A GIS model for personal exposure to PM10 for Dublin commuters

A thesis submitted to the University of Dublin in fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Engineering.

Francesco Pilla
April 2012

Supervisors
Prof. Brian Broderick
Department of Civil Structural and Environmental Engineering, Trinity College Dublin
Prof. Rex Britter
Department of Urban Studies and Planning, Massachusetts Institute of Technology, USA
DECLARATION

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University’s open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

___________________________
Francesco Pilla
April 2012
Abstract

The project has focused on a number of scientific issues in the development of the GIS air quality model for the Dublin Area and its Satellite Towns. Firstly, how to model air quality by integrating existing and self-implemented air quality models and a Geographical Information System platform. Secondly, how to generate input data for the models to allow the estimation of air quality at any location in the Dublin Area. These issues have been addressed by implementing various tools: a method to calculate the background concentration levels, a method to model the dispersion of pollutants from road traffic in urban street canyons, a method to model the dispersion of pollutants from road traffic in general conditions, and a method to calculate the dispersion of pollutants from point and area sources. Thirdly, an approach to include all the models above into GIS tools has been discussed. Fourthly, an approach for disseminating results from air quality studies has been described and proposed.

As part of this work, the performances of various air quality models are assessed and the most suitable tools for modelling the dispersion of particulate matter in Dublin are selected. This is done in the context of the main objective: the assessment of the personal exposure of subjects moving between different locations in the urban area. This objective implies that a highly accurate solution at one location is not valuable if it is combined with less accurate predictions for other locations. It has to be noted that the only model commercially available that is utilised in this study is OSPM; all the other tools are implemented or modified by the author.

A new model for predicting the PM10 background concentration levels in the Dublin Area and its Satellite Towns is created by the author using machine learning algorithms; to the author's knowledge, it has not previously been reported in the context of air pollution modelling in Ireland. The background model uses artificial neural networks to model the non-linear relation between historical PM10 data recorded at Irish EPA permanent monitoring stations and the set of weather variables recorded by Met Eireann permanent monitoring stations. The background PM10 model well performs in calculating predicted PM10 concentration levels for hourly and daily averages, as shown in Chapter 3. As mentioned above, the model uses machine learning algorithms and thus its performances improve day by day, by increasing the size of the input datasets with the next-day measured concentration levels.

Another innovative aspect of the research presented here is the inclusion of all the modelling tools described above into a GIS platform: this integration aims to provide Local Authorities with a tool to calculate concentration levels of PM and to correlate
these with other thematic layers, such as land use and population density, in order to link localised peaks in air pollutants with particular activities. The GIS model has been validated by modelling the personal exposure to particulate matter of commuters travelling to and from work in Dublin city centre using different routes and different transport modes and comparing the modelled data with measured datasets obtained with mobile sensors and GPS units. To the author's knowledge, a validation test against such heterogeneous dataset hasn't been previously reported.

The work described in this thesis also provides a conceptual framework to assess and model personal exposure to other environmental pollutants, such as other air pollutants and noise, on a GIS platform. It could be also employed as a research tool to assess the link between individual health effects and personal exposure to particulate matter. To date, epidemiological studies do not establish causation and the absence of specific dose-effect relationships for air pollutants makes it difficult to establish limit values for human exposure.
Acknowledgements

I would like to thank the Irish Environmental Protection Agency “STRIVE” programme, which funded this work.

I wish to express my gratitude to my supervisor Dr. Brian Broderick for the amazing support and help throughout this project; I’m also very grateful to Dr. Rex Britter for his guidance during the months in MIT and to Dr. McNabola and Dr. Gill.

Thank you to my girlfriend Colette for her patience and understanding during my darkest hours and to my friend Eoin for working his magic with Matlab and saving me from insanity! Thanks to my friend and associate John for carrying the burden of our company for the last few months.

“Faraway, so close!” as a famous local singer would say: thank you of course to my mum, dad and my brother Federico who always supported me during this adventure.

I would also like to thank all the guys in my office: Andrew, John, Ted, Sean, Davie, Aisling and Mary for their support; a special thanks to Avril and Tara for taking me out for coffees and making me move away from my computer from time to time.

I owe my deepest gratitude to Prof. Carlo Ratti for hosting me in the SENSEable City Laboratory in MIT for 6 months and for letting me collaborate at several very interesting projects.

I’m also extremely indebted to Dr. Jensen, Dr. Ketzel and Dr. Hvidberg for their help with the Operational Street Pollution Model and for supporting me during the time spent in the National Environmental Research Institute in Denmark.

It is a pleasure to thank those who made this thesis possible by providing the data I needed for this work, such as the Irish Environmental Protection Agency, Dublin City Council and the Central Statistics Office Ireland.
When I examine myself and my methods of thought, I come to the conclusion that the gift of fantasy has meant more to me than any talent for abstract, positive thinking.
Table of Contents

A GIS MODEL FOR PERSONAL EXPOSURE TO PM10 FOR DUBLIN COMMUTERS

DECLARATION

Abstract

Acknowledgements

List of Figures

List of Tables

CHAPTER 1 – INTRODUCTION AND OBJECTIVES

1. Introduction

1.1 Particulate Matter

1.1.1 Particulate Matter in Ireland

1.2 Air quality standards

1.3 Health effects of particulates

1.4 Objectives of PhD Research

1.5 Novel aspects of PhD Research

1.6 Overview

CHAPTER 2 – AIR QUALITY MODEL FOR DUBLIN

2. Introduction

2.1 Note about air quality modelling

2.1.1 Gaussian air pollutant dispersion equation

2.1.1.1 The Briggs plume rise equations

2.1.2 Background pollution models

2.2 Existing atmospheric dispersion models

2.3 Summary of models commonly used in Ireland

2.3.1 Screening Models

2.3.1.1 AERSCREEN

2.3.1.2 ADMS-SCREEN

2.3.2 Advanced models

2.3.2.1 AERMOD

2.3.2.2 ADMS 4

2.3.3 Urban street canyon models

2.4 Existing Personal Exposure models
CHAPTER 7 – COMMUTERS’ EXPOSURE MODELLING ............................................. 228

7. Introduction ................................................................................................................ 230
  7.1.2 The outdoor personal PM10 exposure to model ................................................. 232
7.2 Commuting by bus ..................................................................................................... 233
  7.2.1 Measured data ..................................................................................................... 234
  7.2.2 Modelled data .................................................................................................... 235
  7.2.3 Analysis and validation ..................................................................................... 237
7.3 Commuting by bicycle ............................................................................................... 239
  7.3.1 Measured data ..................................................................................................... 240
  7.3.2 Modelled data .................................................................................................... 241
  7.3.3 Analysis and validation ..................................................................................... 244
7.4 Dart commuting ......................................................................................................... 246
  7.4.1 Measured data ..................................................................................................... 247
  7.4.2 Modelled data .................................................................................................... 247
  7.4.3 Analysis and validation ..................................................................................... 248
7.5 Walking ..................................................................................................................... 250
  7.5.1 Measured data ..................................................................................................... 251
  7.5.2 Modelled data .................................................................................................... 252
  7.5.3 Analysis and validation ..................................................................................... 253
7.6 Observations and sensitivity analysis ....................................................................... 255
7.7 Comparison between measured and modelled dose values .................................... 256

CHAPTER 8 – GIS MODEL AND DISSEMINATION OF RESULTS ......................... 258

8. Introduction ................................................................................................................ 260
  8.1 The GIS air quality model ....................................................................................... 260
    8.1.1 GIS air quality model ....................................................................................... 262
    8.1.2 GIS personal exposure model .......................................................................... 264
  8.2 GIS Cloud Computing and Web GIS .................................................................... 265
  8.4 Comparison with existing software packages - AirGIS ........................................ 268

CHAPTER 9 – CONCLUSIONS ..................................................................................... 272

9. Summary ..................................................................................................................... 274
  9.1 Main findings ......................................................................................................... 277
  9.2 Scope for further study ........................................................................................... 278

References ..................................................................................................................... 280
List of Figures

Figure 1: Air quality zones (O’Dwyer, 2011) ................................................................. 30
Figure 2: Air monitoring stations (O’Dwyer, 2011) ......................................................... 30
Figure 3: Annual mean PM10 concentrations at individual stations in 2010 (O’Dwyer, 2011) ................................................................. 31
Figure 4: Annual mean PM10 concentrations 2002-2010 (O’Dwyer, 2011) ..................... 31
Figure 5: 98-percentiles of daily PM10 concentrations in major urban areas, 1985-2010b (O’Dwyer, 2011) ................................................................. 31
Figure 6: Atmosphere layers ......................................................................................... 43
Figure 7: Visualization of a buoyant Gaussian air pollutant dispersion plume ............... 45
Figure 8: Logic diagram for using the Briggs’s equation (Beychok, 2005) ...................... 47
Figure 9: “Sequence of exposure” from source to biologically effective dose (Monn, 2001) .................................................................................. 56
Figure 10: Air quality monitoring stations, Dublin area .................................................. 59
Figure 11: Spatial variation of air pollutants ................................................................... 60
Figure 12: Air quality model .......................................................................................... 60
Figure 13: Urban background pollution model, logic scheme ......................................... 62
Figure 14: Visualisation in ArcGIS of PM10 urban background ...................................... 62
Figure 15: Important point and area source model, logic scheme .................................. 64
Figure 16: Visualisation in ArcGIS of PM10 pollution from important area and point sources .................................................................................. 64
Figure 17: Traffic contribution in street canyons model, logic scheme ......................... 65
Figure 18: Visualisation in ArcGIS of PM10 traffic related levels ................................. 65
Figure 19: Visualisation in ArcGIS of the different layers for PM10 .............................. 67
Figure 20: Visualisation in ArcGIS of the total PM10 concentration levels .................... 67
Figure 21: Personal exposure model flowchart ............................................................... 69
Figure 22: PM10 plot ....................................................................................................... 81
Figure 23: Plot PM10 against average rain ..................................................................... 81
Figure 24: Air quality monitoring stations in the Great Dublin Area (Irish EPA) .......... 83
Figure 25: PM10 monitoring station, Dun Laoghaire .................................................... 84
Figure 26: PM10 time series for Dun Laoghaire ............................................................. 85
Figure 27: PM10 monitoring station, Phoenix Park ......................................................... 85
Figure 28: PM10 time series for Phoenix Park ................................................................. 86
Figure 29: PM10 monitoring station, Rathmines ............................................................. 86
Figure 30: PM10 time series for Rathmines .................................................................... 87
Figure 31: PM10 monitoring station, Tallaght ................................................................. 87
Figure 32: PM10 time series for Tallaght .......................................................... 88
Figure 33: Casement weather station ............................................................... 88
Figure 34: Dublin Airport weather station ...................................................... 89
Figure 35: Phoenix Park weather station ......................................................... 90
Figure 36: Dun Laoghaire PM10 plot before and after removing the outliers .... 92
Figure 37: Tallaght PM10 measurements after simple moving average smoothing... 94
Figure 38: Tallaght PM10 measurements after exponential moving average smoothing.......................................................... 94
Figure 39: Dun Laoghaire PM10 trend-line ...................................................... 97
Figure 40: Phoenix Park PM10 trend-line ....................................................... 97
Figure 41: Rathmines PM10 trend-line ........................................................... 97
Figure 42: Tallaght PM10 trend-line ............................................................... 97
Figure 43: Feed-forward NARX network ....................................................... 113
Figure 44: Closed-loop NARX network .......................................................... 113
Figure 45: NARX network for predictions ...................................................... 113
Figure 46: Dun Laoghaire NARX network scheme ....................................... 114
Figure 47: Phoenix Park NARX network scheme ........................................... 114
Figure 48: Rathmines NARX network scheme .............................................. 114
Figure 49: Tallaght NARX network scheme ................................................. 114
Figure 50: Dun Laoghaire, correlation coefficient between measured and modelled PM10 [\(\mu g/m^3\)] ........................................................................................................... 115
Figure 51: Phoenix Park, correlation coefficient between measured and modelled PM10 [\(\mu g/m^3\)] ........................................................................................................... 115
Figure 52: Phoenix Park, correlation coefficient between measured and modelled PM10 [\(\mu g/m^3\)] ........................................................................................................... 115
Figure 53: Tallaght, correlation coefficient between measured and modelled PM10 [\(\mu g/m^3\)] ........................................................................................................... 115
Figure 54: Error histogram for Dun Laoghaire [\(\mu g/m^3\)] ............................ 116
Figure 55: Error histogram for Phoenix Park [\(\mu g/m^3\)] ............................. 116
Figure 56: Error histogram for Rathmines [\(\mu g/m^3\)] .................................. 117
Figure 57: Error histogram for Tallaght [\(\mu g/m^3\)] .................................... 117
Figure 58: Distribution of errors in the time-series for Dun Laoghaire .......... 118
Figure 59: Distribution of errors in the time-series for Phoenix Park ............. 118
Figure 60: Distribution of errors in the time-series for Rathmines .................. 119
Figure 61: Distribution of errors in the time-series for Tallaght ...................... 119
Figure 62: PM10 before the interpolation, Dun Laoghaire ............................ 120
Figure 63: Pre-processed and modelled PM10, Dun Laoghaire .................... 120
Figure 95: 2D model of Dublin city in ArcGIS.................................154
Figure 96: 3D model of Dublin city in ArcGIS.................................154
Figure 97: ArcGIS streets and buildings data cleared from features not needed by
OSPM..............................................................................154
Figure 98: Dublin city centre street layout in OSPM...........................156
Figure 99: street geometry in OSPM ............................................156
Figure 100: The EmiFact program screen with country setting “Ireland” and scenario
year “2009”........................................................................158
Figure 101: The TrafEdit program screen with day case “Monday-Thursday” and
scenario year “2009”.............................................................159
Figure 102: Pearse Street canyon ...................................................163
Figure 103: Location details of background and monitoring site...............163
Figure 104: Local effects of CO and NOx ........................................164
Figure 105: Speed versus emission characteristics for CO......................165
Figure 106: Speed versus emission characteristics for NOx......................165
Figure 107: Diurnal profile of CO emission factor ................................166
Figure 108: Diurnal profile variation for CO using CEF and HEF...............166
Figure 109: Diurnal profile of NOx emission factor..............................167
Figure 110: Diurnal profile variation for NOx using CEF and HEF..............168
Figure 111: PM10 urban and sub-urban monitoring station in the Dublin area....169
Figure 112: PM10 background monitor stations positions..........................170
Figure 113: Modelled PM10 background concentration levels for Phoenix Park....170
Figure 114: Temperature hourly variation in March 2010 recorded by Phoenix Park
weather station...................................................................171
Figure 115: Relative humidity hourly variation in March 2010 recorded by Phoenix Park
weather station...................................................................171
Figure 116: Solar radiation hourly variation in March 2010 recorded by Phoenix Park
weather station...................................................................172
Figure 117: Wind speed hourly variation in March 2010 recorded by Dublin Airport
weather station...................................................................172
Figure 118: Wind direction hourly variation in March 2010 recorded by Dublin Airport
weather station...................................................................172
Figure 119: OSPM “Type_C” traffic file...............................................174
Figure 120: Comparison between measured and modelled PM10 daily average levels
at Ballyfermot Road for the month of March 2010..............................174
Figure 121: Area around the location of Ballyfermot monitoring station........175
Figure 122: Correlation between measured and modelled PM10 daily average levels at Ballyfermot Road for the month of March 2010 ................................................................. 176
Figure 123: Comparison between measured and modelled PM10 daily average levels at Winetavern Street for the month of March 2010 ................................................................. 176
Figure 124: PM10 daily averages in March 2010 recorded by Dun Laoghaire monitoring station ................................................................................................................................. 177
Figure 125: Area around the location of Winetavern monitoring station .......... 177
Figure 126: Correlation between measured and modelled PM10 daily average levels at Winetavern Street for the month of March 2010 ................................................................. 178
Figure 127: Hourly atmospheric stability class in March 2010 recorded by Dublin Airport weather station ................................................................................................................................. 178
Figure 128: Hourly rainfall in March 2010 recorded by Phoenix Park weather station ................................................................................................................................. 179
Figure 129: Daily average PM10 concentration levels calculated with OSPM for the 30th of September 2010 for Dublin City centre ................................................................. 181
Figure 130: suburban area in Dublin ................................................................................................................................................................................................. 184
Figure 131: Schematic representation of the GFLSM model (Ganguly, 2008) .... 191
Figure 132: Receptors’ grid ................................................................................................................................................................................................. 197
Figure 133: Output from the GFLSM ArcGIS model ........................................... 197
Figure 134: PM10 Contours ................................................................................................................................................................................................. 198
Figure 135: Gaussian plume model (Turner, 1994) ............................................ 213
Figure 136: Dispersion of pollutants from the industrial sites in Dublin, “slightly unstable” atmospheric stability condition ................................................................. 217
Figure 137: Dispersion of pollutants from the industrial sites in Dublin, "slightly stable" atmospheric stability condition ................................................................................................................................. 217
Figure 138: Houses with central heating in the Dublin area and its Satellite Towns. 222
Figure 139: Data disaggregation procedure ......................................................... 222
Figure 140: Houses with central heating per sq km ............................................ 223
Figure 141: Houses emission (mg/day) per sqkm ................................................ 223
Figure 142: Temperature variation measured by Phoenix Park weather monitoring station in 2010 ................................................................................................................................. 225
Figure 143: Average natural gas daily consumption (per house) variation in 2010 ................................................................................................................................. 225
Figure 144: Average PM10 daily emission (per house) variation in 2010 ................................................................................................................................. 225
Figure 145: Discrete distance versus concentration for a 1sqKm area source in the Dublin city area ................................................................................................................................. 226
Figure 146: Dispersion of pollutants from a 1sqKm residential area in Dublin, "slightly stable" atmospheric stability condition ................................................................................................................................. 226
Figure 147: The Metone Aerocet 531 ................................................................. 231
Figure 148: Examples of the routes included in the monitoring dataset .......... 233
Figure 149: View of the commuting route by bus .............................................. 234
Figure 150: Measured PM10 concentration levels .............................................. 235
Figure 151: PM10 modelled with the Rathmines hourly ANN model .............. 236
Figure 152: Route modelled with OSPM .......................................................... 236
Figure 153: OSPM multi-streets model ............................................................ 236
Figure 154: Modelled PM10 concentration levels .............................................. 237
Figure 155: Difference between measured and modelled PM10 concentration levels .................................................................................................................. 238
Figure 156: Time variation of the measured PM10 concentration levels .............. 238
Figure 157: 5 minute average measured and modelled PM10 concentration levels.239
Figure 158: Regression of measured against modelled values ......................... 239
Figure 159: View of the commuting route by bicycle ........................................ 240
Figure 160: Measured PM10 concentration levels .............................................. 241
Figure 161: PM10 modelled with the area sources model ................................... 242
Figure 162: PM10 modelled with the Phoenix Park hourly ANN model ............ 242
Figure 163: Route modelled with OSPM .......................................................... 243
Figure 164: OSPM multi-streets model ............................................................ 243
Figure 165: Modelled PM10 concentration levels .............................................. 243
Figure 166: Difference between measured and modelled PM10 concentration levels .................................................................................................................. 244
Figure 167: Time variation of the measured PM10 concentration levels .............. 244
Figure 168: 5 minute average measured and modelled PM10 concentration levels.245
Figure 169: Regression of measured against modelled values ......................... 245
Figure 170: View of the commuting route by train ............................................ 246
Figure 171: Measured PM10 concentration levels .............................................. 247
Figure 172: Modelled PM10 concentration levels .............................................. 248
Figure 173: Difference between measured and modelled PM10 concentration levels .................................................................................................................. 249
Figure 174: 5 minute average measured and modelled PM10 concentration levels.249
Figure 175: Regression of measured against modelled values ......................... 250
Figure 176: View of the walking journey office-home ....................................... 251
Figure 177: Measured PM10 concentration levels .............................................. 251
Figure 178: Time-series for measured PM10 ...................................................... 252
Figure 179: OSPM multi-streets model ............................................................ 252
Figure 180: Modelled PM10 concentration levels .............................................. 253
Figure 181: Difference between measured and modelled PM10 concentration levels ................................................................................................................................................................................................. 254
Figure 182: 5 minute average measured and modelled PM10 concentration levels. 254
Figure 183: Regression of measured against modelled values ......................................................... 255
Figure 184: ESRI ArcGIS toolbox, from ESRI website ............................................................................. 261
Figure 185: Example of custom ESRI ArcGIS model .............................................................................. 261
Figure 186: PALM-GIS personal exposure model flowchart ................................................................. 262
Figure 187: ArcGIS Cloud diagram ........................................................................................................ 266
Figure 188: ArcGIS.com is a Web site providing access to GIS data and applications ........................................................................................................................................................................................................... 266
Figure 189: The Schleswig-Holstein online bike routing portal allows tourists to plan and map their itineraries (http://www.sh-tourismus.de/en/planning) ........................................................................ 268
Figure 190: The AirGIS system makes use of digital maps, national and local administrative databases, and a Geographic Information System (Jensen et al., 2001) ................................................................................................................................................................................................................................................. 269
List of Tables

Table 1: Limit values of CAFE Directive 2008/50/EC .................................................. 33
Table 2: PM10 monitoring station in the Dublin area ..................................................... 83
Table 3: Datasets correlations summary ....................................................................... 91
Table 4: R scores before and after the elimination of outliers ....................................... 93
Table 5: R scores before and after the moving average smoothing ............................... 95
Table 6: Best R score per PM10 monitoring station ....................................................... 98
Table 7: R scores for the initial and final models for Dun Laoghaire ............................. 99
Table 8: ANOVA table for Dun Laoghaire ..................................................................... 99
Table 9: Coefficients values for Dun Laoghaire ............................................................ 100
Table 10: Statistics for excluded predictor variables for Dun Laoghaire ......................... 101
Table 11: R scores for the initial and final models for Phoenix Park ............................ 102
Table 12: ANOVA table for Phoenix Park .................................................................... 102
Table 13: Coefficients values for Phoenix Park ............................................................. 103
Table 14: Statistics for excluded predictor variables for Phoenix Park ......................... 104
Table 15: R scores for the initial and final models for Rathmines ................................ 104
Table 16: ANOVA table for Rathmines ........................................................................ 105
Table 17: Coefficients values for Rathmines ................................................................. 105
Table 18: Statistics for excluded predictor variables for Rathmines ............................. 106
Table 19: R scores for the initial and final models for Tallaght .................................... 107
Table 20: ANOVA table for Tallaght ............................................................................. 107
Table 21: Coefficients values for Tallaght ..................................................................... 108
Table 22: Statistics for excluded predictor variables for Tallaght ................................ 108
Table 23: Summary of the networks’ structures ............................................................. 115
Table 24: R scores for the initial and final models for Winetavern Street ................. 125
Table 25: ANOVA table for Winetavern Street ............................................................. 125
Table 26: Coefficients values for Winetavern Street ...................................................... 125
Table 27: Statistics for excluded predictor variables for Winetavern Street ............... 126
Table 28: R scores for the initial and final models for Winetavern Street ................. 131
Table 29: ANOVA table for Winetavern Street ............................................................. 131
Table 30: Coefficients values for Winetavern Street ...................................................... 131
Table 31: Statistics for excluded predictor variables for Winetavern Street ............... 132
Table 32: R scores for the initial and final models for Rathmines .............................. 132
Table 33: ANOVA table for Rathmines ....................................................................... 133
Table 34: Coefficients values for Rathmines ................................................................. 133
Table 35: Statistics for excluded predictor variables for Rathmines ........................ 134
Table 36: Example of road links – OSPM input .......................................................... 155
Table 37: Statistical evaluation of CO using CEF and HEF for OSPM ....................... 166
Table 38: Statistical evaluation of NOx using CEF and HEF for OSPM ..................... 167
Table 39: PM10 background monitoring stations in the Dublin area divided by typology .................................................................................................................................................. 169
Table 40: The Pasquill stability classes (F. Pasquill, 1961) ........................................... 179
Table 41: Stability class analysis of CO data ................................................................. 192
Table 42: Wind Sensitivity analysis of the CO data ....................................................... 192
Table 43: Formulas produced by Briggs (Briggs, 1973) for $\sigma_y(x)$ and $\sigma_z(x)$ ($10^2 < x < 10^4$ m) .......................................................................................................................... 196
Table 44: The Pasquill stability classes (F. Pasquill, 1961) ......................................... 196
Table 45: Dublin power station and industrial point source locations and modelling parameters .......................................................................................................................... 207
Table 46: Dublin power station and industrial sources emission rates (g/s) ............... 208
Table 47: Emissions for Dublin city (T/yr) for each emission year ................................ 211
Table 48: Percentage of total emissions for Dublin city ............................................. 211
Table 49: Formulas produced by Briggs (Briggs, 1973) for $\sigma_y(x)$ and $\sigma_z(x)$ ($10^2 < x < 10^4$ m) ................................................................. 215
Table 50: The Pasquill stability classes (F. Pasquill, 1961) ........................................... 216
Table 51: Measured and modelled dose for the analysed test cases ......................... 257
CHAPTER 1 – INTRODUCTION
AND OBJECTIVES
1. Introduction

Living or attending school near major roadways has been associated with numerous adverse health outcomes in recent years, including asthma (Matthew J, 2004), other respiratory illnesses (Dockery and Pope, 1994) and increased risk of mortality from cardiopulmonary (Pope et al., 2002) and cardiovascular diseases (Pope et al., 2004). While the potential for a large public health impact has been established and well documented, it remains current practice to derive inhabitants’ exposure from the measurements of very sparse fixed monitoring stations even for large size cities. This approach does not fully capture spatiotemporal patterns of air pollutants, and causes the information on an individual’s environmental exposure to be poorly aligned with more detailed information on their personal health.

One of the topics in which the European Commission (European Commission, 2008) has shown a greater concern is the necessity of developing actions that increase knowledge on transport dynamics and related atmospheric pollution to assure the accomplishment of legislation and to inform the population about their exposure levels. It is indeed very important to increase the general public awareness about air pollution: high levels of air pollutants are “invisible killers”, they can’t be seen and therefore, the unaware inhabitants of polluted cities may underestimate them with detrimental effects to their health.

The use of mathematical models to assess ambient air quality in cities has increased during the last decade because of the inherent difficulties and costs involved in monitoring air pollution on several streets within an urban area and also because of the increased computational power of the computers available on the market. These models have been diffusely adopted by Local Authorities with the purpose of assessing and quantifying population exposure to air pollutants and the compliance with regulations (Vardoulakis et al., 2005): they are de facto the only tool available to assess the possible impact of planned developments (Manning et al., 2000) before they are given the green light. However, modelling the dispersion of air pollutants in streets is not a trivial task because urban areas are not homogeneous entities and the highest levels of pollution occur in street canyons where dilution of car exhaust gases is significantly limited by the presence of buildings flanking the street (Berkowicz et al., 1997).

Epidemiological studies (Vardoulakis et al., 2002a) of urban populations have raised concerns over the adverse effects on human health of airborne traffic related particulate matter (PM), and compelled various National and International Agencies
instigated (Harrison et al., 2001) to propose stringent air quality standards (e.g. European Union Council Directive 1999/30/EC). The setting of threshold limits for human exposure is largely based on analysis of data from epidemiological studies that linked exposure to airborne concentrations of particulate matter (PM) to a variety of health effects, particularly increased morbidity and mortality from respiratory and cardiac causes (Harrison and Yin, 2000).

The aim of the research presented in this work is based on the necessity to understand the dynamics of air pollution with the aim of obtaining a spatiotemporally precise estimation of air quality in Dublin and of the city inhabitants’ air pollution exposure through high-resolution modelling (European Commission, 1996). Specifically, this work aims to create an efficient computational methodology for assessing particulate matter concentration levels in the Greater Dublin Area to facilitate the calculation of inhabitants’ exposure to this pollutant while at their habitations and work places and while commuting between them.

### 1.1 Particulate Matter

Airborne particulate matter consists of a complex combination of organic and inorganic substances including uncombusted fuel, dust and metal fragments and represents a hazardous by-product of the combustion process (Schwela and Zali, 1999). As regards urban areas in general, PM can be characterised by dividing it into the following different categories by source:

- **Particulates from Vehicle Exhaust Emissions**: The particulate matter belonging to this category is generated by vehicles and it can be sub-divided into two categories: primary and secondary emissions. Primary particulate emissions are a direct result of the inefficient vehicular combustion process; a small fraction of particulate deposits on the surfaces of the exhaust system and it’s then incorporated in the exhaust gases and them emitted in the external environment (Habibi, 1973). Primary PM emission rates generally decrease with increasing vehicle speed, when fuel combustion becomes more efficient, and are greater for diesel-fuelled vehicles. Secondary particulates are formed in the immediate atmosphere from the condensation of exhaust vapours or by atmospheric chemical reactions (U.S.E.P.A., 1996).

- **Particulates from Vehicular Non-Exhaust Sources**: this category includes a wide range of sources, which are mostly associated with turbulence/resuspension generated by vehicles and deterioration of vehicle components. The fraction of particulate matter originated from these activities
is the PM10 (<10.0\(\mu\)m in size), or “coarse” fraction (U.S.E.P.A., 1996). In Ireland, the resuspension of the coarse fraction is strongly dependent on weather conditions, such as precipitations, and therefore it has a seasonal component, with peaks during dry months (Yin et al., 2005). Yin et al observed that these peaks in re-suspended particulate matter were not evident in non-urban sites around Dublin where the traffic was almost absent, even during dry months. This suggested that vehicle induced turbulence was an important contributory factor in the resuspension of road surface material. Other sources of PM not directly related to exhaust emissions are the wear of tyres during braking and skidding, the deterioration of brake linings during braking and road abrasion (Ketzel et al., 2007). The generation of coarse particulate from these sources increases with the vehicle speed.

- **Particulates from Anthropogenic Sources**: the finer fraction of particulate is generally originated from domestic heating and industrial processes, while the coarse fraction is generated by construction activities and mining (Harrison et al., 2001). In Dublin, after bituminous coal products were banned for heating purposes in 1990, domestic heating was no any longer a major source of coarse particulate (Kabir et al., 2007).

- **Particulate from Natural Sources**: Yin et al discovered that primary marine aerosol (sea salt) was the largest contributor to coarse particulates (56%) originating from natural sources in coastal sites in Ireland, including Dublin (Yin et al., 2005). The author also indicated wind speed as the main weather variable responsible for higher concentrations of coarse particulate from sea salt: higher wind speeds resulted in higher concentrations.

### 1.1.1 Particulate Matter in Ireland

The World Health Organisation has recently concluded that air pollution has an adverse effect on people's health and development (World Health Organisation, 2005). For instance, chronic exposure to higher levels of air pollution of various types including PM have been associated with decrements in lung function in children (Gauderman et al., 2004). A study from Switzerland found increased levels of allergic sensitisation in adults living for periods longer than 10 years in proximity to busy roads (Wyler et al., 2000). Another more recent study applied a multi-scale air quality modelling system to assess the impact of combustion emissions on UK air quality: epidemiological evidence quantitatively related PM2.5 exposure to risk of early death (Yim and Barrett, 2012).
According to “Air Quality in Ireland 2010” (O’Dwyer, 2011) published by the Irish Environmental Agency, levels of PM10 continue to remain of concern in Ireland. This pollutant is amongst those regulated under the Air Quality Framework Directive on ambient air quality assessment and management (European Commission, 1996). In order to establish whether or not the current levels impact upon the health of sensitive individuals, this work will develop a set of complementary techniques to assess personal exposure to PM10 for commenters in the Dublin Area.

For air quality assessment purposes, Ireland has been divided in four zones (Figure 1) and mean concentrations of PM10 have been monitored (in 2010) at 18 stations across the country (Figure 2).

Traffic and domestic heating are the main sources of particulate matter in air in Ireland and particularly impact air quality in areas where the use of bituminous coal is permitted. As a result, levels of particulate matter in smaller towns are similar or higher than those in cities, where solid fuel is banned. As such, it is important to stress that the use of bituminous coal for domestic heating fuel can have a detrimental impact on the environment and air quality. In order to reduce this impact, from 2011 the government is enforcing a regulation that will require that all bituminous coal placed on the market in Ireland for residential use has a sulphur content of no more
than 0.7% (O'Dwyer, 2011). Mean concentrations of PM10 monitored (in 2010) at 18 stations across Ireland are below the annual limit value of 40 μg/m3 (Figure 3).

![Figure 3: Annual mean PM10 concentrations at individual stations in 2010 (O'Dwyer, 2011)](image)

The trends in PM10 annual mean concentrations in the four different zones are shown in Figure 4. In the “Air Quality in Ireland 2010” report (O'Dwyer, 2011), it is noted that traffic is the main source of PM10 in cities, while emissions from domestic solid fuel combustion dominate in smaller towns.

![Figure 4: Annual mean PM10 concentrations 2002-2010 (O'Dwyer, 2011)](image)

![Figure 5: 98-percentiles of daily PM10 concentrations in major urban areas, 1985-2010b (O'Dwyer, 2011)](image)
A decreasing trend is seen in urban zones (Zones A, B and C) since 2003, which is most likely due to the decreases in particulate emissions from traffic due to improved vehicle engine emissions. This negative trend is not noticeable in Zone D where domestic solid fuel emissions are more significant than traffic emissions.

As it can be seen from Figure 5, concentration levels of PM10 across the Dublin, Limerick and Cork monitoring networks have decreased substantially from 1985 to 2010. This large reduction is mainly due to the smoke-control legislation introduced in the greater Dublin area in 1990 which banned the marketing, sale and distribution of bituminous coal. This ban on bituminous coal was extended to Cork city in 1995 and to Arklow, Drogheda, Dundalk, Limerick and Wexford in 1998. In 2000 the ban was extended to Celbridge, Galway, Leixlip, Naas and Waterford, in 2003 it was extended further to include Bray, Kilkenny, Sligo and Tralee and in 2011 it was extended again to include Athlone, Carlow, Clonmel and Ennis (O’Dwyer, 2011).

Despite of the low 98-percentiles of daily PM10 concentrations in Irish major urban areas, episodes of elevated PM10 levels can occur if meteorological conditions which are favourable to poor dispersion of local emissions persist for a number of days or weeks. These conditions are generally associated with the presence of a stable anticyclone over Ireland, i.e. warm weather in the summer and cold weather in the winter. Such conditions are also favourable to transportation of air masses with high concentrations of PM10 from other European countries to Ireland.

As such, this work’s goal is to create a set of integrated GIS-based exposure assessment tools based on monitoring data, spatial analysis, stochastic modelling and deterministic modelling in order to account for the impact of different emission sources and weather conditions on the total PM10 concentrations. Stochastic models will be developed through spatial regression models of the dependency of the ambient concentration measurements obtained within the project and influencing variables such as housing and traffic density, described in the GIS environment and adjusting for local area deprivation. The synthesis of these techniques will lead to a valuable resource to be used in the assessments of health impacts, and in the development of wider exposure-reduction strategies.

1.2 Air quality standards

In order to protect the health, vegetation and ecosystems of the European Countries and their citizens, the European Commission enforced several directives which set down air quality standards in Ireland and the other member states for a wide variety of
pollutants. The principles to this approach were established in 1996 with the Air Quality Framework Directive (European Commission, 1996). The limits for specific pollutants were then laid down in four "daughter" directives:

- **1st Daughter Directive**: Sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead;
- **2nd Daughter Directive**: Carbon monoxide and benzene;
- **3rd Daughter Directive**: Ozone;
- **4th Daughter Directive**: Polyaromatic hydrocarbons, arsenic, nickel, cadmium and mercury in ambient air.

The Framework Directive and the first, second and third Daughter Directives were then replaced by the Ambient Air Quality and Cleaner Air for Europe (CAFE) Directive 2008/50/EC (European Commission, 2008). The limits and target values for particulate matter are relevant to this work and they are outlined in Table 1 below.

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>Limit Value Objective</th>
<th>Averaging Period</th>
<th>Limit Value ug/m3</th>
<th>Basis of Application of the Limit Value</th>
<th>Limit Value Attainment Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM10</td>
<td>Protection of human health</td>
<td>24 hours</td>
<td>50</td>
<td>Not to be exceeded more than 35 times in a calendar year</td>
<td>1 Jan 2005</td>
</tr>
<tr>
<td>PM10</td>
<td>Protection of human health</td>
<td>calendar year</td>
<td>40</td>
<td>Annual mean</td>
<td>1 Jan 2005</td>
</tr>
<tr>
<td>PM2.5 - Stage 1</td>
<td>Protection of human health</td>
<td>calendar year</td>
<td>25</td>
<td>Annual mean</td>
<td>1 Jan 2015</td>
</tr>
<tr>
<td>PM2.5 - Stage 2</td>
<td>Protection of human health</td>
<td>calendar year</td>
<td>20</td>
<td>Annual mean</td>
<td>1 Jan 2020</td>
</tr>
</tbody>
</table>

Table 1: Limit values of CAFE Directive 2008/50/EC


1.3 Health effects of particulates

The increasing number of motor vehicles is a major source of particulate air pollution, particularly in densely populated urban areas. The characteristics of traffic related
particulate matter are influenced by the type of fuel used by the vehicle, the age of the engine and its operating condition, and whether or not the engine has been fitted with an appropriate emission control device (Schwela and Zali, 1999); this wide spectrum of parameters of the source results into a highly varied nature of particulate matter, which makes adverse health effects difficult to quantify. The effects on health of particulate matter depend on a wide range of factors, which can be broadly grouped into the following headings:

- **Composition:** Fine particles may belong to primary or secondary particulates and they are composed substantially of sulphates and nitrates, generally in the form of ammonium salts (NH₄NO₃ and ((NH₄)₂SO₄), together with organic and elemental carbon and a range of trace metals (Harrison et al., 2001). These particles, along with the chemicals transported on their surfaces, can be highly acidic and potentially a source of lung inflammation upon inhalation (Seaton et al., 1995).

- **The particle size:** This is a key factor that determines exposure and human dosage. Oberdörster et al. observed that particle surface area and dosimetry parameters such as deposition, clearance, retention, and translocation and dissolution of inhaled particles in and to different lung compartments were important for the persistence of particles in the lung and correlated with adverse pulmonary effects (Oberdörster et al., 1994). Epidemiological studies carried out by Donaldson et al. (1998) have shown a correlation between environmental particulate air pollution (PM10) and a range of endpoints of lung morbidity and also mortality, and an ultrafine fraction of the PM10 particles is considered to be the principal mediator of these effects (Donaldson et al., 1998). Schwela and Zali observed that with normal nasal breathing, larger particles of aerodynamic diameter between 10-100μm tended to be deposited in the extra-thoracic part of the respiratory tract; particles in the range 5-10μm were deposited in the proximity of the fine airways; fine particles (<2.5μm) tended to be deposited in the distal lung airway and airspace (Schwela and Zali, 1999). Recent epidemiological studies completed by Oberdörster have shown that ultrafine particles are causally involved in adverse responses seen in sensitive humans, such as the elderly with pre-existing respiratory and cardiovascular diseases (Oberdörster, 2000).

- **Length of exposure and dosage:** Dockery et al. found a 14-20% increase in lung cancer mortality for every 10μg/m³ increase in particulate air pollution and positively associated exposure to fine particulates with death from lung cancer.
and cardiopulmonary disease (Dockery et al., 1993). Dockery and Pope found a connection between acute respiratory effects and exposure to particulate matter (Dockery and Pope, 1994). Seaton et al. linked long term exposure to particulate air pollution to extended hospitalisation (Seaton et al., 1995); while Pope et al. proved a positive connection between long term exposure to particulate matter and lung cancer (Pope et al., 1995) and also cardiopulmonary mortality (Pope et al., 2002, Pope et al., 2004). Chronic rat inhalation studies have shown that a number of different particle types can induce significant adverse effects, including impaired lung clearance, chronic pulmonary inflammation, pulmonary fibrosis, and lung tumours (Oberdörster, 1996). Exposure to particulate matter has been linked to acute reactions, such as irritative bronchitis (or industrial or environmental bronchitis), and chronic reactions, such as chronic bronchitis and pulmonary emphysema (Schwela and Zali, 1999). Jerrett et al. carried out a spatial analysis of PM pollution and mortality in Los Angeles found specificity in cause of death, with PM2.5 associated more strongly with ischemic heart disease than with cardiopulmonary or all-cause mortality (Jerrett et al., 2005). Studies completed in Ireland also showed a correlation between exposure to particulate matter and increased mortality: Clancy et al. reported a decrease of respiratory and cardiovascular deaths after the ban of marketing and distribution of bituminous coal products for domestic heating in Ireland in 1990 (Clancy et al., 2002); Kabir et al. associated a decrease in lung cancer mortality rates with a significant decline in mean annual ‘black smoke’ concentrations after the same ban (Kabir et al., 2007).

1.4 Objectives of PhD Research

Nowadays, the majority of large cities in the world have problems related to air quality, not only in developed countries but also in cities of developing areas. The mother directive 1996/62/EC and daughter directives established the possibility of using modelling techniques to assess air quality. As such, this study aims to create a software package to model the presence, dispersion and dilution of air pollutants, and specifically particulate matter, in Dublin. The objectives of this work are the followings:

- Assess the performances of various air quality models and select the most suitable for modelling the dispersion of particulate matter in Dublin;
- Generate a model to predict the background PM concentration levels for the Dublin area;
• Integrate all the models in a GIS platform in order to create a tool for calculating concentration levels of PM and correlating these with other thematic layers, such as land use and population density, in order to link localised peaks in air pollutants with particular activities;

• Build a GIS air quality model to assess the personal exposure to particulate matter of individuals at specific locations and while travelling in the Dublin area;

• Validate the personal exposure model with measurements gathered with mobile sensors at different locations and on different routes throughout the city;

• Provide a conceptual framework to assess and model personal exposure to other environmental pollutants, such as other air pollutants and noise, on a GIS platform;

• Provide a web interface for the model for dissemination purposes;

• Provide a tool to assess the link between health effects and personal exposure to particulate matter.

1.5 Novel aspects of PhD Research

The contributions of this research to the scientific field of air quality include the followings:

• A GIS air quality model to assess personal exposure to particulate matter of individuals at and travelling between different locations in the Dublin area. Existing modelling tools not created for this purpose do not possess the required balance of detail and efficiency

• A conceptual framework to assess and model personal exposure to other environmental pollutants, such as other air pollutants and noise, on a GIS platform: the model is flexible and able to incorporate other environmental pollutants and it will be ready to be use with other layers in order to find hidden inter-correlations between them;

• A model to predict the background PM concentration levels for the Dublin area created using machine learning algorithms; as such the neural network model will “learn” from experience and its performances will improve with the size of the training dataset, i.e. by including measurements day by day;

• A tool to assess the link between health effects and personal exposure: to date, epidemiological studies do not establish causation and the absence of
specific dose-effect relationships for air pollutants makes it difficult to establish limit values for human exposure;

- A dissemination GIS based tool to increase the Irish public awareness about their exposure to particulate matter.

1.6 Overview

An overview of existing air quality models and of the modelling procedure adopted in this work is presented in Chapter 2. In Chapter 3, different approaches to model background air pollution are presented along with a detailed account of the selected modelling procedure: the datasets employed in this task are treated and the results from the modelling procedure are presented and analysed. Chapter 4 gives an overview of the existing software packages widely used to model the dispersion of pollutants in “urban street canyons”. OSPM is chosen for this work because of its low computational times and also because of the confidence derived from the extensive use of the model within the Department of Civil Structural and Environmental Engineering in Trinity College Dublin, where this study was completed. Modelling results are validated against measurements obtained from permanent monitoring stations located in different areas in Dublin. Chapter 5 evaluates different dispersion models for calculating the traffic related air pollution in sub-urban areas, where the streets are not classified as urban street canyons, and proposes a GIS-based model for this task. A GIS model for point and area source is then described in Chapter 6. Chapter 7 presents the personal exposure GIS model which is tested and validated with different sets of data gathered by commuters during their trip to work with various transport modes. Chapter 8 proposes a framework to make the personal exposure model available to the general public for dissemination and consultation purposes. The concluding remarks and observations are presented in Chapter 9; furthermore, the potentials and limitations of the model are discussed and future research needs are identified.
CHAPTER 2 – Air Quality Model for Dublin
2. Introduction

Air pollution is a major environmental and health problem affecting industrialized and developing countries around the world. The exposure to air pollutants such as particulate matter have harmful effects on humans, mainly respiratory difficulties for sensitive people, and also detrimental consequences on ecosystems, like the damage to agricultural crops and to aquatic and terrestrial ecosystems.

