CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	New tools and methods for analysis of the sources and spatial distribution of air pollution
Authors	Byrne, Rósín
Publication date	2024
Original Citation	Byrne, R. 2024. New tools and methods for analysis of the sources and spatial distribution of air pollution. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2024, Rósín Byrne. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-10-28 03:30:10
Item downloaded from	https://hdl.handle.net/10468/16924



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

New Tools and Methods for Analysis of the Sources and Spatial Distribution of Air Pollution

A thesis submitted to

Ollscoil na hÉireann, Corcaigh
National University of Ireland, Cork



UCC
Coláiste na hOllscoile Corcaigh
University College Cork, Ireland

for the degree of

Doctor of Philosophy

by

Rósín Byrne

based on research carried out at

School of Chemistry

&

Environmental Research Institute

Supervisors: *Dr. Stig Hellebust & Professor John Wenger*
Head of School: *Professor Anita Maguire*

2024

Contents

Contents ..... **i**

Declaration ..... **vi**

Acknowledgments ..... **vii**

Abstract ..... **viii**

List of Abbreviations and Acronyms ..... **x**

1 Introduction **1**

1.1 The Atmosphere **1**

1.2 Atmospheric Aerosols **4**

1.2.1 Sources 4

1.2.2 Chemical Composition 6

1.2.3 Health Impacts 7

1.2.4 Environmental Impacts 10

1.3 The Current State of Air Quality Assessments **13**

1.3.1 Historical Developments in Air Quality Assessments 13

1.3.2 The Legislative Basis of Air Quality Monitoring 20

1.3.3 Overview of Studies of Particulate Matter in Ireland 29

1.4 The Changing Landscape of Air Quality Assessments **34**

1.4.1 Development of Particulate Matter Sensors 34

1.4.2 Challenges Associated with Sensor Measurements 36

1.4.3 Calibrations for Sensor Data 39

1.4.4 Moving Beyond the Calibrations 41

1.5 Aims of Research 51

1.6 Overview of Thesis 52

2 Instrumentation and Methodology 54

2.1 Monitoring Locations 54

2.1.1 Sensor Network Locations 54

2.1.2 Additional Measurement Locations 59

2.2 Instrument Specifications 61

2.2.1 PurpleAir PA-II SD 61

2.2.2 Clarity Node-S 64

2.2.3 Beta-Attenuation Monitor 66

2.2.4 Fine Dust Monitor 69

2.2.5 Multi-Wavelength Aethalometer 70

2.3 Data Analysis Techniques 76

2.3.1 Multiple Linear Regression 76

2.3.2 Cluster Analysis 77

2.3.3 Spectral Analysis 85

2.3.4 Source Apportionment of Equivalent Black Carbon 90

3 Highly Local Sources and Large Spatial Variations in $PM_{2.5}$ Across a City: Evidence from a City-Wide Sensor Network in Cork, Ireland 93

3.1 Abstract 93

3.2 Introduction 95

3.3 Materials and Methods 99

3.3.1 Air Quality Sensors 99

3.3.2 Data Specification and Preparation 100
3.3.3 Air Quality Sensor Performance Evaluation 102
3.4 Results and Discussion 107
3.4.1 Spatial and Temporal Variation in $PM_{2.5}$ Pollution 107
3.4.2 Identification and Estimation of Local Emissions 112
3.4.3 Meteorological Factors 122
3.4.4 Comparison with Source Apportionment Studies 125
3.5 Conclusion 127
3.6 Appendix 1 129
4 Spatial Analysis of $PM_{2.5}$ Using a Concentration Similarity Index Applied to Air Quality Sensor Networks 134
4.1 Abstract 134
4.2 Introduction 136
4.3 Methodology 139
4.3.1 Data Collection, Preprocessing, and Calibration 139
4.3.2 Dungarvan $PM_{2.5}$ Sensor Network 139
4.3.3 Cork City $PM_{2.5}$ Sensor Network 141
4.3.4 Meteorological Measurements 142
4.4 Development of the Concentration Similarity Index 144
4.4.1 Original Function Application 144
4.4.2 Function Parameter Optimisation and Introduction of PM Limit 148
4.4.3 Development and Testing of the Modified Equation 150
4.4.4 Application to Sensor Networks and Analysis of Spatial Trends 157
4.5 Results and Discussion 158

4.5.1 Dungarvan $PM_{2.5}$ Sensor Network 158

4.5.2 Cork City $PM_{2.5}$ Sensor Network 168

4.5.3 Assessment of the Spatial Representativeness of Air Quality Monitoring Locations 175

4.6 Conclusion 178

4.7 Appendix 2 181

5 Spectral Analysis of Air Quality Sensor Network Data for Local Source Contribution Estimates of $PM_{2.5}$ 197

5.1 Abstract 197

5.2 Introduction 199

5.3 Methodology 202

5.3.1 Data Collection and Preprocessing 202

5.3.2 Air Quality Sensor Performance Evaluation 206

5.3.3 Fourier Transform and Periodogram Representation 207

5.3.4 Source Contribution Estimates 209

5.4 Results and Discussion 214

5.4.1 Sensor Performance Evaluation 214

5.4.2 Cork City 2021 Sensor Network 220

5.4.3 Cork City 2023 Sensor Network 226

5.4.4 Dungarvan Sensor Network 230

5.5 Limitations 236

5.6 Conclusion 237

5.7 Appendix 3 240

6 The Changing Nature of Black Carbon and $PM_{2.5}$ Emissions over Multiple Winters Characterised by Spectral Analysis and Aethalometer Data 251

6.1 Abstract 251

6.2 Introduction 253

6.3 Methodology 256

6.3.1 Campaign and Data Specification 256

6.3.2 Diurnal Profile Assessment Metrics 257

6.3.3 Source Apportionment 258

6.3.4 Spectral Analysis 259

6.4 Results and Discussion 262

6.4.1 Meteorological Variations 262

6.4.2 Temporal Variations in the Ambient Aerosol 264

6.4.3 Source Apportionment of eBC 275

6.4.4 Variations in the Relationship of eBC and $PM_{2.5}$ 283

6.4.5 Spectral analysis of $PM_{2.5}$ 287

6.5 Limitations 291

6.6 Conclusion 293

6.7 Appendix 4 296

7 Conclusion 301

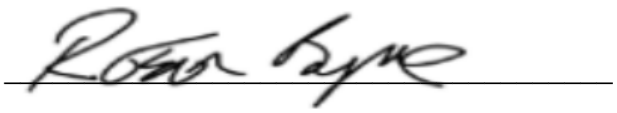
7.1 Summary 301

7.2 Perspectives 305

References 310

Declaration

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism and intellectual property.



Rósín Byrne

Acknowledgments

I would like to thank the Environmental Protection Agency, Ireland and the EU LIFE Programme for providing funding for this research and the opportunity to contribute to the LIFE EMERALD project (LIFE19 GIE/IE/001101).

To the CRAC group, I'm so glad to have gotten to know each of you over the past few years and I hope we can have many more pizza nights! To those specifically in the ERI, you made our time in the office all the better, and I will miss our summer lunches in the sun!

To my supervisors, Dr. Stig Hellebust and Professor John Wenger, my deepest thanks for everything you have done to help me along this journey. From my first time working with you both as an intern, to undertaking the PhD and every step along the way, it has always been such an enjoyable experience. This work would not have been possible without your expertise, guidance, and especially your patience.

Thank you to my parents, John and Teresa, and my siblings for your never-ending support. And finally, to Ciarán, your continued patience, understanding, and love along the way has made the world of difference. I am forever grateful to you.

Abstract

Recent improvements in air quality sensors (AQSs) have presented an affordable and scalable way to enhance air quality monitoring. However, fully exploiting these new data sources requires novel data-driven methodologies to address the inherent uncertainties in the measurements reported by AQSs. This work goes beyond calibration exercises to highlight the potential of AQS networks for air quality assessments, particularly after demonstrating good sensor-reference correlations and minimal inter-sensor variation. Using three distinct datasets from AQS networks in Cork City and Dungarvan, Ireland, a range of analytical approaches have been developed and applied to understand the local spatiotemporal variability of $PM_{2.5}$ in these urban areas.

A novel method leveraging high temporal resolution data from PurpleAir sensors linked rapid, short-lived concentration fluctuations to local emissions. The analysis revealed stark spatial contrasts in the proportion of local emissions across Cork City. On average, half of the total $PM_{2.5}$ exposure originated mainly from local emissions and occurred in just 30% of the time. Discrete Fourier transform spectral analysis was also used to estimate local $PM_{2.5}$ contributions, yielding comparable results when applied to the same dataset. Analysis of the temporal changes to the contribution of local emissions in Cork City between 2021 and 2023 suggested that legislative changes may be positively influencing the proportion of locally emitted $PM_{2.5}$ in winter, however, this effect is not uniform across the city.

A new metric, the concentration similarity index (CSI), was developed and optimised for AQS network data. The CSI was used to assess the spatial representativeness of Environmental Protection Agency (EPA) monitoring locations in both Dungarvan and

Cork City, revealing moderate exposure representation. However, significant spatial variability within each sensor network indicated strong differences in the degree of representation by the central monitoring locations and emphasised that location type, rather than geographical proximity, plays a key role in air quality representation.

Additionally, a frequency-based method was integrated with the analysis of winter EPA $PM_{2.5}$ data and aethalometer-derived equivalent black carbon (eBC) data in a small Irish town, to investigate the changing impact of local emissions over time. Analysis of data from 2018 to 2024 showed considerable reductions in winter pollution levels, suggesting an overall improvement in air quality in the area. Analysis of the Ångström absorption exponent indicated an evolving emissions profile, with an increase in wood or peat/turf burning in recent years, while spectral analysis showed general reductions in both regional/background and local particulate pollution.

This research offers insight into cost-effective, spatially dense air quality monitoring solutions that can inform vital pollution mitigation strategies. The new methodologies developed across all datasets, demonstrate the potential for integrating AQS networks and data-driven techniques into air quality monitoring and assessment.

List of Abbreviations and Acronyms

AAE	Absorption Ångström Exponent
AOD	Aerosol Optical Depth
API	Application Programming Interface
AQG	Air Quality Guideline
AQS	Air Quality Sensor
ATN	Attenuation
ATOFMS	Aerosol Time-of-flight Mass Spectrometer
BAM	Beta-attenuation Monitor
BB	Biomass Burning
BC	Black Carbon
CAFE	Clean Air for Europe
CCN	Cloud Condensation Nuclei
CF	Correction Factor
CMB	Chemical Mass Balance
CSI	Concentration Similarity Index
CSO	Central Statistics Office
CV	Coefficient of Variation
DFT	Discrete Fourier Transform
eBC	Equivalent Black Carbon

EC	European Commission
EEA	European Environment Agency
EMERALD	Emissions Modelling and Forecasting of Air in Ireland
EPA	Environmental Protection Agency
EU	European Union
FA	Factor Analysis
FAIRMODE	The Forum for Air Quality Modelling
FCM	Fuzzy C-means
FF	Fossil Fuel
FT	Fourier Transform
HC	Hierarchical Clustering
HR-TOF-AMS	High-resolution Time-of-flight Aerosol Mass Spectrometer
IADS	Intelligent Aerosol Drying System
IN	Ice Nuclei
IR	Infrared
JRC	Joint Research Centre
LED	Light-emitting Diode
LSCE	Local Source Contribution Estimate
MAC	Mass Absorption Cross-section
MAE	Mean Absolute Error
MISR	Multi-angle Imaging Spectroradiometer

MLR	Multiple Linear Regression
MODIS	Moderate Resolution Imaging Spectroradiometer
MTU	Munster Technological University
NMF	Non-negative Matrix Factorisation
OA	Organic Aerosol
OLS	Ordinary Least Squares
OPC	Optical Particle Counter
PA	PurpleAir
PM	Particulate Matter
PMF	Positive Matrix Factorisation
PTFE	Polytetrafluoroethylene
RPI	Relative Peak Increase
SF	Solid Fuel
SMPS	Scanning Mobility Particle Sizer
TEOM	Tapered Element Oscillating Microbalance
Tr	Traffic
UCC	University College Cork
UV	Ultra-violet
WB	Wood Burning
WHO	World Health Organisation

1 Introduction

1.1 The Atmosphere

Atmospheric dynamics are driven by a variety of chemical and physical processes. The lowest layer in the atmosphere, the troposphere, is turbulent and contains the majority of the water vapour, and essentially all clouds and precipitation. As the warmer air rises in the troposphere, strong vertical mixing occurs and species emitted at the Earth's surface can rise to the tropopause within a few days or less, depending on meteorological conditions (Finlayson-Pitts & Pitts, 2000). A multitude of species are present in the troposphere, with a variety of sources, sinks, and competitive reaction pathways. Many of the components in the air directly impact human health and ecosystems. Air pollutants may include gases, liquids, or solids suspended in the atmosphere in concentrations sufficiently high above their ambient levels as to affect human, animal, or vegetation health (Seinfeld & Pandis, 2016). There are numerous sources of air pollution including vehicular exhausts, solid fuel burning and natural processes, which release pollutants such as particulate matter (PM), nitrogen oxides, sulfur dioxide, and volatile organic compounds (VOCs) (Table 1.1). Air pollutants can contribute to issues such as urban smog and acid deposition. The atmospheric lifetimes, concentration, and distribution of air pollutants vary spatially and temporally, due to influences from local sources, atmospheric transformations, meteorological conditions, and topography.

Table 1.1: *Summary of the main sources, sinks, lifetimes, concentrations, and distributions for a variety of key atmopsheric pollutants (Finlayson-Pitts & Pitts, 2000; Seinfeld & Pandis, 2016).*

Pollutant	Main Sources	Main Sinks	Lifetime	Typical Concentration	Distribution
Particulate Matter (PM)	Combustion, industrial emissions, natural sources (dust, sea spray, fires)	Deposition, coagulation	Hours to weeks	10–100 μg m^{-3}	Regional to global, hotspots in urban or arid regions
Black Carbon (BC)	Incomplete combustion (biomass, fossil fuels)	Deposition, coagulation	Days to weeks	~1 μg m^{-3}	Local hotspots, transported globally in smaller quantities
Nitrogen Oxides (NOx)	Vehicle emissions, power plants, industrial processes	Reaction to form HNO_3 (via OH)	Hours to days	0.1–1 ppb	Urban and industrialised areas; dispersed globally in lower concentrations

Table 1.1 continued.

Pollutant	Main Sources	Main Sinks	Lifetime	Typical Concentration	Distribution
Sulfur Dioxide (SO_2)	Fossil fuel combustion, volcanic activity	Reaction to form sulfates (via OH)	Hours to days	0.1–1 ppb	Higher near industrial sources, declining regionally due to regulation
Ozone (O_3)	Secondary pollutant (NOx + VOCs in sunlight)	Reaction with NO, deposition, photolysis	Hours to days (in troposphere)	10–50 ppb	Global, higher in polluted areas and downwind regions
Volatile Organic Compounds (VOCs)	Solvents, vehicle emissions, natural sources (biogenic)	Reaction with OH radicals	Minutes to days	Variable	Local to regional, with hotspots near sources

1.2 Atmospheric Aerosols

Aerosols are relatively stable suspensions of solid or liquid particles in a gas (Tomasi & Lupi, 2017). The suspended particles, also known as particulate matter (PM), range in sizes from approximately 0.002 μm to 100 μm in aerodynamic diameter (Finlayson-Pitts & Pitts, 2000). Larger particles, similar in size to fine drizzle or very fine sand, do not remain suspended for significant periods of time due to their tendency to fall quickly from the atmosphere. Particles in the 0.002 μm to 10 μm size range are the most significant with respect to atmospheric and physical chemistry (Finlayson-Pitts & Pitts, 2000).

Aerosols can influence atmospheric processes in different ways, such as through the absorption and scattering of light, or serving as condensation surfaces for water vapour (Kolb & Worsnop, 2012). They can also strongly affect human health, exacerbating or causing noncommunicable diseases (Pope & Dockery, 2012). The degree of environmental and health impact is determined by parameters such as the size, concentration, and chemical composition of the particles. These particle characteristics are used to establish the origin of particles for effective reduction and pollution mitigation strategies.

