**Proof-of-Concept Inverse Micro-Scale Dispersion Modelling for Fugitive Emissions
Quantification in Industrial Facilities**

by

Ian M. Joynes

A thesis submitted to The Faculty of Graduate Studies and Research

in partial fulfilment of the degree requirements of

Master of Applied Science in Mechanical Engineering

Ottawa-Carleton Institute for

Mechanical and Aerospace Engineering

Department of Mechanical and Aerospace Engineering

Carleton University

Ottawa, Ontario, Canada

December 2013

Copyright © 2013 – Ian M. Joynes

The undersigned recommend to
the Faculty of Graduate Studies and Research
acceptance of the thesis

**Proof-of-Concept Inverse Micro-Scale Dispersion Modelling for Fugitive Emissions
Quantification in Industrial Facilities**

Submitted by **Ian M. Joynes**

in partial fulfilment of the requirements for the degree of
Master of Applied Science in Mechanical Engineering

Dr. Matthew Johnson, Thesis Supervisor

Dr. Amir Hakami, Thesis Supervisor

Dr. Metin Yaras, Chair, Department of Mechanical and Aerospace Engineering

Carleton University

2013

Abstract

A proof-of-concept study was conducted to locate and quantify fugitive emissions within a gas processing plant with inverse micro-scale dispersion modelling. A synthetic wind field and receptor observations were generated for a simplified model of a gas processing plant. Source reconstruction was performed with an objective function that measured the misfit of candidate emission source distributions. The objective function gradient was evaluated with adjoint sensitivity analysis, and the candidate emission source distribution that minimized the objective function was found using the L-BFGS-B optimization algorithm. The inverse dispersion model was tested under non-ideal conditions such as multiple emission sources, unfavourable receptor coverage, and the addition of observational noise. In an effort to mitigate the prediction of false emission sources, objective function regularization and emission source filtering were investigated. It was found that a combination of regularization and filtering achieved the best estimates of the emission locations and total emission rates.

Acknowledgements

Principally, I would like to thank my supervisors Prof. Matthew Johnson and Prof. Amir Hakami. I certainly would not have succeeded without their support and guidance. Additionally, I owe much gratitude to Carol Brereton and Mike Shives for their assistance with running CFD simulations.

I would also like to thank the other members of Prof. Johnson's research group who provided companionship and support during my graduate studies. Specifically, I'd like to thank, in no particular order, Stephen Schoonbaert, Graham Ballachey, Jan Gorski, Darcy Corbin, Clifton Pereira, Robin Devillers, Brian Crosland, Patrizio Vena, James McEwen, and Adam Coderre. Additionally, I would also like to thank the other students from the different research groups who became my colleagues and friends.

Finally, I owe my success to Katelyn Geraldini and my wonderful family who have been there to support me every step of the way.

Table of Contents

Abstract	iii
Acknowledgements	iv
List of Tables	x
List of Figures	xi
Nomenclature	xiv
Chapter 1 Introduction	1
1.1 Review of Inverse Dispersion Modelling	3
1.1.1 Lagrangian Back-Trajectory Methods	3
1.1.2 Sequential Data Assimilation	5
1.1.2.1 Bayesian Inference	5
1.1.2.2 Kalman Filter	7
1.1.3 Variational Data Assimilation	9
1.1.3.1 Global Optimization	10
1.1.3.2 Local Optimization	12
1.2 Inverse Dispersion Modelling for Fugitive Emissions	13
1.2.1 Transport Modelling Method	13
1.2.2 Emission Source Parameterization	14
1.2.3 Inversion Method	15
1.3 Thesis Overview and Objectives	16
1.4 Thesis Organization	18
Chapter 2 Inverse Modelling	20
2.1 Dispersion Model	21

2.2	Objective Function	22
2.3	Numerical Optimization	25
2.3.1	General Gradient-based Methods	27
2.3.2	L-BFGS-B Method	29
2.4	Objective Function Gradient	32
2.4.1	Review of Differentiation Techniques	32
2.4.2	Forward Sensitivity Method	35
2.4.3	Adjoint Sensitivity Method	37
2.5	Summary	40
Chapter 3	Dispersion Modelling	42
3.1	Governing Transport Equation	43
3.1.1	Initial and Boundary Conditions	44
3.1.2	Turbulent Mixing and Mass Diffusion	46
3.2	Numerical Solution	48
3.2.1	Boundary Value Problem	49
3.2.1.1	The Finite Element Method of Weighted Residuals	52
3.2.1.2	Interpolation Functions	53
3.2.1.3	Upwinding	58
3.2.1.4	Boundary Conditions	60
3.2.1.5	Elemental Integration	62
3.2.1.6	Global Assembly	62
3.2.2	Initial Value Problem	63
3.2.2.1	The θ -Method	64
3.3	Dispersion Model	66
Chapter 4	Inverse Dispersion Model Setup	68

4.1	CFD Modelling	70
4.1.1	Spatial Domain Discretization	71
4.1.2	Boundary Conditions	73
4.1.2.1	Inflow	73
4.1.2.2	Outflow	75
4.1.2.3	Buildings Surfaces	75
4.1.2.4	Lateral Surfaces	75
4.1.3	Initial Conditions	76
4.1.4	Transient Simulation	77
4.2	Inverse Dispersion Model Parameters.....	78
4.2.1	Initial Candidate Solution	79
4.2.2	Background Emission Sources	79
4.2.3	L-BFGS-B Parameters	80
4.2.3.1	Inverse Hessian Approximation.....	80
4.2.3.2	Box Bound Constraints	81
4.2.3.3	Termination Criteria.....	82
4.2.4	IDM Parameter Summary	84
4.3	Inverse Dispersion Model Test Cases	84
4.3.1	Receptor Locations	85
4.3.2	Single-Source Fugitive Emissions Test Case	87
4.3.3	Multiple Source Fugitive Emissions Test Case	88
4.3.4	Receptor Measurement Noise	89
4.3.5	Objective Function Regularization	91
4.3.6	Emission Source Filter	92
4.4	Measures of Model Fitness	94

4.4.1	Objective Function Reduction	95
4.4.2	Gradient Norm Reduction.....	95
4.4.3	Coefficient of Determination	96
4.4.4	Total Emission Rate.....	97
Chapter 5	Inverse Dispersion Model Results	99
5.1	Single-Source Test Case Results.....	99
5.1.1	Predicted Emission Source Distribution	99
5.1.1.1	Objective Function Regularization	103
5.1.2	Convergence	108
5.1.2.1	Objective Function Reduction	108
5.1.2.2	Coefficient of Determination	110
5.1.3	Total Emission Rate.....	113
5.1.3.1	Emission Source Filtering.....	114
5.1.4	Single-Source Summary	118
5.2	Multi-Source Test Case Results	121
5.2.1	Predicted Emission Source Distribution	121
5.2.1.1	Objective Function Regularization	123
5.2.2	Convergence	128
5.2.2.1	Objective Function Reduction	128
5.2.2.2	Coefficient of Determination	130
5.2.3	Total Emission Rate.....	132
5.2.3.1	Emission Source Filtering.....	134
5.2.4	Multi-Source Summary.....	138
5.3	Summary	142
Chapter 6	Conclusions and Recommendations	144

6.1	Conclusions	144
6.1.1	Observational Noise.....	145
6.1.2	Objective Function Regularization	146
6.1.3	Emission Source Filtering.....	147
6.2	Recommendations and Future Work.....	147
6.2.1	Three-Dimensional Model	147
6.2.2	Objective Function Regularization	149
6.2.3	Emission Source Filtering.....	150
6.2.4	Receptor Observation Uncertainty.....	150
6.2.5	Flow Field Inversion	151
6.2.6	Receptor Placement	152
	References.....	153
	Appendix A Continuous Sensitivity Analysis	165
	Appendix B Mesh Quality	178
	Appendix C ANSYS CFX Convergence	180
	Appendix D Receptor Locations.....	184
	Appendix E Multi-Source Locations	185
	Appendix F Objective Function Gradient Convergence.....	186
	Appendix G Multi-Source Convergence	191

List of Tables

Table 3.1: Summary of advection-diffusion equation boundary conditions.....	60
Table 3.2: Stability of θ -methods	65
Table 4.1: L-BFGS-B parameter search space boundary types	81
Table 4.2: Summary of IDM parameter settings	84
Table 5.1: Peak location error and full area half maximum (FAHM) of volumetric emission rate for the final iteration of the single-source test case	118
Table 5.2: Summary of the influence of source filtering and regularization on emission rate of the final candidate emission sources for the single-source test case	119
Table 5.3: Emission source location error in [m] for the multi-source test case	139
Table 5.4: Emission source FAHM in [m ²] for the multi-source test case	140
Table 5.5: Summary of the influence of objective regularization and emission source filtering on the final candidate emission sources for the multi-source test case.....	141
Table D.1: Synthetic receptors locations	184
Table E.1: Multi-source location and emission rate	185

List of Figures

Figure 2.1: Flow chart overview of the inverse dispersion modelling procedure	41
Figure 3.1: Two-dimensional domain of an industrial gas processing plant	44
Figure 3.2: Shape function of a single node over the connected triangular elements	54
Figure 4.1: Two-dimensional airshed domain of an industrial gas processing plant.....	69
Figure 4.2: Discretized computational airshed domain	72
Figure 4.3: Time-variation of inflow wind speed and direction	74
Figure 4.4: Wind rose of inflow wind conditions	74
Figure 4.5: Initial conditions of flow field from ANSYS CFX	77
Figure 4.6: Synthetic receptor locations	86
Figure 4.7: Specified known distribution of the single-source test case.....	88
Figure 4.8: Multiple volumetric source position, 2.33× magnification	89
Figure 4.9: Sample of synthetic receptor observations for the single-source test case.....	90
Figure 4.10: Sample of synthetic receptor observations for the multi-source case	91
Figure 4.11: Receptor FOV PDF and source filters.....	93
Figure 5.1: Iterative progression of candidate emission source distributions for the single-source test case without objective function regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)	100
Figure 5.2: Final emission source distributions (log scale) for the single-source test cases without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)	102
Figure 5.3: Effect of regularization on final candidate emission source distributions (linear scale) for the single-source test case	104
Figure 5.4: Effect of regularization on final candidate emission source distributions (log scale) for the single-source test case	106

Figure 5.5: Convergence history of the objective function minimization for the single-source test case.....	109
Figure 5.6: Convergence history of the coefficient of determination of modelled receptor observations for the single-source test case	111
Figure 5.7: Convergence history of the total emission rate for the single-source test case	113
Figure 5.8: Final emission source distributions with emission source filters for the single-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)	115
Figure 5.9: Convergence history of filtered total emission rate for the single-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)	116
Figure 5.10: Convergence history of filtered total emission rate for the single-source test case with regularization	117
Figure 5.11: Final candidate emission source distributions for the multi-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$).....	122
Figure 5.12: Effect of regularization on the final candidate emission source distributions for the multi-source test case	124
Figure 5.13: Effect of regularization on the final candidate emission source distributions for the multi-source test case	126
Figure 5.14: Convergence history of objective function minimization for the multi-source test case	129
Figure 5.15: Convergence history of coefficient of determination of receptor observations for multi-source case.....	131
Figure 5.16: Convergence history of the total emission rate for the multi-source test case	133
Figure 5.17: Emission source filters for the final emission source distribution of the multi-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)	135

Figure 5.18: Convergence history filtered total emission rate for the multi-source test case	136
Figure 5.19: Convergence history of adjusted filtered total emission rate for the multi-source test case.....	138
Figure B.1: Probability density function of mesh quality	178
Figure C.1: Convergence history of RMS relative residual error of initial flow conditions	181
Figure C.2: Convergence history of maximum relative residual error of initial flow conditions	181
Figure C.3: Convergence history of RMS relative residual error of transient wind simulation sampled at 4 s increments	183
Figure C.4: Convergence history of max relative residual error of transient wind simulation sampled at 4 s increments	183
Figure F.1: Convergence history of objective function gradient for the single-source test case.....	187
Figure F.2: Convergence history of objective function gradient for the multi-source test case.....	189
Figure G.1: Iterative progression of candidate emission source distributions (linear scale) for the multi-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$).....	192
Figure G.2: Iterative progression of candidate emission source distributions (log scale) for the multi-source test without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$).....	194

Nomenclature

Latin Symbol	Description	Units	First Use Eq.	Pg.
A	entire dispersion model coefficient matrix	$[\text{m}^3 \cdot \text{s}^{-1}]$	(3.46)	66
A_n	dispersion model coefficient matrix for the <i>n</i> th time-step	$[\text{m}^3 \cdot \text{s}^{-1}]$	(3.41)	65
a	arbitrary vector	[-]	(4.3)	83
<i>a_i</i>	<i>i</i> th entry of the arbitrary vector a	[-]	(4.3)	83
B	error-covariance matrix of the volumetric emission rate	$[\text{m}^3 \cdot \text{s}^{-1}]$	(2.2)	23
B_n	dispersion model coefficient matrix for the <i>n</i> th time-step	$[\text{m}^3 \cdot \text{s}^{-1}]$	(3.43)	65
<i>b_i</i>	<i>i</i> th element of the dispersion model forcing vector	$[\text{kg} \cdot \text{s}^{-1}]$	(2.22)	37
b	entire dispersion model forcing vector	$[\text{kg} \cdot \text{s}^{-1}]$	(3.47)	66
b_n	dispersion model forcing vector for the <i>n</i> th time-step	$[\text{kg} \cdot \text{s}^{-1}]$	(3.44)	65
C	global finite element capacitance matrix	$[\text{m}^3]$	(3.37)	62
C^(e)	local finite element capacitance matrix	$[\text{m}^3]$	(3.28)	57
C_n	global finite element capacitance matrix for the <i>n</i> th time-step	$[\text{m}^3]$	(3.39)	64
<i>C_μ</i>	<i>k-ε</i> turbulence model constant Standard <i>k-ε</i> : <i>C_μ</i> = 0.09 (Jones and Launder, 1972) RNG <i>k-ε</i> : <i>C_μ</i> = 0.085 (Yakhot et al., 1992)	[-]	(3.6)	47
<i>c</i>	pollution mass concentration	$[\text{kg} \cdot \text{m}^{-3}]$	(3.1)	43
<i>c₀</i>	inital distribution of pollutant concentration	$[\text{kg} \cdot \text{m}^{-3}]$	(3.2)	44
<i>c_i</i>	<i>i</i> th element of a vector of pollutant concentration	$[\text{kg} \cdot \text{m}^{-3}]$	(2.20)	37
<i>c_i^(e)</i>	<i>i</i> th element of an elemental vector of pollutant	$[\text{kg} \cdot \text{m}^{-3}]$	(3.20)	56

Latin Symbol	Description	Units	First Use Eq.	Pg.
$\bar{c}$	Reynolds-averaged pollutant concentration	$[\text{kg}\cdot\text{m}^{-3}]$	(3.4)	46
$\overline{c^0}$	average of all known receptor observations	$[\text{kg}\cdot\text{m}^{-3}]$	(4.6)	97
$\overline{c'\mathbf{u}'}$	Reynolds-averaged mass flux from turbulent mixing	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(3.4)	46
$\mathbf{c}$	discretized global pollutant concentration	$[\text{kg}\cdot\text{m}^{-3}]$	(-)	6
$\mathbf{c}^0$	column vector of known receptor observations of pollutant mass concentration	$[\text{kg}\cdot\text{m}^{-3}]$	(1.1)	5
$\mathbf{c}_0$	initial discretized global pollutant concentration	$[\text{kg}\cdot\text{m}^{-3}]$	(3.45)	66
$\mathbf{c}_n$	discretized global pollutant concentration at the n th time-step	$[\text{kg}\cdot\text{m}^{-3}]$	(3.38)	64
$\mathbf{c}^{(e)}$	discretized local pollutant concentration	$[\text{kg}\cdot\text{m}^{-3}]$	(3.20)	56
D	total isotropic mass diffusivity	$[\text{m}^2\cdot\text{s}^{-1}]$	(3.9)	48
D_{AB}	binary molecular mass diffusivity (methane in air: $D_{AB} \approx 2.2\times 10^{-5} \text{ m}^2\cdot\text{s}^{-1}$)	$[\text{m}^2\cdot\text{s}^{-1}]$	(3.9)	48
D_t	isotropic turbulent eddy mass diffusivity	$[\text{m}^2\cdot\text{s}^{-1}]$	(3.8)	48
D_{tx}, D_{ty}, D_{tz}	turbulent eddy mass diffusivity in the Cartesian coordinate directions	$[\text{m}^2\cdot\text{s}^{-1}]$	(3.5)	46
$\mathbf{D}$	diagonal mass diffusivity matrix	$[\text{m}^2\cdot\text{s}^{-1}]$	(3.1)	43
$\mathbf{D}_t$	turbulent eddy mass diffusivity matrix	$[\text{m}^2\cdot\text{s}^{-1}]$	(3.4)	46
$\mathbf{D}_i^{(e)}$	mass diffusivity at the i th node of an element	$[\text{m}^2\cdot\text{s}^{-1}]$	(3.22)	56
$\mathcal{D}^{(e)}$	rank 3 tensor of nodal mass diffusivities of a single finite element	$[\text{m}^2\cdot\text{s}^{-1}]$	(3.15)	55
$epsmch$	machine precision estimated by the L-BFGS-B numerical optimization algorithm (2.22×10^{-16})	$[-]$	(4.1)	82
f	objective function	$[\text{kg}^2\cdot\text{m}^{-6}]$	(2.2)	23
f^0	observational term of objective function	$[\text{kg}^2\cdot\text{m}^{-6}]$	(2.3)	23

Latin Symbol	Description	Units	First Use Eq.	Pg.
$f_{\min}^o$	estimated minimum of the observational term of the objective function	$[\text{kg}^2 \cdot \text{m}^{-6}]$	(-)	110
$f_{\min,10\% \text{ noise}}^o$	estimated minimum of the observational term of the objective function with 10% observational error	$[\text{kg}^2 \cdot \text{m}^{-6}]$	(-)	109
$f^r(\mathbf{x})$	regularization term of the objective function	$[\text{kg}^2 \cdot \text{m}^{-6}]$	(2.4)	23
$\tilde{f}_k$	Taylor series approximation of the objective function (f), by the k th iteration of the gradient-based numerical optimization algorithm	$[\text{kg}^2 \cdot \text{m}^{-6}]$	(2.10)	28
$\mathbf{f}_{\Gamma}^{(e)}$	local finite element boundary condition vector	$[\text{kg} \cdot \text{s}^{-1}]$	(3.34)	61
$\mathbf{F}$	flux		(-)	50
f_{actr}	L-BFGS-B relative reduction of the objective function termination threshold ($f_{actr} = 0$)	[-]	(4.1)	82
g	arbitrary continuous objective function term	$[\text{kg}^2 \cdot \text{s}^{-1} \cdot \text{m}^{-9}]$	(A.1)	165
H	building height	[m]	(-)	69
$\bar{H}$	average building height	[m]	(-)	69
H	continuous observation function	$[\text{m}^{-3} \cdot \text{s}^{-1}]$	(A.3)	166
$\mathbf{H}$	observation matrix	[-]	(-)	6
$\mathbf{H}_k$	approximation of the inverse Hessian matrix	$[\text{s}^2]$	(2.12)	30
$\mathbf{H}_k^0$	initial approximation of the inverse Hessian matrix	$[\text{s}^2]$	(2.12)	30
$\mathbf{I}$	identity matrix	[-]	(1.5)	8
k	iteration number	[-]	(1.1)	5
k	turbulence kinetic energy per unit mass	$[\text{m}^2 \cdot \text{s}^{-2}]$	(3.6)	47
$\mathbf{K}^{(e)}$	local finite element stiffness matrix	$[\text{m}^3 \cdot \text{s}^{-1}]$	(3.28)	57
$\bar{\mathbf{K}}$	modified global finite element stiffness matrix	$[\text{m}^3 \cdot \text{s}^{-1}]$	(3.37)	62
$\bar{\mathbf{K}}^{(e)}$	modified local finite element stiffness matrix	$[\text{m}^3 \cdot \text{s}^{-1}]$	(3.36)	61
$\mathbf{K}_k$	Kalman filter gain matrix	$[\text{s}^{-1}]$	(1.3)	8

Latin Symbol	Description	Units	First Use Eq.	Pg.
L	characteristic length	[m]	(3.32)	59
$\mathbf{lb}$	L-BFGS-B parameter for the lower bounds of the parameter search space ($\mathbf{lb}$ = zero-vector)	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(-)	81
$\mathbf{M}$	dispersion model for Kalman filter	[s]	(1.3)	8
m	number of iterate updates used by the L-BFGS-B numerical optimization algorithms to approximate the inverse Hessian matrix of the objective function ($m = 5$)	[-]	(-)	30
$maxit$	maximum number of optimization iterations used by the L-BFGS-B algorithm ($maxit = 200$)	[-]	(-)	84
n	time-step number	[-]	(-)	63
$\hat{\mathbf{n}}$	outward normal unit vector of the domain boundary	[-]	(3.3)	45
$n_{\mathbf{c}}$	length of vector of discretized global pollutant concentration ($\mathbf{c}$) ($n_{\mathbf{c}} = n_{\mathbf{x}} \cdot n_t = 82,618,997$)	[-]	(-)	24
$n_{\mathbf{c}^o}$	length of vector of known receptor observations ($\mathbf{c}^o$) ($n_{\mathbf{c}^o} = n_{\mathbf{r}} \cdot n_t = 22,525$)	[-]	(-)	24
n_e	number of elements in mesh ($n_e = 174,772$)	[-]	(4.7)	97
n_f	number of objective function (f) outputs ($n_f = 1$)	[-]	(-)	22
n_r	number of receptors ($n_r = 25$)	[-]	(A.3)	166
n_t	number of time-steps ($n_t = 901$)	[-]	(3.45)	66
$n_{\mathbf{x}}$	number of emission source parameters ($\mathbf{x}$) ($n_{\mathbf{x}} = 91,697$)	[-]	(-)	3
$\mathbf{nbd}$	L-BFGS-B parameter for the boundary type of the parameter search space ($\mathbf{nbd}$ = one-vector)	[-]	(-)	81
N_i	finite element shape function	[-]	(-)	53
$\mathbf{N}$	column vector of all finite element shape functions for a single element	[-]	(3.12)	54
$\mathbf{P}_k$	Kalman filter error-covariance matrix	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(1.2)	7

Latin Symbol	Description	Units	First Use Eq.	Pg.
$\mathbf{P}_{k+1}$	Kalman filter corrected future error-covariance matrix	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(1.5)	8
$\mathbf{p}_k$	Search direction of a gradient-based numerical optimization algorithm	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(2.9)	28
$p_k(\mathbf{c}^0 \mathbf{x})$	Bayesian joint likelihood function	[-]	(1.1)	5
$p_k(\mathbf{x})$	Bayesian prior joint probability function	[-]	(1.1)	5
$p_k(\mathbf{x} \mathbf{c}^0)$	Bayesian posterior joint probability function	[-]	(1.1)	5
p_{FOV}	Receptor network field of view probability density function	[-]	(4.5)	93
Pe_L	Péclet number	[-]	(3.32)	59
$pgtol$	projected gradient tolerance for the L-BFGS-B numerical optimization algorithm ($pgtol = 0$)	$[\text{kg}\cdot\text{s}\cdot\text{m}^{-3}]$	(4.2)	83
Q	total emission rate	$[\text{kg}\cdot\text{s}^{-1}]$	(4.7)	97
Q^*	known total emission rate	$[\text{kg}\cdot\text{s}^{-1}]$	(-)	87
$\mathbf{R}$	error-covariance matrix of receptor observations		(1.4)	8
$\mathbf{r}$	coordinate position vector	[m]	(-)	20
r^2	coefficient of determination	[-]	(4.6)	97
$r_{\text{max},10\% \text{ noise}}^2$	estimated maximum coefficient of determination with 10% observational noise	[-]	(-)	112
Re_L	Reynolds number	[-]	(3.32)	59
s	continuous volumetric emission rate	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(-)	20
s^b	continuous background volumetric emission rate	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(2.5)	24
$\mathbf{s}$	discretized global volumetric emission rate	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(-)	21
$\mathbf{s}^{(e)}$	discretized local volumetric emission rate	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(3.19)	56
$\mathbf{s}^b$	discretized background volumetric emission rate ($\mathbf{s}^b = \text{zero-vector}$)	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(2.2)	23
$\mathbf{s}_k$	change in candidate solution between L-BFGS-B iterations	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(2.12)	30

Latin Symbol	Description	Units	First Use Eq.	Pg.
Sc	Schmidt number	[-]	(3.7)	47
Sc_t	turbulent Schmidt number ($Sc_t = 0.9$)	[-]	(3.8)	47
t	time	[s]	(-)	20
t_0	initial time ($t_0 = 0$ s)	[s]	(-)	43
t_f	final time ($t_f = 900$ s)	[s]	(-)	43
t_n	time of the n th time-step	[s]	(-)	63
t_i^r	i th sample time of the receptor network	[s]	(A.3)	166
Δt	time-step size ($\Delta t = 1$ s)	[s]	(-)	63
$\mathbf{u}$	flow velocity vector field	[m·s ⁻¹]	(3.1)	43
$\mathbf{U}^{(e)}$	matrix of nodal flow velocities of a single finite element	[m·s ⁻¹]	(3.25)	56
U	magnitude of flow velocity	[m·s ⁻¹]	(3.32)	59
$\mathbf{ub}$	L-BFGS-B parameter for the upper bounds of the parameter search space	[kg·s ⁻¹ ·m ⁻³]	(-)	81
u, v, w	flow velocity in x, y and z Cartesian directions	[m·s ⁻¹]	(-)	43
$u_i^{(e)}, v_i^{(e)}, w_i^{(e)}$	flow velocity in x, y and z Cartesian directions at the i th node of an element	[m·s ⁻¹]	(3.23)	56
$\mathbf{V}_k$	intermediate matrix for the L-BGFS inverse Hessian matrix update	[-]	(2.12)	30
W_i	finite element weight function	[-]	(3.10)	52
$\mathbf{W}$	column vector of finite element weight functions	[-]	(3.26)	57
x, y, z	Cartesian spatial coordinates	[m]	(-)	20
x_j^r, y_j^r, z_j^r	Cartesian spatial coordinates of the j th receptor	[m]	(A.3)	166
x_i	i th emission source parameter	[kg·s ⁻¹ ·m ⁻³]	(-)	20
$\mathbf{x}$	column vector of emission source parameters	[kg·s ⁻¹ ·m ⁻³]	(-)	3
$\mathbf{x}_0$	initial candidate emission source parameters ($\mathbf{x}_0$ = zero-vector)	[kg·s ⁻¹ ·m ⁻³]	(-)	24

Latin Symbol	Description	Units	First Use Eq.	Pg.
$\mathbf{x}_k$	iterate of emission source parameters	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(-)	26
$\mathbf{x}^*$	optimal volumetric source emission rate parameters	$[\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}]$	(-)	22
$\mathbf{y}_k$	change in objective function gradient between iterations	$[\text{kg}\cdot\text{s}\cdot\text{m}^{-3}]$	(-)	30
$\mathbf{0}$	zero-matrix	$[\text{m}^3\cdot\text{s}^{-1}]$	(3.46)	66

Greek Symbol	Description	Units	First Use Eq.	Pg.
α	finite element weight function upwinding parameter	[-]	(-)	59
α_k	step size of the search direction of a gradient-based numerical optimization algorithm	[-]	(2.9)	28
β	objective function regularization parameter	$[\text{s}^2\cdot\text{m}^{-3}]$	(2.2)	23
Γ	boundary of spatial domain (Ω)	$[\text{m}^2]$	(-)	43
Γ_{inlet}	inlet boundary of spatial domain (Ω)	$[\text{m}^2]$	(-)	43
Γ_{outlet}	outlet boundary of spatial domain (Ω)	$[\text{m}^2]$	(-)	43
Γ_{wall}	wall boundary of spatial domain (Ω)	$[\text{m}^2]$	(-)	43
$\Gamma^{(e)}$	boundary of a single finite element ($\Omega^{(e)}$)	$[\text{m}^2]$	(-)	49
$\Gamma_{\text{inlet}}^{(e)}$	inlet boundary of a single finite element ($\Omega^{(e)}$)	$[\text{m}^2]$	(-)	60
$\Gamma_{\text{outlet}}^{(e)}$	outlet boundary of a single finite element ($\Omega^{(e)}$)	$[\text{m}^2]$	(-)	60
$\Gamma_{\text{wall}}^{(e)}$	wall boundary of a single finite element ($\Omega^{(e)}$)	$[\text{m}^2]$	(-)	60
γ_i	intermediate parameter computed by the L-BFGS algorithm	[-]	(-)	30
δ	Dirac delta function	[-]	(A.4)	167
ε	turbulence dissipation rate	$[\text{m}^2\cdot\text{s}^{-3}]$	(3.6)	47

Greek Symbol	Description	Units	First Use Eq.	Pg.
θ	θ -method weight ($0 \leq \theta \leq 1$)	[-]	(3.38)	64
Λ	alternate objective function or Lagrange function	$[\text{kg}^2 \cdot \text{m}^{-6}]$	(2.23)	38
λ	continuous adjoint variable	$[\text{kg} \cdot \text{m}^{-6}]$	(A.16)	171
λ	discrete adjoint variable	$[\text{kg} \cdot \text{s} \cdot \text{m}^{-6}]$	(2.23)	38
μ_k	mean of multivariate Gaussian distribution for the Kalman filter	$[\text{kg} \cdot \text{s}^{-1} \cdot \text{m}^{-3}]$	(1.2)	7
ν	kinematic viscosity	$[\text{m}^2 \cdot \text{s}^{-1}]$	(3.7)	47
ν_t	turbulent eddy kinematic viscosity	$[\text{m}^2 \cdot \text{s}^{-1}]$	(3.6)	47
ρ_k	intermediate parameter computed by the L-BFGS algorithm	$[\text{m}^6 \cdot \text{kg}^{-2}]$	(2.12)	30
σ	standard deviation of a Gaussian distribution	$[\text{kg} \cdot \text{m}^{-3}]$	(-)	90
φ	arbitrary scalar function of space ($\mathbf{r}$) and time (t)	[-]	(A.18)	171
ψ	arbitrary scalar function of space ($\mathbf{r}$) and time (t)	[-]	(A.18)	171
Ψ	arbitrary vector function of space ($\mathbf{r}$) and time (t)	[-]	(A.19)	171
Ω	space domain	$[\text{m}^3]$	(-)	43
Ω_t	finite airshed space and time domain	$[\text{m}^3]$	(-)	43
$\Omega^{(e)}$	space domain of a single finite element	$[\text{m}^3]$	(-)	49

Subscript & Superscript	Description
$(\)^b$	background
$(\)^{(e)}$	elemental quantity
$(\)_i$	index number of vector element
$(\)_{\text{inlet}}$	inlet quantity
$(\)_k$	iteration number
$(\)_n$	time-step number

Subscripts & Superscripts	Description
$()^o$	observation or observational quantity
$()_{\text{outlet}}$	outlet quantity
$()_{\text{proj}}$	projected
$()^r$	receptor quantity
$()_t$	turbulent quantity
$()_{\text{wall}}$	wall quantity
$()^*$	optimal or correct
$\widehat{()}$	unit vector
$\overline{()}$	average or modified quantity
$()'$	deviation from the average

Operator	Description
$() \cdot ()$	dot product
$()^{-1}$	inverse
$()^T$	transpose
$ () $	determinant or absolute value
$\ () \ _{\infty}$	maximum norm
$\frac{d()}{dt}$ or $(\dot{ })$	ordinary time derivative
$\frac{\partial ()}{\partial t}$	partial time derivative
$\nabla ()$	gradient
$\nabla_{\mathbf{x}} ()$	gradient with respect to the parameters of vector $\mathbf{x}$
$\nabla \cdot ()$	divergence
$\nabla^2 ()$	Hessian matrix

Acronym	Description	First Use
AD	automatic differentiation	33
ADE	advection-diffusion equation	43
AERMOD	atmospheric dispersion modeling system	42
AMR	automatic mesh refinement	149
BFGS	Broyden-Fletcher-Goldfarb-Shanno numerical optimization algorithm	13
BVP	boundary value problem	49
CALPUFF	commercial Gaussian puff dispersion model	42
CAPP	Canadian Association of Petroleum Producers	1
CAS	computer algebra system	32
CDF	cumulative distribution function	178
CMAQ	Community Multiscale Air Quality modelling system	42
CFD	computational fluid dynamics	45
DI&M	directed inspection and maintenance	1
DFP	Davidon-Fletcher-Powell numerical optimization algorithm	13
EPA	United States Environmental Protection Agency	1
FAHM	full area at half maximum	101
FDM	finite difference method	49
FEM	finite element method	51
FOV	field of view	92
FVM	finite volume method	50
FLEXTRA	a Lagrangian particle dispersion model	3
FLEXPART	a Lagrangian particle dispersion model	3
GA	genetic algorithm	10
GHG	greenhouse gas	1
IDM	inverse dispersion model	1
IVP	initial value problem	44

Acronym	Description	First Use
L-BFGS	limited memory BFGS numerical optimization algorithm	13
L-BFGS-B	bound constrained L-BFGS numerical optimization algorithm.....	13
LPDM	Lagrangian particle dispersion model	3
MUST	Mock Urban Setting Test	7
ODE	ordinary differential equation	49
OWL-QN	orthant-wise limited memory quasi-Newton.....	29
PDE	partial differential equation.....	14
PDF	probability density function	5
PI-ORS	path-integrated optical remote sensing	1
RANS	Reynolds-averaged Navier-Stokes	47
RMS	root mean square	77
RNG	Re-Normalization Group	47
SA	simulated annealing	11
SR1	symmetric rank-one	13
SUPG	streamline-upwind Petrov-Galerkin	51
TDLAS	tunable diode absorption spectroscopy	90
TLM	tangent linear model.....	34
URANS	unsteady RANS.....	70
VOC	volatile organic compound.....	1

Chapter 1

Introduction

Fugitive emissions are unintended gas emissions from gas handling equipment (Kirchgessner et al., 1997). In the upstream oil and gas industry, fugitive methane emissions are a significant source of greenhouse gas (GHG) emissions (CAPP, 2004) that could be easily mitigated. However, the location and quantification of these emissions is laborious and expensive (CAPP, 2007). Typically, fugitive emissions are discovered and quantified manually as part of an emissions inventory program (Picard et al., 1998) with a portable volatile organic compound (VOC) analyzer (EPA, 1995). Nevertheless, qualitative methods such as bubble tests, infrared imaging and ultrasonic leak detectors are often used to indicate leaking equipment (CAPP, 2007). Additionally, optical remote sensing techniques (EPA, 2006) such as fence-line measurements using path-integrated optical remote sensing (PI-ORS) have also been used to monitor fugitive emissions (Hashmonay et al., 2001).

Typically, fugitive emissions inventories are performed intermittently as part of a directed inspection and maintenance (DI&M) practice (CAPP, 2007). In principle, a more effective DI&M program would include semi-continuous emissions monitoring by a network of optical remote sensors placed throughout the domain of an upstream oil and gas facility (Schoonbaert and Johnson, 2011). Observations from these receptors would be assimilated by an inverse micro-scale atmospheric dispersion model (IDM) to estimate

the locations and emission rates of emission sources that best explain the known receptor observations. This is an ill-posed problem with many candidate emission sources that could explain receptor observations (Newsam and Enting, 1988; Enting and Newsam, 1990; Mulholland and Seinfeld, 1995). For example a weak emission source near a receptor could produce the same observations as a stronger emission source further away (Thomson et al., 2007). However, the ill-posed quality of this inverse problem can be reduced with multiple observations from multiple receptors, under variable wind conditions (Thomson et al., 2007).

Data assimilation is a modelling technique that incorporates a limited number of observations of a system's state into a model of that system. With these observations, and knowledge of the system's physics, the state of the entire system can be estimated. Much of the mathematical theory of data assimilation has already been demonstrated in numerical weather prediction (Fisher et al., 2008) and hydrological forecasting (Ghil and Malanotte-Rizzoli, 1991) where the accurate estimation of the system's initial state is critical to the quality of forecasts. Emission source estimation through inverse dispersion modelling takes the requirements of the data assimilation theory further, and seeks to constrain the emission source parameters that produce a system state (e.g. pollutant concentration field) that best explains the observations (e.g. pollutant concentration measurements by receptors). Emission source estimation has been previously studied at macro (Houweling et al., 1999) and meso (Yumimoto and Uno, 2006) atmospheric scales, and there is no theoretical restriction on application of similar techniques at the micro atmospheric scales such as an upstream oil and gas processing facility. In fact, source

identification has already been successful in urban scale studies such as 2003 Oklahoma City Joint Urban field campaign (Chow et al., 2008).