Nowadays, the majority of large cities in the world have problems related to air quality, not only in developed countries but also in cities of developing areas. In the UK, Vardoulakis et al. (Vardoulakis et al., 2007) identified two hundred air pollution hotspots caused by heavy volumes of traffic in urban street canyons. One of the topics in which the European Commission has shown a greater concern is the necessity of developing actions that allow increasing the knowledge on transport and dynamics of atmospheric pollutants to assure the accomplishment of legislation and to inform the population about their levels. At the same time, the mother directive 1996/62/EC and daughter directives establish the possibility of using modelling techniques to assess air quality.

The aim of the research presented in this thesis is based on the necessity to understand the dynamics of air pollution with the aim of obtaining a precise estimation of the air quality through high-resolution modelling: this will result in a personal exposure model for Dublin’s inhabitants. The precision and accuracy of the utilised air quality models will fit with this objective and the required degree of complexity of the models will be decided according to the circumstances. Air quality models (AQMs) are mathematical simulations of the atmospheric processes responsible for air pollution and they are also an essential instrument to improve the knowledge of the behaviour of various compounds in the atmosphere, especially the relation between emissions, atmospheric transport, chemistry and deposition. Dispersion models and data machine learning algorithms are used to estimate or to predict the downwind concentration of air pollutants emitted from sources such as industrial plants, vehicular traffic or accidental chemical releases. Such models are important to governmental agencies tasked with protecting and managing the ambient air quality. They also serve to assist in the design of effective control strategies to reduce emissions of harmful air pollutants and for health impact studies of personal exposure.

Dispersion models vary depending on the mathematics used to develop the model, but all require the input of data that may include:
Meteorological conditions such as wind speed and direction, the Pasquill atmospheric stability class (amount of atmospheric turbulence), the temperature, the rainfall rate, the atmospheric pressure, cloud cover, solar radiation, etc.

Emissions parameters such as source location and height, type of source (i.e., fire, pool or vent stack) and exit velocity, exit temperature and mass flow rate.

Terrain elevations at the source location and at the receptor location(s), such as nearby homes, schools, businesses and hospitals.

The location, height and width of any obstructions (such as buildings or other structures) in the path of the emitted gaseous plume, surface roughness or the use of a more generic parameter “rural” or “city” terrain.

Many of the modern AQMs include a pre-processor module for the input of meteorological and other data, and many also include a post-processor module for graphing the output data and/or plotting the area impacted by the air pollutants on maps. The models use this specific input data to run various algorithms to estimate the dispersion of pollutants between the source and receptor. The models output is in the form of a predicted time-averaged concentration at the receptor. These predicted concentrations are added to suitable background concentrations and compared with the relevant ambient air quality standard or guideline. It has to be noted that the estimation of these background values and the choice of air quality standard in fact are often the most important steps.

2.1 Note about air quality modelling

The Earth's atmosphere is divided in layers (Figure 6) of different extension:

- Troposphere: it’s the layer closest to the Earth's surface and it extends from sea-level to a height of about 18 km and contains about 80% of the mass of the overall atmosphere;
- Stratosphere: it extends from 18 km to about 50 km;
- Mesosphere: it extends from 50 km to about 80 km.

There are other layers above 80 km, but they are insignificant with respect to atmospheric dispersion modelling. The troposphere is divided in 2 parts: a lower part, the atmospheric boundary layer (ABL), which extends from the Earth’s surface to about 2 km in height and an upper part, the free troposphere, which extends up to the 18 km height of the troposphere. The division of the troposphere in these two sub-
layers is due to the fact that the air temperature in the troposphere decreases with increasing altitude until it reaches what is called the inversion layer, after which the air temperature increases with increasing altitude. The inversion layer represents the border between the 2 troposphere sub-layers.

![Atmosphere layers](image)

**Figure 6: Atmosphere layers**

Typically during daytime a convective boundary layer of turbulent air grows and can be up to 2 km in thickness (Schnelle and Dey, 2000). Near sunset, the turbulence weakens to form a neutral residual layer below the temperature inversion. At night the air generally cools over the surface to form a stable boundary layer below the residual layer (Stull, 1998). The night-time boundary layer can be as low as tens of metres (Schnelle and Dey, 2000). A stable boundary layer can also form during the day or night when air is carried across a relatively cooler land or water surface.

The ABL is the most important part of the atmosphere with respect to the emission, transport and dispersion of airborne pollutants. The part of the ABL between the Earth's surface and the bottom of the inversion layer is known as the mixing layer. Almost all of the airborne pollutants emitted into the ambient atmosphere are transported and dispersed within the mixing layer. Some of the emissions penetrate the inversion layer and enter the free troposphere above the ABL.

### 2.1.1 Gaussian air pollutant dispersion equation

The technical literature on air pollution dispersion is quite extensive and dates back to 1936 when Bosanquet and Pearson (Bosanquet and Pearson, 1936) derived an air pollutant plume dispersion equation which did not assume Gaussian distribution nor did it include the effect of ground reflection of the pollutant plume. The assumption of Gaussian distribution for the vertical and crosswind dispersion of the plume and also
the effect of ground reflection of the plume were then included by Sir Graham Sutton in 1947 (Sutton, 1947). In response to the Great Smog of 1952, the British Parliament introduced the Clean Air Act 1956, followed by another one in 1968 (Clean Air Act 1968). These stringent environmental control regulations provided the stimulus that generated an immense growth in the use of air pollutant plume dispersion calculations between the late 1960s and today. A great many computer programs for calculating the dispersion of air pollutant emissions were developed during that period of time and they were called "air dispersion models". The basis for most of those models was the Complete Equation For Gaussian Dispersion Modelling Of Continuous, Buoyant Air Pollution Plumes shown below in Equation 2.1 (Beychok, 2005, Turner, 1994):

\[
C = \frac{Q}{u} \frac{f}{\sigma_y \sqrt{2\pi}} \frac{g_1 + g_2 + g_3}{\sigma_z \sqrt{2\pi}} \tag{Equation 2.1}
\]

Where:

- Crosswind dispersion parameter: \( f = \exp\left[-y^2/(2\sigma_y^2)\right] \)
- Vertical dispersion parameter: \( g = g_1 + g_2 + g_3 \)
- Vertical dispersion with no reflections: \( g_1 = \exp\left[-(z - H_e)^2/(2\sigma_z^2)\right] \)
- Vertical dispersion for reflection from the ground: \( g_2 = \exp\left[-(z + H_e)^2/(2\sigma_z^2)\right] \)
- Vertical dispersion for reflection from an inversion aloft:

\[
g_3 = \sum_{m=1}^{\infty} \left\{ \exp\left[-(z - H_e - 2mL)^2/(2\sigma_z^2)\right] + \exp\left[-(z + H_e + 2mL)^2/(2\sigma_z^2)\right] \right\}
\]

\[+ \exp\left[-(z + H_e - 2mL)^2/(2\sigma_z^2)\right] + \exp\left[-(z - H_e + 2mL)^2/(2\sigma_z^2)\right] \tag{Equation 2.2}\]

- Concentration of emissions, in g/m³ (at any receptor located \( x \) meters downwind from the emission source point, \( y \) meters crosswind from the emission plume centreline, \( z \) meters above ground level) : \( C \)
- source pollutant emission rate, in g/s: \( Q \)
- horizontal wind velocity along the plume centreline, m/s: \( u \)
- height of emission plume centreline above ground level, in m: \( H_e \)
- vertical standard deviation of the emission distribution, in m: \( \sigma_z \)
- horizontal standard deviation of the emission distribution, in m: \( \sigma_y \)
- height from ground level to bottom of the inversion aloft, in m: \( L \)

The above equation not only includes upward reflection from the ground, it also includes downward reflection from the bottom of any inversion lid present in the atmosphere.
The sum of the four exponential terms in $g_3$ converges to a final value quite rapidly. For most cases, the summation of the series with $m = 1$, $m = 2$ and $m = 3$ will provide an adequate solution.

$\sigma_z$ and $\sigma_y$ are functions of the atmospheric stability class (i.e., a measure of the turbulence in the ambient atmosphere) and of the downwind distance to the receptor. The two most important variables affecting the degree of pollutant emission dispersion obtained are the height of the emission source point and the degree of atmospheric turbulence. As a general rule, the more turbulence, the better the degree of dispersion.

2.1.1.1 The Briggs plume rise equations

The Gaussian air pollutant dispersion equation presented above requires the input of the pollutant plume’s centreline height above ground level ($H_e$). $H_e$ is the sum of the actual physical height of the pollutant plume’s emission source point $H_s$ plus the plume rise due the plume’s buoyancy $\Delta h$ and it is described by the following Equation 2.3:

$$H_e = H_s + \Delta h$$  \hspace{1cm} (Equation 2.3)

Figure 7 shows a graphical representation of the pollutant plume’s centreline height above ground level of a buoyant Gaussian air pollutant dispersion plume.

The plume’s buoyancy $\Delta h$ is determined in most of the air dispersion models developed between the late 1960s and the early 2000s using the Briggs equations.
G.A. Briggs first published his plume rise observations and comparisons in 1965 (Briggs, 1965). Then he compared many of the plume rise models then available in the literature at the CONCAWE (a Dutch organization) symposium (Briggs, 1968) and in a publication edited by Slade (Slade, 1968). In 1969, Briggs proposed a set of plume rise equations which have become widely known as "the Briggs equations" (Briggs, 1969). Subsequently, Briggs modified his 1969 plume rise equations in 1971 and in 1972 (Briggs, 1972, Briggs, 1971).

Briggs divided air pollution plumes into these four general categories:

- Cold jet plumes in calm ambient air conditions
- Cold jet plumes in windy ambient air conditions
- Hot, buoyant plumes in calm ambient air conditions
- Hot, buoyant plumes in windy ambient air conditions

Although Briggs proposed plume rise equations for each of the above plume categories, it is important to emphasize that "the Briggs equations" which become widely used are those that he proposed for hot buoyant plumes in windy ambient air conditions. These equations are based on observations and data involving plumes from typical combustion sources such as the flue gas stacks from steam-generating boilers burning fossil fuels in large power plants. Therefore the stack exit velocities were probably in the range of 6 to 30 m/s with exit temperatures ranging from 120 to 260 °C.

The plume’s buoyancy $\Delta h$ can be calculated using the Briggs’s formulae following the logic diagram (Beychok, 2005) presented in Figure 8 below:
Figure 8: Logic diagram for using the Briggs' equation (Beychok, 2005)

Where:
- $\Delta h$ is the plume rise measured in m;
- $F$ is the buoyancy factor, in $m^4s^{-3}$;
- $x$ is the downwind distance from plume source, in m;
- $x_f$ is the downwind distance from the plume source to the point of maximum plume rise, in m;
- $u$ is the wind speed at actual stack height, in m/s;
- $s$ is the stability parameter, in $s^{-2}$.

2.1.2 Background pollution models

In common usage, the term background level refers to the concentration observed in remote areas relatively unaffected by local pollution sources. As such “background” concentrations are defined as those ambient pollutant concentrations in an air shed that would occur in the absence of local anthropogenic emissions. Thus “background” is the sum of local natural emissions and those pollutants (generally long-lived species) conveyed into an area that are of both natural and anthropogenic origin.

For the purposes of this thesis and in a policy relevant context, the background concentration must take into account contributions from uncontrollable sources that can affect local concentrations. In this context, background particulate matter can be
defined as the concentration resulting from local natural sources, long range transported natural and anthropogenic sources.

It is important to define an approach as accurate as possible to calculate values for background concentrations for particulate matter because these values substantially contribute to the overall pollutant levels. In general, environmental data are typically very complex to model due to underlying interrelations between numerous variables of different type which yield a complicated combination of relations. Aiming at the simplification of statistical complexity, many attempts to model the former interrelations have been undertaken. For the purpose of this work, two different approached to model background PM10 concentrations are discussed:

- **Time series analysis:** A time series is a collection of observations of well-defined data items obtained through repeated measurements over time. Time series data have a natural temporal ordering. This makes time series analysis distinct from other common data analysis problems, in which there is no natural ordering of the observations. A time series model will generally reflect the fact that observations close together in time will be more closely related than observations further apart (http://en.wikipedia.org/wiki/Time_series);

- **Artificial neural networks:** “An Artificial Neural Network (ANN) is a mathematical representation of the operation of biological neural networks, in other words it is an emulation of a biological neural system. ANNs are highly adaptive to non-parametric data distributions and make no prior hypotheses about the relationships between the variables. ANNs are also less sensitive to error term assumptions and they can tolerate noise, chaotic components and heavy tails better than most of the others computational methods. An ANN consists of an input layer, one or more layers of neurons (mathematical entities whose behaviour is governed by a predefined function) and an output layer. Each ANN has to be trained first, by using a number of data instances as inputs, and by trying to “fit” to some other data instances (output), commonly in terms of predicting the (future) behaviour of parameters of interest with the aid of their history” (Kyriakidis et al., 2009).

These two modelling approaches are presented and analysed in further detail in Chapter 3.
2.2 Existing atmospheric dispersion models

The European Topic Centre on Air and Climate Change, which is part of the European Environment Agency (EEA), maintains an online Model Documentation System (MDS) that includes descriptions and other information for almost all of the dispersion models developed by the countries of Europe, which are currently (October 2011) 142. For the purpose of this work, only the AQMs included and recommended in the Irish EPA Guidance Note (Air Dispersion Modelling from Industrial Installations Guidance Note) published in 2010 (Marnane et al., 2010) are taken into consideration. The air quality models described in this document can be divided into 2 categories:

1. Gaussian plume models;
2. Lagrangian puff models.

The hypotheses of the Gaussian plume model are described in Chapter 2.1.1 Gaussian air pollutant dispersion equation); this model assumes that the concentration within the plume is proportional to the emission rate and inversely proportional to the wind speed at the point of release. Therefore, at wind speeds close to zero, the predicted concentration approaches infinity and the Gaussian representation of the plume is no longer valid. Another problem arises from the steady state assumption, according to which the meteorological conditions between the source and the receptor are constant for each modelled hour (steady state assumption). This assumption may no longer be valid for the following two conditions (Marnane et al., 2010):

1. **Non-uniform meteorological conditions**: If there are regions within the modelling domain that are complex due to significant terrain features or changes in land use (e.g. urban / rural interface or a valley or coastal region), then the assumption that the wind field within the entire modelling domain is no longer correct: this observation is mostly valid for the Dublin area;

2. **Dispersion over large distances**: Gaussian plume models calculate concentrations in a straight line from source to receptor within the modelling domain for each hour. However they do not take into account the time taken for the plume to travel from source to receptor. Concentration calculations under low wind speeds may be made at receptors tens of kilometres from the source when it would be physically impossible for the plume to travel this distance. Furthermore, at large distances the steady-state assumption is unlikely to be consistent with reality. For this reason Gaussian plume models should only be used for predicting maximum concentrations within about 10km
from the source. As such, this approach is no relevant for the modelling of personal exposure concentrations.

The Lagrangian puff models mathematically follow pollution plume parcels (also called particles) as the parcels move in the atmosphere and they model the motion of the parcels as a random walk process. The Lagrangian model then calculates the air pollution dispersion by computing the statistics of the trajectories of a continuous plume represented by a series of discrete parcels of pollutant material. A Lagrangian model uses a moving frame of reference as the parcels move from their initial location. The total concentration at the receptor is then calculated based on the contribution of all nearby puffs. Puff models may also use the Gaussian distribution to describe the dispersion of pollutants within each puff. This approach, on the other side, has a higher degree of accuracy than the Gaussian dispersion models and therefore, it is more relevant to personal exposure concentrations modelling which is the object of the study described here.

2.3 Summary of models commonly used in Ireland

This section deals with the air quality models commonly used in Ireland and it presents notes from the Irish EPA Air Dispersion Modelling from Industrial Installations Guidance Note (Marnane et al., 2010). It has to be noted that this document is the only guidance available in Ireland: as such, it is reported here even if this guidance is aimed at industrial emissions and it doesn't address the modelling of many urban sources such as traffic and residential.

The AQMs described below are divided in:

1. Screening models;
2. Advanced models;
3. Urban street canyon models.

2.3.1 Screening Models

Screening dispersion models are conservative analytical modelling methods for estimating extreme upper bound concentrations, which are calculated considering a standard set of meteorological conditions and simplified assumptions/representations of source-receptor geometry. The primary purpose of screening modelling is to quickly and easily eliminate sources, whose impacts are low enough that they will not pose a threat to ambient air quality standards or ambient increment standards, from further analysis. Another purpose of screening models is to identify the "worst-case"
operating scenario for a particular source or group of sources that require a more refined assessment to demonstrate compliance with ambient air quality standards.

Screening modelling tends to be easy-to-run, quick and conservative: it tends to over-predict air contaminant concentrations because of the use of standardized "worst-case" meteorological data. As such screening dispersion models may be used to demonstrate compliance with ambient air quality standards or to indicate the proposed operating scenario most likely to produce the maximum predicted air contaminant concentrations to be further analysed via refined modelling.

They have limited capabilities for defining the receptor locations, the pollutant averaging times and output options. Should the results of a screening model predict an exceedance of the air quality standards then a more advanced model should be used.

2.3.1.1 AERSCREEN

AERSCREEN is a steady-state Gaussian plume model which uses worst-case meteorological data to predict the ambient pollutant concentrations resulting from single continuous emission sources and it's based on AERMOD. It replaces SCREEN3 for regulatory screening air dispersion model in the USA. The model produces estimates of "worst-case" 1-hour concentrations for a single source, without the need for hourly meteorological data, and also includes conversion factors to estimate "worst-case" 3-hour, 8-hour, 24-hour, and annual concentrations (U.S.E.P.A., 2011). AERSCREEN is intended to produce concentration estimates that are equal to or greater than the estimates produced by AERMOD with a fully developed set of meteorological and terrain data, but the degree of conservatism will vary depending on the application.

The input requirements and model outputs are the followings (U.S.E.P.A., 2011):

- **Stack details**: Emission rate (g/s), stack height (m), stack diameter (m), gas exit velocity (m/s) or volume flow (m³/s), temperature (K);
- **Meteorology inputs**: Min / max ambient temperature, min wind speed, anemometer height, surface characteristics;
- **Land use inputs**: Urban / Rural Option. Surface characteristics;
- **Terrain inputs**: Digital elevation model (DEM) file to run AERMAP;
- **Building downwash inputs**: Building dimensions and stack location relative to building;
- **Output concentrations**: 1-Hour. Conversion to other averaging times using adjustment factors;
- **Output files**: Range of file types for subsequent analysis and contour plotting.

The model can be downloaded free from the USEPA website:


A simplified version of this model is used for modelling the dispersion from area sources, which is described in Chapter 6.

**2.3.1.2 ADMS-SCREEN**

ADMS-SCREEN is a Gaussian type steady-state plume model which uses worst-case and internal meteorological data to predict the ambient pollutant concentrations resulting from single continuous point sources (C.E.R.C., 2004). The plume rise, dispersion and building downwash calculations performed by the model use the algorithms of the advanced air dispersion model ADMS 3 (C.E.R.C., 2000). The model is designed to predict ambient concentrations that are equal to or greater than those predicted by ADMS 3.

The input requirements and model outputs are the followings:

- **Stack details**: Emission rate (g/s), stack height (m), stack diameter (m), gas exit velocity (m/s) or volume flow (m³/s), temperature (°C);
- **Meteorology inputs**: None for short-term calculations. Specify site location (UK only) for long-term calculations;
- **Terrain inputs**: None, simple terrain only;
- **Building downwash inputs**: Building dimensions and stack location relative to building;
- **Output concentrations**: 1-Hour, 24-hour, annual average concentrations and percentiles;
- **Output files**: Range of file types for subsequent analysis and contour plotting.

ADMS-SCREEN is available to purchase from Cambridge Environmental Research Consultants (CERC). Further details are provided on the CERC website (http://www.cerc.co.uk/).

**2.3.2 Advanced models**

Advanced air dispersion models are usually based on more modern scientific theories and more complex mathematical formulations than screening dispersion models. Advanced models can assess the impact of complex distributions of multiple sources and numerous buildings. Detailed input data regarding meteorology, land use and terrain are required by these models in order to allow them to represent the
atmospheric processes contributing to pollutant dispersion. Significant data pre-processing is often required to prepare the input files used by these models.

Advanced models may have limitations in their ability to assess certain scenarios (such as calm hours, terrain downwash and coastal fumigation). An outline of the advanced dispersion models in general use in Ireland is provided below, including a discussion on the input requirements. These models are mainly used in Ireland because they are suitable for modelling emissions from industrial sources, in fact their use for such assessment is recommended in the Irish EPA Air Dispersion Modelling from Industrial Installations Guidance Note (Marnane et al., 2010), which is the only guidance available in Ireland.

2.3.2.1 AERMOD

AERMOD is a steady-state Gaussian plume model which can simulate dispersion from multiple sources using up-to-date concepts regarding boundary layer characterisation and dispersion (U.S.E.P.A., 2005, U.S.E.P.A., 2004a). There are two input data processors that are regulatory components of the AERMOD modelling system: AERMET (U.S.E.P.A., 2004c), a meteorological data pre-processor that incorporates air dispersion based on planetary boundary layer turbulence structure and scaling concepts, and AERMAP (U.S.E.P.A., 2004b), a terrain data pre-processor that incorporates complex terrain using USGS Digital Elevation Data. Other non-regulatory components of this system include: AERSCREEN, AERSURFACE (a surface characteristics pre-processor) and BPIPPRIME (a multi-building dimensions program incorporating the GEP technical procedures for PRIME (Schulman et al., 2000) applications).

AERMET is a meteorological pre-processor which produces two types of meteorological input files required by AERMOD:

- a surface file which contains various meteorological and surface scalar parameters;
- a profile file which consists of meteorological data at more than one height for use when undertaking an on-site monitoring programme.

The input requirements for AERMET include surface characteristics (such as surface roughness, Bowen ratio, and albedo) and hourly meteorological data (wind speed, wind direction, cloud cover and temperature). AERMET requires user input data on site-specific surface characteristics. Recent guidance by the U.S.E.P.A. regarding the implementation of AERMOD provides the methods for determining the correct surface characteristics required by AERMET (U.S.E.P.A., 2009b).
The AERMAP terrain pre-processor (U.S.E.P.A., 2009a, U.S.E.P.A., 2004b) is used to prepare the terrain information required by AERMOD for complex terrain scenarios. AERMAP sets a hill height scale, which is the height that has the greatest influence on dispersion, for each individual receptor modelled by AERMOD. AERMAP requires terrain information for the modelling domain in the form of a Digital Elevation Model (DEM) file.

The model has U.S.E.P.A. regulatory status (U.S.E.P.A., 2005) and is available to download from: http://www.epa.gov/scram001/dispersion_prefrec.htm.

AERMOD is not believed to be suitable for the modelling tasks part of this study because of the unavailability of the dataset required to AERMET in order to perform the calculations and also because of the complexity and time involved in setting up the weather pre-processor.

2.3.2.2 ADMS 4

ADMS 4 is an advanced steady-state Gaussian type plume model which can simulate dispersion from multiple sources (C.E.R.C., 2007). A puff model for the assessment of instantaneous releases of pollutants is also included. ADMS 4 uses an in-built meteorological pre-processor developed by the UK Met Office and also includes a terrain convertor utility for preparation of terrain data in ADMS 4 format.

ADMS 4 is available to purchase from Cambridge Environmental Research Consultants (http://www.cerc.co.uk/). The package includes a graphical user interface.

Modelling of terrain requires the preparation of an input file with x- and y-coordinates and terrain heights for the modelling domain. The ADMS Terrain Convertor utility can be used to convert Irish National Grid Digital Terrain Model (DTM) files into the required format for the model.

2.3.3 Urban street canyon models

The central parts of most of the cities are characterised by close spacing of buildings; this particular layout is called urban street canyon, which is a relatively narrow street between buildings that line up continuously along both sides. An important characteristic of such configuration is the aspect ratio or height/width ratio, which is the ratio of the mean height of the buildings to the width of the street. The ratio is especially crucial when it comes to determining the type of airflow in the canyon and it may cause a stable circulatory vortex being established within the street canyon if specific wind conditions are met. The combination of vehicle emissions and reduced dispersion in these circumstances can lead to high levels of pollution (Buckland and
 Middleton, 1999), which in turn pose a substantial health risk to pedestrians. When the wind blows across a street canyon, a vortex might be generated if particular conditions are met.

There are numerous models used in urban street canyon applications, which have been proved to be useful in air quality and traffic management, urban planning, interpretation of monitoring data, pollution forecasting, population exposure studies, etc. The models most commonly used for this task are the followings:

1. STREET;
2. CPBM;
3. AEOLIUS;
4. SLAQ;
5. ADMS Urban;
6. OSPM.

These models are discussed and presented in further detail in Chapter 4.

2.4 Existing Personal Exposure models

The basic concepts that are currently used in exposure assessments were developed in the early 1980s by Duan (Duan, 1982) and Ott (Ott, 1982). The term exposure, in the epidemiological domain, is defined as “the event that occurs when a person comes in contact with the pollutant”. A person's individual lifestyle, geographic location, socio-economic conditions and genetic predispositions will impact an exposure outcome, along with the real world circumstances of the interaction event itself (Doore, 2010). In this definition of exposure, a toxic agent is an entity that contacts a target, which is the entity that receives an exposure or a dose (e.g., a human, a population group) (Lioy, 1990). This thesis focuses on environmental exposure defined here as human exposure to PM10 originating from manmade sources or harmful levels of natural elements within a person's environment; for the purpose of this thesis, the agent is PM10 and the targets are the commuters in the Great Dublin area.

This is a definition of an instantaneous contact between a person i (the commuter) and a pollutant with concentration c (PM10), at a particular time t. This definition refers to a contact with a pollutant, but it is not necessary that the person inhales or ingests the pollutant. When the duration of exposure is also taken into consideration, the result is an “average exposure”, calculated by averaging the pollutant concentration over the specified period (e.g. hourly/daily/annual mean). In air pollution epidemiology,
the unit “concentration” is most commonly used. It refers to the “average exposure” to which the population has been exposed over a specific time (Monn, 2001).

The definition of personal exposure presented here refers to concentrations of pollutants in the environment surrounding the target (commuter). However, once the pollutant has crossed the body’s boundary and reaches the target tissue, the concept of “dose” is used (Ott, 1982). “Dose” is the amount of material absorbed or deposited in the body for an interval of time and is measured in units of mass (or mass per volume of body fluid in a biomarker measurement).

Dose can be determined as “average intake” (National Research Council (N.R.C.), 1991) when data on exposure values are available: this can be calculated by dividing intake by time (unit: mass×time−1).

![Diagram](image)

**Figure 9:** “Sequence of exposure” from source to biologically effective dose (Monn, 2001)

Various research projects have been carried out to define the variables that affect the most the personal exposure to air pollutants.

A series of experiments completed by Harvard School of Public Health demonstrated that it is important to acknowledge that personal–ambient relationships are greatly dependent on ambient conditions (e.g., season, meteorology) (Sarnat et al., 2006). The EXPOLIS project (Edwards et al., 2006) along with work carried out by Wallace (Wallace, 2001) proved that socio-demographic and environmental factors which define time–activity subgroups also define quantifiable differences in personal exposure. Graham and McCurdy (Graham and McCurdy, 2004) performed a series of statistical tests on the Consolidated Human Activity Database (CHAD), developed by the US EPA National Exposure Research Laboratory (NERL): they identified age and gender as principal attributes to explain statistically significant differences in time–activity patterns followed by weather parameters, physical activity level and weekend/weekday classification.

There are several stochastic and deterministic models used in personal exposure assessment to air pollutants applications. The models most commonly used for this task are the followings:

- AirGIS: it is a human exposure modelling system for traffic air pollution developed by the National Environmental Research Institute (NERI) for use in
Danish air pollution epidemiologic studies and human exposure studies. The model system has a high spatial resolution (address) and high time resolution (hour). The system is named AirGIS and integrates dispersion models, available digital maps (buildings, roads and address points), Danish national administrative databases on buildings and populations, traffic data, and a geographic information system (GIS). The Danish Operational Street Pollution Model (OSPM) is used to calculate exposures. Apart from modelling exposure at address level, the system is also able to model exposure under transport provided GPS data is available (Jensen, 2006).

- **EXPAND**: Based on the measured results of the EXPOLIS project (Edwards et al., 2006), the model utilises as input values data on the spatial location of the population, time-microenvironment activity data and computed spatial pollutant concentration distributions (Kousa et al., 2002).

- **pCNEM**: it generates a sequence of pollutant concentrations to which a randomly selected individual is exposed over time. The generation is a complex stochastic process that follows the randomly selected individual in his activities over the period of the simulation. The individual is thought of as visiting one *microenvironment* (*ME*) after another as he or she is involved in his or her activities through time (Zidek et al., 2005).

- **EMI**: The U.S. Environmental Protection Agency's Office of Research and Development, National Exposure Research Laboratory (NERL) and collaborators developed this tool for enhancing cumulative risk assessments. It predicts individual exposures for multiple air pollutants from ambient concentrations, meteorology, and questionnaire information such as building characteristics, occupant behaviour related to building operation, indoor sources, and time-activity patterns. This research is being conducted in response to the needs of numerous air pollution cohort health studies, including the National Children's Study (led by the Eunice Kennedy Shriver National Institute of Child Health and Human Development), and complements existing population-based exposure models (e.g., SHEDS-PM) (Zartarian and Schultz, 2010).

- **SHEDS**: the Stochastic Human Exposure and Dose Simulation (SHEDS-PM) model is a population exposure model for particulate matter (PM). SHEDS-PM is a probabilistic model that estimates the population exposure to particulate matter by randomly sampling from various input distributions. Additional model inputs include demographic data for the population being modelled and human activity pattern data from EPA's Consolidated Human Activity Database.
(CHAD). Model outputs include distributions of daily total PM exposures in various microenvironments (indoors, in vehicles, outdoors), and the contribution from PM of ambient origin to daily total PM exposures in these microenvironments (Burke et al., 2001)

2.5 Definition of the air quality model

Ireland does not have serious outdoor air quality problems. This is largely due to the eradication of the burning of coal in many urban areas during the 1980s and the early 1990s (http://www.epa.ie/whatwedo/monitoring/air/). The biggest threat now facing the Country’s air quality is emissions from road traffic: in fact this source of pollutants may cause some locations and thus individuals to experience poor air quality, which may have adverse health effects.

Living or attending school near major roadways has been associated with numerous health outcomes in recent years, including asthma exacerbation (Gordian et al., 2005), other respiratory illnesses (Brauer et al., 2002), and excess risk of mortality from cardiopulmonary disease (Hoek et al., 2002), stroke (Maheswaran and Elliott, 2003), or all causes (Finkelstein et al., 2004). While this growing observational literature has been interpretable and robust, with relative risks that indicate a large public health impact, simulation studies (Baxter et al., 2009) have shown significant exposure misclassification associated with the use of proximity measures relative to “gold standard” air pollution exposure estimates.

While these studies and related work have provided insight about exposure patterns and health effects, and can also characterize air pollution hot spots for public health interventions, significant limitations remain. The above-mentioned studies leveraged temporally rich and spatially dense monitoring data, yet lacked sufficient spatial resolution to capture micro-scale concentration patterns – the model by Jerrett et al. (Jerrett et al., 2005) relied on only 23 sites to capture spatial patterns across a huge metropolitan area, and the model by Gryparis et al. (Gryparis et al., 2007) determined essentially uniform EC concentrations within Boston given a small number of monitoring sites from which a spatial surface could be determined.
As such, the very sparse permanent air quality monitoring stations around the Dublin area (Figure 10) are not believed to be able to produce a sufficient spatial resolution for the pollutants levels: the data collected by these monitoring stations are therefore not representative of the pollutants concentrations at street level. This detail is crucial when accurate information on the exposure of individuals is required: scientific studies on the health impacts of air quality need to pair estimates of individual exposure to air pollution with the assessment of individual health outcomes. The most effective way for Ireland to improve on the assessment of population exposure to air pollution is through developing a robust air quality model for the country. This work deals specifically with the creation of an air quality model for the Great Dublin Area and its Satellite Towns.

2.5.1 The model

A graphic representation of the spatial variation of the different air pollution components that contribute to pollutant concentrations at street level is presented in Figure 11.
The final street level concentration will result from the sum of the following elements:

1. Urban background concentrations, sum of the regional background and the urban increment;
2. Contribution from important point and area sources;
3. Contribution from traffic.

These three elements will be modelled separately and then the resultant modelled air pollutant’s concentrations will be added up to derive the air pollutant’s concentration at street level (Figure 12).
For the personal air pollution exposure model developed here, the sum presented in Figure 12 is performed and displayed using ArcGIS by assigning each component to a different layer.

The modelling process for each component is briefly described below and subsequently presented in detail in the Chapters 3-6.

### 2.5.1.1 Urban background concentrations

The urban background air pollutant’s concentration levels are constituted by the sum of the background concentrations and the distributed contributions of the city itself, such as the overall traffic in the city, area sources and regional pollution.

This modelling of this element employs data recorded by the Irish Environmental Protection Agency’s air quality permanent monitoring stations around the Greater Dublin Area. The stations categorised as “urban background” using the criteria proposed by the European Environmental Agency (Van Dingenen et al., 2004) will be employed for this task. Among those criteria are the distance of the station from large pollution sources such as cities, power plants and major motorways, and the traffic volume:

- **Natural background**: distance from large pollution sources >50 km;
- **Rural background**: distance from large pollution sources 10–50 km;
- **Near-City background**: distance from large pollution sources 3–10 km;
- **Urban background**: <2500 vehicles/day within a radius of 50 m.

Using a neural network modelling approach, historical air pollution data and weather data is analysed to identify the relationship between the two, which will then be used in combination with forecasted weather data to model and predict the required urban background concentration levels (Figure 13). The results are then imported to ArcGIS for visualisation purposes (Figure 14) and for further analysis and combination with the other components of street level pollution. The implementation of this approach to background modelling is presented in Chapter 3.
2.5.1.2 Contribution from important point and area sources

This element of the modelling procedure is discussed in detail in Chapter 5 and will account for the following typologies of air pollution sources:

- **Point source**: A point source is a single, identifiable source of air pollutant emissions, which could be either elevated or at ground-level. It includes sources like combustion furnace flue gas stacks, power stations, steel works, foundries, incinerators, wood and pulp processors, paper mills, refineries, chemical production, etc. (Kibble and Harrison, 2005). Quoting the U.S.E.P.A. “Example Application of Modelling Toxic Air Pollutants in Urban Areas” (U.S.E.P.A., 2002): “Point sources involve the release of emissions from a
well-defined stack or vent, at known physical stack parameters and operational conditions. Consequently, characterizing point sources for modelling is fairly straightforward. The basic model inputs for any point source are: stack location coordinates, the physical stack parameters (height above ground level and inside diameter at stack exit); operational conditions (gas velocity and temperature at stack exit); building dimensions (height, width, depth), and emission rate.

- **Area source**: Those facilities or activities whose individual emissions do not qualify them as point sources are called area sources. An area source is a two-dimensional source of diffuse air pollutant emissions. Area sources represent numerous facilities or activities that individually release small amounts of a given pollutant, but collectively can release significant amounts of a pollutant. Quoting the U.S.E.P.A. “Example Application of Modelling Toxic Air Pollutants in Urban Areas” (U.S.E.P.A., 2002): “Area and Other sources are smaller stationary sources that emit less than 10 tons per year of a single air pollutant or less than 25 tons per year of a combination of air toxics. Typical examples of area sources include neighbourhood dry cleaners and gas stations. Though emissions from individual area sources are often relatively small, collectively their emissions can be of concern particularly where large numbers of sources are located in heavily populated areas. Examples of other stationary sources include wildfires and prescribed burning, which have emissions that are being addressed through the burning policy agreed to by the EPA and the USDA”. It includes also the emissions from sources like a landfill or the evaporated vapours from a large spill of volatile liquid.

- **Volume source**: A volume source is a three-dimensional source of diffuse air pollutant emissions. Essentially, it is an area source with a third (height) dimension. Quoting the U.S.E.P.A. “Example Application of Modelling Toxic Air Pollutants in Urban Areas” (U.S.E.P.A., 2002): “There are two basic types of volume sources: surface-based or ground-level sources that may also be modelled as area sources, and elevated sources. As with area sources, emissions from volume sources are assumed to be of neutral buoyancy. The effective emission height of a surface-based volume source, such as a surface rail line, is usually set equal to zero. An example of an elevated volume source is an elevated conveyor with an effective emission height set equal to the height of the conveyor. A source may be defined as a volume source for modelling when its emissions can be considered to occur over a certain area
and within a certain depth of space. The important parameters used to characterize volume sources for dispersion modelling are location, release height and initial lateral and vertical dimensions". Examples of volume sources are the fugitive gaseous emissions from piping flanges, valves and other equipment at various heights within industrial facilities such as oil refineries and petrochemical plants. Another example would be the emissions from an automobile paint shop with multiple roof vents or multiple open windows.

The Gaussian dispersion modelling approach (Figure 15) used for the calculation of the pollution concentration levels from such sources is presented in Chapter 5. The
output of these calculations will be imported in ArcGIS for visualisation purposes and further analysis. An example of the visualisation in ArcGIS of the PM10 contribution from important point and area sources is presented in Figure 16.

2.5.1.3 Contribution from traffic

Figure 17: Traffic contribution in street canyons model, logic scheme

Figure 18: Visualisation in ArcGIS of PM10 traffic related levels

Nowadays large cities such as Dublin experience the problem of high volumes traffic on their main arteries: this huge amount of motor vehicles on the road emits pollutants which have a substantial impact on the air quality. The construction of high rise buildings to provide accommodation for a growing urban populations further
contributes to a deterioration in air quality in urban streets, as canyons with high aspect ratios reduce the rate of air exchange between roof and street level (So et al., 2005, Li et al., 2007). The recirculation of traffic generated pollutants and its contribution to the overall street level concentration is modelled taking into account this street canyon effect. The main principles of the street canyon model are depicted in Figure 17 (Berkowicz, 2000a) for the case of a wind direction nearly perpendicular to the street canyon. A receptor point in a leeward position is affected by the direct plume showing considerably higher concentrations than a receptor in windward position being exposed to the less concentrated recirculating air.

The model accounts for mechanical turbulence produced by the moving traffic (TPT) which acts in addition to the turbulence created by the roof level wind. This leads to a faster dispersion of the direct plume but also to an improved air exchange at roof level between the street canyon and the background air.

This modelling procedure is explored in detail in Chapter 4 and an example of visualisation in ArcGIS of the results for PM10 is presented in Figure 18.

### 2.5.2 Overall pollutant concentration at street level in ArcGIS

When the individual air pollutant concentration contributions have been calculated (as briefly described in the paragraphs above), they are imported in ArcGIS and dedicated layers are created. The capabilities of this software, such as the possibility of selecting only some data features and combining different thematic layers, are very important for unveiling patterns of air pollution and possible connection with particular activities or successions of events. This analysis will be presented in Chapter 6. An example of the manipulation of different thematic layers is presented in Figure 19.

The three layers of air pollutants concentration levels presented above can be combined in ArcGIS and the final street level concentration calculated. A visualisation of such calculation is presented in Figure 20.

This procedure for storing concentration levels in ArcGIS is extremely useful for the personal exposure calculation algorithm, which is also implemented in ArcGIS and is presented in Chapter 7.
Figure 19: Visualisation in ArcGIS of the different layers for PM10

Figure 20: Visualisation in ArcGIS of the total PM10 concentration levels
2.6 Definition of the personal exposure model

As mentioned above, exposure assessment is an integrated part of health risk assessment and management at a national, regional and local scale but presently no Irish exposure models exist to support these areas. Furthermore, health effects related to air pollution are important topics of epidemiological studies and exposure models could benefit air pollution epidemiological studies. It is indeed necessary to estimate air pollution levels in space and time and people’s contact to these levels to determine exposures which could be used in health risk assessment and management and in epidemiological studies (Jensen et al., 2001). It is also important to assess the impact of control measures in health risk assessment and management.

As such, the exposure model should have the following capabilities:

- Reproduction of the high spatial and temporal variability of ambient air pollution;
- Ability to estimate past, present and future population exposure at any location within the Great Dublin Area (for this project);
- Integration with existing air quality models;
- Marginal input data generation by relying on available data sources;
- Applicability to air pollution epidemiological studies;
- Applicability to health risk assessment and management;
- Integration with a GIS platform.

Compared to existing exposure models, the developed exposure model will add a geographic dimension to exposure modelling with a very high spatial resolution. GIS is chosen as a tool for exposure modelling since digital maps and administrative databases are available for Dublin and will be complete for the entire country within a few years. Because of this set of capabilities, GIS is regarded as a key decision supporting tool for Local Authorities in most developed countries like Ireland. As such, this study proposes to integrate the developed air quality dispersion modelling methodology for the personal exposure to particulate matter within the GIS package ESRI ArcGIS.

A conceptual diagram of the exposure model is presented in Figure 21.
The work presented here aims to create a set of integrated GIS-based exposure assessment tools based on monitoring data, spatial analysis, stochastic modelling and deterministic modelling. Stochastic models is developed through spatial regression models of the dependency of the ambient concentration measurements obtained within the project and influencing variables such as housing and traffic density, described in the GIS environment and adjusting for local area deprivation. The synthesis of these techniques leads to a valuable resource to be used in the assessments of health impacts, and in the development of wider exposure-reduction strategies.

The presented work also seeks to establish a methodology to perform air pollution exposure-response assessments in Ireland. It is also designed to lead to improved air quality action plans being developed across Ireland in the future leading to improved environmental awareness amongst the general public and local authorities alike. Furthermore, the PhD work yields a reliable framework within which epidemiological studies may be performed in the future, thus the impact of the proposed research is likely to extend beyond air pollution in the assessment of other environmental pollutants.

As such, the GIS personal exposure model is created following the steps listed below:

- Creation of a GIS air modelling system: Air Quality Models (AQM) for background, traffic, point and area related pollution are integrated within a GIS
platform. This creates a tool for calculating concentration levels of PM and correlating these with other thematic layers, such as land use and population density, in order to link localised peaks in air pollutants with particular activities;

- Creation of an emissions inventory: A crucial task is to build a detailed but accessible air emissions inventory for Dublin, suitable for use in air quality modelling studies;
- Data acquisition: city geometry, weather and traffic data are acquired in order to be used as input for the AQMs;
- Validation: the personal exposure model is validated using personal exposure data acquired during a measurement campaign on commuters in Dublin carried out as part of another PhD project, described in detail in Chapter 7;
- Outreach and dissemination: ArcGIS Online is chosen as the system for creating and sharing maps and geographic information. Even users with no GIS experience can easily upload their own information and interact with the available models and datasets.
CHAPTER 3 - Background PM10 modelling
3. Introduction

Understanding the background level of air pollution is of importance to policy makers, as it represents the fraction of pollution which is beyond direct control. In common usage, the term background level refers to the concentration observed in remote areas relatively unaffected by local pollution sources. The terms “background”, “natural background” and “baseline” concentrations are used widely, and sometimes with subtle nuances in meaning, in the literature. In Atmospheric Chemistry background concentration (level) is formally defined as:

“The background concentrations of relatively long-lived molecules, methane, carbon dioxide, halocarbons (CF3Cl, CF2Cl2, etc.) and some other species continue to rise due to anthropogenic input, so the composition of background air is undergoing continual change. Background concentration of a given species is sometimes considered to be the concentration of that impurity in a given air mass when the contribution from anthropogenic sources under study is absent” (McNaught et al., 2005).

As such “background” concentrations are defined as those ambient pollutant concentrations in an air shed that would occur in the absence of local anthropogenic emissions. Thus “background” is the sum of local natural emissions and those pollutants (generally long-lived species) conveyed into an area that are of both natural and anthropogenic origin. This definition is somewhat broader than some definitions that exclude all anthropogenic sources.

In defining the nature of “background”, an important related concept is that of “chemical weather”, a term that is gaining increasing acceptance and assuming greater importance in the Atmospheric Chemistry community. Chemical Weather is defined as: “local, regional and global distributions of important trace gases and aerosols and their variabilities on time scales of minutes to hours to days, particularly in light of their various impacts, such as on human health, ecosystems, the meteorological weather and climate” (Lawrence et al., 2005).

This definition of “Chemical Weather” recognises the tremendous short-term variability of the atmospheric composition, resulting from the strong influence of meteorological variability, and regional and temporally varying emissions. In this context “chemical weather” implies that “background” concentrations are not constant but instead are temporally variable.