1.2.1 Sources

Atmospheric aerosols are comprised of primary particles, which are emitted directly to the atmosphere, and secondary particles, which are formed in the atmosphere through chemical reactions or physical processes involving precursor gases. PM is produced through three main processes – bulk-to-particle conversion, gas-to-particle conversion, and combustion processes (Tomasi & Lupi, 2017). Bulk-to-particle

conversion involves the fragmentation of larger materials into airborne particles, such as mineral dust from the Earth's crust, sea-salt particles from marine water, and biological aerosols from plant debris, pollen, and animals. Gas-to-particle conversion occurs when condensable vapours (both organic and inorganic compounds) lead to the formation of new particles or the growth of existing ones. Combustion processes are high-temperature processes which are a combination of both bulk-to-particle and gas-to-particle conversions, producing particles with unique shapes and composition not achievable by the other two methods alone (Tomasi & Lupi, 2017).

Aerosols originate from both natural and anthropogenic sources. Natural sources include pollen and spores, wind-born dust, sea spray, wildfires, and volcanic eruptions. Anthropogenic sources include traffic related emissions (fossil fuel combustion, brake and tyre wear, particle resuspension on roads), waste and biomass burning, industrial activities, and wind-forced mineral dust in agriculturally exploited areas (Tomasi & Lupi, 2017). The larger particles from anthropogenic emissions (>5 μm diameter) are mainly primary particles rather than secondary particles. However, gas-to-particle conversion and combustion are the main sources of anthropogenically derived smaller particles. (Finlayson-Pitts & Pitts, 2000; Hobbs, 2000). While anthropogenic aerosols are typically concentrated in urban environments, rural areas can experience significant pollution due to agricultural activities and solid fuel combustion for domestic heating (Chen et al., 2017; Han et al., 2020; Zhang et al., 2021). The long-range transport of PM can also impact both urban and rural areas (Putaud et al., 2004).

The atmospheric residence times, or lifetimes, of aerosol particles can vary from hours to weeks, depending on their properties and the prevailing meteorological conditions (Raes et al., 2000; Williams et al., 2002). Aerosols can be carried in airflows and transported over regional, intercontinental, and even global, scales in the atmosphere.

Regional transport is likely if the aerosol is produced by gas-to-particle conversion due to the time the conversion takes and the relatively small size of the particles leading to longer residence times in the atmosphere (Hobbs, 2000). Saharan dust storms frequently result in high dust loadings over Europe and the dust particles can also travel as far as the Amazon Basin and the Caribbean (Alastuey et al., 2016; Niedermeier et al., 2014; Resch et al., 2008).

1.2.2 Chemical Composition

Once airborne, particles can change size and composition through physical processes such as condensation and evaporation of species, nucleation, particle coagulation, and through chemical reactions, such as oxidation or activation in the presence of water supersaturation to become fog and cloud droplets (Pandis et al., 1995). The main chemical components of PM, by mass, generally include sulfate, nitrate, ammonium, sea salt, mineral dust, organic compounds, and black or elemental carbon (Pöschl, 2005). However, the components and relative contributions can vary significantly depending on location, season, and meteorological conditions. As PM is often a complicated mixture of particles from various sources it is generally categorised according to the aerodynamic diameter of the particles, where PM_{10}, $PM_{2.5}$, and PM_1 refer to particles with an aerodynamic diameter less than 10 μm, 2.5 μm, and 1 μm, respectively. Particulate matter can be measured and expressed in terms of particle counts per size bin per cubic centimetre of air, or in terms of mass concentration, i.e. mass per unit volume of air ($\mu g\ m^{-3}$).

Carbonaceous aerosol refers to the aerosols containing carbon as a fraction of the total aerosol. Black carbon (BC) is a distinct fraction of the carbonaceous aerosol in the

atmosphere and is primarily formed during the incomplete combustion of fossil fuels and biomass. BC is a highly absorbing refractory material and is resistant to chemical transformation (Bond et al., 2013). The term BC is mainly used when referring to the chemical composition, however the concentration of BC is typically measured entirely through optical methods. In line with the recommendations by Petzold et al. (2013) to emphasise the measurement technique, the optically measured BC concentration is referred to as equivalent BC (eBC). The term eBC refers to BC concentrations determined using a combination of optical methods and a suitable mass absorption cross-section (MAC) value. Employing the term eBC underlines that the concentration measurements were possibly influenced by certain factors affecting the absorption of light of the BC particles. For instance, a large agglomerate, a large former agglomerate (a collapsed particle), and a small individual carbon spherule all absorb light differently for the same mass of elemental carbon (Bond et al., 2006). If a BC core gathers a coating with specific optical properties, it can enhance the absorption of the BC core by acting as a lens to gather the light rays towards the core from a larger cross section (Bond et al., 2006; Lack & Cappa, 2010).

1.2.3 Health Impacts

Air pollution poses significant risks to human health. Despite the increase in knowledge about the impacts of air pollution on morbidity and mortality, the number of global deaths attributable to ambient $PM_{2.5}$ continues to increase (Fuller et al., 2022; Sang et al., 2022). The World Health Organisation (WHO) estimates that approximately 7 million deaths occur annually as a consequence of air pollution (World Health Organisation, 2024). The Global Burden of Disease study in 2021 concluded that 8% of the total disease burden in 2021 was caused by PM air pollution,

making it a leading contributor among all risk factors . The percent of deaths attributable to air pollution varies with cause of death/injury (Table 1.2).

Table 1.2: *Top five causes of death attributable to air pollution, categorised by health outcome, with the percentage of total deaths linked to air pollution for each group.*

Cause of Death/Injury	**Percent of Deaths Attributable to Air Pollution %**
Chronic obstructive pulmonary disease	48
Chronic respiratory diseases	40
Neonatal preterm birth	31
Intracerebral haemorrhage	30
Lower respiratory infections	30

The health impacts arising from particulate matter are mainly due to their small size. Figure 1.1 shows how particle diameter size and the respective lung penetration capabilities differ (Encyclopaedia of the Environment, 2019). Larger particles (5−30 μm) are generally removed in the head or upper respiratory tract with deposition in the nasopharyngeal region, while smaller particles (< 5 μm), such as sea spray and dust particles, and the smaller fossil fuel combustion and gas-to-particle conversion related particles, can travel further into the body after inhalation and to be deposited at the gas-exchange alveolar region (Finlayson-Pitts & Pitts, 2000). Due to its prevalence, $PM_{2.5}$, in particular, poses severe health risks as the ultrafine particles ($PM_{0.1}$) can pass

the air-blood tissue barrier, entering the blood and lymphatic circulatory system to potentially reach any organ in the body.

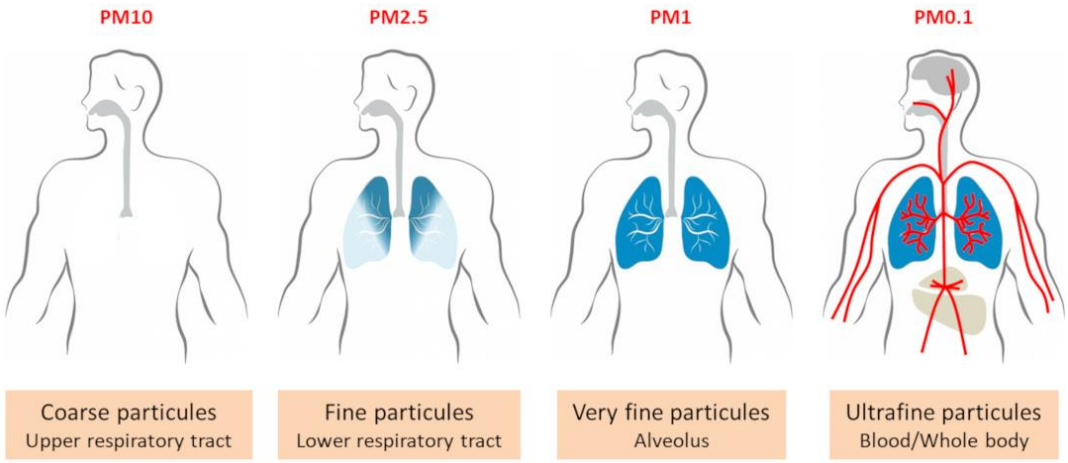


Figure 1.1: *Differences in the lung penetration depth of particles according to diameter size (Encyclopaedia of the Environment, 2019).*

The WHO International Agency for Cancer Research (IACR) categorises overall air pollution and specifically particulate pollution as carcinogenic to humans (WHO IARC, 2020). Studies have found links between particulate pollution exposure and various cancers, diabetes, cardiovascular problems, and asthma (Cesaroni et al., 2014; Donaldson et al., 2001; Krämer et al., 2010; Pope et al., 2002; Raaschou-Nielsen et al., 2013a; Samoli et al., 2013; Weinmayr et al., 2010). Exposure to PM has also been suggested as a possible risk factor for vascular dementia and Alzheimer's disease (Jung et al., 2015; Oudin et al., 2016). Numerous studies have reported statistical associations between maternal exposure to ambient pollution and negative pregnancy outcomes (pregnancy loss, premature births, low birth weight, congenital heart defects) in both rural and urban areas, although conclusive evidence is somewhat limited (Bachwenkizi et al., 2022; Bravo & Miranda, 2022; Hyder et al., 2014; Kloog

et al., 2012; Wojtyla et al., 2020; Xue et al., 2019; Zhang et al., 2016). Bové et al. (2019) found that BC particles can accumulate on the foetal side of the placenta, suggesting that the human placental barrier is not entirely impermeable to particles, and this could lead to particle transportation towards the foetus. Evidence also suggests that BC may pose higher levels of toxicity than overall $PM_{2.5}$ (Grahame & Schlesinger, 2010; Janssen et al., 2011).

While chronic exposure to elevated levels of ambient PM poses severe health risks (Cesaroni et al., 2014; Pope et al., 2002; Pope & Dockery, 2012; Sanyal et al., 2018), numerous studies have also linked increased morbidity and mortality due to respiratory and cardiovascular diseases with acute exposure to PM (Fajersztajn et al., 2017; Orellano et al., 2020; Weinmayr et al., 2010). The lack of spatially dense PM monitoring has been cited as a limitation to studies investigating the links between PM and adverse health outcomes and accurately assessing human exposure to ambient pollution (Anderson et al., 2012; Bai et al., 2022; Fiordelisi et al., 2017). Nonetheless, different types of studies conducted in different areas continue to find similar results regarding the negative effects of PM.

1.2.4 Environmental Impacts

Both natural and anthropogenic aerosols can have considerable environmental impacts influenced by their chemical and physical properties. Aerosol effects on climate are generally classified as direct or indirect radiative forcing effects. Radiative forcing refers to a change in the energy flow of solar radiation and terrestrial radiation in the atmosphere, i.e., a measure of a change in the Earth's radiation balance (Pöschl, 2005). These changes can occur from both natural and anthropogenic changes in the

composition of the atmosphere, the Earth's surface properties, or solar activity. Negative radiative forcing results in a cooling effect, while positive radiative forcing produces a warming effect. While most aerosols contribute a dominant negative effect, this is partially offset by a warming contribution due to the solar radiation absorption capabilities of BC particles (IPCC, 2014).

Direct radiative effects are caused by the light scattering and absorption capabilities of aerosols. The optical properties of aerosols, such as the scattering and absorption coefficient, refractive index, and the extinction cross-section, and thus the direct radiative efficiency, are determined by the particle size, structure, and chemical composition, and the wavelength of the incoming light (Bellouin et al., 2020; Pöschl, 2005). Scattering of radiation causes a cooling effect due to an increase in the Earth's albedo, i.e. an increase in the Earth's reflected radiation compared to the total incoming solar radiation. In contrast, a purely absorbing aerosol, such as black carbon, contributes to atmospheric warming by absorbing solar radiation, and to a lesser extent, infra-red (IR) radiation emitted by the Earth's surface, trapping heat in the atmosphere and producing a warming effect. While this leads to localised heat trapping in the atmosphere, the effect is distinct from the global greenhouse effect caused by long-lived greenhouse gases. The semi-direct effect is an additional consequence of absorbing aerosols, which involves changes in cloud properties due to the absorption of solar radiation by certain aerosols, leading to the evaporation of cloud droplets (Lohmann & Feichter, 2005). The droplet evaporation can alter the moisture distribution leading to changes in the relative humidity. The temperature increase from absorbing aerosols can reduce the relative humidity in the heated region as relative humidity and temperature are inversely related.

Indirect radiative effects occur when aerosols influence cloud microphysics, affecting the cloud reflectivity and persistence. Aerosols can act as cloud condensation nuclei (CNN) and ice nuclei (IN) to form large concentrations of small cloud droplets, prolonging the time it takes for cloud water existing in small drops to convert into larger masses that can precipitate. The suppression of precipitation is stronger in shallower clouds with warmer top temperatures. Aerosol CNN can also enhance the reflection of solar radiation by the clouds (Givati & Rosenfield, 2004).

Certain aerosols can cause additional atmospheric and health impacts. For instance, BC particles can catalyse the formation of nitrous acid in polluted air masses, which is an important precursor for the hydroxyl radical, the main tropospheric oxidising agent (Ammann et al., 1998). BC-catalysed SO_2 oxidation in the presence of NO_2 and ammonia has also been shown to occur in regional haze (Zhang et al., 2020).

1.3 The Current State of Air Quality Assessments

The need for effective air quality monitoring has become increasingly important as awareness around the associated health risks of poor ambient air quality grows. Additionally, anti-smoking campaigns, the health-related impacts of the COVID-19 pandemic, and concerns about indoor damp-related issues have heightened awareness regarding indoor air quality. Assessing ambient air quality involves a range of methodologies to monitor and analyse pollutant levels, with traditional approaches employing fixed, ground-based monitoring equipment. Advancements in technology have allowed for remote sensing and mobile monitoring units, enhancing the spatial and temporal resolution of measured data. Additionally, the development of source apportionment techniques to quantitatively determine the contribution of important PM sources has allowed for more effective and efficient air quality management strategies.

1.3.1 Historical Developments in Air Quality Assessments

While air pollution is often regarded as a modern-day issue, poor air has been recognised as a problem since early society, despite the nature of the gases and particles being unknown (Fowler et al., 2020). Residents of ancient Rome complained about the effects of air pollution – "As soon as I escaped from the oppressive atmosphere of the city, and from that awful odour of reeking kitchens which, when in use, pour forth a ruinous mess of steam and soot, I perceived at once that my health was mending" (Mustard & Gummere, 1917). The rapid industrialisation in London led to reports describing the clouds of smoke, suggesting the relocation of certain industrial activities to the countryside to reduce the scale of the problem in the city

(Fowler et al., 2020). However, there was still little understanding of the underlying atmospheric chemistry. Air quality monitoring and assessments have evolved significantly over time, driven not only by advances in technology but also by an increasing awareness of the negative health impacts of air pollution.

1.3.1.1 Particulate Matter Measurements

In 1882, meteorologist John Aitken began investigating soot particles in the air, leading to the development of a method for counting the number of dust particles in a sample of air and the systematic observation of aerosols in a variety of U.K. and continental Europe locations (Aitken, 1889, 1895). His particle counting instrument involved rapidly expanding a sample of air in a sealed chamber causing the temperature to decrease, which in turn caused the moisture in the sample to condense around any particles present due to the relative humidity exceeding 100% during the rapid cooling. The resulting droplets made the particles visible and countable under a microscope. This work was a vital step in early aerosol research, allowing for the assessment of particle concentrations and highlighted the importance of CNN in the formation of clouds and fog.

Beyond the pioneering work of Aitken, the measurement of PM has undergone significant evolution. The development of gravimetric analysis, involving the collection of airborne particles on a filter and weighing the filter before and after collection, provided a simple yet effective way to measure the total mass concentration of PM (Beamish, 1949; Dunmore et al., 1964; Feeney et al., 1984; Goetz, 1953; Vaughan et al., 1989). Early applications of this method involved collecting large particles by gravitational settling in clean, open-mouth containers over a period of

several weeks. However, bird droppings, insects, algae, and fungi were often found to make up a large portion of the measured mass (Chow, 1995). While early gravimetric analysis helped lay a foundation for subsequent air quality monitoring efforts, it was initially limited by its inability to provide information on particle size distribution or chemical composition. Further advancements regarding size-selective sampling techniques, such as cyclone particle inlets, allowed for the size differentiation of incoming particles. Determining the net gain in weight from deposition of aerosol is complicated by various processes including the small weight of the particle deposit compared to the weight of the filter, the adsorption of gases onto the filter, retention of water by the filter and the deposited sample, electrostatic charges resulting from the attraction of the filter to balance surfaces, and potential contamination during filter handling (Watson et al., 2017). Due to the typically long averaging times associated with gravimetric samplers to collect enough mass, they cannot show short-term temporal patterns. However, gravimetric analysis is a well-developed and common method for aerosol sampling that continues to effectively provide highly accurate measurements of particle mass concentration (Aarhaug et al., 2024; Chatoutsidou & Lazaridis, 2022; Sousa et al., 2024).