1.1 Review of Inverse Dispersion Modelling

Generalized inverse problems have diverse applications in applied mathematics, with much overlap in their mathematical theory. Inverse modelling constrains the column vector $\mathbf{x}$ of n_x model parameters that are required to achieve a desired response from a model, such as the source emission rates required to explain pollutant concentration observations at downstream receptors. Inverse dispersion modelling methods can be separated into three categories: Lagrangian back-trajectory (Ashbaugh et al., 1985; Keeler and Samson, 1989; Stohl, 1996), sequential, and variational (Ide et al., 1997). The following sections are dedicated to a review of these methods for inverse dispersion modelling applications.

1.1.1 Lagrangian Back-Trajectory Methods

Lagrangian particle dispersion models (LPDMs) such as FLEXTRA (Stohl et al., 1995) and FLEXPART (Stohl et al., 1998) are able to establish source-receptor relationships by modelling the paths of fluid parcels through a finite airshed domain. These models can either be run forward in time to determine the regions of the airshed domain influenced by known emission sources, or backwards in time (back-trajectories) to determine the regions of the airshed domain that influence known receptor observations. These

methods are attractive because they are not susceptible to numerical diffusion like Eulerian models (Stohl et al., 2005).

Alone, a single pathline traced out by a fluid parcel in a Reynolds-averaged wind field represents a non-diffusive solution to the pollution transport equation (3.1). Unfortunately, this non-diffusive solution is unable to account for the significant diffusive effects from turbulent mixing within the atmospheric boundary layer (Stohl et al., 2002). For this reason, LPDMs model the diffusive effects of turbulence in a stochastic manner by evaluating pathlines for a large¹ number of fluid parcels, and randomly perturbing the flow velocity field in the vicinity of each parcel proportionately to the local turbulence intensity. This act of evaluating an ensemble of many back-trajectories is analogous to computing an Eulerian retroplume (Stohl et al., 2002; Issartel and Baverel, 2003).

The pollutant concentration measurement associated with each of the back-trajectories are used to infer the regions of the airshed domain that are more likely to contain emission sources. Brereton (2010) studied four back-trajectory methods for the estimation of emission source locations (potential source contribution function, concentration weighted trajectory, residence time weighted concentration and quantitative transport bias analysis) with a synthetic test case of fugitive emissions within an upstream gas processing facility.

¹ FLEXPART evaluates 40,000 pathlines for each release (Seibert and Frank, 2004; Cooper et al., 2010).

1.1.2 Sequential Data Assimilation

Sequential data assimilation methods “use a probabilistic framework and give estimates of the whole system state sequentially by propagating information only forward in time” (Bertion et al., 2003). These methods can be extended to inverse modelling applications for the estimation of the model parameters (e.g. emission sources) that produce a system state (e.g. pollutant concentration field), which explain the known observations (e.g. pollution concentration measurements by receptors). Sequential methods such as Bayesian inference (Lee, 1989) and Kalman filtering (Kalman, 1960) form a joint probability density function (PDF) of the model parameters $\mathbf{x}$ rather than a single point estimate. These joint PDFs offer a measure of the uncertainty within the estimation of the most likely values of the model parameters $\mathbf{x}$ for the given observations.

1.1.2.1 Bayesian Inference

Bayesian inference, or recursive Bayesian filtering is a sequential inverse modelling method based on recursive application of Bayes’ theorem (Dakins et al., 1996). When applied to inverse atmospheric dispersion modelling, this method iteratively estimates a posterior joint PDF of the emission source parameters $\mathbf{x}$ (e.g. source location and emission rate) that best explain the pollutant concentration observations $\mathbf{c}^o$ from receptors. The k th iteration of Bayesian inference for inverse atmospheric dispersion modelling is given by (1.1),

$$p_{k+1}(\mathbf{x}) = p_k(\mathbf{x}|\mathbf{c}^o) = \frac{p_k(\mathbf{x})p_k(\mathbf{c}^o|\mathbf{x})}{\int p_k(\mathbf{x})p_k(\mathbf{c}^o|\mathbf{x})d\mathbf{x}} \quad (1.1)$$

where $p_k(\mathbf{x})$ is the prior joint PDF of the candidate volumetric source emission rate parameters. $p_k(\mathbf{x}|\mathbf{c}^o)$ is the posterior joint PDF of the candidate emission source parameters. $p_k(\mathbf{c}^o|\mathbf{x})$ is the joint “likelihood function that shows the fitness of the calculated $[\mathbf{H}\mathbf{c}(\mathbf{x})]$ and observed concentrations $[\mathbf{c}^o]$ ” (Zheng and Chen, 2011), where $\mathbf{H}$ is an observation matrix that simulates receptor measurements from the modelled pollutant concentration field $\mathbf{c}(\mathbf{x})$. Finally, $\int p_k(\mathbf{x})p_k(\mathbf{c}^o|\mathbf{x})d\mathbf{x}$ is the marginal likelihood function that normalizes the posterior joint distribution such that it represents a probability density. Evaluation of the marginal likelihood function does not always have a closed form solution, and numerical integration techniques often are required.

Typically, the initial prior joint PDF ($p_0(\mathbf{x})$) is assumed to be a uniform distribution (Qian et al., 2003). Iterative application of (1.1) with the previous iteration’s posterior joint PDF assumed to be the prior joint PDF of the next iteration ($p_{k+1}(\mathbf{x}) = p_k(\mathbf{x}|\mathbf{c}^o)$) improves the estimate of the joint PDF of the model parameters $\mathbf{x}$. Convergence is achieved because the likelihood function ($p_k(\mathbf{c}^o|\mathbf{x})$) assigns probability proportional to the fitness of the model response from a sample of $\mathbf{x}$ to the observations ($\mathbf{c}^o$) (Bergin and Milford, 2000). The posterior joint PDF ($p_k(\mathbf{x}|\mathbf{c}^o)$) of a given iteration is typically approximated with a Monte Carlo simulation (Bergin and Milford, 2000) with Markov-chain sampling from the prior joint PDF ($p_k(\mathbf{x})$) and likelihood function ($p_k(\mathbf{c}^o|\mathbf{x})$).

Bayesian inference techniques have been widely applied to emission source inversion problems. Sohn et al. (2003) presented a technique to respond to a sudden pollutant release in an office building. A Monte Carlo method was used to quantify the effects of uncertainty in the modelling parameters on the “expected reduction in uncertainty in end-

point predictions resulting from increment information about the ... building parameters or release details” (Sohn et al., 2003). Bergin and Milford (2000) used Monte Carlo Bayesian inference to estimate the most likely values and the uncertainties of 51 model parameters on observed ozone and other secondary compound concentrations in response to a 25% reduction in specific motor vehicle emissions. Keats et al. (2007) used Bayesian inference to estimate the most likely location, emission rate, and start/end times of a single emission source and the respective uncertainties of the parameters for the Mock Urban Setting Test (MUST) and the Joint Urban 2003 atmospheric dispersion study. Chow et al. (2008) also successfully identified the emission source in Joint Urban 2003 field campaign with Bayesian inference.

1.1.2.2 Kalman Filter

The Kalman filter (Kalman, 1960) is another sequential method that is used for either data assimilation to estimate the state of a system (Lyster et al., 1997) or for inverse modelling to estimate model parameters (Hartley and Prinn, 1993; Haas-Laursen et al., 1996). The scope of this discussion on the Kalman filter will be restricted to inverse modelling applications. Like Bayesian inference, the Kalman filter also provides a joint PDF of model parameters $\mathbf{x}$ in the form of a multivariate Gaussian distribution (1.2),

$$p_k(\mathbf{x}) = p(\mathbf{x}; \boldsymbol{\mu}_k, \mathbf{P}_k) = (2\pi)^{-\frac{n_x}{2}} |\mathbf{P}_k|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_k)^T \mathbf{P}_k^{-1} (\mathbf{x} - \boldsymbol{\mu}_k)\right) \quad (1.2)$$

where $\boldsymbol{\mu}_k$ and $\mathbf{P}_k$ are the mean column vector and covariance matrix, respectively of the estimated multivariate Gaussian distribution of the model parameters $\mathbf{x}$ for the k th

iteration of the Kalman filter. The Kalman filter assumes that a linear model $\mathbf{M}$ governs the mapping of candidate emission source parameters $\mathbf{x}$ to synthetic receptor observations, and that the error of both the observations and emission source parameters contains only white noise (Gaussian distributed). The mean vector $\boldsymbol{\mu}_k$ of the distribution of emission source parameters ($\mathbf{x}$) is estimated with the recursive formula (1.3),

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k + \mathbf{K}_k(\mathbf{c}^o - \mathbf{M}\boldsymbol{\mu}_k) \quad (1.3)$$

where $\mathbf{K}_k$ is the Kalman gain matrix given by (1.4), which maps the residual error between the known receptor observations $\mathbf{c}^o$ and the modelled mean receptor response $\mathbf{M}\boldsymbol{\mu}_k$ into an adjustment on the mean vector ($\boldsymbol{\mu}_k$) of the emission source parameters.

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{M}^T (\mathbf{M} \mathbf{P}_k \mathbf{M}^T + \mathbf{R})^{-1} \quad (1.4)$$

Additionally, $\mathbf{R}$ is the error-covariance matrix of the receptor observations $\mathbf{c}^o$, and the estimate of the error-covariance matrix $\mathbf{P}_k$ of the emission source parameters is updated for the next iteration of the Kalman filter with (1.5).

$$\mathbf{P}_{k+1} = (\mathbf{I} - \mathbf{K}_k \mathbf{M}) \mathbf{P}_k \quad (1.5)$$

Sequential application of (1.3) - (1.5) allow the estimate of the joint PDF of the emission source parameters to update and improve the estimate of the most likely values of $\mathbf{x}$ given the mean $\boldsymbol{\mu}_k$ and covariance $\mathbf{P}_k$.

Performing the update (1.5) on the covariance matrix of the emission source parameters “is the most difficult part of the Kalman filter method” (Rao, 2007) since it requires a matrix solution to a system of linear equations. This limitation restricts the size of problems that can currently be studied with a conventional Kalman filter. The ensemble

Kalman filter (Evensen, 1994) avoids the computational effort of determining the covariance matrix of the posterior multivariate Gaussian distribution by instead approximating the posterior PDF with an ensemble simulation and computing the covariance matrix of the sample.

Another limitation of the standard Kalman filter is the restriction to linear state transition models ($\mathbf{M}$). Atmospheric transport mechanisms such as turbulence are non-linear in nature, “... leading to ill-conditioned covariance matrices and poorly observable systems” (Rao, 2007) that are better suited for extended Kalman filtering. The extended Kalman filter (Gelb, 1974) approximates the covariance matrix ($\mathbf{P}_{k+1}$) of the posterior joint PDF with an intermediate multivariate Taylor series expansion of the non-linear state transition and observation models.

Kalman filters have been successfully applied to environmental modelling by Mulholland and Seinfeld (1995) in a study to determine the required adjustments to the California Air Resources Board’s CO emissions inventory to better match observations.

1.1.3 Variational Data Assimilation

Variational data assimilation methods for inverse modelling formulate a scalar valued objective function $f(\mathbf{x})$ (also referred to as a loss, cost, or fitness function). In the context of inverse atmospheric dispersion modelling, this objective function measures the misfit between known, and modelled receptor observations resulting from candidate emission sources, parametrically defined by the emission source parameters ($\mathbf{x}$). Variational methods iteratively constrain the emission source parameters ($\mathbf{x}$) such that they are in the

best agreement with the known observations by minimizing the objective function ($f(\mathbf{x})$). Numerical optimization algorithms are used to generate improved candidate solutions. These algorithms can fall into two categories: global and local optimization, and each of these categories can be further subdivided into specific algorithms.

1.1.3.1 Global Optimization

Global numerical optimization algorithms attempt to constrain the model parameters ($\mathbf{x}$) to the global minimum of the objective function ($f(\mathbf{x})$). There are many families of global numerical optimization algorithms, but the genetic algorithm and simulated annealing are typically favored.

Genetic Algorithm

Genetic algorithms (GAs) were first described by Holland (1975). There are many alternate implementations, but they all share a common structure. A GA borrows concepts from evolutionary biology by treating the model parameters ($\mathbf{x}$) as genetic information or genes, and attempting to evolve the optimal set of model parameters through a process analogous to natural selection. GAs generate an initial population of candidate model parameters and evaluates their fitness with an objective function. The candidate parameters are stochastically selected proportionately to their fitness for recombination and random mutation to create the next generation of candidate model parameters. The stochastic sampling and random mutation helps a GA to reject local minimums and converge toward the global minimum of the objective function

(Cartwright and Harris, 1993). A hybrid GA and Nelder-Mead downhill simplex (Nelder and Mead, 1965) method was implemented by Allen et al. (2007) and Haupt et al. (2007) to constrain the location and emission rate of a synthetic emission source as well as the wind direction.

Simulated Annealing

Another popular global optimization algorithm is simulated annealing (SA), which was independently introduced by Kirkpatrick et al. (1983) and Černý (1985). This method draws an analogy to the annealing of metals, where in the slow cooling of a metal from a high initial temperature produces the optimal microstructure. SA uses the concept of a slowly reducing temperature to mimic the threshold on the probability of accepting worse candidate solutions. By stochastically accepting worse candidate solutions, SA is able to reject sub-optimal solutions in favour of improved candidates in future iterations. As the method progresses and the objective function's "temperature" is reduced, and worse candidate solutions are less likely to be accepted. In each iteration, a new candidate solution is generated from a modification of the previous candidate, until the problem's "temperature" has been reduced to zero. In theory, if the problem's "temperature" is reduced according to a sufficiently slow annealing schedule, the algorithm will converge to the global minimum of the objective function. Although for most applications the computational time required for a sufficiently slow annealing schedule is impractical. Thomson et al. (2007) performed emission source reconstruction with "a random search algorithm in conjunction with simulated annealing". Thomson et al. (2007) also studied

the robustness of their algorithm when observational data contained Gaussian noise or concentration offset biases, and the influence of objective function regularization.

1.1.3.2 Local Optimization

Local numerical optimization algorithms attempt to converge to a local minimum of an objective function in the neighborhood of an initial candidate solution, and as such are typically deterministic in nature. Local numerical optimization algorithms can fall into one of two categories: direct-search and gradient-based.

Direct-search

Direct-search (also known as derivative-free, black-box or 0th-order) optimization methods such as hill climbing (Nolle, 2006), coordinate descent (Nocedal and Wright, 2006), pattern search (Hooke and Jeeves, 1961), Nelder-Mead (Nelder and Mead, 1965), or Rosenbrock (Rosenbrock, 1960) only use value of the objective function. These methods are well suited for objective functions that are discontinuous and/or have few parameters.

Gradient-based

Alternatively, gradient-based optimization algorithms are more favourable for optimization problems with a continuous objective function and many parameters. These methods use local evaluations of the gradient (and sometimes Hessian) of the objective

function to generate improved candidate solutions. There are numerous implementations of gradient-based numerical optimization algorithms, but some notable examples are gradient descent (Cauchy, 1847), Newton's method (Nocedal and Wright, 2006), and quasi-Newton methods such as symmetric rank-one (SR1) (Conn et al., 1991), Davidon-Fletcher-Powell (DFP) (Davidon, 1991; Fletcher and Powell, 1963), Broyden-Fletcher-Goldfarb-Shanno (BFGS) (Fletcher, 1970) and their limited memory variants (Liu and Nocedal, 1989; Andrew and Gao, 2007). Nocedal and Wright (2006) provide an excellent review of these gradient-based numerical optimization algorithms.

For emission source characterization with many parameters the unbounded and bounded limited memory BFGS algorithms, L-BFGS (Liu and Nocedal, 1989) and L-BFGS-B (Byrd et al., 1994) respectively, are typically favored for their efficient performance (Zhang, 2007). The L-BFGS numerical optimization algorithm has been applied by Elbern et al. (2000) to locate and quantify synthetic emission sources.

1.2 Inverse Dispersion Modelling for Fugitive Emissions

1.2.1 Transport Modelling Method

The building-scale nature of fugitive emissions has considerable influence on the selection of a transport model suitable for resolving pollutant concentrations from candidate emission sources. Conventional atmospheric dispersion modelling is typically studied on scales ranging from local (~10 km) to global (Penenko et al., 2002). For

studies on these large scales, the influence of urban structures on the atmospheric flow field are typically not considered. These flow field simplifications are exploited by box (Ruby, 1987) and Gaussian dispersion (Sutton, 1932; Slade, 1968) models. However, this simplification is not acceptable for modelling fugitive emissions within the domain of a gas processing facility. Industrial gas processing domains typically span an area of $< 1 \text{ km}^2$ and the influence of urban structures on the flow field and turbulent mixing must be considered (Madala et al., 2012). The partial differential equations (PDEs) that govern pollution transport (3.1) must be solved numerically, and since the local arrangement of buildings typically occur in irregular patterns and orientations, unstructured grids are required to discretize the computational domain which presents a significant numerical modelling challenge. Finite volume and finite element methods are typically favoured for numerically solving the Navier-Stokes, turbulence, and pollution transport equations within an unstructured grid.

1.2.2 Emission Source Parameterization

The emission source function must be described by a fixed number of parameters, that the inverse dispersion model attempts to fit to known receptor observations. In situations where a limited number of emission sources can be assumed (e.g. accidental releases), then the parameters of interest are the location, emission rate and duration of the sources, as demonstrated by Allen et al. (2007) and Chow et al. (2008). However, in practice the number fugitive emission sources is unknown. To remain unbiased about the number of emission sources within the domain, a more suitable emission source function

parameterization is the volumetric emission rate (source emission rate per unit volume) from each and every node/element of the discretized domain used by the pollution transport model, as demonstrated by Hartley and Prinn (1993). This emission source parameterization was the chosen for the present work, resulting in a large number of parameters that the IDM would have to constrain. This choice of parameterization also constrains the type of inverse modelling method that can be used to iteratively generate candidate emission source distributions.

1.2.3 Inversion Method

While sequential inversion methods such as Bayesian inference or Kalman filtering provide an estimate of the uncertainty within their estimate of the model parameters, they can also have considerable computational cost for problems with many parameters (Zheng and Chen, 2010), making them intractable for fugitive emissions identification. Similarly, global optimization methods such as GAs and SAs are also ill-suited for applications with many parameters. Finally, Lagrangian back-trajectory methods were not considered because they were previously studied by Brereton (2010) for fugitive emissions applications.

Gradient-based local optimization methods were favoured because they provided an efficient means of constraining many emission source parameters and the gradient of a scalar valued objective function could be efficiently evaluated with the adjoint sensitivity method (see Section 2.4.3).

The adjoint method has previously been applied to pollution transport problems to determine the distribution of chemical species in the atmosphere (Elbern et al., 1997) and to determine the emissions sources required to explain known receptor observations (Elbern et al., 2000). The resulting sensitivities can then be used by a gradient descent numerical optimization algorithm to iteratively improve the candidate model input parameters $\mathbf{x}_{k+1}$ (e.g. volumetric source emission rates) by minimizing the objective function.

Of the gradient-based numerical optimization methods, the L-BFGS-B algorithm (Byrd et al., 1994) was selected for its ability to efficiently approximate the inverse Hessian matrix from a limited history of solution updates and enforce bound constraints on the emission rates to prevent the prediction of sinks (i.e. negative emission sources).

1.3 Thesis Overview and Objectives

A proof-of-concept inverse dispersion modelling study was conducted to locate and quantify fugitive emissions within a simplified two-dimensional model of a gas processing plant. Two synthetic test cases (a single-source, and a multiple (10) source) were studied, and the inverse dispersion model was tested under non-ideal conditions with the addition of observational noise. Full implementations of the inverse dispersion model for quantifying fugitive emissions in a full-scale gas plant would also require estimations of the atmospheric flow field from limited anemometer observations. However, this was beyond the scope of this current proof-of-concept work.

For the present study, a synthetic two-dimensional flow field was generated with computational fluid dynamics modelling of realistic wind conditions. A custom two-dimensional pollution transport model was developed to simulate the transient concentration field and receptor observations from emissions sources. Synthetic receptor observations were generated from known emission source distributions, and source reconstruction was attempted using only these receptor observations and the known flow field.

A scalar valued objective function was formed from the residual error between the known receptor observations and modelled receptor observations from candidate emission source. The distribution of emission sources that minimized the objective function were found with a gradient descent numerical optimization algorithm. The gradient of the objective function was efficiently evaluated with an adjoint sensitivity method. The use of the adjoint sensitivity method required the development of a custom two-dimensional adjoint sensitivity transport model.

The adjoint sensitivity approach to emission source inversion problems has previously been applied at macro atmospheric scales, including for the European Tracer Experiment (Robertson and Langner, 1998; Sofiev and Atlaskin, 2007) and the Comprehensive Nuclear-Test-Ban Treaty (Seibert, 2000; Wotawa et al., 2003). However, to the author's knowledge the application of the adjoint sensitivity method to the micro atmospheric scales of fugitive emissions from an industrial plant is novel.

1.4 Thesis Organization

The remainder of the thesis is organized as follows.

Chapter 2: Inverse Modelling - This chapter discusses the formulation of an objective function that measures the misfit between known and modelled receptor observations resulting from candidate emission sources. The gradient-based numerical optimization algorithm L-BFGS-B is presented to minimize the objective function, and the discrete adjoint sensitivity method is presented to efficiently evaluate the gradient of the objective function.

Chapter 3: Dispersion Modelling - This chapter presents the theory for modelling the atmospheric transport of an inert pollutant within a known flow field, by numerically solving the advection-diffusion equation with a finite element method.

Chapter 4: Inverse Dispersion Model Setup - This chapter presents the CFD modelling that was conducted to generate a synthetic two-dimensional flow field, as well as the single and multi-source inverse dispersion modelling test cases. Finally, this chapter also presents the criteria by which the quality of candidate emission sources were assessed.

Chapter 5: Inverse Dispersion Model Results - This chapter presents the resulting emission source distributions predicted by the inverse dispersion model for single and multi-source test cases. The robustness of the inverse dispersion model is assessed with the addition of observational noise. Observational noise caused an increase in the prediction of false emission sources, which produced a significant over estimate of the total emission rate of the predicted sources. Objective function regularization and

emission source filtering techniques were explored to mitigate these false emission sources.

Chapter 6: Conclusions and Recommendations - This chapter presents a final summary of the results and recommendations for future studies in this area.

Chapter 2

Inverse Modelling

Forward or direct modelling determines the response of a model (i.e. the outputs) to a given set of parameters (i.e. inputs) represented by the column vector $\mathbf{x}$ which contains n_x elements and has the structure $\mathbf{x} = [x_1 \ \cdots \ x_{n_x}]^T$. By contrast, inverse modelling solves the opposite problem, and determines the values of the model parameters $\mathbf{x}$ required to achieve a specific or optimum response from the model. Inverse atmospheric dispersion modelling determines the values of the emission source parameters $\mathbf{x}$ (such as emission rates and source locations) that define the spatial (Cartesian coordinate vector $\mathbf{r} = [x \ y \ z]^T$) and temporal (t) distribution of emission sources $s(\mathbf{r}, t; \mathbf{x})$ of a pollutant, and best explains known receptor observations $\mathbf{c}^o$ of the pollutant's concentration. For the presented studies, the parametrically defined emission source function $s(\mathbf{r}, t; \mathbf{x})$ was constrained to model steady volumetric emission rates (i.e. $s(\mathbf{r}; \mathbf{x}) = s(\mathbf{r}, t; \mathbf{x})$). To remain unbiased about the number of emission sources within a simplified model of an industrial gas processing facility, the emission source parameters $\mathbf{x}$ were chosen to represent a pollutant's volumetric emission rates at the nodes of a discretized airshed domain, resulting in $n_x = 91,697$. These emission rates were interpolated within the discretized model of the airshed domain with finite element shape functions (see Section 3.2.1.2).

Finally, synthetic receptor observations $\mathbf{c}^0$ were generated for known distributions of steady emission sources with a custom micro-scale dispersion model (see Chapter 3). In an envisioned future application, these known receptor observations $\mathbf{c}^0$ would be obtained from physical sensors placed throughout an industrial site of interest.

2.1 Dispersion Model

The solution to the time dependant PDE that models the transport of a pollutant (3.1) from emission sources parametrically defined by $\mathbf{x}$ in a finite airshed domain can be numerically approximated as the solution to a system of linear equations (2.1) known as a dispersion model (see Section 3.3).

$$\mathbf{A}\mathbf{c} = \mathbf{b} \quad (2.1)$$

Here $\mathbf{c} = \mathbf{c}(\mathbf{x})$ is a column vector of length n_c , representing a discretized approximation of the pollutant concentration field resulting from the candidate sources. $\mathbf{A}$ is a sparse, square coefficient matrix defined by the atmospheric flow field physics and domain discretization, and $\mathbf{b} = \mathbf{b}(\mathbf{x})$ is a load vector resulting from the discretized representation of the candidate emission sources $\mathbf{s}(\mathbf{x})$ and boundary conditions of the dispersion model. Chapter 3 provides an overview of a method to evaluate a coefficient matrix $\mathbf{A}$ and load vector $\mathbf{b}(\mathbf{x})$ appropriate for dispersion modelling applications.

The solution of (2.1) will give $\mathbf{c}(\mathbf{x})$ as an approximation of the spatial and temporal distributions of the pollutant's concentration resulting from the candidate emission sources $\mathbf{s}(\mathbf{x})$. Synthetic receptor observations can be extracted from this modelled

pollutant concentration field ($\mathbf{c}(\mathbf{x})$) as point samples (or line averages in the case of fence line monitoring) with an observation matrix $\mathbf{H}$. The residual error between the modelled ($\mathbf{H}\mathbf{c}(\mathbf{x})$) and known ($\mathbf{c}^0$) receptor observations gives insight into quality of the candidate emission sources ($\mathbf{s}(\mathbf{x})$). An objective function provides a method to weight and accumulate these residual errors so that quality of candidate emission sources ($\mathbf{s}(\mathbf{x})$) can be quantified.

2.2 Objective Function

An inverse modelling problem can be formulated as an optimization problem with a smooth, scalar valued ($n_f = 1$), objective function $f(\mathbf{x})$ in $\mathbb{R}^{n_x}$ (also known as a loss, cost or fitness function). The objective function measures the misfit between the known system response and a modelled response for a candidate set of model parameters $\mathbf{x}$, and is formulated such that its global minimizer $\mathbf{x}^*$ results in a model response that is closest to the desired response.

In the context of inverse atmospheric dispersion modelling, the objective function measures the misfit between known receptor observations $\mathbf{c}^0$ (a column vector of length n_{c^0}) and the synthetic receptor observations ($\mathbf{H}\mathbf{c}(\mathbf{x})$) from the modelled pollutant concentration field $\mathbf{c}(\mathbf{x})$ that result from candidate emission sources $\mathbf{s}(\mathbf{x})$. The values of the emission source parameters $\mathbf{x}$ that are the global minimizer of the objective function, define the distribution of emission sources that best explain the known receptor observations. A candidate objective function $f(\mathbf{x})$ suitable for inverse dispersion modelling is given by (2.2).

$$f(\mathbf{x}) = \frac{1}{2}(\mathbf{H}\mathbf{c}(\mathbf{x}) - \mathbf{c}^o)^T \mathbf{R}^{-1}(\mathbf{H}\mathbf{c}(\mathbf{x}) - \mathbf{c}^o) + \frac{\beta}{2}(\mathbf{s}(\mathbf{x}) - \mathbf{s}^b)^T \mathbf{B}^{-1}(\mathbf{s}(\mathbf{x}) - \mathbf{s}^b) \quad (2.2)$$

This objective function is a positive quadratic functional that is continuously differentiable and is the sum of an observational term ($f^o(\mathbf{x})$):

$$f^o(\mathbf{x}) = \frac{1}{2}(\mathbf{H}\mathbf{c}(\mathbf{x}) - \mathbf{c}^o)^T \mathbf{R}^{-1}(\mathbf{H}\mathbf{c}(\mathbf{x}) - \mathbf{c}^o) \quad (2.3)$$

and a regularization term ($f^r(\mathbf{x})$):

$$f^r(\mathbf{x}) = \frac{\beta}{2}(\mathbf{s}(\mathbf{x}) - \mathbf{s}^b)^T \mathbf{B}^{-1}(\mathbf{s}(\mathbf{x}) - \mathbf{s}^b) \quad (2.4)$$

A variant of this objective function was suggested by Ide et al. (1997) for the estimation of initial conditions for forecasting models. The observational term represents a weighted summation of the square of the residual error between the known receptor observations ($\mathbf{c}^o$) and the modelled receptor observations ($\mathbf{H}\mathbf{c}(\mathbf{x})$) from a candidate distributions of emission sources. Regularization improves the underdetermined or ill-conditioned quality of an inverse model, by augmenting the objective function with additional “soft” constraints (Giles and Pierce, 2000) such as minimization of variance of a candidate solution (distribution of emission sources, $\mathbf{s}(\mathbf{x})$) from an assumed prior, typically called the background ($\mathbf{s}^b$) (Marchuk, 1994; Seibert, 2001). In the present work, it was assumed that there were no known emission sources within the domain and the background emission source distribution ($\mathbf{s}^b$) was thus chosen to be the zero-vector. The IDM would allow for the detection of anomalies (i.e. the presence of unknown emission sources) that deviated from this assumption. Additionally, Ide et al. (1997)

recommend choosing the emission source background to match the initial estimate of emission source parameters $\mathbf{x}_0$ used by a numerical optimization algorithm such that $\mathbf{s}^b = \mathbf{s}(\mathbf{x}_0)$, but this is not a requirement.

The sparse observation matrix $\mathbf{H}$ has the dimension $n_{c^o} \times n_c$, and simulates receptor observations from the modelled concentration field $\mathbf{c}(\mathbf{x})$ for comparison with the known receptor observations $\mathbf{c}^o$. Each row of $\mathbf{H}$ represents a weighted average of the nodal concentrations that influence a single receptor observation from the model. In the presented studies the weights are derived from the finite element shape functions (see Section 3.2.1.2) of the element that contains the receptor of interest.

The positive-definite matrix $\mathbf{R}^{-1}$ is a weight matrix of dimension $n_{c^o} \times n_{c^o}$, with $\mathbf{R}$ given as the error-covariance matrix of the observed and modelled concentrations (Sandu et al., 2005). For the purposes of this research $\mathbf{R}^{-1}$ was chosen to be the identity matrix.

The positive-definite matrix “ $\mathbf{B}^{-1}$ is an a priori weight matrix, with $\mathbf{B}$ meant to approximate the error-covariance matrix” (Ide et al., 1997) of the prior estimate of the distribution of nodal source volumetric emission rates $\mathbf{s}^b$. For the presented studies, $\mathbf{B}^{-1}$ was chosen to be a sparse matrix that penalized the deviation of candidate emission sources $\mathbf{s}(\mathbf{x})$ proportionately to the size of the elements of the discretized airshed domain such that (2.5) is satisfied,

$$(\mathbf{s}(\mathbf{x}) - \mathbf{s}^b)^T \mathbf{B}^{-1} (\mathbf{s}(\mathbf{x}) - \mathbf{s}^b) = \int_{\Omega} \left(s(\mathbf{r}; \mathbf{x}) - s^b(\mathbf{r}) \right)^2 d\Omega \quad (2.5)$$

where $s(\mathbf{r}; \mathbf{x})$ and $s^b(\mathbf{r})$ are continuous representations of the volumetric emission rates of $\mathbf{s}(\mathbf{x})$ and $\mathbf{s}^b$ respectively, at the spatial coordinate $\mathbf{r}$.

The positive constant β is a regularization parameter that determines importance of minimizing the regularization term relative to the observational term. In this application β has the physical dimension $[s^2 \cdot m^{-3}]$, but its value is subjective. However, the magnitude of the regularization parameter required to achieve equal weighting of the observational and regularization terms can be estimated with (2.6),

$$\beta = \frac{(\mathbf{H}\mathbf{c}(\mathbf{x}_0) - \mathbf{c}^0)^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{c}(\mathbf{x}_0) - \mathbf{c}^0)}{(\mathbf{s}(\mathbf{x}^*) - \mathbf{s}^b)^T \mathbf{B}^{-1} (\mathbf{s}(\mathbf{x}^*) - \mathbf{s}^b)} \quad (2.6)$$

which is an adaptation of the regularization parameter introduced by Addepalli et al. (2009). However, this estimation is only possible if the correct emission source distribution $\mathbf{s}(\mathbf{x}^*)$ is known.

Finally., the 1/2 factors on both the observational and regularization terms of the objective function are convenient scaling factors that simplify the algebra of evaluating the gradient of the objective function with respect to the emission source parameters $\mathbf{x}$.

The values of the emission source parameters that minimize the objective function, and in turn best explain the known receptor observations ($\mathbf{c}^0$) can be found with a numerical optimization algorithm.

2.3 Numerical Optimization

The values of the emission source parameters $\mathbf{x}$ that minimize the objective function $f(\mathbf{x})$ are regarded as the optimum solution and are denoted as $\mathbf{x}^*$. This optimal solution parametrically defines the distribution of emission sources $\mathbf{s}(\mathbf{x}^*)$ that best explain the known receptor observations $\mathbf{c}^0$, and can be found with a numerical optimization

algorithm. Numerical optimization algorithms start with an initial candidate solution $\mathbf{x}_0$ “... and generate a sequence of improved estimates $[\mathbf{x}_k]$ (called ‘iterates’) until they terminate” (Nocedal and Wright, 2006). Upon termination, it is expected that the final iterate is an approximation of an objective function minimizer. There are generally two families of numerical optimization algorithms, global and local.

Global numerical optimization algorithms, such as genetic algorithms (GA) (Holland, 1975) and simulated annealing (SA) (Kirkpatrick et al., 1983; Černý, 1985), generate new iterates from combinations of previous iterates and random perturbations, with the intent that iterates which produce improvements are more likely to be selected for combination. Due to the random perturbations, global methods are not dependent on their initial candidate solution (Zheng and Chen, 2011) and can converge along different search paths for the same starting point. Additionally, iterates which produce increases in the objective function can be accepted to avoid convergence on local minimas in favour of the global minimizer (Cartwright and Harris, 1993). Finally, these methods do not require information about the gradient of the objective function, which means that they are well suited for combinatorial problems, discontinuous objective functions and discrete search spaces.

Alternatively, local numerical optimization algorithms are dependent on their initial candidate solution ($\mathbf{x}_0$) and will converge to a local minimum within the neighborhood of that starting point. Typically, these algorithms are run multiple times with different starting conditions ($\mathbf{x}_0$) to find better local minimizers (Kirkpatrick et al., 1983). Within the family of local numerical optimization algorithms there are generally two classes of methods, direct-search methods and gradient-based methods. Direct-search optimization

methods require only the value of the objective function and can accommodate objective functions that are not continuous, differentiable or in situations where evaluating the objective function's gradient is too computationally expensive. These methods use multiple local evaluations of the objective function to determine the direction of decent and converge to a local minimum. The required number of objective function evaluations per iteration scales with the number of optimization parameters n_x , making direct methods inefficient for optimization problems with many parameters ($n_x > 1000$ (Liu and Nocedal, 1989)). Alternatively, gradient-based methods improve convergence for large optimization problems by also using the gradient and sometimes the Hessian (curvature) of the objective function. Due to the large number of optimization parameters used in the presented studies a gradient-based numerical optimization algorithm was selected.

2.3.1 General Gradient-based Methods

Consider the scalar valued ($n_f = 1$) objective function $f(\mathbf{x})$ for which the gradient $\nabla f(\mathbf{x})$ (2.7) and Hessian $\nabla^2 f(\mathbf{x})$ (2.8) are available.

$$\nabla f(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x})}{\partial x_1} \quad \dots \quad \frac{\partial f(\mathbf{x})}{\partial x_{n_x}} \right]^T \quad (2.7)$$

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_{n_x}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_{n_x} \partial x_1} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_{n_x} \partial x_{n_x}} \end{bmatrix} \quad (2.8)$$

The objective function can be minimized with a gradient-based numerical optimization algorithm that generates a sequence of iterates $\mathbf{x}_k$, which converge towards a local minimum. For the current application, $\mathbf{x}_k$ represents a candidate distribution of volumetric emission rates from fixed locations within the domain. A method for generating the next iterate ($\mathbf{x}_{k+1}$) in this sequence is (2.9),

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \quad (2.9)$$

where $\mathbf{p}_k$ is a line search direction that will locally reduce the objective function $f(\mathbf{x}_k)$, and the step size α_k must be appropriately chosen with a line search method (Nocedal and Wright, 2006). There are a few suitable methods for determining a line search direction $\mathbf{p}_k$, such as the steepest decent (Nocedal and Wright, 2006) (also known as gradient descent (Cauchy, 1847)) in which $\mathbf{p}_k = -\nabla f(\mathbf{x}_k)$. This choice of line search direction can be favourable because it does not require knowledge of the Hessian matrix, which is often too computationally expensive to evaluate and store for an objective function with many parameters. However, this line search direction can have slow convergence (Nocedal and Wright, 2006), and does not give a sense of an appropriate step size α_k to initialize a line search. The Newton direction is another popular line search direction, which is derived from a second-order Taylor series expansion $\tilde{f}_k(\mathbf{p}_k)$ of the objective function about $\mathbf{x}_k$ (2.10).

$$\tilde{f}_k(\mathbf{p}_k) = f(\mathbf{x}_k) + \mathbf{p}_k^T \nabla f(\mathbf{x}_k) + \frac{1}{2} \mathbf{p}_k^T \nabla^2 f(\mathbf{x}_k) \mathbf{p}_k \quad (2.10)$$

The minimum of this quadratic approximation (2.10) of the objective function can be located by differentiating $\tilde{f}_k(\mathbf{p}_k)$ with respect to the line search direction ($\mathbf{p}_k$), and

setting the resulting gradient ($\nabla \tilde{f}_k(\mathbf{p}_k)$) to the zero-vector to obtain the system of linear equations given by (2.11).

$$\mathbf{p}_k = -(\nabla^2 f(\mathbf{x}_k))^{-1} \nabla f(\mathbf{x}_k) \quad (2.11)$$

Typically, the Newton line search direction (2.11) is scaled such that $\alpha_k = 1$ is an appropriate step size. However, if necessary the step size can be adjusted with a line search.