For the purposes of this thesis and in a policy relevant context, the background concentration must take into account contributions from uncontrollable or
unidentifiable sources that can affect local concentrations. In this context, the background level can be defined as the concentration resulting from local natural sources, long range transported natural and anthropogenic sources that cannot be efficiently modelled on a source-by-source basis.

It is important to define an accurate approach to calculate background concentrations because these values substantially contribute to the overall pollutant levels that affect an individual's personal exposure. It has also to be noted that the required precision of the background modelling should reflect the extent to which the background concentrations affects the personal exposure (compared to the other components) and the accuracy with which we can define the controlling inputs (e.g. weather at different locations).

3.1 Options for modelling: literature review

In general, environmental data are typically very complex to model due to underlying interrelations between numerous variables of different types which yield a complicated combination of relations. Aiming at the simplification of statistical complexity, many attempts to model the former interrelations have been undertaken. For the purpose of this work, two different approached to model background PM10 concentrations are evaluated:

- Time series analysis;
- Artificial neural networks.

3.1.1 Time series analysis

Time series analysis and linear regression methods have been applied for decades and are well known and understood.

Robeson and Stein (Robeson and Steyn, 1990) created three statistical models to estimate daily maximum ozone (O3) concentrations in the lower Fraser Valley of British Columbia using measured concentrations from two monitoring stations during the time period 1978–1985. The three models are a univariate deterministic/stochastic model, a univariate autoregressive integrated moving average (ARIMA) model, and a bivariate temperature and persistence based regression model. The three models as well as a persistence forecast are tested by comparison with O3 concentrations observed during 1986; it is concluded that the bivariate model is superior to both univariate models. The ARIMA model has nearly the same predictive capability as the bivariate model while the mixed deterministic/stochastic model performs the worst.
The authors concluded that the traditional time series technique of decomposing a series into a trend, a cycle and a stochastic component was not appropriate for O3 air quality forecasting.

Milionis and Davies (Milionis and Davies, 1994) evaluated regression and stochastic models for air pollution and analysed their advantages and disadvantages. They reviewed general theoretical aspects of these models, focusing on the points of particular interest to air pollution modelling. They argued that, from both the conceptual and the practical points of view, stochastic models are preferred to regression models. They claimed that regression models used in air pollution studies didn’t deal with the mechanisms of atmospheric diffusion because of the lack of theoretical or a priori information about the functional form of the effect of emissions, meteorology, or the past history of pollutant concentrations. Stochastic models, such as the Box-Jenkins approach, were argued to have more advantages compared with regression methods for modelling air pollution.

Kukkonen et al. (Kukkonen et al., 2001a) developed a model for evaluating the mass-based concentrations of urban particulate matter. The basic model assumption was that local vehicular traffic was responsible for a substantial fraction of the street-level concentrations of both PM10 and NOx, either due to primary emissions or resuspension from street surfaces. The modelling system utilised data from an air quality monitoring network in the Helsinki Metropolitan Area. A linear relationship between the measured urban PM10 data against those of NOx in various urban surroundings was determined. The model performance was evaluated against the measured PM10 data from the monitoring by analysing the agreement between the measured and predicted hourly concentration time series, utilising the values of the fractional bias and the so-called index of agreement. The model predicted relatively well the yearly mean concentrations of PM10; however, its performance was substantially worse in predicting the hourly time series of the year.

Fuller et al. (Fuller et al., 2002) devised an empirical model to predict concentrations of PM10 at background and roadside locations in London. Factors to calculate primary PM10 and PM2.5 concentrations were derived from annual mean NOX, PM2.5 and PM10 measurements across London and south east England. These factors were used to calculate daily means for the primary and non-primary PM10 fractions for the London area. The model accurately predicted daily mean PM10 and EU Directive Limit values across a range of sites from kerbside to rural. The model relied on the currently observed ratio between NOX and PM10. This ratio remained constant over the 4 years antecedent the completion of the study.
3.1.2 Artificial Neural Network

"An Artificial Neural Network (ANN) is a mathematical representation of the operation of biological neural networks, in other words it is an emulation of a biological neural system. ANNs are highly adaptive to non-parametric data distributions and make no prior hypotheses about the relationships between the variables. ANNs are also less sensitive to error term assumptions and they can tolerate noise, chaotic components and heavy tails better than most of the others computational methods. An ANN consists of an input layer, one or more layers of neurons (mathematical entities whose behaviour is governed by a predefined function) and an output layer. Each ANN has to be trained first, by using a number of data instances as inputs, and by trying to “fit” to some other data instances (output), commonly in terms of predicting the (future) behaviour of parameters of interest with the aid of their history" (Kyriakidis et al., 2009).

The literature about the use of Artificial Neural Network for modelling PM10 concentration is vast and it details an approach which has been extensively adopted worldwide.

Gardner and Dorling (Gardner and Dorling, 1998) were among the first researchers to apply the multilayer perceptron, one type of artificial neural network, in the atmospheric sciences.

Pérez et al. (Pérez et al., 2000) measured hourly average concentrations of PM2.5 at a fixed point in the downtown area of Santiago (Chile) from May to September, in the years 1994 and 1995. They predicted concentrations at any hour of the day, by fitting a function of the 24 hourly average concentrations measured on the previous day. Then they compared the predictions produced by three different methods: multilayer neural networks, linear regression and persistence. Overall, the neural network gave the best results.

Kukkonen et al. (Kukkonen et al., 2003a) evaluated five neural network models, a linear statistical model and a deterministic modelling system for the prediction of urban NO2 and PM10 concentrations. The model evaluation work considered the sequential hourly concentration time series of NO2 and PM10, which were measured at two stations in central Helsinki, from 1996 to 1999. The models utilised selected traffic flow and pre-processed meteorological variables as input data. An imputed concentration dataset was also created, in which the missing values were replaced, in order to obtain a harmonised database that is well suited for the inter-comparison of
models. The results showed that neural network models generally perform better than the statistical linear model, for predicting both NO2 and PM10 concentrations.

Slini et al. (Slini et al., 2004) employed hourly and daily time series of CO, NO2 and O3, as well as a variety of meteorological variables in various multi-layer perceptron models, in order to provide reliable air quality forecasts, using as a test case the city of Athens, Greece.

Hooyberghs et al. (Hooyberghs et al., 2005) developed a neural network tool to forecast the daily average PM10 concentrations in Belgium one day ahead. This research was based upon measurements from ten monitoring sites during the period 1997–2001 and upon ECMWF (European Centre for Medium-Range Weather Forecasts) simulations of meteorological parameters. The most important input variable found was the boundary layer height. This brought them to the conclusion that day-to-day fluctuations of PM10 concentrations in Belgian urban areas were to a large extent driven by meteorological conditions and to a lesser extent by changes in anthropogenic sources.

Ordieres et al. (Ordieres et al., 2005) analysed and benchmarked a neural network approach to the prediction of average PM2.5 concentrations. Results from three different topologies of neural networks were compared so as to identify their strengths and weaknesses: Multilayer Perceptron (MLP), Radial Basis Function (RBF) and Square Multilayer Perceptron (SMLP). Moreover, two classical models were built (a persistence model and a linear regression), so as to compare their results with the ones provided by the neural network models. The results clearly demonstrated that the neural approach not only outperformed the classical models but also showed fairly similar values among different topologies.

Corani (Corani, 2005) used several statistical approaches to predict ozone and PM 10 in the Milan area. In particular, feed-forward neural networks (FFNNs) were compared with two alternative approaches derived from machine learning: pruned neural networks (PNNs) and lazy learning (LL). PNNs constitute a parameter-parsimonious approach, based on the removal of redundant parameters from fully connected neural networks; LL, on the other hand, is a local linear prediction algorithm, which performs a local learning procedure each time a prediction is required. All the three approaches were tested in the prediction of ozone and PM10. No strong differences were found between the forecast accuracies of the different models; nevertheless, LL provided the best performances on indicators related to average goodness of the prediction.
(correlation, mean absolute error, etc.), while PNNs were superior to the other approaches in detecting of the exceedances of alarm and attention thresholds.

Slini et al. (Slini et al., 2006) examined the PM10 concentrations data sets for the years 1994–2000 collected by a network of monitoring stations operated for the city of Thessaloniki. In order to provide with an operational air quality forecasting module for PM10, statistical methods were investigated and applied. CART and Neural Network (NN) methods were capable of capturing PM10 concentration trends. Methods studied (including linear regression and principal component analysis) demonstrate promising operational forecasting capabilities.

Perez and Reyes (Perez and Reyes, 2006) developed an integrated artificial neural network model to forecast the maxima of 24-hour average of PM10 concentrations 1 day in advance and they applied it to the case of five monitoring stations in the city of Santiago, Chile. Inputs to the model were concentrations measured until 7 PM at the five stations on the present day plus measured and forecast values of meteorological variables. Forecast values using the neural model were compared with the results obtained with a linear model with the same input variables and with persistence. The neural model was more accurate, although the importance of a good choice of input variables was observed.

Grivas and Chaloulakou (Grivas and Chaloulakou, 2006) evaluated the potential of various developed neural network models to provide reliable predictions of PM10 hourly concentrations. The modelling study used the data collected from 4 measurement locations within the Greater Athens during the period of 2001–2002. Artificial neural network models were developed using a combination of meteorological and time-scale input variables. The results of the neural network models were rather satisfactory, with values of the coefficient of determination ($R^2$) for independent test sets ranging between 0.50 and 0.67 for the four sites and values of the index of agreement between 0.80 and 0.89. The performance of examined neural network models was superior in comparison with multiple linear regression models that were developed in parallel ($R^2$ for regression models ranging between 0.29 and 0.35). Their performance was also found adequate in the case of high-concentration events, with acceptable probabilities of detection and low false alarm rates.

Tzima et al. (Tzima et al., 2007) tested various computational intelligence (CI) methods and tools for assessing PM10 concentration values in the Thessaloniki area (Greece), towards knowledge extraction and parameter forecasting. Both goals were achieved, as the resulting evidence supported and expanded existing findings.
concerning the air quality profile of the area, while the achieved prediction was highly accurate. In addition, the need for investigating prevailing weather patterns emerged, as a mean to help in upgrading the air quality profiling of the area.

Karatzas et al. (Karatzas et al., 2008) adopted artificial neural networks (ANNs) for the operational forecasting of the 8-hour running average for Ozone, 24 hours in advance, for two locations in Athens, Greece. Results verified the ability of the methods to analyze and model this knowledge domain and to forecast the levels of key parameters that provide direct input to the environmental decision making process.

Riga et al. (Riga et al., 2009) investigated the potential of data mining algorithms in air pollution modelling and short-term forecasting problems. In this direction, various data mining methods were adopted for the qualitative forecasting of concentration levels of air pollutants or the quantitative prediction of their values (through the development of different classification and regression models respectively) in five locations of the greater Athens area. The quality of experimental results was also assessed, in order to discover the best performing algorithm (or set of algorithms). The results obtained from the thorough statistical analysis applied on all developed models, revealed that in the task of forecasting air quality in Athens: (i) J48, LMT, JRip, Decision Table and REPTree (belonging to the categories of “Tree” and “Rule-based classifiers”) were the best performing classification algorithms, while (ii) M5P, REPTree and M5Rules (belonging to the same categories as the above) were the best performing regression algorithms. On the other hand, prediction models based on SMOreg and Linear Regression (from the “Function classifiers” category) performed statistically worse. Finally, Random Forest was proved to be weak in cases of generalized evaluation with supplied test sets from different monitoring stations.

Karatzas et al. (Karatzas et al., 2009) supported Neural Networks and Principal Component Analysis with a set of mathematical tools including statistical analysis, fast Fourier transformations and periodograms, for the investigation and forecasting of photochemical pollutants time series in Athens, Greece. Results verified the ability of the methods to analyse and model this knowledge domain and to forecast the levels of key parameters that provide direct input to the environmental decision making process.

Tzima et al. (Tzima et al., 2009) adopted ZCS-DM – Zeroth-level Classifier System for Data Mining – for the production of a set of air quality prediction rules in the urban domain. This method was evaluated against methods widely used in the target domain, namely multi-layer perceptron (MLP), support vector machines (SVM), linear
discriminant analysis (LDA) and classification trees (C4.5). Overall, the obtained results revealed the potential of ZCS-DM as a data mining tool to be used for air quality forecasting, and pointed to its insensitivity for missing values (cf. MLP, SVM, LDA) and the understandability of produced models as its greater advantages.

Tzima et al. (Tzima et al., 2011) investigated the potential of various machine learning algorithms as tools for air quality (AQ) episode forecasting and assessed them using an evaluation procedure, tailored to the task at hand. Among the algorithms employed in the experimental phase, the main focus was on ZCS-DM, an evolutionary rule-induction algorithm specifically designed to tackle problems with skewed class distributions. Obtained experimental results revealed the potential of rule-based algorithms for urban AQ forecasting, and pointed towards ZCS-DM as the most suitable algorithm for the target domain, providing the best trade-off between model performance and understandability.

3.2 Chosen method: Artificial Neural Network

Machine learning is a subfield of artificial intelligence (AI) concerned with the development of algorithms and techniques that allow computers to "learn". These algorithms may be useful to address shortcomings in statistical forecast guidance for particulate matter PM10.

Simple linear regression methods provide reasonable skill for example for \( \text{O}_3 \) forecasting, due to the very strong and nearly linear ozone-temperature relationship, but linear regression methods have shown limited skill in forecasting PM10. PM10 (Figure 22 and Figure 23) is not statistically "well-behaved": its distribution is skewed and there is no strong association with any particular weather variable. Methods such as ANN and classification and regression trees (CART) are capable of addressing non-linear problems posed by PM10.

It should be noted that the generating processes and factors influencing the temporal distribution of PM10 are very complex (Van Der Wal and Janssen, 2000): the concentration, composition, and particle size of suspended particulate matter at a given site are determined by a number of factors, such as meteorological properties of the atmosphere, topographical influences, emission sources and particle parameters (Slini et al., 2006). Furthermore, the concentration levels of PM10 are affected by a number of environmental parameters such as wind speed, surface temperature, humidity, surface atmospheric pressure and dew point temperature (Lee et al., 2003, Lu and Fang, 2002, Sánchez et al., 1990).
This chapter presents the work carried out for modelling the background concentration levels of PM10 for the area of Dublin. It also describes an ANN model for local total PM10 concentration levels. For this purpose, time series of daily mean background concentration levels of PM10 (μg m\(^{-3}\)) recorded at different permanent monitoring stations around the city area (Chapter 3.3) and a range of meteorological variables (Chapter 3.4) such as atmospheric stability class, daily maximum, mean and minimum air temperature (°C), dew point at the surface level (°C), daily maximum and mean wind speed (m s\(^{-1}\)) have been employed for the investigation of statistical methods of air quality forecasting. The choice of meteorological variables was based upon selecting those that were generally available and widely accepted as being associated to elevated PM10 levels. The PM10 and weather data is presented in details in Chapters 3.3 and 3.4.
The structure of the forecasting models was based on key meteorological parameters for the first 6 years (2003 – 2009), while their performance was compared with the observed concentration levels of PM$_{10}$ for the year 2010, in order to assess the model forecasting efficacy.

All the weather variables are pre-processed, in order to detect relationships between the former variables and predictions of the PM$_{10}$ concentration levels. The method used for this task is the “classic” linear regression. Linear regression analysis (LRA) is a statistical methodology that utilises the relation between two or more quantitative variables so that one variable can be predicted from the others, according to the Equation 3.1:

$$Y = a + \sum_{i=1}^{p} b_i X_i$$

(Equation 3.1)

where $Y$ is the PM10 concentration, $b_i$ is the $i^{th}$ regression coefficient, $X_i$ is the $i^{th}$ variable, $p$ is the number of regression variables included and $a$ is a constant representing the contribution from non-specified sources.

The choice of an appropriate set of predictors is a central problem for a regression model, because each variable contributes to a different extent to reducing the remaining variation in $Y$, the response variable, after allowance is made for the contributions of other predictor variables that have been included in the model (Slini et al., 2006).

The weather variables which are identified as the best predictors for PM10 background concentration levels for the studied case are inputted to the Artificial Neural Network (ANN) model. ANN makes use of a complex combination of weights and functions to convert input variables into an output (prediction). It can be employed to examine relationships in complex non-linear data sets in the same way as conventional statistical techniques, but without many of the parametric restrictions about the nature of the data relationships. In the current study, the Levenberg-Marquardt Algorithm (Levenberg, 1944, Marquardt, 1963), a type of feed-forward neural network, is adapted and it is presented and described in detail in Chapter 3.5.3.

### 3.3 Background PM10 time series

Figure 24 shows the locations of current monitoring sites (green dots) and sites where assessment reports have been prepared (blue squares) in Dublin.
Of the monitoring stations in the Dublin area shown in Figure 24, the following measure PM10 (Table 2):

<table>
<thead>
<tr>
<th>Station name</th>
<th>Typology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashtown Grove</td>
<td>(not active)</td>
</tr>
<tr>
<td>Ballyfermott</td>
<td>Suburban Traffic</td>
</tr>
<tr>
<td>Blanchardstown</td>
<td>Suburban Traffic</td>
</tr>
<tr>
<td>Clontarf</td>
<td>(not active)</td>
</tr>
<tr>
<td>Coleraine Street</td>
<td>Urban Traffic</td>
</tr>
<tr>
<td>College Green</td>
<td>Urban Traffic</td>
</tr>
<tr>
<td>Dun Laoghaire</td>
<td>Suburban Background</td>
</tr>
<tr>
<td>Knocklyon</td>
<td>(not active)</td>
</tr>
<tr>
<td>Marino</td>
<td>(not active)</td>
</tr>
<tr>
<td>Phoenix Park</td>
<td>Suburban Background</td>
</tr>
<tr>
<td>Rathmines</td>
<td>Urban Background</td>
</tr>
<tr>
<td>Ringsend</td>
<td>Urban Traffic</td>
</tr>
<tr>
<td>Tallaght</td>
<td>Suburban Background</td>
</tr>
<tr>
<td>Winetavern Street</td>
<td>Urban Traffic</td>
</tr>
<tr>
<td>Wood Quay</td>
<td>(not active)</td>
</tr>
</tbody>
</table>

Table 2: PM10 monitoring station in the Dublin area
For the purpose of modelling PM10 background concentration levels, the measurements of the following monitoring stations are considered as suitable:

- Dun Laoghaire
- Phoenix Park
- Rathmines
- Tallaght

These stations provide daily averages for PM10 concentration levels. The availability, quality and extent of the datasets are different from station to station.

The data is available for downloading at:
http://erc.epa.ie/safer/iso19115/displayAllAttachments.jsp?isoID=69

3.3.1 Dun Laoghaire - Suburban Background

The PM10 monitoring station in Dun Laoghaire is located in proximity of Lower Glenageary Road (Figure 25).

![Figure 25: PM10 monitoring station, Dun Laoghaire](image)

It has been active since the 1st of January 2008 and the data is available from that date until the 31st of December 2010. As it can be seen from the time series for this station presented in Figure 26, there are minor gaps in the datasets.
3.3.2 Phoenix Park - Suburban Background

The PM10 monitoring station in Phoenix Park is located in proximity of Ordnance Road (Figure 27).

![Figure 27: PM10 monitoring station, Phoenix Park](image)

It has been active since the 1st of January 1996 and the data is available from that date until the 31st of December 2010. As it can be seen from the time series for this station presented in Figure 28, there is a major gap between the 1st of January 1999 and the 31st of December 2000 along with other minor gaps.
3.3.3 Rathmines - Urban Background

The PM10 monitoring station in Rathmines is located in proximity of Wynnefield Road (Figure 29).

Figure 29: PM10 monitoring station, Rathmines

It has been active since the 1st of January 1996 and the data is available from that date until the 31st of December 2010. As it can be seen from the time series for this station presented in Figure 30, there are three major gaps between the 1st of January

![PM10 - Rathmines](image)

Figure 30: PM10 time series for Rathmines

### 3.3.4 Tallaght - Suburban Background

The PM10 monitoring station in Tallaght is located in proximity of Old Bawn Road (Figure 31). It has been active since the 1st of January 1996 and the data is available from that date until the 31st of December 2010.

![PM10 monitoring station, Tallaght](image)
As it can be seen from the time series for this station presented in Figure 32, there are two major gaps between the 14th of July 2008 and the 11th of September 2008 and between the 7th of May 2009 and the 2nd of June 2009 along with other minor gaps.

![Figure 32: PM10 time series for Tallaght](image)

3.4 Weather data

Weather data was obtained from Met Eireann (http://www.met.ie/climate/climate-data-information.asp) for the following three synoptic stations:

- Casement;
- Dublin Airport;
- Phoenix Park.

Casement weather station is located in proximity of Casement airport (Figure 33) on the west side of the city. The information available for this station is hourly data for: wind speed, wind direction, temperature, humidity, dew point, atmospheric pressure and rainfall. Solar radiation is not available so the dataset is integrated with the solar radiation measured at the Dublin Airport weather station.
Dublin Airport weather station is located in proximity of Dublin airport (Figure 34) on the North side of the city. The information available for this station is hourly data for: wind speed, wind direction, temperature, humidity, dew point, atmospheric pressure, rainfall and solar radiation.

Phoenix Park weather station is located in Phoenix Park in proximity of Ordnance Road (Figure 35) close to the centre of the city. The information available for this
station is hourly data for: atmospheric stability classes, temperature, humidity, dew point, atmospheric pressure, rainfall and solar radiation. Wind speed and direction are not available for this station, so two different datasets are created by integrating the Phoenix Park dataset with wind data from Casement and Dublin Airport weather stations.

![Figure 35: Phoenix Park weather station](image)

### 3.5 Modelling procedure

#### 3.5.1 Data pre-processing

Data quality is one of the fundamental issues influencing the performance of any data investigation algorithm and the quality of its results (Kyriakidis et al., 2009). In the investigation chain, the data selection phase is followed by the pre-processing phase, which results in increased data quality. Pyle (Pyle, 1999) illustrates all the pre-processing steps that lead to this result, which can be grouped in the following actions:

1. Identifying and removing of the outliers;
2. Data smoothing;
3. Handling of missing data;
4. Data de-trending.

These pre-processing steps are applied to the background PM10 concentration time-series measured at the monitoring stations listed in the previous sections. After each
step, the correlation with the weather parameters data is tested with backward linear regression; this operation is carried out for each weather monitoring station (Table 3).

<table>
<thead>
<tr>
<th>PM10 station</th>
<th>Weather station</th>
<th>Anemometer</th>
<th>Atmospheric Stability Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dun Laoghaire</td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td>Phoenix Park</td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td>Rathmines</td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td>Tallaght</td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>Phoenix Park</td>
</tr>
</tbody>
</table>

Table 3: Datasets correlations summary

As mentioned above, the Phoenix Park dataset doesn’t include wind data and therefore it is taken into account twice, by being combined with both the Dublin Airport and the Casement anemometers data. The daily average value of the Atmospheric Stability class is obtained from Phoenix Park monitoring station and is combined with all the weather datasets. The following weather parameters are taken into account: wind speed, wind direction, temperature, humidity, dew point, atmospheric pressure, rainfall and solar radiation. For each of these parameters, the daily average, maximum and minimum values are calculated.

3.5.1.1 Identifying and removing of the outliers

This step is necessary to avoid smoothing functions and the replacement of missing values functions being affected by outliers. In statistics, an outlier (Barnett and Lewis, 1994) is an observation that is numerically distant from the rest of the data. Grubbs (Grubbs, 1969) defined an outlier as:

“An outlying observation, or outlier, is one that appears to deviate markedly from other members of the sample in which it occurs”.

Outliers result in large errors and consequently large weight updates because of the large deviation from the norm. They can occur by chance in any distribution, but they
are often indicative either of measurement error or that the population has a heavy-tailed distribution. The correct identification of the outliers is very important because their presence could affect the performance/results of the model and lead to inaccurate prediction (Engelbrecht, 2007), but it requires a good knowledge of the scientific domain they represent (here air pollution) and the characteristics of the geographic area of interest (here Dublin).

For the purpose of this work, a common method to quantitatively identifying outliers in the PM10 time-series is evaluated, which involves the deletion of values located more than a certain number (two) of standard deviations away from the mean (Figure 36).

![Figure 36: Dun Laoghaire PM10 plot before and after removing the outliers](image)

The chosen method for the outliers’ identification is quite crude but, as it can be seen in the R scores reported in Table 4: R scores before and after the elimination of outliers below, it increased the correlation between the PM10 and the weather dataset for all stations (Table 4). The R-value are calculated by comparing PM10 time-series with the complete set of weather variables.

The quantity R, called the linear correlation coefficient, measures the strength and the direction of a linear relationship between two variables. The value of R ranges between -1 and +1, where the + and – signs are used for positive linear correlations and negative linear correlations. A correlation greater than 0.8 is generally described as strong, whereas a correlation less than 0.5 is generally described as weak, although these values can vary based upon the "type" of data being examined and the field of study.
<table>
<thead>
<tr>
<th>PM10 station</th>
<th>Weather station</th>
<th>Anemometer</th>
<th>$R$ unmodified</th>
<th>$R$ no outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dun Laoghaire</td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>0.571</td>
<td>0.609</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>0.567</td>
<td>0.604</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Dublin Airport</td>
<td>0.557</td>
<td>0.596</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>0.561</td>
<td>0.600</td>
</tr>
<tr>
<td>Phoenix Park</td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>0.694</td>
<td>0.707</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>0.694</td>
<td>0.706</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Dublin Airport</td>
<td>0.540</td>
<td>0.611</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>0.549</td>
<td>0.620</td>
</tr>
<tr>
<td>Rathmines</td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>0.720</td>
<td>0.739</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>0.720</td>
<td>0.736</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Dublin Airport</td>
<td>0.691</td>
<td>0.703</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>0.669</td>
<td>0.706</td>
</tr>
<tr>
<td>Tallaght</td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>0.617</td>
<td>0.641</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>0.616</td>
<td>0.640</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Dublin Airport</td>
<td>0.603</td>
<td>0.638</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>0.610</td>
<td>0.643</td>
</tr>
</tbody>
</table>

Table 4: R scores before and after the elimination of outliers

Even if this pre-processing step improves the correlation between the PM10 and weather datasets for all the testing cases, it is believed that such a crude approach might delete essential information from the PM10 time series, such as high concentration levels due to particular combination of weather variables.

3.5.1.2 Data smoothing

In statistics, to smooth a data set is to create an approximating function that attempts to capture important patterns and real trends in the data, while eliminating noise or other fine-scale structures/rapid phenomena. If noise is not removed, replacing missing values functions may be affected (Kyriakidis et al., 2009). Smoothing may also deal with missing values, if they do not represent a large percentage of the dataset. In some cases, seasonal variation are so strong that do not allow for any trend or periodicity indications: smoothing can remove seasonality and may help in revealing lagged fluctuations within the data set (Xiong et al., 2006). Many different algorithms are used in smoothing. One of the most common smoothing algorithms is the "moving average". Moving Average is a type of finite impulse response filter used to analyse a set of data points by creating a series of averages of different subsets of the full data set. Given a series of numbers and a fixed subset size, the moving average can be obtained by first taking the average of the first subset. The fixed subset size is then shifted forward, creating a new subset of numbers, which is averaged. This process is repeated over the entire PM10 data series.
For the purpose of this chapter, two different smoothing algorithms were applied and compared:

a. **Simple moving average:** this is the unweighted mean of an equal number of data either side of a central value. This ensures that variations in the mean are aligned with the variations in the data rather than being shifted in time. The selected period for this task is 7 days, in order to reproduce the weekly variability of air pollution.

b. **Exponential moving average:** An exponential weighted moving average is a type of infinite impulse response filter that applies weighting factors which decrease exponentially. The weighting for each older data point decreases exponentially, never reaching zero.

![Figure 37 : Tallaght PM10 measurements after simple moving average smoothing](http://en.wikipedia.org/wiki/Moving_average)

![Figure 38 : Tallaght PM10 measurements after exponential moving average smoothing](http://en.wikipedia.org/wiki/Moving_average)
The plots in Figure 37 and Figure 38 represent the graphical outcomes for two algorithms described below, when applied to the PM10 background concentration levels for the Tallaght monitoring station.

The R scores resulting from the two different algorithms are reported in the table below (Table 5). It should be noted that the correlation between the PM10 and the weather dataset increases after applying the simple moving average smoothing to the PM10 time-series for all the stations except Rathmines. The exponential moving average smoothing has on the other hand a negative effect on the correlation, resulting in lower R scores. This may happen because it uses PM10 data points which were a long way away from the time being considered.

<table>
<thead>
<tr>
<th>PM10 station</th>
<th>Weather station</th>
<th>Anemometer</th>
<th>R unmodified</th>
<th>R simple smoothing</th>
<th>R expon. smoothing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dun Laoghaire</strong></td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>0.571</td>
<td>0.653</td>
<td>0.454</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>0.567</td>
<td>0.654</td>
<td>0.454</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Dublin Airport</td>
<td>0.557</td>
<td>0.633</td>
<td>0.448</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>0.561</td>
<td>0.640</td>
<td>0.454</td>
</tr>
<tr>
<td><strong>Phoenix Park</strong></td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>0.694</td>
<td>0.718</td>
<td>0.565</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>0.694</td>
<td>0.720</td>
<td>0.558</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Dublin Airport</td>
<td>0.540</td>
<td>0.649</td>
<td>0.449</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>0.549</td>
<td>0.646</td>
<td>0.445</td>
</tr>
<tr>
<td><strong>Rathmines</strong></td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>0.720</td>
<td>0.738</td>
<td>0.630</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>0.720</td>
<td>0.738</td>
<td>0.622</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Dublin Airport</td>
<td>0.691</td>
<td>0.704</td>
<td>0.573</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>0.669</td>
<td>0.696</td>
<td>0.563</td>
</tr>
<tr>
<td><strong>Tallaght</strong></td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>0.617</td>
<td>0.662</td>
<td>0.493</td>
</tr>
<tr>
<td></td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>0.616</td>
<td>0.672</td>
<td>0.496</td>
</tr>
<tr>
<td></td>
<td>Dublin Airport</td>
<td>Dublin Airport</td>
<td>0.603</td>
<td>0.674</td>
<td>0.487</td>
</tr>
<tr>
<td></td>
<td>Casement</td>
<td>Casement</td>
<td>0.610</td>
<td>0.671</td>
<td>0.486</td>
</tr>
</tbody>
</table>

Table 5: R scores before and after the moving average smoothing
3.5.1.3 Handling of missing data

It is common in air quality data to have missing values, because of a failure of the monitoring station, or other exogenous reasons. The PM10 time-series recorded at Dun Laoghaire, Phoenix Park, Rathmines and Tallaght monitoring stations have gaps of various extents, as presented in Chapter 3.3. It is important to deal with this problem of the datasets because some algorithms that may be used in the data mining procedure require that the missing values should be replaced or removed: most data processing methods require for complete data metrics before performing any further analysis (Bishop, 1996).

There are several interpolation methods that deal with missing values (Linear, Piecewise Cubic Hermite, Cubic Spline, Nearest Neighbour, etc.). Kyriadis et al. (Kyriakidis et al., 2009) performed a small scale experiment in order to investigate the performance of those methods in air quality data. A time series with no missing values was used, comprising of hourly values. In this series, 8.5% of the data in the first step, and 12.5% of the data in a next step were removed to imitate a missing values case. The missing values were imputed to the data set, on the basis of the missing values of an air quality time series. The results of the study suggested that linear interpolation methods have the best performance, in terms of minimum error and higher correlation. On this basis, the linear interpolation method was selected and it was applied to replace the missing values in the studied PM10 data sets.

Even though this pre-processing step didn’t result in any improvements in the linear correlation coefficient R values, it was necessary to perform it in order to prepare the datasets for the Artificial Neural Network model.

3.5.1.4 De-trending

Trend in a time series is a slow, gradual change in some property of the data over the whole time interval under investigation. Trend is sometimes loosely defined as a long term change in the mean, but can also refer to a change in other statistical properties. De-trending is the mathematical operation of removing trends from the data under investigation. De-trending is often applied to remove a feature thought to distort or obscure the relationships of interest. In order to de-trend, constant and straight like de-trending methods may be employed. Constant de-trending removes the mean of the data to create zero mean data. Straight line de-trending finds linear trends (in the least-squared sense) and the removes them (Kyriakidis et al., 2009).
Figure 39: Dun Laoghaire PM10 trend-line

Figure 40: Phoenix Park PM10 trend-line

Figure 41: Rathmines PM10 trend-line

Figure 42: Tallaght PM10 trend-line
Figure 39 to Figure 42 illustrate the trend in the four PM10 monitoring stations time series considered for this analysis. The trends are all characterised by a decreasing behaviour.

Straight-line de-trending was applied to the 4 PM10 time-series and the resultants were then linearly regressed against the set of weather parameters used in the previous analysis. The result of such procedure was a general marked decrease in the linear correlation coefficients $R$ for all four PM10 monitoring stations. As such the de-trending step was omitted from the pre-processing procedure for these datasets.

### 3.5.2 Sensitivity analysis

Mesin et al. (Mesin et al., 2010) demonstrated that local air quality forecasting can be made on the basis of meteorological and air pollution time series. The data used in their work contained redundant information, thus Mesin et al. used a partial mutual information criterion to select the regressors which carried the greatest non redundant information to be used in building a prediction model. An application was also shown regarding the forecast of PM10 concentration one day in advance, by employing the selected regressors within an artificial neural network. Table 6 shows the best R scores obtained from the correlation between the pre-processed dataset of each PM10 monitoring station and the weather data.

<table>
<thead>
<tr>
<th>PM10 station</th>
<th>Weather station</th>
<th>Anemometer</th>
<th>$R - 7d$ smooth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dun Laoghaire</td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>0.720</td>
</tr>
<tr>
<td>Phoenix Park</td>
<td>Phoenix Park</td>
<td>Casement</td>
<td>0.722</td>
</tr>
<tr>
<td>Rathmines</td>
<td>Phoenix Park</td>
<td>Dublin Airport</td>
<td>0.740</td>
</tr>
<tr>
<td>Tallaght</td>
<td>Dublin Airport</td>
<td>Dublin Airport</td>
<td>0.676</td>
</tr>
</tbody>
</table>

Table 6: Best $R$ score per PM10 monitoring station

#### 3.5.2.1 Dun Laoghaire

The initial model included the following 26 variables: year, month, day, average atmospheric stability class, wind speed (average, minimum and maximum), wind direction (average), temperature (average, minimum and maximum), humidity (average, minimum and maximum), dew point (average, minimum and maximum), atmospheric pressure (average, minimum and maximum), rainfall (average and maximum) and solar radiation (average and maximum).

Fourteen different regression models were generated by excluding progressively the single variable with least impact on the global linear correlation coefficient value: the dependent variable PM10 was regressed against the complete group of the above $n (= 26)$ predictor variables and the $R$ score was calculated, then the predictor variable
with the lowest correlation was eliminated and the procedure repeated for (n-1) predictor variables. This culling procedure was repeated until all the predictor variables with minor correlation with PM10 were eliminated. The backward linear regression calculation and culling procedure were performed using IBM SPSS Statistics software package.

The final model consisted of the following predictor variables: (Constant), year, month, average atmospheric stability class, maximum and average humidity, average cosine of the wind direction, average wind speed, average temperature, maximum rainfall, average and maximum dew point, average solar radiation and minimum atmospheric pressure. As it could be noted from the R scores in Table 7, the final model with 13 predictor variables has similar performance to the initial model. The aim of the culling procedure was to obtain a model with similar performances to the initial one, but with less variables and thus reduced calculation times. These variables appear to be the best predictors for PM10 background concentration levels for Dun Laoghaire.

<table>
<thead>
<tr>
<th>Number of variables</th>
<th>R</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>.720</td>
<td>4.02</td>
</tr>
<tr>
<td>13</td>
<td>.718</td>
<td>4.01</td>
</tr>
</tbody>
</table>

Table 7: R scores for the initial and final models for Dun Laoghaire

The ANOVA table for the two models is presented below.

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Regression</td>
<td>16770</td>
<td>26</td>
<td>69645</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>15560.235</td>
<td>962</td>
<td>16.175</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>32330.372</td>
<td>988</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Regression</td>
<td>16659.395</td>
<td>13</td>
<td>1281.492</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>15670.977</td>
<td>975</td>
<td>16.073</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>32330.372</td>
<td>988</td>
<td></td>
</tr>
</tbody>
</table>

Table 8: ANOVA table for Dun Laoghaire

Table 9 shows the correlation coefficients for each included variable. The cosine coefficient (wind direction) is positive and it has a high value, which means that positive important contributions to PM10 values are due to positive value of cosine, i.e. for wind blowing from the sea, pointing to cross boundaries contributions. The maximum rainfall variable has a negative contribution to PM10 levels, as expected. The month parameter is correlated to the measured PM10 levels for this site, which points to a seasonal behaviour in the pollution level.
<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>526.634</td>
<td>334.233</td>
<td>1.576</td>
<td>.115</td>
</tr>
<tr>
<td>year</td>
<td>-.348</td>
<td>.166</td>
<td>-.049</td>
<td>-2.095</td>
</tr>
<tr>
<td>month</td>
<td>-.202</td>
<td>.047</td>
<td>-.120</td>
<td>-4.340</td>
</tr>
<tr>
<td>avg stab</td>
<td>1.057</td>
<td>.543</td>
<td>.048</td>
<td>1.945</td>
</tr>
<tr>
<td>avg cos</td>
<td>3.257</td>
<td>.231</td>
<td>.345</td>
<td>14.108</td>
</tr>
<tr>
<td>avg speed</td>
<td>.198</td>
<td>.080</td>
<td>.081</td>
<td>2.481</td>
</tr>
<tr>
<td>avg temp</td>
<td>2.796</td>
<td>.913</td>
<td>2.552</td>
<td>3.061</td>
</tr>
<tr>
<td>avg hum</td>
<td>.749</td>
<td>.185</td>
<td>.949</td>
<td>4.047</td>
</tr>
<tr>
<td>max hum</td>
<td>-.174</td>
<td>.050</td>
<td>-.115</td>
<td>-3.453</td>
</tr>
<tr>
<td>max dew</td>
<td>-3.104</td>
<td>.922</td>
<td>-2.621</td>
<td>-3.366</td>
</tr>
<tr>
<td>avg dew</td>
<td>-.256</td>
<td>.134</td>
<td>-.213</td>
<td>-1.905</td>
</tr>
<tr>
<td>min press</td>
<td>.134</td>
<td>.012</td>
<td>.311</td>
<td>11.551</td>
</tr>
<tr>
<td>max rain</td>
<td>-.270</td>
<td>.087</td>
<td>-.072</td>
<td>-3.114</td>
</tr>
<tr>
<td>avg solar</td>
<td>-.024</td>
<td>.007</td>
<td>-.123</td>
<td>-3.664</td>
</tr>
</tbody>
</table>

Table 9: Coefficients values for Dun Laoghaire

In Table 9, the presented coefficients have the following meanings:

- **B**: regression coefficient used to compute the regression equation;
- **Std. Error**: a measure of how much the value of a test statistic varies from sample to sample. It is the standard deviation of the sampling distribution of the statistic;
- **Beta coefficient**: it's the regression coefficient when all variables are expressed in standardised (z-score) form. Transforming the independent variables to standardised form makes the coefficients more comparable since they are all in the same units of measure;
- **t**: statistic used to test the null hypothesis that there is no linear relationship between a dependent variable and an independent variable or, in other words, that a regression coefficient is equal to 0. When the significance level is small (less than 0.10) the coefficient is considered significant.

Table 10 presents the statistics for the excluded predictor variables.
<table>
<thead>
<tr>
<th>Excluded Variables</th>
<th>Beta In</th>
<th>t</th>
<th>Sig.</th>
<th>Partial Correlation</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg press</td>
<td>.029</td>
<td>.269</td>
<td>.788</td>
<td>.009</td>
<td>.043</td>
</tr>
<tr>
<td>min temp</td>
<td>-.093</td>
<td>-.878</td>
<td>.380</td>
<td>-.028</td>
<td>.044</td>
</tr>
<tr>
<td>min speed</td>
<td>.015</td>
<td>.327</td>
<td>.744</td>
<td>.010</td>
<td>.235</td>
</tr>
<tr>
<td>max press</td>
<td>.033</td>
<td>.543</td>
<td>.587</td>
<td>.017</td>
<td>.133</td>
</tr>
<tr>
<td>max solar</td>
<td>-.042</td>
<td>-.520</td>
<td>.603</td>
<td>-.017</td>
<td>.074</td>
</tr>
<tr>
<td>max speed</td>
<td>-.029</td>
<td>-.469</td>
<td>.640</td>
<td>-.015</td>
<td>.129</td>
</tr>
<tr>
<td>day</td>
<td>.008</td>
<td>.356</td>
<td>.722</td>
<td>.011</td>
<td>.986</td>
</tr>
<tr>
<td>avg rain</td>
<td>.038</td>
<td>.832</td>
<td>.406</td>
<td>.027</td>
<td>.232</td>
</tr>
<tr>
<td>avg dir</td>
<td>-.015</td>
<td>-.601</td>
<td>.548</td>
<td>-.019</td>
<td>.833</td>
</tr>
<tr>
<td>min hum</td>
<td>.029</td>
<td>.495</td>
<td>.621</td>
<td>.016</td>
<td>.144</td>
</tr>
<tr>
<td>max temp</td>
<td>.066</td>
<td>.527</td>
<td>.598</td>
<td>.017</td>
<td>.031</td>
</tr>
<tr>
<td>min dew</td>
<td>-.130</td>
<td>-.999</td>
<td>.318</td>
<td>-.032</td>
<td>.030</td>
</tr>
<tr>
<td>avg sin</td>
<td>.038</td>
<td>1.432</td>
<td>.153</td>
<td>.046</td>
<td>.691</td>
</tr>
</tbody>
</table>

Table 10: Statistics for excluded predictor variables for Dun Laoghaire

In Table 10, the presented coefficients have the following meanings:

- **Beta In**: the standardised regression coefficient that would result if the variable were entered into the equation at the next step;
- **Sig.**: the conditional probability that a relationship as strong as the one observed in the data would be present, if the null hypothesis were true. It is often called the p-value. Typically a value of less than 0.05 is considered significant;
- **Partial Correlation**: the correlation that remains between two variables after removing the correlation that is due to their mutual association with the other variables. The correlation between the dependent variable and an independent variable when the linear effects of the other independent variables in the model have been removed from both;
- **Tolerance**: a statistic used to determine how much the independent variables are linearly related to one another (multicollinear). The proportion of the variable's variance not accounted for by other independent variables in the equation. A variable with very low tolerance contributes little information to a model, and can cause computational problems. It is calculated as 1 minus R squared for an independent variable when it is predicted by the other independent variables already included in the analysis.

The following variables are the best predictors for the PM10 background concentration levels in Dun Laoghaire: time, wind speed and direction, average
atmospheric stability class, average and maximum humidity, average temperature, average and maximum dew point, minimum atmospheric pressure, maximum rainfall, average global radiation. The global R score for the model suggests that the PM10 levels for this monitoring site might be affected by other variables that have not been considered, e.g. the traffic in the surrounding area.

### 3.5.2.2 Phoenix Park

The initial model included the same variables initially included in the Dun Laoghaire model and the same culling procedure adopted for that dataset was used here.

The final model consisted of the following predictor variables: (Constant), year, month, average atmospheric stability class, humidity, average cosine and sine of the wind direction, minimum wind speed, average rainfall, maximum and average dew point, average solar radiation, maximum and average temperature and minimum atmospheric pressure. As it could be noted from the R scores in Table 11, the final model with 16 predictor variables has similar performance to the initial model. These variables appear to be the best predictors for PM10 background concentration levels for Phoenix Park.

<table>
<thead>
<tr>
<th>Number of variables</th>
<th>R</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>.722</td>
<td>3.258</td>
</tr>
<tr>
<td>16</td>
<td>.720</td>
<td>3.253</td>
</tr>
</tbody>
</table>

Table 11: R scores for the initial and final models for Phoenix Park

The ANOVA table for the two models is presented below.