The introduction of continuous PM monitors, such as the beta-attenuation monitor (BAM) (Met One Instruments, 2008), the tapered element oscillating microbalance (TEOM) (Thermo Fisher, 2018), and optical light scattering based monitors, enabled real-time measurements of PM concentrations. These devices offer improved temporal resolution, facilitating the study of diurnal variations in PM levels. The development of remote sensing techniques allowed for the measurement of PM on a global scale with satellite-based instruments, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Multi-angle Imaging SpectroRadiometer

(MISR). These instruments provide data on aerosol optical depth, which can be used to estimate ground-level PM concentrations (Kahn et al., 2005; Kahn & Gaitley, 2015).

1.3.1.2 Source Apportionment

Advancements in chemical analysis methods enabled the detailed characterisation of PM composition and identification of the sources of PM pollution. Insights into the relative contributions of different sources, such as traffic and solid fuel burning related emissions and industrial activities, is vital for pollution mitigation and exposure reduction strategies (Hopke, 2016).

Source apportionment models typically begin with observed ambient particle concentrations at a receptor site and aim to apportion the observed measurements between several source types based on knowledge of the compositions of the source and receptor materials (Henry et al., 1984). The chemical mass balance (CMB) model was one of the first quantitative source apportionment methods developed (Miller et al., 1972; Winchester & Nifong, 1971). The CMB model uses the chemical composition of collected PM samples where elements are associated with particular source types to develop a mass balance for airborne particles and known source profiles are used to estimate the contribution of each source to the observed PM.

Multivariate receptor models such as factor analysis (FA) have been successfully developed and applied for PM source apportionment analysis (Blifford & Meeker, 1967; Hopke et al., 1976; Laamanen & Partanen, 1969). FA identifies factors, or sources, which explain the variance in the observed data. The correlation matrix for a matrix of size $m \times n$, where m is the number of samples (e.g. days or hours) and n is the number of variables (e.g. different pollutant concentrations), is computed to

understand the relationship between the variables. The objective is to explain the maximum variance in the data using a smaller number of factors than the original variables, and the factors can be linked to specific emission sources. Early applications did not provide quantitative source apportionment of PM mass or of specific elemental concentrations, however this problem was solved with further developments on alternative forms of FA (Hopke, 1988 and references therein). Positive matrix factorisation (PMF), a variant of FA, also decomposes the observed pollutant data matrix into factors representing different sources, however it accounts for uncertainties in the data and introduces constraints, such as non-negativity of the factor contributions, ensuring the outputs are physically meaningful (Paatero & Tapper, 1994). In contrast to CMB analysis, FA methods do not require prior knowledge of the source profiles, making them useful for exploratory analysis. However, they typically require detailed chemical composition, and additional information, such as co-located meteorological parameters, which can be resource intensive. In addition, these methods are often difficult to employ with high spatial resolution due to the expense of deploying multiple high-cost instruments with high spatial density.

In the context of BC measurements, multi-wavelength aethalometers offer a way to apportion BC to two different sources, i.e. fossil fuel combustion and biomass burning. Aethalometers operate on the principle of the continuous measurement of the attenuation of a beam of light transmitted through a filter (Hansen et al., 1984) and can provide high time-resolution measurements (1-minute intervals). Multi-wavelength aethalometers operate on the same principle but measure the attenuation over a range of wavelengths, typically in the range of 370–950 nm. By measuring the light attenuation at the different wavelengths and pre-defining a MAC value for each wavelength, the concentration of BC is derived. As a consequence of incorporating

multiple wavelength measurements and the differences in absorption efficiency of BC particles of different origins, a source apportionment model was developed for multi-wavelength aethalometer data (Sandradewi et al., 2008b). BC derived from the burning of fossil fuels typically shows stronger absorption in the near-IR region, whereas biomass burning derived BC has higher absorption in the ultra-violet (UV) region.

1.3.1.3 Spatial Representativeness Assessment

Spatial representativeness refers to the spatial extent over which measurements at a specific location reflect the pollutant concentrations and population exposure (EC, 2018). Assessing the spatial representativeness of PM is another aspect of air quality assessments vital for accurate and effective air quality monitoring and management, and the limited numbers of monitoring stations and low spatial representativeness of existing ones is a barrier for accurate pollution exposure assessments and often leads to bias and exposure misclassifications (de Preux et al., 2023). While the European Union (EU) states that air pollutant sampling locations should be representative of similar locations outside of their immediate vicinity and it recognises spatial representativeness as a core principle for monitoring network design, there is, however, no unified agreement in scientific literature or in EU regulations clearly defining spatial representativeness and how to determine the representativeness of monitoring stations (Kracht et al., 2018).

The issue of providing a harmonised approach to assess spatial representativeness is complex. It is difficult to specifically define how the similarity between pollutant concentrations obtained at a monitoring location and those experienced in the broader surrounding area should be assessed (EC, 2018), and because data are representative

only in the context of a specific application or area, it is difficult to derive a fully objective definition for “representativeness” (Wieringa, 1996). Pollution variability is dependent on the specific characteristics of each pollutant, such as the nature of the source and pollutant dispersion tendencies in the atmosphere (due to meteorological conditions and chemical reactivity). Therefore, an area may be well represented by a monitoring station in terms of one pollutant but not of another. The forum for air quality modelling (FAIRMODE), an initiative of the European Environment Agency (EEA) and the European Commission (EC) Joint Research Centre (JRC), is actively promoting discussions concerning spatial representativeness in order to address this gap (EEA & JRC, 2007; Janssen et al., 2023). FAIRMODE have recognised the variation in spatial representativeness assessment estimates when different approaches are taken to investigate the same monitoring locations. An example of this is illustrated in Figure 1.2, which shows spatial representativeness estimates from six different modelling approaches (Kracht et al., 2017).

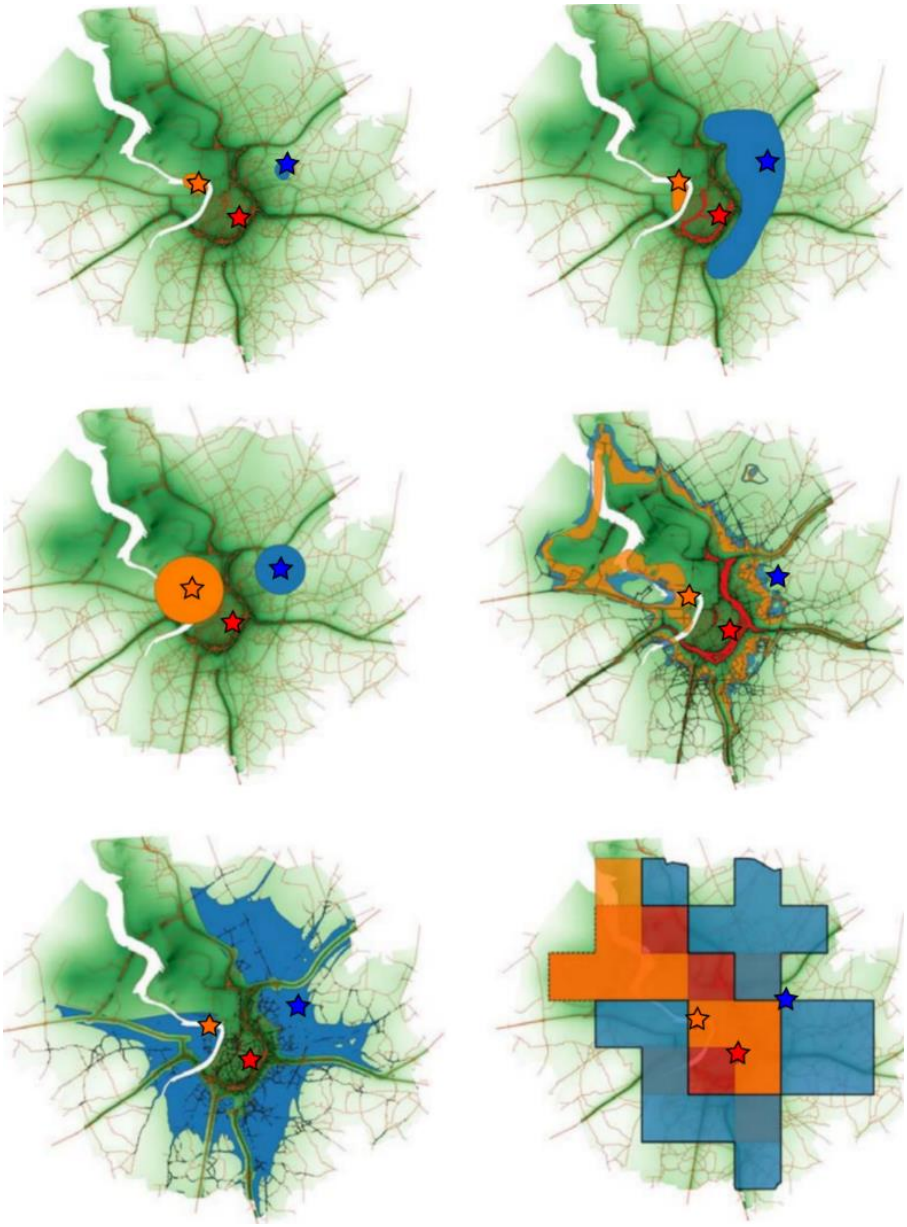


Figure 1.2: *Six different estimates of spatial representativeness of NO_2 in Antwerp, Belgium. Coloured areas delineate pollution exposure represented by the concentrations at the respective monitoring locations (stars). The results emphasise the variations in representativeness depending on the method used. Stations include the urban background sites Linkeroever (orange) and Schoten (blue), and the traffic station Borgerhout (red). Adapted from Kracht et al. (2017).*

1.3.2 The Legislative Basis of Air Quality Monitoring

1.3.2.1 Air Quality Guidelines and Regulations in Europe

The first regulations and laws to reduce air pollution were introduced following a number of extreme air pollution episodes during the 1930s–1960s, including the

famous winter smogs in Meuse Valley, Belgium and London, U.K., which were both caused by coal combustion (Brimblecombe, 2012; Haldane, 1931). These events highlighted the urgent need for policies to protect public health, leading to the first attempts to curb air pollution. However, more comprehensive and structured frameworks emerged later, driven by growing scientific evidence on the health impacts of air pollutants.

In 1987, the WHO introduced its first air quality guidelines (AQGs) for Europe, and while not legally binding, the guidelines are evidence-based recommendations aimed at protecting public health by setting threshold limits for specific air pollutants, including PM, NO_2, and SO_2. The guidelines offered benchmarks for European countries to structure their national air pollution mitigation strategies (WHO, 1987). Specific guideline limits for $PM_{2.5}$ were introduced in the 2005 update, and these limits have subsequently been revised and updated periodically, reflecting advances in scientific understanding and changes in air quality across Europe (Table 1.3) (WHO, 2005). In 2021, the WHO, in recognising the need for continued efforts to mitigate pollution levels, set stricter pollutant concentration targets. For $PM_{2.5}$, the annual mean guideline limit was reduced from 10 μg m^{-3} to 5 μg m^{-3}, and the mean 24-hour AQG was reduced from 25 μg m^{-3} to 15 μg m^{-3}.

Table 1.3: *Progression of the WHO Air Quality Guidelines (AQGs) for $PM_{2.5}$, showing the recommended annual and 24-hour guideline levels in the 2005 and 2021 updates, reflecting advancements in scientific understanding of the health impacts of $PM_{2.5}$ exposure.*

Year	Guideline Level for Annual $PM_{2.5}$ ($\mu g\ m^{-3}$)	Guideline Level for 24-hour $PM_{2.5}$ ($\mu g\ m^{-3}$)
2005	10	25
2021	5	15

The EU implemented legally binding air quality standards through the Clean Air for Europe (CAFE) Directive (2008/50/EC) in May 2008. The directive consolidated existing air quality legislation into a single, comprehensive framework. A primary – objective of the directive was to establish limit values for key pollutants, including $PM_{2.5}$, PM_{10}, O_3, and NO_x, and it defined maximum allowable concentrations and established timelines for EU Member States to comply with these standards (The European Parliament and of the Council of the European Union, 2008). For $PM_{2.5}$, the CAFE directive set a mean annual concentration limit of 25 $\mu g\ m^{-3}$ to be met by 2015, and a lower annual mean limit of 20 $\mu g\ m^{-3}$ to be met by 2020, which is considerably less stringent than the updated WHO guideline of 5 $\mu g\ m^{-3}$ (EEA, 2024a). Notably, the directive did not establish a 24-hour exceedance value for $PM_{2.5}$, in contrast for the WHO's 15 $\mu g\ m^{-3}$ target. However, for PM_{10}, the directive imposes a 24-hour mean limit of 50 $\mu g\ m^{-3}$, with a maximum of 35 exceedances permitted annually. The directive introduced the concept of exposure reduction targets and average exposure indicators with the aim to gradually reduce population exposure to harmful pollutants over time, beyond simply focusing on compliance with fixed concentration limits.

In order to ensure compliance and promote a harmonised approach to air quality assessments, the directive mandates that EU Member States establish and maintain ambient air quality monitoring networks, outlining how and where air quality should be measured. Standardised methods are required for pollution monitoring to ensure the data collected across the EU is comparable and of similarly high quality. Member States must divide their territories into zones, ensuring coverage in urban, suburban, and rural areas, with the aim to allow for a comprehensive understanding of both regional and local air pollution levels and enable targeted interventions according to standardised methods and criteria (The European Parliament & the Council of the European Union, 2008). The design of air quality monitoring networks is an important factor in accurate exposure assessments, and often differences in design can lead to assessments that are difficult to compare (Duyzer et al., 2015). The EU directives require member states to locate sampling points both "where the highest concentrations occur" and other areas "which are representative of the exposure of the general population". However, it is not always clear if EU Member States monitor effectively according to these guidelines (EC, 2019). As the assessment of spatial representativeness of monitoring networks is not currently defined in EU legislation, the effectiveness of monitoring network design and suitability to assess population exposures can be hindered (Kracht et al., 2018).

Despite efforts to reduce emissions and improve air quality across Europe, significant discrepancies persist between the WHO recommended limits and the EU legally binding thresholds. According to the European Environment Agency (EEA), an estimated 96% of the EU's urban population is continually exposed to $PM_{2.5}$ levels above those recommended by the WHO, even though emissions have been steadily

decreasing (EEA, (2024b)). In contrast, less than 1% of the EU's urban population is exposed to $PM_{2.5}$ levels above the annual EU limit value (Figure 1.3).



Figure 1.3: *Proportion of the EU urban population exposed to $PM_{2.5}$ concentrations above EU standards and WHO guidelines in 2022. The EU standard reflects annual exposure above 20 μg m^{-3}. WHO guideline exposures reflect annual exposure above 5 μg m^{-3}. Adapted from EEA (2024b).*

Analysis into the effectiveness of the CAFE directive found that there is an overall downward trend in air pollution levels across Europe, and that fewer Member States report exceedances compared to previous years (EC, 2019). However more stringent compliance values were recommended, with a particular focus on $PM_{2.5}$, which poses significant risks to human health even at lower concentrations. In October 2022, the EC proposed a revision of the CAFE directive to provide more ambitious targets for air pollution levels and to better align with the current recommendations of the WHO (EC, 2022). In April 2024, the European Parliament formally adopted a revised provisional political agreement with EU Member States on the new measures to improve air quality throughout the EU (Halleux, 2024). In line with continually

emerging scientific evidence, the updated directive includes new limits, target values, and long-term objectives that are more stringent than the 2008 directive. $PM_{2.5}$ is recognised as a main pollutant, and the annual limit value is reduced by more than half, to 10 μg m^{-3}, and a 24-hour mean limit value is introduced for the first time (25 μg m^{-3}, not to be exceeded more than 18 times in one year). Progressive targets will be introduced for 2030, 2040, and 2050. In addition, there are mechanisms included for regular reviews of the standards and limits in line with the latest scientific information. The revision also seeks to provide better support for local authorities within Member States in their attempts to achieve cleaner air, with an emphasis on strengthening air quality monitoring and modelling.