Unfortunately, the Newton line search direction (2.11) requires the evaluation and storage of the Hessian matrix or its inverse. A class of algorithms known as limited memory quasi-Newton methods, such as the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) (Liu and Nocedal, 1989), the bound constrained L-BFGS (L-BFGS-B) (Byrd et al., 1994), or the orthant-wise limited memory quasi-Newton (OWL-QN) (Andrew and Gao, 2007) approximate the inverse Hessian matrix from a limited history of iterate updates. For the presented studies the L-BFGS-B numerical optimization algorithm was selected to avoid the computational expense of evaluating and storing a large Hessian matrix.

2.3.2 L-BFGS-B Method

The L-BFGS-B (Byrd et al., 1994) numerical optimization algorithm is the bound constrained extension to the limited memory BFGS (L-BFGS) method (Nocedal, 1980). L-BFGS-B also supports unconstrained optimization, has similar performance to its predecessor with a small increase to its memory requirements, and “could be considered

to supersede L-BFGS” (Zhu et al., 1997). The formal description of the L-BFGS algorithm is presented by Nocedal (1980), and Nocedal and Wright (2006) and performance benchmark tests are presented by Liu and Nocedal (1989) and Gilbert and Lemaréchal (1989). The L-BFGS algorithm approximates the inverse Hessian matrix $\mathbf{H}_k \cong (\nabla^2 f(\mathbf{x}_k))^{-1}$ from a limited history of the m most recent updates to the correction pairs $\mathbf{s}_k = \mathbf{x}_k - \mathbf{x}_{k-1}$ and $\mathbf{y}_k = \nabla f(\mathbf{x}_k) - \nabla f(\mathbf{x}_{k-1})$, with a variant of the BFGS update formula (2.12) (Nocedal and Wright, 2006),

$$\begin{aligned} \mathbf{H}_k = & (\mathbf{V}_{k-1}^T \cdots \mathbf{V}_{k-m}^T) \mathbf{H}_k^0 (\mathbf{V}_{k-m}^T \cdots \mathbf{V}_{k-1}^T) \\ & + \rho_{k-m} (\mathbf{V}_{k-1}^T \cdots \mathbf{V}_{k-m+1}^T) \mathbf{s}_{k-m} \mathbf{s}_{k-m}^T (\mathbf{V}_{k-m+2}^T \cdots \mathbf{V}_{k-1}^T) \\ & + \rho_{k-m+1} (\mathbf{V}_{k-1}^T \cdots \mathbf{V}_{k-m+2}^T) \mathbf{s}_{k-m+1} \mathbf{s}_{k-m+1}^T (\mathbf{V}_{k-m+2}^T \cdots \mathbf{V}_{k-1}^T) \\ & \cdots \\ & + \rho_{k-1} \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T \end{aligned} \quad (2.12)$$

where $\mathbf{V}_k = \mathbf{I} - \rho_k \mathbf{y}_k \mathbf{s}_k^T$ and $\rho_k = (\mathbf{y}_k^T \mathbf{s}_k)^{-1}$. It should be noted that there is little interest in evaluating this approximation of the inverse Hessian matrix ($\mathbf{H}_k$). Instead, the interest is in determining the line search direction ($\mathbf{p}_k$), given as $\mathbf{p}_k = -\mathbf{H}_k \nabla f(\mathbf{x}_k)$. This matrix-vector product can be efficiently evaluated with a two-loop recursion algorithm (Nocedal, 1980) as follows.

L-BFGS Line Search Direction Algorithm (Adapted from (Nocedal and Wright, 2006))

```

 $\mathbf{p}_k \leftarrow -\nabla f(\mathbf{x}_k);$ 
for  $i = k-1, k-2, \dots, k-m$ 
     $\gamma_i \leftarrow \rho_i \mathbf{s}_i^T \mathbf{p}_k;$ 
     $\mathbf{p}_k \leftarrow \mathbf{p}_k - \gamma_i \mathbf{y}_i;$ 
end (for)
 $\mathbf{p}_k \leftarrow \mathbf{H}_k^0 \mathbf{p}_k;$ 
for  $i = k-m, k-m+1, \dots, k-1$ 
     $\beta \leftarrow \rho_i \mathbf{y}_i^T \mathbf{p}_k;$ 
     $\mathbf{p}_k \leftarrow \mathbf{p}_k + (\gamma_i - \beta) \mathbf{s}_i;$ 
end (for)
stop with  $\mathbf{p}_k = -\mathbf{H}_k \nabla f(\mathbf{x}_k);$ 

```

Nocedal and Wright (2006) recommend choosing $\mathbf{H}_k^0 = (\mathbf{s}_{k-1}^T \mathbf{y}_{k-1})(\mathbf{y}_{k-1}^T \mathbf{y}_{k-1})^{-1} \mathbf{I}$ as the initial approximation of the inverse Hessian matrix $\mathbf{H}_k$. This choice produces an inverse Hessian matrix approximation that is of a similar scale to the true inverse Hessian matrix. The full L-BFGS algorithm is as follows.

L-BFGS Algorithm (Adapted from (Nocedal and Wright, 2006))

```

choose initial candidate solution  $\mathbf{x}_0$ ;
choose number of update histories  $m > 0$ ;
 $k \leftarrow 0$ ;
 $\mathbf{s}_0 \leftarrow -\nabla f(\mathbf{x}_0)$ ;
 $\mathbf{y}_0 \leftarrow \nabla f(\mathbf{x}_0 + \mathbf{s}_0) - \nabla f(\mathbf{x}_0)$ ;
repeat
     $\mathbf{H}_k^0 \leftarrow (\mathbf{s}_{k-1}^T \mathbf{y}_{k-1})(\mathbf{y}_{k-1}^T \mathbf{y}_{k-1})^{-1} \mathbf{I}$ ;
    compute  $\mathbf{p}_k$  with the L-BFGS Line Search Direction Algorithm;
    determine an appropriate  $\alpha_k$  with a line search;
    if  $k > m$ 
        discard the vector pair  $\{\mathbf{s}_{k-m}, \mathbf{y}_{k-m}\}$  from storage;
    end (if)
     $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$ ;
     $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ ;
     $\mathbf{y}_k = \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)$ ;
     $k \leftarrow k + 1$ ;
until convergence.
stop with  $\mathbf{x}_k$  as the minimizer of  $f(\mathbf{x})$ ;

```

“The main weakness of the L-BFGS method is that it converges slowly on ill-conditioned problems-specifically, on problems where the Hessian matrix contains a wide distribution of eigenvalues” (Nocedal and Wright, 2006).

For the presented studies the L-BFGS-B algorithm was used to constrain the lower bound of the emission rate of candidate emission source to zero. This constraint prevented the prediction of unphysical negative emission sources (i.e. sinks). Additionally, the L-BFGS-B algorithm has good performance on large scale optimization problems with

limited storage. However, the evaluation of the gradient of an objective function with many parameters can require considerable effort.

2.4 Objective Function Gradient

The gradient of the scalar valued objective function $f(\mathbf{x})$ is a column vector of partial derivatives of the objective function with respect to each of the $n_{\mathbf{x}}$ elements of $\mathbf{x}$ (2.7). In the context of inverse dispersion modelling, the gradient is a measure of the sensitivity of the objective function to each of the emission source parameters. To locally reduce the objective function and achieve a model response that is a closer match to the known observations, the emission source parameters must be adjusted in to opposition direction of the gradient. Unfortunately, evaluating the gradient of objective function is not a trivial task, and can require considerable computational effort for objective functions with many parameters.

2.4.1 Review of Differentiation Techniques

There are four families of methods to determine the gradient of a function: symbolic differentiation, finite difference methods (and their complex variable variants), automatic differentiation, and manual differentiation (Bischof et al., 1994).

Symbolic differentiation uses a computer algebra system (CAS) such as Maple, Macsyma, or Reduce to apply the rules of differentiation to the elementary functions that form the function of interest. However, these methods are unable to accommodate

common programming constructs such as branches, loops or subroutines, requiring considerable effort to reduce an existing code into manageable pieces for a symbolic manipulator (Bischof et al., 1994).

Alternatively, the finite difference method (Press et al., 1997) (also known as divided difference method) treats the function as a “black-box” and approximates its derivatives with secants, by observing the function’s changes to the independent perturbation of each of the function’s parameters. This method can require considerable computational effort for functions with many parameters. Additionally, this method is prone to truncation errors if the perturbation size is too large, or round-off/cancellation errors if the perturbation size is too small (Tortorelli and Michaleris, 1994). Complex variable finite difference methods (Lyness and Moler, 1967; Squire and Trapp, 1998) avoid these cancellation errors, but are unable to reduce the computational expense of evaluating the result of a perturbation of every parameter.

Automatic differentiation (AD) (also known as algorithmic differentiation) is a source code translation tool that generates an additional source code to evaluate the derivatives of the original function’s source code, with respect to its parameters. This technique exploits “the fact that every function, no matter how complicated, is executed on a computer as a (potentially very long) sequence of elementary operations such as additions, multiplications and elementary functions” (Bischof et al., 1994). AD applies the chain rule in a mechanical fashion to the composition of these elementary functions to generate a derivative-code in the native language of the original function’s source code, such that “in the absence of floating-point exceptions, [the derivative-code] computes the values of the analytical derivatives accurate up to machine precision” (Bischof et al.,

1994). One very attractive feature of AD is that changes made to the original code are automatically implemented in the derivative computing code (Giering, 1999). Typically, AD processors can operate in either forward or reverse accumulation mode. The reverse accumulation mode is favourable for functions with have fewer outputs than inputs ($n_f < n_x$), such as the presented objective function (2.2). While AD seems ideal for the evaluation of the objective function's gradient, it was not implemented in presented studies.

Finally, the most developer intensive method for evaluating derivatives of a function is to manually differentiate it, and develop an additional source code to evaluate its derivatives. For large or non-linear codes, this procedure is quite difficult and prone to errors (Bischof et al., 1994). Typically, the derivatives of a linearized version of the model are computed instead, and the linearized derivative computing model is referred to as the tangent linear model (TLM) (Rao, 2007). Additionally, changes to the source code which evaluate the function must be manually implemented in the derivative computing source code too. Despite these limitations, the gradient of the presented objective function (2.2) was evaluated with a manually implemented derivative computing source code.

The development of an objective function gradient source code required the implicit differentiation of the dispersion model. An important distinction must be made when differentiating the dispersion model, differentiation of the original PDE that governs the model (3.1) will give the sensitivity of the physical system (see Appendix A), while differentiation of its discretized form (2.1) will give the sensitivity of the specific numerical implementation of the model. Differentiation of the original PDE results in

additional PDE(s) that also must be discretized and numerically solved. This numerical approximation introduces truncation errors into the derivatives of the continuous the form of the dispersion model, which in turn cause disagreement between the objective function and the estimate of its gradient. Alternatively, the sensitivities of the discretized form of the dispersion model (2.1) are only subject to the numerical errors introduced by the linear equation solver, and as such are in closer agreement to the objective function. Since the objective function (2.2) was based on the discretized solution to the dispersion model (2.1), it was deemed appropriate to differentiate the dispersion model in its discrete form rather than its continuous PDE form (3.1).

2.4.2 Forward Sensitivity Method

The sensitivity of the objective function (2.2) with respect to the emission source parameters ($\mathbf{x}$) is represented by a row vector of objective function partial derivatives (2.13).

$$\frac{df(\mathbf{x})}{d\mathbf{x}} = \left[\frac{\partial f(\mathbf{x})}{\partial x_1} \quad \dots \quad \frac{\partial f(\mathbf{x})}{\partial x_{n_x}} \right] \quad (2.13)$$

The gradient of the objective function (2.14) is simply the transpose of the sensitivity ($df(\mathbf{x})/d\mathbf{x}$) (2.13).

$$\nabla f(\mathbf{x}) = \left(\frac{df(\mathbf{x})}{d\mathbf{x}} \right)^T \quad (2.14)$$

Differentiating the objective function (2.2) with respect to the vector of emission source parameters ($\mathbf{x}$) via chain rule gives (2.15) as a system of algebraic equations,

$$\frac{df(\mathbf{x})}{d\mathbf{x}} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{c}} \frac{d\mathbf{c}(\mathbf{x})}{d\mathbf{x}} + \frac{\partial f(\mathbf{x})}{\partial \mathbf{s}} \frac{d\mathbf{s}(\mathbf{x})}{d\mathbf{x}} \quad (2.15)$$

where $\partial f(\mathbf{x})/\partial \mathbf{c}$ (2.16) and $\partial f(\mathbf{x})/\partial \mathbf{s}$ (2.17) are row vectors of the sensitivity of the objective function to the discretized concentration ($\mathbf{c}(\mathbf{x})$) and emission source ($\mathbf{s}(\mathbf{x})$) vectors, and can be determined with matrix differentiation (Magnus and Neudecker, 1988).

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{c}} = \frac{1}{2}(\mathbf{H}\mathbf{c}(\mathbf{x}) - \mathbf{c}^0)^T \mathbf{R}^{-T} \mathbf{H} + \frac{1}{2}(\mathbf{H}\mathbf{c}(\mathbf{x}) - \mathbf{c}^0)^T \mathbf{R}^{-1} \mathbf{H} \quad (2.16)$$

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{s}} = \frac{\beta}{2}(\mathbf{s}(\mathbf{x}) - \mathbf{s}^b)^T \mathbf{B}^{-T} + \frac{\beta}{2}(\mathbf{s}(\mathbf{x}) - \mathbf{s}^b)^T \mathbf{B}^{-1} \quad (2.17)$$

Since $\mathbf{R}^{-1}$ and $\mathbf{B}^{-1}$ are both positive-definite matrices (as defined in Section 2.2), they are also symmetric, such that $\mathbf{R}^{-1} = \mathbf{R}^{-T}$ and $\mathbf{B}^{-1} = \mathbf{B}^{-T}$. This symmetry allows (2.16) and (2.17) to be simplified as (2.18) and (2.19) respectively.

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{c}} = (\mathbf{H}\mathbf{c}(\mathbf{x}) - \mathbf{c}^0)^T \mathbf{R}^{-1} \mathbf{H} \quad (2.18)$$

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{s}} = \beta(\mathbf{s}(\mathbf{x}) - \mathbf{s}^b)^T \mathbf{B}^{-1} \quad (2.19)$$

The $d\mathbf{c}(\mathbf{x})/d\mathbf{x}$ term of (2.15) is an $n_c \times n_x$ matrix that represents the sensitivity of the discretized concentration field ($\mathbf{c}(\mathbf{x})$) to emission source parameters ($\mathbf{x}$), has the structure (2.20),

$$\frac{d\mathbf{c}(\mathbf{x})}{d\mathbf{x}} = \begin{bmatrix} \frac{\partial c_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial c_1(\mathbf{x})}{\partial x_{n_x}} \\ \vdots & \ddots & \vdots \\ \frac{\partial c_{n_c}(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial c_{n_c}(\mathbf{x})}{\partial x_{n_x}} \end{bmatrix} \quad (2.20)$$

and is determined by differentiating the dispersion model (2.1) with respect to the emission source parameters ($\mathbf{x}$) to obtain (2.21),

$$\frac{d\mathbf{c}(\mathbf{x})}{d\mathbf{x}} = \mathbf{A}^{-1} \frac{d\mathbf{b}(\mathbf{x})}{d\mathbf{x}} \quad (2.21)$$

where $d\mathbf{b}(\mathbf{x})/d\mathbf{x}$ is an $n_c \times n_x$ matrix which has the structure (2.22).

$$\frac{d\mathbf{b}(\mathbf{x})}{d\mathbf{x}} = \begin{bmatrix} \frac{\partial b_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial b_1(\mathbf{x})}{\partial x_{n_x}} \\ \vdots & \ddots & \vdots \\ \frac{\partial b_{n_c}(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial b_{n_c}(\mathbf{x})}{\partial x_{n_x}} \end{bmatrix} \quad (2.22)$$

The bulk of the computational effort in evaluating the sensitivity of the objective function to the emission source parameters ($df(\mathbf{x})/d\mathbf{x}$) (2.15) involves evaluating the sensitivity of the modelled pollutant concentration field to the emission source parameters ($d\mathbf{c}(\mathbf{x})/d\mathbf{x}$) (2.21). In the presented studies $d\mathbf{c}(\mathbf{x})/d\mathbf{x}$ is a very large dense matrix that is infeasible to evaluate. The computational complexity of evaluating the gradient of the objective function can be reduced with an adjoint method.

2.4.3 Adjoint Sensitivity Method

The adjoint sensitivity method is an alternative method of determining the gradient of a function. This method mimics the reverse Jacobian accumulation mode of automatic

differentiation, and is computationally efficient when the number of the function's outputs is lower than the number of inputs ($n_f < n_x$).

Following the example of Hughes (2009), the adjoint form of the objective function's gradient is derived from an alternate, but equivalent objective function, $\Lambda(\mathbf{x})$ (2.23). This alternate objective function (typically termed the Lagrange function) augments the original objective function $f(\mathbf{x})$ with the inner product of a column vector of adjoint variables $\boldsymbol{\lambda}$ (also known as Lagrange multipliers, conjugate variables, or co-state variables (Hughes, 2009)) of length n_c and the residual error of the governing dispersion model (2.1) (also referred to as the problem constraints).

$$\Lambda(\mathbf{x}) = f(\mathbf{x}) - \boldsymbol{\lambda}^T (\mathbf{A}\mathbf{c}(\mathbf{x}) - \mathbf{b}(\mathbf{x})) \quad (2.23)$$

Since the residual error of the dispersion model is expected to be the zero-vector, this alternate objective function $\Lambda(\mathbf{x})$ (2.23) is numerically equivalent to the original $f(\mathbf{x})$ (2.2), as are its derivatives with respect to the emission source parameters $(\mathbf{x})$. Differentiating (2.23) with respect to the emission source parameters $(\mathbf{x})$ gives (2.24).

$$\frac{d\Lambda(\mathbf{x})}{d\mathbf{x}} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{c}} \frac{d\mathbf{c}(\mathbf{x})}{d\mathbf{x}} + \frac{\partial f(\mathbf{x})}{\partial \mathbf{s}} \frac{d\mathbf{s}(\mathbf{x})}{d\mathbf{x}} - \boldsymbol{\lambda}^T \left(\mathbf{A} \frac{d\mathbf{c}(\mathbf{x})}{d\mathbf{x}} - \frac{d\mathbf{b}(\mathbf{x})}{d\mathbf{x}} \right) \quad (2.24)$$

Rearrangement of (2.24) to collect the $d\mathbf{c}(\mathbf{x})/d\mathbf{x}$ terms gives (2.25).

$$\frac{d\Lambda(\mathbf{x})}{d\mathbf{x}} = \boldsymbol{\lambda}^T \frac{d\mathbf{b}(\mathbf{x})}{d\mathbf{x}} + \frac{\partial f(\mathbf{x})}{\partial \mathbf{s}} \frac{d\mathbf{s}(\mathbf{x})}{d\mathbf{x}} + \left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{c}} - \boldsymbol{\lambda}^T \mathbf{A} \right) \frac{d\mathbf{c}(\mathbf{x})}{d\mathbf{x}} \quad (2.25)$$

If $\boldsymbol{\lambda}$ is chosen to satisfy the system of linear equations given by (2.26),

$$\mathbf{A}^T \boldsymbol{\lambda} = \left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{c}} \right)^T \quad (2.26)$$

then the $(\partial f(\mathbf{x})/\partial \mathbf{c} - \boldsymbol{\lambda}^T \mathbf{A})$ term of (2.25) will evaluate to the zero-vector, making it unnecessary to evaluate $d\mathbf{c}(\mathbf{x})/d\mathbf{x}$, which was deemed to be the most computationally intensive part of determining the sensitivity of the objective function to the emission source parameters $(df(\mathbf{x})/d\mathbf{x})$ via the forward sensitivity method (see Section 2.4.2). The solution of the adjoint variables $(\boldsymbol{\lambda})$ (2.26) requires the same computational effort as the dispersion model (2.1), and its computational complexity is independent of number of function parameters (n_x) , and instead scales with the number of function outputs (n_f) . The adjoint form of the sensitivity of the objective function (2.25) can be simplified as (2.27),

$$\frac{df(\mathbf{x})}{d\mathbf{x}} = \frac{d\Lambda(\mathbf{x})}{d\mathbf{x}} = \boldsymbol{\lambda}^T \frac{d\mathbf{b}(\mathbf{x})}{d\mathbf{x}} + \frac{\partial f(\mathbf{x})}{\partial \mathbf{s}} \frac{d\mathbf{s}(\mathbf{x})}{d\mathbf{x}} \quad (2.27)$$

where the adjoint variables $(\boldsymbol{\lambda})$ are subject to the constraint of (2.26).

As shown in (2.14), the gradient of the objective function is simply the transpose of the sensitivity row vector $(df(\mathbf{x})/d\mathbf{x})$, which gives (2.28).

$$\nabla f(\mathbf{x}) = \left(\frac{d\mathbf{b}(\mathbf{x})}{d\mathbf{x}} \right)^T \boldsymbol{\lambda} + \left(\frac{d\mathbf{s}(\mathbf{x})}{d\mathbf{x}} \right)^T \left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{s}} \right)^T \quad (2.28)$$

This adjoint form of the objective function's gradient can be used by a gradient-based numerical optimization algorithm such as L-BFGS-B to determine values of the emission source parameters $(\mathbf{x})$ that minimize the objective function $(f(\mathbf{x}))$.

2.5 Summary

An inverse dispersion modelling problem was reformulated as an optimization problem with a scalar valued objective function $f(\mathbf{x})$ with many parameters (2.2). The L-BFGS-B numerical optimization algorithm was presented to constrain the values of the emission source parameters ($\mathbf{x}$) that minimized the objective function and improved agreement between the known ($\mathbf{c}^0$) and modelled ($\mathbf{Hc}(\mathbf{x})$) receptor observations. Finally, an adjoint sensitivity method was presented to efficiently evaluate the objective function gradient ($\nabla f(\mathbf{x})$) for use by the L-BFGS-B. Figure 2.1 shows a schematic overview of the numerical optimization procedure for the inverse dispersion model.

Figure 2.1: Flow chart overview of the inverse dispersion modelling procedure

Chapter 3

Dispersion Modelling

The transport of pollution in the atmosphere is governed by range of phenomena such as “... emissions of primary pollutants, atmospheric chemical transformations, advection by the wind field, turbulent diffusion and dry deposition” (Sportisse, 2001). A variety of numerical models such as AERMOD (Cimorelli et al., 2004), CALPUFF (Scire et al., 1990), and CMAQ (Byun and Schere, 2004) have been developed to predict the dispersion of pollutants within the atmosphere at the global to meso-scales. At these scales, conventional dispersion models do not consider the influence that urban structures can have on a wind field or pollution transport. However, at the micro-scale of fugitive emissions from an industrial gas processing facility, buildings can have a significant influence on the transport of pollutants.

This chapter presents the partial differential equation (PDE) that models unsteady pollution transport, and a numerical method to solve this PDE and approximate the pollution concentration field from known emission sources. This custom micro-scale atmospheric dispersion model was used to study the pollutant concentration field and receptor observations resulting from candidate emission sources within a complex domain of an industrial gas processing facility.

3.1 Governing Transport Equation

The unsteady transport of a single, inert, neutrally buoyant pollutant, in an atmospheric flow is governed by advection with the wind and diffusion from turbulent mixing. The mass concentration of the pollutant $c = c(\mathbf{r}, t; \mathbf{x})$ in $[\text{kg} \cdot \text{m}^{-3}]$ at the Cartesian spatial coordinate $\mathbf{r} = [x \ y \ z]^T$ and time t can be modelled by a parabolic (Kuzmin, 2010; Stockie, 2011) PDE known as the advection-diffusion equation (ADE) (1.4) (McRae et al., 1982; Stockie, 2011) (also known as the convection-diffusion, or generic scalar transport equation),

$$\frac{\partial c}{\partial t} + \nabla \cdot (c\mathbf{u} - \mathbf{D}\nabla c) = s \quad (3.1)$$

where $\mathbf{u} = \mathbf{u}(\mathbf{r}, t) = [u(\mathbf{r}, t) \ v(\mathbf{r}, t) \ w(\mathbf{r}, t)]^T$ is the flow velocity vector field $[\text{m} \cdot \text{s}^{-1}]$. $\mathbf{D} = \mathbf{D}(\mathbf{r}, t)$ is the diagonal mass diffusivity matrix $[\text{m}^2 \cdot \text{s}^{-1}]$ (McRae et al., 1982). $s = s(\mathbf{r}, t; \mathbf{x})$ is the parameterized volumetric source emission rate of the pollutant $[\text{kg} \cdot \text{m}^{-3} \cdot \text{s}^{-1}]$, and $\mathbf{x}$ is the column vector of $n_{\mathbf{x}}$ emission source parameters (see Chapter 2) that define candidate emission source distributions.

The solution of this mass conservation equation (3.1) is sought within the finite airshed domain Ω_t . This computational domain is formed from the temporal domain $[t_0, t_f]$ and the spatial domain Ω , which comprises the air surrounding the buildings of the gas processing plant, within the boundary $\Gamma = \partial\Omega$. The boundary Γ of the spatial domain is formed from the open inflow (Γ_{inlet}) and outflow (Γ_{outlet}) surfaces and also includes solid surfaces through which there is no contaminant transport such as the ground and building surfaces (Γ_{wall}). Figure 3.1 shows the two-dimensional spatial domain Ω and boundary Γ

of the studied gas processing facility. Due to the transient nature of wind, the boundaries which are considered inlets and outlets will vary with time. However, for the inflow conditions specified in Section 4.1.2.1, the Southern boundary was always an inlet, the Northern boundary was always an outlet, and the Eastern and Western boundary switched between inlets and outlets depending on the originating direction of the inflow wind as specified in Figure 4.3.

Figure 3.1: Two-dimensional domain of an industrial gas processing plant

3.1.1 Initial and Boundary Conditions

To restrict the solution of the ADE (3.1) to a unique distribution of a pollutant's concentration (c), extra information is required in the form of initial conditions (3.2),

$$c(\mathbf{r}, t_0; \mathbf{x}) = c_0(\mathbf{r}; \mathbf{x}) \quad (3.2)$$

and boundary conditions (3.3) (Skiba and Parra-Guevara, 2000),

$$\begin{aligned} (c\mathbf{u} - \mathbf{D}\nabla c) \cdot \hat{\mathbf{n}} &= 0 & \text{on } \Gamma \text{ for } \mathbf{u} \cdot \hat{\mathbf{n}} < 0 & \quad (\text{inlet}) \\ (\mathbf{D}\nabla c) \cdot \hat{\mathbf{n}} &= 0 & \text{on } \Gamma \text{ for } \mathbf{u} \cdot \hat{\mathbf{n}} \geq 0 & \quad (\text{outlet/wall}) \end{aligned} \quad (3.3)$$

where $\hat{\mathbf{n}} = \hat{\mathbf{n}}(\mathbf{r})$ is the outward normal unit vector of the domain boundary Γ . These boundary conditions (3.3) specify that at an inlet ($\mathbf{u} \cdot \hat{\mathbf{n}} < 0$) there is no net transport of the pollutant, and at a wall or outlet ($\mathbf{u} \cdot \hat{\mathbf{n}} \geq 0$), there is no diffusional transport (i.e. no pollution transport through a wall where the flow velocity normal to the wall is also zero ($\mathbf{u} \cdot \hat{\mathbf{n}} = 0$), and only advective transport at an outlet). It should be noted that these boundary conditions (3.3), do not account for wet or dry deposition of the pollutant on the ground. Additionally, the initial condition (3.2) of the ADE (3.1) is a function of the emission source parameters $\mathbf{x}$, to accommodate the influence of emission sources on the initial distribution of a pollutant's concentration within the domain.

Before the ADE (3.1) can be solved to model the pollutant concentration field $c(\mathbf{r}, t, \mathbf{x})$ resulting from a candidate distribution emission sources $s(\mathbf{r}, t; \mathbf{x})$, the flow velocity vector field ($\mathbf{u}$) and mass diffusivity ($\mathbf{D}$) must be determined. The ADE (3.1) assumes that all spatial and temporal scales of turbulent fluctuations within the flow field are resolved. This is an unreasonable expectation of the current computational fluid dynamics (CFD) hardware/software for atmospheric flows. Since the aim of this thesis is a proof-of-concept study of inverse micro-scale atmospheric dispersion modelling for fugitive emissions in the upstream oil and gas industry, the flow field properties ($\mathbf{u}$ and $\mathbf{D}$) were obtained from a CFD model of a synthetic, unsteady, Reynolds-averaged, two-dimensional flow through a simplified model of the gas processing plant depicted in Figure 3.1. The diffusional effects of turbulent mixing were incorporated into the dispersion model (3.1) with Reynolds-averaging and turbulence modelling.

3.1.2 Turbulent Mixing and Mass Diffusion

It is unreasonable to expect all spatial and temporal scales of turbulent fluctuations in an atmospheric flow to be fully resolved. Instead, obtaining a time-smoothed or Reynolds-averaged (Reynolds, 1895; Bird et al., 1960) representation of the flow physics is more feasible. The mixing effects of turbulence can be modelled as mass diffusion with an empirical correlation such as the eddy diffusivity model (3.4) (Bird et al., 1960).

$$\overline{c'\mathbf{u}'} = -\mathbf{D}_t \nabla \bar{c} \quad (3.4)$$

where $\overline{c'\mathbf{u}'}$ is the averaged turbulent mass flux vector and $\bar{c}$ is the averaged pollutant concentration that result from a Reynolds decomposition (Wilcox, 1993) of the ADE (3.1). $\mathbf{D}_t$ is a diagonal matrix representing the turbulent eddy diffusivity (McRae et al., 1982), with the structure given by (3.5) for Cartesian coordinates,

$$\mathbf{D}_t = \mathbf{D}_t(\mathbf{r}, t) = \begin{bmatrix} D_{tx}(\mathbf{r}, t) & 0 & 0 \\ 0 & D_{ty}(\mathbf{r}, t) & 0 \\ 0 & 0 & D_{tz}(\mathbf{r}, t) \end{bmatrix} \quad (3.5)$$

where D_{tx} , D_{ty} , and D_{tz} are the turbulent eddy diffusivities in the respective Cartesian coordinate directions x , y , and z . The eddy diffusivity model (3.4) is analogous to Fick's first law of diffusion (Fick, 1855; Bird et al., 1960) or Fourier's law of heat conduction (Turns, 1996).

If turbulent diffusion is assumed to be isotropic ($D_t = D_{tx} = D_{ty} = D_{tz}$), then the turbulent eddy diffusivity D_t can be estimated from the turbulent eddy kinematic

viscosity ν_t of the Reynolds-averaged Navier-Stokes (RANS) solution of the flow field. Jeong et al. (2002) recommend the Re-Normalization Group (RNG) k - ε turbulence model (Yakhot and Orszag, 1986; Yakhot et al., 1992) for simulations of wind around bluff structures (e.g. buildings). This is a two equation, isotropic turbulence model that estimates the turbulence kinetic energy per unit mass k with the units $[\text{m}^2 \cdot \text{s}^{-2}]$ and eddy dissipation rate ε with the units $[\text{m}^2 \cdot \text{s}^{-3}]$ (ANSYS, 2006a).

At high Reynolds numbers, the turbulent eddy kinematic viscosity (ν_t) can be estimated from the k and ε fields predicted by the k - ε turbulence model with (3.6) (Jones and Launder, 1972),

$$\nu_t = C_\mu \frac{k^2}{\varepsilon} \quad (3.6)$$

where C_μ is an empirical turbulence model constant ($C_\mu = 0.09$ for standard k - ε (Jones and Launder, 1972) and $C_\mu \approx 0.085$ for RNG k - ε (Yakhot et al., 1992)).

The turbulent eddy kinematic viscosity ν_t is related to the turbulent mass diffusivity D_t through the turbulent Schmidt number Sc_t . The generic Schmidt number Sc (3.7), is a non-dimensional ratio of momentum diffusivity (or kinematic viscosity ν) to mass diffusivity D , and assumed to be of the order unity (Bird et al., 2007).

$$Sc = \frac{\nu}{D} \quad (3.7)$$

Combination of (3.6) and (3.7) gives (3.8) as an estimate of the turbulent mass diffusivity (D_t),

$$D_t = \frac{v_t}{Sc_t} = \frac{C_\mu}{Sc_t} \frac{k^2}{\varepsilon} \quad (3.8)$$

where $Sc_t = 0.9$ is the turbulent Schmidt number (Launder and Sharma, 1974).

In addition to diffusion effects of turbulent mixing, the transport of a pollutant can also be influenced by molecular diffusion. The isotropic mass diffusivity D_{AB} of the binary mixture of air and a pollutant can be estimated with the Chapman-Enskog kinetic theory (Bird et al., 2007). Since turbulent and molecular diffusion are both modelled as linear processes, their combined effects can also be modelled by simply summing the respective diffusivities into a single mass diffusivity D (3.9).

$$D = D_t + D_{AB} \quad (3.9)$$

This total mass diffusivity (D) was used by the custom dispersion model in place of the diagonal mass diffusivity matrix ($\mathbf{D}$). Preliminary CFD modelling has shown that within the wakes of turbulent mixing downstream of buildings, the turbulent mass diffusivity dominates (D_t). However, outside of these wakes molecular diffusion (D_{AB}) is the dominant mass diffusion mechanism.

3.2 Numerical Solution

It is challenging to solve the ADE (3.1) within a complex airshed domain such as the industrial gas processing plant shown in Figure 3.1. Exact solutions of the ADE (3.1) exist only for idealized models of pollution transport. However, the approximate solutions to the ADE can be obtained numerically.

The numerical solution of the ADE (3.1) is tackled in two stages. First, as a boundary value problem (BVP) by numerically integrating the ADE (3.1) over a discretized representation of the airshed domain to obtain a system of ordinary differential equations (ODEs) in time (see Section 3.2.1). The ODEs are then considered as an initial value problem (IVP) which are numerically solved by discretizing and integrating them in time to form a sparse system of linear equations (see Section 3.2.2), for which the solution can be approximated with a sparse linear equation solving routine.

3.2.1 Boundary Value Problem

Generally, the numerical solution of a BVP begins by discretizing the spatial domain Ω and boundary Γ , into an assembly of simple elements $\Omega^{(e)}$ and boundaries $\Gamma^{(e)}$. Typically, this discrete approximation of the spatial domain and boundary is referred to as a mesh (Berg et al., 2008). The mesh provides a stencil upon which the spatial derivatives of the PDE of interest are approximated, and a piecewise approximation of the PDE's solution is constructed. Three popular methods for the piecewise approximation of the solution to the PDE are the finite difference, finite volume and finite element methods.

The finite difference method (FDM) approximates each spatial derivative term of the governing PDE with a truncated Taylor series expansion about each node of the mesh to form a coupled system of algebraic difference equations. This Taylor series approximation of the spatial derivatives typically favours structured meshes that align with the coordinate system of the PDE of interest. Additionally, the FDM is the most

mature of the classical numerical BVP solution methods, and has had considerable stability analysis (Crank and Nicolson, 1947; Charney et al., 1950). Due to the structured nature of the discretized finite difference domains, these linear systems typically manifest as banded matrices which have efficient storage structures and solution algorithms. Additionally, due to the alignment of the mesh with the coordinate system of the PDE, the linear system can be directionally decoupled providing an additional reduction in computational complexity of their numerical solution (Marchuk, 1986). Atmospheric dispersion modelling applications of the FDM are best suited for the global to meso-scales (Klug et al., 1992; Carmichael et al., 2003; Hourdin et al., 2006) due to the ease of discretizing airshed domains on these scales with a structured, orthogonal mesh. Unfortunately, creating a structured and orthogonal mesh for micro-scale atmospheric dispersion modelling applications can be very challenging due to the irregular placement and orientations of buildings, which form a portion of the mesh domain's boundaries.