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Regression</td>
<td>12947.454</td>
<td>26</td>
<td>497.979</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>11888.586</td>
<td>1120</td>
<td>10.615</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>24836.040</td>
<td>1146</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Regression</td>
<td>12875.258</td>
<td>16</td>
<td>804.704</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>11960.782</td>
<td>1130</td>
<td>10.585</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>24836.040</td>
<td>1146</td>
<td></td>
</tr>
</tbody>
</table>

Table 12: ANOVA table for Phoenix Park

Table 13 shows the value for each coefficient. The sine and cosine parameters are positive, which outlines a contribution to PM10 for wind coming mainly from the N-E sector so from the traffic on Chesterfield Avenue and from the D15 area. The atmospheric stability appears to give an important contribution to the pollution levels, with the temperature, atmospheric pressure and solar radiation parameters. The
The month parameter is correlated with the PM10 levels, which outlines a seasonal behaviour of the pollution levels.

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>409.753</td>
<td>213.357</td>
<td>1.920</td>
<td>.055</td>
</tr>
<tr>
<td>year</td>
<td>-.286</td>
<td>.106</td>
<td>-.057</td>
<td>-2.688</td>
</tr>
<tr>
<td>month</td>
<td>-.084</td>
<td>.035</td>
<td>-.061</td>
<td>-2.377</td>
</tr>
<tr>
<td>avg stab</td>
<td>1.062</td>
<td>.400</td>
<td>.061</td>
<td>2.656</td>
</tr>
<tr>
<td>avg cos</td>
<td>3.501</td>
<td>.182</td>
<td>.461</td>
<td>19.192</td>
</tr>
<tr>
<td>avg sin</td>
<td>.500</td>
<td>.226</td>
<td>.052</td>
<td>2.214</td>
</tr>
<tr>
<td>min speed</td>
<td>.227</td>
<td>.069</td>
<td>.093</td>
<td>3.275</td>
</tr>
<tr>
<td>avg temp</td>
<td>2.567</td>
<td>.727</td>
<td>2.781</td>
<td>3.532</td>
</tr>
<tr>
<td>max temp</td>
<td>.595</td>
<td>.106</td>
<td>.696</td>
<td>5.618</td>
</tr>
<tr>
<td>avg hum</td>
<td>.626</td>
<td>.144</td>
<td>.979</td>
<td>4.346</td>
</tr>
<tr>
<td>max hum</td>
<td>-.090</td>
<td>.041</td>
<td>-.074</td>
<td>-2.170</td>
</tr>
<tr>
<td>min hum</td>
<td>.071</td>
<td>.022</td>
<td>.194</td>
<td>3.242</td>
</tr>
<tr>
<td>avg dew</td>
<td>-3.311</td>
<td>.718</td>
<td>-3.340</td>
<td>-4.611</td>
</tr>
<tr>
<td>max dew</td>
<td>-.201</td>
<td>.106</td>
<td>-.199</td>
<td>-1.893</td>
</tr>
<tr>
<td>min press</td>
<td>.115</td>
<td>.008</td>
<td>.331</td>
<td>14.129</td>
</tr>
<tr>
<td>avg rain</td>
<td>-.915</td>
<td>.475</td>
<td>-.042</td>
<td>-1.925</td>
</tr>
<tr>
<td>avg solar</td>
<td>-.020</td>
<td>.005</td>
<td>-.119</td>
<td>-3.871</td>
</tr>
</tbody>
</table>

Table 13: Coefficients values for Phoenix Park

<table>
<thead>
<tr>
<th>Model</th>
<th>Beta In</th>
<th>t</th>
<th>Sig.</th>
<th>Partial Correlation</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tolerance</td>
</tr>
<tr>
<td>min dew</td>
<td>-.094</td>
<td>-.805</td>
<td>.421</td>
<td>-.024</td>
<td>.031</td>
</tr>
<tr>
<td>max rain</td>
<td>-.013</td>
<td>-.285</td>
<td>.776</td>
<td>-.008</td>
<td>.212</td>
</tr>
<tr>
<td>max glo</td>
<td>.028</td>
<td>.759</td>
<td>.448</td>
<td>.023</td>
<td>.313</td>
</tr>
<tr>
<td>max speed</td>
<td>-.054</td>
<td>-.741</td>
<td>.459</td>
<td>-.022</td>
<td>.080</td>
</tr>
<tr>
<td>avg speed</td>
<td>.037</td>
<td>.360</td>
<td>.719</td>
<td>.011</td>
<td>.041</td>
</tr>
<tr>
<td>day</td>
<td>.047</td>
<td>.832</td>
<td>.405</td>
<td>.025</td>
<td>.135</td>
</tr>
<tr>
<td>min temp</td>
<td>.012</td>
<td>.590</td>
<td>.555</td>
<td>.018</td>
<td>.987</td>
</tr>
<tr>
<td>avg press</td>
<td>-.124</td>
<td>-1.104</td>
<td>.270</td>
<td>-.033</td>
<td>.034</td>
</tr>
<tr>
<td>max press</td>
<td>.056</td>
<td>1.081</td>
<td>.280</td>
<td>.032</td>
<td>.161</td>
</tr>
<tr>
<td>avg dir</td>
<td>.061</td>
<td>1.597</td>
<td>.111</td>
<td>.047</td>
<td>.290</td>
</tr>
</tbody>
</table>

Table 14 presents the statistics for the excluded predictor variables.
The following variables are the best predictors for the PM10 background concentration levels in Phoenix Park: time, average atmospheric stability class, humidity, average cosine and sine of the wind direction, minimum wind speed, average rainfall, maximum and average dew point, average solar radiation, maximum and average temperature and minimum atmospheric pressure. As at Dun Laoghaire, the global R score for the model suggests that the PM10 levels for this monitoring site might be affected by other variables, i.e. the traffic in the surrounding area.

3.5.2.3 Rathmines

The initial model included the same variables initially included in the Dun Laoghaire model and the same culling procedure adopted for that dataset was used here. The final model consisted of the following predictor variables: (Constant), year, month, maximum solar radiation, humidity, average atmospheric stability class, average cosine of the wind direction, average rainfall, minimum atmospheric pressure, maximum and average temperature and average dew point. As it could be noted from the R scores in Table 15, the final model with 14 predictor variables has similar performance to the initial model. These variables appear to be the best predictors for PM10 background concentration levels for Rathmines.

<table>
<thead>
<tr>
<th>Model</th>
<th>Beta In</th>
<th>t</th>
<th>Sig.</th>
<th>Partial Correlation</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>min dew</td>
<td>-.094</td>
<td>-.805</td>
<td>.421</td>
<td>-.024</td>
<td>.031</td>
</tr>
<tr>
<td>max rain</td>
<td>-.013</td>
<td>-.285</td>
<td>.776</td>
<td>-.008</td>
<td>.212</td>
</tr>
<tr>
<td>max glo</td>
<td>.028</td>
<td>.759</td>
<td>.448</td>
<td>.023</td>
<td>.313</td>
</tr>
<tr>
<td>max speed</td>
<td>-.054</td>
<td>-.741</td>
<td>.459</td>
<td>-.022</td>
<td>.080</td>
</tr>
<tr>
<td>avg speed</td>
<td>.037</td>
<td>.360</td>
<td>.719</td>
<td>.011</td>
<td>.041</td>
</tr>
<tr>
<td>day</td>
<td>.047</td>
<td>.832</td>
<td>.405</td>
<td>.025</td>
<td>.135</td>
</tr>
<tr>
<td>min temp</td>
<td>.012</td>
<td>.590</td>
<td>.555</td>
<td>.018</td>
<td>.987</td>
</tr>
<tr>
<td>avg press</td>
<td>-.124</td>
<td>-1.104</td>
<td>.270</td>
<td>-.033</td>
<td>.034</td>
</tr>
<tr>
<td>max press</td>
<td>.056</td>
<td>1.081</td>
<td>.280</td>
<td>.032</td>
<td>.161</td>
</tr>
<tr>
<td>avg dir</td>
<td>.061</td>
<td>1.597</td>
<td>.111</td>
<td>.047</td>
<td>.290</td>
</tr>
</tbody>
</table>

Table 14: Statistics for excluded predictor variables for Phoenix Park

<table>
<thead>
<tr>
<th>Number of variables</th>
<th>R</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>.740</td>
<td>4.833</td>
</tr>
<tr>
<td>14</td>
<td>.739</td>
<td>4.819</td>
</tr>
</tbody>
</table>

Table 15: R scores for the initial and final models for Rathmines

The ANOVA table for the two models is presented below.
Table 16: ANOVA table for Rathmines

Table 17 shows the value for each coefficient. Again the cosine parameter of the wind direction is positive, which outlines a contribution to the PM10 levels from the North where the city centre is located. The atmospheric stability seems to have an evident impact also here. Temporal parameters such as year and month have a significant correlation, which may suggest that PM10 background concentration levels may have a seasonal component.

Table 17: Coefficients values for Rathmines
Table 18 presents the statistics for the excluded predictor variables.

The following variables are the best predictors for the PM10 background concentration levels in Rathmines: time, maximum solar radiation, humidity, average atmospheric stability class, average cosine of the wind direction, average rainfall, minimum atmospheric pressure, maximum and average temperature and average dew point.

3.5.2.4 Tallaght

The initial model included the same variables initially included in the Dun Laoghaire model and the same culling procedure adopted for that dataset was used here.
The final model consisted of the following predictor variables: (Constant), year, average atmospheric stability class, maximum solar radiation, average cosine of the wind direction, atmospheric pressure, maximum and average humidity, minimum and average dew point, average rainfall, maximum and average temperature. As it could be noted from the R scores in Table 19, the final model with 15 predictor variables has similar performance to the initial model. These variables appear to be the best predictors for PM10 background concentration levels for Tallaght.

<table>
<thead>
<tr>
<th>Number of variables</th>
<th>R</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>.676</td>
<td>4.297</td>
</tr>
<tr>
<td>15</td>
<td>.672</td>
<td>4.296</td>
</tr>
</tbody>
</table>

Table 19: R scores for the initial and final models for Tallaght

The ANOVA table for the two models is presented below.

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression 1</td>
<td>16406.738</td>
<td>26</td>
<td>631.028</td>
<td>34.172</td>
</tr>
<tr>
<td>Residual 1</td>
<td>19444.938</td>
<td>1053</td>
<td>18.466</td>
<td></td>
</tr>
<tr>
<td>Total 1</td>
<td>35851.676</td>
<td>1079</td>
<td>18.458</td>
<td></td>
</tr>
<tr>
<td>Regression 2</td>
<td>16212.516</td>
<td>15</td>
<td>1080.834</td>
<td>58.557</td>
</tr>
<tr>
<td>Residual 2</td>
<td>19639.161</td>
<td>1064</td>
<td>18.458</td>
<td></td>
</tr>
<tr>
<td>Total 2</td>
<td>35851.676</td>
<td>1079</td>
<td>18.458</td>
<td></td>
</tr>
</tbody>
</table>

Table 20: ANOVA table for Tallaght

Table 21 shows the value for each coefficient. The cosine coefficient of the wind direction is positive, which points to a contribution to the measured PM10 from the North, i.e. the traffic from the N81. The average rainfall has a negative impact on the pollution levels as expected. The average temperature has also a positive impact on the PM10 concentration levels, suggesting that secondary PM10 may give a positive contribution to the overall concentration levels. The particulate matter pollution here has no seasonal component as suggested by the exclusion of the “month” coefficient.
<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>497.491</td>
<td>331.846</td>
<td>1.499</td>
<td>.134</td>
</tr>
<tr>
<td>year</td>
<td>-.319</td>
<td>.166</td>
<td>-.045</td>
<td>-1.927</td>
</tr>
<tr>
<td>avg stab</td>
<td>1.345</td>
<td>.534</td>
<td>.062</td>
<td>2.519</td>
</tr>
<tr>
<td>avg dir</td>
<td>-.007</td>
<td>.002</td>
<td>-.081</td>
<td>-3.237</td>
</tr>
<tr>
<td>avg cos</td>
<td>3.464</td>
<td>.237</td>
<td>.364</td>
<td>14.639</td>
</tr>
<tr>
<td>avg temp</td>
<td>2.292</td>
<td>1.023</td>
<td>1.967</td>
<td>2.241</td>
</tr>
<tr>
<td>max temp</td>
<td>-.590</td>
<td>.144</td>
<td>-.537</td>
<td>-4.094</td>
</tr>
<tr>
<td>avg hum</td>
<td>.549</td>
<td>.191</td>
<td>.680</td>
<td>2.869</td>
</tr>
<tr>
<td>max hum</td>
<td>-.130</td>
<td>.051</td>
<td>-.089</td>
<td>-2.561</td>
</tr>
<tr>
<td>avg dew</td>
<td>-1.947</td>
<td>1.022</td>
<td>-1.538</td>
<td>-1.905</td>
</tr>
<tr>
<td>min dew</td>
<td>-.313</td>
<td>.132</td>
<td>-.259</td>
<td>-2.376</td>
</tr>
<tr>
<td>avg press</td>
<td>-.248</td>
<td>.134</td>
<td>-.531</td>
<td>-1.856</td>
</tr>
<tr>
<td>max press</td>
<td>.158</td>
<td>.080</td>
<td>.311</td>
<td>1.972</td>
</tr>
<tr>
<td>min press</td>
<td>.208</td>
<td>.068</td>
<td>.485</td>
<td>3.056</td>
</tr>
<tr>
<td>avg rain</td>
<td>-2.593</td>
<td>.725</td>
<td>-.094</td>
<td>-3.576</td>
</tr>
<tr>
<td>max solar</td>
<td>.008</td>
<td>.003</td>
<td>.117</td>
<td>2.796</td>
</tr>
</tbody>
</table>

Table 21: Coefficients values for Tallaght

Table 22 presents the statistics for the excluded predictor variables.

<table>
<thead>
<tr>
<th>Model</th>
<th>Beta In</th>
<th>t</th>
<th>Sig.</th>
<th>Partial Correlation</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>day</td>
<td>.004</td>
<td>.164</td>
<td>.870</td>
<td>.005</td>
<td>.982</td>
</tr>
<tr>
<td>min temp</td>
<td>.080</td>
<td>.578</td>
<td>.563</td>
<td>.018</td>
<td>.027</td>
</tr>
<tr>
<td>min hum</td>
<td>.023</td>
<td>.342</td>
<td>.732</td>
<td>.010</td>
<td>.116</td>
</tr>
<tr>
<td>avg sin</td>
<td>.016</td>
<td>.310</td>
<td>.756</td>
<td>.010</td>
<td>.203</td>
</tr>
<tr>
<td>avg solar</td>
<td>-.052</td>
<td>-.555</td>
<td>.579</td>
<td>-.017</td>
<td>.059</td>
</tr>
<tr>
<td>max dew</td>
<td>-.135</td>
<td>-1.004</td>
<td>.315</td>
<td>-.031</td>
<td>.028</td>
</tr>
<tr>
<td>avg speed</td>
<td>.029</td>
<td>.907</td>
<td>.365</td>
<td>.028</td>
<td>.517</td>
</tr>
<tr>
<td>max speed</td>
<td>.039</td>
<td>1.248</td>
<td>.212</td>
<td>.038</td>
<td>.530</td>
</tr>
<tr>
<td>max rain</td>
<td>.060</td>
<td>1.280</td>
<td>.201</td>
<td>.039</td>
<td>.232</td>
</tr>
<tr>
<td>min speed</td>
<td>.035</td>
<td>1.205</td>
<td>.228</td>
<td>.037</td>
<td>.612</td>
</tr>
<tr>
<td>month</td>
<td>-.038</td>
<td>-1.528</td>
<td>.127</td>
<td>-.047</td>
<td>.819</td>
</tr>
</tbody>
</table>

Table 22: Statistics for excluded predictor variables for Tallaght

The following variables are the best predictors for the PM10 background concentration levels in Tallaght: year, average atmospheric stability class, maximum solar radiation, average cosine of the wind direction, atmospheric pressure, maximum
and average humidity, minimum and average dew point, average rainfall, maximum and average temperature.

### 3.5.3 ANN models

An Artificial Neural Network is a mathematical representation of the operation of biological neural networks, in other words it is an emulation of a biological neural system. The principal characteristics of ANNs are:

- Highly adaptive to non-parametric data distributions;
- No prior hypotheses about the relationships between the variables;
- Very low sensitivity to error term assumptions;
- Higher tolerance to noise, chaotic components and heavy tails than most alternative computational methods.

An ANN consists of an input layer, one or more layers of neurons (mathematical entities whose behaviour is governed by a predefined function) and an output layer. Each ANN has to be trained first (A), by using a number of data instances as inputs, and by trying to “fit” to some other data instances (output), commonly in terms of predicting the (future) behaviour of parameters of interest with the aid of their history (Kyriakidis et al., 2009). Then the performance of the model needs to be assessed (B).

(A) **Training**

The training function used for modelling the PM10 background concentrations at the different measurement stations for considered in this work is the Levenberg-Marquardt Algorithm (Levenberg, 1944, Marquardt, 1963). This algorithm provides a numerical solution to the problem of minimising a function, generally nonlinear, over a space of parameters of the function. The Levenberg-Marquardt Algorithm (LMA) interpolates between the Gauss–Newton Algorithm (Fletcher, 1987, Björck, 1996) and the method of gradient descent, which is a first order optimisation algorithm. The LMA is more robust than the Gauss–Newton Algorithm, which means that in many cases it finds a solution even if it starts very far away from the final minimum.

The primary application of the Levenberg–Marquardt algorithm is in the least squares curve fitting problem: given a set of m empirical datum pairs of independent and dependent variables \((x, y)\), it optimises the parameters \(\beta\) of the model curve \(f(x, \beta)\) so that the sum of the squares of the deviations becomes minimal.
\[
S(\beta) = \sum_{i=1}^{m} [y_i - f(x_i, \beta)]^2
\]

(Equation 3.2)

Like other numeric minimization algorithms, the Levenberg–Marquardt algorithm is an iterative procedure. In each iteration step, the parameter vector \( \beta \) is replaced by a new estimate, \( \beta + \delta \). To determine \( \delta \), the functions \( f(x_i, \beta+\delta) \) are approximated by their linearization:

\[
f(x_i, \beta + \delta) \approx f(x_i, \beta) + J_i \delta
\]

(Equation 3.3)

Where:

\[
J_i = \frac{\partial f(x_i, \beta)}{\partial \beta}
\]

(Equation 3.4)

\( J_i \) is the local gradient of \( f \) with respect to \( \beta \) at \( x_i \).

At its minimum, the sum of squares \( S(\beta) \), the gradient of \( S \) with respect to \( \delta \) is zero. The first-order approximation of \( f(x_i, \beta+\delta) \) in the equation above gives:

\[
S(\beta + \delta) \approx \sum_{i=1}^{m} (y_i - f(x_i, \beta) - J_i \delta)^2
\]

(Equation 3.5)

Or in vector notation,

\[
S(\beta + \delta) \approx \|y - f(\beta) - J\delta\|^2
\]

(Equation 3.6)

Taking the derivative with respect to \( \delta \) and setting the result to zero gives:

\[
(J^T J)\delta = J^T [y - f(\beta)]
\]

(Equation 3.7)

where \( J \) is the Jacobian matrix whose \( i^{th} \) row equals \( J_i \), and where \( f \) and \( y \) are vectors with \( i^{th} \) component \( f(x_i, \beta) \) and \( y_i \), respectively. This is a set of linear equations which can be solved for \( \delta \).

Levenberg (Levenberg, 1944) replaced the previous equation (Equation 3.7) with a "damped version":

\[
(J^T J + \lambda I)\delta = J^T [y - f(\beta)]
\]

(Equation 3.8)
where I is the identity matrix, giving as the increment \( \delta \), to the estimated parameter vector \( \beta \).

The (non-negative) damping factor \( \lambda \) is adjusted at each iteration. If the reduction of \( S \) is rapid, a smaller value can be used, bringing the algorithm closer to the Gauss–Newton algorithm, whereas if an iteration gives insufficient reduction in the residual, \( \lambda \) can be increased, giving a step closer to the gradient descent direction. Note that the gradient of \( S \) with respect to \( \beta \) equals \(-2(J^T[y - f(\beta)])^T\). Therefore, for large values of \( \lambda \), the step will be taken approximately in the direction of the gradient. If either the length of the calculated step \( \delta \), or the reduction of sum of squares from the latest parameter vector \( \beta + \delta \), fall below predefined limits, iteration stops and the last parameter vector \( \beta \) is considered to be the solution.

Levenberg’s algorithm has the disadvantage that if the value of damping factor \( \lambda \) is large, inverting \( J^TJ + \lambda I \) is not used at all. Marquardt improved the algorithm by scaling each component of the gradient according to the curvature so that there is larger movement along the directions where the gradient is smaller. This avoids slow convergence in the direction of small gradient. Therefore, Marquardt replaced the identity matrix \( I \) with the diagonal matrix consisting of the diagonal elements of \( J^TJ \), resulting in the Levenberg–Marquardt algorithm:

\[
(J^TJ + \lambda \text{diag}(J^TJ))\delta = J^T[y - f(\beta)]
\]

(Equation 3.9)

(B) Performance assessment of the training procedure (Beale et al., 2011)

Once the network weights and biases are initialized, the network is ready for training. The multilayer feed-forward network can be trained for function approximation (nonlinear regression) or pattern recognition. The training process requires a set of examples of proper network behaviour - network inputs \( p \) and target outputs \( t \).

The process of training a neural network involves tuning the values of the weights and biases of the network to optimize network performance. The default performance function for feed-forward networks is mean square error (mse), the average squared error between the networks outputs and the target outputs \( t \). It is defined as follows:

\[
F = \text{mse} = \frac{1}{N} \sum_{i=1}^{N} (e_i)^2 = \frac{1}{N} \sum_{i=1}^{N} (t_i - a_i)^2
\]

(Equation 3.10)

In the default mean square error performance function, each squared error contributes the same amount to the performance. The error weighting object needs to have the
same dimensions as the target data. In this way, errors can be weighted according to time step, sample number, signal number or element number.

3.6 Background PM10 models results

The PM10 background concentration levels are modelled using the Matlab Neural Network Toolbox which has the capabilities to use non-linear auto-regression with external input networks (NARX). The NARX network is a dynamical neural architecture commonly used for input-output modelling of nonlinear dynamical systems. When applied to time series prediction, the NARX network is designed as a feed-forward time delay Neural Network (Menezes Jr and Barreto, 2008). It predicts series y(t), which in this instance is the PM10 concentrations, given a certain number of past values of y(t) and another series x(t), which here are the weather variables (Figure 43).

In Figure 43, 10 is the number of predictor variables, 1 is the number of target variables, 1:2 is the delay, w is a scalar weight and b is a scalar bias. In one-step-ahead prediction tasks, ANN models are required to estimate the next sample value of a time series, without feeding back it to the model's input regressor; in other words, the input regressor contains only actual sample points of the time series. The NARX network can also be applied to multi-step-ahead or long-term prediction by feeding back the model's output to the input regressor for a fixed but finite number of time steps (Menezes Jr and Barreto, 2008). For the purpose of this work, the PM10 background concentration levels prediction are made for a time frame variable between one and seven steps/days ahead, according to the availability and reliability of the forecasted weather data.

The standard NARX network is a two-layer feed-forward network, with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer (Beale et al., 2011). This network also uses tapped delay lines to store previous values of the x(t) and y(t) sequences. Note that the output of the NARX network, y(t), is fed back to the input of the network through delays (Figure 44), since y(t) is a function of y(t−1), y(t−2), ..., y(t−d).
However, for efficient training and for predictions this feedback loop is successively opened. From Figure 45, it can be seen that the network is identical to the previous open-loop network, except that one delay has been removed from each of the tapped delay lines. The output of the network is then $y(t+1)$ instead of $y(t)$.

The PM10 background concentration levels data and the weather data are inputted to the Matlab NARX model, which divides them into sub-datasets for training, validation and testing purposes according to the following ratios: 70% for training, 15% for validation and 15% for testing. All these steps are performed automatically by the algorithm prepared with the Matlab Neural Network toolbox.

The followings sections show the results obtained by applying the modelling algorithm to the PM10 data collected at the different monitoring stations and the pre-processed weather data described above.
3.6.1 Background pollution ANN models

The procedure described above is applied to the background monitoring stations dataset discussed in the previous section. The schemes of the NARX networks used for modelling the PM10 background concentration levels for these stations are presented in Figure 46 to Figure 49 below.

The networks’ structures are summarised in Table 23 below.
The performances of the four models are very high, with a correlation coefficient $R$ close to the unit for the four cases (Figure 50 to Figure 53). The plots in Figure 50 to
Figure 53 graphically show the correlation between modelled and measured values for PM10 (µg/m³): as it can be noted, the correlation is very high for all the models and the output slightly diverges from the target for all of them, with a few sporadic outliers.

The error histograms in Figure 54 to Figure 57 show an almost Gaussian distribution of the errors around the zero for all the models. The histogram gives an indication of outliers, which are data points where the fit is significantly worse than the majority of data. For the Dun Laoghaire case (Figure 54), while most errors fall between -6 and 6, there is a point with an error of 30.63 [µg/m³].

![Error Histogram with 20 Bins](image)

**Figure 54: Error histogram for Dun Laoghaire [µg/m³]**

For the Phoenix Park case (Figure 55), most errors fall between -4.2 and 4.2 and there are no important outliers.

![Error Histogram with 20 Bins](image)

**Figure 55: Error histogram for Phoenix Park [µg/m³]**
For the Rathmines case (Figure 56), while most errors fall approximately between -7 and 7, there are some possible outliers with errors of -15.4, 16.25 and 18.01.

![Error Histogram for Rathmines (µg/m³)](image)

Finally for the Tallaght case (Figure 57), most errors fall between -7 and 7 and there are some possible outliers with errors of -18.27, -14.92 and 13.62.

![Error Histogram for Tallaght (µg/m³)](image)

The outliers identified in the various error histograms presented above can also be recognised in the time series response for the four models presented in Figure 58 to Figure 61.

For the Dun Laoghaire case (Figure 58), the outlier can be found around the time value 10. This is explained if it’s considered that the fraction of the PM10 data around
the outlier was interpolated to fill a gap in the original dataset, as it can be seen in Figure 62.

![Figure 58: Distribution of errors in the time-series for Dun Laoghaire](image)

For the Phoenix Park case, it can be observed from the distribution of errors varying with time shown in Figure 59 that there are no errors of significant magnitude.

![Figure 59: Distribution of errors in the time-series for Phoenix Park](image)

For the Rathmines case, the outliers identified in the error histogram are also visible in the distribution of errors in time in Figure 60 below around the time step 50 and 350: the errors are in correspondence of important peaks, which might not be explained by the weather variables included in the model and it might be traffic-related.
For the Tallaght case, the outliers identified in the error histogram can be identified around the time-steps 50 and 650 (Figure 61).

3.6.2 ANN background models results

This section presents the results from the ANN background models presented above. For the Dun Laoghaire case, Figure 62 and Figure 63 show the PM10 plots before pre-processing in green, pre-processed in navy and modelled in red. The plot in Figure 62 shows a good correlation between the pre-processed PM10, which was used as target in the ANN model, and the modelled PM10. The major discrepancies are circled with double black lines and they seem to happen where the missing data (Figure 62) was interpolated. The ANN model is not able to reproduce many of the larger peaks in the pre-processed PM10 dataset for Dun Laoghaire. However it has to be noted that these peaks are probably not weather related but they might depend on particularly high local traffic events.
The results from the ANN background model for Phoenix Park are presented in Figure 64, with a comparison between the pre-processed PM10 and the modelled PM10. The plot below shows a good correlation between the pre-processed PM10, which was used as target in the ANN model, and the modelled PM10. The ANN model is able to reproduce most of the peaks of the pre-processed PM10 dataset for Phoenix Park. A few peaks are not reproduced, but it may be possible that these peaks may be due to regional pollution effects.
Figure 64: Pre-processed and modelled PM10, Phoenix Park

For the Rathmines case, a comparison between the pre-processed PM10 and the modelled PM10 is shown in Figure 65. The plot below shows a good correlation between the two variables. The ANN model is able to reproduce most of the peaks of the pre-processed PM10 dataset for Rathmines. A few peaks are not reproduced, but it has to be noted that they might due to particularly high local traffic events or to regional pollution and particularly stagnant weather conditions.

Finally for the Tallaght case, Figure 66 and Figure 67 show the PM10 plots before pre-processing in green, pre-processed in navy and modelled in red.
The plots above show a good correlation between the pre-processed PM10 and the modelled PM10. The major discrepancies are circled with double black lines and they seem to happen where the missing data (Figure 67) was interpolated. Again, the ANN model fails to reproduce some of the peaks of the pre-processed PM10 dataset for Tallaght.

### 3.7 Urban Traffic case: Winetavern Street

The PM10 concentration levels measured by a monitoring station classified as “urban traffic” by the Irish EPA are analysed and modelled in this section. The station is located on Winetavern Street and it was chosen because of its position which makes it an ideal example of the air pollution in the city centre (Figure 68). A NARX model is
built also for this typology of monitoring station as a tool to quantify the average air pollution in the area surrounding a station in an urban environment. As such, the model gives an indication of the pollution in the area, not the pollution at street level for each road link, which varies for each street according to traffic and road geometry specific for the link and it is calculated using a more advanced street pollution model, as discussed in Chapter 4.

Figure 68: PM10 monitoring station, Winetavern Street

It has been active since the 1st of January 2001 and the data is available from that date until the 31st of December 2010. As it can be seen from the time series for this station presented in Figure 69, there are minor gaps in the datasets.

Figure 69: PM10 time series for Winetavern Street
3.7.1 Data pre-processing

The procedure used in the previous section for pre-processing the PM10 monitored levels is also applied to the dataset acquired by this station. The pre-processing steps suggested by Pyle (Pyle, 1999), which can be grouped in the following actions:

1. Identifying and removing of the outliers;
2. Data smoothing;
3. Handling of missing data;
4. Data de-trending.

are again utilised in order to enhance the quality of the PM10 concentration levels data. After each step, the correlation with the weather parameters data is tested with backward linear regression; this operation is carried out for each weather monitoring station. As a result of such procedure, it was found that the Winetavern Street PM10 data have the highest correlation coefficient with Casement weather data and wind data. The correlation coefficient R was found to have a value of 0.729.

The outcome of the pre-processing procedure applied to the PM10 time series is shown in Figure 70 below.

![Figure 70: PM10 pre-processed time series for Winetavern Street](image)

3.7.2 Sensitivity analysis

The initial model included the same variables initially included in the background pollution models described in the previous section and the same culling procedure adopted for those datasets was used here.
The ANOVA table for the two models is presented below.

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>105159.713</td>
<td>26</td>
<td>4044.604</td>
<td>151.547</td>
</tr>
<tr>
<td>Residual</td>
<td>92930.123</td>
<td>3482</td>
<td>26.689</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>198089.836</td>
<td>3508</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td>105038.771</td>
<td>20</td>
<td>5251.939</td>
<td>196.868</td>
</tr>
<tr>
<td>Residual</td>
<td>93051.065</td>
<td>3488</td>
<td>26.677</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>198089.836</td>
<td>3508</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 25: ANOVA table for Winetavern Street

The final model consisted of the following predictor variables: (Constant), year, month, average atmospheric stability class, average wind direction, average and maximum solar radiation, atmospheric pressure, average and maximum humidity, wind speed, average rainfall, maximum and average temperature and average dew point. As it
could be noted from the R scores in Table 24, the final model with 20 predictor variables has similar performance to the initial model. These variables appear to be the best predictors for PM10 concentration levels for Winetavern Street.

Table 26 shows the value for each coefficient. The cosine coefficient (wind direction) is positive and the sine is negative and they have high values, which means that positive important contributions to PM10 values are due to wind blowing from South-East, i.e. from the crossroad between High Street and Nicholas Street.

The maximum rainfall variable has a negative contribution to PM10 levels as expected. The average temperature has also a positive impact on the PM10 concentration levels, suggesting that secondary PM10 may give a positive contribution to the overall concentration levels. The month parameter is correlated to the measured PM10 levels for this site, which points to a seasonal behaviour of the pollution.

Table 27 presents the statistics for the excluded predictor variables.

<table>
<thead>
<tr>
<th>Excluded Variables</th>
<th>Beta In</th>
<th>t</th>
<th>Sig.</th>
<th>Partial Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>min hum</td>
<td>.011</td>
<td>.338</td>
<td>.735</td>
<td>.006</td>
</tr>
<tr>
<td>min temp</td>
<td>-.028</td>
<td>-.525</td>
<td>.600</td>
<td>-.009</td>
</tr>
<tr>
<td>min dew</td>
<td>-.013</td>
<td>-.258</td>
<td>.797</td>
<td>-.004</td>
</tr>
<tr>
<td>max dew</td>
<td>-.050</td>
<td>-.867</td>
<td>.386</td>
<td>-.015</td>
</tr>
<tr>
<td>max rain</td>
<td>-.020</td>
<td>-.903</td>
<td>.367</td>
<td>-.015</td>
</tr>
<tr>
<td>day</td>
<td>-.017</td>
<td>-1.435</td>
<td>.151</td>
<td>-.024</td>
</tr>
</tbody>
</table>

Table 27: Statistics for excluded predictor variables for Winetavern Street

The following variables are the best predictors for the PM10 concentration levels in Winetavern Street: average atmospheric stability class, global radiation, wind speed direction, year, atmospheric pressure, month, average and maximum humidity, average rainfall, maximum and average temperature and average dew point. The global R score for the model suggests that the PM10 levels for this monitoring site might be affected by other variables, i.e. the traffic in the surrounding area.

3.7.3 Winetavern PM10 ANN model results

The procedure described above is applied to Winetavern dataset. A scheme of the NARX network used for modelling the PM10 concentration levels for Winetavern is presented in Figure 71 below.
The network consists of 19 predictors, 14 hidden neurons and 2 delays in the hidden layer. The performance of the model is very high, with a correlation coefficient $R$ close to the unit: the output of the model slightly diverges from the target (Figure 72).
Figure 74: Distribution of errors in the time-series

Figure 75: PM10 before the interpolation

Figure 76: Pre-processed and modelled PM10
he error histogram in Figure 73 shows an almost Gaussian distribution of the errors around the zero. The histogram gives another proof of the existence of outliers, which are data points where the fit is significantly worse than the majority of data. In this case, while most errors fall between -7.7 and 7.7, there are training points with errors of -16.95, -12.34, 15.34, 19.95 and 26.87. These outliers are also visible in the testing regression plot in Figure 72. The outliers can also be identified in the time series response in Figure 74 around the time value 600, 800, 2600 and 2800. This is explained if it’s considered that the fraction of the PM10 data around the outlier was interpolated to fill a gap in the original dataset, as it can be seen in Figure 62.

Figure 75 and Figure 76 show the PM10 plots before pre-processing in green, pre-processed in navy and modelled in red.

The plots above show a good correlation between the pre-processed PM10, which was used as target in the ANN model, and the modelled PM10. The major discrepancies are circled with double black lines and they seem to happen where the missing data (Figure 76) was interpolated. The ANN model is not able to reproduce some of the peaks of the pre-processed PM10 dataset for Winetavern. However it has to be noted that these peaks are probably not weather related but they might depend on local effects, such as traffic, and regional pollution. The model is also able to reproduce the general trend and the seasonal variation of the PM10 concentration levels also measured by the monitoring station in Winetavern Street.

3.8 Hourly average model for PM10

The PM10 concentration levels measured by the monitoring stations in the Dublin area are available only as daily averages and therefore they don’t provide a sufficiently detailed description of the pollution for the personal exposure model which is one of the goals of this work. On the other hand, the monitoring station provide hourly averages for NOx, NO and NO2 concentration levels. A few monitoring stations provide both daily PM10 averages and hourly nitrates averages, which is an ideal situation for finding a relation between the two datasets. Two stations are selected for this task, Winetavern Street and Rathmines, a “traffic pollution” and a “background pollution” monitoring station. As such, a modelling strategy to obtain an hourly distribution profile for the PM10 concentration levels from the hourly NOx, NO and NO2 time-series is devised. The hourly distribution profile for PM10 is necessary as input for the street pollution model presented in Chapter 4 and also for the personal exposure model discussed in Chapter 7.
Two different NARX models are created:

1. Hourly PM10 traffic: using the data from Winetavern Street monitoring station;
2. Hourly PM10 background: using the data from Rathmines monitoring station.

3.8.1 Sensitivity Analysis

The pre-processed PM10 datasets for the two models are regressed against the NOx, NO and NO2 datasets and also against the weather data and wind data, which showed to have the highest correlation coefficient with the particulate matter measurements (Chapter 3.7).

Both initial models included the following variables: year, month, day, average atmospheric stability class, wind speed (average, minimum and maximum), wind direction (average), temperature (average, minimum and maximum), humidity (average, minimum and maximum), dew point (average, minimum and maximum), atmospheric pressure (average, minimum and maximum), rainfall (average and maximum), solar radiation (average and maximum) and daily averages for NOx, NO and NO2.

3.8.1.1 Winetavern Street – traffic PM10 hourly model

Seven different regression models were generated by excluding each time a single variable with minimum impact on the global linear correlation coefficient value: the dependent variable PM10 was regressed against the complete group of the above mentioned n (29) predictor variables and the R score was calculated, then the predictor variable with less correlation was eliminated and this procedure was repeated for (n-1) predictor variables. This culling procedure was repeated until all the predictor variables with minor correlation with PM10 were eliminated.

The final model consisted of the following predictor variables: (Constant), year, month, day, average atmospheric stability class, average wind direction, average and maximum solar radiation, average and minimum atmospheric pressure, average and maximum humidity, wind speed, average rainfall, average temperature, average dew point and daily averages for NOx, NO and NO2. As it could be noted from the R scores in Table 28, the final model with 23 predictor variables has identical performance to the initial model. These variables appear to be the best predictors for the PM10 concentration levels for Winetavern Street.
### Table 28: R scores for the initial and final models for Winetavern Street

<table>
<thead>
<tr>
<th>Number of variables</th>
<th>R</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>.727</td>
<td>4.816</td>
</tr>
<tr>
<td>23</td>
<td>.727</td>
<td>4.814</td>
</tr>
</tbody>
</table>

The ANOVA table for the two models is presented below.

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>74704.121</td>
<td>29</td>
<td>2576.004</td>
<td>111.034</td>
</tr>
<tr>
<td>1 Residual</td>
<td>66630.793</td>
<td>2872</td>
<td>23.200</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>141334.914</td>
<td>2901</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td>74624.763</td>
<td>23</td>
<td>3244.555</td>
<td>139.976</td>
</tr>
<tr>
<td>2 Residual</td>
<td>66710.151</td>
<td>2878</td>
<td>23.179</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>141334.914</td>
<td>2901</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 29: ANOVA table for Winetavern Street

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>1053.124</td>
<td>73.583</td>
<td>14.312</td>
<td>.000</td>
</tr>
<tr>
<td>year</td>
<td>-.617</td>
<td>.037</td>
<td>-16.889</td>
<td>.000</td>
</tr>
<tr>
<td>month</td>
<td>-.230</td>
<td>.028</td>
<td>-8.131</td>
<td>.000</td>
</tr>
<tr>
<td>day</td>
<td>-.029</td>
<td>.010</td>
<td>-2.885</td>
<td>.004</td>
</tr>
<tr>
<td>avg stab</td>
<td>1.551</td>
<td>.394</td>
<td>3.941</td>
<td>.000</td>
</tr>
<tr>
<td>avg dir</td>
<td>-.006</td>
<td>.003</td>
<td>-1.916</td>
<td>.055</td>
</tr>
<tr>
<td>avg cos</td>
<td>3.661</td>
<td>.190</td>
<td>19.313</td>
<td>.000</td>
</tr>
<tr>
<td>avg sin</td>
<td>-1.058</td>
<td>.424</td>
<td>-2.497</td>
<td>.013</td>
</tr>
<tr>
<td>avg speed</td>
<td>-.373</td>
<td>.168</td>
<td>-2.221</td>
<td>.026</td>
</tr>
<tr>
<td>max speed</td>
<td>.157</td>
<td>.093</td>
<td>1.692</td>
<td>.091</td>
</tr>
<tr>
<td>min speed</td>
<td>.204</td>
<td>.104</td>
<td>1.956</td>
<td>.051</td>
</tr>
<tr>
<td>avg temp</td>
<td>4.192</td>
<td>.662</td>
<td>6.329</td>
<td>.000</td>
</tr>
<tr>
<td>avg hum</td>
<td>.821</td>
<td>.132</td>
<td>6.226</td>
<td>.000</td>
</tr>
<tr>
<td>max hum</td>
<td>-.115</td>
<td>.033</td>
<td>-3.448</td>
<td>.001</td>
</tr>
<tr>
<td>avg dew</td>
<td>-4.609</td>
<td>.671</td>
<td>-6.869</td>
<td>.000</td>
</tr>
<tr>
<td>avg press</td>
<td>-.555</td>
<td>.094</td>
<td>-5.903</td>
<td>.000</td>
</tr>
<tr>
<td>max press</td>
<td>.343</td>
<td>.055</td>
<td>6.250</td>
<td>.000</td>
</tr>
<tr>
<td>min press</td>
<td>.343</td>
<td>.050</td>
<td>6.904</td>
<td>.000</td>
</tr>
<tr>
<td>avg rain</td>
<td>-2.477</td>
<td>.540</td>
<td>-4.585</td>
<td>.000</td>
</tr>
<tr>
<td>avg solar</td>
<td>-.047</td>
<td>.009</td>
<td>-5.098</td>
<td>.000</td>
</tr>
<tr>
<td>max solar</td>
<td>.014</td>
<td>.004</td>
<td>3.584</td>
<td>.000</td>
</tr>
<tr>
<td>NOX DayAvg ppb</td>
<td>.041</td>
<td>.005</td>
<td>8.193</td>
<td>.000</td>
</tr>
<tr>
<td>NO DayAvg ppb</td>
<td>-.013</td>
<td>.007</td>
<td>-1.818</td>
<td>.069</td>
</tr>
<tr>
<td>NO2 DayAvg ppb</td>
<td>.056</td>
<td>.006</td>
<td>9.102</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 30: Coefficients values for Winetavern Street
Table 30 shows the value for each coefficient. The NOx and NO2 have positive coefficients while the NO shows a negative correlation with the PM10 measured levels.

Table 31 presents the statistics for the excluded predictor variables.

![Table 31: Statistics for excluded predictor variables for Winetavern Street](image)

3.8.1.2 Rathmines – background PM10 hourly model

Fifteen different regression models were generated by excluding each time a single variable with minimum impact on the global linear correlation coefficient value: the dependent variable PM10 was regressed against the complete group of the above mentioned n (29) predictor variables and the R score was calculated, then the predictor variable with less correlation was eliminated and this procedure was repeated for (n-1) predictor variables. This culling procedure was repeated until all the predictor variables with minor correlation with PM10 were eliminated.

The final model consisted of the following predictor variables: (Constant), year, average atmospheric stability class, humidity, cosine of the wind direction, maximum solar radiation, average rainfall, minimum atmospheric pressure, average wind speed, average dew point, average and daily averages for NOx, NO and NO2. As it could be noted from the R scores in Table 32, the final model with 15 predictor variables has similar performance to the initial model. These variables appear to be the best predictors for the PM10 concentration levels for Rathmines.

![Table 32: R scores for the initial and final models for Rathmines](image)

The ANOVA table for the two models is presented below.
Table 33: ANOVA table for Rathmines

Table 34 shows the value for each coefficient. The NOx and NO2 again have positive coefficients while the NO shows a negative correlation with the PM10 measured levels.