1.3.2.2 Air Quality Measurement Standards in Europe

The EU has established specific standards and classifications for the monitoring of PM to ensure accurate and consistent air quality measurements that meet the data quality objectives set out in the CAFÉ directive. PM monitors can be classed as either reference grade monitors or reference-equivalent grade monitors. The reference grade classification applies to those monitors that meet the highest standards of accuracy and reliability as defined by the EU and are used for regulatory purposes and are typically subject to rigorous testing and certification processes. Gravimetric samplers are the primary reference grade method for PM measurements according to the EU standards. The two EU standards for reference grade monitors are EN 12341 (Ambient air – Standard gravimetric measurement method for the determination of PM_{10} or $PM_{2.5}$ mass concentration of suspended particulate matter) and EN 14907 (Ambient air quality – Standard gravimetric measurement method for the determination of the $PM_{2.5}$

mass fraction of suspended particulate matter), where the latter standard is similar to the former but has a specific focus on $PM_{2.5}$.

Continuous monitoring samplers are typically certified as reference-equivalent grade monitors. To obtain this certification the monitors must be validated against reference grade monitors and demonstrated to produce measurements under specific conditions that meet key performance criteria for accuracy, precision, linearity, and stability. The EU reference-equivalent grade standard is EN 16450 (Automated measuring systems for the measurement of the concentration of particulate matter (PM_{10}; $PM_{2.5}$).

1.3.2.3 Air Quality Legislation in Ireland

In Ireland, air quality legislation has evolved significantly, with legislative framework being shaped by national needs and EU directives. The Air Pollution Act 1987 aimed to control emissions from industrial sources, domestic heating, and transport by establishing monitoring networks and setting air quality standards (The Government of Ireland, 1987). Under this act, it was the role of the local authorities to undertake whatever measures they consider necessary to prevent, limit, and regulate air pollution in their area, through various monitoring and enforcement efforts. The Retail Sale of Solid Fuel Regulations (1989) mandated that all solid fuel retailers in Dublin provide low-smoke alternative fuels to existing fuel types (The Government of Ireland, 1989). The Marketing, Sale, and Distribution of Fuels Regulations (1990), commonly referred to as the "smoky coal ban", prohibited the sale of certain high-polluting fuels, such as bituminous coal, in Dublin (The Government of Ireland, 1990). Cork City was included as a restricted area for solid fuel sale in 1995, and further amendments included additional cities and towns where the population exceeded 10,000. The Solid

Fuels Regulations (2022) replaced the 1990 regulations, with the objective of further limiting the impact of domestic solid fuel combustion by extending the restrictions nationwide. Retailers are required to adhere to certain criteria regarding the solid fuel they sell such as smoke emission rates and sulfur content for coal and part-biomass products, as well as moisture content for 100% biomass products. The regulations also included stricter fuel requirements to be introduced in 2025.

The EU's CAFE directive was transposed into Irish law in 2011, introducing national standards for ambient air pollution, establishing air quality zones, and requiring action plans for areas with pollution exceedances (The Government of Ireland, 2011). The Environmental Protection Agency (EPA) is now the primary body responsible for air quality monitoring in Ireland, enforcing regulations, and providing air quality reports. The National Clean Air Strategy published in 2023 outlined a comprehensive approach for improving air quality across Ireland and ensuring compliance with EU and national standards (The Government of Ireland, 2023).

There are four air quality zones in Ireland, Zones A–D (EPA, 2024). Zone A and Zone B relate to Dublin City and surrounding suburban areas, and Cork City and surrounding suburban areas, respectively. Zone C consists of other cities and large towns (population >15, 000), and Zone D is rural Ireland, i.e. the rest of the State excluding those locations in Zones A–C. The number of reference/reference-equivalent grade monitoring stations per zone are listed in Table 1.4 and shown in Figure 1.4. The national ambient air quality monitoring network consists of a total of 116 monitoring stations, with $PM_{2.5}$ monitored at 75 locations in 2022 (EPA, 2023). There are 31 local monitoring sites which are established in partnership with local authorities to provide indicative measurements in smaller communities than those covered with the monitoring locations in Zone C (EPA, 2015). There are also 10

locations included in the European monitoring and evaluation programme (EMEP), which continuously monitor for transboundary air pollution.

Table 1.4: *Summary of regulatory monitoring in air quality monitoring zones in Ireland.*

Zone	Areas covered	Number of monitoring stations
A	Dublin City and surrounding suburban areas	19
B	Cork City and surrounding suburban areas	6
C	Limerick, Galway, Waterford, Drogheda, Dundalk, Bray, Navan, Ennis, Tralee, Kilkenny, Carlow, Naas, Sligo, Newbridge, Mullingar, Wexford, Letterkenny, Athlone, Celbridge, Clonmel, Balbriggan, Greystones, Leixlip, Portlaoise.	21
D	Rural Ireland	20

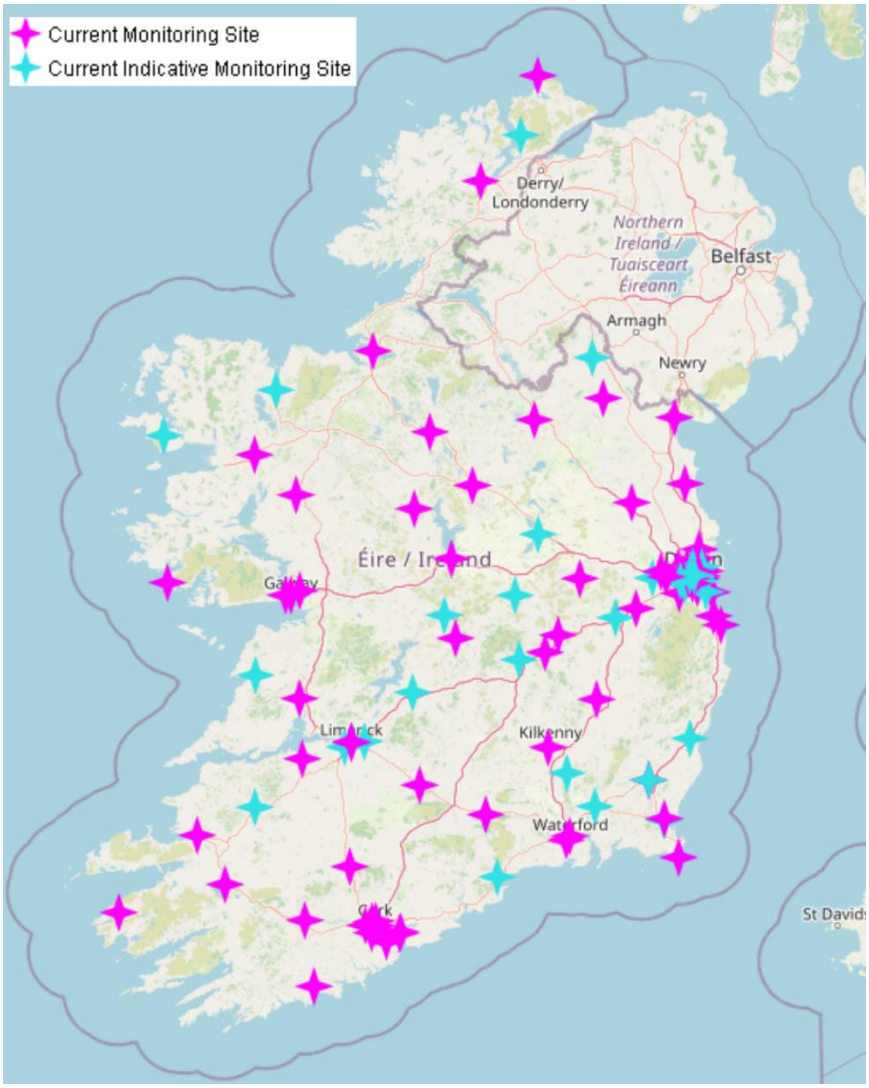


Figure 1.4: *Map of EPA regulatory air quality monitoring stations. Marker colours represent the type of monitoring site – pink markers are reference stations, and blue markers are indicative stations (EPA, 2024).*

1.3.3 Overview of Studies of Particulate Matter in Ireland

The in-depth examination of PM in Ireland gained attention due to its impact on air quality and subsequent health effects on the population, thus the sources, composition, spatial distribution, and health implications of PM pollution in Ireland have been investigated in numerous studies.

A report prepared for the Irish EPA by Jennings et al. (2000) found that local sources accounted for at least 50% of the $PM_{2.5}$ mass measured at urban roadside, urban centre, and urban background locations. Yin et al. (2005) collected 24-hour PM_{10} and $PM_{2.5}$ samples at five different locations (roadside, urban, rural, and coastal) around Ireland. The concentrations of $PM_{2.5}$ at urban sites were found to be around four times the concentrations measured at rural background sites. When long-range transport sources were significant, PM levels were associated with polluted air masses from the U.K. and continental Europe. However, higher levels of PM were mainly dominated by local sources.

During the year 2008, Hellebust et al. (2010) identified the main source groups of ambient $PM_{2.5}$ in Cork City as road transport (19%), domestic solid fuel burning (14%) and oil-fired domestic and industrial boilers (31%). Healy et al. (2010) used an aerosol time-of-flight mass spectrometer (ATOFMS) to measure the size resolved chemical composition of single particles in Cork harbour for three weeks in August 2008. Particle number, elemental carbon, organic carbon, sulfate, and $PM_{2.5}$ were measured semi-continuously at the site and PMF was used to apportion $PM_{2.5}$ mass. Vehicular traffic (23%), marine (14%), long-range transport (13%), various combustion (11%), domestic solid fuel combustion (5%), and shipping traffic sources (1.5%) were found to contribute to the measured $PM_{2.5}$ mass.

In February 2009, an ATOFMS and a high-resolution time-of-flight aerosol mass spectrometer (HR-TOF-AMS) were used to quantify the contribution of certain solid fuel types (namely coal, peat, and wood) to ambient PM measurements at the Port of Cork (Dall'Osto et al., 2013a). Four instances were recorded where locally emitted $PM_{2.5}$ caused the port measurements to exceed those recorded in the city centre. Additionally, solid fuels were found to contribute 28% to the total PM_1 aerosol mass.

In a following study, $PM_{2.5}$ sources were identified using PMF during wintertime at an urban background location in Cork City in 2009 (Dall'Osto et al., 2014). Five sources were identified – regional domestic solid fuel burning (24–27%), local urban domestic solid fuel burning (22–23%), road vehicle emissions (15–20%), secondary aerosols from regional anthropogenic sources (9–13%), and secondary aged/processed aerosols related to urban anthropogenic sources (21–26%). Despite restrictions on the use of smoky fuels in place at the time, solid fuel burning was identified as the major source of ambient $PM_{2.5}$ (46–50%) in wintertime in Cork. Lin et al. (2017) investigated the primary nonrefractory PM_1 emissions from the burning of typical commercially available solid fuels (peat, coal, wood) in Galway. During periods influenced by marine air masses, local sources dominated. Peat burning found to be a significant contributor (39%), followed by oil (21%), coal (17%), and wood (11%).

Lin et al. (2018) showed PM_1 levels exceeded 300 $\mu g\ m^{-3}$ in Dublin during the 2016/2017 winter heating season. Frequent exceedances (20% of the measurement period) of the WHO 24-hour mean guideline value (25 $\mu g\ m^{-3}$) were also found. Source apportionment of eBC using a multi-wavelength aethalometer during a particular pollution event showed 95% of eBC was attributed to biomass burning/solid fuel.

Non-refractory PM_1 was measured and characterised in Galway during June 2016 (Lin et al., 2019a). The major component was organic aerosol (OA) (54%), and secondary organic and inorganic aerosol together accounted for 91% of the total non-refractory PM_1, indicating the summer aerosol load was mainly comprised of aged aerosol originating from secondary formation during long-range/regional transport. In the 2016 winter, a contrasting result was found with PM_1 characterisation at four sites across Ireland (Dublin City, Birr, Carnsore Point, Mace Head) (Lin et al., 2019b). Dublin City was the most polluted ($\bar{x}_{PM1}$ = 8.6 $\mu g\ m^{-3}$, with a range of

0.5−146.8 μg m^{-3}) and it was mainly carbonaceous aerosol. Birr had a mean PM_1 of 4.8 μg m^{-3} and a range of 0.5–63.0 μg m^{-3}. Despite having less than 1% the population of Dublin, the mean concentration was around half that of Dublin (56%). Local emissions from residential heating were the dominant contributors (65–74% of the OA) at the Dublin and Birr sites. Carnsore Point and Mace Head (regional background sites), both showed lower mean PM_1 concentrations (0.7−2.2 μg m^{-3}). Carnsore Point PM_1 was dominated by secondary aerosol, accounting for 71% of the OA. Mace Head showed a similar chemical composition, but the PM_1 concentration was more than three times lower due to geographical differences (further from mainland Europe and greater effects of the Atlantic Ocean).

Lin et al. (2020) investigated the chemical composition of PM_1 and sources of OA at a roadside site in Dublin City Centre, and a residential site in the Dublin suburbs during Autumn 2018. The measurement period consisted of a 1-week non-heating period (10−17 September 2018) and a 1-week heating period (27 October−4 November 2018). During both periods, BC was the dominant component of PM_1 at the kerbside (38–55%), while OA dominated at the residential location (46–64%). The BC measured at the kerbside during the non-heating period followed a daily and weekly cycle consistent with vehicular emissions patterns. However, at the residential site, traffic emissions were found to have a minor impact on the PM_1 composition, due to its distance from traffic sources and wind distribution effects.

Wenger et al. (2020) measured $PM_{2.5}$ in three towns (Killarney, Enniscorthy, and Birr) during the winters between 2014 and 2016. These three towns had no natural gas supply, were outside of the solid fuel regulations in place at the time, and so residents were likely to use solid fuels (wood, peat, and coal) for domestic space heating. $PM_{2.5}$ was found to peak at all locations in the evening time and episodes of high $PM_{2.5}$ levels

were found when wind speeds were low. Killarney and Enniscorthy exceeded the 2005 WHO 24-hour mean guideline value (25 μg m^{-3}) 11% and 42% of total measurement campaign days, respectively. Residential solid fuel burning was found to be the dominant source for $PM_{2.5}$ in all locations with contributions of 72%, 82%, 60% for Killarney, Enniscorthy, Birr, respectively. The sources of combustion particles were differentiated, and peat was found to be the main contributor of the $PM_{2.5}$ mass, corresponding to contributions of 31%, 27%, 26% in Killarney, Enniscorthy, and Birr, respectively. Wood was the second largest combustion source (17%, 21%, 23% in Killarney, Enniscorthy, and Birr, respectively) followed by coal (16%, 17%, 5% in Killarney, Enniscorthy, and Birr, respectively).

1.3.3.1 Knowledge gaps

Overall, PM studies in Ireland show a consistent contribution of local emission sources to the wintertime PM measurements with increased $PM_{2.5}$ levels shown in the evening time in both urban and rural areas. Moreover, the reduced contribution of local emissions and overall reduction in $PM_{2.5}$ measurements during summer months highlights the seasonal nature of $PM_{2.5}$ pollution in Ireland and the influence of solid fuel burning. While these studies clearly show that extensive research that has been conducted in Ireland in relation to the sources of PM, to date there have been no high spatial resolution air quality assessments in a single town area in Ireland. Previous air quality assessments typically monitored aerosols at a limited number of locations due to the resources required by such characterisation and source apportionment campaigns. Additionally, studies focussing on the high frequency changes (on the order of minutes) in $PM_{2.5}$ concentrations are limited.

1.4 The Changing Landscape of Air Quality Assessments

In recent times, a variety of sensing technologies have emerged allowing for changes in the monitoring and evaluation of air quality. The advent of these new systems has enabled broader audiences to participate in air quality discussions and made the general public more aware of air pollution (Morawska et al., 2018; Nguyen et al., 2021; Snyder et al., 2013). Although no exact definition for low-cost air quality sensors (AQSs) exists in the literature and the term "low-cost" is subjective and may vary regionally and contextually, the devices are characterised by their relative affordability compared to the purchasing and operating costs of reference grade instruments, their compactness, and the ease of deployment compared to traditional air quality monitoring infrastructure (Karagulian et al., 2019). AQSs can offer a cost-effective solution for air quality monitoring by providing valuable and insightful data with relatively minimal investment to achieve greater spatial coverage. Typically, commercial sensing devices consists of an integrated set of hardware with individual sensors packaged in weatherproof housing and include the necessary electronics for local data storing options, such as with an SD card, or data transmission via Wi-Fi or cellular connection, thus enabling remote data access during deployment (Castell et al., 2017). Commercial devices are typically mains powered or solar powered and often contain additional temperature, relative humidity, and pressure sensors. PM sensors are the focus of this thesis.