The finite volume method (FVM) (Versteeg and Malalasekera, 2007) approximates the computational domain as an assembly of many control volumes and enforces a flux balance between control volumes that share an interface. This flux balance is ensured by integrating the PDE of interest over each control volume and transforming the volume integral of the divergence of flux ($\mathbf{F}$) term ($\int_{\Omega(e)} \nabla \cdot (\mathbf{F}) d\Omega$) into a boundary integral of the flux ($\int_{\Gamma(e)} (\mathbf{F} \cdot \hat{\mathbf{n}}) d\Gamma$) via the divergence theorem (A.19). For this reason, the FVM is typically favoured for conservation equations such as the ADE (3.1). However, the numerical solution obtained with the FVM represent volume averages, resulting in solution discontinuities at control volume interfaces. Additionally, the FVM is unable to

exactly accommodate Neumann or Robin boundary conditions such as those presented for the ADE (3.3).

Finally, the finite element method (FEM) discretizes the computational domain into an assembly of elements, and approximates the weak solution of a BVP with respect to a family of appropriately chosen weight and interpolation functions. The weight functions of the standard Galerkin FEM (i.e. Bubnov-Galerkin) (Mikhlin, 1964) are ill-suited for advection-dominated BVPs such as pollution transport (Huebner et al., 2001). Brooks and Hughes (1982) proposed a streamline-upwind Petrov-Galerkin (SUPG) stabilization technique for advection-dominated transport problems that is able to achieve stable solutions by adding artificial diffusion in the flow direction only (Donea and Huerta, 2003). Finally, the FEM can exactly accommodate Neumann and Robin boundary conditions (3.3), and the elemental interpolation functions of the method allow for the interpretation of the numerical solution of the BVP as a piecewise continuous function.

The FVM and FEM are both able to accommodate unstructured meshes, making them suitable for micro-scale dispersion modelling. The FEM was selected to numerically solve the BVP portion of the ADE (3.1), and obtain an IVP as a system of coupled ODEs that model the evolution of the mesh's nodal pollutant concentrations resulting from the candidate distribution of emission sources and a time-varying flow field.

3.2.1.1 *The Finite Element Method of Weighted Residuals*

A BVP such as the ADE (3.1) can be spatially discretized and reformulated as a variational problem with the FEM of weighted residuals. This method seeks a solution that minimizes the inner product of an elemental weight function $W_i(\mathbf{r}, t)$ (also known as test function) and the residual error of the discretized BVP. This weight function W_i is dependent on the unsteady flow within the element, and as such is dependent on both space ($\mathbf{r}$) and time (t). A weak solution is sought for this modified discrete problem, rather than the original continuous problem. If the discrete elements are small enough and the weight functions are appropriately chosen, the weak solution is a reasonable approximation of the solution of the original BVP. Following the example of Huebner et al. (2001) for the application of the FEM of weighted residuals to advection-diffusion type PDEs a numerical method suitable for solving the ADE with steady emission sources within a two-dimensional domain, that has been discretized into a triangular mesh (3-node triangular elements) is presented.

The FEM formulation of the ADE (3.1) on a single element $\Omega^{(e)}$ of the discretized domain gives (3.10),

$$\int_{\Omega^{(e)}} W_i \left(\frac{\partial c}{\partial t} + \nabla \cdot (c\mathbf{u} - \mathbf{D}\nabla c) - s \right) d\Omega = 0 \quad (3.10)$$

where the weight function $W_i = W_i(\mathbf{r}, t) = 0$ when evaluated outside the domain of the element of interest, and is subject to the constraints of $W_i(\mathbf{r}, t) = 1$ when evaluated at the i th node of the element, and $W_i(\mathbf{r}, t) = 0$ at all other nodes of the element. The compact support of the weight function $W_i(\mathbf{r}, t)$ constrains the spatial integration of (3.10) to the

element of interest. Evaluation of (3.10) with the application of the divergence theorem gives (3.11).

$$\begin{aligned} \int_{\Omega^{(e)}} W_i \frac{\partial c}{\partial t} d\Omega + \int_{\Omega^{(e)}} (\nabla W_i) \cdot (\mathbf{D} \nabla c - c \mathbf{u}) d\Omega \\ = \int_{\Omega^{(e)}} W_i s d\Omega + \int_{\Gamma^{(e)}} W_i (c \mathbf{u} - \mathbf{D} \nabla c) \cdot \hat{\mathbf{n}} d\Gamma \end{aligned} \quad (3.11)$$

The original second-order PDE has been reduced to a first-order PDE, which is a weak form of the ADE. Conveniently, a term which can accommodate Robin boundary conditions (3.3) has been introduced $(\int_{\Gamma^{(e)}} W_i (c \mathbf{u} - \mathbf{D} \nabla c) \cdot \hat{\mathbf{n}} d\Gamma)$, and for a two-dimensional element $\nabla W_i = \nabla W_i(\mathbf{r}, t) = \left[\frac{\partial W_i(\mathbf{r}, t)}{\partial x} \quad \frac{\partial W_i(\mathbf{r}, t)}{\partial y} \right]$.

3.2.1.2 Interpolation Functions

The spatial ($\mathbf{r}$) variation of the flow velocity ($\mathbf{u}(\mathbf{r}, t)$), mass diffusivity ($\mathbf{D}(\mathbf{r}, t)$), pollutant volumetric emission rate ($s(\mathbf{r}; \mathbf{x})$), pollutant mass concentration ($c(\mathbf{r}, t; \mathbf{x})$), and related derivatives ($\partial c(\mathbf{r}, t; \mathbf{x}) / \partial t$ and $\nabla c(\mathbf{r}, t; \mathbf{x})$) can be approximated within the element of interest with interpolation functions $\mathbf{N} = \mathbf{N}(\mathbf{r})$ (also known as shape or trial functions). Where $\mathbf{N}(\mathbf{r})$ represents a column vector of interpolation functions $N_i = N_i(\mathbf{r})$, which correspond to the interpolation weight factor for the i th node of the element. Each interpolation function $N_i(\mathbf{r})$ is subject to the constraints of $N_i(\mathbf{r}) = 1$ when evaluated at the i th node of the element, $N_i(\mathbf{r}) = 0$ when evaluated at all other nodes of the element, and $\sum N_i(\mathbf{r}) = 1$. Huebner et al. (2001) outline the appropriate

choice of interpolation functions for a variety of element types. For a 3-node triangular element the appropriate interpolation functions $\mathbf{N}(\mathbf{r})$ are given by (3.12),

$$\mathbf{N}(\mathbf{r}) = \begin{bmatrix} N_1(\mathbf{r}) \\ N_2(\mathbf{r}) \\ N_3(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.12)$$

where x_i and y_i are the Cartesian coordinates of the i th node of the element, and $\mathbf{r} = [x \ y]$. Figure 3.2 shows the resulting shape functions for a single node over its connecting triangular elements.

Figure 3.2: Shape function of a single node over the connected triangular elements

The gradient of the interpolation functions ($\nabla \mathbf{N} = \nabla \mathbf{N}(\mathbf{r})$) (3.13) is obtained by differentiating the shape functions (3.12) with respect to the x and y spatial coordinates of position vector $\mathbf{r}$.

$$\nabla \mathbf{N}(\mathbf{r}) = \begin{bmatrix} \frac{\partial N_1(\mathbf{r})}{\partial x} & \frac{\partial N_1(\mathbf{r})}{\partial y} \\ \frac{\partial N_2(\mathbf{r})}{\partial x} & \frac{\partial N_2(\mathbf{r})}{\partial y} \\ \frac{\partial N_3(\mathbf{r})}{\partial x} & \frac{\partial N_3(\mathbf{r})}{\partial y} \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (3.13)$$

It can be seen that the shape function gradient ($\nabla \mathbf{N}(\mathbf{r})$) of a 3-node triangular element are not dependant on the spatial coordinates x and y , which means that $\nabla \mathbf{N}(\mathbf{r})$ is constant within a triangular element.

The spatial ($\mathbf{r}$) distributions of flow velocity ($\mathbf{u}(\mathbf{r}, t)$), mass diffusivity ($\mathbf{D}(\mathbf{r}, t)$), pollutant source volumetric emission rate ($s(\mathbf{r}; \mathbf{x})$) and pollutant mass concentration ($c(\mathbf{r}, t; \mathbf{x})$), and related derivatives ($\dot{c}(\mathbf{r}, t; \mathbf{x})$ and $\nabla c(\mathbf{r}, t; \mathbf{x})$) in (3.11) can be approximated with the interpolation function $\mathbf{N}(\mathbf{r})$ and the elemental nodal values of the flow velocity ($\mathbf{U}^{(e)}(t)$) (3.14), mass diffusivity ($\mathcal{D}^{(e)}(t)$) (3.15), pollutant mass concentration ($\mathbf{c}^{(e)}(t; \mathbf{x})$, $\dot{\mathbf{c}}^{(e)}(t; \mathbf{x})$) (3.16), (3.17), (3.18) and pollutant source volumetric emission rate ($\mathbf{s}^{(e)}(\mathbf{x})$) (3.19).

$$\mathbf{u}(\mathbf{r}, t) \cong \left(\mathbf{U}^{(e)}(t) \right) \mathbf{N}(\mathbf{r}) \quad (3.14)$$

$$\mathbf{D}(\mathbf{r}, t) \cong \left(\mathcal{D}^{(e)}(t) \right) \mathbf{N}(\mathbf{r}) \quad (3.15)$$

$$c(\mathbf{r}, t; \mathbf{x}) \cong \left(\mathbf{N}(\mathbf{r}) \right)^T \mathbf{c}^{(e)}(t; \mathbf{x}) \quad (3.16)$$

$$\dot{c}(\mathbf{r}, t; \mathbf{x}) \cong \left(\mathbf{N}(\mathbf{r}) \right)^T \dot{\mathbf{c}}^{(e)}(t; \mathbf{x}) \quad (3.17)$$

$$\nabla c(\mathbf{r}, t; \mathbf{x}) \cong \left(\nabla \mathbf{N}(\mathbf{r}) \right)^T \mathbf{c}^{(e)}(t; \mathbf{x}) \quad (3.18)$$

$$s(\mathbf{r}; \mathbf{x}) = (\mathbf{N}(\mathbf{r}))^T \mathbf{s}^{(e)}(\mathbf{x}) \quad (3.19)$$

For a 3-node triangular element the structure of $\mathbf{c}^{(e)}$, $\dot{\mathbf{c}}^{(e)}$, $\mathbf{D}^{(e)}$, $\mathbf{U}^{(e)}$ and $\mathbf{s}^{(e)}$ are given by (3.20) - (3.24),

$$\mathbf{c}^{(e)} = \mathbf{c}^{(e)}(t; \mathbf{x}) = [c_1^{(e)}(t; \mathbf{x}) \quad c_2^{(e)}(t; \mathbf{x}) \quad c_3^{(e)}(t; \mathbf{x})]^T \quad (3.20)$$

$$\dot{\mathbf{c}}^{(e)} = \dot{\mathbf{c}}^{(e)}(t; \mathbf{x}) = \left[\frac{\partial c_1^{(e)}(t; \mathbf{x})}{\partial t} \quad \frac{\partial c_2^{(e)}(t; \mathbf{x})}{\partial t} \quad \frac{\partial c_3^{(e)}(t; \mathbf{x})}{\partial t} \right]^T \quad (3.21)$$

$$\mathbf{D}^{(e)} = \mathbf{D}^{(e)}(t) = (\mathbf{D}_1^{(e)} \quad \mathbf{D}_2^{(e)} \quad \mathbf{D}_3^{(e)}) \quad (3.22)$$

$$\mathbf{U}^{(e)} = \mathbf{U}^{(e)}(t) = \begin{bmatrix} u_1^{(e)}(t) & u_2^{(e)}(t) & u_3^{(e)}(t) \\ v_1^{(e)}(t) & v_2^{(e)}(t) & v_3^{(e)}(t) \end{bmatrix} \quad (3.23)$$

$$\mathbf{s}^{(e)} = \mathbf{s}^{(e)}(\mathbf{x}) = [s_1^{(e)}(\mathbf{x}) \quad s_2^{(e)}(\mathbf{x}) \quad s_3^{(e)}(\mathbf{x})]^T \quad (3.24)$$

where $c_i^{(e)}$, $u_i^{(e)}$, $v_i^{(e)}$, $\mathbf{D}_i^{(e)}$, and $s_i^{(e)}$ reflect the pollutant concentration, two-dimensional flow velocity, mass diffusivity and volumetric emission rate at i th node of the element.

Substitution of (3.14) - (3.19) into (3.11) gives (3.25),

$$\begin{aligned} \int_{\Omega^{(e)}} (W_i \mathbf{N}^T) d\Omega \dot{\mathbf{c}}^{(e)} + \int_{\Omega^{(e)}} (\nabla W_i) \left((\mathbf{D}^{(e)} \mathbf{N})(\nabla \mathbf{N})^T - \mathbf{U}^{(e)} \mathbf{N} \mathbf{N}^T \right) d\Omega \mathbf{c}^{(e)} \\ = \int_{\Omega^{(e)}} (W_i \mathbf{N}^T) d\Omega \mathbf{s}^{(e)} + \int_{\Gamma_N^{(e)}} (W_i (c\mathbf{u} - \mathbf{D} \nabla c) \cdot \hat{\mathbf{n}}) d\Gamma \end{aligned} \quad (3.25)$$

Eq. (3.25) represents a single ODE that relates the evolution of all the nodal values of a pollutant's concentration within an element. However, a single ODE is insufficient to constrain the pollutant's nodal concentration to a unique solution. Additional ODEs can be derived by evaluating (3.25) with an elemental weight function W_i for each node of the element, resulting in 3 coupled ODEs for a 3-node triangular element. The elements of the coefficient matrices that form these ODEs can be simultaneously determined by

substituting the node specific weight function W_i with a column vector $\mathbf{W}$ of all the weight functions within the element to give (3.26),

$$\begin{aligned} & \left(\int_{\Omega^{(e)}} \mathbf{W} \mathbf{N}^T d\Omega \right) \dot{\mathbf{c}}^{(e)} \\ & + \left(\int_{\Omega^{(e)}} (\nabla \mathbf{W}) \left((\mathcal{D}^{(e)} \mathbf{N}) (\nabla \mathbf{N})^T - \mathbf{U}^{(e)} \mathbf{N} \mathbf{N}^T \right) d\Omega \right) \mathbf{c}^{(e)} \\ & = \left(\int_{\Omega^{(e)}} \mathbf{W} \mathbf{N}^T d\Omega \right) \mathbf{s}^{(e)} + \left(\int_{\Gamma^{(e)}} \mathbf{W} ((c\mathbf{u} - \mathbf{D} \nabla c) \cdot \hat{\mathbf{n}}) d\Gamma \right) \end{aligned} \quad (3.26)$$

where for a triangular element $\mathbf{W} = \mathbf{W}(\mathbf{r}, t) = [W_1(\mathbf{r}, t) \quad W_2(\mathbf{r}, t) \quad W_3(\mathbf{r}, t)]^T$, and the structure of the gradient of the weight function $\nabla \mathbf{W}$ is given by (3.27).

$$\nabla \mathbf{W} = \nabla \mathbf{W}(\mathbf{r}, t) = \begin{bmatrix} \frac{\partial W_1(\mathbf{r}, t)}{\partial x} & \frac{\partial W_1(\mathbf{r}, t)}{\partial y} \\ \frac{\partial W_2(\mathbf{r}, t)}{\partial x} & \frac{\partial W_2(\mathbf{r}, t)}{\partial y} \\ \frac{\partial W_3(\mathbf{r}, t)}{\partial x} & \frac{\partial W_3(\mathbf{r}, t)}{\partial y} \end{bmatrix} \quad (3.27)$$

For compactness, (3.26) can be rewritten as (3.28) (Huebner et al., 2001),

$$\mathbf{C}^{(e)} \dot{\mathbf{c}}^{(e)} + \mathbf{K}^{(e)} \mathbf{c}^{(e)} = \mathbf{C}^{(e)} \mathbf{s}^{(e)} + \mathbf{f}_{\Gamma}^{(e)} \quad (3.28)$$

where the elemental capacitance (or dampening) $\mathbf{C}^{(e)} = \mathbf{C}^{(e)}(t)$ and stiffness $\mathbf{K}^{(e)} = \mathbf{K}^{(e)}(t)$ matrices and boundary condition vector $\mathbf{f}_{\Gamma}^{(e)} = \mathbf{f}_{\Gamma}^{(e)}(t; \mathbf{x})$ are given as (3.29), (3.30), and (3.31) respectively.

$$\mathbf{C}^{(e)} = \mathbf{C}^{(e)}(t) = \int_{\Omega^{(e)}} \mathbf{W} \mathbf{N}^T d\Omega \quad (3.29)$$

$$\mathbf{K}^{(e)} = \mathbf{K}^{(e)}(t) = \int_{\Omega^{(e)}} (\nabla \mathbf{W}) \left((\mathcal{D}^{(e)} \mathbf{N})(\nabla \mathbf{N})^T - \mathbf{U}^{(e)} \mathbf{N} \mathbf{N}^T \right) d\Omega \quad (3.30)$$

$$\mathbf{f}_{\Gamma}^{(e)} = \mathbf{f}_{\Gamma}^{(e)}(t; \mathbf{x}) = \int_{\Gamma^{(e)}} \mathbf{W}((\mathbf{c}\mathbf{u} - \mathbf{D}\nabla c) \cdot \hat{\mathbf{n}}) d\Gamma \quad (3.31)$$

The influence of boundary condition vector $\mathbf{f}_{\Gamma}^{(e)}$ only needs to be considered at the boundaries of the entire discretized domain, rather than the boundaries of individual elements, since the pollutant flux $(\mathbf{c}\mathbf{u} - \mathbf{D}\nabla c)$ between elements that share an interface will be cancelled due to the opposing directions of their outward normal unit vectors ($\hat{\mathbf{n}}$) along their shared boundary.

3.2.1.3 Upwinding

The selection of weight functions ($\mathbf{W}$) appropriate for the type of element and PDE is critical to obtaining an accurate numerical approximation of the solution. The application of the conventional Bubnov-Galerkin FEM ($\mathbf{W} = \mathbf{N}$) (Mikhlin, 1964) to model advection-dominated physics such as pollution transport can result in a numerical solution that exhibits spurious, unbounded oscillations. These unbounded oscillations can be suppressed with sufficient mesh refinement (Huyakorn, 1977; Huebner et al., 2001) such that the elemental mass diffusion Péclet number (Pe_L) (3.32) (a non-dimensional ratio of the rate of mass advection to the rate of mass diffusion) does not exceed 2 (Huebner et al., 2001; Versteeg and Malalasekera, 2007),

$$\text{Pe}_L = \frac{UL}{D} = \text{Re}_L \cdot \text{Sc} \quad (3.32)$$

where U is the magnitude of the flow velocity, D is the mass diffusivity and L is a characteristic size of the element of interest. Gresho and Lee (1979) are proponents of the use of conventional the Bubnov-Galerkin FEM for advection-dominated problems. Gresho and Lee (1979) argue that a mesh which exhibits non-physical solution oscillations is simply insufficiently refined for the application. While this is a valid observation, it is not practical, since the mesh refinement required to satisfy the Péclet number condition ($\text{Pe}_L < 2$) of the conventional Bubnov-Galerkin FEM may lead to formidably large systems of ODEs that have unreasonable computational requirements to solve numerically.

An alternative method to reduce the numerical instability of the advective term of the BVP is a streamline-upwind Petrov-Galerkin (SUPG) FEM (Brooks and Hughes, 1982). The Petrov-Galerkin method of weighted residuals selects weight functions from a different class than the interpolation functions ($\mathbf{W} \neq \mathbf{N}$). The SUPG weight functions place more weighting on the element upwind of the node of interest than the element downwind (Donea and Huerta, 2003). While sufficient upwinding ensures a solution free of unbounded spurious oscillations, it also adds artificial diffusion which degrades the solution (Donea and Huerta, 2003).

The two-dimensional triangular element SUPG weight/interpolation functions suggested by Huyakorn (1977) for the steady-state solution of the ADE were implemented. Huyakorn (1977) introduced the upwinding weight parameter α , which assumes a value between 0 and 1. An upwinding weight of $\alpha = 1$ corresponds to full-upwinding, and

$\alpha = 0$ is the conventional Bubnov-Galerkin scheme. Optimal values of α are known for steady-state one-dimensional transport problems. However, there is no consensus for the optimal upwinding parameter selection in transient, two-dimensional transport problems on a triangular mesh. To ensure a stable solution, full-upwinding ($\alpha = 1$) was enforced, with no regard given to the artificial diffusion that was introduced by this assumption.

3.2.1.4 Boundary Conditions

The terms of the boundary condition vector $\mathbf{f}_\Gamma^{(e)}$ (3.31) can be estimated from the boundary conditions given by the ADE (3.3). Table 3.1 gives a summary of the boundary condition, flow condition and the resulting boundary condition vector ($\mathbf{f}_\Gamma^{(e)}$) for each boundary type of the discretized domain: inlets ($\Gamma_{\text{inlet}}^{(e)}$), outlets ($\Gamma_{\text{outlet}}^{(e)}$) and walls ($\Gamma_{\text{wall}}^{(e)}$).

Table 3.1: Summary of advection-diffusion equation boundary conditions

Boundary Type	Flow Condition on Boundary	Advection-Diffusion Equation Boundary Condition	Boundary Condition Vector
Inlet	$\mathbf{u} \cdot \hat{\mathbf{n}} < 0$	$(c\mathbf{u} - \mathbf{D}\nabla c) \cdot \hat{\mathbf{n}} = 0$	$\mathbf{f}_{\Gamma_{\text{inlet}}}^{(e)} = \text{zero-vector}$
Wall	$\mathbf{u} \cdot \hat{\mathbf{n}} = 0$	$(\mathbf{D}\nabla c) \cdot \hat{\mathbf{n}} = 0$	$\mathbf{f}_{\Gamma_{\text{wall}}}^{(e)} = \text{zero-vector}$
Outlet	$\mathbf{u} \cdot \hat{\mathbf{n}} > 0$	$(\mathbf{D}\nabla c) \cdot \hat{\mathbf{n}} = 0$	$\mathbf{f}_{\Gamma_{\text{outlet}}}^{(e)} = \int_{\Gamma_{\text{outlet}}^{(e)}} \mathbf{W}(c\mathbf{u} \cdot \hat{\mathbf{n}}) d\Gamma$

From Table 3.1 it can be seen that the outlet boundary has the only non-zero boundary condition vector ($\mathbf{f}_{\Gamma_{\text{outlet}}}^{(e)}$). However, the outlet boundary condition vector ($\mathbf{f}_{\Gamma_{\text{outlet}}}^{(e)}$) depends on the unknown pollutant concentration field (c). The spatial distribution of the flow velocity vector field ($\mathbf{u}$) and pollutant concentration field (c) can be approximated

along an elemental boundary with the interpolation functions (3.14) and (3.16), respectively. Substitution of these approximations of pollutant concentration (c) and flow velocity vector field ($\mathbf{u}$) into the outlet boundary condition vector ($\mathbf{f}_{\Gamma_{\text{outlet}}}^{(e)}$) gives (3.33).

$$\mathbf{f}_{\Gamma_{\text{outlet}}}^{(e)} = \mathbf{f}_{\Gamma_{\text{outlet}}}^{(e)}(t; \mathbf{x}) = \int_{\Gamma_{\text{outlet}}^{(e)}} (\mathbf{W} \hat{\mathbf{n}}^T \mathbf{U}^{(e)} \mathbf{N} \mathbf{N}^T) d\Gamma \mathbf{c}^{(e)} \quad (3.33)$$

For compact notation, the general elemental boundary condition vector $\mathbf{f}_{\Gamma}^{(e)}$ can be evaluated with (3.34) as a matrix-vector product of a square elemental boundary stiffness matrix ($\mathbf{K}_{\Gamma}^{(e)}$), and the element's nodal pollutant concentration vector ($\mathbf{c}^{(e)}$),

$$\mathbf{f}_{\Gamma}^{(e)} = \mathbf{K}_{\Gamma}^{(e)} \mathbf{c}^{(e)} \quad (3.34)$$

where $\mathbf{K}_{\Gamma}^{(e)}$ is the square elementary boundary stiffness matrix given by (3.35).

$$\begin{aligned} \mathbf{K}_{\Gamma}^{(e)} &= \mathbf{K}_{\Gamma}^{(e)}(t) = \int_{\Gamma_{\text{outlet}}^{(e)}} \mathbf{W} \hat{\mathbf{n}}^T \mathbf{u}^{(e)} \mathbf{N} \mathbf{N}^T d\Gamma \quad \text{for outlets} \\ \mathbf{K}_{\Gamma}^{(e)} &= \text{zero} - \text{matrix} \quad \text{for inlets/walls/element interfaces} \end{aligned} \quad (3.35)$$

Substituting (3.35) into (3.28) gives (3.36) the modified the system of first-order ODEs (Donea and Huerta, 2003),

$$\mathbf{C}^{(e)} \dot{\mathbf{c}}^{(e)} + \bar{\mathbf{K}}^{(e)} \mathbf{c}^{(e)} = \mathbf{C}^{(e)} \mathbf{s}^{(e)} \quad (3.36)$$

where $\bar{\mathbf{K}}^{(e)} = \mathbf{K}^{(e)} - \mathbf{K}_{\Gamma}^{(e)}$. The coefficients of this modified system of elemental ODEs are obtained by evaluating the elemental integrals within $\mathbf{C}^{(e)}$ (3.29), $\mathbf{K}^{(e)}$ (3.30), and $\mathbf{K}_{\Gamma}^{(e)}$ (3.35).

3.2.1.5 *Elemental Integration*

To obtain the coefficients of the elemental ODEs (3.36), the elemental integrals within capacitance $\mathbf{C}^{(e)}$ (3.23), stiffness $\mathbf{K}^{(e)}$ (3.24), and boundary stiffness $\mathbf{K}_\Gamma^{(e)}$ (3.35) matrices must be evaluated. Evaluating these integrals analytically may be quite tedious, especially if the boundaries of the element being integrated over are not aligned with the spatial coordinate system of the BVP, such as a triangular element. However, since these are integrals are bounded by the element, they can be approximated with a numerical quadrature algorithm (Press et al., 1997). Huebner et al. (2001) presents the Gauss-Legendre quadrature algorithms appropriate for integrating $\mathbf{C}^{(e)}$, $\mathbf{K}^{(e)}$ and $\mathbf{K}_\Gamma^{(e)}$ over a triangular element.

3.2.1.6 *Global Assembly*

All of the individual elemental ODEs (3.36) can be assembled into a single global system of ODEs (3.37),

$$\mathbf{C}\dot{\mathbf{c}} + \bar{\mathbf{K}}\mathbf{c} = \mathbf{C}\mathbf{s} \quad (3.37)$$

where the coefficients of the global capacitance $\mathbf{C} = \mathbf{C}(t)$, stiffness $\bar{\mathbf{K}} = \bar{\mathbf{K}}(t)$ matrices are functions of time (t) and relate all the nodal pollutant concentrations $\mathbf{c} = \mathbf{c}(t; \mathbf{x})$ to the steady nodal emission rates $\mathbf{s} = \mathbf{s}(\mathbf{x})$. Due to the nature of the compact support of the elemental weight ($\mathbf{W}$) and interpolation ($\mathbf{N}$) functions, the global capacitance ($\mathbf{C}$) and stiffness ($\mathbf{K}$) matrices are sparse, and can take advantage of efficient

storage/computational routines. This global system of ODEs represent an IVP whose solution can also be numerically approximated.

3.2.2 Initial Value Problem

The evolution of the nodal pollutant concentrations ($\mathbf{c}$) can be determined by solving the IVP (3.37) created from the application of the FEM to the ADE (3.1). A numerical solution to this system of coupled, first-order ODEs can be obtained from a wide variety of methods such as the fractional step Runge-Kutta methods (Butcher, 2008), the multistep Adams-Bashforth and Adams-Moulton methods (Hairer et al., 1993), or the Rosenbrock methods (Rosenbrock, 1963) for stiff systems.

The nodal pollutant concentrations $\mathbf{c} = \mathbf{c}(t; \mathbf{x})$ can be discretized in time to give $\mathbf{c}_n = \mathbf{c}(t_n; \mathbf{x})$, where t_n is the n th time-step, such that $n \geq 0$. Future nodal pollutant concentrations $\mathbf{c}_{n+1}$ are predicted by the recursion formula of the ODE's numerical integration algorithm, where $\mathbf{c}_{n+1} = \mathbf{c}(t_{n+1}; \mathbf{x})$, $t_{n+1} = t_n + \Delta t$, and Δt is the time-step size. “The recursion formulas make it possible for the solution to be ‘marched out’ in time, starting from the initial conditions at time $[t = t_0]$ and continuing step by step until reaching the desired duration $[t = t_f]$ ” (Huebner et al., 2001).

3.2.2.1 The θ -Method

The θ -method (a member of the Runge-Kutta family of ODE numerical integrating methods) is a single-step, finite difference method for solving first-order ODEs. The recursive formula of the θ -method (3.38) uses a parameter θ to obtain a weighted average of $\dot{\mathbf{c}}$ at the start ($\dot{\mathbf{c}}_n$) and end ($\dot{\mathbf{c}}_{n+1}$) of a time-step,

$$\frac{\mathbf{c}_{n+1} - \mathbf{c}_n}{\Delta t} = \theta \dot{\mathbf{c}}_{n+1} + (1 - \theta) \dot{\mathbf{c}}_n \quad (3.38)$$

where $0 \leq \theta \leq 1$ and for the system of ODEs of interest (3.37), $\dot{\mathbf{c}}_n$ and $\dot{\mathbf{c}}_{n+1}$ are given by (3.39) and (3.40) respectively,

$$\dot{\mathbf{c}}_n = \mathbf{C}_n^{-1}(\mathbf{C}_n \mathbf{s} - \bar{\mathbf{K}}_n \mathbf{c}_n) \quad (3.39)$$

$$\dot{\mathbf{c}}_{n+1} = \mathbf{C}_{n+1}^{-1}(\mathbf{C}_{n+1} \mathbf{s} - \bar{\mathbf{K}}_{n+1} \mathbf{c}_{n+1}) \quad (3.40)$$

where the time discretized capacitance $\mathbf{C}_n = \mathbf{C}(t_n)$, $\mathbf{C}_{n+1} = \mathbf{C}(t_{n+1})$ and stiffness $\bar{\mathbf{K}}_n = \bar{\mathbf{K}}(t_n)$, $\bar{\mathbf{K}}_{n+1} = \bar{\mathbf{K}}(t_{n+1})$ matrices are derived from the flow field conditions at $t = t_n$ and $t = t_{n+1}$.

Choosing $\theta = 0.5$ for the recursive formula of the θ -method (3.38) gives a variant the Crank-Nicolson scheme (Crank and Nicolson, 1947), which is unconditionally stable and second order accurate in time (Huebner et al., 2001; Donea and Huerta, 2003). Other popular choices of θ are summarized in Table 3.2.

Table 3.2: Stability of θ -methods

θ	Method Name	Explicit / Implicit	Order of Accuracy	Stability
0	Forward Euler	Explicit	1 st	conditional
1/2	Crank-Nicolson (Crank and Nicolson, 1947)	Implicit	2 nd	unconditionally stable with oscillatory decay.
2/3	Galerkin (Zienkiewicz, 1971)	Implicit	1 st	unconditionally stable with oscillatory decay.
1	Backward Difference Euler	Implicit	1 st	unconditionally stable with monotonic decay.

The nodal pollutant concentration at a future time-step ($\mathbf{c}_{n+1}$) can be isolated from the recursive formula of the θ -method (3.38) and expressed as the solution to the linear system (3.41),

$$\mathbf{A}_{n+1}\mathbf{c}_{n+1} = \mathbf{B}_{n+1}\mathbf{c}_n + \mathbf{b}_{n+1} \quad (3.41)$$

where $\mathbf{A}_{n+1}$, $\mathbf{B}_{n+1}$, and $\mathbf{b}_{n+1}$ are given by (3.42), (3.43) and (3.44), respectively.

$$\mathbf{A}_{n+1} = \frac{1}{\Delta t}\mathbf{C}_{n+1} + \theta(\bar{\mathbf{K}}_{n+1}) \quad (3.42)$$

$$\mathbf{B}_{n+1} = \frac{1}{\Delta t}\mathbf{C}_{n+1} - (1 - \theta)\mathbf{C}_{n+1}\mathbf{C}_n^{-1}\bar{\mathbf{K}}_n \quad (3.43)$$

$$\mathbf{b}_{n+1} = \mathbf{C}_{n+1}\mathbf{s} \quad (3.44)$$

The evolution of the time discretized nodal pollutant concentration ($\mathbf{c}_n$) can be modelled by recursive application of the time-stepping formula (3.41) from the initial nodal

pollution concentration $\mathbf{c}_0$ whose value is given by the initial condition (3.2) of the original PDE.

3.3 Dispersion Model

For compact notation, the time-stepping formula (3.41) which models the evolution of the nodal pollutant concentration, can be represented by a single system of linear equations (2.1) (shown in Section 2.1 and reproduced below),

$$\mathbf{A}\mathbf{c} = \mathbf{b} \quad (2.1)$$

where $\mathbf{c} = \mathbf{c}(\mathbf{x})$ is a column vector given by (3.45), and comprises the column vectors of time discretized nodal pollutant concentrations ($\mathbf{c}_n$) for all time-steps ($0 \leq n \leq n_t$).

$$\mathbf{c} = \mathbf{c}(\mathbf{x}) = [\mathbf{c}_0(\mathbf{x}) \quad \cdots \quad \mathbf{c}_{n_t}(\mathbf{x})]^T \quad (3.45)$$

The square coefficient matrix $\mathbf{A}$ has a block structure given by (3.46),

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_1 & \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 & \mathbf{A}_2 & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{n_t-2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{n_t-1} & \mathbf{A}_{n_t-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_{n_t} & \mathbf{A}_{n_t} \end{bmatrix} \quad (3.46)$$

where the sparse square matrices $\mathbf{A}_n$ and $\mathbf{B}_n$ are given by (3.42) and (3.43) respectively, and $\mathbf{I}$ and $\mathbf{0}$ are the identity and zero-matrix respectively, with the same dimension as $\mathbf{A}_n$ or $\mathbf{B}_n$. The forcing vector $\mathbf{b}$ is a column vector with the structure given by (3.47).

$$\mathbf{b} = \mathbf{b}(\mathbf{x}) = [\mathbf{b}_0(\mathbf{x}) \quad \cdots \quad \mathbf{b}_{n_t}(\mathbf{x})]^T \quad (3.47)$$

The formula for evaluating $\mathbf{b}_n$ is given by (3.44) and is only applicable for $n > 0$. $\mathbf{b}_0$ is chosen to assume the value of the initial conditions ($\mathbf{b}_0 = \mathbf{c}_0$).

This compact representation of the dispersion model (2.1) is convenient for determining the functional form of the sensitivity of the dispersion model to the emission source parameters $\mathbf{x}$, as discussed in Section 2.4. It should be noted that although (2.1) suggests that the sparse matrix $\mathbf{A}$ should be inverted, this action is not advisable, as $\mathbf{A}^{-1}$ is most likely a dense matrix that is too large to store in the memory of modern computers. Also, the coefficients of $\mathbf{A}^{-1}$ may be skewed by numerical errors. For some cases pre-computation of the LU factorization of $\mathbf{A}$ may greatly reduce the computational effort required to solve (2.1), however this approach also is not advised for this application since the LU factorization of $\mathbf{A}$ will most likely produce dense block matrix that is also too large to store in computer memory. An incomplete LU factorization will produce a sparser LU decomposition of $\mathbf{A}$, by dismissing matrix entries below a desired threshold, but this method introduces additional numerical modelling errors and may not provide a significant reduction in computational complexity. Instead, it is recommended that the future nodal pollution concentrations ($\mathbf{c}_{n+1}$) are determined with the time-stepping formula given by (3.41).

Chapter 4

Inverse Dispersion Model Setup

The applicability of inverse micro-scale atmospheric dispersion modelling for the location and quantification of fugitive methane emissions in the upstream oil and gas industry was studied with synthetic test cases. This chapter reviews the parameters of the IDM, the conditions of the synthetic test cases, and the methods of evaluating the quality the candidate emission sources.