Table 34: Coefficients values for Rathmines

Table 35 presents the statistics for the excluded predictor variables.
Excuded Variables | Beta In | t | Sig. | Partial Correlation | Collinearity Statistics |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>max msl</td>
<td>-.040</td>
<td>-.773</td>
<td>.440</td>
<td>-.024</td>
<td>.140</td>
</tr>
<tr>
<td>avg glo</td>
<td>-.012</td>
<td>-.178</td>
<td>.859</td>
<td>-.005</td>
<td>.077</td>
</tr>
<tr>
<td>min temp</td>
<td>-.050</td>
<td>-.615</td>
<td>.539</td>
<td>-.019</td>
<td>.054</td>
</tr>
<tr>
<td>max rain</td>
<td>-.021</td>
<td>-.513</td>
<td>.608</td>
<td>-.016</td>
<td>.209</td>
</tr>
<tr>
<td>avg press</td>
<td>-.098</td>
<td>-1.068</td>
<td>.286</td>
<td>-.033</td>
<td>.043</td>
</tr>
<tr>
<td>min speed</td>
<td>-.015</td>
<td>-.394</td>
<td>.694</td>
<td>-.012</td>
<td>.254</td>
</tr>
<tr>
<td>max speed</td>
<td>-.043</td>
<td>-.804</td>
<td>.422</td>
<td>-.025</td>
<td>.125</td>
</tr>
<tr>
<td>avg sin</td>
<td>.016</td>
<td>.775</td>
<td>.438</td>
<td>.024</td>
<td>.809</td>
</tr>
<tr>
<td>month</td>
<td>-.029</td>
<td>-1.203</td>
<td>.229</td>
<td>-.037</td>
<td>.610</td>
</tr>
<tr>
<td>max temp</td>
<td>.112</td>
<td>1.017</td>
<td>.309</td>
<td>.031</td>
<td>.030</td>
</tr>
<tr>
<td>max dew</td>
<td>-.075</td>
<td>-.812</td>
<td>.417</td>
<td>-.025</td>
<td>.043</td>
</tr>
<tr>
<td>min dew</td>
<td>-.068</td>
<td>-.768</td>
<td>.443</td>
<td>-.024</td>
<td>.047</td>
</tr>
<tr>
<td>avg dir</td>
<td>-.030</td>
<td>-1.503</td>
<td>.133</td>
<td>-.046</td>
<td>.890</td>
</tr>
<tr>
<td>day</td>
<td>-.030</td>
<td>-1.563</td>
<td>.118</td>
<td>-.048</td>
<td>.991</td>
</tr>
</tbody>
</table>

Table 35: Statistics for excluded predictor variables for Rathmines

3.8.2 ANN hourly models

The procedure described in the previous sections is applied to Winetavern and Rathmines datasets integrated with NOx, NO and NO2 measurements. A scheme of the NARX network used for modelling the PM10 concentration levels is presented in Figure 77 for Winetavern and in Figure 78 for Rathmines. The network for Winetavern consists of 17 predictors and 14 hidden neurons and 2 delays in the hidden layer, while the network for Rathmines consists of 13 predictors and 9 hidden neurons and 2 delays in the hidden layer. The performances of both models are very high, with correlation coefficients R close to the unit: the outputs of both models slightly diverge from the targets (Figure 81 and Figure 82).

For the Winetavern case, it is possible to identify some outliers (data points where the fit is significantly worse than the majority of data) which are also visible in the histogram below. The error histogram in Figure 79 shows an almost Gaussian distribution of the errors around the zero with most errors between -7 and 7, with 3 important outliers at -27.48, -13.7 and 16.15.
The outliers can also be identified in the time series response in Figure 80. This is explained if it’s considered that the fraction of the PM10 data around the outlier was interpolated to fill a gap in the original dataset.
Figure 80: Winetavern, distribution of errors in the time-series

Figure 81: Winetavern, correlation coefficient between measured and modelled PM10 [µg/m³]  

Figure 82: Rathmines, correlation coefficient between measured and modelled PM10 [µg/m³]

Figure 83: Rathmines, error histogram [µg/m³]
For the Rathmines case, the error histogram in Figure 83 shows the existence of outliers; in this case, while most errors fall between -4 and 4, there are training points with errors of -9.7, 10.8 and 9.3. The outliers can also be identified in the time series response in Figure 84, around the time step 80 and 720.

![Response of Output Element 1 for Time-Series 1](image)

Figure 84: Rathmines, distribution of errors in the time-series

### 3.8.3 ANN hourly models - results

The traffic (Winetavern) and the background (Rathmines) hourly PM10 models are generated by using the hourly values for all the variables and by applying the NARX models to them. The outputs for Winetavern and Rathmines are shown in Figure 85 and Figure 86 respectively: the two plots which display the hourly PM10 concentration levels distribution for a week.

![Hourly distribution](image)

Figure 85: Hourly distribution for PM10 and NO2 for Winetavern
Both the modelled time-series for the hourly variation of PM10 concentration levels have similar trends and behaviours of the NO2 time-series. Unfortunately, the Irish Environmental Protection Agency doesn’t provide also hourly time-series for PM1); as such, it is not possible to validate these two models by comparing the modelled data to actual measurements.

Nonetheless, the creation of the background PM10 model is particularly important and necessary for modelling the pollution for the personal exposure model. The commuting time-frame is only a fraction of the day and it might last for less than 1 hour; therefore, it is important to have a more refined estimate of the pollution than the daily average.

### 3.9 GIS application

As an example of possible GIS application, PM10 background levels are modelled for the four background monitoring stations described above using the ANN models. The resulting concentration levels are imported into ArcGIS and a shapefile is generated to calculate the overall PM10 concentrations in the Dublin area around the four monitoring stations. As such, a raster (a rectangular grid of pixels) is generated by interpolating the PM10 background modelled values for the four stations. The interpolation technique used for this task is the Natural Neighbour (Sibson, 1981). This procedure supports an estimate of the modelled PM10 background concentration levels at each point in Dublin. An example of this is shown in Figure 87 below, where
The PM10 background concentration levels are calculated, interpolated and graphically presented using ArcGIS.

![Raster of interpolated values of PM10 background concentration levels in Dublin](image)

Figure 87: Raster of interpolated values of PM10 background concentration levels in Dublin

This procedure is useful to have an estimate of the PM10 background pollution at a distance from any of the four air quality monitoring stations: the generated raster includes calculated values for each point of the grid. While the ANN models are capable to replicate the variation in time of PM10 pollution, the interpolation between the modelled data for each station at each time step can add also a spatial dimension to the variation of the concentration levels of the pollutants in the Great Dublin area.
CHAPTER 4 - Urban traffic pollution modelling
4. Introduction

The contribution of road traffic to the ambient concentrations of various air pollutants can vary dramatically in urban areas. The 2003 London Atmospheric Emissions Inventory (G.L.A, 2006) estimated that road traffic accounts for over 51% of nitrous oxides and 73% particulate matter (PM10) levels. As a rule the closer people are to traffic, the more pollution they inhale. At pollution “hot spots” such as very busy roads, levels can be much higher than the urban average. An air quality modelling methodology aimed at evaluating the personal exposure of individuals in urban areas should include specific provision these important conditions.

The central part of Dublin (Westland Row, on the west side of Trinity college in Figure 88) is characterised by close spacing of buildings; this particular layout is called urban street canyon and it may cause a stable circulatory vortex being established within the street canyon if specific wind and geometry conditions are met.

A street canyon is a relatively narrow street between buildings that line up continuously along both sides. The combination of vehicle emissions and reduced dispersion in these circumstances can lead to high levels of pollution (Buckland and Middleton, 1999), which in turn pose a substantial health risk to pedestrians. When the wind blows across a street canyon, a vortex might be generated if particular conditions are met.

Oke (Oke, 1988) provided a systematic classification of flow regimes in urban street canyons (Figure 89), according to which the canyon geometry is defined mainly by the ratio H/W, where H is the average height of the canyon walls and W is the canyon width. As such the flow types are characterised by three regimes depending on the canyon geometry:

1. **isolated roughness flow**: for H/W < 0.3 (widely spaced buildings), the flow fields associated with the buildings do not interact;
2. **wake interference flow**: for 0.3 < H/W < 0.7, the wake created by the upwind building is distributed by the downwind building creating a downward flow along the windward face of this building;
3. **skimming flow**: for H/W > 0.7, a stable circulatory vortex is established in the canyon and the ambient flow is decoupled from the street flow.

It has to be noted that this classification of flow regimes provided by Oke (Figure 89) is a result of wind tunnel experiments and as such qualitative and different from “real life” conditions.
Wind speed and direction are the two major driving factors behind dispersion of pollutants within a street canyon and it is generally accepted that a street canyon vortex will not form at low wind speeds. Formation of a vortex within the canyon may occur at an ambient wind velocity above 1.5 - 2.0 m/s (Berkowicz et al., 1997), the formation of a stable vortex is much less likely at lower wind speeds, i.e. winds speed below 2 m/s (Vardoulakis et al., 2005, Kukkonen et al., 2003b). A street canyon vortex
is formed with wind directions perpendicular to the street canyon within a 30° scope and as the angle of the wind approaches that of the street axis, the vortex disappears (Kukkonen et al., 2001a). Therefore, a street vortex will only form when there is significant wind speed and the wind direction is near perpendicular to the street axis. Low wind speed and winds parallel to the street axis result in the worst dispersion conditions (Vardoulakis et al., 2002a). In the presence of a vortex the local airflow at street level is in the opposite direction to that at roof level (Figure 90), resulting in higher concentrations of pollutants on the leeward side of the street compared with the windward side (Manning et al., 2000). The windward side is here defined as the side of the street to which the wind blows whilst the leeward side is the side the wind blows from (Figure 90).

![Urban street canyon model](image)

Figure 90: Urban street canyon model (Berkowicz, 2000a)

4.1 Modelling pollutant dispersion in a street canyon: literature review

There are numerous models used in urban street canyon applications, which have been proved to be useful in air quality and traffic management, urban planning, interpretation of monitoring data, pollution forecasting, population exposure studies, etc. The models most commonly used for this task are the followings:

- **STREET**: Johnson et al. (Johnson et al., 1973) created this relatively simple Gaussian type diffusion simulation model for calculating urban carbon monoxide (CO) concentrations as a function of local meteorology and the distribution of traffic. STREET consists in a single box model, together with some simplified assumptions concerning initial dispersion and car induced
turbulence, which formed part of a multipurpose urban diffusion model for inert pollutants (APRAC);

- CPBM: Yamartino and Wiegand (Yamartino and Wiegand, 1986) developed the Canyon Plume Box Model, which combines a Gaussian plume model for the direct impact of pollutants emitted in the street, with a box model that accounts for the additional impact of pollutants trapped within the wind vortex formed inside the canyon;

- AEOLIUS: Buckland (Buckland, 1998) created this semi-empirical model employing concepts and techniques previously used for the development of CPBM. AEOLIUS calculates concentrations of exhaust gases on both sides of a canyon assuming three different contributions: (a) the contribution from the direct flow of pollutants from the source to the receptor, (b) the recirculation component due to the flow of pollutants around the vortex generated within the recirculation zone of the canyon, and (c) the urban background contribution;

- SLAQ: The Street Level Air Quality (SLAQ) model was developed by Micallef and Colls (Micallef and Colls, 1999); it's a semi-empirical emission dispersion model for predicting the temporal and spatial distribution of airborne particulate matter in street canyons. This model is mainly based on CPBM and it is provided with emission modules, meteorological pre-processors, modules for within-canyon processes, dispersion modules, and modules accounting for external influences;

- ADMS Urban: ADMS-Urban is basically an extension of the ADMS2 industrial model, combining a line and area source model with a street canyon model, a chemistry model for nitrogen, sulphur and ozone and a traffic emissions database (Robins et al., 1997). The street canyon analysis is based on the OSPM formulation (described below). For improved predictions, the model takes into account advanced algorithms the height dependence of wind speed, turbulence and stability. A validation study has been reported by Robins (Robins et al., 1997);

- OSPM: The Operational Street Pollution Model (OSPM) is a practical street pollution model developed by the National Environmental Research Institute (NERI), Denmark. It is based on a simplified description of flow and dispersion conditions in street canyons. It can be successfully applied to the prediction of street traffic pollution for single street canyons, multi-street networks or travel routes (Berkowicz, 2000a).
4.1.1 Critical comparison

This section presents a critical comparison between the urban street canyon models listed above and provides an argumentation for the choice of the most suitable model for the purpose of this work.

The six considered models are based on the assumption that pollutant concentrations at the roadside consist of two components: the urban background concentration and the concentration increment due to vehicle emissions generated within the specific street. The also all apply a Gaussian plume model for the direct contribution and a box model for the recirculating part of the pollutants in the street.

The considered models follow different approaches to calculate the concentrations of pollutants on the two sides of a street, as described below:

- **STREET**: on the leeward side, concentrations are assumed to be inversely proportional to the distance between the line source and the receptor point; on the windward side, the vertical decrease of concentrations due to entrainment of fresh air through the top of the canyon is taken into account;
- **CPBM and SLAQ**: on the leeward side, the total impact is calculated by adding the direct plume to the recirculated fraction; on the windward side, where the only contribution arises from the recirculation component, the dilution of the concentrations due to roof-level intrusion and entrainment of clean air is also taken into account. In particular, the plume-recirculation model is used only when the vortex advection dominates over diffusion;
- **AEOLIUS**: on the leeward side of the street, concentrations are calculated as the sum of the direct and recirculation contributions; on the windward side only the direct contribution of emissions generated outside the recirculation zone are taken into account;
- **ADMS-Urban and OSPM**: on the leeward side the direct contribution is calculated taking into account the emissions from traffic in the recirculation zone only; on the windward side, only contributions from the emissions outside the recirculation zone are taken into account (Berkowicz et al., 1997).

The models have also different approaches when the wind flow is near-parallel to the street axis or the wind speed is near zero (when cross-canyon turbulence dominates over advective transport), as follows:

- **CPBM, STREET and SLAQ**: a simpler non-vortex plume model is used (Gualtieri, 2010);
AEOLIUS, ADMS-Urban and OSPM: the concentrations on both sides of the canyon are considered equal.

Vardoulakis (Vardoulakis et al., 2003) tested the CPBM model against the STREET model and the first showed to perform better than the latter for those meteorological conditions for which STREET was not specifically designed. CPBM uses an empirical model to calculate the turbulent sigma parameters representing the standard deviation of flow velocities about the mean flow. This empirical model takes into account wind generated turbulence as well as thermal effects induced by solar radiation and moving vehicles (Vardoulakis et al., 2003). AEOLIUS empirically derives the mechanical turbulence in the street due to the wind and vehicle traffic. SLAQ accounts for non-uniformity in along-canyon emission rates and to model a street that has several intersections along its length by scaling exhaust emission rates along the street for each of the lanes according to the direction of traffic flow to account for modal operation of vehicles near signalised intersections. It also accounts for thermal turbulence due to environmental surface sensible heat, vehicle-generated heat and particle settling and estimation of wind direction standard deviation (Micallef and Colls, 1999). ADMS Urban models major roads as line sources whereas minor roads as area sources. OSPM achieves the best computational performances between the assessed models because of the parameterization of the flow and dispersion conditions in street canyons, which was deduced from extensive analysis of experimental data and model tests. The results of these tests were used to further improve the model performance, especially with regard to different street configurations and a variety of meteorological conditions (Berkowicz, 2000a).

It has to be noted that all the models considered here are not integrated or compatible with GIS: this is an important factor to base the choice on, because it may have an impact on the personal exposure model performance. In particular, it is important that the urban street canyon model is able to acquire the streets and buildings geometry data from GIS. This is a crucial characteristic which is required in order to avoid the manual input of such data in the dispersion model, which would be extremely time-demanding. A compatibility tool to acquire city geometry data from GIS is available for OSPM as part of the AirGIS system (Jensen et al., 2001).

The OSPM is selected for application with the urban air quality modelling framework developed in this thesis. This choice is guided by the successful application of the model in previous research projects in Trinity College Dublin (Ganguly, 2008) and also by the possibility of successfully integrate it with GIS. A more complete description of the OSPM is provided in the following section.
4.2 Operational Street Pollution Model (OSPM): General Principles

The calculation of the dispersion of gas exhausted can be efficiently carried out with OSPM. It’s a computer based semi-empirical dispersion model and its reduced degree of complexity leads to shorter computing times. The practical implication is that it allows local authorities to investigate areas of high atmospheric pollution quickly and, if necessary, designate them as local air quality management sites (Buckland and Middleton, 1999). It uses a plume equation to derive the direct contribution and a box model to calculate the recirculation of pollutant concentrations from traffic travelling in a street canyon (Vignati et al., 1999). The model is based on the six following assumptions (Berkowicz, 2000a):

1. A vortex is formed in the street whenever a wind component perpendicular to the street exists. Vortex length is dependent on upwind (leeward) building height. The length of the vortex is also dependent on the wind speed and wind direction.
2. The upwind receptor receives contributions from traffic emissions within the recirculation zone (i.e., the vortex area), the recirculation pollution and a portion of the emissions outside the recirculation zone. These contributions are wind speed and wind direction dependant.
3. The downwind receptor receives the contribution mainly from the recirculation zone and if the vortex does not occupy the entire street, traffic emissions from outside the recirculation zone.
4. As wind speed approaches calm conditions, or wind direction is parallel to the street, concentrations on both sides are equal.
5. The direct contribution in the recirculation zone is calculated by a plume model that assumes linear dispersion of pollutants with distance. Emissions are assumed to be uniformly distributed across the street.
6. A box model describes the recirculation component. Concentrations are calculated assuming equality of the incoming and outgoing pollution fluxes.

Emissions within the plume are assumed to come from a number of infinitesimal line sources aligned perpendicular to the wind direction at street level (Aquilina and Micallef, 2004). The model assumes traffic emissions are uniformly distributed across the street canyon and that traffic volume and average vehicle speed are constant over a particular calculation hour (Berkowicz et al., 1997), as such traffic is assumed to be a continuous stream.
The box model of the recirculation vortex is assumed to be trapezoidal in shape, again pollutants are assumed to be well mixed within the box, and ventilation of the recirculation zone (Figure 91) takes place through the edges of the trapezoid, but can be limited by the presence of a downwind building (Berkowicz et al., 1997). The wind direction at street level is assumed to be opposite to that at roof level and the wind speed at street level is calculated from the synoptic roof level wind speed and direction and the surface roughness of the surrounding buildings (Holmes and Morawska, 2006).

The turbulence in the street is modelled as two separate components: ambient turbulence, which depends on wind speed, and traffic-induced turbulence (Berkowicz, 200a). Traffic induced turbulence dominates in the street canyon when wind speeds are low, below 2m/s, and because highest concentrations occur in near calm conditions (i.e. reduced turbulence due to low wind speeds and increased traffic emissions due to low average vehicle speeds), appropriate modelling of these conditions is crucial (Ketzel et al., 2003). The effects of thermal stratification on turbulence are not considered and the stability class is assumed to be always neutral.

The main argument for such an assumption is that atmospheric stability is determined by both mechanical stresses and buoyant forcing but, for a typical urban setting with a given thermal stability or sensible heat flux, building induced mechanical stresses can become so dominant to drive the resulting stability toward nearly neutral.

![Figure 91: The geometry of the recirculation zone; (a) the recirculation zone totally inside the canyon and (b) the downwind building intercepting the recirculation zone (Berkowicz et al., 1997)]
The model requires inputs of hourly average traffic data, meteorological and urban background concentration data, and the geometry of the street being modelled (the street width, length, and azimuth and also the varying height of buildings within the street).

4.3 OSPM: Mathematical Modelling

The OSPM program calculates pollution in a street canyon based on three different contributions: direct pollution from the traffic to the receptor ($C_d$), the recirculation component of pollutants ($C_r$), and the urban background contribution ($C_b$). As such the total concentration is given by Equation 4.1 (Vardoulakis et al., 2002b).

$$C = C_d + C_r + C_b \quad \text{(Equation 4.1)}$$

The direct contribution from traffic is given by Equation 4.2.

$$C_d = \sqrt{\frac{2Q}{\pi W \sigma_w}} F \quad \text{(Equation 4.2)}$$

where $Q$ is the emission rate, $W$ is the canyon width and $\sigma_w$ is the vertical velocity fluctuation due to mechanical turbulence generated by wind and traffic. Equation 4.3 shows the mathematical definition of the vertical velocity fluctuation:

$$\sigma_w = \sqrt{(\alpha u)^2 + \sigma^2_{wo}} \quad \text{(Equation 4.3)}$$

where $\alpha$ is a proportionality constant (empirically given the value 0.1), $u$ is the wind speed at street level and $\sigma_{wo}$ is the traffic induced turbulence:

$$\sigma_{wo} = b \left(\frac{NV S^2}{W}\right)^{1/2} \quad \text{(Equation 4.4)}$$

In Equation 4.4, $b$ is the aerodynamic drag co-efficient (empirically given the value 0.3), $N$ is the traffic units per unit time, $V$ is the average vehicle speed, $S^2$ is the road surface area occupied by a single vehicle and $W$ is the canyon width (Vardoulakis et al., 2002b).

The simplified box model employed in OSPM for calculating the recirculation co-efficient of the pollutants is shown in Figure 92.
Figure 92: Assumed geometry of the recirculation zone (Buckland, 1998)

Figure 93: Formation of the recirculation zone in a street canyon highlighting the angle of the roof-level wind direction relative to the street axis (Buckland, 1998)

The inflow of pollutant is equivalent to $QL_r/L$ while the outflow can be separated in diffusion through the upper edge and advection through the side edge. Equating the inflow and the outflow gives the mathematical definition of the recirculation component of pollutants:

$$C_r = \frac{Q}{W} \frac{L_r}{\sigma_{wt} L_t + U_L L_{S1} + u L_{S2}}$$

(Equation 4.5)

In Equation 4.5, $L_r$, $L_t$, $L_{S1}$ and $L_{S2}$ are the dimensions of the trapezoidal recirculation zone, and $\sigma_{wt}$ is the ventilation velocity defined by the following equation:

$$\sigma_{wt} = \sqrt{(\lambda U)^2 + F_{roof} \sigma^2_{w0}}$$

Equation 4.6)
U is the roof-top wind speed which represents the rate at which material is assumed to disperse or ventilate out of the canyon, and λ and $F_{\text{root}}$ are both proportionality constants given the empirical values 0.1 and 0.4 respectively (Vardoulakis et al., 2002b).

The width of the recirculation zone cannot exceed that of the street canyon and is defined by the following Equation 4.7:

$$L_r = F_{\text{vortex}}Hr\sin\phi$$  \hspace{1cm} \text{(Equation 4.7)}

Where $F_{\text{vortex}}$ is a proportionality constant given the empirical value of 2, H is the height of the canyon, $\phi$ is the angle of the roof-level wind direction with respect to street axis (Figure 93), and r is a wind speed dependant factor reflecting the strength of the vortex.

The conditions governing the value of r are detailed in Equation 4.8 below:

$$r = \begin{cases} 
1 & \text{if } U > U_{\text{critical}} \\
U/U_{\text{critical}} & \text{otherwise} 
\end{cases}$$  \hspace{1cm} \text{(Equation 4.8)}

The critical wind speed, $U_{\text{critical}}$, for the formation of a vortex in a street is empirically defined as 2m/s (Vardoulakis et al., 2002b).

4.4 OSPM: Data Inputs

OSPM is a parameterised computer model and therefore requires a substantial amount of inputted data (Figure 94) to run an analysis. These can be categorised as:

1. Background concentration of each pollutant;
2. Streets geometry;
3. Emission factors;
4. Traffic data;
5. Meteorology data.
The various data sources and methods employed to assemble these inputs within this project are described in the following sections.

![Diagram of data sources and methods](image)

**Figure 94: Inputs required by OSPM (Berkowicz, 2000a)**

### 4.4.1 Background concentration of air pollutants

OSPM requires that the background concentration are known before any modelling of concentrations of pollutants can start, because it has an effect on the calculated total concentrations. In this study, background concentrations are acquired and modelled following the procedure described in Chapter 3.

### 4.4.2 Streets and buildings geometry

Streets and buildings data for the city of Dublin were supplied by Dublin City Council in GIS format; as such the initial main challenge in using OSPM in this project was to import these street and buildings data into the environmental software. The task was highly complex because of the size of the dataset: almost 11000 road links and more than 40000 buildings. Each road link contained the following information: street name, number of lines, direction of each line, link length, link width, coordinates of the centre of the link and of the two ends, link azimuth, average hourly traffic, traffic speed, etc. Each building contained the following information: coordinates of the vertices, building height, number of stores, address, etc. (Figure 95 and Figure 96).
Figure 95: 2D model of Dublin city in ArcGIS

Figure 96: 3D model of Dublin city in ArcGIS

Figure 97: ArcGIS streets and buildings data cleared from features not needed by OSPM
The task was made even harder because of the very low compatibility between OSPM and the GIS platform. The necessity to avoid a manual input of the data in OSPM is evident and therefore an automated process was created. First, all data features not required by OSPM were removed from the ArcGIS streets and buildings dataset (Figure 97). Each road was then divided in two links wherever it was bisected by a cross-road and segmented in links shorter than 100m. Each building was then associated with the road link on which it is located.

<table>
<thead>
<tr>
<th>street ID</th>
<th>x_center</th>
<th>y_center</th>
<th>traffic type</th>
<th>daily traffic</th>
<th>traffic speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1_2008</td>
<td>316790</td>
<td>232912</td>
<td>Type_C.trf</td>
<td>3368</td>
<td>40</td>
</tr>
<tr>
<td>2_2008</td>
<td>316733</td>
<td>232957</td>
<td>[None]</td>
<td>200</td>
<td>40</td>
</tr>
<tr>
<td>3_2008</td>
<td>314958</td>
<td>232585</td>
<td>Type_C.trf</td>
<td>33521</td>
<td>40</td>
</tr>
<tr>
<td>4_2008</td>
<td>314881</td>
<td>232592</td>
<td>[None]</td>
<td>200</td>
<td>40</td>
</tr>
<tr>
<td>7_2008</td>
<td>314608</td>
<td>232758</td>
<td>[None]</td>
<td>200</td>
<td>40</td>
</tr>
<tr>
<td>8_2008</td>
<td>316656</td>
<td>232842</td>
<td>Type_C.trf</td>
<td>3368</td>
<td>40</td>
</tr>
<tr>
<td>14_2008</td>
<td>314947</td>
<td>233734</td>
<td>Type_C.trf</td>
<td>13096</td>
<td>40</td>
</tr>
<tr>
<td>15_2008</td>
<td>315036</td>
<td>233766</td>
<td>[None]</td>
<td>200</td>
<td>40</td>
</tr>
<tr>
<td>16_2008</td>
<td>314617</td>
<td>233478</td>
<td>Type_C.trf</td>
<td>5441</td>
<td>40</td>
</tr>
</tbody>
</table>

Table 36: Example of road links – OSPM input

An extract from the outcome of the procedure described above is presented in Table 36, where the data is ready for OSPM. As it can be seen in Table 36, some of the road links have the value “[None]” under the header “traffic type”: OSPM excludes the roads with less than 200 traffic units per day from the calculation process. The results of such a culling procedure are shown in Figure 98 below, representing the road network included into OSPM calculations.
Figure 98: Dublin city centre street layout in OSPM

Figure 99: street geometry in OSPM

Figure 99 displays the OSPM street geometry window: each building façade span defined as occupying in a wind sector centred at the link middle point and of angular
width equal to the difference between the azimuths to the façade corners. Each road link can be explored and edited, including the buildings, using this window.

4.4.3 Emission factors

Emission factors can be calculated, edited and visualised by using the emission factor calculation program EmiFact. Emissions from road traffic are approximated as precisely as possible through emission inventory tools such as COPERT IV for the European fleet. In the COPERT methodology, emission factors are specified for numerous vehicle sub-categories covering the majority of European vehicle types (Berkowicz et al., 2003). Emission variation can be affected by a number of factors, such as:

- Vehicle parameters: vehicle class, year, mileage, fuel delivery, etc.;
- Fuel parameters: oxygen content, fuel type and volatility, etc.;
- Vehicle operating conditions: the starting mode of the vehicle (cold/hot), mean vehicle speed, driver behaviour, etc.;
- Vehicle operating environment: humidity, ambient temperature and road.

The key parameters in the EmiFact program are the country setting and the scenario year, which are connected and remand to the appropriate Vehicle List and Fuel List files:

- Vehicle list file: contains a breakdown of the national fleet share for all vehicle types under the European Emission Standard Classifications, which are based on a combination of engine technology, engine capacity and fuel type e.g. petrol passenger cars, Euro IV, 1.4-2.0l;
- Fuel list file: contains data on actual fuel composition along with expressions for the fuel composition correction used for calculation of the emission factors.

These files are used to calculate the emission factors for each vehicle in the national fleet (Berkowicz et al., 2003). Figure 100 shows EmiFact main screen, where the national vehicle fleet share (in percentages) under the various European Emission Standard Classifications is displayed. The fuel tabs above the main table display the fraction of the vehicle fleet made up of each fuel type; petrol, diesel, liquid petroleum gas (LPG). Below the main table is a graphical representation of the emission factor (g/km) against speed (km/hour). To the right of the graph the user can calculate an emission factor for each pollutant for each vehicle class. By selecting the vehicle class and engine capacity from the respective drop-down lists and the pollutant from the pollutant drop-down list, the user can simply adjust the speed value using the up or
down arrow next to the ‘Vehicle Speed (km/hour)’ box to calculate the specific emission factor of a pollutant at a particular speed in the ‘Emission (g/km)’ box. The expression used to calculate the emission factor is shown in the bottom right hand portion of the EmiFact screen (Figure 100), where the parameter ‘V’ is the average vehicle speed (in km/hr). The user can also select the ambient temperature and the percentage Cold Starts to give a more accurate representation of the vehicle emissions in the street being modelled.

The traffic emissions in OSPM are based on the traffic flow in the street (vehicles/hour) and the emission factors (g/vehicle/km). The vehicle emission factors are based on the average vehicle speed which is assumed to be constant over the particular calculation hour; the traffic emissions are thus calculated depending on the Country setting and the selected calculation year. The selected calculation year determines also the vehicle fleet composition, which is calculated using the figures included in the Irish Transport report released each year by the Central Statistics Office (CSO) Ireland.

4.4.4 Traffic data

The traffic data editor program, TrafEdit, allows users to create and modify hourly input traffic files for use with OSPM. The main table in TrafEdit (Figure 101) displays
the traffic volume of each vehicle classification for each hour of the day. The following items can be edited/entered in this table:

- the average speed of vehicles for a particular hour, over short distances (Speed Short) and over longer distances (Speed Long);
- the average percentage of vehicular 'Cold Starts', where the vehicle starts from a stationary position, over a particular hour;
- the Average Daily Traffic (ADT) Split-up table, where the average annual traffic data is split up by vehicle types;
- the total ADT, the number of vehicles per day as an average over the year.

The TrafEdit program (Figure 101) allows the user to save and print the graphs of the average diurnal traffic (vehicles/hour) and average diurnal emission of each pollutant modelled (g/km/hour), shown to the left of the main table, which are automatically created as the user enters values into the main table.

![Figure 101: The TrafEdit program screen with day case “Monday-Thursday” and scenario year “2009”](image)

A traffic file must be created for the whole year which is subdivided into eight ‘Day Cases’ representing days of the week during different months of the year, which can be selected from the Day Case drop-down list. The Day Cases in TrafEdit are as follows:

- Day Case 1: Monday - Thursday [January - June; August - December];
- Day Case 2: Friday [January - June; August - December];
- Day Case 3: Saturday [January - June; August - December];
- Day Case 4: Sunday [January - June; August - December];
- Day Case 5: Monday - Thursday [July];
- Day Case 6: Friday [July];
- Day Case 7: Saturday [July];
- Day Case 8: Sunday [July].

The user has the option of choosing pre-defined daily traffic data contained within the program (associated with Denmark) or inputting user-defined daily traffic data (vehicles per hour). For the purpose of this work, the pre-defined traffic file “Type_C” is used along with the average hourly traffic for each link provided by Dublin City Council. This traffic data is the result of a modelling procedure and a series of assumptions for the minor links based on the data measured by on-street detectors on the main road links.

4.4.5 Meteorology data

OSPM requires hourly values for meteorological and urban background concentration data. As with inputting traffic data, the user has the option of utilising pre-defined meteorological data (pertaining to the climatic conditions of various regions in Denmark) or providing user-defined meteorological and background concentration data. The dataset used in this work included the followings:

- Wind speed and direction;
- Global radiation;
- Temperature;
- Relative humidity;
- PM10 background concentration.

The wind speed and direction are assumed to be representative of the conditions above roof-top level in the street being modelled (Berkowicz et al., 2003). Temperature and global radiation are used to calculate the rate of the modelled chemical reaction between NO and O₃ to produce NO₂, and thus are not strictly necessary as inputs when modelling PM₁₀ concentrations. The urban background concentrations should represent the background roof-top pollution level of the city. The weather data was obtained from the monitoring stations reported in Chapter 3 and the PM10 background concentration levels were modelled following the procedure described in the same chapter.
4.5 OSPM: known model limitations

One of the main disadvantages of the OSPM model is its poor performance under low wind speed conditions, primarily due to the simplifying assumptions the model makes regarding atmospheric processes. At low wind speeds the street vortex disappears and the model’s performance becomes questionable (Berkowicz et al., 1997). The stability of the atmosphere is not well represented by programs like OSPM (Manning et al., 2000). Traffic induced turbulence is assumed to dominate in the street canyon at low wind speeds. The effect of vehicle induced turbulence in OSPM is empirically derived and OSPM is unable to model intermittent fluctuations of wind flow. For this reason OSPM is not recommended for calculating concentrations on timescales shorter than one hour (Holmes and Morawska, 2006).

It could be argued therefore that semi-empirical street canyon models are oversimplistic (Manning et al., 2000). For instance, the model fails to account for non-continuous traffic flow and non-uniform traffic speed. The model also makes no adjustments for excess pollutants created by queuing traffic at traffic lights, intersections, and pedestrian crossings, and the increased ventilation supplied by side streets to the street canyon. The model does allow a percentage of “Cold Starts” to be inputted but, the accuracy of a generic percentage in capturing the complex driving conditions within an urban area is debatable.

OSPM models traffic-related air pollution, including particulates, but focuses largely on direct and recirculated vehicle exhaust emissions. Although vehicle induced turbulence, which can influence resuspension of road dust, is modelled by the program, non-exhaust sources of particulate emissions are largely neglected. In this way the program only captures the primary and secondary exhaust emissions of vehicles and could be prone to underestimation of ambient PM$_{10}$ levels. On the other hand, the factors affecting PM$_{10}$ concentrations, in particular the coarse fraction of particulates, are largely site specific and incorporating their contribution to ambient PM$_{10}$ concentrations into the OSPM program would lead to greater complexity, longer computational time, and require more input data with a greater degree of accuracy.

From his study of AELIOUS, a UK semi-empirical dispersion model directly derived from OSPM, Manning concluded that the model appeared unable to reproduce the peaks in pollution that occur within the street canyon. This is not unexpected as models rarely capture extreme events given their coefficients are heavily biased towards mean concentrations (Manning et al., 2000).
Models such as OSPM are based on numerous simplifying assumptions and this can be viewed as a serious weakness of the semi-empirical approach. Uncertainty in modelling simulations can be considered as the sum of three components: uncertainty due to errors in the model physics, uncertainty due to input data errors, and uncertainty due to stochastic atmospheric processes (e.g. turbulence) (Vardoulakis et al., 2002b). Model uncertainty could be reduced by creating more physically realistic and computationally efficient algorithms and using more accurate monitoring instruments set up at representative locations to eliminate some of the effects of input data errors. However, stochastic turbulent fluctuations are a natural characteristic of the atmosphere and cannot be resolved in the model.

4.6 OSPM: Local model validation

The OSPM model was extensively validated by Ganguly (Ganguly, 2008) as part of his PhD Thesis project in Civil Engineering, in Trinity College Dublin, Ireland. Part of his work consisted of modelling CO and NOx concentrations on Pearse Street, an urban street canyon in the heart of Dublin city centre. In this context, two of the urban street canyon models described above, namely STREET and OSPM, were investigated. Model performance was assessed through comparison with measured CO and NOx concentrations. Modelling analysis was carried out for weekdays during the monitoring period, which extended over an eight-month period for which measured traffic flows and background pollutant concentrations were available. Both models considered are semi empirical in nature, but STREET can be considered an analytical model whereas OSPM is more numerically intensive.

The results of this work are summarised and presented in the following sections of this chapter.

4.6.1 Overview of the validation site

Pearse Street is a four-lane one-way route in the centre of Dublin City, with an approximately East-West orientation and an average daily traffic flow of 60,000 vehicles. The height of the canyon is 16m and width is 21m. An overall layout of the Pearse Street canyon is shown in Figure 102.
Continuous monitoring of hourly roadside CO and NOx concentrations was performed by Trinity College Dublin, hourly meteorological conditions were acquired from Dublin Airport weather station and hourly traffic flows on the street were measured by the traffic control system operated by Dublin City Council. Hourly background concentrations were obtained from an urban air quality monitoring station operated by Dublin City Council at Winetavern Street, which is located approximately 1 km west of Pearse St, and approximately 100m from the nearest trafficked street. A location map outlining both the streets is shown in Figure 103.
All these recorded parameters were used in computing the modelled CO and NOx concentrations. Comparison with measured pollutants concentrations was then carried out based on an eight months data set from May 2006 to December 2006.

In this context the monitored and background data were used for determining the local effects of CO and NOx by subtracting the background data from the monitored data to determine the local effects of CO and NOx concentration. The average monitored and background concentrations were 118.36ppb and 46.91ppb respectively for NOx and 0.67ppm and 0.30ppm for CO. Hence the background concentrations constituted about 40% of the monitored concentrations for NOx and 45% for CO with the remaining being attributed to local effects, which for NOx are primarily traffic emissions. Figure 104 represents a graphical overview of the local traffic impacts (Ganguly, 2008).

![Graphs showing local effects of CO and NOx](image)

**Figure 104: Local effects of CO and NOx**

### 4.6.2 OSPM model results

The composite emission factors used to assess the performances of OSPM at the Pearse street study sites were calculated using vehicle fleet characteristics data for the study year of 2005. Separate daytime and night-time emission factors were employed to allow for the varying effect of congestion on vehicle speeds. The speed versus the composite emission factor (CEF) relationships for CO and NOx are shown in Figure 105 and Figure 106 (Ganguly, 2008) respectively.
Improved model performance can be achieved by capturing the temporal dependence of the traffic source emissions. Because traffic flows followed a strong diurnal pattern, it was possible to derive hourly emission factors (HEFs) representing the average emission rates for each hour of the day. The main criteria considered for HEFs was the speed of vehicle. Since there were no HGVs passing through the study site (as per local traffic management plan) the dependency of hourly emission factors on HGVs was not considered. A speed measurement campaign was carried out on Pearse Street to obtain the average speed profile for the 0700-1800 hours period.

4.6.2.1 CO results

The diurnal profile of CO emission factor is shown in Figure 107 (Ganguly, 2008).
Table 37 presents a statistical comparison of modelled (OSPM) and monitored CO concentrations using CEFs and HEFs (Ganguly, 2008).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Monitored</th>
<th>OSPM (CEF)</th>
<th>OSPM (HEF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.67</td>
<td>0.68</td>
<td>0.68</td>
</tr>
<tr>
<td>standard deviation</td>
<td>0.47</td>
<td>0.62</td>
<td>0.66</td>
</tr>
<tr>
<td>index of agreement</td>
<td>1.00</td>
<td>0.70</td>
<td>0.71</td>
</tr>
<tr>
<td>normalised mean square error</td>
<td>0.00</td>
<td>0.67</td>
<td>0.69</td>
</tr>
<tr>
<td>Pearson’s correlation coeff. R</td>
<td>1.00</td>
<td>0.52</td>
<td>0.55</td>
</tr>
<tr>
<td>fractional bias</td>
<td>0.00</td>
<td>0.01</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 37: Statistical evaluation of CO using CEF and HEF for OSPM

The fractional bias and index of agreement values suggest that there was substantial agreement between measured and predicted values with emission factors. Pearson’s correlation coefficients obtained with both models are relatively high: this suggests that more of the factors affecting short-term concentration variations are correctly represented in the OSPM model. The results imply that relatively a high percentage of
the modelled results were substantially error-free, indicating that the modelling approach followed is suitable for larger scale applications such as that performed in the present project. The

Figure 108 presents the average diurnal variations of the measured and modelled concentrations of CO for OSPM using CEF and HEF. The monitored CO profile displays morning and evening peaks coinciding with periods of maximum travel demand. The evening peak is higher than the morning peak because traffic congestion on Pease Street is usually worst at this time of day, causing average vehicle velocities to fall and unit emissions to rise. This variation is successfully captured by both OSPM models, indicating that the model can capture the variation in personal exposure of individuals located on-street or in adjacent buildings at different times of the day.

4.6.2.2 NOx results

The diurnal profile of NOx emission factor is shown in Figure 109 (Ganguly, 2008).

![Diurnal profile of NOx emission factor](image)

Figure 109: Diurnal profile of NOx emission factor

As performed for CO, a statistical analysis of NOx modelling results was carried out for both CEFs and HEFs using the OSPM model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Monitored</th>
<th>OSPM (CEF)</th>
<th>OSPM (HEF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>117.91</td>
<td>140.96</td>
<td>141.40</td>
</tr>
<tr>
<td>standard deviation</td>
<td>88.54</td>
<td>97.07</td>
<td>101.38</td>
</tr>
<tr>
<td>index of agreement</td>
<td>1.00</td>
<td>0.80</td>
<td>0.79</td>
</tr>
<tr>
<td>normalised mean square error</td>
<td>0.00</td>
<td>0.39</td>
<td>0.42</td>
</tr>
<tr>
<td>Pearson’s correlation coeff. R</td>
<td>1.00</td>
<td>0.65</td>
<td>0.65</td>
</tr>
<tr>
<td>fractional bias</td>
<td>0.00</td>
<td>0.18</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 38: Statistical evaluation of NOx using CEF and HEF for OSPM
The OSPM tended to slightly over predict the mean measured concentrations with both CEFs and HEFs, as indicated by the fractional bias results in Table 38, and the Pearson’s correlation coefficients are not as high as for CO.

Figure 110: Diurnal profile variation for NOx using CEF and HEF

The monitored NOx profile displays morning and evening peaks coinciding with periods of maximum travel demand. The evening peak is slightly lower than the morning peak and this variation is successfully captured by the model using CEF (Figure 110). When compared with monitored data, the concentrations calculated by Ganguly (Ganguly, 2008) using OSPM successfully predicted observed variations in air quality. From the statistical parameters presented it is quite evident the use of CEF and HEF were both successful in predicting long-term average concentrations, but that an model performance varied slightly by pollutant, which is likely to be a reflection of emission factor accuracy. As such, in the present project, average hourly emission factors are used for modelling the PM10 concentrations in urban street canyons.

4.7 OSPM: Urban area model validation

The OSPM model is further tested and its results validated against the data collected at the Environmental Protection Agency’s permanent PM10 monitoring stations in Dublin city (Figure 111). The model can only be validated on a daily basis because of the resolution of the available measured data. The procedure employed illustrated in the following sections of this Chapter, where the PM10 daily average concentration levels are modelled for March 2010 for two different locations within the Great Dublin Area:

- Winetavern Street: this location was chosen as a test implementation of the urban configuration of the model;
- Ballyfermott Street: location was chosen as a test implementation of the sub-urban configuration of the model.
The Operational Street Pollution Model requires the following inputs:

- Daily averages for the PM10 background concentration levels;
- Hourly values for temperature, humidity, solar radiation, wind speed and direction;
- Hourly values for traffic composition and volume.

The preparation of these datasets is described and discussed in the following sections of this chapter.

### 4.7.1 Model set-up

#### 4.7.1.1 Background concentration levels

The modelling procedure used to determine the background concentration at different locations in Dublin is discussed in detail in Chapter 3. The background PM10 monitoring stations active in the Dublin area are listed and presented according to their typology in Table 39 below.

<table>
<thead>
<tr>
<th>Station name</th>
<th>Typology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dun Laoghaire</td>
<td>Suburban Background</td>
</tr>
<tr>
<td>Phoenix Park</td>
<td>Suburban Background</td>
</tr>
<tr>
<td>Rathmines</td>
<td>Urban Background</td>
</tr>
<tr>
<td>Tallaght</td>
<td>Suburban Background</td>
</tr>
</tbody>
</table>

Table 39: PM10 background monitoring stations in the Dublin area divided by typology
An overview and detailed description of the station locations is given in Chapter 3. Figure 112 graphically summarises their distribution within the Dublin area.

The closest background air quality monitoring station to Winetavern Street is Phoenix Park, which belongs to the “sub-urban background” typology, according to the Irish Environmental Protection Agency’s classification presented in Chapter 3. This background monitoring station is also the most suitable to model the PM10 background concentration levels in the area around Ballyfermot road, because of its proximity. A plot of the time-series of the modelled PM10 concentration levels for the
station is shown in Figure 113 below. The modelled background levels range between 9.5 and 16 µg/m³. This time-series is used as an input to the urban traffic model, which is assessed and validated in the following sections.