1.4.1 Development of Particulate Matter Sensors

Low-cost $PM_{2.5}$ sensors are typically nephelometers or optical particle counters (OPCs), both of which operate on the principle of particle light scattering to determine

particle concentrations. The basic operation of OPCs consists of a fan drawing a sample of ambient air into a chamber. The air passes a laser beam in the chamber and photodetectors detect the scattered light, and the intensity of the light is used to infer the size of the particle. The reflection of light is measured as a pulse by the detector, and the length of the pulse determines the size of the particle, while the number of pulses determines the particle count. The working principle of an OPC is depicted in Figure 1.5. OPCs provide detailed data regarding the particle size distribution and particle counts which is then converted into particle mass concentration (Nguyen et al., 2021). Nephelometers, on the other hand, do not measure individual particles, but instead measure the overall light scattering caused by all particles in the air sample. The intensity of the scattered light is proportional to the particulate mass concentration in the air, assuming a known particle type and distribution. Nephelometers provide scattering coefficient data, a measure of the total light scattering due to particles, and this data is used to derive a PM mass concentration estimate.

The exact mechanism of particle detection employed by commonly used PM sensors, such as those manufactured by Plantower (PMS5003 and PMS6003), continues to be debated. While the sensors report particle size distribution data, as expected from an OPC, studies carrying out laboratory and field evaluations of the response of the PMS5003 to aerosols have reported the sensors behave more like a nephelometer, responding to an ensemble of particles as opposed to counting and sizing induvial particles (He et al., 2020; Ouimette et al., 2022). Contrastingly, Ouimette et al. (2023) found that the PMS5003 does not have the specific characteristics to operate as a nephelometer-type sensor and concluded the sensor instead functions as an “imperfect OPC” that is capable of detecting and counting particles with diameters down to 0.3 µm. While the PMS5003 was found to detect particles individually, limitations of

the sensor were noted. The sensor was found incapable of sizing the particles correctly due to particles of different diameters potentially producing the same photodiode power, depending on the location they intersect the laser beam.

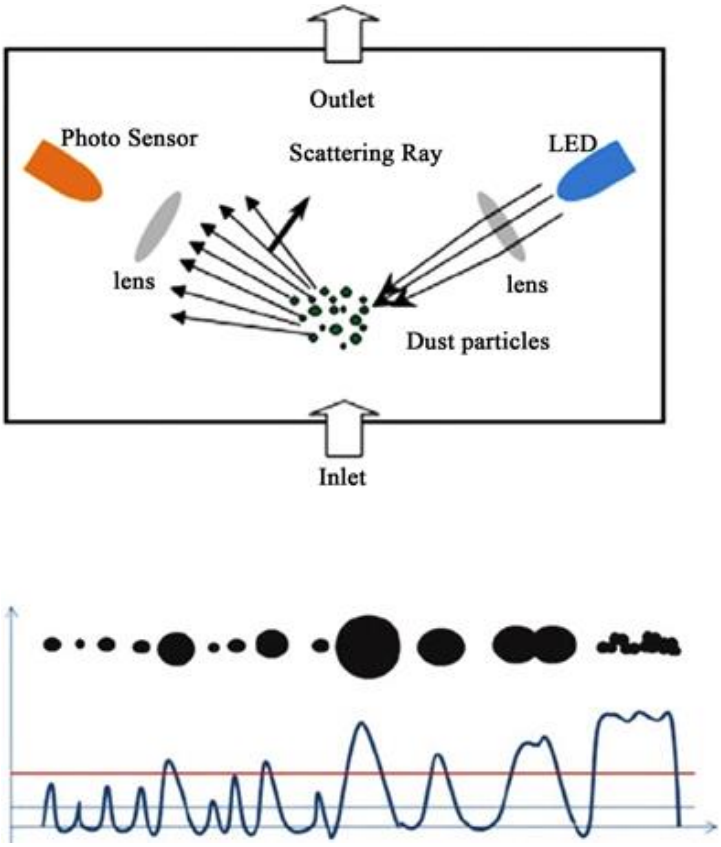


Figure 1.5: *Working principle of OPC sensor for $PM_{2.5}$ measurements, reproduced from Nguyen et al. (2021).*

1.4.2 Challenges Associated with Sensor Measurements

Given the lower cost and decreased sophistication of measurement technique for AQSs compared to the standard air quality monitoring methods, there are certain drawbacks associated with the quality of the output data. Factors that can affect the absolute $PM_{2.5}$ measurements from AQSs include particle composition, size, and hygroscopicity, air temperature, and particularly, ambient relative humidity (Di Antonio et al., 2018; He

et al., 2020; Magi et al., 2020; Tryner, L'Orange, et al., 2020; Tryner, Mehaffy, et al., 2020; Zou et al., 2021). Molina Rueda et al. (2023) found that typical particulate AQS were capable of measuring ultrafine PM (PM_1) with good correlations against reference data ($R^2 \geq 0.83$). However, the performance worsened for larger particles with poorer linear fits for the $PM_{1\text{-}2.5}$ size fraction ($0.65 \leq R^2 \leq 0.76$), although the data could still be improved with field calibrations. Wang et al. (2015) found the slope between the sensor and reference data increased with increasing particle size. Hygroscopic growth occurs when ambient humidity increases, causing water vapour to condense on aerosol particles (Cabada et al., 2004). The condensation directly affects the mass concentration measurements by causing the particle size to increase, and also increases the water content in the aerosol phase, enhancing the particles' ability to scatter light which in turn affects the mass readings reported by the optical sensors (Malings, et al., 2020a). Jayaratne et al. (2018) observed significant increases in mass concentrations at relative humidity levels greater than 75%, with episodes of fog causing a significant effect on sensor readings. Crilley et al. (2018) co-located 14 AQSs with reference OPCs and found significant effects on AQS measurements above relative humidity levels of 85%. A following study found that when AQSs reported elevated levels compared to the reference OPCs, it was mainly due to bulk aerosol hygroscopicity under different relative humidity conditions due to aerosol composition (Crilley et al., 2020).

While several studies have evaluated the short-term performance of $PM_{2.5}$ AQSs, there have been fewer studies investigating the long-term drift in AQS measurements (Clements et al., 2017). deSouza et al. (2023) found when comparing pairs of PMS5003 sensors the proportion of diverging paired measurements increased linearly and averaged to approximately 4% after being in operation for three years, with

increased degradation found in hot and humid climate zones. Additionally, the AQS bias, the difference between corrected AQS data and reference data, increased after an operational period of 3.5 years. As calibration models have been shown to be location specific, it is also likely that a calibration drift overtime is influenced by a potential change in the aerosol profile of a location, due to seasonal or behavioural changes (Malings, et al., 2020a).

As AQSs tend to measure particle mass concentration indirectly, there can be issues with the particle count to mass concentration conversion due to certain assumptions and generalisations made about particle density and size. Often times, the conversion algorithm is a proprietary algorithm and the method description for converting particle numbers into mass concentration is not disclosed, as is the case with the Plantower sensors. The Plantower sensors have two data correction factor (CF) streams, *CF_1* and *CF_ATM*, which represent the output after the application of different algorithms to convert the particle count data. The algorithms are described as being suitable for "factory" and "atmospheric" use, respectively, and make different assumptions based on the density of the particles. Studies have identified the *CF_1* data as the principal data stream (Holder et al., 2020; Wallace, 2022, 2023b). Wallace, 2023a closely approximated the *CF_1* algorithm, with the best-fitting model for $PM_{2.5}$ found capable of matching the *CF_1* algorithm following the form:

$$PM_{2.5} = a(N1 + N2) + b(N3) + d \quad \textbf{(1.1)}$$

where *N1*, *N2*, and *N3* are particle numbers in the size bins 0.3–0.5 µm, 0.5–1µm, and 1–2.5 µm, respectively, and *a*, *b*, and *d* are variable parameters included to fit the data. This insight into the algorithm showed a lack of flexibility if changes in the relative proportion of particles in the *N1* and *N2* categories changes, due to the use of a single

parameter to multiply the sum of the particle numbers. Another issue noted was that the parameter *d* was negative, thus causing negative values at low PM concentrations, which are subsequently replaced by zero values, despite the particle numbers never reporting as zero. These findings were also corroborated in a following study (Wallace, 2023b). Additionally, Kelly et al. (2017) observed that the *CF_1* data stream showed a non-linear comparison to reference data above 40 μg m^{-3} and the *CF_1* and *CF_ATM* measurements diverged at this point, indicating the *CF_ ATM* values have a correction factor applied for non-linear behaviour. Additional correction factors for the PMS5003 data have been proposed and incorporated into the PurpleAir (PA) map, which hosts real-time PM measurements as the PA devices use Plantower PMS5003 sensors. These correction factors aim to provide more accurate measurements compared to the *CF_1* and *CF_ATM* data streams. Barkjohn et al. (2021) developed a US-wide correction factor that incorporated wildfire smoke and dust storm events. Castell et al. (2017) found higher agreement between AQS and reference measurements at low traffic sites compared to high traffic sites and concluded that a location specific particle number to mass concentration conversion factor may improve the sensor agreement.

1.4.3 Calibrations for Sensor Data

If AQS data is to be used as an indication of the absolute concentration values, it is generally recommended that sensors be calibrated across a full range of meteorological conditions and expected pollutant concentrations. Studies show that the performance of AQSs can be specific to each brand of sensor, but also specific to the measurement location and season (Hagan & Kroll, 2020; Kaur & Kelly, 2023b; Sayahi et al., 2019; Sousan et al., 2016; Wang et al., 2015; Zamora et al., 2020, 2023; Zusman et al., 2020). Castell et al. (2017) found that laboratory calibrations were not sufficient to reflect real

world conditions when applied to field measurements. Thus, in situ (or field) calibration methods are needed to improve the accuracy of the reported AQS $PM_{2.5}$ measurements. Field calibrations involve co-locating a sensor with a reference (or reference-equivalent) monitor and using statistical calibration methods to adjust the AQS $PM_{2.5}$ measurements, producing "true" concentration values that more accurately reflect the real-world $PM_{2.5}$ concentrations represented by the reference data. Additional parameters can also be included to improve the accuracy of the calibration model, such as relative humidity and temperature.

A common form of calibration is a multiple linear regression (MLR) model, as shown in Equation 1.2, that describes the dependant variable ($\hat{Y}$), as a linear combination of input variables ($X_1, X_2, \ldots X_n$) and a constant term (c_0) (further information is provided in Section 2.3.1). In field calibrations, the reference data is the dependent variable, AQS data is included as an input variable, and the model outputs "predicted" y-values representing the calibrated sensor data.

$$\hat{Y} = c_0 + c_1X_1 + c_2X_2 + \ldots c_nX_n \quad \textbf{(1.2)}$$

MLR models optimise the best fit for the predicted values based on minimising the distance between them and the dependant variable. The models are relatively easy to form and apply to data, not requiring high levels of computational resources or efficiency. In addition, the interpretable model coefficients ($c_1, c_2, \ldots c_n$) allow for understanding the influence of the input variables.

Numerous studies have shown MLR models produce good $PM_{2.5}$ data calibrations. Malings et al. (2020b) found that the correction of PA data showed little improvement of the correlation with reference data after calibration. However, both the bias and error with the PA data were improved. Datta et al. (2020) found that raw AQS $PM_{2.5}$

data tended to overestimate co-located reference data by as much as 25 μg m^{-3} ($\Delta_{mean} \approx 5$ μg m^{-3}), however after calibration the data was more in line with the reference measurements ($\Delta_{mean} < 2$ μg m^{-3}). McFarlane et al. (2021) developed a MLR calibration model to reduce the mean absolute error (MAE) between PA measurements and reference data from 14.8 μg m^{-3} to 3.4 μg m^{-3} and improve the correlation (R^2 = 0.96). Barkjohn et al. (2021) assessed a U.S-wide PA sensor network and after developing a MLR correction model, reduced the 24-hour averaged $PM_{2.5}$ sensor data root mean square error (RMSE) from 8 μg m^{-3} to 3 μg m^{-3} when compared to reference data at various locations. Raheja et al. (2023) tested various calibration procedures and improved both the correlation and RMSE for PurpleAir SD-II and Clarity Node-S devices with a MLR model including relative humidity and temperature parameters. The calibrated R^2 and RMSE for the PA data were 0.85 and 2.88 μg m^{-3}, respectively, and 0.73 and 3.77 μg m^{-3} for the Clarity data, respectively.

1.4.4 Moving Beyond the Calibrations

Despite the biases associated with AQS measurements, numerous studies have found high correlations with reference data and low inter-sensor variations, indicating AQSs can consistently track trends in $PM_{2.5}$ concentrations (Ardon-Dryer et al., 2020; Zikova et al., 2017). Gao et al. (2015) found high correlations (R^2 = 0.97–0.98) in pairwise comparisons of raw one-minute AQS data, and the hourly averaged data showed similar correlations (R^2 = 0.97–1.00). Additionally, the 1 h averaged data showed high correlations with a BAM (R^2 = 0.85–0.90). Pope et al. (2018) found low inter-sensor variation with three co-located sensors, with a coefficient of variance of 8.8 ± 2.0%. Sayahi et al. (2019) observed that four Plantower sensors showed good correlations

with reference data ($R^2 > 0.858$) and low inter-sensor variation ($R^2 > 0.970$). Malings et al. (2020b) found high consistency between PMS5003 sensors, with mean inter-sensor disagreement below 2.5 $\mu g\ m^{-3}$ and correlations above 0.9. Prior to data calibration, Crilley et al. (2018) observed good agreement between multiple AQSs and between sensors and reference instruments. Zusman et al. (2020) found high levels of precision between pairs of Plantower sensors (mean $R^2 = 0.98$), and similarly high correlation was found with comparison to regulatory-grade data ($R^2 = 0.96$). Masic et al. (2020) tested eight AQSs during periods of strong and mild urban pollution. Very little inter-sensor variation was found, with R^2 greater than 0.99 between any two sensors. Good correlations ($R^2 > 0.9$) with the reference data were also observed during periods of high urban pollution, however the sensor data tended to over represent ambient $PM_{2.5}$ levels by approximately 30% during periods of mild pollution. Kelly et al. (2017) showed that Plantower sensors followed the $PM_{2.5}$ concentration trends of reference-equivalent monitors during the build-up and dissipation of several high PM episodes. However, the sensors tended to overestimate concentrations when they exceeded 10 $\mu g\ m^{-3}$. McFarlane et al. (2021) found a high correlation between raw PA and reference data ($R^2 = 0.88$). Raheja et al. (2023) conducted a large-scale AQS intercomparison study in Ghana, including two PurpleAir SD-II and 16 Clarity Node-S devices. The individual measurement channels in the PA devices were found to report similar measurement ranges, and both the PA and Clarity devices overestimated the reference measurements, with the latter being biased high to a greater extent. The raw PA measurements were well correlated with a co-located reference monitor (Teledyne T640) ($R^2 = 0.82$, RMSE = 17.24 $\mu g\ m^{-3}$), while the raw Clarity measurements showed a moderate correlation ($R^2 = 0.69$, RMSE = 17.08 $\mu g\ m^{-3}$).

1.4.4.1 Sensors for Air Quality Assessments

Deployment of AQSs in networks is emerging as a very useful approach for obtaining information on the spatiotemporal patterns of pollutants in their immediate surroundings (Awokola et al., 2020; Kumar et al., 2015). The low-cost nature of the sensors is particularly important as it facilitates air quality assessments in regions with limited access to regulatory monitoring due to a lack of financial and human resources.