A simplified three-dimensional model of an actual gas processing plant was obtained. This model has previously been used to construct airshed domains for forward and inverse dispersion model studies (Brereton, 2010; Joynes et al., 2011; Brereton and Johnson, 2011; Joynes et al., 2012; Carol A. Brereton and Johnson, 2012; Carol A Brereton and Johnson, 2012). For the proof-of-concept studies presented in this thesis, the airshed domain was constrained to two-dimensions (as shown in Figure 4.1) to reduce the computational complexity of the IDM.

Figure 4.1: Two-dimensional airshed domain of an industrial gas processing plant

The airshed domain depicted in Figure 4.1a spans an area of approximately 2 km^2 . The red box highlights the region of the domain occupied by the buildings (shown in grey), which is shown at higher magnification in Figure 4.1b. This airshed domain was constructed with guidelines suggested by Franke et al. (2007), who suggested that a three-dimensional domain for urban wind modelling should be sized such that the inlets and outlets are at a distance of at least $6H$ and $15H$ respectively from the buildings, where H is the height of the buildings. For a two-dimensional model, the building height (H) has no influence on the flow field. However, the depicted domain was conservatively sized such that the inlet and outlets were a distance of at least $45\bar{H}$ from buildings, where $\bar{H} \approx 10 \text{ m}$ (Brereton, 2010) is the average height of buildings of the modelled gas plant (neglecting the two 122 m tall stacks). Brereton (2010) studied the same gas plant model with a three-dimensional domain that spanned an area of 1 km^2 , and found that it was a sufficient size for synthetic fugitive emissions studies.

Since the dispersion of a pollutant was assumed to have no influence on the flow field, the IDM was decoupled from the numerical solution of the flow equations (Navier-Stokes and RNG k - ϵ turbulence model). This model decoupling required the pre-computation of the synthetic atmospheric flow fields such as the velocity vector field ($\mathbf{u}$) and turbulent eddy kinematic viscosity (ν_t), which were obtained using computational fluid dynamics (CFD) modelling.

4.1 CFD Modelling

A synthetic two-dimensional flow field was simulated with known inlet boundary conditions. Operational implementations of the IDM will further require estimations of the atmospheric flow field from limited anemometer observations, although this was beyond the scope of the current work. Unsteady Reynolds-averaged Navier-Stokes (URANS) CFD modelling was conducted on isothermal (25°C) air flow within the two-dimensional domain depicted in Figure 4.1 with ANSYS CFX, a commercial CFD software package. Synthetic wind velocity ($\mathbf{u}$) and turbulent eddy kinematic viscosity (ν_t) fields were obtained for fifteen minutes of physical time. CFD modelling required the airshed domain to be discretized into an assembly of geometrically simple control volumes, within which a numerical approximation of the flow field was obtained. Key input parameters for the CFD simulations are separately discussed in the following sections.

4.1.1 Spatial Domain Discretization

The two-dimensional airshed domain depicted in Figure 4.1 was discretized with ANSYS ICEM CFD, a mesh generation software package. As recommended by the “ANSYS CFX-Solver Modeling Guide” (ANSYS, 2006b), the two-dimensional domain was extruded by one cell to create a three-dimensional analogue of the airshed domain suitable for CFD modelling with ANSYS CFX. The computational domain was discretized into an assembly of triangular prisms with refinements near the buildings, where the steepest flow field gradients were expected. Figure 4.2 shows the discretized airshed domain at four magnifications, to highlight the resolution of the mesh.

Figure 4.2: Discretized computational airshed domain

This mesh was 14 m thick in the lateral direction, and contained 174,772 triangular prisms with 183,394 nodes. A two-dimensional horizontal cross-section through the discretized airshed domain contains 174,772 triangles and 91,697 nodes. For unit completeness of physical quantities such as concentration (c , [$\text{kg}\cdot\text{m}^{-3}$]) and volumetric emission rate (s , [$\text{kg}\cdot\text{s}^{-1}\cdot\text{m}^{-3}$]), the two-dimensional triangular domain was considered to have unit thickness.

In addition to the resolution, the shape of the triangular elements also determines the quality of the mesh, with equilateral triangles being ideal. The quality of the triangular mesh is quantified in Appendix B, and was determined to be sufficient for the current application.

4.1.2 Boundary Conditions

4.1.2.1 Inflow

Realistic inflow wind velocity data were obtained from the National Wind Technology Centre M2 Tower in Colorado (NREL, 2011). This wind tower measures the wind speed and direction at 2, 5, 10, 20 50 and 80 m elevations, and average and standard deviation data are recorded for one minute intervals. The summary statistics of the probe located at the 2 m elevation were used to generate fifteen minutes of synthetic wind at 2 second increments. The synthetic wind speed and direction was used as the inflow boundary condition for the wind simulation in ANSYS CFX. Figure 4.3 shows the synthetic time-varying wind speed and direction. These data are summarized in a wind rose plot shown in Figure 4.4, which indicates the prevailing wind speed and direction of the synthetic wind were predominantly from the south.

Figure 4.3: Time-variation of inflow wind speed and direction

Figure 4.4: Wind rose of inflow wind conditions

4.1.2.2 Outflow

The boundary conditions of the outflow surfaces of the domain were set to a constant zero gauge static pressure. With this “... outflow boundary condition the derivatives of all flow variables are forced to vanish, corresponding to fully developed flow. Therefore this boundary should be ideally far enough away from the built area to not have any fluid entering into the computational domain ...” (Franke et al., 2007).

4.1.2.3 Buildings Surfaces

On the boundaries of buildings, Franke et al. (2007) recommend using the no-slip boundary condition for the wind velocity, and wall functions to estimate the shear stress at smooth walls. The wall function approach reduces the mesh refinement requirements, by assuming a logarithmic flow velocity profile in the direction normal to a wall (Franke et al., 2007). Since there was little interest in accurate prediction of the flow boundary layers around the buildings, the wall function approach was deemed appropriate in the presented CFD studies.

4.1.2.4 Lateral Surfaces

As recommended by the “ANSYS CFX-Solver Modeling Guide” (ANSYS, 2006b), symmetry boundary conditions were applied to the lateral surfaces of the airshed domain to achieve an accurate two-dimensional solution of the Navier-Stokes equations.

4.1.3 Initial Conditions

The initial conditions to the transient wind simulation were supplied from a prior pseudo-transient simulation in which the inflow conditions were fixed to the corresponding initial inflow conditions of the subsequent fully transient simulation. Because of the inherent unsteadiness in the flow associated with vortex shedding downstream of the buildings, a converged steady-state solution to the prior steady-inlet simulation with a maximum relative error less than 5×10^{-4} (the maximum recommend relative error (ANSYS, 2006b)) could not be expected. Instead, the flow field solution obtained after 300 iterations (or pseudo time-steps) was used as the initial conditions to the subsequent transient wind simulation. Appendix C.1 shows the convergence history of the initial steady-inlet simulation reported by ANSYS CFX. Figure 4.5a depicts the spatial distribution of the flow streamlines of the initial conditions, revealing significant recirculations within the wakes of buildings as anticipated. Figure 4.5b depicts the spatial distribution of the turbulent eddy kinematic viscosity (ν_t) on a logarithmic scale.

Figure 4.5: Initial conditions of flow field from ANSYS CFX

The turbulent eddy kinematic viscosity (ν_t) of mixing wakes downstream of the buildings apparent in Figure 4.5b does not extend to the outlet boundaries of the domain, which suggests that the 300 iterations used to evaluate the initial distributions of the flow variables was shorter than ideal. Fortunately, this oversight did not significantly affect the results of the atmospheric dispersion model, since the turbulent mixing wakes are developed in the regions close to the building where the receptors and emission sources are expected, and these wakes will continue to develop ahead of the plumes from candidate emission sources during the transient simulations.

4.1.4 Transient Simulation

As noted above, the fifteen minute transient simulation was initialized from the flow field predicted by the prior steady-inlet simulation. A time-step of 0.2 seconds was selected and a residual error target of 1×10^{-5} was enforced on the root mean square (RMS) of the relative residual error on the mass and momentum equations. Appendix C.2 shows the

convergence histories of the entire transient simulation. The resulting wind velocity ($\mathbf{u}$) and turbulent eddy kinematic viscosity (ν_t) fields were used by the IDM to predict the pollutant concentration fields from candidate emission sources.

4.2 Inverse Dispersion Model Parameters

The resulting velocity vector ($\mathbf{u}$) and turbulence eddy kinematic viscosity (ν_t) fields predicted by ANSYS CFX were exported at one second increments in the Enight Gold ASCII format for further processing by the IDM. To adequately model the transport of a non-reacting pollutant within the airshed domain, it was necessary to estimate the distribution of the pollutant's isotropic mass diffusivity (D). The mass diffusivity of the pollutant was assumed to be driven by a combination of turbulent mixing and binary molecular diffusion. Turbulence driven mass diffusion was assumed to follow the eddy diffusivity model (turbulent eddy kinematic viscosity, ν_t) with rescaling by the turbulent Schmidt number (Sc_t), as discussed in Section 3.1.2. The binary molecular mass diffusivity of methane (the pollutant of interest) in air was estimated to be $D_{AB} \approx 2.2 \times 10^{-5} \text{ m}^2 \cdot \text{s}^{-1}$ using the Chapman-Enskog kinetic theory (Turns, 1996). The total effective mass diffusivity (D) was assumed to be the sum of these two mass diffusivities.

The flow velocity ($\mathbf{u}$) and mass diffusivity (D) fields were used to construct the global capacitance ($\mathbf{C}_n$) and stiffness ($\bar{\mathbf{K}}_n$) matrices of a finite element atmospheric dispersion model with the formulas presented in Section 3.2, and ultimately the system of linear equations of the dispersion model (see Section 3.3), which predicted the pollutant concentration fields resulting from candidate emission sources. The IDM used synthetic

receptors to extract receptor observations from the pollutant concentration field, which were then used to evaluate the objective function and its gradient. Candidate emission sources that minimized the objective function were generated by the L-BFGS-B numerical optimization algorithm. Several IDM parameters were fixed for all the presented studies including the flow field, the initial assumed emission source distribution, the background emission source distribution, and the L-BFGS-B control parameters. These are reviewed in the following sections.

4.2.1 Initial Candidate Solution

Numerical optimization algorithms require an initial candidate solution ($\mathbf{x}_0$) that can be iteratively improved. For the presented studies, an initial distribution of no emission sources was assumed. Thus, for the chosen parameterization of the emission source function $s(\mathbf{r}, t; \mathbf{x}_k)$, the initial values of the emission source parameters ($\mathbf{x}_0$) were all zero.

4.2.2 Background Emission Sources

The regularization term of the objective function (2.2) penalizes emission sources that deviate from an assumed prior distribution known as a background ($\mathbf{s}^b$). Since it was assumed that there were no known emission sources within the domain, the background emission source distribution ($\mathbf{s}^b$) was chosen to be the zero-vector. The IDM would allow for the detection of anomalies (i.e. the presence of unknown emission sources) that deviated from this assumption. This also matched Ide et al. (1997) recommendation of

choosing the emission source background to match the initial estimate of emission source distribution used by a numerical optimization algorithm such that $\mathbf{s}^b = \mathbf{s}(\mathbf{x}_0)$.

4.2.3 L-BFGS-B Parameters

The box constraint variant (L-BFGS-B) (Byrd et al., 1994) of the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) (Nocedal, 1980) numerical optimization algorithm was used to generate candidate distributions of emission sources. There are a number parameters that determine the behaviour of the L-BFGS-B algorithm, such as the number of update histories (m) used to approximate the inverse Hessian matrix ($\mathbf{H}_k$) the boundary constraints on the search space of the optimization parameters within $\mathbf{x}$, and the termination criteria. These are discussed below.

4.2.3.1 *Inverse Hessian Approximation*

The L-BFGS-B numerical optimization algorithm approximates the inverse Hessian matrix $\mathbf{H}_k$ from a limited history of the m previous updates to the optimization parameters $\mathbf{x}_k$ and objective function gradients $\nabla f(\mathbf{x}_k)$. Nocedal and Wright (2006) suggest that between 3 and 20 previous updates are sufficient to approximate the inverse Hessian matrix. Liu and Nocedal (1989) and Gilbert and Lemaréchal (1989) observed that increasing m beyond 5 had little improvement on the convergence of the L-BFGS algorithm relative to the increase in storage requirements. For this reason the inverse

Hessian matrix was approximated using 5 update histories in all IDM studies presented in this thesis.

4.2.3.2 Box Bound Constraints

The L-BFGS-B numerical optimization algorithm accommodates the optional enforcement of simple box boundary constraints on the search space of the objective function's parameters $\mathbf{x}$. The algorithm offers four types of boundary constraints on the search space of each optimization parameter, which are identified with the integer-valued vector $\mathbf{nbd}$. Each entry nbd_i of $\mathbf{nbd}$ classifies the boundary type of the search space of the corresponding optimization parameter x_i of $\mathbf{x}$. A listing of the boundary types offered by the L-BFGS-B algorithm is presented in Table 4.1.

Table 4.1: L-BFGS-B parameter search space boundary types

nbd_i	Description
0	The search space of x_i is unbounded.
1	The search space of x_i is only lower bounded.
2	The search space of x_i is lower and upper bounded.
3	The search space of x_i is only upper bounded.

The values associated with these optional boundaries are defined by a vector of the lower boundaries ($\mathbf{lb}$) and a vector of upper boundaries ($\mathbf{ub}$) of the parameter search space, where lb_i and ub_i are the respective lower and upper bounds of the search space for the parameter x_i . Given that the optimization parameters $\mathbf{x}$ represent the volumetric emission rate at the nodes of the discretized domain, only a lower boundary ($\mathbf{nbd} = \text{one-vector}$) of a zero volumetric emission rate was enforced on the search space of all parameters. This

search space constraint assumed that there were no sinks (i.e. negative sources) within the airshed domain. Since no upper limit was imposed on the volumetric emission rates, the IDM was free to predict emission sources of any strength.

4.2.3.3 Termination Criteria

Under normal operational conditions there are three parameters that control the termination of the L-BFGS-B numerical optimization algorithm: the relative reduction of the objective function (*factr*), the relative reduction of the maximum norm of the projected objective function gradient (*pgtol*), and the maximum number of iterations (*maxit*).

Objective Function Reduction (factr)

The L-BFGS-B numerical optimization algorithm will terminate if there is an insufficient reduction in the value of the objective function between consecutive iterates. This occurs when the inequality (4.1) is satisfied (Zhu et al., 1997).

$$\frac{(f(\mathbf{x}_k) - f(\mathbf{x}_{k+1}))}{\max(|f(\mathbf{x}_{k+1})|, |f(\mathbf{x}_k)|, 1)} \leq factr \cdot epsmch \quad (4.1)$$

In Eq. (4.1), *factr* is a non-dimensional user control parameter that sets the threshold of what is considered a significant relative reduction of the objective function, and *epsmch* represents the machine precision as predetermined by the L-BFGS-B algorithm (Zhu et al., 1997) (*epsmch* $\approx 2.22 \times 10^{-16}$ for the presented studies).

For the current simulations, it was deemed undesirable for the algorithm to prematurely terminate due to a slowly reducing objective function. Thus, this termination test was practically disabled by setting $factr = 0$. The L-BFGS-B numerical optimization algorithm could then only terminate due to this condition “...if the objective function remains unchanged after one iteration” (Zhu et al., 1997).

Projected Gradient Reduction (pgtol)

The L-BFGS-B algorithm will also terminate when the maximum norm of the projected gradient of the objective function ($\nabla_{\text{proj}} f(\mathbf{x}_{k+1})$) (see Section 4.4.2) is below the user controlled threshold $pgtol$ (4.2),

$$\|\nabla_{\text{proj}} f(\mathbf{x}_{k+1})\|_{\infty} \leq pgtol \quad (4.2)$$

where $\|\cdot\|_{\infty}$ is the maximum norm function (also known as the uniform norm, supremum norm, the Chebyshev norm, or the infinity norm) defined as (4.3).

$$\|\mathbf{a}\|_{\infty} = \max \{|a_1|, \dots, |a_n|\} \quad (4.3)$$

Zhu et al. (1997) comment that this termination criteria is difficult to invoke if $pgtol$ is less “...than square root of the machine precision”. It was also deemed undesirable for the algorithm to prematurely terminate due to a shallow projected gradient. Thus, this termination criteria was also effectively disabled by setting $pgtol = 0$ (Zhu et al., 1997). Under this condition this termination criteria would only be invoked if the projected gradient of the objective function was the zero-vector, which would correspond to a local minimum of the objective function within the boundaries of the parameter search space.

Maximum Iterations (maxit)

The last of the standard L-BFGS-B termination controls is the maximum number of optimization iterations, which is set by the user control parameter *maxit*. Preliminary convergence studies showed that 200 iterations were sufficient to converge to a consistent estimate of emission sources, and this criterion was enforced for all IDM studies presented in this thesis.

4.2.4 IDM Parameter Summary

The IDM parameter settings that were implemented for all cases discussed in this thesis are summarized in Table 4.2.

Table 4.2: Summary of IDM parameter settings

Parameter	Description
$\mathbf{x}_0$ = zero-vector	No emission sources case used for the initial candidate emission source
$\mathbf{s}^b$ = zero-vector	No emission sources case used for the background emission source
$m = 5$	Number of update histories used to approximate the inverse Hessian matrix.
$\mathbf{nbd}$ = one-vector	Search space of all emission source parameters was only lower bounded.
$\mathbf{lb}$ = zero-vector	Lower bound of zero on the search space of all emission source parameters.
$\mathbf{ub}$ = undefined	No upper bound on the search space of any emission source parameter.
$f_{actr} = 0$	Terminate if the value of objective function remains unchanged over consecutive iterations.
$p_{gtol} = 0$	Terminate if converged to a local minimum within the acceptable search space of the objective function.
$maxit = 200$	Terminate after 200 iterations.

4.3 Inverse Dispersion Model Test Cases

The synthetic, two-dimensional flow field obtained from the CFD modelling of wind within the gas plant was used to generate synthetic receptor observations from known

distributions of single and multiple emission sources. Using only these synthetic receptor observations and the flow field, emission source reconstructions were attempted. The robustness of the IDM was evaluated with tests in which simulated measurement noise was added to the synthetic receptor observations. Objective function regularization and emission source filtering was also studied in an attempt to improve the quality of candidate emission sources predicted by the IDM.

For the presented test cases, 200 optimization iterations were performed on the candidate emission sources with the L-BFGS-B numerical optimization algorithm. Each iteration required a solution of a pollution transport model (see Section 3.3) with the candidate emission source to evaluate the objective function, and a solution of an adjoint sensitivity transport model (see Section 2.4.3) to efficiently evaluate the gradient of the objective function. These computations represented the bulk of the computational effort, and for the presented implementation, the test cases required approximately one month to complete 200 optimization iterations. However as shown in Chapter 5, reasonable emission source distributions were obtained within 50 iterations (approximately one week of computational time).

4.3.1 Receptor Locations

For each of the test cases presented in this thesis, twenty-five synthetic receptors were distributed within the computational domain. The locations of these receptors were chosen such that they aligned with nodes of the discretized two-dimensional airshed

domain. The positioning of the synthetic receptors relative to the buildings is shown in Figure 4.6 as black dots.

Figure 4.6: Synthetic receptor locations

The exact location of these synthetic receptors within the two-dimensional airshed domain is listed in Appendix D. These synthetic receptors sampled the pollutant mass concentration field as point measurements in space and time at one second intervals, which aligned with the time-steps of the dispersion model and was implemented with the sparse observation matrix $\mathbf{H}$ as discussed in Section 2.2.

The optimal placement of the receptors for a given flow condition or the cost versus benefits of additional receptors was not considered within the scope of this thesis. These are significant design problems that will need to be addressed for operational applications of inverse dispersion modelling for fugitive emissions. However, previous studies seem to suggest that 31 (Brereton, 2010; Carol A Brereton and Johnson, 2012) and 64 (Whitenett et al., 2003) receptors are suitable.

4.3.2 Single-Source Fugitive Emissions Test Case

The first fugitive emissions test case considered a single, steady emission source, shown in Figure 4.7 centred in the hollow black circle labeled A. This emission source aligned with a node of the discretized domain, and was chosen for its favourable receptor coverage for the wind direction (see Section 4.1.2.1). This synthetic emissions source had a constant emission rate of $Q^* = 4.5$ g/s, which was classified as a “large” emission source by Brereton (2010), and is typical in the oil and gas industry (based on a private communication with David Picard, Clearstone Engineering Ltd.). In reality, it is not possible to numerically simulate a point emission source within a discretized domain. Instead, an equivalent volumetric emission source was defined as a tent function over a cluster elements as shown in Figure 4.7b. Synthetic observations were generated for the twenty-five receptors for the flow conditions presented in Section 4.1.

Figure 4.7: Specified known distribution of the single-source test case

4.3.3 Multiple Source Fugitive Emissions Test Case

A ten source test case was also studied as shown in Figure 4.8, where emission sources identified by black hollow circles labeled A – J were each modelled with a constant emission rate of 5 g/s, for a combined total emission rate of $Q^* = 50$ g/s. Appendix E lists the locations and emission rates of these ten emission sources. Similar to the single-source test case, the volumetric emission rate of the ten sources were each defined as a tent function over a small cluster of adjacent elements and synthetic observations were computed for the twenty-five receptors using the identical flow conditions presented in Section 4.1.

Figure 4.8: Multiple volumetric source position, 2.33× magnification

4.3.4 Receptor Measurement Noise

Operational applications of IDM contain modelling errors that reduce the effectiveness of the inverse model in determining emission sources. These errors can manifest as numerical modelling errors (e.g. truncation and round-off errors), errors in the initial data (e.g. initial concentration distribution), errors in the flow field parameters (e.g. wind and turbulence field estimates) (Penenko et al., 2002), and errors in the observations (e.g. instrument measurement errors). All of these sources of error influence the ability of an IDM to reliably identify emission sources. The robustness of the IDM was studied with synthetic observational measurement noise.

An atmospheric methane receptor is currently being designed by the Energy and Emissions Group at Carleton University (Schoonbaert and Johnson, 2011). This instrument will be able to detect atmospheric concentrations of methane using tunable diode laser absorption spectroscopy (TDLAS). The minimum standard deviation of the Gaussian distributed measurement noise of this TDLAS receptor is estimated to be $\sigma = 19.2 \text{ mg} \cdot \text{m}^{-3}$ and is expected to increase linearly by 1% of the measured pollutant concentration. To study the robustness of the IDM, synthetic Gaussian noise was added to the synthetic receptor observations. The Gaussian noise was generated with the same minimum standard deviation as the TDLAS receptor under development ($\sigma = 19.2 \text{ mg} \cdot \text{m}^{-3}$), but increased by 10% of the measured pollutant concentration, as a “worst-case” scenario. Figure 4.9 shows a sample of receptor observations from the single-source test case, with and without observational noise.

Figure 4.9: Sample of synthetic receptor observations for the single-source test case

Ten percent Gaussian noise was also added to the synthetic receptor observations of the multi-source test case. Figure 4.10 compares a sample of those receptor observations, with and without observational noise.

Figure 4.10: Sample of synthetic receptor observations for the multi-source case

Both Figure 4.9 and Figure 4.10 show that while the 10% Gaussian noise added small fluctuations to the synthetic receptor observations, the character of the large scale fluctuations of the receptor observations was still retained.

4.3.5 Objective Function Regularization

Objective function regularization was studied as a potential means to prevent the prediction of false or unnecessarily complex emission source distributions. The regularization term of the objective function penalized the deviation of emission sources from an assumed prior background distribution ($\mathbf{s}^b$). A regularization parameter β was introduced to weight the importance of minimizing the regularization term over the observational term of the objective function. The influence of the regularization parameter β on the emission source predictions by the IDM was studied for the range $\beta = [0, 1, 10, 100, 1000] \text{ s}^2 \cdot \text{m}^{-3}$, and the results are presented in Chapter 5.

4.3.6 Emission Source Filter

While objective function regularization provides a means of penalizing the prediction of false emission sources, an alternate method, categorized as emission source filtering, was also studied. Regions of the domain with poor observability were defined, such that candidate emission sources predicted in these regions could be deemed unreliable and subsequently ignored or filtered-out. The emission source filters were constructed from the field of view (FOV) of the receptor network. The receptor network's FOV was determined by treating the receptors as emission sources with constant unit emission rates. These unit emissions were transported from the receptors backwards in time, and against the flow with the continuous form of the adjoint advection-diffusion equation (4.4) (Keats et al., 2007b),

$$-\frac{\partial \lambda}{\partial t} - \mathbf{u} \cdot \nabla \lambda - \nabla \cdot (\mathbf{D}^T \nabla \lambda) = H \quad (4.4)$$

where λ is the continuous adjoint variable, and H is a function that models unit emission rates from the receptors. The derivation of the adjoint advection-diffusion equation and the associated initial and boundary conditions are shown in Appendix A.3. The adjoint variable (λ) can be considered as a retroplume (Issartel and Baverel, 2003) or an ensemble of all back-trajectories (Rao, 2007) from the receptor network. The adjoint advection-diffusion equation was numerically solved with a routine very similar to the forward dispersion model (see Chapter 3), with the exception that emissions were propagated backwards in time and against the flow direction. Pudykiewicz (1998) suggested a method for determining the PDF for the receptor network's FOV for observing an instantaneous emission source, which was later implemented by Penenko et

al. (2002). The Pudykiewicz (1998) receptor network FOV PDF was modified for use with a steady emission source as presented in (4.5).

$$p_{\text{FOV}} = \frac{\int_{t_0}^{t_f} \lambda \, dt}{\int_{\Omega} \int_{t_0}^{t_f} \lambda \, dt \, d\Omega} \quad (4.5)$$

The PDF of receptor network's FOV (p_{FOV}) for the presented flow conditions (see Section 4.1) and receptor arrangement (see Section 4.3.1) is shown in Figure 4.11a, with a logarithmic scale.

Figure 4.11: Receptor FOV PDF and source filters

Distributions of the adjoint variable in the outer 200 m perimeter of the domain revealed significant numerical errors due to the coarse spatial discretization, as shown in Figure 4.11b. Data in this outer region well away from the domain of interest surrounding the buildings were ignored and Figure 4.11c shows a reimplementation of the receptor network FOV PDF that has been constrained to the 1 km² central domain with sufficient mesh refinement. Contour lines for the 0.1, 1, 5, and 10th percentiles of the receptor network FOV PDF are superimposed on Figure 4.11c and Figure 4.11d. These contours were used to define different emission source filters, where regions with FOV coverage falling below the specified 0.1, 1, 5, and 10th percentile values of the PDF were disregarded. Emission source filtering was implemented as post-processing step to the candidate emission sources, although the filtering could also be implemented within the IDM by constraining both the upper (**ub**) and lower (**lb**) bounds of volumetric emission rates from mesh nodes outside the filter boundaries to zero. Results using the emission source filters are presented in Chapter 5 for both the single and multi-source test cases with and without observational noise.

4.4 Measures of Model Fitness

The quality of the emission sources predicted by the IDM can be quantified with a variety of techniques, such as the relative reduction of the objective function, the reduction of the norm of the objective function's gradient, the coefficient of determination (r^2) of the modelled receptor observations, and the total emission rate of the predicted emission

source distribution. These measures of model fitness are reviewed in the following sections.

4.4.1 Objective Function Reduction

The relative reduction of the objective function ($f(\mathbf{x}_k)/f(\mathbf{x}_0)$) gives a non-dimensional measure of the improvement of candidate emission sources over successive iterations. Alternatively, the convergence of only the observations can be studied by considering only the observational term of the objective function ($f^o(\mathbf{x}_k)/f^o(\mathbf{x}_0)$). These measures of model fitness are case specific and depend on the duration of the data assimilation, the number of receptors, and the initial emission source parameter guess $\mathbf{x}_0$, which were all fixed for the presented studies.

4.4.2 Gradient Norm Reduction

Analysis of the objective function's gradient $\nabla f(\mathbf{x}_k)$ is another measure of the quality of candidate emission sources. Gradient-based numerical optimization algorithms such as L-BFGS-B attempt to converge to a local minimum of the objective function such that the gradient is equal to the zero-vector. Deviations from this zero-vector can be quantified with a vector norm of the objective function's gradient such as the maximum norm ($\|\cdot\|_\infty$) (see (4.3)). However, if bounds are placed on the parameter search space then the numerical optimization algorithm may be prevented from converging to a local minimum, resulting in a non-zero vector norm of objective function's gradient. This can

incorrectly suggest that the candidate emission sources can be further improved when in fact the parameters are constrained by the search boundaries. The projected objective function gradient $\nabla_{\text{proj}}f(\mathbf{x}_k)$ mitigates this misinformation by ignoring the gradient components from parameters that are constrained by the bounds of their search space. The L-BFGS-B numerical optimization algorithm also uses this measure of model fitness as a termination criteria (see Section 4.2.3.3) (Zhu et al., 1997). Normalization of the norm of the objective function's gradient, by the norm of the gradient from the initial emission source guess ($\|\nabla f(\mathbf{x}_k)\|_\infty / \|\nabla f(\mathbf{x}_0)\|_\infty$ or $\|\nabla_{\text{proj}}f(\mathbf{x}_k)\|_\infty / \|\nabla_{\text{proj}}f(\mathbf{x}_0)\|_\infty$) provides a non-dimensional measure of the convergence of candidate emission sources. Similar to the reduction of the objective function, the value of gradient norm is case specific.

4.4.3 Coefficient of Determination

While the iterative reduction of the objective function and norm of its gradient each provide a sense of the improvement of candidate emission sources, their values are case specific and depend on a number of factors such as the number of receptors (n_r), the duration of the data assimilation ($[t_0, t_f]$) and the initial guess emission sources ($\mathbf{x}_0$). An alternate non-dimensional measure of model fitness that is not case specific is the coefficient of determination (r^2) of the model receptor observations and the candidate emission sources and is given by (4.6).

$$r^2(\mathbf{x}) = 1 - \frac{(\mathbf{H}\mathbf{c}(\mathbf{x}) - \mathbf{c}^0)^T (\mathbf{H}\mathbf{c}(\mathbf{x}) - \mathbf{c}^0)}{(\bar{\mathbf{c}}^0 - \mathbf{c}^0)^T (\bar{\mathbf{c}}^0 - \mathbf{c}^0)} \quad (4.6)$$

Similar to observational term of the objective function, the coefficient of determination penalizes deviations of the receptor observations ($\mathbf{H}\mathbf{c}(\mathbf{x})$) from a candidate distribution of emission sources from the known receptor observations ($\mathbf{c}^0$). Perfect agreement between the modelled and known receptor observations, will give $r^2 = 1$. It should be noted that it is possible for the coefficient of determination to be negative, which indicates that the average of the known observations ($\bar{\mathbf{c}}^0$) is a better fit to the known observations ($\mathbf{c}^0$) than the modelled observations ($\mathbf{H}\mathbf{c}(\mathbf{x})$) from a candidate distribution of emission sources.

4.4.4 Total Emission Rate

Finally, the total emission rate Q of a candidate distribution of emission sources also provides a sense of convergence. Normalization of the total emission rate of the candidate emission sources by known total emission rate Q^* gives a non-dimensional measure of error within the estimated emission rate of the sources. Unlike the previous measures of model fitness, this measure is only possible with a synthetic test case, where the total emission rate of the sources is known. The total emission rate is obtained by integrating the steady candidate volumetric emission rate profile $s(\mathbf{r}, t; \mathbf{x}_k)$ over the entire computational domain Ω (4.7).

$$Q = \int_{\Omega} s \, d\Omega = \sum_{i=1}^{n_e} \int_{\Omega_i^{(e)}} s \, d\Omega = \sum_{i=1}^{n_e} \int_{\Omega_i^{(e)}} \mathbf{N}^T \, d\Omega \, \mathbf{s}^{(e)} \quad (4.7)$$

For the presented studies this spatial integration was evaluated with the summation of elemental integrals of the piecewise continuous volumetric source emission rate function. The choice of elements and their shape functions $\mathbf{N}_i$ will dictate the elemental numerical integration scheme and required order.

Chapter 5

Inverse Dispersion Model Results

This chapter reviews the results of the IDM for the location and quantification of single and multiple fugitive emission sources. The sensitivity of the IDM to observational noise is also reviewed, and finally the mitigation of false emission sources and their contribution to the total emission rate estimate was studied with objective function regularization and emission source filtering. The parameters of these synthetic test cases are outlined in Section 4.3.

5.1 Single-Source Test Case Results

5.1.1 Predicted Emission Source Distribution

The IDM iteratively predicted candidate emission source distributions that minimized the objective function, and hence the error between the modelled and known receptor observations. Figure 5.1 shows a magnified view of the candidate emission source distributions predicted at the 25th, 50th, 100th and 200th iterations of the IDM for the single-source fugitive emissions test case, with and without observational noise, and without objective function regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$).

Figure 5.1: Iterative progression of candidate emission source distributions for the single-source test case without objective function regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)

Figure 5.1: Iterative progression of candidate emission source distributions for the single-source test case without objective function regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$) (cont.)

The peak of the candidate volumetric emission rate profiles converged to the exact location (same mesh node) as the known emission source (shown as a hollow black circle) at iteration 57 for the case with observational noise and iteration 59 without observational noise. Additionally, the spatial distributions of the candidate volumetric emission rate (quantified by the full area at half maximum, FAHM) became more concentrated with each iteration, giving more confidence in the estimate of the emission source location. The final (200th iteration) predicted candidate emission sources achieved $\text{FAHM} = 93 \text{ m}^2$ (14 elements), regardless of observational noise. This suggests that 10% observational noise did not adversely affect the ability of the IDM to accurately predict the correct location of the emission source. However, it should be noted that these final candidate emission sources were still more diffuse than the known emission source shown in Figure 4.7, which has $\text{FAHM} = 36 \text{ m}^2$ (6 elements).

Figure 5.2 shows the final (200th iteration) candidate emission source distributions over the entire airshed domain on a logarithmic scale, which highlights the simultaneous prediction of several weaker false emission sources. These false emission sources were mostly identified in regions of the domain that had poor observability for the receptor network and the wind direction.

Figure 5.2: Final emission source distributions (log scale) for the single-source test cases without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)

The general “keyhole shaped” outline of the false emission sources matches the bounds of the region of the domain with good receptor network coverage for the prevailing wind direction, as illustrated by the emission source filters (shown in Figure 4.11d). Candidate emission sources in the outer regions of domain have little influence on the receptor measurements and as such can exist at practically no penalty to the observational term of the objective function. The addition of Gaussian noise to the receptor observations encouraged the prediction additional weak false emission sources. While these false emission sources seem trivial because they are 1 to 2 orders of magnitude lower the

dominant emission source, they are spread over a significant area of the domain. It will become apparent in Section 5.1.3 that these false emissions sources can cause a significant overestimation of the total emission rate of the known emission source, when all predicted emission sources are integrated over the entire domain. Two different approaches were considered to mitigate the influence of these false emission sources on the estimate of the total emission rate, objective function regularization (introduced in Section 2.2) and emission source filtering (introduced in Section 4.3.6).

5.1.1.1 Objective Function Regularization

As discussed in Section 2.2, objective function regularization is an inverse modelling technique that assigns a penalty to optimization parameters that deviate from an assumed prior. Regularization as implemented by the IDM penalized candidate emission sources as an area weighted sum of the squared deviations from an assumed background ($\mathbf{s}^b$) of no emission sources. The regularization term was scaled by the positive regularization parameter β , which served as a unit converter and determined the relative importance of minimizing either the residual error of the receptor observations or the deviation of the candidate emission sources from the background. Figure 5.3 shows a magnified view of the final predicted (200th iteration) emission source distributions for the single-source test case with and without observational noise, and with the regularization parameter β chosen to be 0, 1, 10, 100 or 1000 $\text{s}^2\cdot\text{m}^{-3}$.

Figure 5.3: Effect of regularization on final candidate emission source distributions (linear scale) for the single-source test case

Figure 5.3: Effect of regularization on final candidate emission source distributions (linear scale) for the single-source test case (*cont.*)

Figure 5.3 shows that regularization encouraged the prediction of much more diffuse emission sources, quantified by the FAHM value noted on each figure panel. This trait is undesirable in inverse atmospheric dispersion modelling because it gives less certainty in the estimates of the emission source location. This trend of objective function regularization encouraging the prediction of diffuse candidate emission sources is even

more apparent when considering the entire domain on a logarithmic scale, as shown in Figure 5.4.

Figure 5.4: Effect of regularization on final candidate emission source distributions (log scale) for the single-source test case

Figure 5.4: Effect of regularization on final candidate emission source distributions (log scale) for the single-source test case (*cont.*)

In general, while objective function regularization did mitigate the prediction of some of the false emission sources by making them more diffuse, it also introduced additional weak false emission sources upwind of the known emission source location. Nevertheless, the nodal location of the peak candidate volumetric emission rate in Figure 5.3 and Figure 5.4 still exactly matched the nodal location of the known emission source (shown as a hollow black circle). A suitable range of the regularization parameter appears to be $0 \leq \beta < 10 \text{ s}^2 \cdot \text{m}^{-3}$ for the single-source test case, since this range offered a reasonable compromise between false emission source mitigation and diffusion of the true emission source prediction.