4.7.1.2 Weather

The hourly time-series of weather variables required by the OSPM model are: temperature (°C), relative humidity (fraction of 1), solar radiation (kWh/m²), wind speed (knots) and wind direction (degrees from the true north). The weather stations’ used for modelling the PM10 concentration levels at Winetavern Street and Ballyfermot Road with OSPM are Phoenix Park (temperature, humidity and solar radiation) and Dublin Airport (wind speed and direction), described in Chapter 3. The hourly time-series for these weather variables are presented and summarised in the plots from Figure 114 to Figure 118 below.

![Temperature plot](image1)

Figure 114: Temperature hourly variation in March 2010 recorded by Phoenix Park weather station

![Humidity plot](image2)

Figure 115: Relative humidity hourly variation in March 2010 recorded by Phoenix Park weather station
4.7.1.3 Traffic volume and composition

The traffic volumes on streets in the vicinity of Winetavern Street and Ballyfermot Road were obtained from Dublin City Council, who provided the hourly average traffic volumes on each road link throughout most of the Dublin urban area.
The traffic composition is obtained from the Irish Bulletin of Vehicle and Driver Statistics (Department of Transport, 2010) for the year 2010, which is available for downloading from the Irish Department of Transport website: http://www.transport.ie. The report provides a wide range of statistics about the Irish car fleet composition and also separates the vehicles by category, engine size, year of production, mileage, etc, which are all extremely valuable for modelling the road traffic emissions. It was not possible to retrieve information about the traffic composition variation for each road link of the Dublin Area because such data was not available from any of the Irish Statutory Bodies at the time of this work. As such, the hourly variation of the volumes for each vehicle category for each road link is obtained by using the OSPM predefined “Traffic Variation” files (Berkowicz et al., 2003). A pre-defined traffic file is data set with information on the relative distribution of traffic flow over time for all types of vehicles, such as e.g. passenger cars, vans, trucks and buses, indicated as the fraction of the daily total. This file provides a different relative distribution of vehicle types for different days of the week and months of the year - e.g. a case for Fridays in July. Clearly there is scope for including more locally-relevant data should these become available later.

In order to use the pre-defined traffic data for a particular street, the average daily traffic must be specified (vehicles per day as an average over the year), as must an average travel speed for the street. This data was obtained, as mentioned before, from Dublin City Council. The travel speed is an average speed for a street section of length 100-200 meter, close to the receptors.

The pre-defined traffic file used for Winetavern Street and Ballyfermot Road is the "Type_C" (Figure 119) which is recommended for "Larger residential street - distribution street in residential area", "Fordelingsveje i boligområder" from the OSPM manual (Berkowicz et al., 2003).

Figure 119 shows the window with these traffic patterns, which can all be modified using TrafEdit, a Traffic Data Editor extension included into the OSPM software package.
4.7.2 Model results

The Operational Street Pollution Model was run for Ballyfermot Road and Winetavern Street for each hour in the month of March 2010 and the daily average PM10 concentration levels calculated and compared with the correspondent measured levels from the permanent monitoring station.

The comparison between the time-series of measured and modelled levels for Ballyfermot Road is shown in Figure 120 below.

![Figure 120: Comparison between measured and modelled PM10 daily average levels at Ballyfermot Road for the month of March 2010](image-url)
As it can be noticed from the plot above, the two time-series have similar trends and behaviours, except that the OSPM model was not able to reproduce in magnitude the two peaks on the 4th and 9th of March. This can be explained by bearing in mind one of the model limitations: the Operational Street Pollution Model doesn’t take into account the contributions from nearby roads when modelling the dispersion of pollutants from a road link. It should be observed that these two peaks occur with winds blowing from North, North-East and East, which suggests that they may be due to the contribution from the nearby N4 Chapelizod Bypass, which is a heavy traffic motorway on the North-East of Ballyfermot Road (Figure 121). As such, the monitored peaks may be explained by a conjunction of particularly high volumes of traffic on the N4 and favourable weather conditions, i.e. the wind blowing from the motorway to the monitoring station. OSPM is not designed to take into account such conditions generating from nearby roads.

Despite of the discussed model limitation, the coefficient of correlation between measured and modelled concentration levels is reasonably high, with an $R^2$ equal to 0.71 (Figure 122).
Figure 122: Correlation between measured and modelled PM10 daily average levels at Ballyfermot Road for the month of March 2010

The comparison between the time-series of measured and modelled levels for Winetavern Street is shown in Figure 123 below.

Figure 123: Comparison between measured and modelled PM10 daily average levels at Winetavern Street for the month of March 2010

The time series of modelled and measured concentrations display very similar trends and behaviour, but again the OSPM model fails to reproduce in magnitude the three peaks on the 4th, 9th and 16th of March. As mentioned before, OSPM doesn’t take into account the contributions from nearby roads while modelling the dispersion of pollutants for a road link. Furthermore, these are the days when high concentrations were observed at Ballyfermot too. This suggests that it might have been a regional effect, i.e. high background concentrations across all of Dublin. This is confirmed by the time series (Figure 124) of PM10 daily averages measured by Dun Laoghaire...
monitoring station in March 2010: the peaks around the 4\textsuperscript{th}, the 9\textsuperscript{th} and the 16\textsuperscript{th} are evident also here.

![Graph showing PM10 daily averages in March 2010 recorded by Dun Laoghaire monitoring station.](image1.png)

Figure 124: PM10 daily averages in March 2010 recorded by Dun Laoghaire monitoring station

![Map showing area around the location of Winetavern monitoring station.](image2.png)

Figure 125: Area around the location of Winetavern monitoring station

The PM10 concentration levels recorded in Winetavern Street are greatly affected by the contribution from the road traffic on the adjacent Quays (N1 and N4 along the river) as observed in Chapter 3. It should be observed that the first two peaks happen for wind predominantly blowing from North, which suggests that they may be due to the contribution from the Quays, which is a heavy traffic road on the North of
Winetavern Street (Figure 125). The third peak happens for wind predominantly blowing from the South, which is where the busy crossroad between the N81 and High Street (circled in yellow) is located. The monitored peaks on the 4th and 9th of March may be thus explained by a conjunction of particularly high volumes of traffic on the Quays and favourable weather conditions, while the peak on the 16th may be due to heavy traffic in correspondence of the crossroad between the N81 and High Street.

Despite of the discussed model limitation, the coefficient of correlation between measured and modelled concentration levels is reasonably high, with an $R^2$ equal to 0.7 (Figure 126).

![Winetavern Street- correlation between measured and modelled PM10 in March 2010](image1)

**Figure 126:** Correlation between measured and modelled PM10 daily average levels at Winetavern Street for the month of March 2010

![Atmospheric Stability Class](image2)

**Figure 127:** Hourly atmospheric stability class in March 2010 recorded by Dublin Airport weather station

The fact that the peaks in measured PM10 concentration levels on the 4th and 9th of March are recorded by both monitoring stations suggests that a more accurate
analysis of the weather conditions for these two days is necessary. The time series for the hourly variation of the atmospheric stability class recorded by Dublin Airport weather station is reported in Figure 127 below. The values range between 1 and 6, which correspond to the Pasquill stability classes reported in Table 40.

The stability categories were developed by Pasquill in 1961 (F.Pasquill, 1961) and are based on wind speed, insolation, and extent of cloud cover and they give an indication about the amount of turbulence in the ambient atmosphere, which has a major effect on the dispersion of air pollution (Beychok, 2005) Atmospheric stability/turbulence was categorised into six stability classes listed in Table 40 below.

<table>
<thead>
<tr>
<th>Stability class</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>very unstable</td>
</tr>
<tr>
<td>2</td>
<td>unstable</td>
</tr>
<tr>
<td>3</td>
<td>slightly unstable</td>
</tr>
<tr>
<td>4</td>
<td>neutral</td>
</tr>
<tr>
<td>5</td>
<td>slightly stable</td>
</tr>
<tr>
<td>6</td>
<td>stable</td>
</tr>
</tbody>
</table>

Table 40: The Pasquill stability classes (F.Pasquill, 1961)

Stability is related to the tendency of the atmosphere to resist or enhance vertical motion and thus turbulence, which is by far the most important parameter affecting dilution of a pollutant. A neutral atmosphere neither enhances nor inhibits mechanical turbulence, an unstable atmosphere enhances turbulence, whereas a stable atmosphere inhibits mechanical turbulence. In general, the more unstable the atmosphere, the greater the dilution. The atmospheric stability is predominantly “stable” on the around the 4th and 9th of March, which explains why the peaks are recorded by both monitoring stations on the same days, even if the locations are in different part of Dublin.

Figure 128: Hourly rainfall in March 2010 recorded by Phoenix Park weather station
The magnitude of the peaks may be also due to the complete lack of precipitations during the first half of March, as shown in Figure 128.

4.8 OSPM: Merits and faults

The results presented in the previous section and the generally good observed correlations between measured and monitored PM10 concentration levels demonstrate that OSPM performs well in modelling the dispersion of pollutants from the road links selected as test cases. Another merit of the model is the short calculation times: it was possible to model the daily average concentration levels for a month for a link in approximately 15-20 seconds with an old Pentium 4 machine with 2 Gigabytes of RAM. Moreover, the preparation of the datasets necessary to run the OSPM, such as the weather file and the street file, is quite rapid if the pre-determined traffic distribution file is adopted. The model was also run to calculate daily average PM10 concentration levels for the whole Dublin City centre, which includes approximately 10,000 road links. On the same machine, the calculation time was approximately one hour for a simulation length of 24 hours. The results of this simulation were imported into ArcGIS and are presented in Figure 129.

On the other hand OSPM works well for single road and multiple roads but it is not recommended to run it for an entire city because, when it calculates the concentration of pollutants along one road, it doesn't take into account the contributions of the neighbouring roads. This is evident from the tests carried out for Ballyfermot Road and Winetavern Street, presented in the previous section. It is also clear that it is necessary to analyse carefully the weather conditions during the modelled period because anomalous (for Ireland) weather conditions, such as long periods of “stable” atmospheric turbulence and absent rainfall, could reduce pollutant dispersion and removal and result in peaks in the concentration levels that the model is not able to predict. It has also to be borne in mind that OSPM doesn't perform well in open areas, such as areas with spare houses or semi-detached houses with gardens, such as big part of Dublin. Finally the model considers streets with less than 200 units of traffic per day as background streets and they are not included into the calculation process.
Figure 129: Daily average PM10 concentration levels calculated with OSPM for the 30th of September 2010 for Dublin City centre.
CHAPTER 5 – Suburban traffic pollution modelling
5. Introduction

This chapter describes the modelling of PM10 concentration levels resulting from road traffic in the suburban areas of Dublin. The modelling procedure described here is also applied to the road links in the urban area of Dublin which cannot be classified as "street canyons" because of their geometry (buildings / street width ratio) and therefore the OSPM model is not suitable for. Dublin city is characterised by different street geometries depending on the area, but most of them have a low building height / street width ratio because of the low rise houses and the presence of front gardens and footpaths flanked by trees. An example of this set-up is presented in Figure 130 below, which shows the area around the Stillorgan Flyover Bridge in proximity of University College Dublin (UCD).

![Figure 130: suburban area in Dublin](image)

5.1 Overview of existing models

An overview of the models that may be suitable for modelling the dispersion of air pollutants from road traffic in such areas is presented in this section. The models are separated in two categories: screening and advanced models.
5.1.1 Screening Models

As mentioned in the previous sections, screening models produce estimates of "worst-case" concentrations scenarios and they are usually applied before a refined air quality model to determine if refined modelling is needed. They tend to be easy-to-run, quick and conservative: they tend to over-predict air contaminant concentrations which are calculated considering a standard set of meteorological conditions and simplified assumptions/representations of source-receptor geometry. They have limited capabilities for defining receptor locations, pollutant averaging times and output options. Should the results of a screening model predict an exceedance of the air quality standards then a more advanced model should be used.

The following models are described in the sections below:

1. Design Manual for Roads and Bridges (DMRB): This screening method was formulated by the former UK Department of Transport and it was originally published in August 1994; a revision was made at the end of 1997 and 1999, which included features also applicable to urban situations. The method is a simple procedure originally based on tables and nomograms and it gives a preliminary indication of air quality near roads;

2. Calculation of Air pollution from Road traffic (CAR): it is a simple parameterized model for the determination of air quality alongside roads (including street canyons) in cities (Eerens et al., 1993). This is a screening model formulated to support the implementation of air quality legislation in the Netherlands and it is a parameterisation of the Gaussian TNO traffic model proposed by Eerens et al.;

3. General Finite Line Source Model (GFLSM): it is an analytical model for prediction of the dispersion of pollutants downwind of a straight finite length of roadway (Luhar and Patil, 1989). The mathematical approach is based on a coordinate transformation which reconciles the local highway alignment with the wind direction.

5.1.1.1 Critical comparison

The DMRB method requires information on vehicle flow, vehicle speed and receptor-road distances and is more suited to rural motorways and trunk roads than urban traffic conditions. It contains an useful database of vehicular emission factors for future years (Ganguly, 2008). The model assumes a fixed set of meteorological conditions (a low wind speed and neutral atmospheric stability; the wind direction is assumed to be evenly distributed around the compass). Assumptions about background pollutant levels are included in the calculations.
CAR takes into account the configuration of buildings and trees along the road and photochemistry. Emission factors and the wind speed are defined by the user. The model supports the implementation of air quality decrees under the Air Pollution Act by provincial and municipal authorities.

The main disadvantage of the GFLSM is the limited receptor co-ordinates for which concentrations can be calculated: it requires the receptor to be located at 90 degrees to the segment of the road considered and the length of line source should be at least three times the distance between the receptor location and road (Gokhale and Khare, 2004). However, if these conditions are met, the analytical solution employed is exact and the accuracy of the results depends only on the accuracy of the input parameters employed, including source strength and metrological conditions.

GFLSM has the advantage of the simplicity of its application as it is an analytical solution of the Gaussian equation (Ganguly, 2008). For this reason, it is possible to fully integrate the GFLSM with ArcGIS by including it in a custom GIS tool (ArcTool).

### 5.1.2 Advanced models

Advanced air dispersion models are usually based on more modern scientific theories and more complex mathematical formulations than screening dispersion models. Advanced models can assess the impact of complex distributions of multiple sources and numerous buildings. Detailed input data regarding meteorology, land use and terrain are required by these models in order to allow them to represent the atmospheric processes contributing to pollutant dispersion. Significant data pre-processing is often required to prepare the input files used by these models.

The following models are presented in the sections below:

1. ADMS Urban: ADMS-Urban is basically an extension of the ADMS2 industrial model, combining a line and area source model with a street canyon model, a chemistry model for nitrogen, sulphur and ozone and a traffic emissions database (Robins et al., 1997). It was presented in more detail in Chapter 4;

2. CAR-FMI: Harkonen et al. developed the CAR-FMI (Contaminants in the Air from a Road – Finnish Meteorological Institute) model for regulatory purposes in 1995 (Harkonen et al., 1995). The model evaluates atmospheric dispersion and chemical transformation from a network of line sources in local scale;

3. CALINE 4: CAIifornia LINE Source Dispersion Model was developed by the California Department of Transportation (Caltrans) in 1988 (Benson, 1988). It is the standard modelling program used by Caltrans to assess carbon monoxide impacts near transportation facilities;
4. HIWAY2: The computer model HIWAY 2 was created by W.B.Petersen in the 1980. It is used to estimate the concentrations of nonreactive pollutants from highway traffic and applied to determine air pollution concentrations at receptor locations downwind of 'at grade' and 'cut section' highways located in relatively uncomplicated terrain (Petersen, 1980);

5. STREET: it is an empirical screening model which predicts concentrations on both the windward and leeward sides of the road. It includes an empirical parameter constant ‘K’ which should be validated for each individual application site. Qin and Kot (Qin and Kot, 1993) estimated the value of K to be 6 for their work in an asymmetric street canyon in China whereas Bogo et al. (Bogo et al., 2001) reportedly used a K value of 8 for a similar study in Buenos Aires. Vardoulakis et al. (Vardoulakis et al., 2002b) used a K value of 7 for application in a street canyon in Paris. The model is presented in further detail in Chapter 4.

5.1.2 Critical comparison

The CAR model requires information on emissions and meteorological parameters on an hourly basis. The dispersion model is based on the GFLSM approach (Luhar and Patil, 1989) which prevents need of numerical solutions as required in HIWAY and CALINE4 (Kukkonen et al., 2001b). The modelling system includes an emission model, a dispersion model, statistical analysis of the computed time series of concentrations and a graphical Windows-based user interface. The meteorological data for the model is evaluated by the FMI meteorological pre-processor.

For the CAR model, the Gaussian dispersion parameters are functions of boundary layer variables, and their dependence on source height is also accounted (Kukkonen et al., 1997): as such, the data preparation time for this model can be considerably longer than for the others assessed here.

Both the CAR and the CALINE 4 models include the influence of a finite mixing height on the plume dispersion. The HIWAY2 model was developed for situations where horizontal wind flow occurs, thus it cannot consider complex terrain or large obstructions, such as buildings or large trees, to the flow.

The limitations of these models are that they all adopt the Gaussian plume model, which assumes stationary meteorology and emissions during the dispersion. They also do not allow for the influence on dispersion of individual buildings and obstacles, or inhomogeneous terrain and they can be used in moderately flat regions.
5.2 Previous Research on Highway Dispersion Modelling in Ireland

This section gives an overview of a number of recent studies into transported related air pollution conducted in Ireland and it is adapted from (Ganguly, 2008).

An air quality modelling systems (AQMS) developed by Reynolds and Broderick (Reynolds and Broderick, 2000) for metropolitan areas in Europe was applied to Dublin, with the primary objective of incorporating temporal and spatial distributions of traffic and air quality data. The AQMS consisted of three modules: a traffic module, an emission module and dispersion module integrated within a GIS system.

Budd (Budd, 2004) modelled the concentrations of CO, NOx and PM10 at the M4 motorways at Leixlip and (in association with researchers in NUI Galway) at a busy roundabout in intersection in Galway using DMRB and CALINE4 models. The background concentrations were taken to be the average of hourly pollutant concentrations when the traffic flow had least influence. The diurnal variation profiles followed the traffic profile at the motorway site. The DMRB model over-predicted pollutant concentrations at both sites as expected because it uses a worst case meteorological scenario.

Ganguly and Broderick (Ganguly and Broderick, 2009) used an updated version of the DMRB model and observed that this version of the screen DMRB model was successful in predicting reduced concentrations for NOx and particulate matter at the motorway. In the roundabout problem, certain modifications were made in relation to the earlier version suggested in the DMRB manual. This led to improved but still over-predicted pollutant concentrations estimates. Ganguly and Broderick (Ganguly and Broderick, 2009, Ganguly, 2008) further modelled the CO concentrations using GFLSM and showed that it was a better modelling option than CALINE4.

O’Donoghue (O’Donoghue, 2004) carried out a hydrocarbon monitoring campaign at the M50 site in Dublin. This campaign was conducted for approximately eight months. Hydrocarbon peak concentrations were shown to be higher in winter than in summer and heavily dependent on wind speeds. Traffic related compounds such as acetylene, ethene and aromatic hydrocarbons showed strong diurnal profiles associated with M50 traffic flow. Highest concentrations were shown to occur with parallel or near parallel wind directions relative to the motorway orientation. High traffic flow and lower speeds contributed to this observation. Analysis of concentrations obtained from non-motorway directional winds showed no obvious source effects, confirming that hydrocarbon concentration in the area was are primarily sourced from the M50 roadway. O’Donoghue (O’Donoghue, 2004) used the CALINE4 model to predict the
observed concentrations and reported that CALINE4 modelled results were a good match for the observed hydrocarbon concentrations. Ganguly and Broderick (Ganguly and Broderick, 2009, Ganguly, 2008) showed that the use of GFLSM model achieved better predictions of these hydrocarbon concentrations than CALINE4.

Delaney (Delaney, 2006) carried out an assessment study at a semi-rural site (Monasterevin, Co. Kildare) from October 2004 to January 2005 with the prime objective of assessing the change in traffic impacts due to opening of a new town bypass. The pollutants studied were CO, NOx HC’s and PM10. Predictions by the Gaussian model CALINE4 were compared with the observed changes in pollutant concentrations. Ganguly and Broderick (Ganguly, 2008) used the GFLSM also model to model these changes.

5.3 General Finite Line Source Model

The GFLSM is chosen for the modelling of suburban traffic related PM10 concentration levels because of the experience and satisfactory results obtained in the past years in the Civil Engineering Department of Trinity College Dublin (TCD) by Ganguly and Broderick (2008, 2009). GFLSM has been chosen for the GIS air quality model generated as part of this work also because of the simplicity of its application as it is an analytical solution of the Gaussian equation (Ganguly, 2008). For this reason, it is possible to fully integrate the GFLSM with ArcGIS by including it in a custom GIS tool (ArcTool), as discussed in detail in this Chapter.

The General Finite Line Source Model (GFLSM) for vehicular pollution prediction was evaluated by Luhar and Patil (Luhar and Patil, 1989) using statistical analysis of predicted and observed CO concentrations in Mumbai, and particulate concentrations in New York. In the CO evaluation, the GFLSM was seen to perform better than the GM, CALINE3 and HIWAY2 models. In the particulate evaluation, the GFLSM model outperformed the GM model. Khare and Sharma (Khare and Sharma, 1999) assessed the performance of the GFLSM by comparing predicted CO concentrations with measured values at different traffic intersections in Delhi. Khare and Sharma modified the GFLSM model to create the DFLSM by removing the error function term and obtained improved results. Nagendra et al. (Nagendra et al., 2004) applied the GFLSM at busy traffic intersections at Bangalore, India and found that model performance was good when compared with monitored data. Gokhale and Patil (Gokhale and Patil, 2004) used the GFLSM model to predict concentrations of size separated particular matter below 10μm size from vehicular exhausts using a modified GFLSM model for particulates and obtained a good correlation between observed and
predicted values for size ranges below 4.7μm. Wang et al. (Wang et al., 2006) measured simultaneously the concentrations of carbon monoxide CO and fine particles PM$_{2.5}$ pollutants and the traffic and ambient atmospheric conditions at three selected local urban road sites in Hong Kong. A high level of agreement was found between the measured and modelled with GFSLM CO and PM$_{2.5}$ data. The assessment also demonstrated that the developed local general finite line source model had the capability of reasonably predicting the characteristics of gaseous and particle pollutant dispersion from on-road vehicles for the local urban air quality. Ganguly and Broderick (Ganguly, 2008) assessed the relative performance of two atmospheric dispersion models, GFSLM and CALINE4, using carbon monoxide levels measured adjacent to a four lane motorway. The CALINE4 model had already been validated for Irish motorway conditions and the performance of the GFSLM was looked at under similar conditions. The performance of the model was tested using composite and hourly emission factors calculated with the COPERT III methodology. The model performed best for neutral stability conditions and wind speeds greater than 0.5 m/s. The modelled and predicted data were fitted to statistical distributions and it was found that the modelled and predicted data follow the Weibull distribution (Weibull, 1951).

5.3.1 The roadside dispersion model GFSLM

The General Finite Line Source Model is based on the Gaussian diffusion equation and it is formulated so that it could be used for any orientation of wind direction with roadway; it also does not have the infinite line source constraint of the General Motors (GM) model Luhar and Patil (Luhar and Patil, 1989). The GFSLM is also modified to predict particulate concentrations by incorporating some simple corrections. Its solution is based on the modification of an equation derived by Csanady (Csanady, 1972) that allowed receptor concentrations due to emissions from a finite length line source to be calculated for perpendicular winds. The modification by Luhar and Patil (Luhar and Patil, 1989) extended this solution to include all possible wind road angles. The main advantage of this model lies in the simplicity of its application. The main disadvantage of the model is the limited receptor coordinates for which concentrations can be calculated (Ganguly, 2008). The GFSLM requires the receptor to be located at 90 degrees to the segment of road considered, and the length of the line source should be at least three times the distance between the receptor location and road (Gokhale and Khare, 2004). The GFSLM uses the following Equation 5.1 to calculate the contribution of a line source to ambient concentrations (Luhar and Patil, 1989):
\[ C = \frac{Q_1}{2\sqrt{2\pi}\sigma_z(u \sin \theta + u_0)} \left[ \exp \left( -\frac{(z-h)^2}{2\sigma_z^2} \right) + \exp \left( -\frac{(z+h)^2}{2\sigma_z^2} \right) \right], \]

\[ x \left[ \text{erf} \left( \frac{\sin \theta (p - y) - x \cos \theta}{\sqrt{2}\sigma_y} \right) + \text{erf} \left( \frac{\sin \theta (p + y) - x \cos \theta}{\sqrt{2}\sigma_y} \right) \right] \]

(Equation 5.1)

where \( C \) is the receptor concentration, \( Q_1 \) is the source strength per unit length, \( u \) is the average wind speed, \( \theta \) is the angle between the wind direction and the road varying between 0-180 degrees, \( x \), \( y \) and \( z \) are the receptor co-ordinates relative to an origin located at the midpoint of the line source, \( h \) is the effective source height, \( p \) is the half length of the line source, \( \text{erf} \) is the error function and \( \sigma_z \) and \( \sigma_y \) are the vertical and horizontal dispersion coefficients. The term \( u_0 \) is a wind speed correction to account for the effects of traffic wake and assumes different values for different stability classes as suggested in the GM model (Chock, 1978). The vertical and horizontal dispersion coefficients are dependent on the downwind distance \( x \) and the Pasquill (F.Pasquill, 1961) stability class. The Briggs urban dispersion coefficients are employed for \( \sigma_z \) and \( \sigma_y \) (Zannetti, 1990).

Measured traffic and meteorological data, and roadway geometry are used as the input parameters to the model. The emission rate, \( Q_1 \) is determined from composite vehicle emission factors calculated using the COPERT methodology (Kouridis et al., 2000); as described later. A schematic representation of the GFLSM is shown in Figure 131.

![Figure 131: Schematic representation of the GFLSM model (Ganguly, 2008)](image)

From the Luhar and Patil equation reported above it is clearly evident that the concentration values predicted using the GFLSM model are heavily influenced by the ‘y’ coordinate of the receptor.

191
5.3.2 Sensitivity analysis of GFLSM model

Sensitivity analysis of the GFLSM model was carried out by Ganguly (Ganguly, 2008) by performing statistical computations on the predicted data obtained under different stability categories and wind speed conditions. The results are shown in Table 41 and Table 42.

<table>
<thead>
<tr>
<th>Parameters/Stability class</th>
<th>Unstable</th>
<th>Stable</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>IA</td>
<td>0.39</td>
<td>0.47</td>
<td>0.53</td>
</tr>
<tr>
<td>NMSE</td>
<td>0.40</td>
<td>0.43</td>
<td>0.10</td>
</tr>
<tr>
<td>R</td>
<td>0.30</td>
<td>0.14</td>
<td>0.49</td>
</tr>
<tr>
<td>FB</td>
<td>-0.16</td>
<td>-0.34</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Table 41: Stability class analysis of CO data

<table>
<thead>
<tr>
<th>Parameters/Wind speed</th>
<th>&lt;0.5m/s</th>
<th>0.5-2m/s</th>
<th>&gt;2m/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>IA</td>
<td>0.35</td>
<td>0.39</td>
<td>0.50</td>
</tr>
<tr>
<td>NMSE</td>
<td>0.68</td>
<td>0.23</td>
<td>0.12</td>
</tr>
<tr>
<td>R</td>
<td>-0.09</td>
<td>0.11</td>
<td>0.50</td>
</tr>
<tr>
<td>FB</td>
<td>-0.40</td>
<td>0.10</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 42: Wind Sensitivity analysis of the CO data

For the stability class analysis, Ganguly sorted the entire monitored and predicted data sets according to the stability conditions pertaining for each hour and then he considered three classifications: unstable, stable and neutral. Ganguly also observed the followings:

- The model performed best under neutral conditions;
- The relatively large under-prediction of mean concentration during stable conditions, indicated by the FB value of -0.34, was attributable to the use of a constant background concentration;

As regards the wind sensitivity analysis study, the monitored and modelled data were statistically analysed after being sorted for three different wind speed classes: less than 0.5m/s, between 0.5 and 2m/s and greater than 2 m/s. This test resulted in the following observations:

- The GFLSM model performed best for wind speeds above 2m/s;
- The GFLSM performed least well for wind speeds less than 0.5m/s;
- The low R value observed for the low wind speeds was attributed to the poor definition of wind direction under these conditions, which affects model reliability;
The FB value displayed during low wind conditions was in line with that observed for stable conditions, which are normally associated with low wind speeds.

5.4 The GFLSM ArcGIS model

The creation of an air quality model for mobile sources is a complex task and involves the completion of different procedures steps (Federal Highway Administration, 2000) which can be summarised into the following broad actions:

1. Estimation of traffic activity and fleet composition: this can be achieved by using a traffic model or traffic counters.
2. Estimation of vehicular emission rates in grams per unit time or distance (or both) computed using an emission rate model.
3. Computation of ambient pollutant concentration using a pollutant dispersion model.

The procedure described above is followed for the creation of the sub-urban traffic pollution model and it is described in detail in the following sections.

5.4.1 Vehicles Emission Factors

The combined assessment of the traffic activity/fleet composition and the vehicular emission rates gives as output a spatial and temporal estimation of emissions, which leads to the development of an emission inventory (Barlow et al., 2007, Boulter et al., 2007). In this project, traffic activity is quantified using traffic count data obtained by the Irish National Road Authority and Dublin County Council, and COPERT 4, a specialised roadway emission model, is used to calculate the emission factors for roadway traffic.

5.4.1.1 COPERT 4

COPERT 4 (Ntziachristos et al., 2009) is a software package which allows the user to calculate air pollutant and greenhouse gas emissions from road transport. It became the official version of COPERT since November 2006, replacing COPERT III. The new version 9.0 of the software, which has been available since October 2011, covers all vehicle technologies up to Euro 6 for passenger cars, Euro VI for heavy duty vehicles, and Euro 3 for motorcycles. The development of COPERT is coordinated by the European Environment Agency (EEA), in the framework of the activities of the European Topic Centre for Air Pollution and Climate Change Mitigation. The European Commission's Joint Research Centre manages the scientific development
of the model. COPERT has been developed for official road transport emission inventory preparation in EEA member countries and its methodology is part of the EMEP/EEA air pollutant emission inventory guidebook (European Environment Agency (E.E.A.), 2009) for the calculation of air pollutant emissions and is consistent with the 2006 IPCC Guidelines (Task Force on National Greenhouse Gas Inventories (T.F.I.), 2006) for the calculation of greenhouse gas emissions. The EMEP/EEA air pollutant emission inventory guidebook (formerly called the EMEP CORINAIR emission inventory guidebook) provides guidance on estimating emissions from both anthropogenic and natural emission sources. It is designed to facilitate reporting of emission inventories by countries to the UNECE Convention on Long-range Transboundary Air Pollution and the EU National Emission Ceilings Directive. The use of a software tool to calculate road transport emissions allows for a transparent and standardized, hence consistent and comparable data collecting and emissions reporting procedure, in accordance with the requirements of international conventions and protocols and EU legislation (http://www.emisia.com/copert/).

The emission equations and methodology embedded in this programme were employed to estimate composite PM10 emission factors for the traffic conditions encountered in this study. The COPERT methodology covers a wide range of vehicles, divided into five primary categories, and sub-divided by model year, emission-reduction technology, engine volume, weight and fuel type. Detailed information about the Irish car fleet composition for the years 2009, 2010 and 2011 were gathered from the Central Statistics Office (CSO) and from the Irish Bulletin of Vehicle and Driver Statistics for the various years, obtained from Department of the Environment, Heritage and Local Government (DoEHLG). This was necessary in order to calculate representative composite emissions factors.

A weighted emission factor of 0.01g/km for PM10 was estimated by using the characteristics of the traffic fleet and the estimated annual kilometres travelled by each vehicle type (Reynolds and Broderick, 2000) and by assuming a mean vehicle speed of 40km/hr (Boulter, 2009).

5.4.2 GFLSM excel model

A GFLSM model is implemented using Microsoft Excel. The Excel file is divided in a number of worksheets, within which different parts of the calculations are performed.

The worksheets are the followings:

1. **Streets**: this sheet holds the characteristics and information for all the street-links necessary for the modelling calculations, such as the link centre
coordinates, the length and width of the link, the angle from the true north and the traffic per hour;

2. **Input**: the user can input here all the weather data, such as wind speed (m/s) and direction from the true north, ambient temperature (°C), stability class and also the distance between each node of the receptors’ grid, at which the pollutant’s concentration levels are calculated;

3. **Grid**: a grid of 30 x 30 nodes is built around the centre of the link. The coordinates \((x_i, y_i)\) of each \(i\) node are modified and updated according to the position \((x, y)\) of the centre of the link and to the input spacing between the nodes;

4. **Rotated**: the grid is rotated according to the link angle from the true north;

5. **GFLSM**: the GFLSM equation below (Equation 5.2) is used to calculate the pollutant’s concentrations at each node of the grid (Luhar and Patil, 1989):

\[
C = \frac{1}{2\sqrt{2\pi}\sigma_z(u\sin\theta + u_0)} \left[ \exp \left( -\frac{(z-h)^2}{2\sigma_z^2} \right) + \exp \left( -\frac{(z+h)^2}{2\sigma_z^2} \right) \right],
\]

\[
x \left[ \text{erf} \left( \frac{\sin(\theta(p-y) - x\cos\theta)}{\sqrt{2}\sigma_y} \right) + \text{erf} \left( \frac{\sin(\theta(p+y) - x\cos\theta)}{\sqrt{2}\sigma_y} \right) \right]
\]

(Equation 5.2)

The term \(u_0\) is a wind speed correction to account for the effects of traffic wake and assumes different values for different stability classes as suggested in the GM model (Chock, 1978). The emission rate, \(Q_1\), was determined from composite vehicle emission factors calculated using the COPERT4 methodology (Ntziachristos et al., 2009) as described in the previous section. The vertical and horizontal dispersion coefficients \(\sigma_v\) and \(\sigma_z\) are dependent on the downwind distance \(x\) and the Pasquill (F. Pasquill, 1961) stability class. The Briggs Method (Briggs, 1973) was used for calculating the lateral and vertical dispersion coefficient because it offers the best approximation of the coefficients’ values for dispersion in urban environments (Hanna et al., 1982). Briggs used theoretical concepts of the related formulas to obtain a set of formulas which can be used in common practices. According to these formulas, \(\sigma_v\) and \(\sigma_z\) are proportional to \(x\) at all stability conditions and they are independent of release height and roughness. The values of \(\sigma_v\) and \(\sigma_z\) in urban conditions are given in Table 43 (Hanna et al., 1982).
The stability categories were developed by Pasquill in 1961 (F.Pasquill, 1961) and are based on wind speed, insolation, and extent of cloud cover. The amount of turbulence in the ambient atmosphere has a major effect on the dispersion of air pollution plumes (Beychok, 2005) because turbulence increases the entrainment and mixing of unpolluted air into the plume and thereby acts to reduce the concentration of pollutants in the plume. It is therefore important to categorize the amount of atmospheric turbulence present at any given time. The atmospheric turbulence was categorised into six stability classes as follows:

<table>
<thead>
<tr>
<th>Pasquill Stability class</th>
<th>$\sigma_y$</th>
<th>$\sigma_z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-B</td>
<td>$0.32^*x(1+0.0004^*x)^{-1/2}$</td>
<td>$0.24^*x(1+0.001^*x)^{1/2}$</td>
</tr>
<tr>
<td>C</td>
<td>$0.22^*x(1+0.0004^*x)^{-1/2}$</td>
<td>$0.20^*x$</td>
</tr>
<tr>
<td>D</td>
<td>$0.16^*x(1+0.0004^*x)^{-1/2}$</td>
<td>$0.14^*x(1+0.0003^*x)^{1/2}$</td>
</tr>
<tr>
<td>E-F</td>
<td>$0.11^*x(1+0.0004^*x)^{-1/2}$</td>
<td>$0.08^*x(1+0.00015^*x)^{-1/2}$</td>
</tr>
</tbody>
</table>

Table 43: Formulas produced by Briggs (Briggs, 1973) for $\sigma_y(x)$ and $\sigma_z(x)$ ($10^2 < x < 10^4$ m)

6. **Output:** this worksheet holds the coordinates of each receptor node with the correspondent concentration level of the pollutant, ready to be imported in to ArcGIS.

**5.4.3 Evolution of the model: the GFLSM ArcGIS model**

The output of the calculations performed in Microsoft Excel can be easily imported in to ArcGIS. Once the Excel table is imported into ArcGIS by using the “Add XY data” function, the receptors’ grid rotated according to the link direction is displayed on the screen as it is shown in Figure 132. Each point of the grid contains the following data:

- Coordinates;
- Distance from the street;
- Concentration level.
The next step consists of converting the table file into a layer. This step is necessary for the flowing one: interpolate between the concentrations values at each grid node and generate a raster. The output of this procedure is illustrated in Figure 133.

The final steps involves the generation of the PM10 contours (Figure 134), which are a continuous line on a map that represents equal level PM10 concentration levels.
The main drawback of this procedure is that it is time consuming because it needs to be repeated for each road link and each combination of weather variables, such as temperature, atmospheric stability class, wind speed and direction.

A fully integrated GFLSM ArcGIS model tool is under testing. The modelling steps are similar to the ones used in the Excel model, but the iteration of the procedure for each link and weather conditions is carried out automatically by ArcGIS through a Python code. Python is one of the most user-friendly programming languages available today. It is open source, runs on just about every platform, and can be used to program many kinds of GIS software, including ArcGIS. Scripts are used to automate the time-consuming and complex GIS workflow illustrated above. ArcGIS also allows the user to create a custom user script tool by coding the workflow and the succession of commands. The custom tool can be then easily called and used by any ArcGIS user, like all the others tools.
CHAPTER 6 – Point and Area sources modelling
6. Introduction

This chapter describes the elements of the modelling procedure which account for the following typologies of air pollutants sources:

- **Point source**: A point source is a single, identifiable source of air pollutant emissions, which could be either elevated or at ground-level. It includes sources like combustion furnace flue gas stacks, power stations, steel works, foundries, incinerators, wood and pulp processors, paper mills, refineries, chemical production, etc. (Kibble and Harrison, 2005). Quoting the U.S.E.P.A. “Example Application of Modelling Toxic Air Pollutants in Urban Areas” (U.S.E.P.A., 2002): “Point sources involve the release of emissions from a well-defined stack or vent, at known physical stack parameters and operational conditions. Consequently, characterizing point sources for modelling is fairly straightforward. The basic model inputs for any point source are: stack location coordinates, the physical stack parameters (height above ground level and inside diameter at stack exit); operational conditions (gas velocity and temperature at stack exit); building dimensions (height, width, depth), and emission rate”.

- **Area source**: Those facilities or activities whose individual emissions are not strong enough to be modelled as point sources can be grouped together and treated as area sources. An area source is a two-dimensional source of diffuse air pollutant emissions. Area sources represent numerous facilities or activities that individually release small amounts of a given pollutant, but collectively can release significant amounts of a pollutant. Quoting the U.S.E.P.A. “Example Application of Modelling Toxic Air Pollutants in Urban Areas” (U.S.E.P.A., 2002): “Area and Other sources are smaller stationary sources that emit less than 10 tons per year of a single air pollutant or less than 25 tons per year of a combination of air toxics. Typical examples of area sources include neighbourhood dry cleaners and gas stations. Though emissions from individual area sources are often relatively small, collectively their emissions can be of concern particularly where large numbers of sources are located in heavily populated areas.”. It includes also the emissions from sources like a landfill or the evaporated vapours from a large spill of volatile liquid. As regards this study, the area sources identified in the Greater Dublin area belong mostly to the house heating category;
**Volume source**: A volume source is a three-dimensional source of diffuse air pollutant emissions. Essentially, it is an area source with a third (height) dimension. Quoting the U.S.E.P.A. “Example Application of Modelling Toxic Air Pollutants in Urban Areas” (U.S.E.P.A., 2002): “There are two basic types of volume sources: surface-based or ground-level sources that may also be modelled as area sources, and elevated sources. A source may be defined as a volume source for modelling when its emissions can be considered to occur over a certain area and within a certain depth of space. Release area, base elevation and area are needed for modelling. The important parameters used to characterize volume sources for dispersion modelling are location, release height and initial lateral and vertical dimensions”. Examples of volume sources are the fugitive gaseous emissions from piping flanges, valves and other equipment at various heights within industrial facilities such as oil refineries and petrochemical plants. Another example would be the emissions from an automobile paint shop with multiple roof vents or multiple open windows. No important volumes sources have been identified in the Greater Dublin area.

A literature review of the possible modelling approaches and a detailed description of the model used for the purpose of this work and its results are presented in the following sections.

### 6.1 Overview of existing models

An overview of existing models for calculating the pollutants dispersion from point and area sources is presented below. These models are the following:

- **AUSPLUME**: it is a screening Gaussian-plume model and the de facto standard Gaussian-plume model in New Zealand and Australia for small, steady-state, near-field applications. It is the Australian EPA approved regulatory model;
- **IS CST3**: the Industrial Source Complex Short Term Model (IS CST3) is a steady-state Gaussian plume model which can be used to assess pollutant concentrations from wide variety of sources associated with an industrial complex. It is especially designed to support the US EPA's regulatory modelling programs;
- **AERMOD**: it is a “near-field, steady-state” guideline model. It uses boundary-layer similarity theory to define turbulence and dispersion coefficients as a continuum, rather than as a discrete set of stability classes;
CTDMPLUS: it is a US EPA regulatory model developed specifically for tall point-sources in areas of complex terrain. CTDMPLUS is a steady-state plume model containing algorithms that enable a more physically realistic description of vertical dispersion and air-flow around complex terrain features;

ADMS 3: the Atmospheric Dispersion Modelling System (ADMS 3) is an advanced atmospheric pollution dispersion model for calculating concentrations of atmospheric pollutants emitted either continuously from point, line, volume and area sources, or intermittently from point sources. It was developed by Cambridge Environmental Research Consultants (CERC) in collaboration with the UK Meteorological Office and the University of Surrey.

To predict the impact of a source, AUSPLUME and ISCST3 require information about the sources, meteorological data, receptor information and topographic data. It employs a GUI through which the user may easily edit and execute the model. Both models are very easy to use and quick to run, and the output is easily interpreted. However, they are principally designed for calculating impacts in regions of flat terrain. Whether designed for flat or complex terrain, Gaussian-plume models are best used for near-field applications where the steady-state meteorology assumption is most likely to apply.

AERMOD is a “near-field, steady-state” guideline model. It uses boundary-layer similarity theory to define turbulence and dispersion coefficients as a continuum, rather than as a discrete set of stability classes. Variation of turbulence with height allows a better treatment of dispersion from different release heights. Also, dispersion coefficients for unstable conditions are non-Gaussian, to represent the high concentrations that can be observed close to a stack under convective conditions.

CTDMPLUS is a steady-state plume model containing algorithms that enable a more physically realistic description of vertical dispersion and air-flow around complex terrain features. In the past CTDMPLUS has been successfully used in New Zealand, but it is not frequently used any more due to its highly specialised meteorological data requirements and its applicability only to tall point sources (Bluett et al., 2004). The model simulates the flow and plume distortion near user-selected, three-dimensional terrain features, yet retains simplicity by applying flow-distortion corrections to flat-terrain, Gaussian, and bi-Gaussian pollutant distributions.

ADMS 3 includes algorithms which take into account downwash effects of nearby buildings within the path of the dispersing pollution plume and effects of complex terrain. The system also includes a meteorological data input pre-processor (Beychok,
2005) and it is capable of simulating passive or buoyant continuous plumes as well as short duration puff releases

AERMOD, CTDPLUS and ADMS-3 require considerably more time than AUSPLUME and ISCST3 for data preparation and model set-up because they include algorithms which take into account the atmospheric turbulence and its variation with height.