Gao et al. (2015) used a network of AQSs in Xi'an, China, an area with significant $PM_{2.5}$ concentration issues. The AQSs successfully identified the monitoring location with the highest $PM_{2.5}$ concentrations, showing potential as an aid for pollution hotspot monitoring. A network of five field calibrated PA sensors were used to obtain $PM_{2.5}$ measurements in Lomé, Togo over a two-year period (Raheja et al., 2021). Lomé is a coastal city in western Africa and at the time of the study there were no published ambient $PM_{2.5}$ measurement data and Togo did not have national air quality monitoring standards. As a result, the Lomé population (~1.4 million people) had no access to up-to-date $PM_{2.5}$ exposure information despite being in a region of the world particularly known for ambient $PM_{2.5}$ levels exceeding the WHO guidelines, as indicated by satellite data, air quality models, and preliminary monitoring infrastructure. Across a range of sites, including the urban centre, the coastal region, and the city outskirt regions, more than 87% of days exceeded the 2021 WHO daily $PM_{2.5}$ guideline ($\bar{x}_{PM2.5}$ = 23.2 μg m^{-3}). Overall, the AQS network followed seasonal trends and patterns experienced in the region during the Harmattan season (a period between November and March during which regional wind from the Sahara desert brings dry and dusty ambient conditions over West Africa) while the diurnal and weekly patterns were shown to be largely influenced by local influences. In Cape Coast, Ghana, Gameli Hodoli et al. (2020) used two AQSs to assess the use of factory calibrated AQSs for

air quality monitoring when local regulatory monitoring infrastructure is unavailable, i.e. negating the need for a site-specific data calibration. The sensors were shown to be suitable and effective for the application, providing high-temporal resolution data in a previously unmonitored region. Raheja et al. (2023) found with a network of calibrated Clarity devices that the four-month mean daily $PM_{2.5}$ in the city of Accra, Ghana, was 23.4 $\mu g\ m^{-3}$. $PM_{2.5}$ levels were found to be 1.2 times higher in the mornings during Harmattan periods compared to non-Harmattan levels. Multiple additional studies have also reported significantly high levels of $PM_{2.5}$ using AQSs in underrepresented areas (McFarlane et al., 2021; Pope et al., 2018; Singh et al., 2021).

As well as providing air quality information to locations which have little or no monitoring infrastructure, AQSs can also supplement and improve existing air quality infrastructure due to the increased spatial and temporal resolution which can be achieved. Multiple studies have also utilised AQSs for the assessment of wildfire related $PM_{2.5}$ and environmental health surveillance purposes (Barkjohn et al., 2022; Connerton et al., 2023; Holder et al., 2020; Kelleher et al., 2018). Additionally, large-scale sensor deployments are becoming a viable option for local authorities to enhance their air quality monitoring capabilities, often times involving the public for citizen science initiatives. Public health groups are also utilising sensor networks to enhance exposure studies. The Born in Bradford research programme initiated a project to evaluate the effectiveness of the clean air zone in Bradford, U.K., with a focus on reducing children's exposure to pollution by conducting AQS measurements (Bradford Royal Infirmary, 2024). It also aims to assess the impact of the clean air zone on birth outcomes and cardiovascular health in both adults and children by linking hospital and GP attendances with pollution exposure levels before and after the introduction of the zone. The Breathe London project, run by the Environmental

Research Group at Imperial College London, was initiated in 2019. The extensive sensor network uses over 100 Clarity sensors distributed across the city of London, including in schools and hospitals, and incorporates reference co-location sites for data quality assessments (Environmental Research Group & Imperial College London, 2024). Cork City Council set up a PA sensor network of 15 sensors across Cork City in 2019, with calibrated data hosted by an online dashboard (Cork City Council & University College Cork, 2022). The Dún Laoghaire-Rathdown County Council has deployed over 10 AQSs throughout the Dún Laoghaire-Rathdown area in Dublin (Dún Laoghaire-Rathdown County Council, 2024). More and more local authorities, community groups, and public health bodies are recognising the need for high spatial air quality measurements, particularly to inform the public of the local air pollution conditions.

1.4.4.2 Additional Applications of $PM_{2.5}$ Sensors

In addition to providing enhanced spatiotemporal information on the absolute PM levels across a region, there is significant potential for the data produced from AQS networks to be used for more in-depth PM analysis, although research into these applications is limited compared to calibration and basic data analysis applications (Carotenuto et al., 2023).

Malings et al. (2020a) explored the integration of AQSs with satellite observations to estimate ground measurements. After development of the conversion from satellite aerosol optical depth (AOD) to surface $PM_{2.5}$ concentrations where reference-grade data and spatially dense AQS data was available in the U.S.A, the method was employed in various locations in sub-Saharan Africa. During method development, the

AQS data was found to provide the same level of accuracy in surface $PM_{2.5}$ estimates as the reference data. However, having multiple AQSs deployed over an area proved beneficial for more robust conversions. After the integration of local AQSs and AOD datasets in Rwanda, a 40% improvement in the accuracy of surface $PM_{2.5}$ estimates was found compared to using only the closest reference monitor.

Frederickson et al. (2024) used AQSs deployed vertically up to 25 m on the exterior of a building to investigate the vertical transport of pollution from the road below. While the $PM_{2.5}$ AQSs exhibited low inter-sensor variability (minimum $R^2 = 0.89$), no significant vertical gradient was detected, supported by simulation models. However, the vertical concentration profile was found to be sensitive to stratification, and under certain atmospheric conditions, concentrations measured with the higher AQSs showed levels similar to those on the ground floor. In a more expansive monitoring set-up, Hao et al. (2022), measured the vertical concentration profile up to 623 m in Shanghai. $PM_{2.5}$ generally showed a decreasing trend from near ground level to a height of 105 m, and then gradually increased, reflecting the impact of regional transport and the production of secondary PM. It was concluded that regionally transported pollution in Shanghai is particularly impactful in autumn above 415 m.

There has been an increase in the number of studies investigating the use of AQS systems for pollutant source identification. Source identification and quantification typically involves separating the pollutant concentration measurements into underlying regional background sources and local emission sources (Pey et al., 2010). Typically, the procedure follows the basic principles of the incremental approach, whereby the total pollutant concentration is a combination of regional, urban, and local emissions (Lenschow et al., 2001). Heimann et al. (2015) explored the quantification of regional and local contributions to total pollution concentrations using a dense

network (32 nodes) of high temporal resolution (10 seconds) carbon monoxide (CO) measurements. Through the application of a mathematical baseline extraction method, CO measurements were separated into local emission contributions, local emission build-up, and regional contributions. Using a regression model-based approach to quantify the incremental contributions to ambient $PM_{2.5}$, Lung et al. (2020) assessed the exposure contribution of various community sources in a town in Taiwan. Seasonal components were found to be significant (~5 $\mu g\ m^{-3}$), while at one location various activities associated with a nearby market (a laundry, a steak house) were also found to contribute (~4 $\mu g\ m^{-3}$) to the ambient $PM_{2.5}$ concentrations. Traffic locations with stop-and-go traffic were more impacted (~4 $\mu g\ m^{-3}$) than those sites with traffic passing by (~3 $\mu g\ m^{-3}$). F. D. Pope et al. (2018) used the incremental approach with three AQSs deployed in Kenya – two urban locations in Nairobi (an urban background site and a roadside site) and one rural location upwind of Nairobi. With comparison of the urban background and urban roadside measurements, the urban background and roadside increments for $PM_{2.5}$ (~33 $\mu g\ m^{-3}$, ~43 $\mu g\ m^{-3}$, respectively) were quantified. Both urban sites exhibited distinct diurnal patterns, with peaks in the PM concentrations during morning and evening rush hour times. Vehicular traffic was identified as a dominant source of $PM_{2.5}$ in the city (~48%).

Hagan et al. (2019) used non-negative matrix factorisation (NMF) to decompose the matrix of several one-minute AQS measurements of both gases (SO_2, NO_2, O_3, CO) and PM, and environmental variables (relative humidity and temperature) into smaller source and contribution matrices in order to characterise the particulate pollution sources. The data revealed a factor related to combustion-related activities, due to the high CO signal and additional contributions from SO_2, NO_2, and small particles, with a distinct diurnal pattern. The factor information from the AQS data was then used in

conjunction with chemical composition data obtained with research-grade instruments, which aided in the identification of a secondary inorganic aerosol factor likely to have strong regional influences. Similarly, Yang et al. (2022) used NMF applied to normalised AQS data, co-located with research-grade instruments, to resolve pollution sources over the 2020/2021 winter in Atlanta, Georgia, U.S.A. The resolved factors identified a factor correlating with primary organic aerosol during morning and evening rush hours, indicating it is a combustion related factor. Factors relating to daytime and nighttime secondary organic aerosol were also identified.

Bousiotis et al. (2021) investigated the use of an AQS, specifically an OPC, for source identification of particles and identifying the conditions that affect particle concentrations. Using cluster analysis, six particle profiles (clusters) were identified with both the OPC and a co-located scanning mobility particle sizer (SMPS). The OPC was found to agree well the SMPS in terms of source identification for observed particles. While the OPC differentiated sources of particles with sizes greater than 1 μm well, it was limited for the differentiation relating to smaller particles while the SMPS is more capable at the smaller size ranges. While the OPC was not as precise as the SMPS when separating more complex pollution sources within an urban environment where the variation of very small particles is crucial for identifying particle and emission sources, the study showcased that the OPC can identify high levels of pollution associated with larger particle sizes, mainly due to aged pollutants, indicating that the ultrafine particle range can have a detectable effect in larger particle distributions. Additionally, it was noted that a dense network of low-cost OPCs can be significantly more beneficial and cost-effective for the study of multiple pollution sources or hot spot identification than the same cost-benefit of employing multiple SMPS instruments. In a following study, Bousiotis et al. (2022) employed PMF

receptor analysis for source apportionment using AQS data (PM, NO, NO_2, and O_3 measurements) and BC measurements. Factors were found relating to local urban pollution with calm conditions, marine sources, and urban background sources. The PMF results from the AQS data agreed well with the results from reference data and showed good consistency with the previous cluster analysis results.

Frederickson et al. (2022) used spectral analysis of AQS concentrations timeseries, employing the Fourier transform (FT) to transform the timeseries into the frequency domain, thereby allowing for the characterisation of different source types (regional, urban, and local) according to their periodicity. By transforming NO_2 and $PM_{2.5}$ measurements, the study found that across the network roughly 54% of NO_2 was from local sources, while local $PM_{2.5}$ accounted for an estimated 40% of the total $PM_{2.5}$.

Gameli Hodoli et al. (2023) utilised cluster analysis and polar pollution plots, representing the mean AQS pollutant concentration measured as a function of wind speed and direction, to identify source types of PM in Ghana. High $PM_{2.5}$ concentrations were associated with low wind speeds, indicative of local sources contributing to the total $PM_{2.5}$. In addition, high north-northwestern winds contributed to elevated PM_1 and $PM_{2.5}$ levels, and northeastern winds contributed to high PM_{10} concentrations, indicating different sources for finer and coarser particles. The identification of unpaved and busy roads to the north and east causing wind-blown resuspended dust and traffic emissions, cooking sources to the north-east, and a lack of identifiable major sources to the south corroborated the findings.

Overall, the growing trend of AQS applications and methodologies for advanced analysis shows the potential for applying AQS networks to both supplement existing regulatory monitoring infrastructure, and to address gaps in underserved areas.

However, additional work is necessary to further the efforts to integrate AQS networks and data applications into air quality monitoring. Continued monitoring with AQSs in additional locations is crucial for understanding their behaviour and reliability in different environments.

1.5 Aims of Research

The aims of this work can be summarised as follows:

- Explore the reliability of $PM_{2.5}$ AQSs by conducting inter-sensor variation analysis and co-location assessments with reference monitors.
- Develop timeseries-based methods for analysis of $PM_{2.5}$ sensor network data and provide spatial representativeness assessments and local source contributions estimates for selected locations in Ireland. The requirement for data calibrations is explored with respect to the different methodologies used.
- Integrate $PM_{2.5}$ and BC data with a spectral analysis-based source contribution estimate method to examine the annual trends in wintertime air pollution at a rural town in Ireland.
- Enhance the understanding of local air pollution dynamics, allowing for targeted interventions and informed policymaking. The adoption of AQSs and appropriate data analysis approaches aligns with the growing need for innovative and accessible tools in air quality research, contributing to a more comprehensive and nuanced understanding of $PM_{2.5}$ dynamics.

1.6 Overview of Thesis

Chapter 1 presents an introduction to atmospheric aerosols, with a focus on $PM_{2.5}$ and BC, along with the relevant sources, chemical composition, health and environmental effects. The historical developments of air quality monitoring and assessment, as well as previous studies relating to the monitoring of $PM_{2.5}$ in Ireland, are discussed. Applications of AQSs ranging from calibrations and basic monitoring to more advanced techniques are also presented.

Chapter 2 provides an overview of the operation of the $PM_{2.5}$ sensors, reference instrumentation, and aethalometer used for the research described in this thesis. Sampling location information is provided and background information regarding the data analysis techniques employed throughout the research is also described.

Chapter 3 provides details of a novel source contribution discrimination technique developed on a network of AQSs in Cork City. Following data validation against co-located reference data, the contribution of local and non-local sources to the overall $PM_{2.5}$ concentrations is estimated across the AQS network. The seasonal and spatial variability of $PM_{2.5}$ across the city is investigated.

Chapter 4 details the development of a new metric, the Concentration Similarity Index, to assess the time-averaged similarity of AQS timeseries. The index is developed on extensively validated data and applied to AQS network data in Dungarvan, Co. Waterford, and Cork City. Cluster analysis is used on the similarity results to assess the spatial variation in each location. Using the Concentration Similarity Index, this chapter also evaluates the spatial representativeness of regulatory monitoring locations within the population centres.

Chapter 5 investigates an additional source contribution discrimination technique, using spectral analysis, providing comparisons to the results from Chapter 3 in Cork City, along with an additional application in Dungarvan.

Chapter 6 provides details of the temporal variation of air quality in Enniscorthy, Co. Waterford over multiple winter periods using meteorological data, reference-grade $PM_{2.5}$ measurements, and eBC measurements. Source apportionment is carried out with eBC data, while integrating the method developed in Chapter 5 for additional source analysis.

Chapter 7 presents a summary of the work presented in the thesis, with integrated conclusions based on the results obtained in the various locations. Additionally, scope for future research and implications for policy making and legislation are discussed.

2 Instrumentation and Methodology

2.1 Monitoring Locations

2.1.1 Sensor Network Locations

This thesis uses data obtained from two AQS networks in Ireland. The first is a network in Cork City, which was established around 2020 and is still in operation. The second is a network in Dungarvan, which was set up for a limited period as part of a research project, LIFE Emerald, co-ordinated by the EPA (EPA, 2022).

2.1.1.1 Cork City Sensor Network

Cork City, the second largest city in the Republic of Ireland, is located near the south coast, with an estimated population of 224,000 (CSO, 2022). The city centre is situated between two channels of the River Lee, which converge at the east of the city and flow past the Port of Cork towards Cork Harbour. In 2019, the city's boundary was expanded to include Cork Airport, Douglas, Ballincollig, Blarney, and surrounding areas, increasing the population by approximately 85,000, and the land area from 39 km^2 to 187 km^2 (Local Government Act 2019, 2019).

Solid fuel sales have been restricted in Cork City since 1995. The city now has six regulatory monitoring locations – UCC, Munster Technological University (MTU), South Link Road, Heatherton Park, Glanmire Road, and Cork Airport. The latter two stations were established in 2022 and 2023, respectively. Prior to 2022, $PM_{2.5}$ was measured at UCC and Heatherton Park; with the new additions, $PM_{2.5}$ is now monitored at four sites. An EPA report in 2022 found that neither UCC nor Heatherton

Park exceeded the EU annual mean limit for $PM_{2.5}$ (EPA, 2023). However, both sites exceeded the WHO 24-hour AQG on multiple occasions – 19 days and 17 days for UCC and Heatherton Park, respectively (EPA, 2023).

Cork City Council operates a network of PA sensors (PurpleAir, Inc.) distributed across the city, including areas added in the 2019 boundary expansion. This AQS network is the first local authority-managed air quality monitoring system incorporating AQSs in Ireland, and one of the first citywide PA networks in Europe. This thesis uses $PM_{2.5}$ data collected from these sensors for two periods – 01 January 2021 to 31 December 2021, and 01 January 2023 to 31 December 2023. Across the two periods, 16 locations were monitored; however, some units experienced technical difficulties that limited data collection. Details of the specific devices used in each period are given in the relevant chapters to follow, and all of the locations are shown in Figure 2.1. Reference $PM_{2.5}$ data from the UCC air quality monitoring station (BAM-1020, Met One Instruments) was sourced from the EPA. Local meteorological data from Cork Airport, approximately 5.5 km south of the city centre, was provided by Met Éireann.