5.1.2 Convergence

5.1.2.1 Objective Function Reduction

Each iteration of the IDM generates a new candidate emission source distribution that further minimizes the objective function by improving the agreement between the modelled and known receptor observations. The regularization term of the objective function adds a penalty to emission source distributions that deviate from an assumed prior. Figure 5.5 shows the convergence history, as well as the influence of the regularization parameter (β) and the addition of Gaussian distributed observational noise on the minimization of the objective function normalized by its initial value ($f_{\text{no noise}}(\mathbf{x}_0) = 7.98 \times 10^{-6} \text{ kg}^2 \cdot \text{m}^{-6}$) for the single-source test case.

Figure 5.5: Convergence history of the objective function minimization for the single-source test case

Figure 5.5a shows the convergence of the full objective function ($f(\mathbf{x}_k)$, including both the observational and regularization term), while Figure 5.5b show the convergence of only the observational term ($f^o(\mathbf{x}_k)$) of the objective function. The minimization of the objective function was significantly influenced by the addition 10% observational noise and by the regularization parameter (β). Observational noise impeded the minimization of the objective function by introducing errors which were presumably unexplainable by any candidate steady emission sources. With added observational noise, the initial value of the objective function (i.e. calculated objective function value for an initial guess of no emission sources), was 59% higher than the corresponding value without added observational noise. The added receptor noise produced unexplainable errors that gave the observational term of the objective function a non-zero global minimum. This minimum ($f_{\text{min, 10% noise}}^o = 4.54 \times 10^{-6} \text{ kg}^2 \cdot \text{m}^{-6}$, shown as a grey dashed line in Figure 5.5b)

can be estimated by evaluating the objective function with noisy observations and the known emission source. Figure 5.5c shows the convergence history of the observational term of the objective function relative to its expected global minimum ($f^o(\mathbf{x}_k) - f_{\min}^o$). Figure 5.5c thus demonstrates that the observational term of the objective function is being minimized toward its relative global minimum, regardless of observational noise

Since the initial guess of no emission sources matched the assumed background ($\mathbf{s}^b$), the initial value of the objective function was not dependant on the regularization parameter (β) (i.e. $f(\mathbf{x}_0) = f^o(\mathbf{x}_0)$). Figure 5.5 also shows that the increasing the regularization parameter (β) of the objective function encouraged early convergence toward weaker emission sources that had reduced capacity of reproduce the known receptor observations, because the penalty of deviating from the assumed prior emission source distribution was too great. These same trends can be observed when considering the gradient of the objective function (see Appendix F) as a measure of the convergence of a candidate solution to a minima of the objective function.

5.1.2.2 Coefficient of Determination

While the objective function and its gradient (Appendix F) provide a general trend of a reducing residual error between the modelled and known receptor observations, their values are case specific and only provide a relative sense for how well candidate emission sources explain the known receptor observations. By contrast, the coefficient of determination (r^2) of the modelled receptor observations is a more descriptive measure of the quality of a candidate emission source distribution (see Section 4.4.3). Figure 5.6a

shows the convergence history of the coefficient of determination for the modelled receptor observations produced from the candidate emission source distributions for the single-source test case.

Figure 5.6: Convergence history of the coefficient of determination of modelled receptor observations for the single-source test case

A coefficient of determination of one ($r^2 = 1$) is an upper limit and indicates perfect agreement between modelled and known receptor observations. Figure 5.6a shows that after 200 iterations, the single-source cases without observational noise (solid lines) and limited regularization ($\beta \leq 1 \text{ s}^2 \cdot \text{m}^{-3}$) resulted in modelled receptor observations that were in almost perfect agreement with the synthetic known observations ($r^2 \geq 0.987$). Increasing the regularization parameter (β) of the objective function produced candidate emission source distributions whose receptor observations were less correlated with the known receptor observations.

The IDM was less successful in predicting candidate emissions sources that could explain the random fluctuations of the observational noise (dashed lines), resulting in reduced coefficients of determination relative to the cases with perfect observations. Similar to the estimated lower limit of the objective function (see Section 5.1.2.1), an expected upper limit of the coefficient of determination for the noisy observations can be estimated using the modelled receptor observations from the known emission source. For the single-source test case with observational noise this expected upper limit was $r_{\max, 10\% \text{ noise}}^2 = 0.627$, shown as a grey dashed line in Figure 5.6a. In principal, it is possible for a candidate emission source distribution to result in a coefficient of determination that exceeded this expected upper limit if it can reproduce some the random noise fluctuations in the known receptor observations. Figure 5.6b shows the convergence history of the coefficient of determination normalized by its expected maximum ($r_{\max, \text{no noise}}^2 = 1$ and $r_{\max, 10\% \text{ noise}}^2 = 0.627$). While the candidate emission sources predicted with noisy observations were unable to reproduce the random fluctuations at the receptors, they nonetheless produced candidate receptor observations that approached the maximum expected correlation value. In this manner, the IDM performed similarly well regardless of observational noise.

Finally, Figure 5.6 shows that increasing the regularization parameter (β) of the objective function lead faster convergence. However, the resulting emission sources had diminished capacity to explain the receptor observations, regardless of observational noise. From Figure 5.6 it is believed that a suitable regularization parameter range is $0 \leq \beta < 10 \text{ s}^2 \cdot \text{m}^{-3}$, since within this range the coefficients of determination approached the expected upper limits, regardless of observational noise.

5.1.3 Total Emission Rate

The total emission rate of a candidate emission source distribution was determined by integrating the volumetric source emission rate function $s(\mathbf{r}, t; \mathbf{x}_k)$ over the entire spatial domain (see Section 4.4.4). Figure 5.7 shows the convergence history of the total emission rate for the single-source test case, normalized by the known total emission rate ($Q^* = 4.5$ g/s).

Figure 5.7: Convergence history of the total emission rate for the single-source test case

Figure 5.7 shows that objective function regularization with perfect observational data (solid lines) was unable to significantly improve the overestimate of the total emission rate. While, minimal regularization ($\beta = 1 \text{ s}^2 \cdot \text{m}^{-3}$) arguably improved the total emission rate overestimate from 14% to 12%, stronger regularization ($\beta \geq 10 \text{ s}^2 \cdot \text{m}^{-3}$) led to more significant overestimates of between 22% and 44%. However, regularization was able to

significantly improve the unstable divergence of the predicted total emission rate with noisy observational data (dashed lines). The diverging total emission rate predicted for the case with observational noise and no regularization (red dashed line) can be attributed to the effectively unbounded prediction of weak false emission sources in regions of the domain with poor receptor network coverage, as the IDM seeks to find an emission source distribution to explain the fluctuating character of the noisy receptor observations. It is believed that regularization was able to penalize the prediction of the false sources and induce convergence to consistent overestimates of the total emission rate. The most aggressive regularization ($\beta = 1000 \text{ s}^2 \cdot \text{m}^{-3}$) was able to reduce emission rate to a converged overestimate of 40%. However, although objective function regularization encouraged faster convergence to a consistent total emission rate estimate, this came at the expense of predicting of very diffuse emission sources that reduced the confidence in the estimates of the emission source location.

5.1.3.1 Emission Source Filtering

Emission source filtering was also studied as either an alternative or a complementary method to objective function regularization for mitigating the influence of weak false emission sources on the overestimation of the total emission rate. Emission source filtering (see Section 4.3.6) considers regions of the domain with poor observability by the receptor network for the given flow direction as unknowable, and ignores the emission rate contribution from candidate emission sources in these regions. Figure 5.8 shows the emission source filters presented in Figure 4.11d superimposed on top of the

final (200th iteration) predicted emission source distributions for the single-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$), with and without observational noise.

Figure 5.8: Final emission source distributions with emission source filters for the single-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)

All of the filters shown in Figure 5.8 (ranging from the 0.1% - 10% thresholds) effectively exclude the false emission sources while still capturing the true emission source. The convergence history of the estimated total emission rates obtained by excluding the emission sources predicted outside of the bounds of the various filters is shown in Figure 5.9.

Figure 5.9: Convergence history of filtered total emission rate for the single-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)

Figure 5.9 shows that for the single-source test case with perfect observational data and no regularization, all of the filters were able to improve the emission rate overestimate from 14% to 8%. Emission source filtering also had a favourable influence on the diverging overestimation of the total emission rate for the cases with added observational noise. Minimal filtering (0.1%) was able to achieve a consistent emission rate overestimate of 89%, and additional filtering further improved the overestimate. The most aggressive filter (10%) gave the best total emission rate overestimate of 23%.

Finally, the combination of objective function regularization and emission source filtering was also considered. Figure 5.10 shows the convergence history of the total emission rate estimates with both objective function regularization and emission source filtering.

Figure 5.10: Convergence history of filtered total emission rate for the single-source test case with regularization

The most aggressive objective function regularization case ($\beta = 1000 \text{ s}^2 \cdot \text{m}^{-3}$) was not considered because it produced emission source distributions that were too diffuse. Figure 5.10 suggests that the 10% filter with $\beta < 1 \text{ s}^2 \cdot \text{m}^{-3}$ provides the best estimates of the total emission rate for the case of perfect receptor observations, and $\beta < 10 \text{ s}^2 \cdot \text{m}^{-3}$ is best for the case with 10% observational noise.

5.1.4 Single-Source Summary

The IDM was able to accurately estimate the location and total emission rate of a synthetic test case with a single emission source. While the final candidate emission source distributions were more diffuse than the known emission source, the peak nodal location of the final candidate volumetric emission rates exactly matched the nodal location of the known emission source. Table 5.1 gives a summary of the peak location error and the full area at half maximum (FAHM) of the known and final (200th iteration) candidate emission source distributions for the single-source test case with various levels of objective function regularization.

Table 5.1: Peak location error and full area half maximum (FAHM) of volumetric emission rate for the final iteration of the single-source test case

β [s ² ·m ⁻³]	No Observational Noise		10% Observational Noise	
	Location Error [m]	FAHM [m ²]	Location Error [m]	FAHM [m ²]
0	0 (0)	93 (14)	0 (0)	93 (14)
1	0 (0)	93 (14)	0 (0)	93 (14)
10	0 (0)	187 (27)	0 (0)	187 (27)
100	0 (0)	371 (51)	0 (0)	371 (51)
1000	0 (0)	656 (87)	0 (0)	656 (87)
known	-	36 (6)	-	36 (6)

() indicates the emission source location error or FAHM in number of elements

It is apparent from Table 5.1 and Figure 5.3 that objective function regularization encouraged the prediction of diffuse emission sources which gave less confidence in estimate of the emission source location. To maintain a compact estimate of the emission

source distribution, objective function regularization should be limited to $0 \leq \beta < 10 \text{ m}^2 \cdot \text{s}^{-3}$. Table 5.1 also shows that observational noise had no noticeable influence of the IDM's ability to locate the emission source. Table 5.2 gives a summary of the influence of emission source filtering and objective function regularization on final estimates of the total emission rate of the single-source test case with and without 10% observational noise.

Table 5.2: Summary of the influence of source filtering and regularization on emission rate of the final candidate emission sources for the single-source test case

$\beta [\text{s}^2 \cdot \text{m}^{-3}]$	0	1	10	100	1000
No Observational Noise					
$f(\mathbf{x}_{200})/f_{\text{no noise}}(\mathbf{x}_0)$	6.72×10^{-3}	2.48×10^{-2}	9.38×10^{-2}	3.03×10^{-1}	6.41×10^{-1}
$f^0(\mathbf{x}_{200})/f_{\text{no noise}}^0(\mathbf{x}_0)$	6.72×10^{-3}	1.20×10^{-2}	3.99×10^{-2}	1.76×10^{-1}	4.90×10^{-1}
r^2	0.993	0.987	0.957	0.812	0.476
Q/Q^*	1.14 [†]	1.12	1.22	1.45	1.34
$Q_{0.1\%}/Q^*$	1.08	1.10	1.19	1.33	1.14
$Q_{1\%}/Q^*$	1.08	1.10	1.18	1.33	1.14
$Q_{5\%}/Q^*$	1.08	1.10	1.18	1.30	1.11
$Q_{10\%}/Q^*$	1.08	1.10	1.17	1.27	1.07
10% Observational Noise					
$f(\mathbf{x}_{200})/f_{\text{no noise}}(\mathbf{x}_0)$	5.75×10^{-1}	5.94×10^{-1}	6.65×10^{-1}	8.82×10^{-1}	1.22×10^0
$f^0(\mathbf{x}_{200})/f_{\text{no noise}}^0(\mathbf{x}_0)$	5.75×10^{-1}	5.82×10^{-1}	6.09×10^{-1}	7.50×10^{-1}	1.07×10^0
$(f^0(\mathbf{x}_{200}) - f_{\min}^0)/f_{\text{no noise}}^0(\mathbf{x}_0)$	5.58×10^{-3}	1.25×10^{-2}	3.95×10^{-2}	1.81×10^{-1}	5.05×10^{-1}
r^2	0.623	0.619	0.601	0.508	0.296
$r^2/r_{\text{max},10\% \text{ noise}}^2$	0.994	0.987	0.959	0.811	0.472
Q/Q^*	3.98 [†]	2.33	1.86	1.68	1.40
$Q_{0.1\%}/Q^*$	1.89	1.80	1.67	1.54	1.19
$Q_{1\%}/Q^*$	1.64	1.61	1.56	1.51	1.18
$Q_{5\%}/Q^*$	1.35	1.33	1.35	1.40	1.13
$Q_{10\%}/Q^*$	1.23	1.23	1.26	1.33	1.08

[†] Not converged to a standard deviation $\leq 1\%$ of the final value over the last 25 iterations

$$f_{\text{no noise}}(\mathbf{x}_0) = f_{\text{no noise}}^0(\mathbf{x}_0) = 7.98 \times 10^{-6} \text{ kg}^2 \cdot \text{m}^{-6}$$

$$f_{\min,10\% \text{ noise}}^0 = 4.56 \times 10^{-6} \text{ kg}^2 \cdot \text{m}^{-6}$$

$$r_{\text{max},10\% \text{ noise}}^2 = 0.627$$

$$Q^* = 4.5 \text{ g/s}$$

From Table 5.2 and Figure 5.7 it is apparent that observational noise encouraged the prediction of false emission sources that contributed to a diverging overestimate of the

total emission rate. Regularization was able to partially mitigate this overestimation; however as previously stated, regularization also encouraged the prediction of diffuse emission sources which gave less confidence in the estimate of the emission source location. However, regularization did give faster convergence to a consistent emission source distribution and should not be outright rejected. Emission source filtering was able improve the diverging total emission rate overestimates to converged overestimates that were closer to the known total emission rate.

The results of the single-source test case suggest that objective function regularization and emission source filtering should be combined to confidently locate and quantify fugitive emissions. Based on these results, objective function regularization should be limited to $0 \leq \beta < 10 \text{ s}^2 \cdot \text{m}^{-3}$ and emission source filtering should be 10%. These combinations gave a suitable compromise between convergence to a consistent result within 200 iterations, confident location of the emission source, and more accurate quantification of the total emission rate. It should be noted that the single-source was located in a region of the domain that had favourable receptor coverage, and 10% filtering was able to aggressively penalize the false emission sources without influencing the correct emission source. The effect of filtering on emission sources with poor receptor coverage was further considered in the multi-source test case.

5.2 Multi-Source Test Case Results

5.2.1 Predicted Emission Source Distribution

The ability of the IDM to confidently locate and quantify multiple (ten) simultaneous fugitive emission sources was also considered. It is important to note that the IDM makes no assumptions about the number of emission sources originally used to generate the synthetic receptor observations, and is free to predict as many emission sources as required to minimize the objective function. Figure 5.11 shows the final (200th iteration) candidate emission source distributions for the multi-source test case, with and without 10% observational noise and without objective function regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$). The iterative convergence of the candidate emission source distributions are shown in Appendix G.

Figure 5.11: Final candidate emission source distributions for the multi-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)

Similar to the single-source case, the candidate emission sources were more diffuse than the known synthetic emission sources. However, Figure 5.11 shows that the peak volumetric emission rates generally aligned with the locations of the known emission sources (shown as hollow black circles, and labeled A - J). Table 5.3 and Table 5.4 (see Section 5.2.4) quantify the peak location error and FAHM for each of the known emission sources. A notable exception to this trend is the very diffuse candidate emission

source predicted around emission source C. This diffuse emission source distribution gives little confidence in the estimate of its location. This particularly poor prediction is directly attributable to poor sensitivity of the receptor network to potential emission sources in this region of the domain, due to the prevailing direction of the wind, as was evident in the development of the source filters discussed in Section 4.3.6. The IDM was able to confidently identify the locations of emission sources located more centrally in the domain.

Figure 5.11c and d show the final candidate emission source distributions for the multi-source case with and without 10% observational noise on a logarithmic scale to highlight the prediction of false emission source in regions of the domain that have poor observability to the receptor network. Similar to the single-source case, the addition of observational noise caused the prediction of a greater number of false emission sources which contributed to an overestimate of the total emission rate (see Section 5.2.3).

5.2.1.1 Objective Function Regularization

Objective function regularization was again implemented in an attempt to mitigate the prediction of false emission sources. Similar to the single-source case, the regularization parameter β was chosen to be 0, 1, 10, 100 or $1000 \text{ s}^2 \cdot \text{m}^{-3}$. Figure 5.12 shows magnified views of the final (200th iteration) candidate emission source distributions for the multi-source source test case with and without observational noise and regularization.

Figure 5.12: Effect of regularization on the final candidate emission source distributions for the multi-source test case

Figure 5.12: Effect of regularization on the final candidate emission source distributions for the multi-source test case (*cont.*)

As shown in Figure 5.12 and quantified in terms of FAHM in Table 5.4 (see Section 5.2.4), objective function regularization encouraged the prediction of much more diffuse emission sources, similar to what was seen for the single-source test case. This trait is undesirable in inverse atmospheric dispersion modelling because it gives less certainty in estimate of the emission source locations. Additionally, Figure 5.13 shows the influence of objective function regularization on the final (200th iteration) candidate emission

source distributions over the entire domain on a logarithmic scale to highlight the weak false emission source.

Figure 5.13: Effect of regularization on the final candidate emission source distributions for the multi-source test case

Figure 5.13: Effect of regularization on the final candidate emission source distributions for the multi-source test case (*cont.*)

Similar to the single-source test case, objective function regularization introduced additional weak false emission sources within the region of favourable observability to the receptor network. From Figure 5.12 and Figure 5.13 it is believed that a suitable range of the regularization parameter is $0 \leq \beta < 10 \text{ s}^2 \cdot \text{m}^{-3}$ for the multi-source test case, since this range offered a reasonable compromise between false emission source mitigation and diffusion of the emission source estimates.

5.2.2 Convergence

5.2.2.1 Objective Function Reduction

Figure 5.14 shows the convergence history, as well as the influence of the regularization parameter (β) and observational noise on the minimization of the objective function normalized by its initial value ($f_{\text{no noise}}(\mathbf{x}_0) = 1.51 \times 10^{-4} \text{ kg}^2 \cdot \text{m}^{-6}$) for the multi-source test case.

Figure 5.14: Convergence history of objective function minimization for the multi-source test case

Similar to the single-source case, Figure 5.14a and b shows that the minimization of the objective function was significantly influenced by observational noise and regularization. Added Gaussian observational noise once again introduced modelling errors which were unexplainable by any candidate steady emission sources, resulting in a non-zero global minimum of the observational term of the objective function. The value of this minimum ($f_{\text{min}, 10\% \text{ noise}}^o = 8.16 \times 10^{-6} \text{ kg}^2 \cdot \text{m}^{-6}$, shown as a grey dashed line in Figure 5.14b) was estimated using the same procedure as for the single-source case. The objective function for the multi-source case with observational noise and without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$, dashed red line) also approached its estimated global minimum. Figure 5.14c shows the convergence history of the observational term of the objective function relative to its expected global minimum ($f^o(\mathbf{x}_k) - f_{\text{min}}^o$), which shows that the

observational term was minimized toward its relative global minimum, regardless of observational noise

Variation of the regularization parameter (β) in the multi-source case showed the same trends as the single-source case. Since the initial guess of no emission sources matched the assumed background ($\mathbf{s}^b$), the initial value of the objective function was not dependant on the regularization parameter (β) (i.e. $f(\mathbf{x}_0) = f^o(\mathbf{x}_0)$). While regularization did speed convergence to consistent emission source solutions, it also impeded the minimization of the objective function by penalizing candidate emissions sources that deviated from an assumed prior more than they improved the misfit between modelled and known receptor observations. Similar trends can be observed when considering the gradient of the objective function (see Appendix Gradient) as a measure of the convergence of a candidate solution to a minimum of the objective function.

5.2.2.2 *Coefficient of Determination*

Figure 5.15a shows the convergence histories of the coefficient of determination (r^2) of the modelled and known observations for the multi-source test case. The candidate emission sources predicted without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$) in the absence of observational noise (solid red line) approached perfect correlation ($r^2 = 1$). For the multi-source case with 10% added observational noise, the expected upper limit of the coefficient of determination was estimated to be $r_{\text{max}, 10\% \text{ noise}}^2 = 0.932$ (shown as a grey dashed line) using the same technique discussed with the single-source case in Section 5.1.2.2. With no regulation ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$, dashed red line), the coefficient of

determination for multi-source test case with added observational noise also approached its expected upper limit.

Figure 5.15: Convergence history of coefficient of determination of receptor observations for multi-source case

Figure 5.15b shows the convergence history of the coefficient of determination normalized by its expected upper limit ($r_{\text{max, no noise}}^2 = 1$ and $r_{\text{max, 10% noise}}^2 = 0.932$). Similar to the single-source test case, although the candidate emission sources predicted with noisy observations were unable to reproduce the random fluctuations in the observations, they were able to produce receptor observations that achieved the expected maximum coefficient of determination.

Increases in the regularization parameter (β) led to predictions of emission sources that had diminished ability to explain the known receptor observations. With $\beta \leq 10 \text{ s}^2 \cdot \text{m}^{-3}$, the negative influence on the correlation was relatively small. However, with $\beta \geq 100 \text{ s}^2 \cdot \text{m}^{-3}$, the coefficient of determination was significantly reduced, indicating that

regularization penalized the emission source distributions too aggressively such that they were unable to reproduce the known receptor observations. However the regularization did speed convergence to consistent emission source distributions.

Figure 5.15 suggests that a suitable range of values for the regularization parameter is $0 \leq \beta \leq 10 \text{ s}^2 \cdot \text{m}^{-3}$, with which the coefficients of determination approached their expected upper limits, regardless of observational noise.

5.2.3 Total Emission Rate

The total emission rate of the candidate emission source distribution was again determined by integrating the candidate volumetric source emission rate $s(\mathbf{r}, t; \mathbf{x}_k)$ over the entire airshed domain (see Section 4.4.4). While the candidate emission sources (see Section 5.2.1.1) were more diffuse than the known emission sources, the total emission rates were converged to within 20% the known total emission rate ($Q^* = 50 \text{ g/s}$) in all scenarios calculated after 200 iterations. Figure 5.16 shows the convergence history of the total emission rates for the candidate emission sources for all scenarios of the multi-source case.

Figure 5.16: Convergence history of the total emission rate for the multi-source test case

After 200 iterations, the total emission rates of the candidate emission sources predicted with perfect observations ranged from an 11% overestimate to a 13% underestimate of the known synthetic total emission rate ($Q^* = 50 \text{ g/s}$). Observational noise increased the estimated total emission rates by no more than 6%. As discussed previously (see Section 5.2.1.1), the addition of regularization encouraged the prediction of weaker and more diffuse sources. However, regularization also tended to reduce the calculated total emission rate, and with $\beta = 100 \text{ s}^2 \cdot \text{m}^{-3}$ led to a 3% overestimate of the known synthetic total emission rate ($Q^* = 50 \text{ g/s}$), and with $\beta = 1000 \text{ s}^2 \cdot \text{m}^{-3}$ led to a 13% underestimate, independent of whether observational noise was considered. Finally, similar to the single-source test case, objective function regularization encouraged convergence to a consistent total emission rate estimate at the expense of predicting diffuse emission

sources and reducing the confidence in the estimates of emission the emission source locations.

5.2.3.1 Emission Source Filtering

The issue of poor emission rate convergence due to the influence of false emission sources in regions with poor observability was also investigated through the use of emission source filters (as derived in Section 4.3.6). Figure 5.17a and b shows the emission source filters superimposed on the final (200th iteration) predicted multi-source distribution with no regularization, with and without added observational noise.

Figure 5.17: Emission source filters for the final emission source distribution of the multi-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)

From Figure 5.17c and d it can be seen that emission source C falls between the bounds of the 1% and 5% filters. This implies that emission source C is on the boundary of the field of view of the receptor network, which explains the difficulty that the IDM had in confidently locating emission source C with the same precision as the other sources. From Figure 5.17e,f it can be seen that emission source H also had poor observability to the receptor network. The total emission rate achieved by applying the various source filters to the source distribution is shown in Figure 5.18.

Figure 5.18: Convergence history filtered total emission rate for the multi-source test case

Figure 5.18a shows that a filter threshold of 0.1% or 1% was sufficient to remove the influence of the false emission sources on the overestimate of the total emission rate and enable convergence to a consistent total emission rate estimate. However, further increases in the source filter threshold (5% - 10%) caused emission sources C and H to be excluded, which produced underestimates of the overall emission rate for the entire domain. Since emission sources C and H are located in regions of the domain that are in a sense unknowable, normalizing by the total known emission rate ($Q^* = 50$ g/s) is not

necessarily a fair criterion for assessing the performance of the model within the restricted bounds of the filter. Figure 5.18b instead shows the convergence history of the filtered total emission rate normalized by the total emission rate of the known emission sources within the boundaries of the respective filters. From Figure 5.18b it can be seen that the 5% filter gave the best estimate of the total source emission rate within its restricted field of view. The 10% filter increased the overestimate of the total emission rate because it ignored the emission rate contribution from the synthetic source H, but did not ignore all of the diffused volumetric emission rate distribution predicted for source H.

Finally, similar to the single-source test case, the combination of objective function regularization and emission source filter was also considered. Figure 5.19 shows the convergence history of the total emission rate estimates with both objective function regularization and emission source filtering. The plotted results suggest that the 5% filter with $\beta \leq 10 \text{ s}^2 \cdot \text{m}^{-3}$ provided the best estimate of the total emission rate for the multi-sources case, both with and without observational noise.

Figure 5.19: Convergence history of adjusted filtered total emission rate for the multi-source test case

5.2.4 Multi-Source Summary

The IDM was able to estimate the locations and total emission rate of multiple emission sources with the proviso that they were within the field of view of the receptor network. Similar to the single-source test case, the final candidate emission source distributions were more diffuse than the known synthetic emission sources. Observational noise, objective function regularization, and emission source filtering were explored to determine their influence on emission source location and quantification.

Similar to the single-source test case, the addition of 10% observational noise to the multi-source test case had little influence on the location the known fugitive emissions. However, observational noise caused the prediction of additional, diffuse false emission sources that spanned a large portion of the domain on the edge of the FOV of the receptor network. These false sources caused an overestimate of the known total emission rate, which regularization and emission source filtering were able to mitigate.

Table 5.3 and Table 5.4 show the influence of regularization on the source location error and the full area at the half maximum (FAHM) of source volumetric emission rate, respectively.

Table 5.3: Emission source location error in [m] for the multi-source test case

Source	$\beta \text{ [s}^2 \cdot \text{m}^3]$									
	No Observational Noise					10% Observational Noise				
	0	1	10	100	1000	0	1	10	100	1000
A	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
B	0 (0)	0 (0)	0 (0)	0 (0)	21.95 (6)	0 (0)	0 (0)	0 (0)	0 (0)	21.95 (6)
C	9.82 (2)	9.82 (2)	9.82 (2)	14.70 (3)	14.70 (3)	9.82 (2)	9.82 (2)	9.82 (2)	14.70 (3)	14.70 (3)
D	4.50 (1)	4.50 (1)	4.49 (1)	4.49 (1)	4.49 (1)	4.50 (1)	4.50 (1)	4.49 (1)	4.49 (1)	4.49 (1)
E	0 (0)	0 (0)	0 (0)	4.49 (1)	4.20 (1)	0 (0)	0 (0)	0 (0)	7.20 (2)	4.20 (1)
F	0 (0)	0 (0)	0 (0)	3.79 (1)	0 (0)	0 (0)	0 (0)	0 (0)	3.79 (1)	0 (0)
G	0 (0)	0 (0)	0 (0)	8.40 (1)	15.88 (4)	0 (0)	0 (0)	0 (0)	8.40 (2)	15.88 (4)
H	3.93 (1)	3.93 (1)	3.93 (1)	6.35 (2)	6.35 (2)	3.93 (1)	3.93 (1)	3.93 (1)	6.35 (2)	6.35 (2)
I	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
J	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

() indicates the emission source location error in number of elements

Table 5.4: Emission source FAHM in [m²] for the multi-source test case

Source	Known	β [s ² ·m ⁻³]									
		No Observational Noise					10% Observational Noise				
		0	1	10	100	1000	0	1	10	100	1000
A	36	257	257	269	371	591	244	244	269	371	591
	(6)	(36)	(36)	(38)	(51)	(80)	(34)	(34)	(38)	(51)	(80)
B	69	132	132	132	274	134	132	132	132	274	134
	(6)	(13)	(13)	(13)	(34)	(34)	(13)	(13)	(13)	(34)	(19)
C	58	1428	1493	1819	2858	3467	1363	1421	1845	2668	3526
	(6)	(156)	(164)	(211)	(353)	(427)	(148)	(156)	(214)	(329)	(434)
D	51	418	418	467	753	1597	418	418	433	745	1532
	(6)	(50)	(50)	(57)	(97)	(222)	(50)	(50)	(52)	(95)	(208)
E	48	433	433	458	1157	4529	412	433	449	1256	4709
	(6)	(55)	(55)	(58)	(217)	(1199)	(52)	(55)	(57)	(229)	(1196)
F	65	65	65	110	235	542	65	65	110	235	533
	(7)	(7)	(7)	(14)	(29)	(77)	(7)	(7)	(14)	(29)	(75)
G	49	635	655	722	1386	3182	671	687	797	1454	3154
	(7)	(81)	(83)	(94)	(197)	(414)	(84)	(87)	(105)	(208)	(410)
H	20	110	110	115	304	345	110	110	115	309	340
	(7)	(34)	(34)	(37)	(209)	(346)	(34)	(34)	(37)	(218)	(317)
I	34	189	189	202	313	768	189	189	189	313	766
	(6)	(27)	(27)	(29)	(41)	(127)	(27)	(27)	(27)	(41)	(114)
J	69	302	302	431	1133	2110	320	320	431	1134	2109
	(6)	(35)	(35)	(51)	(138)	(303)	(35)	(35)	(51)	(138)	(299)

() indicates the emission source FAHM in number of elements

It was concluded that regularization with $\beta < 100 \text{ s}^2 \cdot \text{m}^{-3}$ consistently gave the best estimates of the emission source locations and $\beta < 10 \text{ s}^2 \cdot \text{m}^{-3}$ predicted more concentrated emission source distributions, giving more confidence in their estimates of emission source location.

Table 5.5 summarizes the influence of objective regularization and emission source filtering on the objective function, the coefficient of determination, and the total emission rate for the final (200th iteration) candidate emission sources predicted by the IDM for the multi-source test case.

Table 5.5: Summary of the influence of objective regularization and emission source filtering on the final candidate emission sources for the multi-source test case

β [$\text{s}^2 \cdot \text{m}^{-3}$]	0	1	10	100	1000
No Observational Noise					
$f(\mathbf{x}_{200})/f_{\text{no noise}}(\mathbf{x}_0)$	1.27×10^{-2}	1.77×10^{-2}	4.51×10^{-2}	1.41×10^{-1}	3.41×10^{-1}
$f^0(\mathbf{x}_{200})/f_{\text{no noise}}^0(\mathbf{x}_0)$	1.27×10^{-2}	1.44×10^{-2}	2.26×10^{-2}	7.93×10^{-2}	2.26×10^{-1}
r^2	0.983	0.981	0.970	0.894	0.698
Q/Q^*	1.11	1.09	1.06	1.02 [†]	0.87
$Q_{0.1\%}/Q_{0.1\%}^*$	1.08	1.08	1.05	0.98	0.79
$Q_{1\%}/Q_{1\%}^*$	1.06	1.05	1.03	0.95	0.77
$Q_{5\%}/Q_{5\%}^*$	1.02	1.02	1.01	0.96	0.80
$Q_{10\%}/Q_{10\%}^*$	1.08	1.08	1.06	1.00	0.83
10% Observational Noise					
$f(\mathbf{x}_{200})/f_{\text{no noise}}(\mathbf{x}_0)$	6.60×10^{-2}	7.03×10^{-2}	9.90×10^{-2}	1.97×10^{-1}	3.98×10^{-1}
$f^0(\mathbf{x}_{200})/f_{\text{no noise}}^0(\mathbf{x}_0)$	6.60×10^{-2}	6.66×10^{-2}	7.55×10^{-2}	1.34×10^{-1}	2.82×10^{-1}
$(f^0(\mathbf{x}_{200}) - f_{\text{min}}^0)/f_{\text{no noise}}^0(\mathbf{x}_0)$	1.21×10^{-2}	1.27×10^{-2}	2.17×10^{-2}	8.02×10^{-2}	2.28×10^{-1}
r^2	0.918	0.917	0.906	0.833	0.649
$r^2/r_{\text{max},10\% \text{ noise}}^2$	0.984	0.983	0.971	0.893	0.696
Q/Q^*	1.17	1.14	1.09	1.03	0.87
$Q_{0.1\%}/Q_{0.1\%}^*$	1.10	1.10	1.07	0.99	0.79
$Q_{1\%}/Q_{1\%}^*$	1.07	1.07	1.04	0.96	0.77
$Q_{5\%}/Q_{5\%}^*$	1.03	1.02	1.01	0.97	0.80
$Q_{10\%}/Q_{10\%}^*$	1.08	1.08	1.06	1.00	0.83

$$\begin{aligned}
 f_{\text{no noise}}(\mathbf{x}_0) &= f_{\text{no noise}}^0(\mathbf{x}_0) = 1.51 \times 10^{-4} \text{ kg}^2 \cdot \text{m}^{-6} \\
 f_{\text{min},10\% \text{ noise}}^0 &= 8.16 \times 10^{-6} \text{ kg}^2 \cdot \text{m}^{-6} \\
 r_{\text{max},10\% \text{ noise}}^2 &= 0.933 \\
 Q^* &= Q_{0.1\%}^* = Q_{1\%}^* = 50 \text{ g/s} \\
 Q_{5\%}^* &= 45 \text{ g/s} = 45 \text{ g/s} \\
 Q_{10\%}^* &= 40 \text{ g/s}
 \end{aligned}$$

From Table 5.5 and Figure 5.16, it is apparent that observational noise encouraged the prediction of false emission sources that contributed to overestimates of the total emission rate, although the overestimates were not as pronounced as in the single-source test case.

Regularization was able to partially mitigate this overestimation and enable faster convergence to a consistent total emission rate estimate. However regularization also encouraged the prediction of diffuse emission sources that gave less confidence in the estimated source locations. By comparison, emission source filtering was able mitigate

the divergence of the total emission rate overestimates, resulting in converged overestimates that were closer to the known total emission rate. As with the single-source test case, normalizing the coefficient of determination (r^2) by its expected maximum ($r_{\max, \text{no noise}}^2 = 1$ and $r_{\max, 10\% \text{ noise}}^2 = 0.933$) resulted in almost perfect agreement between the no noise and 10% noise cases for all conditions (see Table 5.5).

The results of the multi-source test case once again suggest that a combination of objective function regularization and emission source filtering are most effective in confidently locating and quantify fugitive emission sources. For the conditions tested, objective function regularization of $0 \leq \beta < 10 \text{ s}^2 \cdot \text{m}^{-3}$ and 5% emission source filter gave the best compromise between confidently locating the emission sources and accurately quantifying the total emission rate.

5.3 Summary

The IDM was able to locate and quantify the total emission rate of synthetic fugitive emissions on the condition that they were within the FOV of the receptor network. The candidate emission sources predicted after 200 iterations were more diffuse than known emission sources that produced synthetic observations. The addition of observational noise did not influence the IDM's ability to confidently locate the known emission sources. However observational noise encouraged the prediction of additional diffuse false emission sources which contributed to diverging overestimates of the total emission rate. Objective function regularization was able to partially mitigate some of these false emission sources, allowing quicker convergence to a consistent estimate of the total

emission rate, at the expense of diffusing the true emission sources and reducing the confidence in the estimate of the emission source locations. Emission source filtering was also effective at mitigating the total emission rate overestimate from false emission source by ignoring the total emission rate contribution from candidate sources outside the receptor network FOV. This strategy also ignored some of the true emission sources which were outside of the receptor network FOV.