6.1.6 Limitations of Gaussian-plume models

The following limitations of steady-state Gaussian models should be considered and weighed against their advantages before employing this type of model in any dispersion study (Bluett et al., 2004):

a) Causality effects: Gaussian-plume models assume pollutant material is transported in a straight line instantly (like a beam of light) to receptors that may be several hours or more in transport time away from the source. This means that plume models cannot account for causality effects. This feature becomes important with receptors at distances more than a couple of kilometres from the source.

b) Low wind speeds: Gaussian-plume models have low performances during low wind speed or calm conditions due to the inverse wind speed dependence of the steady-state plume equation, and this limits their application. It has to be noted that in many circumstances these conditions produce the worst-case dispersion results for many types of sources.

c) Straight-line trajectories: In moderate terrain areas, these models will typically overestimate terrain impingement effects during stable conditions because they do not account for turning or rising wind caused by the terrain itself. CTDM and SCREEN are designed to address this issue.

d) Spatially uniform meteorological conditions: Gaussian steady-state models have to assume that the atmosphere is uniform across the entire modelling domain, and that transport and dispersion conditions exist unchanged long enough for the material to reach the receptor. In the atmosphere, truly uniform conditions rarely occur. Water bodies, hills and other terrain features, differences in land use, surface characteristics, and surface moisture all produce inhomogeneities in the structure of the boundary layer which can affect pollutant transport and dispersion. Convective conditions are one example of a non-uniform meteorological state that Gaussian-plume models cannot emulate. Luhar et al. (Luhar et al., 2000) and Hurley (Hurley et
al., 2005) demonstrated that under prediction can occur in the near field for tall stacks (>100 m) under convective conditions. The notable exception to this is AERMOD, which has a specially developed, “add-on” probability density function to replace the Gaussian concentration distribution during convective atmospheric conditions.

e) **No memory of previous hour’s emissions**: in calculating each hour’s ground-level concentration, the plume model has any memory of the contaminants released during the previous hours. This limitation is especially important for the proper simulation of morning inversion break-up, fumigation and diurnal recycling of pollutants over cities.

### 6.2 List of important point and area sources in Dublin

This section deals with the emissions from power stations and industrial plants in the Dublin area.

#### 6.2.1 Power stations and industrial emissions

A total of nine major power station and industrial sources in the Dublin city area were identified and incorporated into this study. Their location and modelling parameters are reported in Table 45, and their emission rates are reported in Table 46. The new Ringsend CCGT power generating plant commenced operation in 2002 and the main industrial emission source in this area of Dublin, the Irish Glass Bottle facility closed in 2002. The power station at North Wall and the CCGT plant at Poolbeg both operate on natural gas with distillate oil only used during emergency conditions, if the supply of natural gas is interrupted. The main generating plant at Poolbeg power station operates on a combination of fuel oil and natural gas.

The industrial sector within the Dublin urban area is comprised mainly of relatively small premises situated within industrial estates within the suburbs. There are no longer any large industrial emission sources within the port or inner city areas, with much of this sector comprising warehousing, distribution, high technology or light engineering activities. Any industries emitting significant quantities of atmospheric pollutants are licenced under the Integrated Pollution Prevention Control (IPPC) licensing legislation by the Irish Environmental Protection Agency. Detailed procedures concerning the IPPC licensing process are set out in the EPA Acts 1992 to 2011 and the associated licensing regulations (Irish.E.P.A., 1992). In the majority of cases where emission limits are specified, these relate to process emissions, in particular of VOCs, or boiler emissions. The latter sources generally have very low
emissions and relate only to NOx, as the industrial facility would normally be burning natural gas. These industrial sources have been included in the emissions inventories as part of the area source emissions, accounted for by annual fuel usage estimates. As expected, since most burn natural gas, the particulate emissions from these sources are low and so are unlikely to be significant in the context of overall urban air quality.

Since the mid-1980s, there has been a major shift in fuel usage patterns within both the industrial and commercial sector due to the replacement of oil by natural gas in the majority of large industrial and the larger commercial users, such as hospitals and shopping centres. There are now only a very small number of users burning distillate oil for process and space heating requirements. For the purpose of this work, it has been assumed that all significant commercial premises burn natural gas (Dublin City Council, 2009).

Information concerning industrial and commercial fuel usage, based on fuel usage data supplied by Bord Gais for their sales areas for the Dublin urban area and provided by the Irish E.P.A., gives a basis for estimating emissions from the industrial and large commercial activities. These statistics give the annual volume of natural gas sales for a number of zones within each area and by using appropriate emission factors these data can be converted into emission rates.

<table>
<thead>
<tr>
<th>Source</th>
<th>Location (Grid Ref)</th>
<th>Height (m)</th>
<th>Diameter (m)</th>
<th>Temperature (°C)</th>
<th>Velocity (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Power stations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Wall CCGT</td>
<td>319142, 234794</td>
<td>70</td>
<td>5.5</td>
<td>200</td>
<td>24</td>
</tr>
<tr>
<td>North Wall GT</td>
<td>319216, 234771</td>
<td>65</td>
<td>5.7</td>
<td>480</td>
<td>34</td>
</tr>
<tr>
<td>Poolbeg A</td>
<td>320612, 233723</td>
<td>207</td>
<td>4.1</td>
<td>105</td>
<td>22</td>
</tr>
<tr>
<td>Poolbeg B</td>
<td>320693, 233726</td>
<td>207</td>
<td>6.2</td>
<td>139</td>
<td>11</td>
</tr>
<tr>
<td>Poolbeg CCGT1</td>
<td>320839, 233690</td>
<td>75</td>
<td>5.2</td>
<td>80</td>
<td>27</td>
</tr>
<tr>
<td>Poolbeg CCGT2</td>
<td>320862, 233691</td>
<td>75</td>
<td>5.2</td>
<td>80</td>
<td>27</td>
</tr>
<tr>
<td>Ringsend</td>
<td>319600, 233700</td>
<td>70</td>
<td>6.5</td>
<td>90</td>
<td>26</td>
</tr>
<tr>
<td><strong>Industrial sources</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irish Glass Bottle 1-4</td>
<td>319148, 233543</td>
<td>26</td>
<td>1.2</td>
<td>760</td>
<td>5.8</td>
</tr>
<tr>
<td>Guinness Brewery B1-4</td>
<td>314100, 234100</td>
<td>46</td>
<td>1.5</td>
<td>130</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 45: Dublin power station and industrial point source locations and modelling parameters
### Table 46: Dublin power station and industrial sources emission rates (g/s)

<table>
<thead>
<tr>
<th>Source</th>
<th>NOx  1999</th>
<th>NOx  2004</th>
<th>NOx  2010</th>
<th>PM   1999</th>
<th>PM   2004</th>
<th>PM   2010</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Power stations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Wall CCGT</td>
<td>7.2</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Wall GT</td>
<td>2.4</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poolbeg A</td>
<td>40</td>
<td>40</td>
<td>17</td>
<td>3.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poolbeg B</td>
<td>59</td>
<td>59</td>
<td>33</td>
<td>5.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poolbeg CCGT1</td>
<td>26</td>
<td></td>
<td>2.44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poolbeg CCGT2</td>
<td>26</td>
<td></td>
<td>2.44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ringsend</td>
<td>-</td>
<td>34</td>
<td>34</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total power stations</strong></td>
<td><strong>162</strong></td>
<td><strong>196</strong></td>
<td><strong>146</strong></td>
<td><strong>15</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Industrial sources</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irish Glass Bottle 1</td>
<td>1.7</td>
<td>-</td>
<td>0.3</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irish Glass Bottle 2</td>
<td>1.7</td>
<td>-</td>
<td>0.3</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irish Glass Bottle 3</td>
<td>1.3</td>
<td>-</td>
<td>0.2</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irish Glass Bottle 4</td>
<td>1.3</td>
<td>-</td>
<td>0.2</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guiness Brewery B1</td>
<td>1.4</td>
<td></td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guiness Brewery B2</td>
<td>1.4</td>
<td></td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guiness Brewery B3</td>
<td>1.4</td>
<td></td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guiness Brewery B4</td>
<td>0.4</td>
<td></td>
<td>0.044</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total industrial</strong></td>
<td><strong>11</strong></td>
<td><strong>5</strong></td>
<td><strong>5</strong></td>
<td><strong>1.4</strong></td>
<td><strong>0.4</strong></td>
<td><strong>0.4</strong></td>
</tr>
</tbody>
</table>

6.2.2 Domestic emissions

The availability of information on which to base an emissions inventory for the domestic sector is limited. Given the widespread change not only in heating patterns but also in the number of housing developments that have taken place over the past decade within the Dublin urban area, there are substantial gaps in the information available for preparing the emissions inventories. No detailed data are available concerning recent housing constructions on a small area basis at a District Electoral Division (DED) level, as information is only collated by the Department of Environment on a county basis. This means that the number of new and existing dwellings and types of properties within the urban area is limited. The information included in this section was included in an unpublished report written in 2003 and kindly provided by the Irish Environmental Protection Agency.

Emissions from the domestic sector were derived for the Dublin urban area based on disaggregated 1996 National Population Census and fuel usage information from the 1991 Census. The 1991 Census included information on the main fuel used for heating the house and so the number of houses using solid fuel (coal and peat), oil, natural gas and other types could be calculated for each DED within the urban area.
A subsequent Census in 1996 provided an update on the number of households within each DED. However, the 1996 Census did not provide any fuel usage information and, with the 2002 Census data unpublished at the time of the 2003 report, an exercise was carried out to adjust these historical statistics to provide an estimate for 1999.

Due to a major change in fuel usage patterns within the domestic sector since 1991, direct incorporation of the 1991 Census fuel data to derive emissions for 1999 was not appropriate. Since 1991 there has been a dramatic change in the pattern of domestic fuel use as a result of the 1990 legislation prohibiting the sale of bituminous coal (Air Pollution Act, 1987), and a subsequent rapid development expansion in the number of houses using natural gas. Therefore, the 1991 Census data on fuel usage is of limited value in reflecting the spatial pattern of domestic emissions for 1999. Therefore, an adjustment was made in the 1991 Census statistics for the numbers of houses where coal, natural gas and oil are used within each DED in Dublin.

The main change that has occurred in the past decade in emissions from the domestic sector is the substantial reduction in particulates and other pollutants associated with the burning of coal and other solid fuel. The majority of these domestic users converted to natural gas during the 1990s. Based on information available from sales demand figures provided by Bord Gais, the 1991 Census data for houses burning solid fuel was adjusted to take account of this fuel switch in Dublin. The Bord Gais data is presented on the basis of sales areas for each of the urban areas, rather than on a DED or postal area basis. In addition, no details are available concerning the number of new connections made in apartments during the 1990s, as no distinction is made according to dwelling size. However, the information from Bord Gais sales is the only available data on which to base the adjustment between 1991 and 1999.

It was assumed that all new houses built between 1991 and 1996 burnt natural gas and the remaining number of houses on natural gas in each sales area had switched from coal or were existing gas consumers in 1991. More recent information available from Bord Gais relating to 1999 natural gas sales within the domestic sector enabled a further adjustment to be carried out.

It was assumed that the majority of new houses built between 1996 and 1999 were built in the suburbs and the effect of inner city in-filling with new houses would not have a significant impact on emissions within each DED. Many of the new dwellings constructed within the inner city area of Dublin between 1996 and 1999 are
apartments and so the energy usage per dwelling would be low compared to the
typical suburban property. It was assumed that there was no significant change in the
number of houses using oil for heating between 1991 and 1999 and that those
switching to natural gas were replaced by new users. This enabled an estimate of the
number of dwellings or houses burning natural gas, oil or solid fuel within each DED to
be estimated for 1999.

These data concerning the number of houses burning natural gas, oil and solid fuel
usage per DED were disaggregated into 1km by 1km grid squares based on a
computerised methodology to estimate the percentage of each DED within each grid
square. Adjustments were also made to take account of significant areas of open
ground such as parkland, industrial areas and other open spaces contained within the
DEDs in the Dublin urban area. The total number of houses burning each of the three
types of fuel within each grid square was therefore estimated.

To convert the disaggregated domestic fuel usage data to annual average emissions
of NOx and PM10 per house, appropriate emission factors were applied for houses
where the fuel being burnt was natural gas, solid fuel or oil. Energy demand studies
carried out in the 1980s by An Foras Forbartha provide an estimate of actual energy
demand for different types of dwelling size, such as the semi-detached or mid-
terraced houses. These studies reflect the actual energy demand requirements based
on real measurements in different types of houses. The resulting figures of domestic
energy demand take into account the energy content of the fuel being burned and the
efficiency of the combustion compliance.

Based on this, an estimate was derived in terms of gigajoules per year for natural gas,
solid fuel or oil usage for a typical semi-detached or terraced house. The annual
energy demand (GJ/year) for each fuel type per grid square was converted into an
annual emission rate using CORINAIR (European Environment Agency (E.E.A.),
2009) emissions data, coupled with the 2002 UK National Atmospheric Emissions
Inventory database for the various pollutants. This provided estimates of annual
emissions per average house using natural gas, solid fuel or oil in terms of kg/yr. The
annual emission rate from all dwellings within each 1km by 1km grid square for NOx
and PM10 was then calculated. This was finally converted into an emission rate (g/s)
for incorporation into the emissions database.

The numbers of solid fuel- and oil- burning houses have been reduced by 5% and
2.5% per annum from the 1999 values respectively; based on the assumption that
these percentages of houses will convert to natural gas fuel in 2004 and 2010.
6.2.3 Emission totals

Table 47 and Table 48 show, respectively, the total emissions in Dublin resulting from each source type for each pollutant for the three modelled years included in the 2003 report provided by the E.P.A., and these values as percentages of the totals for each pollutant.

<table>
<thead>
<tr>
<th>Source type</th>
<th>Year</th>
<th>NOx</th>
<th>PM10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power stations</td>
<td>1999</td>
<td>5106</td>
<td>473</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>6178</td>
<td>473</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>4614</td>
<td>473</td>
</tr>
<tr>
<td>Other point sources</td>
<td>1999</td>
<td>331</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>145</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>145</td>
<td>13</td>
</tr>
<tr>
<td>Domestic heating</td>
<td>1999</td>
<td>557</td>
<td>276</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>560</td>
<td>223</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>562</td>
<td>180</td>
</tr>
<tr>
<td>Industrial/commercial</td>
<td>1999</td>
<td>367</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>367</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>367</td>
<td>38</td>
</tr>
</tbody>
</table>

Table 47: Emissions for Dublin city (T/yr) for each emission year

<table>
<thead>
<tr>
<th>Source type</th>
<th>Year</th>
<th>NOx</th>
<th>PM10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power stations</td>
<td>1999</td>
<td>43%</td>
<td>41%</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>52%</td>
<td>49%</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>50%</td>
<td>57%</td>
</tr>
<tr>
<td>Other point sources</td>
<td>1999</td>
<td>3%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>Major roads</td>
<td>1999</td>
<td>34%</td>
<td>26%</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>26%</td>
<td>21%</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>24%</td>
<td>14%</td>
</tr>
<tr>
<td>Minor roads</td>
<td>1999</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>3%</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>2%</td>
<td>1%</td>
</tr>
<tr>
<td>Cold-start and hot-soak</td>
<td>1999</td>
<td>10%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>10%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>12%</td>
<td>0%</td>
</tr>
<tr>
<td>Domestic heating</td>
<td>1999</td>
<td>5%</td>
<td>24%</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>5%</td>
<td>23%</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>6%</td>
<td>22%</td>
</tr>
<tr>
<td>Industrial/commercial</td>
<td>1999</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>3%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>4%</td>
<td>5%</td>
</tr>
</tbody>
</table>

Table 48: Percentage of total emissions for Dublin city

Excluding traffic emissions (which have been examined Chapters 4 and 5), the major contributions to total PM10 emissions in Dublin are domestic heating and power stations. The effect of this kind of sources is expected to be relatively important for
modelling the personal exposure of inhabitants in sub-urban areas where road traffic doesn't have a significant impact on the air quality. On the other hand, in areas characterised by the presence of high volumes traffic road links, the effect of the point and area sources on the total pollution may be neglectable.

6.3 Dispersion model for point sources

The point sources identified in the Dublin area are mainly power stations and industrial sources; the dispersion of pollutants from sources belonging to this category is modelled by using the Gaussian plume model with the Briggs Method (Briggs, 1973) for the lateral and vertical dispersion coefficient.

6.3.1 Power stations and industrial sources model

The Gaussian plume model is a (relatively) simple mathematical model that is typically applied to point source emitters, such as coal-burning electricity-producing plants. Occasionally, this model will be applied to non-point source emitters, such as exhaust from automobiles in an urban area. The model is implemented in Microsoft Excel and it is also translated in Python to be used as a tool in ArcGIS.

One of the key assumptions of this model is that over short periods of time (such as a few hours) steady state conditions exist with regard to air pollutant emissions and meteorological changes (Turner, 1994). Air pollution is represented by an idealized plume coming from the top of a stack of some height and diameter. One of the primary calculations is the effective stack height. As the gases are heated in the plant (from the burning of coal or other materials), the hot plume is thrust upward some distance above the top of the stack (the effective stack height). The calculation of this vertical displacement depends on the stack gas exit velocity and temperature, and the temperature of the surrounding air.

Once the plume has reached its effective stack height, dispersion continues in three dimensions. Dispersion in the downwind direction is a function of the mean wind speed along across the plume. Dispersion in the cross-wind direction and in the vertical direction will be governed by the Gaussian plume equations of lateral dispersion. Lateral dispersion depends on a value known as the atmospheric condition, which is a measure of the relative stability of the surrounding air. The model assumes that dispersion in these two dimensions will take the form of a normal Gaussian curve, with the maximum concentration in the centre of the plume, Figure 135.
The "standard" algorithm used in plume studies is the Gaussian plume model, developed in 1932 by O.G. Sutton (Sutton, 1932). The algorithm is as follows:

$$C(x, y, z) = \frac{Q \cdot \exp\left(\frac{-y^2}{2 \cdot \sigma_y^2}\right)}{2 \cdot \pi \cdot u \cdot \sigma_y \cdot \sigma_z} \cdot \left(\exp\left(\frac{-(z-h)^2}{2 \cdot \sigma_z^2}\right) + \exp\left(\frac{-(z+h)^2}{2 \cdot \sigma_z^2}\right)\right)$$

(Equation 6.1)

where:

1. $C(x, y, z)$ is the concentration of the emission (in micrograms per cubic meter) at any point $x$ meters downwind of the source, $y$ meters laterally from the centreline of the plume, and $z$ meters above ground level;
2. $Q$ is the quantity or mass of the emission (in grams) per unit of time (seconds);
3. $u$ is the wind speed (in meters per second);
4. $H$ is the height of the source above ground level (in meters);
5. $\sigma_y$ and $\sigma_z$ are the standard deviations of a statistically normal plume in the lateral and vertical dimensions, respectively.

This algorithm has been shown in a large number of studies to provide reasonable predictions of emission dispersion in a variety of conditions. In this algorithm, we are concerned with dispersion in all three dimensions ($x$, $y$, and $z$):
• longitudinally (in the x direction) along a centreline of maximum concentration running downwind from the source

• laterally (in the y direction) on either side of the centreline, as the pollution spreads out sideways

• vertically (in the z direction) above and below a horizontal axis drawn through the source

The parameters that need to be calculated for a simple Gaussian plume model are as follows:

1. The effective stack height $H_{eff}$;
2. The lateral and vertical dispersion coefficients $\sigma_y$ and $\sigma_z$;
3. The ground-level concentrations $C_{max}$.

### 6.3.1.1 Effective Stack Height

The effective stack height $H_{eff}$ is defined as the final equilibrium height of the centreline of an effluent (smoke) plume from a smokestack or other source (Briggs, 1973). The effective stack height is the sum of the actual physical height of the top of the stack, plus any plume rise due to buoyancy or initial momentum (inertia) of the rising effluent, minus any downwash such as stack downwash, building downwash, or terrain downwash. The downwashes are associated with wake turbulence behind objects on the ground. In other words, the effective stack height $H_{eff}$ is equal to the physical stack height ($h_s$) plus the height of the plume (plume rises, $\Delta h$) determined from where the plume bends over.

$$H_{eff} = h_s + \frac{1.6 \cdot \exp\left(\frac{\ln f_0}{3}\right) \cdot \exp\left(\frac{2 \cdot \ln (3.5 \cdot x_0)}{3}\right)}{u}$$

(Equation 6.2)

Where:

$h_s$=physical height of the stack (in metres)

$U$=wind speed (in metres per second)

$f_0$ and $x_0$ are given by the formulas

$$f_0 = 3.12 \cdot 0.785 \cdot v_0 \cdot d^2 \cdot \frac{t_0 - t_1}{t_0}$$

(Equation 6.3)
And:

If $f_0 > 55$, then $x_0 = 34 \exp(0.4\ln(f_0))$          (Equation 6.4)

If $f_0 \leq 55$, then $x_0 = 14 \exp(0.625\ln(f_0))$  (Equation 6.5)

Where:

$v_0$ = gas velocity (in metres per second)

$d$ = stack diameter (in metres)

$t_0$ = gas exit temperature (in degrees Kelvin)

$t_1$ = ambient temperature (in degrees Kelvin)

6.3.1.2 Lateral and Vertical Dispersion Coefficients

The Briggs Method (Briggs, 1973) was used for calculating the lateral and vertical dispersion coefficient because it offers the best approximation of the coefficients’ values for dispersion in urban environments (Hanna et al., 1982). Briggs used theoretical concepts of the related formulas to obtain a set of formulas which can be used in common practices. According to these formulas, $\sigma_y$ and $\sigma_z$ are proportional to $x$ for all stability conditions and they are independent of release height and roughness. The values of $\sigma_y$ and $\sigma_z$ in urban conditions are given in Table 49 (Hanna et al., 1982).

<table>
<thead>
<tr>
<th>Pasquill Stability class</th>
<th>$\Sigma_y$</th>
<th>$\Sigma_z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-B</td>
<td>$0.32\times(1+0.0004\times x)^{-1/2}$</td>
<td>$0.24\times(1+0.001\times x)^{1/2}$</td>
</tr>
<tr>
<td>C</td>
<td>$0.22\times(1+0.0004\times x)^{-1/2}$</td>
<td>$0.20\times x$</td>
</tr>
<tr>
<td>D</td>
<td>$0.16\times(1+0.0004\times x)^{-1/2}$</td>
<td>$0.14\times(1+0.0003\times x)^{1/2}$</td>
</tr>
<tr>
<td>E-F</td>
<td>$0.11\times(1+0.0004\times x)^{-1/2}$</td>
<td>$0.08\times(1+0.00015\times x)^{1/2}$</td>
</tr>
</tbody>
</table>

Table 49: Formulas produced by Briggs (Briggs, 1973) for $\sigma_y$ and $\sigma_z$ ($10^2 < x < 10^4$ m)

The stability categories were developed by Pasquill in 1961 (F.Pasquill, 1961) and are based on wind speed, insolation, and extent of cloud cover. The amount of turbulence in the ambient atmosphere has a major effect on the dispersion of air pollution plumes (Beychok, 2005) because turbulence increases the entrainment and mixing of unpolluted air into the plume and thereby acts to reduce the concentration of pollutants in the plume (i.e., enhances the plume dispersion). It is therefore important to categorize the amount of atmospheric turbulence present at any given time. The atmospheric turbulence is categorised into six stability classes as follows:
### Stability class

<table>
<thead>
<tr>
<th>Stability class</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>very unstable</td>
</tr>
<tr>
<td>B</td>
<td>Unstable</td>
</tr>
<tr>
<td>C</td>
<td>slightly unstable</td>
</tr>
<tr>
<td>D</td>
<td>neutral</td>
</tr>
<tr>
<td>E</td>
<td>slightly stable</td>
</tr>
<tr>
<td>F</td>
<td>stable</td>
</tr>
</tbody>
</table>

Table 50: The Pasquill stability classes (F.Pasquill, 1961)

6.3.1.3 Ground-Level Concentrations

The equation below (Equation 6.6) is obtained from the previous one by setting \( z = 0 \) and \( y = 0 \); it has to be noted that the concentration at the ground level \( (z=0) \) is needed for assessing the personal exposure of people at street. Furthermore, \( y = 0 \) gives the maximum concentrations (straight downwind).

\[
C_{\text{max}} = \frac{Q}{2 \pi u \sigma_y \sigma_z} \left( \exp\left( -\frac{(H_{\text{eff}})^2}{2 \sigma_z^2} \right) + \exp\left( -\frac{(H_{\text{eff}})^2}{2 \sigma_y^2} \right) \right)
\]

(Equation 6.6)

If either standard deviation \( (\sigma_y \text{ or } \sigma_z) \) is less than 0.01, then \( C_{\text{max}} \) is set to zero.

6.3.2 Model results

The model described above is implemented in Microsoft Excel for each point source identified in the modelled area. The model inputs are the followings:

1. Source parameters: coordinates, stack height, stack diameter, emission rate \((g/s)\), gas exit velocity, gas exit temperature and ambient temperature;
2. Wind speed \((m/s)\) and direction;
3. Receptors grid: a receptor grid is generated for each source, with a editable distance between points and main dimension parallel to the wind direction;
4. Atmospheric Stability Category: very unstable, moderately unstable, slightly unstable, neutral, somewhat stable and stable.

The output is the estimated concentration of the pollutant at ground level in mmg/m³ at each point of the receptor grid. Figure 136 and Figure 137 illustrate the output of such a calculation for a wind speed of 5m/s, wind direction of 135°, ambient temperature of 10°C and for “slightly unstable” and “somewhat stable” atmospheric stability conditions respectively. Once the grid with concentration values at each point is imported into ArcGIS, the values are interpolated using the kriging algorithm and the contours displayed in the figures below are generated.
The downfall of this procedure is that, as with the GFLSM model, it is time consuming because it needs to be repeated for each point source and each combination of weather variables. As such, a fully integrated point source ArcGIS model tool is under testing. The modelling steps are similar to the ones used in the Excel model, but the iteration of the procedure for each point source and weather conditions is carried out.

Figure 136: Dispersion of pollutants from the industrial sites in Dublin, "slightly unstable" atmospheric stability condition

Figure 137: Dispersion of pollutants from the industrial sites in Dublin, "slightly stable" atmospheric stability condition
automatically by ArcGIS through a series of Python scripts. ArcGIS also allows the user to create a custom user script tool by coding the workflow and the succession of commands. The custom tool can be then easily called and used by any ArcGIS user, like all the others tools.

6.4 Dispersion model for area sources

The area sources identified in the Dublin area are mainly domestic residential sources. For the purpose of this work, a simplified approach for calculating the contribution from domestic sources to ambient PM10 concentration levels is used.

Gaussian plume models have been developed for point sources (e.g., stacks as discussed in the previous section), line sources (e.g., roads), and area sources (e.g., residential areas). The basic Gaussian plume model assumes a point source. Line sources can be approximated as a series of point sources. Using a simple transformation, an area source can be approximated by modelling it as a point source located further away from the receptor than it actually is, such that the plume has dispersed to be as wide as the area source when it reaches the true location of the area source. This is called a virtual point source and it is the approach used in the U.S. Environmental Protection Agency screening model SCREEN3. A different approach is to integrate over the area using the “narrow plume approximation”. If the receptor is reasonably near the source and the angle between the wind direction and a line between the source and receptor is not greater than 45°, the values obtained from a virtual point source estimate and a narrow plume approximation calculation will be roughly the same.

6.4.1 Domestic sources model – SCREEN3

The SCREEN3 model was developed to provide an easy-to-use method of obtaining pollutant concentration estimates. The area source algorithm in SCREEN3 is based on the numerical integration algorithm implemented in the Industrial Source Complex (ISC) Short Term area source model (U.S.E.P.A., 1995b). This model allows for rectangular shaped area sources and it is based on a numerical integration over the area in the upwind and crosswind directions of the Gaussian point source plume formula given in the previous section. Areas which have irregular shapes can be simulated by dividing the area source into multiple rectangles that approximate the geometry of the area source (U.S.E.P.A., 1995c).
The ground-level concentration at a receptor located downwind of all or a portion of the source area is given by a double integral in the upwind (x) and crosswind (y) directions as (U.S.E.P.A., 1995b):

\[
X = \frac{Q_A K}{2\pi u_s} \int_x V \frac{1}{\sigma_y \sigma_z} \left\{ \int_y e^{x} \exp \left[ -0.5 \left( \frac{y}{\sigma_y} \right)^2 \right] dy \right\} dx
\]

(Equation 6.7)

where:

\( Q_A \) = area source emission rate (mass per unit area per unit time);

\( K \) = a scaling coefficient to convert calculated concentrations to desired units (default value of 1 x 10^6 for \( Q_A \) in g/(m^2s) and concentration in µg/m^3);

\( u_s \) = wind speed at release height;

\( V \) = vertical term which accounts for the vertical distribution of the Gaussian plume, including the effects of source elevation, receptor elevation, and limited mixing in the vertical.

The vertical term \( V \) in order to reduce computational time without a loss of accuracy, is accounted in the form:

\[
V = \frac{\sqrt{2\pi} \cdot \sigma_z}{z_i}
\]

(Equation 6.8)

\( z_i \) = mixing height at downwind distances where the \( \sigma_i/z_i \) ratio is greater than or equal to 1.6.

The numerical integration algorithm for area sources allows for impacts to be estimated at receptors located within the area source. Since this algorithm involves a numerical estimate of a double integral over the area source, the calculations are very intensive and cannot practically be performed by hand. The SCREEN3 model should therefore be used to estimate maximum short term (1-hr) concentrations from area sources (U.S.E.P.A., 1995c). The SCREEN3 user’s manual also states that concentrations close to an area source will not vary as much as those for point sources in response to varying wind directions, and the meteorological conditions which are likely to give maximum 1-hour concentrations can persist for several hours. As such it recommends to conservatively assuming the maximum 1-hour concentration to apply for averaging periods out to 24 hours. It is also asserted that
SCREEN3 can perform all the single source short-term calculations in the EPA screening procedures document, including:

- Estimating maximum ground-level concentrations and the distance to the maximum;
- Incorporating the effects of building downwash on the maximum concentrations for both the near wake and far wake regions;
- Incorporate the effects of simple elevated terrain on maximum concentrations;
- Estimating 24-hour average concentrations due to plume impaction in complex terrain using the VALLEY model 24-hour screening procedure;
- Modelling simple area sources using a numerical integration approach;
- Modelling the effects of simple volume sources using a virtual point source procedure;
- Calculating the maximum concentration at any number of user-specified distances in flat or elevated simple terrain;
- Examining a full range of meteorological conditions, including all stability classes and wind speeds to find maximum impacts.

However, the SCREEN3 user’s manual (U.S.E.P.A., 1995c) and the EPA’s guideline on air quality models (U.S.E.P.A., 1995a) recommend some precautions in the adoption of SCREEN3, because the model will not be able to explicitly determine maximum impacts from multiple sources, except for the procedure to handle multiple nearby stacks by merging emissions into a single "representative" stack.

6.4.1.1 Area Source Parameters

For an area source, the following parameters must be determined (U.S.E.P.A., 1995c):

- Emission rate: the emission rate for area sources is input as an emission rate per unit area;
- Source release height: the source release height above ground has to be specified;
- Sides Length of rectangular area: the area source is approximated as rectangular and the user is asked to input the larger and smaller sides of it;
- Search through range of wind directions: since the concentration at a particular distance downwind from a rectangular area is dependent on the orientation of the area relative to the wind direction, the SCREEN3 model provides the user with two options for treating wind direction. The user is prompt to choose whether or not to search through a range of wind directions.
By selecting this option, EPA’s SCREEN model will search through a range of wind directions to find the maximum concentration. The range of directions used in the search is determined from a set of look-up tables based on the aspect ratio of the area source, the stability category, and the downwind distance (U.S.E.P.A., 1995c). This is the regulatory default option and should be used for most applications. If the “search through range of wind directions” option is not chosen, a wind direction orientation relative to the long axis of the rectangular area must be specified. This option may be used to estimate the concentration at a particular receptor location relative to the long area.

6.4.2 Derivation of the emission factors

The availability of information on which to base an emissions inventory for the domestic sector is limited. Given the widespread change not only in heating patterns but also in the number of housing developments that have taken place over the past decade within the Dublin urban area, there are substantial gaps in the information available for preparing the emissions inventories. There is available data concerning recent housing constructions on a small area basis at an Electoral Division (ED) level; this information is available for downloading from the Central Statistics Office (C.S.O) Ireland as part of the 2006 and 2011 Census reports:


Emissions from the domestic sector were derived for the Dublin urban area based on the “Private dwellings in permanent housing units in each Province, County and City, with or without central heating” for each ED obtained from the National Population Census 2006 and 2011 and average natural gas usage for each household, information obtained from the Commission for Energy Regulation through Bord Gais.

The main change that has occurred in the past two decades in emissions from the domestic sector is the substantial reduction in particulates and other pollutants associated with the burning of coal and other solid fuel. The 1991 Census included information on the main fuel used for heating the house and so the number of houses using solid fuel (coal and peat), oil, natural gas and other types could be calculated for each ED within the urban areas. The majority of these domestic users converted to natural gas during the 1990s. It was therefore assumed that all houses with central heating in the Dublin area burnt natural gas because it was believed that the number of dwellings or houses burning coal, oil or solid fuel was within the Dublin area would not have a significant impact on emissions within each ED.
The data extracted from the Census 2006 and 2011 concerning the number of houses burning natural gas per ED were georeferenced and imported into ArcGIS (Figure 138).

The data were then disaggregated into 1km by 1km grid squares using ArcGIS to estimate the percentage of each ED within each grid square. Adjustments were also made to take account of significant areas of open ground such as parkland, industrial areas and other open spaces contained within the EDs in the Dublin urban area (Figure 139).

Figure 138: Houses with central heating in the Dublin area and its Satellite Towns.

Figure 139: Data disaggregation procedure
The total number of houses within each grid square was therefore estimated (Figure 140).

Figure 140: Houses with central heating per sq km

Based on the average natural gas usage for each household (obtained from the Commission for Energy Regulation through Bord Gais), an estimate was derived in terms of Kilowatt-hours (kWh) per year for natural gas for a typical dwelling. The annual energy demand (kWh/year) for natural gas per grid square was converted into
an annual emission rate using CORINAIR emissions data, coupled with the 2009 UK National Atmospheric Emissions Inventory (NAEI) database for the various pollutants and with the United States Environmental Protection Agency Emission factor documentation for Natural Gas Combustion (U.S.E.P.A., 1998). This provided estimates of annual emissions per average house using natural gas in terms of kg/yr. The annual emission rate from all dwellings within each 1km by 1km grid square for PM10 was then calculated using ArcGIS. This was finally converted into an emission rate (g/day) for incorporation into the database (Figure 141) and incorporated into a GIS raster.

According to the data provided by Board Gais, the average yearly consumption of natural gas per house is estimated to be around 13800 kWh/year. This was converted in m$^3$ using an average calorific power of roughly 11 kWh/m$^3$ for the natural gas. As such, the average yearly consumption of natural gas was calculated to be 1254.5 m$^3$/year, thus 3.44 m$^3$/day. This volume of natural gas was then converted in emission factor for the average house using the relation provided by the US Environmental Protection Agency (U.S.E.P.A., 1998) for calculating the production of Particulate Matter (mg) for each m$^3$ of natural gas burned. This factor is 121.6 mg of PM per m$^3$ of gas burned. The house average emission per day for therefore found to be 418.03 mg/day.

The variation of natural gas consumption during the year 2010 is calculated using the average daily temperature variation during the same year as input. The plots in Figure 142, Figure 143 and Figure 144 represent respectively the variation of temperature, the variation of gas consumption per house and the variation of PM10 emission per house during the year 2010.

From the Figure 141 above, it can be seen that the emission of PM due to the combustion of natural gas in private houses reaches 1276 grams per square kilometre per day in the Dublin city area with some other peaks in correspondence of the Satellite Towns, such as Dun Laoghaire.

The calculation is carried out for one of the grid squares in the Dublin area as an example and the results are presented below. The concentrations are calculated by applying the SCREEN3 method with the Dublin residential emission rate of 1276 grams per square kilometre per day, with the following weather conditions:

- Atmospheric stability class: slightly unstable;
- Wind speed: 3m/s;
- Wind direction: 90°.
Figure 142: Temperature variation measured by Phoenix Park weather monitoring station in 2010

Figure 143: Average natural gas daily consumption (per house) variation in 2010

Figure 144: Average PM10 daily emission (per house) variation in 2010

The results of such calculations are shown in the plot in Figure 145 below: as it can be noted, the concentration levels reach the maximum at 500 meters distance from the side of the residential square grid and then they gradually drop.
A raster is then created in ArcGIS using again the kriging interpolation algorithm and the outcomes of such procedure are presented in Figure 146.

Figure 146: Dispersion of pollutants from a 1sqKm residential area in Dublin, "slightly stable" atmospheric stability condition
CHAPTER 7 – Commuters’ exposure modelling
7. Introduction

One of the final purposes of the research presented in this thesis is to create an ArcGIS model for Personal Exposure (PALM-GIS) to Particulate Matter for the inhabitants of the Dublin City. The air dispersion models described in the previous chapters are integrated in ArcGIS following the procedure described in Chapter 8. The advantage of this solution is that the iteration of the modelling procedure for different modelling tests and weather conditions is carried out automatically by ArcGIS through a Python code. It is user-friendly, open source, runs on just about every platform, and can be used to program many kinds of GIS software, including ArcGIS. Scripts are used to automate the time-consuming and complex GIS workflow illustrated above. ArcGIS also allows the user to create a custom user script tool by coding the workflow and the succession of commands. The custom tool can be then easily called and used by any ArcGIS user, as with other tools. This integration aims to provide the researchers, local authorities and others with a tool to calculate the concentration levels of PM and to correlate them with other thematic layers, such as land use and population density, in order to link localised peaks in air pollutants with particular activities. As such, the following outcomes are obtained by using dedicated ArcGIS workflows and tools:

1. Modelled background concentration levels;
2. Modelled traffic related concentration levels;
3. Modelled industrial sources related concentration levels;
4. Modelled domestic sources related concentration levels;

The concentration levels are then combined in ArcGIS in order to obtain total concentration levels of at specific locations in the Dublin.

This chapter illustrates the procedure and outcomes from applying the model to the assessment of the personal exposure to PM10 of commuters in Dublin. This modelling employs the various tools and algorithms described in the previous chapters that are included and unified within ArcGIS, as discussed in Chapter 8. Its outcomes are compared and validated against the dataset acquired as part of another PhD project, also belonging to the Personal exposure Activity Location Model (PALM) Project, funded by the EPA STRIVE 2007-2013 (Science, Technology, Research and Innovation for the Environment) Programme. This project comprises the monitoring of the personal exposure, activity and location of a number of different subjects living and working in Dublin and its satellite towns. This data is used to develop the GIS model of personal exposure described in this section, allowing assessment of
individuals or sub-groups of the population and their health risks due to air pollution. The model will also inform individuals of activities or locations in their daily lives with that are associated with high personal exposure to air pollution, allowing them to take individual action to reduce or avoid these health risks.

7.1 The PALM Project

The PALM Project is divided in three main streams or research work, which are developed and investigated as part of three separate PhD projects:

1. The measurement of 24 hours personal exposure to PM10
2. The modelling of indoor exposure to PM10;
3. The modelling of the outdoor exposure to PM10; (described in this thesis).

7.1.1 Measurement of 24 hour personal exposure to PM10:

Background concentration is an unsatisfactory basis on which to investigate the impacts of air pollution on humans. A more relevant measure is the personal exposure of the individual subjects of the health impact assessment. In the PALM project, this has been experimentally quantified by recording personal exposure data over 24-hour periods on a real-time basis. These real-time measurements of the personal exposure of each subject comprise continuous samples obtained at two-minute intervals measured over a 24 hour period. Sampling of the personal exposure of each subject was carried out using a real-time particulate monitor, the Metone Aerocet 531 (Figure 147), which recorded PM1.0, PM2.5, PM10 and particle count. The volunteers were also equipped with a GPS tracking device which continuously monitored their locations and all volunteers kept a log of their activities during sampling (commuting; sleeping; working; shopping etc.). This approach highlighted the relationships between personal exposure, location and activity, while other influencing factors such as traffic and meteorological conditions were also taken into account.

Figure 147: The Metone Aerocet 531
7.1.2 Modelling indoor exposure to PM10

This modelling (performed by PALM project partners NUI Galway) predicts the personal exposure to PM10 of individuals during the time they spend indoors, according to their different indoor activities and locations, and also to the outdoor pollution concentration variations over time. Examples of the activities the model accounts for are: different kind of cooking such as frying or boiling, home activities like hovering and ironing, cleaning, sleeping, etc. The model also accounts for a multitude of variables, such as the operational status of the HVAC (heating, ventilation and air conditioning) systems, whether windows are open or closed. The testing of this modelling involves the replication of the measurements acquired by the Aerocet units during the monitoring campaign by reproducing the same indoor environment conditions. The results of the indoor model depend on the outdoor concentrations at the building location, and these data are obtained from the outdoor modelling methodology described in this thesis.

7.1.3 Modelling outdoor exposure to PM10

One of the tasks carried out as part of the research presented in this thesis is to create a model of Dublin inhabitants' outdoor exposure to PM10, while they are commuting from their homes to the workplaces. The assessment of this model's performance is performed by validating its results against the measurements gathered by the Aerocet units during the monitoring campaign. This is described in the following sections of this Chapter, where the commuters' exposure during their itineraries from their residences to work is modelled for different routes, periods of the year and means of transport and then compared to the relevant measured datasets.

7.1.2 The outdoor personal PM10 exposure to model

In the following sections, the modelling procedure employed to calculate the exposure of commuters during their trips from home to work are presented and discussed. The dataset of measurements includes more than 500 sample days of monitoring with the Aerocet units, with commuting routes converging from all over the Dublin to the city centre (Figure 148).
A number of significant examples are selected to be modelled and used as validation tests for the model. For completeness purpose, the commuters’ datasets to be modelled and analysed are selected with different transport modes and routes through the city. The following examples are chosen:

- Bus: southbound route to Trinity College with bus;
- Bicycle: westbound route to Trinity College with bicycle;
- Train: westbound route to Trinity College with train;
- Walk: northbound route from Trinity College.

### 7.2 Commuting by bus

This dataset was collected on the 24th of June 2010 and it consists of measurements of PM10 concentrations on the route from Temple Road, Rathmines to Dawson Street, in proximity of Trinity College (Figure 149). The commuting route starts with a short walk from Temple Road to a bus stop on Palmerston Park (A), followed by bus journey to Dawson Street of an approximately 20 minutes duration and finishes with a short walk through Trinity College (B).
The route goes through different air quality environments:

- The residential area, characterised by roads flanked by trees and low rise detached and semi-detached buildings with gardens;
- A heavily trafficked road (Rathmines Road), characterised by a building height to street width ratio which is typical of the “urban street canyons” and may lead to the generation of recirculation street vortexes;
- A park in an urban area, characterised by an open green area surrounded by heavily trafficked roads.

### 7.2.1 Measured data

The Aerocet averages the measurements over two minutes periods; these data are then combined with the positions recorded by the GPS unit, (which was carried by the commuter along with the PM10 monitoring device), and imported into ArcGIS (Figure 150).
As it can be seen from Figure 150, the measured PM10 concentration levels range between 15 and 31 µg/m³ and they reach maximum levels on Dawson Street, (where usually the traffic proceeds slowly), and after the bus journey, during the walk from the bus stop to the commuter’s office in Trinity College.

7.2.2 Modelled data

The modelling procedure for this dataset can be divided into three main components:

1. **Hourly PM10 sub-urban traffic model for Rathmines**: an Artificial Neural Network (ANN) model is built following the procedure described in Chapter 3. The hourly variation of PM10 concentration levels is calculated by finding a non-linear relation between the hourly concentrations of NO₂ measured at the Rathmines EPA monitoring station, the set of weather variables described in Chapter 3, such as temperature, atmospheric stability class, atmospheric pressure, humidity, solar radiation, rainfall, wind speed and direction, and historic data for PM10 concentration levels. This model is used for calculating the PM10 concentration levels in the area included into the red rectangle in Figure 151.
Figure 151: PM10 modelled with the Rathmines hourly ANN model

2. **OSPM**: The Operational Street Pollution Model (Berkowicz, 2000a) is the core of the ArcGIS tool used to model PM10 concentrations on the route indicated in Figure 152. OSPM can model pollutant dispersion in urban canyons was run in its multi-street mode (Figure 153).

3. **Hourly PM10 background model**: the hourly variation in PM10 background concentration levels is necessary for the OSPM calculations and it is also used to determine the concentration levels within the park in Trinity College. The hourly model is generated following the procedure described in details in Chapter 3.

The outcomes from the three modelling phases are combined in ArcGIS (Figure 154) in order to compare them with the measured data and also for analysis purposes. The modelled PM10 concentration levels range between 18 and 25 µg/m³ and they reach
their maximum levels on Dawson Street and on Lower Camden Street, which are both heavy traffic roads.