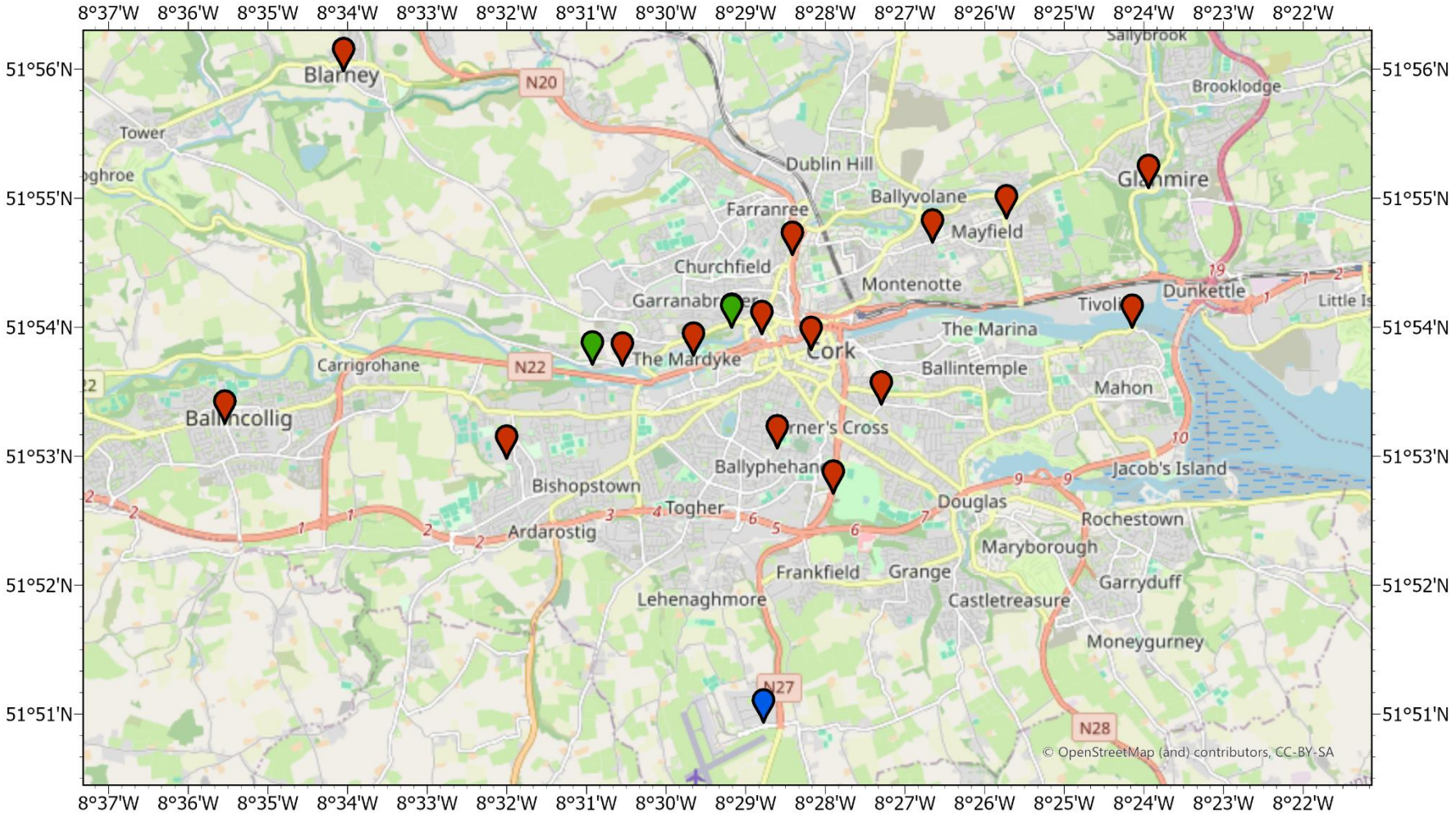


Figure 2.1*: Map of Cork City with sensor network locations (red markers), co-location assessment sites (green markers), and the Met Éireann measurement location at Cork Airport (blue marker).*

2.1.1.2 Dungarvan Sensor Network

Dungarvan is a small town in the south-east of Ireland in Co. Waterford. The 2022 census found Dungarvan had a population of 10,081 (CSO, 2022). The coastal town is situated at the mouth of the Colligan River. The river splits the town into two sides referred to locally as Dungarvan (west) and Abbeyside (east). However, for the purposes of this work the entire town (both east and west) is referred to as Dungarvan. Dungarvan harbour is situated between the two sides of the town, with the 4 km wide Dungarvan Bay extending southward. The solid fuel regulations came into place in Dungarvan in 2022.

There are no regulatory monitoring locations in Dungarvan, however there is an indicative $PM_{2.5}$ monitor (Turnkey, Osiris) located near the library as part of the Local Monitoring Network. This EPA network runs in partnership with local authorities and aims to supplement the national monitoring network with indicative measurements in smaller population areas (EPA, 2015). Dungarvan is included in Zone D of the national monitoring network. A network of 18 Clarity Node-S (Clarity Movement Co., USA) devices was deployed in Dungarvan (Figure 2.2), with measurements spanning from 01 November 2022 to 31 May 2023. Meteorological data was obtained from three Clarity Node-S devices which were fitted with weather station modules. Due to logistical constraints, deployment of these devices was delayed until 21 January 2023. The meteorological parameters measured were wind speed (m s^{-1}), wind direction (°), temperature (°C), relative humidity (%), and air pressure (hPa).

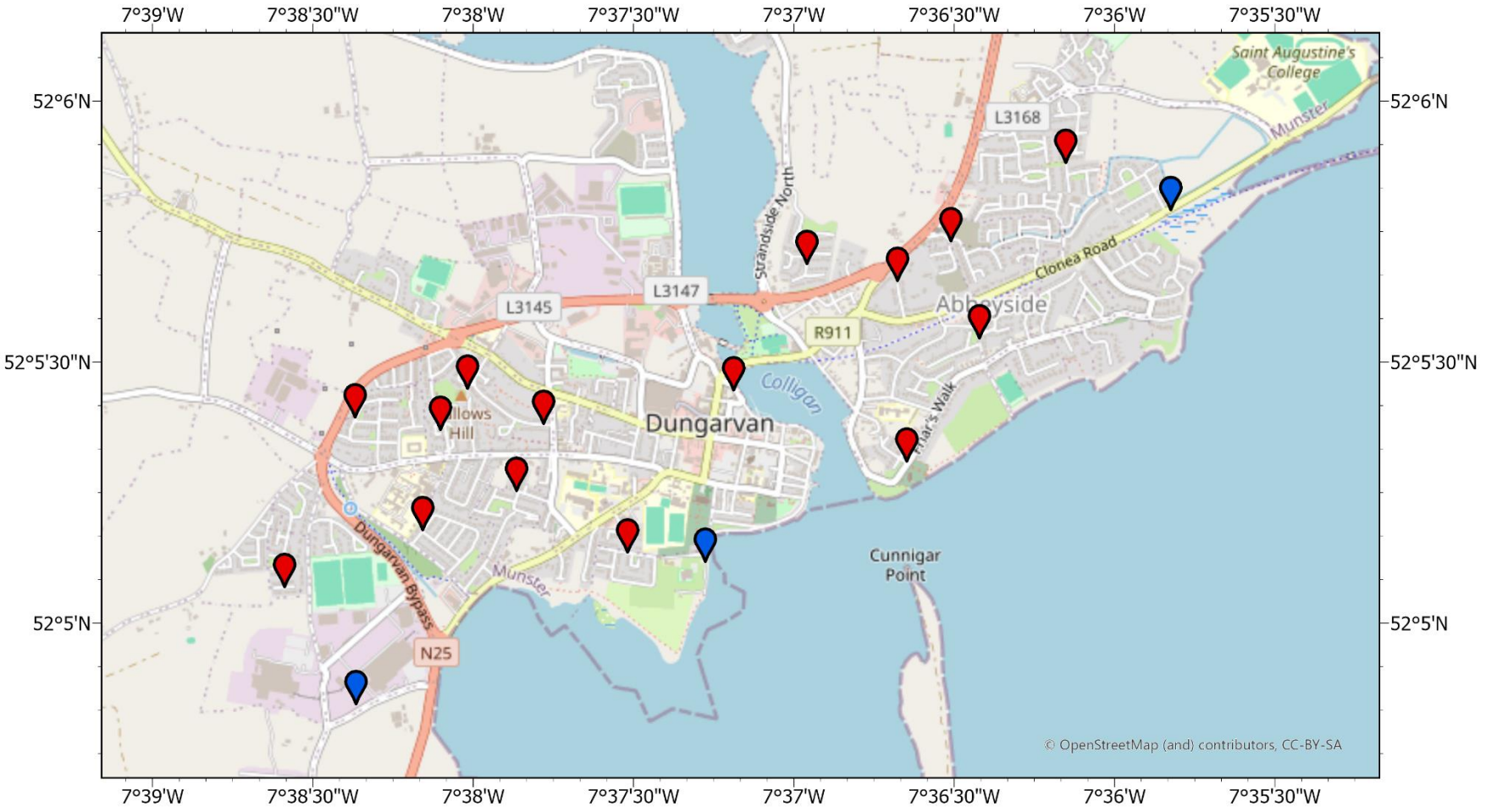


Figure 2.2: *Map of Dungarvan, Co. Waterford with sensor locations that monitored $PM_{2.5}$ only (red markers), and the locations that also monitored meteorological data (blue markers).*

2.1.2 Additional Measurement Locations

2.1.2.1 Co-location Assessments

To support sensor deployment and the proposed methodologies, reference co-location assessments were conducted at two air quality monitoring stations: University College Cork (UCC) (51.900111, -8.486329) and in Edenderry, Co. Offaly (53.342666, -7.047655). The EPA established a regulatory monitoring location situated at the public library in Edenderry in 2021 which includes measurements for $PM_{2.5}$ (Fidas 200, Palas). A sensor network was established as part of the work presented in this thesis in 2021 with the aim of capturing the local pollution dynamics over the 2021/2022 winter period. However, due to logistical constraints, the majority of the winter period was not effectively monitored and therefore the data is not included in the work presented. However, periods of co-location of PA and Clarity devices with the reference $PM_{2.5}$ monitor were carried out and have been used in this thesis. Additional inter-sensor variation analysis was carried out at the Ellen Hutchins Building, Cork City (51.895136, -8.516146).

2.1.2.2 Enniscorthy

Enniscorthy, Co. Wexford, is a small town in the south-east of Ireland, with a population of 12,310 (CSO, 2022). Enniscorthy has a PM reference monitor (Fidas 200, Palas) located at the library. In 2022, the EPA reported that the WHO 24-hour mean AQG was exceeded on 55 days (EPA, 2023).

Concentration measurements for $PM_{2.5}$ and BC were included in this work to provide an updated assessment on the air quality in Enniscorthy, which has previously been found to suffer from significant $PM_{2.5}$ pollution (Wenger et al., 2020). Heffernan

(2022) analysed $PM_{2.5}$ and BC measurements in Enniscorthy between 01 December 2019 and 20 January 2021 in a study of the spatiotemporal variation of ambient carbonaceous aerosol in Ireland. Winter data from 2018 to 2024 was included in order to assess the temporal variation of $PM_{2.5}$ and BC prior to and after the onset of the ban on the marketing, sale, and distribution of bituminous coal, which came into effect in Enniscorthy in late 2020.

$PM_{2.5}$ concentration data was obtained from the EPA, and a multi-wavelength dual spot aethalometer (AE33, Magee Scientific) was used to determine eBC concentrations. Hourly local meteorological data was obtained from Met Éireann at Johnstown Castle, located approximately 23.5 km south-east of the monitoring site in Enniscorthy.

2.2 Instrument Specifications

2.2.1 PurpleAir PA-II SD

The PurpleAir PA-II SD device contains two Plantower PMS5003 laser particle counters to measure ambient $PM_{2.5}$ concentrations (Figure 2.3). The simultaneous measurements provided with PA devices adds redundancy should one sensor fail or give inaccurate readings, and also aids in assessing the health of the sensor by measuring the discrepancy between both sensors. The data from each of the PMS5003 sensors are referred to as channel A and channel B. If the values drift apart, the PA system may downgrade one of the channels, excluding the data from their online map (www.map.purpleair.com). Data from both channels can be retrieved separately or an average can be obtained unless one channel has been downgraded. The PA units also contain a pressure, temperature, and humidity sensor (BME280, Bosch). The units can connect via Wi-Fi and data can be transferred to the online database hosted by PA. Data can be accessed through the PA website or via an Application Programming Interface (API) with user-specific API keys. The sensors take measurements every 10 seconds, which are averaged by the PA software to give two-minute data.

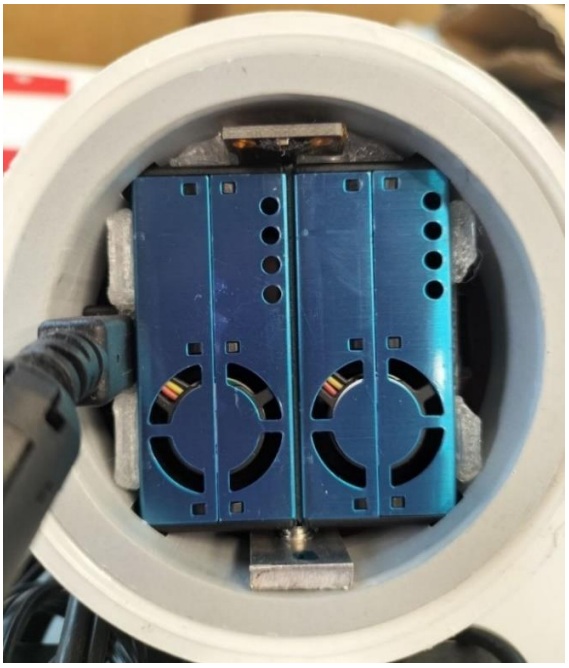


Figure 2.3: *Photo from the bottom of the PA-II SD unit containing two PMS5003 sensors.*

The PMS5003 sensor operates by scattering light at a 90° angle and a photodiode detector that converts the scattered light to a voltage pulse. The photodiode sits below a hole in the housing unit to allow scattered light from the aerosol to impact the photodiode (Ouimette et al., 2022). The light source used in the sensor is a laser with a wavelength range of 680 ± 10 nm (Kelly et al., 2017; Sayahi et al., 2019). The back of the measurement chamber contains a light-trap designed to block any unintended scattering of the laser light (Kelly et al., 2017). Figure 2.4 shows the inner portion of the PMS5003 while the laser is active, and a schematic of the PMS5003 operation is shown in Figure 2.5. A fan draws the ambient air sample from the lower level of the sensor through to the upper part, where they come out through the air flow holes and subsequently pass through the path of the laser, causing the beam to scatter.

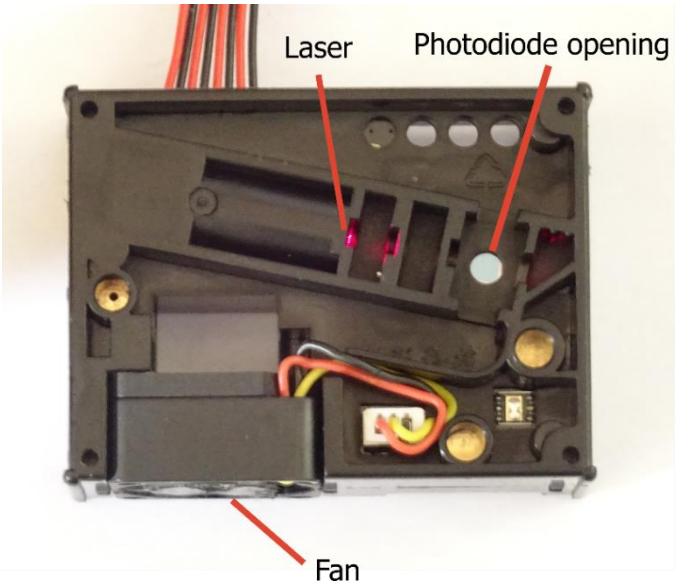


Figure 2.4: *Inner portion of a PMS5003 showing the laser active, the photodiode opening, and the fan.*

Plantower reports an effective detection range of 0–500 µg m^{-3} for the PMS5003 (Plantower, 2022). The uncertainties for the concentration ranges of 0–100 µg m^{-3} and 100–500 µg m^{-3} are ± 10 µg m^{-3} and ± 10%, respectively (Sayahi et al., 2019). The sensors have a maximum range of 1000 µg m^{-3}, and a resolution of 1 µg m^{-3}, and working humidity and temperature ranges of 0–99% and −10–60 °C, respectively (Plantower, 2022). The particle count per decilitre is found by counting the pulses from the scattered light signal and quantified within six overlapping size categories (> 0.3 µm, > 0.5 µm, > 1 µm, > 2.5 µm, > 5 µm and > 10 µm). The conversion from particle count to mass concentration, within various PM fractions (PM_1, $PM_{2.5}$, PM_{10}), is carried out using a proprietary algorithm developed by Plantower (Sayahi et al., 2019; Wallace, 2023a). Plantower provides two mass concentration data series – *CF_1* and *CF_ATM*, which represent "factory" and "atmospheric" conditions, respectively – calculated using proprietary algorithms (Wallace et al., 2021). As previous studies have found the *CF_1* data to be the principal data stream and more suitable to ambient use (Holder et al., 2020; Wallace, 2023a), this data is used in this thesis.