Based on the range of conditions considered in this thesis, the best compromise between the accurate estimates of the emission source locations, total emission rates, and convergence rates should be achieved with a combination of objective function regularization and emission source filtering. Choosing the regularization parameter in the range $0 \leq \beta < 10 \text{ s}^2 \cdot \text{m}^{-3}$ allows confident location of the emission source(s), and in combination with the 5% filter, enables accurate estimation of the total emission rate by ignoring emission rate contributions from false emission sources outside the receptor network FOV.

Chapter 6

Conclusions and Recommendations

6.1 Conclusions

The goal of this thesis was to present a synthetic proof-of-concept study for the location and quantification of fugitive methane emissions in the upstream oil and gas industry using inverse micro-scale atmospheric dispersion modelling. Several modelling simplifications were implemented to make this problem tractable with reasonable computing time scales (~1 month for 200 optimization iterations, with fifteen physical minutes of receptor observations). The study was constrained to a two-dimensional airshed domain by assuming no vertical transport of pollution. Additionally, other transport phenomena such as plume buoyancy and wet/dry deposition deemed to be negligible and were ignored. Finally, the emission sources were assumed to maintain constant emission rates and locations in time. Under these modelling simplifications a custom inverse micro-scale atmospheric dispersion model (IDM) was developed and tested under non-ideal conditions such as multiple emission sources, unfavourable receptor coverage, and the addition of Gaussian distributed observational noise.

The results from the IDM are encouraging. For the presented test cases, the IDM was able to locate and quantify most of the synthetic emission sources. However, the IDM was unable to confidently locate emission sources outside the field of view of the receptor network, and false emission sources were often predicted in regions of the

domain that had poor receptor coverage. Overall, the test case results have illustrated that observational noise, objective function regularization, and emission source filtering are each key factors that influence the robustness of the IDM in confidently locating and quantifying emission sources. These are separately discussed below.

6.1.1 Observational Noise

The addition of 10% Gaussian distributed observational noise did not significantly affect the IDM's ability to confidently locate the emission sources. Observational noise introduced pollution concentration fluctuations which were most likely unexplainable by any steady emission sources. This resulted in the observational term of the objective function with noisy observations having a non-zero global minimum. The objective function values and coefficients of determination suggested that the final emission source distributions predicted with observational noise were significantly worse than those predicted with perfect observations, despite them looking similar on a linear scale. It is apparent that the objective function and coefficient of determination are not comparable between cases with different levels of observational noise. Instead their relative values should be compared by removing the bias that observation noise introduces ($f^o(\mathbf{x}_k) - f_{\min}^o$ and $r^2/r_{\max}^2$). However, the minimum of the observation term of the objective function ($f_{\min}^o$) and the maximum of the coefficient of determination ($r_{\max}^2$) can be estimated for a synthetic test case, for which the correct emission source distribution is already known.

While observational noise did not significantly influence IDM's ability to confidently locate emission sources, it also caused the IDM to significantly overestimate the total emission rate the actual emission sources. This overestimate was caused by several diffuse emission sources predicted in regions of the domain which had poor observability to the receptor network. Objective function regularization and emission source filtering were studied to mitigate this total emission rate overestimate.

6.1.2 Objective Function Regularization

Objective function regularization was implemented in an attempt to prevent the prediction of false emission sources in regions with poor receptor network coverage or due to noise in the observations. The regularization term was chosen to be an area weighted sum of the squared deviations of the candidate emission source distribution from an assumed background of no emission sources. This regularization term was chosen for the current studies because it had the favourable property of being continuously differentiable with respect to the model parameters ($\mathbf{x}$), permitting the use of a gradient-based numerical optimization algorithm such as L-BFGS-B for the minimization of the objective function.

For the presented studies the regularization term of the objective function was scaled by the parameter β , to study the influence of regularization on the candidate emission source distributions predicted by the IDM. Objective function regularization was able to partially mitigate the prediction of false emission sources, and encourage faster convergence to consistent estimates of the total emission rate, at the expense of more

diffuse predictions of the true emission sources, which reduced confidence in the emission source locations. Based on the test cases presented in this thesis, it was concluded that objective function regularization should be limited to $0 \leq \beta < 10 \text{ s}^2 \cdot \text{m}^{-3}$ and performed best when combined with emission source filtering.

6.1.3 Emission Source Filtering

Emission source filtering was also studied in combination with objective function regularization to mitigate the influence of false emission sources on the total emission rate. The airshed domain was divided in two regions, one with favourable observability to the receptor network and the other with unfavourable observability. Candidate emission sources predicted in the unfavourable region were deemed to be “unknowable” and were excluded from the contribution to the total emission rate. For the test cases presented in this thesis, the 5% filter performed the best in combination with objective function regularization ($0 \leq \beta < 10 \text{ s}^2 \cdot \text{m}^{-3}$), where this combination gave the best estimates of the total emission rate while avoiding prediction of overly diffuse emission sources.

6.2 Recommendations and Future Work

6.2.1 Three-Dimensional Model

One of the most significant limiting factors of the current study was the restriction to two-dimensional dispersion modelling. Future implementations of the inverse dispersion

model should consider a three-dimensional domain, to capture the vertical transport of pollutants. Ideally, the three-dimensional model would accommodate an unstructured mesh, solve the Navier-Stokes equations (continuity and momentum), and determine turbulent mass diffusivity with turbulence modelling, in addition to modelling the transport of pollutants with the advection-diffusion equation.

While it is expected that a three-dimensional model would introduce additional mesh nodes and elements that would increase the computational time of each optimization iteration, the number of emission source parameters would not necessarily increase at the same rate. The emission source parameters could be constrained to only represent the potential emission sources from the surfaces of the three-dimensional domain (e.g. ground, buildings, and inlets). Finally, the three-dimensional model must also provide an efficient means for computing the gradient of the objective function with respect to all of the emission source parameters.

A candidate CFD software package that seems to meet these criteria is OpenFOAM. This is a free and open source C++ software package for the numerical solution of PDEs, and is specifically targeted toward fluid dynamics applications. The gradient of the objective function could be efficiently evaluated in two ways within OpenFOAM. The first method would evaluate the objective function gradient with the continuous adjoint method (see Appendix A). OpenFOAM can solve the adjoint advection-diffusion equation with the same computational sophistication as the pollution transport modelling with the advection-diffusion equation. A second method in which OpenFOAM can be used to evaluate the objective function gradient is to use automatic differentiation (AD) with reverse Jacobin accumulation mode. The method would generate an additional

derivative computing source code by mechanically applying chain rule to the arithmetic operations of the original source code for evaluating the objective function. Several AD packages for C++ are available such as ADC, ADOL-C, FADBAD++, CasADi, CppAD, OpenAD, Sacado and TAPENADE.

Evaluating the objective function with the continuous adjoint method allows for the use of adaptive mesh refinement (AMR) procedures such as those suggested by Ivan (2010) to achieve better spatial resolution in regions of high gradients. Guo, Yang, Zhang, Weng, & Fan (2009) demonstrated automatic mesh refinement with an adjoint method to efficiently predict receptor observations from a candidate emission source. However, as with the advection-diffusion equation (3.1), the numerical solution of the adjoint advection-diffusion equation (see Appendix A) will contain truncation errors from the discretization and integration schemes. These errors will cause disagreement between the actual objective function gradient and the estimated gradient. By contrast, automatic differentiation of the objective function already contains the truncation errors, and as such the evaluated gradient is closer to the true gradient. In the limit of an infinite resolution computational grid, both the discrete and continuous adjoint approach “should be consistent and converge to the correct analytic value for the gradient of the objective function” (Giles and Pierce, 2000).

6.2.2 Objective Function Regularization

Alternative objective function regularization techniques such as L_1 -regularization (Tibshirani, 1996), maximum entropy (Press et al., 1997), or the total flux (Thomson et

al., 2007) may also be worth consideration as these alternate regularization terms may converge to sparser emission source distributions. However, these alternate regularization terms (such as L_1 -regularization) may not be C^1 continuous (i.e. the gradient contains discontinuities at the minima) and special numerical optimization algorithms such as the Orthant-wise Limited-memory Quasi-Newton Optimizer (OWL-QN) (Andrew and Gao, 2007) may be required to accommodate the gradient discontinuities of the objective function.

6.2.3 Emission Source Filtering

Rather than applying the emission source filter as a post-process to emission source distributions predicted by the IDM, future implementations could incorporate the emission source filter within the model. The volumetric emission rates from mesh nodes outside of the bounds of the emission source filter could be fixed to zero by the L-BFGS-B optimization algorithm by constraining their upper and lower search bounds to zero. This incorporation of the filter within the model would allow for the quality of the filtered emission source distributions to be quantified with the coefficient of determination, the objective function, and its gradient, in addition to the total emission rate.

6.2.4 Receptor Observation Uncertainty

Operational implementations of an IDM to locate and quantify fugitive emissions will use in situ instruments to take pollutant concentration measurements (i.e. receptor

observations). The uncertainty (standard deviation) of these concentration measurements should ideally be used to weight the contribution of deviations of the model from the known receptor observations to the objective function. The presented objective function (2.2) contains the term $\mathbf{R}^{-1}$ which is the inverse of the error-covariance matrix of the observed concentrations. In the present study, this term was chosen to be the identity matrix, which suggested that the uncertainties in all the receptor observations were constant. Future implementations should estimate the actual error-covariance matrix ($\mathbf{R}$) for the known receptor observations. Since it is not expected for one receptor observation to influence another, $\mathbf{R}$ would be a diagonal matrix, for which the inverse ($\mathbf{R}^{-1}$) is easily evaluated as also a diagonal matrix, whose elements are the reciprocal of the elements of $\mathbf{R}$.

6.2.5 Flow Field Inversion

Operational implementations of the IDM will also require estimations of the atmospheric flow field from limited anemometer observations. Two naive solutions to this problem are immediately apparent. If flow speed and direction are assumed to be constant over the entire inlet to the airshed domain. A scalar valued objective function can be formed to measure the misfit between the anemometer observations and the model predictions for a candidate inlet speed and direction. It is expected that the values of the parameters (speed and direction) that minimize this objective function can be found with similar techniques presented in this thesis for emission source inversion. An alternative approach is to pre-compute a library of flow fields with different inlet conditions. An

artificial neural network could be trained with this library of flow solutions to allow for estimation of the inlet parameters, which define the entire flow field.

6.2.6 Receptor Placement

Finally, a critical factor in emission source location is the field of view of the receptor network and the resultant observability of an emission source. The receptor placement and wind direction are the two factors that most strongly influence field of view of the receptor network. In general, it is desirable to maximize the region of the domain that is covered by the receptor network to maximize the regions that can be confidently searched for potential emission sources, while under the constraint of a limited number of receptors. One suggested method for determining a suitable placement of a fixed number of receptors is to formulate an optimization problem with the coordinates of the receptors as the parameters of the objective function. A candidate objective function for this optimization problem is to maximize the area (or volume in three-dimensions) of the domain containing retroplumes above a predetermined threshold, resulting from unit emission rates from the candidate locations of the receptors. The evaluation of these retro-plumes has already been demonstrated in the development of the emission source filters (see Section 4.3.6). This objective function could be augmented with a penalty term for predicting receptors outside of the computational domain (e.g. inside of a building). For a network of 25 sensors (50 optimization parameters as xy -coordinate pairs) this problem is sufficiently small (few optimization parameters) such that it could be studied with a Genetic-Algorithm, Simulated Annealing or a gradient based optimization algorithm.

References

- Addepalli, B., Sikorski, C., and Pardyjak, E.R. (2009). Source Characterization of Atmospheric Releases using Quasi-Random Sampling and Regularized Gradient Optimization.
- Allen, C., Young, G., and Haupt, S. (2007). Improving pollutant source characterization by better estimating wind direction with a genetic algorithm. *Atmos. Environ.* 41:2283–2289.
- Andrew, G. and Gao, J. (2007). *Scalable Training of L1-Regularized Log-Linear Models*. In 24th international conference on Machine learning.
- ANSYS (2006a). ANSYS CFX-Solver Theory Guide. Canonsburg, PA 15317.
- ANSYS (2006b). ANSYS CFX-Solver Modeling Guide. Canonsburg, PA 15317.
- ANSYS (2009a). ANSYS ICEM CFD 12.1 Help Manual. Canonsburg, PA 15317.
- ANSYS (2009b). ANSYS ICEM CFD 12.1 User Manual. Canonsburg, PA 15317.
- Ashbaugh, L.L., Malm, W.C., and Sadeh, W.Z. (1985). A residence time probability analysis of sulfur concentrations at grand Canyon National Park. *Atmos. Environ.* 19:1263–1270.
- Berg, M., Cheong, O., Kreveld, M., and Overmars, M. (2008). *Computational Geometry: Algorithms and Applications* Third. Berlin, Heidelberg, Springer-Verlag.
- Bergin, M.S. and Milford, J.B. (2000). Application of Bayesian Monte Carlo analysis to a Lagrangian photochemical air quality model. *Atmos. Environ.* 34:781–792.
- Bertion, L., Evensen, G., and Wackernagel, H. (2003). Sequential Data Assimilation Techniques in Oceanography. *Int. Stat. Rev.* 71:223–241.
- Bird, R.B., Stewart, W.E., and Lightfoot, E.N. (1960). *Transport Phenomena*. New York, Wiley.
- Bird, R.B., Stewart, W.E., and Lightfoot, E.N. (2007). *Transport Phenomena* 2nd ed. Wiley.
- Bischof, C., Carle, A., Peyvand, K., and Mauer, A. (1994). The ADIFOR 2.0 System for the Automatic Differentiation of Fortran 77 Programs. . 0–32.

- Brereton, C.A. (2010). Investigation of Trajectory Statistical Methods for Locating Fugitive Emissions Sources on a Building Scale. Carleton University.
- Brereton, C.A. and Johnson, M.R. (2011). *Effect of Wind Modeling Error on Fugitive Emission Location Predictions*. In Air & Waste Management Association 104th Annual Conference. Orlando, FL.
- Brereton, Carol A and Johnson, M.R. (2012). Identifying Sources of Fugitive Emissions in Industrial Facilities using Trajectory Statistical Methods. *Atmos. Environ.* 51:46–55.
- Brereton, Carol A. and Johnson, M.R. (2012). *Steady-state Wind Field Reconstruction for Pollutant Emission Location*. In Air & Waste Management Association (AWMA) 105th Annual Conference. San Antonio, TX.
- Brooks, A.N. and Hughes, T.J.R. (1982). Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Eng.* 32:199–259.
- Butcher, J.C. (2008). *Numerical Methods for Ordinary Differential Equations* Second Ed. John Wiley & Sons.
- Byrd, R.H., Lu, P., Nocedal, J., and Zhu, C. (1994). A Limited Memory Algorithm for Bound Constrained Optimization. Evanston, IL.
- Byun, D. and Schere, K.L. (2004). Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. Research Triangle Park, North Carolina.
- CAPP (2004). A National Inventory of Greenhouse Gas (GHG), Criteria Air Contaminant (CAC) and Hydrogen Sulphide (H₂S) Emissions by the Upstream Oil and Gas Industry - Volume 1, Overview of the GHG Emissions Inventory.
- CAPP (2007). Management of Fugitive Emissions at Upstream Oil and Gas Facilities.
- Carmichael, G.R., Tang, Y., Kurata, G., Uno, I., Streets, D., Woo, J.H., Huang, H., Yienger, J., Lefer, B., Shetter, R., et al. (2003). Regional-scale chemical transport modeling in support of the analysis of observations obtained during the TRACE-P experiment. *J. Geophys. Res.* 108:8823.
- Cartwright, H.M. and Harris, S.P. (1993). Analysis of the distribution of airborne pollution using genetic algorithms. *Atmos. Environ. Part A. Gen. Top.* 27:1783–1791.

- Cauchy, A.-L. (1847). Méthode générale pour la résolution des systèmes d'équations simultanées [Translated: (2010)]. *Compte Rendu des S'eances L'Acad'emie des Sci.* 25:536–538.
- Černý, V. (1985). Thermodynamical Approach to the Traveling Salesman Problem : An Efficient Simulation Algorithm. *J. Optim. Theroy Appl.* 45:41–51.
- Charney, J.G., Fjörtoft, R., and von Neumann, J. (1950). Numerical Integration of the Barotropic Vorticity Equation. *Tellus.* 2:237–254.
- Chow, F.K., Kosović, B., and Chan, S. (2008). Source Inversion for Contaminant Plume Dispersion in Urban Environments Using Building-Resolving Simulations. *J. Appl. Meteorol. Climatol.* 47:1553–1572.
- Cimorelli, A.J., Perry, S.G., Venkatram, A., Weil, J.C., Paine, R.J., Wilson, R.B., Lee, R.F., Peters, W.D., Brode, R.W., and Paumier, J.O. (2004). AERMOD: Description of Model Formulation. Research Triangle Park, North Carolina.
- Conn, A.R., Gould, N.I.M., and Toint, P.L. (1991). Convergence of quasi-Newton matrices generated by the symmetric rank one update. *Math. Program.* 50:177–195.
- Cooper, O.R., Parrish, D.D., Stohl, A., Trainer, M., Nédélec, P., Thouret, V., Cammas, J.P., Oltmans, S.J., Johnson, B.J., Tarasick, D., et al. (2010). Increasing springtime ozone mixing ratios in the free troposphere over western North America. *Nature.* 463:344–348.
- Crank, J. and Nicolson, P. (1947). A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Math. Proc. Cambridge Philos. Soc.* 43:50–67.
- Dakins, M.E., Toll, J.E., Small, M.J., and Brand, K.P. (1996). Risk-based environmental remediation: Bayesian Monte Carlo analysis and the expected value of sample information. *Risk Anal.* 16:67–79.
- Davidon, W.C. (1991). Variable Metric Method for Minimization. *SIAM J. Optim.* 1:1–17.
- Donea, J. and Huerta, A. (2003). *Finite Element Methods for Flow Problems*. Chichester, Wiley.
- Elbern, H., Schmidt, H., and Ebel, A. (1997). Variational data assimilation for tropospheric chemistry modeling. *J. Geophys. Res.* 102:15967–15985.
- Elbern, H., Schmidt, H., Talagrand, O., and Ebel, a. (2000). 4D-variational data assimilation with an adjoint air quality model for emission analysis. *Environ. Model. Softw.* 15:539–548.

- Enting, I.G. and Newsam, G.N. (1990). Atmospheric Constituent Inversion Problems: Implications for Baseline Monitoring. *J. Atmos. Chem.* 11:69–87.
- EPA (1995). Method 21 - Determination of Volatile Organic Compound Leaks.
- EPA (2006). Other Test Method 10 - Optical Remote Sensing for Emission Characterization from Non-point Sources.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* 99:10143.
- Fick, A. (1855). Über Diffusion [Translated: (1995) “On liquid diffusion”. Journal of Membrane Science]. *Poggendorff's Ann. der Phys. und Chemie.* 94:59–86.
- Fisher, M., Nocedal, J., Trémolet, Y., and Wright, S.J. (2008). Data assimilation in weather forecasting: a case study in PDE-constrained optimization. *Optim. Eng.* 10:409–426.
- Fletcher, R. (1970). A new approach to variable metric algorithms. *Comput. J.* 13:317–322.
- Fletcher, R. and Powell, M.J.D. (1963). A Rapidly Convergent Descent Method for Minimization. *Comput. J.* 6:163–168.
- Franke, J., Hellsten, A., Schlünzen, H., and Carissimo, B. (2007). Best practice guideline for the CFD simulation of flows in the urban environment. Brussels. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:BEST+PRACTICE+GUIDELINE+FOR+THE+CFD+SIMULATION+OF+FLOW+IN+THE+URBAN+ENVIRONMENT#0> [November 18, 2010]
- Gelb, A. (1974). *Applied optimal estimation*. Cambridge, M.I.T. Press.
- Ghil, M. and Malanotte-Rizzoli, P. (1991). Data assimilation in meteorology and oceanography. *Adv. Geophys.* 33:141–266.
- Giering, R. (1999). Tangent linear and Adjoint Model Compiler: Users manual.
- Gilbert, J.C. and Lemaréchal, C. (1989). Some numerical experiments with variable-storage quasi-Newton algorithms. *Math. Program.* 45:407–435.
- Giles, M.B. and Pierce, N.A. (2000). An Introduction to the Adjoint Approach to Design. *Flow, Turbul. Combust.* 65:393–415.
- Gresho, P.M. and Lee, R.L. (1979). Don't Suppress the Wiggles - They're Telling You Something. *Comput. Fluids.* 9:223–253.

- Guo, S., Yang, R., Zhang, H., Weng, W., and Fan, W. (2009). Source identification for unsteady atmospheric dispersion of hazardous materials using Markov Chain Monte Carlo method. *Int. J. Heat Mass Transf.* 52:3955–3962.
- Haas-Laursen, D.E., Hartley, D.E., and Prinn, R.G. (1996). Optimizing an inverse method to deduce time-varying emissions of trace gases. *J. Geophys. Res.* 101:22823.
- Hairer, E., Nørsett, S.P., and Wanner, G. (1993). *Solving ordinary differential equations* 2nd ed. Berlin, Springer-Verlag.
- Hartley, D. and Prinn, R. (1993). Feasibility of Determining Surface Emissions of Trace Gases Using an Inverse Method in a Three-Dimensional Chemical Transport Model. *J. Geophys. Res.* 98:5183–5197.
- Hashmonay, R.A., Natschke, D.F., Wagoner, K., Harris, D.B., Thompson, E.L., and Yost, M.G. (2001). Field evaluation of a method for estimating gaseous fluxes from area sources using open-path Fourier transform infrared. *Environ. Sci. Technol.* 35:2309–13.
- Haupt, S.E., Young, G.S., and Allen, C.T. (2007). A Genetic Algorithm Method to Assimilate Sensor Data for a Toxic Contaminant Release. *J. Comput.* 2:85–93.
- Holland, J.H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. Ann Arbor: University of Michigan Press.
- Hooke, R. and Jeeves, T.A. (1961). "Direct Search" Solution of Numerical and Statistical Problems. *J. ACM.* 8:212–229.
- Hourdin, F., Talagrand, O., and Idelkadi, A. (2006). Eulerian backtracking of atmospheric tracers. II: Numerical aspects. *Q. J. R. Meteorol. Soc.* 132:585–603.
- Houweling, S., Kamsink, T., Dentener, F., Lelieveld, J., and Heimann, M. (1999). Inverse modeling of methane sources and sinks using the adjoint of a global transport model. *J. Geophys. Res.* 104:26137–26160.
- Huebner, K.H., Dewhirst, D.L., Smith, D.E., and Byron, T.G. (2001). *The Finite Element Method for Engineers* 4th ed. New York, Wiley.
- Hughes, E. (2009). The Adjoint Method for Derivatives. . 1–7.
- Huyakorn, P.S. (1977). Solution of steady-state, convective transport equation using an upwind finite element scheme. *Appl. Math. Model.* 1:187–195.

- Ide, K., Courtier, P., Ghil, M., and Lorenc, A.C. (1997). Unified Notation for Data Assimilation: Operational, Sequential and Variational. *J. Meteorological Soc. Japan*. 75:181–189.
- Issartel, J.-P. and Baverel, J. (2003). Inverse transport for the verification of the Comprehensive Nuclear Test Ban Treaty. *Atmos. Chem. Phys.* 3:475–486.
- Ivan, L. (2010). Development of High-Order CENO Finite-Volume Schemes with Block-Based Adaptive Mesh Refinement. University of Toronto.
- Jeong, U.Y., Koh, H.-M., and Lee, H.S. (2002). Finite element formulation for the analysis of turbulent wind flow passing bluff structures using the RNG k-epsilon model. *J. Wind Eng. Ind. Aerodyn.* 90:151–169.
- Jones, W.P. and Launder, B.E. (1972). The Prediction of Laminarization with a Two-Equation Model of Turbulence. *Int. J. Heat Mass Transf.* 15:301–314.
- Joynes, I.M., Hakami, A., and Johnson, M.R. (2011). *Adjoint Sensitivity Analysis for Fugitive Emissions Characterization on a Building Scale*. In Air & Waste Management Association 104th Annual Conference. Orlando, FL.
- Joynes, I.M., Hakami, A., and Johnson, M.R. (2012). *Inverse Dispersion Modelling for Fugitive Emissions Quantification within an Industrial Geometry*. In Air & Waste Management Association (AWMA) 105th Annual Conference. San Antonio, TX.
- Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Trans. ASME-Journal Basic Eng.* 82:35–45.
- Keats, A., Yee, E., and Lien, F.-S. (2007a). Bayesian inference for source determination with applications to a complex urban environment. *Atmos. Environ.* 41:465–479.
- Keats, A., Yee, E., and Lien, F.-S. (2007b). Efficiently characterizing the origin and decay rate of a nonconservative scalar using probability theory. *Ecol. Modell.* 205:437–452.
- Keeler, G.J. and Samson, P.J. (1989). Spatial representativeness of trace element ratios. *Environ. Sci. Technol.* 23:1358–1364.
- Kirchgessner, D.A., Lott, R.A., Cowgill, R.M., Harrison, M.R., and Shires, T.M. (1997). Estimate of methane emissions from the U.S. natural gas industry. *Chemosphere*. 35:1365–90.
- Kirkpatrick, S., Gelatt, C.D., and Vecchi, M.P. (1983). Optimization by simulated annealing. *Science*. 220:671–680.

- Klug, W., Graziani, G., Grippa, G., Pierce, D., and Tassone, C. (1992). Evaluation of Long Range Atmospheric Transport Models Using Environmental Radioactivity Data from the Chernobyl Accident: The ATMES Report.
- Kuzmin, D. (2010). A Guide to Numerical Methods for Transport Equations.
- Launder, B.E. and Sharma, B.I. (1974). Application of the Energy-Dissipation Model of Turbulence to the Calculation of Flow Near a Spinning Disc. *Lett. Heat Mass Transf.* 1:131–138.
- Lee, P.M. (1989). *Bayesian statistics: an introduction*. New York, Oxford Press.
- Liu, D.C. and Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. *Math. Program.* 45:503–528.
- Lyness, J.N. and Moler, C.B. (1967). Numerical Differentiation of Analytic Functions. *SIAM J. Numer. Anal.* 4:202–210.
- Lyster, P.M., Cohn, S.E., Ménard, R., Chang, L.-P., Lin, S.-J., and Olsen, R.G. (1997). Parallel Implementation of a Kalman Filter for Constituent Data Assimilation. *Mon. Weather Rev.* 125:1674.
- Madala, S., Satyanarayana, A.N.V., and Krishna Prasad, V. (2012). Micro-scale Dispersion of Air Pollutants over an Urban Setup in a Coastal Region. *Open J. Air Pollut.* 01:51–58.
- Magnus, J.R. and Neudecker, H. (1988). *Matrix differential calculus with applications in statistics and econometrics*. Chichester, Wiley.
- Marchuk, G.I. (1986). *Mathematical Models in Environmental Problems*. Imprint: North-holland.
- Marchuk, G.I. (1994). *Adjoint Equations and Analysis of Complex Systems*. Dordrecht, Kluwer Academic Publishers.
- McRae, G.J., Goodin, W.R., and Seinfeld, J.H. (1982). Numerical Solution of the Atmospheric Diffusion Equation for Chemically Reacting Flows. . 42:1–42.
- Mikhlin, S.G. (1964). *Variational methods in mathematical physics*. New York, Macmillan.
- Mulholland, M. and Seinfeld, J.H. (1995). Inverse air pollution modelling of urban-scale carbon monoxide emissions. *Atmos. Environ.* 29:497–516.
- Nelder, J.A. and Mead, R. (1965). A simplex method for function minimization. *Comput. J.* 7:308–313.

- Newsam, G.N. and Enting, I.G. (1988). Inverse problems in atmospheric constituent studies. I. Determination of surface sources under a diffusive transport approximation. *Inverse Probl.* 4:1037–1054.
- Nocedal, J. (1980). Updating Quasi-Newton Matrices with Limited Storage. *Math. Comput.* 35:773.
- Nocedal, J. and Wright, S.J. (2006). *Numerical Optimization* 2nd ed. Springer.
- Nolle, L. (2006). On a Hill-Climbing Algorithm with Adaptive Step Size: Towards a Control Parameter-Less Black-Box Optimisation Algorithm. *Comput. Intell. Theory Appl.* 38:587–595.
- NREL (2011). National Wind Technology Center M2 Tower.
http://www.nrel.gov/midc/nwtc_m2/
- Penenko, V. V., Baklanov, A., and Tsvetova, E. (2002). Methods of sensitivity theory and inverse modeling for estimation of source parameters. *Futur. Gener. Comput. Syst.* 18:661–671.
- Picard, D.J., Stribny, M., and Harrison, M.R. (1998). Handbook for Estimating Methane Emissions From Canadian Natural Gas Systems. Guelph.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (1997). *Numerical Recipes in C* Second Edi. Cambridge.
- Pudykiewicz, J.A. (1998). Application of Adjoint Tracer Transport Equations for Evaluating Source Parameters. *Atmos. Environ.* 32:3039–3050.
- Qian, S.S., Stow, C.A., and Borsuk, M.E. (2003). On Monte Carlo methods for Bayesian inference. *Ecol. Modell.* 159:269–277.
- Rao, K.S. (2007). Source estimation methods for atmospheric dispersion. *Atmos. Environ.* 41:6964–6973.
- Reynolds, O. (1895). On the Dynamical Theory of Incompressible Viscous Fluids and the Determination of the Criterion. *Philos. Trans. R. Soc. London. A.* 186:123–164.
- Robertson, L. and Langner, J. (1998). Source function estimate by means of a variational data assimilation applied to the ETEX-I tracer experiment. *Atmos. Environ.* 32:4219–4225.
- Rosenbrock, H.H. (1960). An Automatic Method for Finding the Greatest or Least Value of a Function. *Comput. J.* 3:175–184.

- Rosenbrock, H.H. (1963). Some general implicit processes for the numerical solution of differential equations. *Comput. J.* 5:329–330.
- Ruby, M. (1987). Introduction to Atmospheric Dispersion Modeling. Available at: <http://www.envirometrics.com/abstracts/dismodel.pdf>
- Sandu, A., Daescu, D.N., Carmichael, G.R., and Chai, T. (2005). Adjoint sensitivity analysis of regional air quality models. *J. Comput. Phys.* 204:222–252.
- Schey, H.M. (1973). *Div, Grad, Curl and all that; an informal text on vector calculus*. New York, Norton.
- Schoonbaert, S.B. and Johnson, M.R. (2011). Fibre-Based TDLs Sensor Network for Detecting Fugitive Emission Sources. *Air Waste Manag. Assoc. 104th Annu. Conf. Exhib.* 1–6.
- Scire, J.S., Strimaitis, D.G., and Yamartino, R.J. (1990). Model Formulation and User's Guide for the CALPUFF Dispersion Model. Sacramento, California 95814.
- Seibert, P. (2000). *Methods for source determination in the context of the CTBT radionuclide monitoring system*. In Proceedings Informal Workshop on Meteorological Modelling in Support of CTBT Verification. Vienna, pp.1–6.
- Seibert, P. (2001). *Inverse Modelling with a Lagrangian Particle Dispersion Model: Application to Point Releases Over Limited Time Intervals*. In S.-E. Gryning & F. A. Schiermeier, eds. *Air Pollution Modeling and Its Applications XIV*. Kluwer Academic Publishers, pp.381–389.
- Seibert, P. and Frank, A. (2004). Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode. *Atmos. Chem. Phys.* 4:51–63.
- Skiba, Y.N. (2003). On a method of detecting the industrial plants which violate prescribed emission rates. *Ecol. Modell.* 159:125–132.
- Skiba, Y.N. and Parra-Guevara, D. (2000). Industrial pollution transport . Part 1 . Formulation of the problem and air pollution estimates. . 5:169–175.
- Slade, D.H. (1968). *Meteorology and Atomic Energy*.
- Sofiev, M. and Atlaskin, E. (2007). An Example of Application of Data Assimilation Technique and Adjoint Modelling to an Inverse Dispersion Problem Based on the ETEX Experiment. *Air Pollut. Model. Its Appl. XVII.* 438–448.
- Sohn, M.D., Sextro, R.G., Gadgil, A.J., and Daisey, J.M. (2003). Responding to sudden pollutant releases in office buildings: 1. Framework and analysis tools. *Indoor Air.* 13:267–276.

- Sportisse, B. (2001). Box models versus Eulerian models in air pollution modeling. *Atmos. Environ.* 35:173–178.
- Squire, W. and Trapp, G. (1998). Using complex variables to estimate derivatives of real functions.pdf. *SIAM Rev.* 40:110–112.
- Stewart, J. (1991). *Single variable calculus : early transcendentals* 2nd ed. Pacific Grove, California, Brooks/Cole Publishing Company.
- Stewart, J. (2003). *Multivariable Calculus* 5th ed. United States of America, Brooks/Cole.
- Stockie, J.M. (2011). The Mathematics of Atmospheric Dispersion Modeling. *SIAM Rev.* 53:349.
- Stohl, A. (1996). Trajectory statistics-A new method to establish source-receptor relationships of air pollutants and its application to the transport of particulate sulfate in Europe. *Atmos. Environ.* 30:579–587.
- 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. *Atmos. Environ.* 36:4635–4648.
- Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G. (2005). Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys. Discuss.* 5:4739–4799.
- Stohl, A., Hittenberger, M., and Wotawa, G. (1998). Validation of the Lagrangian particle dispersion model FLEXPART against large scale tracer experiment data. *Atmos. Environ.* 32:4245–4264.
- Stohl, A., Wotawa, G., Seibert, P., and Kromp-Kolb, H. (1995). Interpolation Errors in Wind Fields as a Function of Spatial and Temporal Resolution and Their Impact on Different Types of Kinematic Trajectories. *J. Appl. Meteorol.* 34:2149–2165.
- Sutton, O.G. (1932). A Theory of Eddy Diffusion in the Atmosphere. *Proc. R. Soc. London. Ser. A, Contain. Pap. a Math. Phys. Character.* 135:143–165.
- Thomson, L.C., Hirst, B., Gibson, G., Gillespie, S., Jonathan, P., Skeldon, K.D., and Padgett, M.J. (2007). An improved algorithm for locating a gas source using inverse methods. *Atmos. Environ.* 41:1128–1134.
- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B.* 58:276–288.

- Tortorelli, D.A. and Michaleris, P. (1994). Design sensitivity analysis: Overview and review. *Inverse Probl.* 1:71–105.
- Turns, S.R. (1996). *An Introduction to Combustion: Concepts and Applications* 2nd ed. McGraw-Hill Higher Education.
- Versteeg, H.K. and Malalasekera, W. (2007). *An Introduction to Computational Fluid Dynamics: The Finite Volume Method* Second. Glasgow, Pearson Education Limited.
- Whitenett, G., Stewart, G., Atherton, K., Culshaw, B., and Johnstone, W. (2003). Optical fibre instrumentation for environmental monitoring applications. *J. Opt. A Pure Appl. Opt.* 5:S140–S145.
- Wilcox, D.C. (1993). Turbulence Modeling for CFD.
- Wotawa, G., Geer, L. De, Denier, P., Kalinowski, M., Toivonen, H., Amours, R.D., Desiato, F., Issartel, J., Langer, M., Seibert, P., et al. (2003). Atmospheric transport modelling in support of CTBT verification — overview and basic concepts. *Environment*. 37:2529–2537.
- Yakhot, V. and Orszag, S.A. (1986). Renormalization group analysis of turbulence. I. Basic theory. *J. Sci. Comput.* 1:3–51.
- Yakhot, V., Orszag, S.A., Thangam, S., Gatski, T.B., and Speziale, C.G. (1992). Development of turbulence models for shear flows by a double expansion technique. *Phys. Fluids A Fluid Dyn.* 4:1510–1520.
- Yumimoto, K. and Uno, I. (2006). Adjoint inverse modeling of CO emissions over Eastern Asia using four-dimensional variational data assimilation. *Atmos. Environ.* 40:6836–6845.
- Zhang, L. (2007). Large-Scale Simulations Using First and Second Order Adjoints with Applications in Data Assimilation. *Order A J. Theory Ordered Sets Its Appl.*
- Zheng, X. and Chen, Z. (2010). Back-calculation of the strength and location of hazardous materials releases using the pattern search method. *J. Hazard. Mater.* 183:474–81.
- Zheng, X. and Chen, Z. (2011). Inverse calculation approaches for source determination in hazardous chemical releases. *J. Loss Prev. Process Ind.* 24:293–301.
- Zhu, C., Byrd, R.H., Lu, P., and Nocedal, J. (1997). Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans. Math. Softw.* 23:550–560.