The historical data used as input to the background concentration model are recorded by permanent monitoring stations that have been classified as “background” locations within the Dublin Area. It is assumed that the concentrations measured by these monitoring stations include also the contributions from the area sources (such as distributed housing) in the area.

### 7.2.3 Analysis and validation

The measured and modelled PM10 concentration levels are compared and the difference between them is calculated and displayed in ArcGIS, in order to have a better insight into the reasons that might have caused any discrepancies between the two datasets (Figure 155). As it can be seen from Figure 155, the largest differences between the measured and modelled concentration levels are focused mainly around two locations, the bottom of Upper Rathmines Road and the corner between Dawson Street and Nassau Street. These are the locations where the subject boarded and exited the bus during the journey under analysis.
A closer look to the measured data shows a large variation in the measurements recorded during the change of environment, from outside to inside the bus and vice versa (Figure 156).

The measured and modelled data are averaged over 5 minute periods order to minimise the influence of such short-term variations and the two averaged datasets compared. Figure 157 indicates close agreement between the two time series of 5 minute average measured and modelled PM10 concentrations.
Figure 157: 5 minute average measured and modelled PM10 concentration levels

Figure 158 presents a regression analysis of the measured and modelled PM10 concentrations with a resultant value of the coefficient of correlation $R^2$ equal to 0.74. This indicates that the modelling procedure described above performed well in predicting the PM10 concentration levels to which the commuter was exposed during this bus-based journey from home to the work.

Figure 158: Regression of measured against modelled values

7.3 Commuting by bicycle

In this section the different ArcGIS models are applied to calculate the exposure to PM10 exposure of a cyclist commuting to work. This dataset was collected on the 6th of July 2010 and it consists of measurements of PM10 concentrations on the route from Carpenterstown Road, Castleknock (A) to Pearse Street (B), in proximity of Trinity College (Figure 159). The commuting route was completed by bicycle, thus the
Aerocet is not enclosed in any microenvironment, such as car or bus, and it’s in close contact with the pollutant source, the road traffic.

Figure 159: View of the commuting route by bicycle

This route goes through very heterogeneous environments, which require completely different air pollutants dispersion modelling approaches:

- The residential area around Castleknock, characterised by roads flanked by trees and low rise detached and semi-detached buildings with gardens;
- The rural environment from the M50 Flyover and the residential area surrounding Phoenix Park, characterised by very sparse houses, open fields and trees along the road;
- The open park (Phoenix Park), characterised by open spaces with no buildings;
- The heavily trafficked road (the Quays), characterised by a ratio building height street width which is typical of the “urban street canyons” and may lead to the generation of recirculation street vortexes;

### 7.3.1 Measured data

As for the previous test case, the Aerocet measurements are georeferenced and imported into ArcGIS (Figure 160).
Figure 160: Measured PM10 concentration levels

Figure 160 graphically shows the variation of the measured PM10 concentration levels, which range between 14 and 28 µg/m³ and they reach their maximum levels between Lower Ormond Quay, (where traffic is usually slowed down by a traffic light before the bridge), and Pearse Street. The measurements show also a peak between the exit to Phoenix Park and the Quays.

7.3.2 Modelled data

The modelling procedure for this dataset is more time consuming because it involves the use of three different models. Also the input data preparation requires more time than the previous model because of the diversity of the employed tools. The utilised models are the followings:

1. **Area sources model:** the PM10 concentration levels for the part of journey outlined in Figure 161 is calculated with the area model described in Chapter 6.4. The required inputs are the emission factor per square km, which depends on the number of dwellings with central heating, and the hourly variation of PM10 background concentration levels. The contribution from the residential area sources is calculated following the procedure described in Chapter 6. The calculated pollution from this category of sources is low compared to the road traffic contribution. The hourly model for the background concentration is generated following the procedure described in Chapter 3 and relies on the
time series of concentrations measured at the EPA monitoring station in the Phoenix Park.

![Figure 161: PM10 modelled with the area sources model](image1)

2. **Hourly PM10 background model for Phoenix Park**: data mining algorithms are applied to the historical datasets for PM10, NO2 and weather variables recorded by the air quality and weather monitoring stations located in Phoenix Park. Phoenix Park PM10 monitoring station only records daily averages for PM10, but it also records NO2 hourly averages. As such, the NO2 hourly variation and also the weather variables hourly averages are used to generate a PM10 hourly profile following the procedure described in Chapter 3. This model is used to calculate PM10 concentration levels in the area within the red rectangle shown in Figure 162, for the time of day when the analysed commute took place.

![Figure 162: PM10 modelled with the Phoenix Park hourly ANN model](image2)

3. **OSPM and point sources**: The ArcGIS tool running OSPM is used to model the dispersion of particulate matter from road traffic on the route within the red rectangle in Figure 163. The different widths of the links shown in Figure 164...
are a visualisation of larger traffic volumes. The contribution of the point sources located along this part of the route (Guinness factory) is modelled using the procedure described in Chapter 6. In this instance, the contribution is minimal if compared to traffic related pollution because of the weather conditions recorded during the test (i.e. wind not blowing towards the quays).

Figure 163: Route modelled with OSPM

Figure 164: OSPM multi-streets model

Figure 165: Modelled PM10 concentration levels
Once the three tools have performed these calculations, the outcomes are combined in ArcGIS (Figure 165). The modelled total concentration levels, as it can be seen in Figure 165 below, are able to reproduce the gradual increase that is noticeable also for the measured concentration levels in Figure 160; although the measured PM10 concentration levels have a higher variability than the modelled ones, which range between 14 and 25 µg/m³ and reach their maximum level on Pearse Street, a heavily trafficked road.

7.3.3 Analysis and validation

ArcGIS is used again to calculate the difference between the modelled and measured datasets (Figure 166).
As it can be seen from Figure 166, the model performs well for the journey through the residential area in proximity of Castleknock Village, the successive rural area and Phoenix Park. The model though under-predicts the concentration levels just after the park and for the last part of the journey, between Lower Ormond Quay and Pearse Street. These locations correspond with the peaks noted in the measured data and may be attributable to congestion at the traffic lights, which could locally increase vehicle emission rates and result in peaks in the measured data.

A closer look to the measured data shows a large variation between consecutive measurements recorded in the proximity of the two locations discussed above (Figure 167).

![Figure 168: 5 minute average measured and modelled PM10 concentration levels](image)

![Figure 169: Regression of measured against modelled values](image)
Both datasets are again averaged over 5 minute periods to minimise the effects of short-term variations. As it would be expected from the visualisation of the two datasets in ArcGIS, the two time series of 5 minute average measured and modelled PM10 concentration levels display similar positive trends, with the differences discussed above, still clearly noticeable (Figure 168).

The measured and modelled PM10 concentration level values are then regressed against each other (Figure 169), with a resultant value of the coefficient of correlation $R^2$ equal to 0.8, which shows that the modelling procedure described above performed well in predicting the PM10 concentration levels to which the commuter was exposed during their bicycle trip from home to work using the bicycle.

### 7.4 Dart commuting

This dataset was collected on the 7th of April 2010 and it consists of measurements of PM10 concentrations on the Dart route from Coolmine Station (A) to Pearse Station (B) (Figure 170). The challenging aspects of this test case are mainly that the user and the Aerocet are confined for most of the journey in the microenvironment of the Dart which travels at high velocity through different air environments.

![Figure 170: View of the commuting route by train](image)

The route goes through different air quality environments:

- A residential area, characterised by low rise detached and semi-detached buildings with gardens;
- Dublin centre, characterised by a building height to street width ratio typical of “urban street canyons” and may lead to the generation of recirculation street vortexes;
7.4.1 Measured data

The georeferenced PM10 concentration levels are imported into ArcGIS and each measurement point is extruded according to concentration level, in order to perform a visual analysis of the dataset (Figure 171). As it can be seen from Figure 171, the measured PM10 concentration levels present high variability and they range between 11 and 34 \( \mu g/m^3 \). Despite the fact that the Aerocet was confined in the Dart microenvironment, it recorded high levels of PM10 in proximity to the junction between the M50 and the N3, and also while the Dart was travelling relatively close to the Navan Road (N3). After the train journey, the portable monitor recorded peaks in PM10 concentration levels during the walk from Pearse Dart Station to the subject’s office on Pearse Street, in Dublin city centre.

Figure 171: Measured PM10 concentration levels

7.4.2 Modelled data

The modelling procedure for this dataset employed two different modelling methods:

1. **The Dart journey model**: a combination of different models was used to determine the PM10 concentration levels to which the user was exposed during their Dart journey. The task proved to be difficult because the train rails don’t run parallel to street, which would have led to a more direct correlation with road traffic, and they also pass through different and heterogeneous areas. As such, the background PM10 ANN model is used to generate a baseline level to which the outcomes from the area source model are added.
The GFLSM model (Chapter 5) is used to calculate the traffic contributions from the main adjacent roads while the train was stopped at the various train stations with its doors open.

2. The walk to the office model: The Operational Street Pollution Model (Berkowicz, 2000a) was used to model the PM10 concentrations between Pearse Station and the office on Pearse Street. The hourly ANN PM10 background model is used to calculate the hourly time series for the concentration levels, which is a required input for the OSPM model.

The outcomes from the above modelling phases are combined in ArcGIS (Figure 172) in order to compare them with the measured data. The modelled PM10 concentration levels range between 15 and 22.5 µg/m³ and they reach their maximum levels on Westland Row and Pearse Street. As it’s evident from the visualisation in Figure 172, the modelling algorithms are not able to reproduce the peaks in concentrations observed in the proximity of the M50/N3 junction.

![Figure 172: Modelled PM10 concentration levels](image)

7.4.3 Analysis and validation

The measured and modelled PM10 concentration levels are compared and the difference between them is calculated and displayed in ArcGIS, in order to have a better insight into the reasons that might have caused these discrepancies between the two datasets (Figure 173). As it can be seen from Figure 173 below, the Dart model under-predicts the concentration levels for the part of the journey outside
Dublin centre. This might be due to the omission or underestimation of PM emission sources during that part of the trip.

The plot in Figure 174 below shows the two time series of 5 minute average measured and modelled PM10 concentration levels; it is evident, as observed above, that the model under-predicts the concentration levels for the majority of the trip.

Although the two time series have similar trends over the second half of the journey, the coefficient of correlation $R^2$ determined from regression analysis of the modelled and measured data (Figure 175) is equal to 0.42, which indicates that the modelling
procedure described above should be improved. It has to be noted though that, even if the $R^2$ value is low, the mean modelled and measured values are quite close.

![DART - Correlation between measured and modelled PM10 concentration levels](image)

Figure 175: Regression of measured against modelled values

The average $R^2$ and performance of the model can be explained by the fact that the train is moving at high speed between very heterogeneous environments characterised by a multiplicity of sources that are difficult to identify and quantify in a model designed to calculate the personal exposure of individuals distributed over a wide urban area. Moreover, the Dart interior possesses many of the characteristics of an indoor environment, and interior sources and carriage ventilation both affect ambient PM concentrations.

### 7.5 Walking

This test sample is selected because the user commuted to work without using any other transport mode, thus the Aerocet was not confined in any microenvironment and it measured the PM10 concentration levels on the footpath. This dataset was collected on the 28th of September 2010 and it consists of measurements of PM10 concentrations on the route from the office on Pearse Street, Trinity College to Portland Place, Drumcondra (Figure 176). The route goes through a quite homogeneous air quality environment, characterised by a building height to street width ratio which is typical of the “urban street canyons” and may lead to the generation of recirculation street vortexes. This test case was also chosen because it was carried out around 6pm, which is generally “rush hour” in Dublin and therefore a period in there is likely to high exposure to road traffic pollutants time-frame for pedestrians in that part of the city.
7.5.1 Measured data

The Aerocet averages the measurements over two minutes periods; these data were then combined with the positions recorded by the GPS unit, which was carried around by the test user along with the PM10 monitoring device, and imported into ArcGIS (Figure 177).
As it can be seen from Figure 177, the measured PM10 concentration levels range between 23 and 36 µg/m³ and they remain consistently high during the entire walk in the North part of the city. Two peaks are “located” on Middle Gardiner Street, in proximity of traffic lights and they are evident in the time-series presented in Figure 178.

![Figure 178: Time-series for measured PM10](image)

7.5.2 Modelled data

The modelling procedure for this dataset is performed using the Operational Street Pollution Model (OSPM). An Artificial Neural Network (ANN) model is built following the procedure described in Chapter 3.

![Figure 179: OSPM multi-streets model](image)

The hourly variation of PM10 concentration levels is calculated by finding a non-linear relation between the hourly variation of weather variables, such as temperature,
atmospheric stability class, atmospheric pressure, humidity, solar radiation, rainfall, wind speed and direction, and historic data for PM10 concentration levels. The multi-street OSPM model is run for the route presented in Figure 179.

Figure 179: OSPM model run for the route presented in Figure 179

The historical data used as input for the background concentration model are recorded at the permanent monitoring stations that are classified as “background” sites are located within the Dublin city close to the modelled route. As such, it is assumed that the measured levels by these monitoring stations include contributions from local distributed area sources.

The modelling results are compared with the measured data in ArcGIS (Figure 180). The modelled PM10 concentration levels range between 23.6 and 32.5 µg/m³ and they reach their maximum levels on Middle Gardiner Street, which is a heavy traffic road.

7.5.3 Analysis and validation

The measured and modelled PM10 concentration levels are compared and the difference between them is calculated and displayed in ArcGIS, in order to have a better insight into the causes of the discrepancies between the two datasets (Figure 181). As it can be seen from Figure 181 below, the largest differences between the measured and modelled concentration levels are focused mainly in correspondence with the two traffic lights on Middle Gardiner Street, where the measured concentrations include the two large peaks identified above.
The measured and modelled data are averaged over 5 minute periods to minimise the effects of such variations, and the two averaged datasets are compared in Figure 182, which shows very good agreement between them.

The measured and modelled PM10 concentration level values are then regressed against each other (Figure 183), with a resultant coefficient of correlation $R^2$ which is very close to 1, with a value of 0.96, which indicates that the modelling procedure described above performed well in predicting the PM10 concentration levels to which the commuter was exposed during their walk from the work place to home.
7.6 Observations and sensitivity analysis

The sensitivity analysis for the background hourly model and for the Rathmines model is presented in Chapter 3, where the variations of performance of the two ANN models are assessed and evaluated for different sets of predictors (weather variables). It has to be noted that the background model is based on the time-series of historical measurements of very sparse monitoring stations. As such, the model is built under the assumption that the background concentrations are constant over a large area around the monitoring station, while they may differ because of the different sources present in the area. Another source of error for this model is due to the assumptions that the weather conditions and especially wind direction and wind speed are also considered constant over a large area, while wind direction and speed are likely to change for different point of the urban boundary layer.

As regards the calculation of traffic related pollution, average values for the traffic for each link and a predefined variation of the traffic during the 24 hours were used as inputs for OSPM. Pre-defined traffic data are files with information on the relative distribution of vehicle types (such as e.g. passenger cars, vans, trucks and buses) and the fraction of the daily total for each of the vehicle types (Berkowicz et al., 2003). It has to be noted that the outcomes of the model operating under these conditions are still satisfactory, with an $R^2$ close to 75%. The accuracy of the model might be increased if actual traffic data would be used along with the daily and weekly profile for each road link, but this would increase the data preparation time with a possible relatively small increase in the quality of the results. For the purpose of this work, the
pre-defined traffic file “Type_C” was used along with the average hourly traffic for each link provided by Dublin City Council.

As regards the area sources, the modelling procedure is carried out under the assumption that the residential sources are uniformly distributed over the 1sqKm used for the calculations. Also the source characteristics are assumed constant for all the habitations: average values for the emission rate and the source release height are adopted for the calculations. Furthermore, due to the limited availability of information on which to base an emissions inventory for the domestic sector, the number of private dwellings with central heating was derived from data collected at Electoral Division (ED) level as part of the 2006 and 2011 Irish Census reports. The data were then disaggregated into 1km by 1km grid squares and the total number of houses within each grid square was therefore estimated. Then the emission rate per each house was derived from the average yearly consumption of natural gas per dwelling provided by Bord Gais.

These assumptions may result in a less accurate estimate of the contribution from residential sources to the total pollution levels, but it is necessary to guarantee an adequate speed of the calculation procedure. Furthermore, it has to be noted that for the Great Dublin area, the pollution from residential sources is neglectable when compared to the background concentration levels and the contribution from road traffic.

7.7 Comparison between measured and modelled dose values

In personal exposure studies, *inhalation* is considered as a major route of exposure for an individual, who breathes in polluted air which enters the respiratory tract. Identification of the pollutant uptake by the respiratory system can determine how the resulting exposure contributes to the dose. In toxicology, *dose* may refer to the amount of a harmful agent to which an organism is exposed: as such, the mechanism of pollutant uptake by the respiratory system can be used to predict potential health impacts within the human population. In order to define the inhalation exposure, it is assumed that the concentration of the airborne pollutant in the air at the mouth and nose boundary is inhaled by the person.

As such, in order to have a better estimate of the performances of the personal exposure model for the test cases analysed above, the exposure is calculated for the measured and modelled data. The calculation is performed for each presented test case and the results are presented in Table 51 below.
<table>
<thead>
<tr>
<th>Transport mode</th>
<th>Measured exposure</th>
<th>Modelled exposure</th>
<th>Ratio modelled/measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>bus</td>
<td>594.78026</td>
<td>591.3282</td>
<td>0.99</td>
</tr>
<tr>
<td>bicycle</td>
<td>652.9346</td>
<td>610.4161</td>
<td>0.94</td>
</tr>
<tr>
<td>dart</td>
<td>706.9305</td>
<td>623.4418</td>
<td>0.88</td>
</tr>
<tr>
<td>walk</td>
<td>711.2253</td>
<td>702.1068</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 51: Measured and modelled dose for the analysed test cases

It is clear from the results presented in Table 51 that the personal exposure model highly performs in predicting the exposure for the individuals while commuting to work with four different transport modes. It has also to be noted that the performance of the personal exposure model is very satisfactory also for the calculation of the exposure for the commuter travelling to work with the Dart, despite of the low coefficient of correlation between the modelled and measured 5 minute averages, as discussed above.
CHAPTER 8 – GIS model and dissemination of results
8. Introduction

Geographical Information Systems (Wong and Wu, 1996) are essential instruments in a wide range of areas and are extensively used by Businesses, Governments, Educators and Scientists, Environmental and Conservation Organizations, Natural Resource Groups, Researchers, amongst others. A Geographical Information System or Geospatial Information Systems (GIS) is a collection of tools that captures, stores, analyses, manages, and presents data that are linked to geographical locations (Bhat et al., 2011). Because of this set of capabilities, it is regarded as a key decision supporting tool for Local Authorities in most developed countries like Ireland. As such, this study proposes to integrate the developed air quality dispersion modelling methodology for the personal exposure to particulate matter within the GIS package ESRI ArcGIS.

This Chapter reports the challenges encountered when including the air quality models presented in the previous chapters into ArcGIS, and also discusses and proposes possible improvements and further research topics. A framework to use such GIS tools for disseminating results from air quality studies and to make the personal exposure model accessible to the general public for self-assessment of their exposure to particulate matter is also described here.

Furthermore, a strategy to include an indoor air quality model within the GIS personal exposure model and make means of making it available to Local Authorities and the general public for consultation is also proposed in this chapter. The use of advanced technology at the time of this study, such as cloud computing, is proposed and examined in the contest of the employed GIS platform.

8.1 The GIS air quality model

ESRI ArcGIS is a powerful software package for any kind of geographical study and it allows the user to perform simultaneous analysis of multiple datasets, stored as layers, and also to compare these datasets and find hidden relations and correlations between them. ArcGIS is equipped with a set of very useful toolboxes called ArcToolbox (Figure 184), which contains tools and toolsets (a toolset is just an organizational device, like a system folder). Each tool is a pre-set function or a task which can be employed to perform different actions, analysis, conversions, and associations between the selected layers.
The tools included in the ArcToolbox can all be used to build a custom model, which is therefore a series of actions and algorithms that can be connected and concatenated to build the desired process (Figure 185).

The custom model, as mentioned before, can be saved as an ArcTool which can then be run like any other pre-defined tool included in the ArcToolbox. The advantage of
building a model and then storing it as an ArcTool is that in GIS modelling or GIS data management, a tedious series of steps which are needed to be processed on a regular basis can be automated.

This approach is used to integrate all the modelling procedures described in the previous chapters into ArcGIS models for the calculation of ambient pollutant concentrations and personal exposure. These models are then both boxed into two separate ArcTools (Figure 186).

![Figure 186: PALM-GIS personal exposure model flowchart](image)

### 8.1.1 GIS air quality model

The ArcGIS air quality model, because of the complexity of its task, is divided in sub-tools which correspond to individual modelling and data acquisition operations. The interoperability of the sub-tools and their unification into a single tool is under study.

The ArcGIS models generated as part of this work are:

- Urban model: This ArcGIS model can be used to calculate the dispersion of pollutants in an urban environment with street canyons, such as Dublin centre. It runs OSPM as an external application, using a sequence of commands written in Python language, like the one reported below:

```python
# Import system modules
import sys, string, os, arcgisscripting

# Create the Geoprocessor object
gp = arcgisscripting.create(9.3)
```
# Read the parameter values:
# 1: input file
OSMP_exe = gp.GetParameterAsText(0)
    os.system(OSPM_exe)
except:
    print "There was an error!"

The model employs the same data pre-processing and preparation that are required by OSPM: the weather file, the city model with building heights, the traffic for each road link. The background concentration needs also to be pre-calculated and this is achieved with the Matlab ANN model described in Chapter 3. This model is run separately in Matlab but it is possible to generate an executable model from a Matlab script, which could be integrated into an ArcGIS model similarly to OSPM. The weather variables and target files must be manually prepared before starting the model. The ArcGIS urban model then assign a pollutant concentration value to each road link for which a calculation is performed.

- Sub-urban model: This ArcGIS model can be used to calculate the dispersion of pollutants in a sub-urban environment without street canyons as is found throughout most of Dublin and its Satellite Towns. In its current implementation, the model doesn’t perform the pollutant dispersion calculations, which are done by the General Finite Line Source Model (GFLSM) in Excel, but instead performs the data acquisition and visualisation in ArcGIS. An improved version of this ArcGIS model, which consists of a GFLSM fully integrated in ArcGIS, has also been considered. The main unsolved obstacle to this is to find a way to select the correct road links to include in the dispersion model for a given set of receptors, and to run the GFLSM only for those by following an automated procedure. At the moment, the selection of road links is done by manually creating a new shapefile form the original which contains all the road links.

- Point source model: this model is very similar to the sub-urban model described above, except that the important point sources in the Dublin Area are a limited number.

- Indoor air quality model: The core code of this model is developed in NUI Galway as part of the wider PALM project. The principal inputs into the model are the concentrations outside the target building, the air exchange rate (AER), and the sources; these are all in separate text files. The sources file determines the activities that are taking place, thus it is necessary to generate a sources file for each room. Window opening patterns are determined in the
AER file, by adjusting the AER value in the text file. On average, there are a total of 45 AER files for the home environment, as this allows adjusting each home to different configurations. There are individual files then for the AER and a source for each of the additional micro environments. All the initial concentrations, rooms’ sizes, emissions rates, deposition rates, etc. are contained within a Java class file, which the user would have access to as well. The major issue in generating an ArcGIS model from this code lies in its complexity and in the amount of files that need to be adjusted before the modelling procedure is started. At the present time, the code has been written very dynamically so that it can easily be adjusted for different house compositions of house, which results in such a large amount of text files. For example, the model not only needs to know the different activities that are taking place, but also the times at which these activities occur. The same rules apply for the AER rate. A possible solution to partially overcome this issue is to make some assumptions to simply the input procedure.

As mentioned before, a forthcoming research goal is to unify all the ArcGIS models described above, into a single GIS air quality model which would be included in a single ArcTool. This presents several challenges due to the logic and interconnections between the single tasks part of each model.

8.1.2 GIS personal exposure model

The ArcGIS model for personal exposure to PM10 is an extension of the model described in the previous section. Once the concentration levels are calculated within ArcGIS, the calculation of the personal exposure of a commuter on his/her route from habitation to workplace is trivial and can be done using ArcGIS Network Analyst. ArcGIS Network Analyst uses an advanced Network Data Model that allows the use of realistic datasets encompassing multiple attributes such as cost, restriction, and hierarchy. Various cost attributes, such as distance, time, and visibility can be defined according to their analysis requirements. Once the pollutant concentrations of are calculated for each road link on the route, this ArcGIS extension is able to calculate the exposure of the commuter on the route according to the time spent on each link, which is function of the average speed on the same link. It is also possible to calculate the route from a starting point (the habitation) to a destination (the work place) which would minimise the exposure by using a cost function applied to the road network and its characteristics, such as the average speed and the pollutant concentration.
The problem with this model is that it is not intuitive and it requires an advanced knowledge of ArcGIS. As such, a Graphical User Interface is considered which would allow also a non-expert user to run the model by inputting the start and end points of his/her route through the city.

8.2 GIS Cloud Computing and Web GIS

This section is purely speculative and its only aim is to address some research needs and opportunities. The core idea is to make the model developed as part of this study available to expert users (Local Authorities) and to the general public, in order to increase public awareness of air pollution and also have a platform for sharing self-built air quality related datasets and layers. The emerging technology of cloud computing seems ideal for achieving these objectives.

According to NIST, National Institute of Standards and Technology:

“Cloud Computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.” Cloud Computing is often equated with the concept of a utility, in which an organization can “plug-in” to this virtual computing environment and use the computing resources available on an as-required basis (Vaquero et al., 2008).

Different software and datasets can be run and stored on virtual servers in the cloud and they can be accessed via web clients. Cloud computing has five characteristics: on-demand self-service, ubiquitous network access, location-independent resource pooling, rapid elasticity and pay-per-use. It can be also differentiated according to three delivery models: software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS). Finally cloud computing can be defined according also to four deployment models: private, community, public and hybrid (Foster et al., 2008).

For the purposes of this study, cloud computing is believed to be an ideal modus operandi to share the GIS air quality and personal exposure models between Local Authorities and also, with limited capabilities, to the general public. The two models could be available in the Cloud and each user would be able to access different layers of data according to different privilege levels. Different organizations and developers could therefore publish and quickly deploy custom GIS mapping applications (Figure 187).
Figure 187: ArcGIS Cloud diagram

Figure 188: ArcGIS.com is a Web site providing access to GIS data and applications
ArcGIS Online (Figure 188) is a cloud-based system for creating and sharing maps and geographic information. It’s an extremely flexible system which enables the users to create, modify, add contents to maps and then share them with the other users who have access to the Cloud. Even users with no GIS experience could easily upload their own information and interact with the available models and datasets. Because ArcGIS Online is in the cloud, once a user has created a map and shared it, the map would be automatically part of an online catalogue. As such, this procedure would enrich the shared data with more information that could be employed in further studies.

Furthermore, ArcGIS in the Cloud could be an ideal environment for testing the ArcGIS Personal Exposure (PALM-GIS) model described in the previous section. ArcGIS Online is designed to function as complete stand-alone Software as a Service (SaaS) application for geographic information management and it could be used for running the PALM-GIS model on the Web and have a feedback from a wide range of users. The field trial would also allow the acquisition of important information about population activities patterns and mobility in the city. The PALM-GIS model could also be accessible to more expert users who would be able to run it from their desktop, mobile, and web applications and they could create maps and datasets by uploading their data using a browser. It is important to stress that by using ArcGIS online, the PALM-GIS model would be accessible to all the public without the need to have ArcGIS installed on their devices (Figure 188), which would increase its dissemination potential.

An example of this kind of application for ArcGIS Online is the Schleswig-Holstein online bicycle routing portal, which allows tourists to plan and map their itineraries between 2,500 points of interest, taking into account average speed, terrain, and weather forecasts (Figure 189).

ArcGIS Online would be also an ideal tool to share maps and geographic information across workgroups and organizations such as Local Authorities, which would facilitate the collaboration between them. Using this approach, each GIS organization can make its information available to non-specialists. This also enables the integration of information across organizations, providing a strong basis for collaboration (Chappell, 2010).

Furthermore, ArcGIS Online (ArcGIS.com) provides a number of services, including the following:
• Storage of publicly available maps and other GIS information (the maps and layers are provided by both ESRI and general users);
• Applications and tools created by ESRI, which allow the user to examine and work with the layers available or uploaded to ArcGIS.com;
• A central site for finding and accessing GIS applications that use the data on ArcGIS.com, including applications that run on computers in non-ESRI data centers.

ArcGIS.com may be an ideal platform to deploy the PALM-GIS personal exposure model because its services are ready-to-use tools which would enable even the most inexperienced user to perform basic analysis to his personal exposure data and have a better understanding of the health risk connected to the exposure to high pollution levels.

Figure 189: The Schleswig-Holstein online bike routing portal allows tourists to plan and map their itineraries (http://www.sh-tourismus.de/en/planning)

8.4 Comparison with existing software packages - AirGIS

AirGIS (Jensen et al., 2001) is a software package developed by the National Environmental Research Institute in Denmark, which also developed OSPM. The system applies a geographic information system (GIS) and it is based on OSPM, technical and cadastral digital maps and Danish national administrative databases for buildings, cadastres and populations. AirGIS estimates ambient air pollution levels at high temporal and spatial resolutions and it has the following capabilities (Jensen et al., 2001):
- Traffic emissions mapping;
- Air quality levels mapping;
- Human exposure calculation at residence addresses, at workplace addresses and in streets;
- Comparison of mapping and scenario results with air quality limits;
- Impact assessment of traffic air pollution abatement measures.

Compared to existing exposure models, the developed exposure model adds a geographic dimension to exposure modelling with a very high spatial resolution as the address is an exposure indicator.

The key features and advantages of the AirGIS model are:

- The model system allows for exposure estimates at a high spatial resolution (address level) and a high time resolution (hour);
- It's been extensively validated for epidemiological studies (Sørensen et al., 2012), human exposure studies (Raaschou-Nielsen et al., 2011), as well as urban air quality assessment and management (Jensen et al., 2009);
- The system can estimate exposure along a route provided tracking data is available. The position of individuals can be obtained using Global Positioning System (GPS) built into mobile phones. (Hertel et al., 2008).
- The capability to generate street configuration and traffic data for the OSPM model based on digital maps and national databases: this enables the automatic estimation of air quality levels at residence and workplace addresses (Figure 190);

Digital maps
- buildings with height (polygons)
- streets with traffic (polylines)
- addresses with people (points)

National and local databases

Figure 190: The AirGIS system makes use of digital maps, national and local administrative databases, and a Geographic Information System (Jensen et al., 2001)
There are several similarities between the AirGIS model and the modelling methodology created within this study. First of all, the GIS model presented in this project can automatically assess pollution at residence and workplace if, like AirGIS, the geocoded addresses of the locations are provided. Both models take advantage of the capabilities of the Geographic Information System environment and of the tools that allow comparison between different datasets included in layers.

AirGIS operates with ArcGIS 3.x, which is an old version of the ESRI GIS software package and therefore doesn’t have all the tools and functionalities of the latest version. The ArcGIS model presented here is generated using ArcGIS 9.x and also ArcGIS 10.x, the most recent versions of the ESRI ArcGIS packages, with all the benefits that this brings in terms of available analytical tools and functionalities. But the main differences are the following:

- **Background model:** the model used to calculate background concentration levels in AirGIS is the Urban Background Model (UBM) (Berkowicz, 2000b). This model is typically applied to generate an urban background concentration field with a resolution of 1km x 1km; the model requires emission data, regional concentration data and meteorological data to perform the calculations. The UBM is also capable of calculating the urban background concentration from a simplified area source dispersion formula (the SUB method) that takes into account urban vehicle emission density, city dimensions (transport distance) and building height (initial dispersion height) (Berkowicz et al., 2008); this option is only available for very large datasets. In contrast, the background model generated within this project employs historical data on air pollutant concentrations and for a set of weather variables and it is able to generate predictions of hourly and daily averages, as shown in Chapter 3. The model uses machine learning algorithms and thus its performances improve day by day, by increasing the size of the input datasets with the next-day measured concentration levels. The grid resolution of the background concentration levels model depends on the presence of permanent monitoring stations, which would provide the historical data for the surrounding area.

- **Sub-urban model:** AirGIS models the dispersion of pollutants using OSPM as its “core engine”, thus it is designed to perform calculations only for streets that can be classified as urban canyons because of their geometry. OSPM is not able to satisfactorily model the dispersion of pollutants for wide streets with low rise buildings or houses with gardens; as such, AirGIS would not be
recommended for modelling the air quality for most of Dublin City and Its Satellite Towns. The ArcGIS air quality model presented here includes a sub-urban tool which uses its GFLSM as core engine. This tool, as it is shown in Chapter 5, is capable of modelling the dispersion of pollutants from traffic on streets which are not classified as urban canyons.

- **Personal exposure model:** the AirGIS personal exposure model estimates the exposure along a route provided tracking data is available; it requires the user to carry a tracking system with him and also text messages to be sent from the mobile phone to the machine running the model. Even if not practical, this system accounts for the real time position of the user and therefore a more accurate estimate of his/her exposure to air pollution. The ArcGIS personal exposure model under study as part of this work would allow an estimate of the position of the user without real time tracking, according to average speeds for each road link.

- **AirGIS:** is a unique model already fully integrated with ArcGIS, while the GIS air quality model presented here is constituted by different tools which operate separately, even though they are all integrated with ArcGIS. A model which encompasses all the different GIS air modelling tools is under study at the time of writing, as mentioned above.
CHAPTER 9 – Conclusions
9. Summary

The research presented in this thesis focuses on the creation of a GIS air quality model for Dublin City and its Satellite Towns. The aims of the research are to evaluate the features of different air quality models and modelling methods and then to incorporate the most appropriate models into ArcGIS, along with a tool for evaluate the exposure to PM10 of commuters on their way from the residences to the work places.

Geographical Information Systems (Wong and Wu, 1996) are identified as an ideal platform for this study because it is an essential instrument in a wide range of areas of interest and is extensively used by Governments, Research Centres, Environmental and Conservation Organizations, etc. In particular, ESRI ArcGIS is selected for this work because it is widely used and it is a collection of tools that captures, stores, analyses, manages, and presents data that are linked to geographical locations (Bhat et al., 2011). Because of this set of capabilities, it is envisaged as a key decision supporting tool for Local Authorities in most of the developed countries like Ireland. As such, this study proposes to integrate the air quality dispersion models for particulate matter concentrations and personal exposure model to particulate matter within ESRI ArcGIS.

The study and creation of the air quality model is divided into different sub models according to the different scales, sources and dispersion mechanisms:

- Background model;
- Urban street model;
- Sub-urban street model;
- Point and area source dispersion model.

The background model is generated using data mining techniques using historical PM10 data recorded at permanent air quality monitoring stations in Dublin and a set of weather variables recorded by Met Eireann permanent monitoring stations. The background PM10 model performed well in calculating observed hourly and daily average PM10 concentration levels in Dublin, as shown in Chapter 3. The model uses machine learning algorithms and thus its performances improve day by day, by increasing the size of the input datasets with the next-day measured concentration levels. The grid resolution of the background concentration levels model depends on the presence of permanent monitoring stations, which provide the historical data for the surrounding area.
The same methodology was applied to historical PM10 data collected at permanent monitoring stations classified as "suburban traffic" or "urban traffic" and it is found that the model well performs in predicting daily averages for both cases. The model was also able to predict hourly average PM10 concentrations where hourly averages of NOx were available.

The calculated background concentrations are a necessary input for the urban street model, along with weather conditions, traffic volumes and emission rates. One of the most important inputs is the street geometry: a relatively narrow street between buildings that line up continuously along both sides is called urban street canyon and a stable circulatory vortex may be established within the street canyon if specific wind and geometry conditions are met. The combination of vehicle emissions and reduced dispersion in these circumstances can lead to high levels of pollution (Buckland and Middleton, 1999), which in turn pose a substantial health risk to pedestrians. The model selected for this task is the Operational Street Pollution Model, developed by the National Environmental Research Institute, Denmark. The choice of the model is also due to the extensive experience the Department of Civil Engineering in Trinity College Dublin has gained during several research projects. The model was validated for daily averages against measurements carried out at two separate locations, one in Dublin city centre and one in the suburbs, for the month of March 2010. The results correlate well with observed concentrations of PM10.

OSPM was not designed to calculate the dispersion of pollutants outside urban street canyons, thus it would not be appropriate for most of the Dublin area and its Satellite Towns, where the city geometry is characterised by relatively wide streets flanked by trees and sparse low rise buildings with gardens. The General Finite Line Source Model is chosen for modelling suburban traffic related PM10 concentration levels because of the experience and satisfactory results obtained in the past years in the Civil Engineering Department of Trinity College Dublin (TCD). The model is implemented using Microsoft Excel and then translated into Python to work as a tool in ArcGIS.

The power stations and industrial plants in the Dublin area were identified along with their emission rates and incorporated into this study. The dispersion of pollutants from sources belonging to these categories is modelled by using the Gaussian plume model with the Briggs Method (Briggs, 1973) for the lateral and vertical dispersion coefficients. The model is implemented in Microsoft Excel and it is also translated in Python to be used as a tool in ArcGIS. On the other hand, the availability of information on which to base an emissions inventory for the domestic sector is limited.
The area sources identified in the Dublin area are mainly domestic sources. Emissions from the domestic sector were derived for the Dublin urban area based on data on the “Private dwellings in permanent housing units in each Province, County and City, with or without central heating” for each Electoral District obtained from the National Population Census 2006 and 2011, combined with data on average natural gas usage for each household obtained from the Commission for Energy Regulation through Bord Gais. For the purpose of this work, a simplified approach for calculating the contribution from domestic sources to PM10 concentration levels is used. Area sources can be approximated by assuming the source is further away from the receptor than it actually is, such that the plume is already as wide as the area source at the correct distance. This is called a virtual point source and it is the approach used in the U.S. Environmental Protection Agency screening model SCREEN3, which uses a Gaussian plume model.

One of the goals of the research presented in this thesis is to generate a GIS model for personal exposure to particulate matter experienced by the inhabitants of Dublin while commuting to work. The modelled concentration levels resulting from the modelling procedures described above are imported into ArcGIS and combined in order to obtain total concentrations of PM10 for the Dublin Area and its Satellite Towns. The exposure of a set of users while commuting to work on different routes with various transport modes was modelled using ArcGIS and then validated against the measurements taken during a monitoring campaign during which subjects carried particulate monitoring and GPS units. The correlation between the modelled and measured PM10 concentrations was calculated for the different test cases: the correlation was found to be high for the bus, bicycle and walking modes, but low for the train mode.

The air dispersion models described above are then integrated in ArcGIS. The advantage of this solution is that the iteration of the modelling procedure for different modelling aims and weather conditions is carried out automatically by ArcGIS through a Python code. Python is one of the most user-friendly programming languages available today. It is open source, runs on just about every platform, and can be used to program many kinds of GIS software, including ArcGIS. Scripts are used to automate the time-consuming and complex GIS workflow otherwise required. ArcGIS also allows the user to create a custom user script tool by coding the workflow and the succession of commands. The custom tool can be then easily called and used by any ArcGIS user, like all the others tools. The ArcGIS tools created for this study are
under testing to optimise the work flow, to eliminate the eventual bugs and to make them more user-friendly.

A solution for making the tools available to the general public for self-assessment of personal exposure and for disseminating air quality modelling results is identified: ArcGIS online is suggested as an ideal tool for achieving this goal. Furthermore, the use of advanced technology at the time of this study, such as cloud computing, is proposed and examined in the contest of the employed GIS platform.

9.1 Main findings

The project has focused on a number of scientific issues in the development of the GIS air quality model for the Dublin Area and its Satellite Towns. Firstly, how to model air quality by integrating existing and self-implemented air quality models and a Geographical Information System platform. Secondly, how to generate input data for the models to allow the estimation of air quality at any location in the Dublin Area. These issues have been addressed by implementing various tools: a method to calculate the background concentration levels, a method to model the dispersion of pollutants from road traffic in urban street canyons, a method to model the dispersion of pollutants from road traffic in general conditions, and a method to calculate the dispersion of pollutants from point and area sources. Thirdly, an approach to include all the models above into GIS tools has been discussed. Fourthly, an approach for disseminating results from air quality studies has been described and proposed.

As part of this work, the performances of various air quality models are assessed and the most suitable tools for modelling the dispersion of particulate matter in Dublin are selected. This is done in the context of the main objective: the assessment of the personal exposure of subjects moving between different locations in the urban area. This objective implies that a highly accurate solution at one location is not valuable if it is combined with less accurate predictions for other locations. It has to be noted that the only model commercially available that is utilised in this study is OSPM; all the other tools are implemented or modified by the author.

A new model for predicting the PM10 background concentration levels in the Dublin Area and its Satellite Towns is created by the author using machine learning algorithms; to the author’s knowledge, it has not previously been reported in the context of air pollution modelling in Ireland. The background model uses artificial neural networks to model the non-linear relation between historical PM10 data recorded at Irish EPA permanent monitoring stations and the set of weather variables.
recorded by Met Eireann permanent monitoring stations. The background PM10 model well performs in calculating predicted PM10 concentration levels for hourly and daily averages, as shown in Chapter 3. As mentioned above, the model uses machine learning algorithms and thus its performances improve day by day, by increasing the size of the input datasets with the next-day measured concentration levels.

Another innovative aspect of the research presented here is the inclusion of all the modelling tools described above into a GIS platform: this integration aims to provide Local Authorities with a tool to calculate concentration levels of PM and to correlate these with other thematic layers, such as land use and population density, in order to link localised peaks in air pollutants with particular activities. The GIS model has been validated by modelling the personal exposure to particulate matter of commuters travelling to and from work in Dublin city centre using different routes and different transport modes and comparing the modelled data with measured datasets obtained with mobile sensors and GPS units. To the author's knowledge, a validation test against such heterogeneous dataset hasn't been previously reported.

The work described in this thesis also provides a conceptual framework to assess and model personal exposure to other environmental pollutants, such as other air pollutants and noise, on a GIS platform. It could be also employed as a research tool to assess the link between individual health effects and personal exposure to particulate matter. To date, epidemiological studies do not establish causation and the absence of specific dose-effect relationships for air pollutants makes it difficult to establish limit values for human exposure.

9.2 Scope for further study

A first forthcoming research goal is to unify all the ArcGIS models described above, once completed, into a single GIS air quality model which would be included in a single ArcTool. This presents several challenges due to the logic and interconnections between the single tasks part of each model and therefore it requires further research and tests.

A second forthcoming research goal is to optimise and simplify the input procedure of the indoor air quality model in order to generate an ArcGIS model from this code. At present time the code developed as part of another PhD project carried out in NUI Galway requires a large amount of files that need to be adjusted before the modelling procedure is started. The indoor air quality model has been written very dynamically so that it can easily be adjusted for different composition of house,
which results in such a large amount of text files. For example, the model not only needs to know the different activities that are taken place, but also the times that these activities are taking places. The same rules apply for the AER rate. This results in a large number of input files which values that need to be set before any calculation starts.

A third forthcoming research goal is to perfect the Graphical User Interface for the ArcGIS Personal Exposure Model, which would allow also the non-expert user to run the model by inputting the start and end points of his/her route through the city.

The last forthcoming research goal is to make the PALM-GIS model available online on the Cloud through ArcGIS Online: this could be an ideal environment for testing the ArcGIS Personal Exposure model. ArcGIS Online is designed to function as complete stand-alone Software as a Service (SaaS) application for geographic information management and it could be used for running the PALM-GIS model on the Web and have a feedback from a wide range of users. The field trial would also allow the acquisition of important information about population activities patterns and mobility in the city.
References


Bishop, C. 1996. *Neural Networks for Pattern Recognition*, Oxford University Press, USA.


Faulkner, W. B., Shaw, B. W. & Grosch, T. 2008. Sensitivity of Two Dispersion Models (AERMOD and ISCST3) to Input Parameters for a Rural Ground-Level Area Source. 58, 9.


Karatzas, K., Papadourakis, G. & Kyriakidis, I. 2009. Understanding and Forecasting Air Pollution with the Aid of Artificial Intelligence Methods in Athens, Greece. Springer Berlin / Heidelberg.


Information Technologies in Environmental Engineering, Springer Berlin Heidelberg.


Oberdörster, G. 1996. Significance of particle parameters in the evaluation of exposure-dose-response relationships of inhaled particles. Inhalation Toxicology, 8 Suppl, 73-89.


