north and 4.5 m above the actual source. Results shown at a height of 0.5 m for iteration 191. Actual release location is marked by white circle.

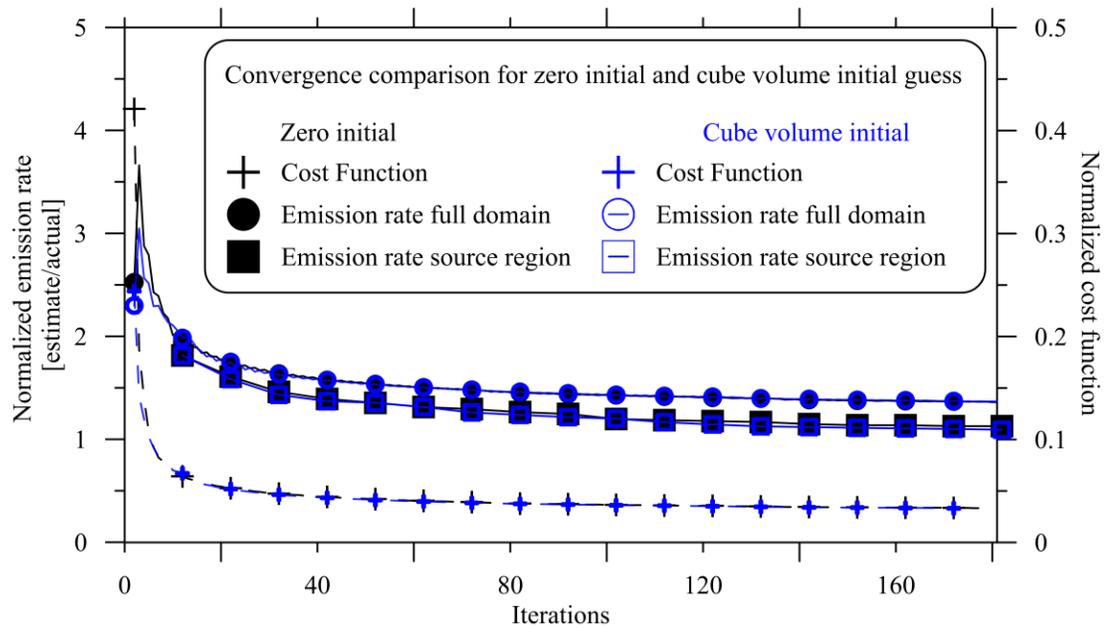


Figure F.8: Predicted emission rates and convergence for PPG with an initial guess of $\sim 10 \text{ m} \times 10 \text{ m} \times 10 \text{ m}$ volumetric source 15 m north and 4.5 m above the actual source.

F.5 Summary of Effects of Different Initial Conditions

The different initial condition cases examined in this Appendix suggest that although the problem is ill-posed and the optimization scheme could be subject to local minima, the algorithm is robust. Correct solutions with comparable accuracy to a zero-initial source guess case were found even when starting with incorrect guessed source magnitudes and locations. Convergence time could be significantly increased however as noted in the case of an initially guessed $10\times$ larger than actual source near the true source location. Initial guessed sources outside the coverage region of the sensors were effectively ignored without impacting convergence times. While the problem remains ill-posed such that there are likely initial guess conditions that would cause the algorithm to fail, all results in this thesis suggest that a starting guess of zero sources in the domain should consistently and efficiently produce useful solutions.