Zienkiewicz, O.C. (1971). *The Finite Element Method in Engineering Science*. McGraw-Hill.

Appendix A

Continuous Sensitivity Analysis

A.1 Objective Function

The objective function presented in Section 2.2 was formulated from the discrete representations of the pollutant concentration $\mathbf{c}(\mathbf{x})$ field and emission source field $\mathbf{s}(\mathbf{x})$. Alternatively, a scalar valued ($n_f = 1$) objective function can be formed from continuous representations of these fields ($c(\mathbf{r}, t; \mathbf{x})$ and $s(\mathbf{r}, t; \mathbf{x})$). The general form of this continuous objective function $f(\mathbf{x})$ is given by (A.1),

$$f(\mathbf{x}) = \int_{t_0}^{t_f} \int_{\Omega} g(c(\mathbf{r}, t; \mathbf{x}), s(\mathbf{r}, t; \mathbf{x})) d\Omega dt \quad (\text{A.1})$$

where $g(c(\mathbf{r}, t; \mathbf{x}), s(\mathbf{r}, t; \mathbf{x}))$ is given by (A.2) as a positive function that comprises an observational term and a regularization term.

$$\begin{aligned} g(c(\mathbf{r}, t; \mathbf{x}), s(\mathbf{r}, t; \mathbf{x})) &= \frac{1}{2} H(\mathbf{r}, t) (c(\mathbf{r}, t; \mathbf{x}) - c^o(\mathbf{r}, t))^2 \\ &+ \frac{\beta}{2} (s(\mathbf{r}, t; \mathbf{x}) - s^b(\mathbf{r}, t))^2 \end{aligned} \quad (\text{A.2})$$

Similar to the discrete objective function (2.2), the observational term $(\frac{1}{2} (H(\mathbf{r}, t) c(\mathbf{r}, t; \mathbf{x}) - c^o(\mathbf{r}, t))^2)$ measures of the misfit between the predicted receptor

measurements ($H(\mathbf{r}, t)c(\mathbf{r}, t; \mathbf{x})$) and the observed (or synthetic reference) receptor observations ($c^o(\mathbf{r}, t)$), and the regularization term $(\frac{\beta}{2}(s(\mathbf{r}, t; \mathbf{x}) - s^b(\mathbf{r}, t))^2)$ improves the underdetermined or ill-conditioned quality of inverse modelling by imposing an additional “soft” constraint that penalizes the variance of the predicted emissions source distribution $s(\mathbf{r}, t; \mathbf{x})$ from an initial (or background) estimate $s^b(\mathbf{r}, t)$ (Marchuk, 1994; Seibert, 2001).

The regularization parameter β ensures that the units of the observational and regularization terms match, and in the present application has units of $[\text{s} \cdot \text{m}^{-3}]$. The value of β is subjective and governs how the inverse model assigns the relative importance of matching predicted and observed receptor readings (i.e. minimizing the observational term) versus satisfying the additional soft constraints of the regularization (e.g. minimizing the variance from an assumed background).

The observation function $H(\mathbf{r}, t)$ simulates receptor measurements from the predicted concentration field $c(\mathbf{r}, t; \mathbf{x})$. This observation function can be formed from Dirac or Radon measures (Penenko et al., 2002). A candidate observation function that mimics point measurements of receptors is given by (A.3) as a product of Dirac measures,

$$H(\mathbf{r}, t) = \sum_{i=0}^{n_t} \sum_{j=1}^{n_r} \delta(t - t_i^r) \delta(x - x_j^r) \delta(y - y_j^r) \delta(z - z_j^r) \quad (\text{A.3})$$

where x_j^r , y_j^r and z_j^r represent the coordinates of the j th receptor in the spatial domain; t_i^r represents i th sample time of the receptors; and $\delta(x)$ is the Dirac delta function, which is

defined as a Gaussian or normal distribution with standard deviation approaching 0 in the limit (A.4).

$$\delta(x) = \lim_{a \rightarrow 0^+} \frac{1}{a\sqrt{\pi}} e^{-\left(\frac{x}{a}\right)^2} \quad (\text{A.4})$$

The Dirac delta function has a convenient sifting property (A.5), which extracts point measurements from a continuous function $c(x)$.

$$\int_{-\infty}^{\infty} c(x) \delta(x - x_i^r) dx = c(x_i^r) \quad (\text{A.5})$$

Finally, the 1/2 factor on both observational and regularization terms of (A.1) are convenient scaling factors that simplify the evaluation of the objective function's gradient.

The emission source parameters that minimize the objective function can be found with a numerical optimization algorithm. For objective functions with many parameters, a gradient-based numerical optimization algorithm such as L-BFGS-B (Byrd et al., 1994) is preferable. However, as the name suggests, a gradient-based optimization algorithm requires the evaluation of the gradient of the objective function.

A.2 Forward Sensitivity Method

The sensitivity of the objective function with respect to all n_x emission source parameters is represented by the following row vector of objective function partial derivatives (A.6),

$$\frac{df(\mathbf{x})}{d\mathbf{x}} = \left[\frac{\partial f(\mathbf{x})}{\partial x_1} \quad \cdots \quad \frac{\partial f(\mathbf{x})}{\partial x_{n_x}} \right] \quad (\text{A.6})$$

and the gradient of the objective function (A.1) is simply the transpose of (A.6) given as (A.7).

$$\nabla f(\mathbf{x}) = \left(\frac{df(\mathbf{x})}{d\mathbf{x}} \right)^T \quad (\text{A.7})$$

Differentiating the objective function (A.1) with respect to the vector of emission source parameters $\mathbf{x}$ gives (A.8).

$$\frac{df(\mathbf{x})}{d\mathbf{x}} = \int_{t_0}^{t_f} \int_{\Omega} \frac{\partial g}{\partial c} \frac{dc(\mathbf{x})}{d\mathbf{x}} + \frac{\partial g}{\partial s} \frac{ds(\mathbf{x})}{d\mathbf{x}} d\Omega dt \quad (\text{A.8})$$

For the specific objective function suggested by (A.2), the chain rule terms are scalar fields (A.9) and (A.10), and row vectors of scalar fields (A.11) and (A.12).

$$\frac{\partial g}{\partial c} = \frac{\partial g(\mathbf{r}, t; \mathbf{x})}{\partial c} = H(\mathbf{r}, t) (c(\mathbf{r}, t; \mathbf{x}) - c^o(\mathbf{r}, t)) \quad (\text{A.9})$$

$$\frac{\partial g}{\partial s} = \frac{\partial g(\mathbf{r}, t; \mathbf{x})}{\partial s} = \beta (s(\mathbf{r}, t; \mathbf{x}) - s^b(\mathbf{r}, t)) \quad (\text{A.10})$$

$$\frac{ds}{d\mathbf{x}} = \frac{ds(\mathbf{r}, t; \mathbf{x})}{d\mathbf{x}} = \left[\frac{\partial s(\mathbf{r}, t; \mathbf{x})}{\partial x_1} \quad \cdots \quad \frac{\partial s(\mathbf{r}, t; \mathbf{x})}{\partial x_{n_x}} \right] \quad (\text{A.11})$$

$$\frac{dc}{d\mathbf{x}} = \frac{dc(\mathbf{r}, t; \mathbf{x})}{d\mathbf{x}} = \left[\frac{\partial c(\mathbf{r}, t; \mathbf{x})}{\partial x_1} \quad \cdots \quad \frac{\partial c(\mathbf{r}, t; \mathbf{x})}{\partial x_{n_x}} \right] \quad (\text{A.12})$$

The elements that form the vector of pollutant concentration field sensitivities to the emission source parameters ($dc/d\mathbf{x}$) (A.12) are determined by differentiating the advection-diffusion equation (3.1) with respect to the each of the n_x emission source

parameters. Equation (A.13) shows the advection-diffusion (3.1) equation differentiated with respect to the i th emission source parameter (x_i),

$$\frac{\partial}{\partial t} \left(\frac{\partial c}{\partial x_i} \right) + \nabla \cdot \left(\frac{\partial c}{\partial x_i} \mathbf{u} - \mathbf{D} \nabla \left(\frac{\partial c}{\partial x_i} \right) \right) = \frac{\partial s}{\partial x_i} \quad (\text{A.13})$$

where, similar to the concentration field $c = c(\mathbf{r}, t; \mathbf{x})$, the sensitivity of the concentration field to the i th emission source parameter $\frac{\partial c}{\partial x_i} = \frac{\partial c}{\partial x_i}(\mathbf{r}, t; \mathbf{x})$ is also a function of space, time, and the emission source parameters.

The unique solution of this pollutant concentration sensitivity transport PDE (A.13) requires initial and boundary conditions. The initial conditions of (A.13) are determined by differentiating the initial conditions of the advection-diffusion equation (3.2) with respect to the i th emission source parameter (x_i) to give (A.14).

$$\frac{\partial c(\mathbf{r}, t_0; \mathbf{x})}{\partial x_i} = \frac{\partial c_0(\mathbf{r}; \mathbf{x})}{\partial x_i} \quad (\text{A.14})$$

If the initial conditions of the advection-diffusion equation are not dependant on the i th emission source parameter, then the initial conditions of pollutant concentration sensitivity PDE are zero.

The boundary conditions of (A.13) are derived by differentiating the boundary conditions of the advection-diffusion equation (3.3) with respect to the i th emission source parameter (x_i) to give (A.15).

$$\begin{aligned}
\left(\frac{\partial c}{\partial x_i} \mathbf{u} - \mathbf{D} \nabla \left(\frac{\partial c}{\partial x_i} \right) \right) \cdot \hat{\mathbf{n}} &= 0 \quad \text{on } \Gamma \text{ for } \mathbf{u} \cdot \hat{\mathbf{n}} < 0 \\
\left(\mathbf{D} \nabla \left(\frac{\partial c}{\partial x_i} \right) \right) \cdot \hat{\mathbf{n}} &= 0 \quad \text{on } \Gamma \text{ for } \mathbf{u} \cdot \hat{\mathbf{n}} \geq 0
\end{aligned} \tag{A.15}$$

The evaluation of $dc/d\mathbf{x}$ (A.12) requires the solution of (A.13) for all of the $n_{\mathbf{x}}$ emission source parameters, which could require considerable effort if $n_{\mathbf{x}}$ is large. If $n_f < n_{\mathbf{x}}$ then it is more efficient to consider the adjoint sensitivity method to evaluate the gradient of the objective function.

A.3 Adjoint Sensitivity Method

The continuous adjoint sensitivity method is an alternative method for determining the gradient of the objective function with respect to the emission source parameters. This method mimics the reverse Jacobian accumulation mode of automatic differentiation (Bischof et al., 1994) and is computationally efficient when the dimension of objective function outputs (scalar-valued f , $n_f = 1$) is lower than the dimension of the inputs (vector-valued $\mathbf{x}$).

The continuous adjoint sensitivity method is derived by augmenting the same scalar-valued objective function (A.1) with the inner product of an adjoint field variable λ and the residual error of the advection-diffusion equation (3.1) (also referred to as the problem constraints) to create an alternate but equivalent objective function $\Lambda(\mathbf{x})$ (A.16).

$$\Lambda(\mathbf{x}) = \int_{t_0}^{t_f} \int_{\Omega} g(c(\mathbf{x}), s(\mathbf{x})) - \lambda \left(\frac{\partial c}{\partial t} + \nabla \cdot (c\mathbf{u} - \mathbf{D}\nabla c) - s \right) d\Omega dt \quad (\text{A.16})$$

It should be noted that the value of this augmented term is zero and has no influence on the value or derivatives of the objective function. Differentiating this augmented objective function with respect to the i th emission source parameter gives (A.17).

$$\begin{aligned} \frac{\partial \Lambda(\mathbf{x})}{\partial x_i} = & \int_{t_0}^{t_f} \int_{\Omega} \frac{\partial g}{\partial c} \frac{\partial c}{\partial x_i} + \frac{\partial g}{\partial s} \frac{\partial s}{\partial x_i} \\ & - \lambda \left(\frac{\partial}{\partial t} \left(\frac{\partial c}{\partial x_i} \right) + \nabla \cdot \left(\frac{\partial c}{\partial x_i} \mathbf{u} - \mathbf{D} \nabla \left(\frac{\partial c}{\partial x_i} \right) \right) - \frac{\partial s}{\partial x_i} \right) d\Omega dt \end{aligned} \quad (\text{A.17})$$

As discussed in the previous section, the function $\partial c / \partial x_i$ represents the sensitivity of concentration to the i th emission source parameter, and is given by the solution to (A.13). However, determining the sensitivity of pollutant concentrations to each of the emission source parameters is usually not practical for large numbers of emission source parameters. Alternatively, it is convenient to rearrange the terms of the sensitivity of the augmented objective function (A.16) with integration by parts (A.18) (Stewart, 1991), Divergence theorem (A.19) (Stewart, 2003), and a vector calculus identity for the divergence of the product of a scalar and vector function (A.20) (Schey, 1973),

$$\int_{t_0}^{t_f} \varphi \frac{d\psi}{dt} dt = \varphi \psi \Big|_{t=t_0}^{t=t_f} - \int_{t_0}^{t_f} \varphi \frac{d\psi}{dt} dt \quad (\text{A.18})$$

$$\int_{\Omega} (\nabla \cdot \Psi) d\Omega = \int_{\Gamma} (\Psi \cdot \hat{\mathbf{n}}) d\Gamma \quad (\text{A.19})$$

$$\nabla \cdot (\varphi \Psi) = \varphi (\nabla \cdot \Psi) + \Psi \cdot (\nabla \varphi) \quad (\text{A.20})$$

where $\varphi = \varphi(\mathbf{r}, t)$, $\psi = \psi(\mathbf{r}, t)$ are arbitrary scalar functions and $\mathbf{\Psi} = \mathbf{\Psi}(\mathbf{r}, t) = [\psi_1(\mathbf{r}, t) \ \psi_2(\mathbf{r}, t) \ \psi_3(\mathbf{r}, t)]^T$ is an arbitrary vector function.

The evolution operator term, $(-\lambda \frac{\partial}{\partial t} (\frac{\partial c}{\partial s_i}))$, of (A.17) can be integrated in time by parts (A.18) to give (A.21).

$$\begin{aligned} & - \int_{t_0}^{t_f} \int_{\Omega} \lambda \frac{\partial}{\partial t} \left(\frac{\partial c}{\partial x_i} \right) d\Omega dt \\ & = - \int_{\Omega} \lambda \frac{\partial c}{\partial x_i} \Big|_{t=t_0}^{t=t_f} d\Omega + \int_{t_0}^{t_f} \int_{\Omega} \frac{\partial \lambda}{\partial t} \frac{\partial c}{\partial x_i} d\Omega dt \end{aligned} \quad (\text{A.21})$$

The flux operator $(-\lambda \nabla \cdot (\frac{\partial c}{\partial x_i} \mathbf{u} - \mathbf{D} \nabla (\frac{\partial c}{\partial x_i})))$ of (A.17) can be integrated in space using the Divergence theorem (A.19), and the vector calculus identity (A.20) applied to give (A.22).

$$\begin{aligned} & - \int_{t_0}^{t_f} \int_{\Omega} \lambda \nabla \cdot \left(\frac{\partial c}{\partial x_i} \mathbf{u} - \mathbf{D} \nabla \left(\frac{\partial c}{\partial x_i} \right) \right) d\Omega dt \\ & = - \int_{t_0}^{t_f} \int_{\Gamma} \lambda \left(\frac{\partial c}{\partial x_i} \mathbf{u} - \mathbf{D} \nabla \left(\frac{\partial c}{\partial x_i} \right) \right) \cdot \hat{\mathbf{n}} d\Gamma dt \\ & \quad + \int_{t_0}^{t_f} \int_{\Omega} \left(\frac{\partial c}{\partial x_i} \mathbf{u} - \mathbf{D} \nabla \left(\frac{\partial c}{\partial x_i} \right) \right) \cdot (\nabla \lambda) d\Omega dt \end{aligned} \quad (\text{A.22})$$

Considering only the diffusion term of the last term of (A.22) gives (A.23).

$$- \int_{t_0}^{t_f} \int_{\Omega} \left(\mathbf{D} \nabla \left(\frac{\partial c}{\partial x_i} \right) \right) \cdot (\nabla \lambda) d\Omega dt \quad (\text{A.23})$$

It should be noted that the diffusivity matrix $\mathbf{D}$ can be moved from the gradient of the pollutant concentration sensitivity ($\nabla(dc/dx_i)$) to the gradient of the adjoint variable ($\nabla\lambda$) with the identity (A.24).

$$\left(\mathbf{D} \nabla \left(\frac{\partial c}{\partial x_i} \right) \right) \cdot (\nabla \lambda) = \left(\nabla \left(\frac{\partial c}{\partial x_i} \right) \right) \cdot (\mathbf{D}^T \nabla \lambda) \quad (\text{A.24})$$

However, since the mass diffusivity matrix is diagonal (see Section 3.1), it is also symmetric ($\mathbf{D} = \mathbf{D}^T$), which means that the transpose operation has no influence, but is retained for the generality of the method.

An additional application of the divergence theorem (A.19) and the vector calculus identity (A.20) to (A.23) and (A.24) gives (A.25).

$$\begin{aligned} & \int_{t_0}^{t_f} \int_{\Omega} \left(\nabla \left(\frac{\partial c}{\partial x_i} \right) \right) \cdot (-\mathbf{D}^T \nabla \lambda) d\Omega dt \\ &= - \int_{t_0}^{t_f} \int_{\Gamma} \frac{\partial c}{\partial x_i} \left(\nabla \cdot (\mathbf{D}^T \nabla \lambda) \right) \cdot \hat{\mathbf{n}} d\Gamma dt \\ &+ \int_{t_0}^{t_f} \int_{\Omega} \frac{\partial c}{\partial x_i} \left(\nabla \cdot (\mathbf{D}^T \nabla \lambda) \right) d\Omega dt \end{aligned} \quad (\text{A.25})$$

Combining the results of (A.17), (A.21), (A.22), and (A.25) yields (A.26).

$$\begin{aligned}
\frac{\partial \Lambda}{\partial x_i}(\mathbf{x}) = & \int_{t_0}^{t_f} \int_{\Omega} \left(\lambda + \frac{\partial g}{\partial s} \right) \frac{\partial s}{\partial x_i} \\
& + \left(\frac{\partial \lambda}{\partial t} + \mathbf{u} \cdot \nabla \lambda + \nabla \cdot (\mathbf{D}^T \nabla \lambda) + \frac{\partial g}{\partial c} \right) \frac{\partial c}{\partial x_i} d\Omega dt \\
& - \int_{\Omega} \lambda \frac{\partial c}{\partial x_i} \Big|_{t=t_0}^{t=t_f} d\Omega \\
& - \int_{t_0}^{t_f} \int_{\Gamma} \left(\lambda \left(\frac{\partial c}{\partial x_i} \mathbf{u} - \mathbf{D} \nabla \frac{\partial c}{\partial x_i} \right) + \frac{\partial c}{\partial x_i} (\mathbf{D}^T \nabla \lambda) \right) \cdot \hat{\mathbf{n}} d\Gamma dt
\end{aligned} \tag{A.26}$$

The second, third and fourth terms of (A.26) require knowledge of the evolution of the sensitivity of the pollutant concentration to the i th emission source parameter $(\partial c / \partial x_i)$. In Section A.2, it was recognized that determining the sensitivity of the pollutant concentration to all emission source parameters would require considerable effort. However, if λ is chosen such that the value of the second, third and fourth terms of (A.26) are zero, then there is no requirement to determine the sensitivity of the pollutant concentration to any of the emission source parameters. The adjoint variable λ can be chosen to be the solution to the adjoint equation (A.27), subject to the final value (A.28) and boundary conditions (A.29), such that the second and fourth terms and part of the third term of (A.26) have no contribution to the computation to the sensitivity of the augmented objective function,

$$-\frac{\partial \lambda}{\partial t} - \mathbf{u} \cdot \nabla \lambda - \nabla \cdot (\mathbf{D}^T \nabla \lambda) = \frac{\partial g}{\partial c} \tag{A.27}$$

$$\lambda(\mathbf{r}, t = t_f) = 0 \tag{A.28}$$

$$\begin{aligned}
(\mathbf{D}^T \nabla \lambda) \cdot \hat{\mathbf{n}} &= 0 & \text{on } \Gamma \text{ for } \mathbf{u} \cdot \hat{\mathbf{n}} < 0 \\
(\lambda \mathbf{u} + \mathbf{D}^T \nabla \lambda) \cdot \hat{\mathbf{n}} &= 0 & \text{on } \Gamma \text{ for } \mathbf{u} \cdot \hat{\mathbf{n}} \geq 0
\end{aligned} \tag{A.29}$$

where the adjoint forcing function $\partial g/\partial c$ is defined in (A.9). A similar adjoint equation was also presented by (Skiba, 2003) for incompressible flow with isotropic mass diffusion. Combining the sensitivity of the augmented objective function (A.26) with the initial condition (A.14) and boundary conditions (A.15) for the sensitivity of the pollutant concentration field to the emission source parameters and the adjoint equation (A.27) with its final condition (A.28) and boundary conditions (A.29), yields (A.30) as a simplified form of the sensitivity of the objective function to the emission source parameter x_i .

$$\begin{aligned}\frac{\partial f}{\partial x_i}(\mathbf{x}) &= \frac{\partial \Lambda}{\partial x_i}(\mathbf{x}) \\ &= \int_{t_0}^{t_f} \int_{\Omega} \left(\lambda + \frac{\partial g}{\partial s} \right) \frac{\partial s}{\partial x_i} d\Omega dt \\ &\quad + \int_{\Omega} \lambda(\mathbf{r}, t_0) \frac{\partial c}{\partial x_i}(\mathbf{r}, t_0; \mathbf{x}) d\Omega\end{aligned}\tag{A.30}$$

It should be noted that the solution of the adjoint equation (A.27) is independent of the number of emission source parameters, and only requires a single solution of λ for all of the objective function sensitivity evaluations (A.30). The gradient of the objective function with respect to the optimization parameters can be represented as (A.31),

$$\nabla f(\mathbf{x}) = \int_{t_0}^{t_f} \int_{\Omega} \left(\lambda + \frac{\partial g}{\partial s} \right) \nabla_{\mathbf{x}} s(\mathbf{x}) d\Omega dt + \int_{\Omega} \lambda(\mathbf{r}, t_0) \nabla_{\mathbf{x}} c_0(\mathbf{x}) d\Omega\tag{A.31}$$

where $\nabla_{\mathbf{x}} s(\mathbf{x}) = \nabla_{\mathbf{x}} s(\mathbf{r}, t; \mathbf{x}) = \left[\frac{\partial s(\mathbf{r}, t; \mathbf{x})}{\partial x_1} \quad \cdots \quad \frac{\partial s(\mathbf{r}, t; \mathbf{x})}{\partial x_{n_x}} \right]^T$ is a column vector of the sensitivity of the emission source function to the emission source parameters $\mathbf{x}$, and

$\nabla_{\mathbf{x}} c_0(\mathbf{x}) = \nabla_{\mathbf{x}} c(\mathbf{r}, t_0; \mathbf{x}) = \left[\frac{\partial c(\mathbf{r}, t_0; \mathbf{x})}{\partial x_1} \quad \cdots \quad \frac{\partial c(\mathbf{r}, t_0; \mathbf{x})}{\partial x_{n_x}} \right]^T$ is a column vector of the sensitivity

of the initial pollutant concentration field to the emission source parameters $\mathbf{x}$. For the case of an initial pollutant concentration field that has no dependence on the emission source parameters $\mathbf{x}$, the vector $\nabla_{\mathbf{x}}c_0(\mathbf{x})$ will be the zero-vector and there is no need to consider the second term of (A.31).

Furthermore, the divergence of the product of the scalar field λ and vector field $\mathbf{u}$ satisfies the identity (A.32).

$$\nabla \cdot (\lambda \mathbf{u}) = (\nabla \lambda) \cdot \mathbf{u} + \lambda (\nabla \cdot \mathbf{u}) \quad (\text{A.32})$$

If the flow is assumed to be incompressible, then the flow velocity vector field $\mathbf{u}$ can be assumed to be divergence free (A.33).

$$\nabla \cdot \mathbf{u} = 0 \quad (\text{A.33})$$

This condition permits (A.32) to be simplified to (A.34).

$$\nabla \cdot (\lambda \mathbf{u}) = \mathbf{u} \cdot (\nabla \lambda) \quad (\text{A.34})$$

Application of (A.34) to the advection term of the adjoint equation (A.27) gives the equivalent representation (A.35).

$$-\frac{\partial \lambda}{\partial t} + \nabla \cdot (-\lambda \mathbf{u} - \mathbf{D}^T \nabla \lambda) = \frac{\partial g}{\partial c} \quad (\text{A.35})$$

It should be noted that (A.35) is similar in form to the advection-diffusion equation (3.1), and be solved with a similar numerical solution algorithm with a sign change on the flow velocity vector field $\mathbf{u}$ where the solution is integrated backwards in time as a final value problem.

Another appealing property of the continuous adjoint sensitivity approach is that adaptive mesh refinement (AMR) procedures such as those suggested by Ivan (2010) can be invoked in the numerical solution of the continuous adjoint equation (A.27) to achieve better spatial resolution in regions of high gradients. Guo et al., (2009) demonstrated automatic mesh refinement with an adjoint method to efficiently predict receptor observations from a candidate emission source.

As with the advection-diffusion equation (3.1), the continuous adjoint variable transport PDE (A.27) can be discretized and integrated in time and space to produce a system of linear equations. This discretization and integration of the continuous adjoint equation (A.27) (or (A.35) for an incompressible flow) produces a system of linear equations and introduces truncation errors into the solution of λ , which in turn affect the gradient of the objective function (A.31). By contrast, the discrete form of the objective function (discussed in Section 2.2) already contains the truncation errors from the numerical solution of the discrete adjoint variable, and as such the gradient computed for the discrete objective function is closer to the true gradient. In the limit of an infinite resolution computational grid, both the discrete and continuous adjoint approach “should be consistent and converge to the correct analytic value for the gradient of the objective function” (Giles and Pierce, 2000).

Appendix B

Mesh Quality

ANSYS ICEM CFD, a mesh generation software package was used to create a triangular prism mesh of the air surrounding the buildings of a simplified gas plant, for modelling a two-dimensional transient wind field and pollution transport. For the purposes of the two-dimensional dispersion model, only the geometry of the triangular faces of the prism elements determine the quality of the mesh. The quality of a triangular element is determined as the minimum ratio of the triangle's height to base length for each side, normalized by $\sqrt{3}/2$ (the height to base ratio of an equilateral triangle) (ANSYS, 2009a). Figure B.1 shows a probability density function (PDF) and cumulative distribution function (CDF) of the quality of the triangular elements in the airshed domain mesh (shown in Figure 4.2).

Figure B.1: Probability density function of mesh quality

A minimum mesh quality of 0.3 is recommended, however 0.15 is sufficient for complicated models (ANSYS, 2009b). It can be seen from Figure B.1 that most of the elements were well above both minimum thresholds. The minimum element quality was 0.24, and that only 0.04% (74 elements) of the 174,772 elements within the mesh fell below the recommended 0.3 threshold.

Appendix C

ANSYS CFX Convergence

ANSYS CFX was used to generate a synthetic transient two-dimensional wind field. The convergence histories of the numerical solution of the Navier-Stokes equations (continuity and momentum) and turbulence model (RNG k - ε) are presented.

C.1 Initial Conditions Solution

The initial conditions to this transient wind field were evaluated with a pseudo-transient simulation in which the inflow conditions were fixed to the corresponding first time-step of the subsequent fully transient simulation. A target RMS relative residual error of 1×10^{-5} , a maximum of 300 iterations, and a pseudo time-step size of 0.2 s were selected for the initial conditions simulation. Figure C.1 and Figure C.2 show the convergence histories of the RMS and maximum relative residual errors respectively, of the flow field obtained from ANSYS CFX for the initial conditions of the transient wind simulation.

Figure C.1: Convergence history of RMS relative residual error of initial flow conditions

Figure C.2: Convergence history of maximum relative residual error of initial flow conditions

The convergence target was not achieved within the 300 iterations, but it is believed this may have been the result of the unsteady nature of vortex shedding from the buildings and the selection of too small a pseudo time-step.

C.2 Transient Solution

A fifteen minute transient simulation was initialized from the flow field predicted by the prior steady-inlet simulation. A time-step of 0.2 seconds was selected with a coefficient loop maximum of 20 iterations, and a residual error target of 1×10^{-5} was enforced on the RMS of the relative residual error on the mass and momentum equations.

Figure C.3 and Figure C.4 show the convergence histories of the RMS and maximum relative residual errors respectively, of the flow fields obtained from ANSYS CFX for the transient wind simulation sampled at 4 s increments. From Figure C.3 it can be seen that the residual error target of 1×10^{-5} of the RMS relative residual error on the mass and momentum equations was achieved within the maximum (20) coefficient loop iterations.

Figure C.3: Convergence history of RMS relative residual error of transient wind simulation sampled at 4 s increments

Figure C.4: Convergence history of max relative residual error of transient wind simulation sampled at 4 s increments

Appendix D

Receptor Locations

Table D.1 lists the locations of the 25 synthetic receptors (shown in Figure 4.6) within the two-dimensional airshed domain.

Table D.1: Synthetic receptors locations

Receptor #	x [m]	y [m]
1	615.7480	750.9140
2	714.3120	738.1900
3	772.8460	741.7840
4	842.0680	729.4080
5	680.9590	778.1650
6	741.7620	812.9090
7	782.4070	878.7470
8	631.4420	799.4190
9	631.8930	870.4950
10	631.9510	939.4020
11	575.3098	786.4570
12	535.1962	826.7420
13	485.1851	878.1710
14	563.2221	730.8690
15	492.4823	731.9390
16	421.8939	729.3590
17	571.1382	666.7270
18	532.9231	631.0900
19	486.3155	579.6044
20	631.3940	658.5680
21	633.5600	590.7171
22	587.4172	536.1152
23	681.6260	680.5890
24	731.1290	630.9400
25	783.2100	581.0241

Appendix E

Multi-Source Locations

Table E.1 lists the locations and emission rates of the 10 synthetic emission sources (shown in Figure 4.8) within the two-dimensional airshed domain from the multi-source test case.

Table E.1: Multi-source location and emission rate

Source	x [m]	y [m]	Q^* [g/s]
A	676.4400	608.6990	5
B	702.7490	775.5500	5
C	400.2103	597.6201	5
D	502.2215	800.1180	5
E	799.3760	799.1280	5
F	601.9760	749.0030	5
G	699.9240	897.9550	5
H	529.0874	678.9380	5
I	560.0289	640.3340	5
J	757.3490	659.2520	5

Appendix F

Objective Function Gradient Convergence

In addition to the iterative reduction in the objective function, the quality of a candidate emission source distribution can also be quantified with the convergence of the objective function gradient ($\nabla f(\mathbf{x}_k)$) to the zero-vector (see Section 4.4.2). Since L-BFGS-B is a gradient-based numerical optimization algorithm, it converges toward a minimum of the objective function, where the gradient is the zero-vector. A vector norm of the gradient vector provides a measure of the deviation of the objective function gradient from the zero-vector. The maximum norm (see (4.3)) of the gradient ($\|\nabla f(\mathbf{x}_k)\|_\infty$) is a convergence measure that is insensitive to the dimension of the optimization problem by measuring the maximum deviation of the gradient from the zero-vector. In fact, the maximum norm of the gradient is the basis of the gradient termination criteria used by the L-BFGS-B numerical optimization algorithm (see Section 4.2.3.3).

F.1 Single-Source Test Case

The convergence history of the maximum norm of the objective function gradient ($\|\nabla f(\mathbf{x}_k)\|_\infty$) for the single-source test case with and without observational noise and regularization is shown in Figure F.1.

Figure F.1: Convergence history of objective function gradient for the single-source test case

Objective function regularization and observational noise change the character of the objective function such that gradient measurements are only comparable across optimization iterations for a given IDM configuration. Figure F.1a suggests that the gradient of the objective function without regularization ($\beta = 0 \text{ s}^2\cdot\text{m}^{-3}$) is still tending toward a minimum ($\|\nabla f(\mathbf{x}_k)\|_{\infty} = 0 \text{ kg}\cdot\text{s}\cdot\text{m}^{-3}$) after 200 iterations. Figure F.1a suggests that the addition of objective function regularization ($\beta > 0 \text{ s}^2\cdot\text{m}^{-3}$) prevents the IDM from converging to a minimum where the gradient is equal to the zero-vector. However the measurement of the gradient's deviation from the zero-vector may be obscured by parameters with large sensitivities, that are constrained by the boundary of their acceptable search space. The gradient contributions of these boundary constrained parameters can be ignored with the projected gradient of the objective function (see Section 4.2.3.3).

The maximum norm of the projected gradient of the objective function ($\|\nabla_{\text{proj}} f(\mathbf{x}_k)\|_{\infty}$) gives an indication of the convergence of the objective function to a local minimum within the acceptable search space of the emission source parameters $\mathbf{x}$. Figure F.1b shows the convergence history of the maximum norm of the projected gradient for the single-source test cases with and without observational noise and regularization. Figure F.1b shows that, across all the configurations of the IDM for the single-source test case, the gradient of the objective function is converging toward the zero-vector. Figure F.1b also shows that objective function regularization also introduced oscillations in the magnitude of the projected gradient, which were amplified by the regularization parameter (β). These oscillations are believed to be caused by the emission source parameters “bouncing” off the lower bounds of their acceptable search space, which may be caused by an inadequate approximation of the inverse Hessian matrix of the objective function by the L-BFGS-B. It should also be noted that the choice of matching the background emission source estimate to the initial source it guess $s^b(\mathbf{r}, t) = s(\mathbf{r}, t; \mathbf{x}_0)$ results in an initial objective function gradient that is insensitive to the regularization term.

F.2 Multi-Source Test Case

The convergence histories of the objective function gradient for all conditions of the multi-source test case are shown in Figure F.2a.

Figure F.2: Convergence history of objective function gradient for the multi-source test case

Figure F.2a shows that the candidate emission source distributions predicted with $0 \leq \beta \leq 10 \text{ s}^2\cdot\text{m}^{-3}$ are still converging to the zero-vector after 200 iterations, regardless of observational noise, while solutions using a regularization parameter $\beta \geq 100 \text{ s}^2\cdot\text{m}^{-3}$ converge to a non-zero gradient. As with the single-source test case, it is believed that this convergence to a non-zero gradient resulted from objective function regularization moving the local minimum of the objective function outside of the acceptable search space of the optimization parameters. The gradient component from emission source parameters which are fixed by their search space boundaries can be ignored with the projected gradient of the objective function ($\nabla_{\text{proj}} f(\mathbf{x}_k)$) (see Section 4.2.3.3) as is plotted in Figure F.2b. The continued reduction in the maximum norm of the projected gradient objective function shown in Figure F.2b confirms that regularization has moved the local minimum of the objective function outside of the acceptable parameter search

space. Additionally, the large oscillations in the norm of the projected gradient for $\beta \geq 100 \text{ s}^2 \cdot \text{m}^{-3}$ suggest that projected gradient is sensitive to the regularization term. These oscillations are again believed to be caused by the source parameters “bouncing” off the lower bounds of their search space, which may be caused by an inadequate approximation of the inverse Hessian matrix of the objective function by the L-BFGS-B.

Appendix G

Multi-Source Convergence

The IDM iteratively predicted candidate emission source distributions that minimized the objective function, and hence minimized the error between the modelled and known receptor observations. The predicted emission source distributions are presented for the multi-source test case on a linear scale (see Section G.1) to highlight the location of the known emission sources, and a logarithmic scale (see Section G.2) to highlight the prediction of weak false emission sources.

G.1 Linear Scale

Figure G.1 shows a magnified sample of the candidate emission source distributions on a linear scale, predicted at the 25th, 50th, 100th and 200th iterations of the IDM for the multi-source test case, with and without observational noise and without objective function regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$).

Figure G.1: Iterative progression of candidate emission source distributions (linear scale) for the multi-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)

Figure G.1: Iterative progression of candidate emission source distributions (linear scale) for the multi-source test case without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$) (cont.)

G.2 Logarithmic Scale

Figure G.2 shows the iterative prediction of all the emission sources within the entire domain on a logarithmic scale for the multiple source test cases without regularization. While the IDM was able to correctly identify the location of the most of synthetic emission sources, several weaker false emission sources were also identified in regions of the domain that have poor observability for the receptor network and the wind direction used in this study.

Figure G.2: Iterative progression of candidate emission source distributions (log scale) for the multi-source test without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$)

Figure G.2: Iterative progression of candidate emission source distributions (log scale) for the multi-source test without regularization ($\beta = 0 \text{ s}^2 \cdot \text{m}^{-3}$) (cont.)