2.2.2 Clarity Node-S

The Clarity Node-S devices measure $PM_{2.5}$ using the Plantower PMS6003 sensor, a newer version of the PMS5003 used in the PA devices. Both sensors offer the same effective range maximum range, resolution, working temperature ranges, and particle counts, while the PMS6003 has a working humidity range of 0–95% (Plantower, 2024). The two sensors operate with the same working principal of light scattering to detect particles in the ambient air. However, the PMS6003 utilises a dual laser system for increased sensor lifespan resulting in a different flow configuration (Figure 2.5) (Plantower, 2024). The "block" feature in the PMS6003 helps avoid possible stagnation and sharp corners (Kaur & Kelly, 2023a). While comparison studies of the two sensor versions are limited, it has been shown that they produce similar responses to particles (Kaur & Kelly, 2023a; Wallace, 2023b).

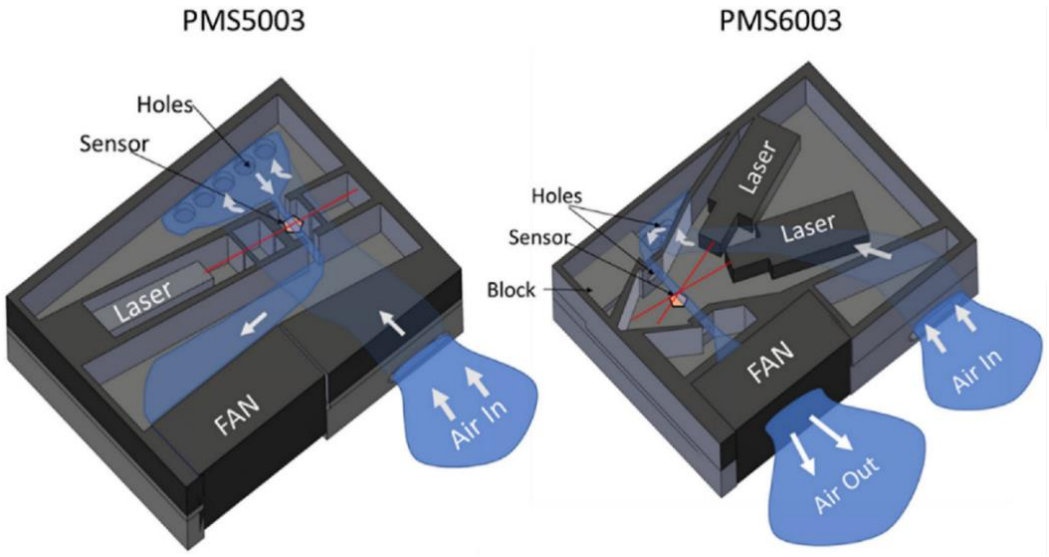


Figure 2.5: *Flow configuration for the PMS5003 and the PMS6003. Adapted from Kaur & Kelly (2023a).*

The Clarity Node-S devices transmit data via a SIM card with cellular connection and are pre-configured to be solar powered, with the battery life lasting approximately 30 days if no solar energy is received (Clarity Movement Co., 2024). Data can be accessed through the Clarity API or on the online dashboard/web app. As standard, the devices operate with a sampling resolution of 15 minutes, which includes a 90 second sampling interval, 900 second sleep time, and additional data upload time. The sleep time can be adjusted remotely with the highest possible data retrieval interval of 2 minutes (1 second sleep time). However, continuous operation at the highest time resolution can cause battery drainage, particularly during the winter months. As a compromise, the sampling resolution was set to 8 minutes to allow for reasonably high temporal resolution while also reducing battery drainage.

The Clarity Node-S devices can also be attached to a wind module (Figure 2.6) containing a solid-state two-axis ultrasonic wind sensor (Clarity Movement Co., 2024). This module, which is also powered by the solar panel attached to the unit, provides wind speed and direction measurements at ranges of 0–60 $m\ s^{-1}$ and 0−359.9°, respectively. The resolution and accuracy for the wind speed measurements are 0.01 $m\ s^{-1}$ and ± 2%, respectively, and 0.1° and ± 3° for wind direction, respectively. For devices connected to a wind module, the sampling resolution was 15 minutes.

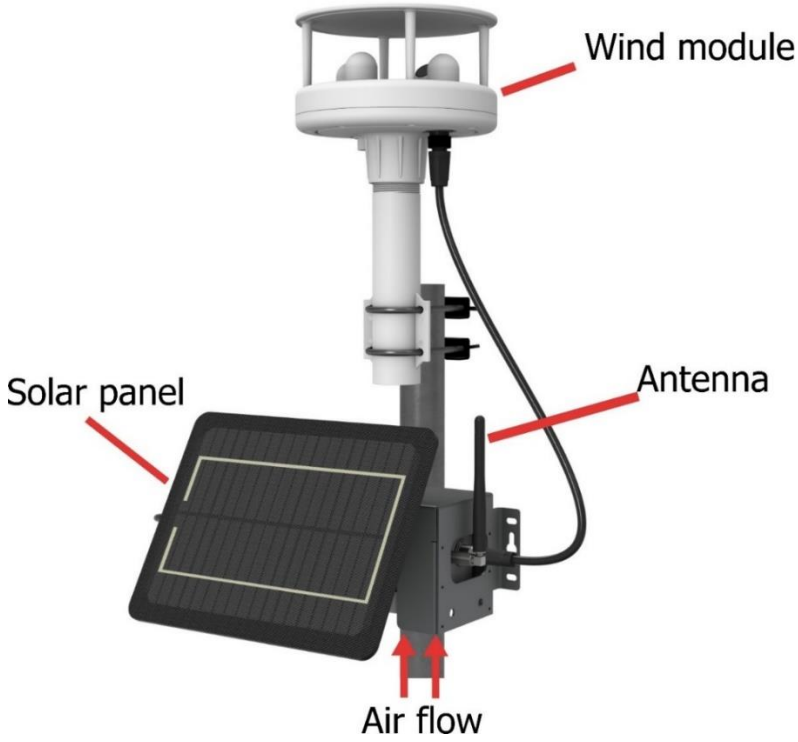


Figure 2.6: *Clarity Node-S device and wind module. Adapted from Clarity Movement Co. (2024).*

2.2.3 Beta-Attenuation Monitor

The BAM-1020, developed by Met One Instruments, automatically measures the mass concentration of ambient PM and is categorised as a reference-equivalent grade monitor within the EU (EN 16450). The instrument operates on the principle of beta radiation attenuation, whereby high-energy electrons, or beta particles, are continuously emitted from a radioactive source (^{14}C) and are detected and counted by a photomultiplier tube scintillation detector within the instrument. When matter is placed between the beta particle source and the detector, the beta particles are absorbed, resulting in a reduction in the number reaching the detector. The magnitude in the reduction is a function of the mass of the absorbing matter (Met One Instruments, 2008).

A schematic of the BAM-1020 is shown in Figure 2.7. A spool of glass fibre filter tape is loaded into the instrument, with an empty spool loaded on the opposite side to collect the used tape. The instrument can be equipped with a PM_{10} inlet and a $PM_{2.5}$ cyclone inlet in order to segregate larger particles out of the sample air stream. BAM units are also typically equipped with a heater module to control the relative humidity in order to reduce any effects associated with particle hygroscopicity (Shukla & Aggarwal, 2022). The BAM-1020 is typically configured to operate on a 1-hour cycle, with a real-time clock controlling the cycle timing. At the beginning of the cycle, the filter tape is advanced with the next clean, unused spot positioned between the beta source and the detector. For exactly 8 minutes, beta particles pass through this clean spot and a baseline attenuation is measured. After this period, the detector stops counting beta particles and moves the tape exactly four windows forward, which positions the same spot used in the previous measurement under the sampling nozzle. The nozzle is automatically lowered onto the filter tape, and an external vacuum pump draws air through the filter tape for exactly 42 minutes at 16.7 L min^{-1}. After the sampling is complete, the tape is moved back four windows, positioning the sample spot between the beta source and the detector. The beta particles are absorbed by particles collected on the filter tape, and the beta particles passing through the filter spot are measured with the detector.

The BAM-1020 operates with a standard measurement range of 0–1000 μg m^{-3} and an effective hourly lower detection limit of 4.0 μg m^{-3}. The ambient temperature range for effective operation is −40–55 °C, and the ambient relative humidity range is 0−90%. Met One Instruments does not quantify an uncertainty associated with the BAM-1020 $PM_{2.5}$ measurements, however an external evaluation showed it is 16% (Hafkenscheid & Vonk, 2014)

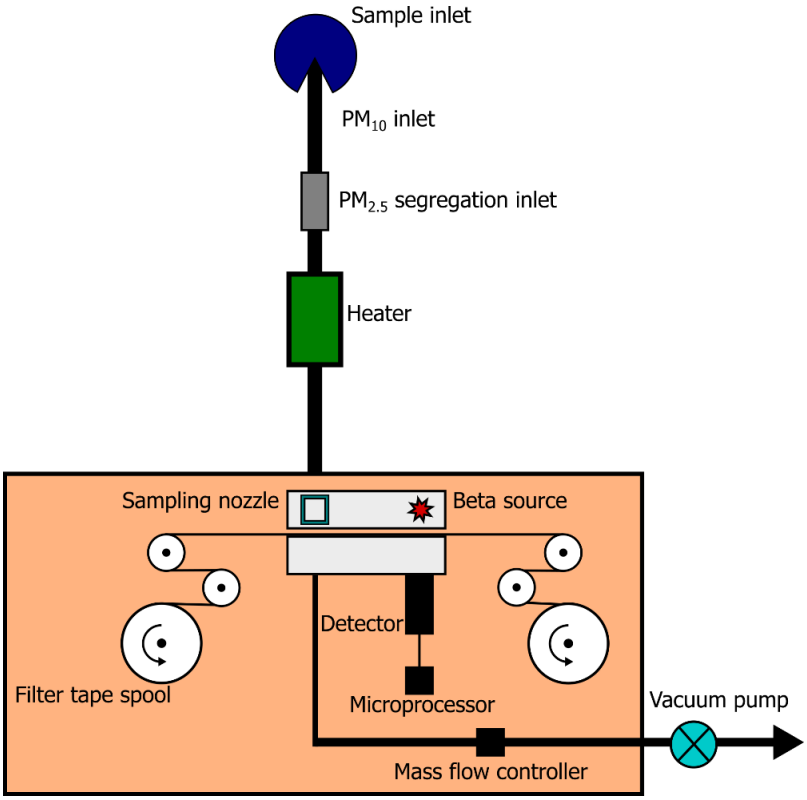


Figure 2.7: *Schematic of the BAM-1020. Filter tape is loaded onto the right spool and wound around the pulley system to travel from left to right through the system.*

2.2.3.1 Mass Concentration Calculation

The attenuation of the beta particle signal leads to a measurable reduction in the beta radiation reaching a detector. The incident attenuation, I_0, and the attenuation of the particle-laden filter tape spot, I, have the following relationship:

$$I = I_0 e^{-\mu x} \tag{2.1}$$

where μ is the absorption cross section of the material absorbing the beta particles ($cm^2\ mg^{-1}$), and x is the mass density of the absorbing matter ($mg\ cm^{-2}$). The absorption cross section is determined during an initial calibration period, and then the predicted mass density is calculated.

When the ambient air is sampled at a constant flow rate, Q (L min^{-1}), for a specified amount of time, Δt (minutes), through a cross-sectional filter surface area, A (cm^2), the ambient mass concentration of PM, C, is calculated with the following equation:

$$C = \frac{A \ln {}^{I_0}/_{I}}{\Delta t \; Q \; \mu} \tag{2.2}$$

where C is measured in μg m^{-3} (Shukla & Aggarwal, 2022).

2.2.4 Fine Dust Monitor

The Palas Fidas 200 is a reference-equivalent certified (EN 16450) instrument for the continuous monitoring of $PM_{2.5}$ within the EU. The principle of operation relies on the measurement technique of optical light scattering on a single particle using a high intensity LED light source, a similar method to OPCs (Palas, 2024). The Fidas 200 employs an optical aerosol spectrometer to determine particle size using Lorenz-Mie light scattering analysis on individual particles. As particles travel through an optically differentiated measurement volume which is uniformly illuminated with polychromatic light, each particle generates a scattered light impulse detected at an angle between 85° and 95°. The particle number is quantified based on the number of scattered light impulses, while the amplitude of each impulse is a measure of the particle size diameter. By incorporating enhanced optics, increasing light density with a white LED light source, the lower detection limit for ambient air has been reduced to 180 nm (Palas, 2024).

The nature of the operation technique allows for a near-real time resolution of PM concentrations (1 minute). The Fidas 200 instrument also contains an Intelligent

Aerosol Drying System (IADS) to dry particles in the incoming sample stream and avoid interference from hygroscopic growth. The Fidas 200 operates with a sample flow of approximately 5 L min^{-1} (Palas, 2024). The measurement range is 0−1500 μg m^{-3}, and the operating temperature is −30–35 °C. The measurement uncertainty associated with $PM_{2.5}$ measurements is 9.7%.

2.2.5 Multi-Wavelength Aethalometer

A dual spot multi-wavelength aethalometer (AE33, Magee Scientific) was deployed in Enniscorthy. The instrument design was originally developed by Hansen et al. (1984), which utilised a single wavelength (550 nm) light source to measure the light attenuation by particles collected on a filter paper. The AE33, however, measures light attenuation across seven wavelengths (Table 2.1) ranging from near-UV (370 nm) to near-IR (950 nm). The transmission of light through the sample filter spot is compared to an unloaded reference filter spot. The instrument also simultaneously compares the light transmission through a second sample spot with different accumulation rates of the ambient sample to mitigate potential errors from non-linearities associated with filter based optical measurements. The non-linear relationship between the attenuation and BC loading at high attenuation values due to filter saturation can cause an underestimation of BC absorption, and consequently, an underestimation of eBC concentration (Cuesta-Mosquera et al., 2021; Drinovec et al., 2015).

The AE33 has two mounts for the measurement tape, an optical chamber, and a bypass filter. The bypass filter removes particles from the ambient air to perform clean air tests. The filter medium is glass fibre-coated polytetrafluoroethylene (PTFE). The type of tape used can influence two parameters, the optical enhancement factor (*C*) and the

tape tangential leakage factor (*Z*), that must be pre-selected when the tape is replaced, and a new sampling cycle begins. Additional parameters can also be adjusted, such as the time base, flow rate, date and time, and the maximum attenuation limit prior to the tape advancing. The AE33 operated continuously on a 60 second time basis with a flow rate of 5 L min^{-1} for all applications in this work.

Figure 2.8 illustrates the operational flow diagram of the AE33. Ambient aerosol enters through an air inlet at the rear of the instrument, passing through a switch valve (*Valve 4*), determining whether the air flows to the filter tape or through the bypass filter for clean air tests. Aerosol flowing towards the filter tape is sampled through three tape spots – *S1*, *S2*, and *Ref* – while light attenuation is measured with the detector with light from a light-emitting diode (LED) source. Different flow rates pass through spots *S1* and *S2* and are regulated by an orifice and flowmeters, with disturbances to the air flow which may impact the optical measurements minimised by a pulse damper. After each sampling cycle, the filter tape advances for the next cycle.

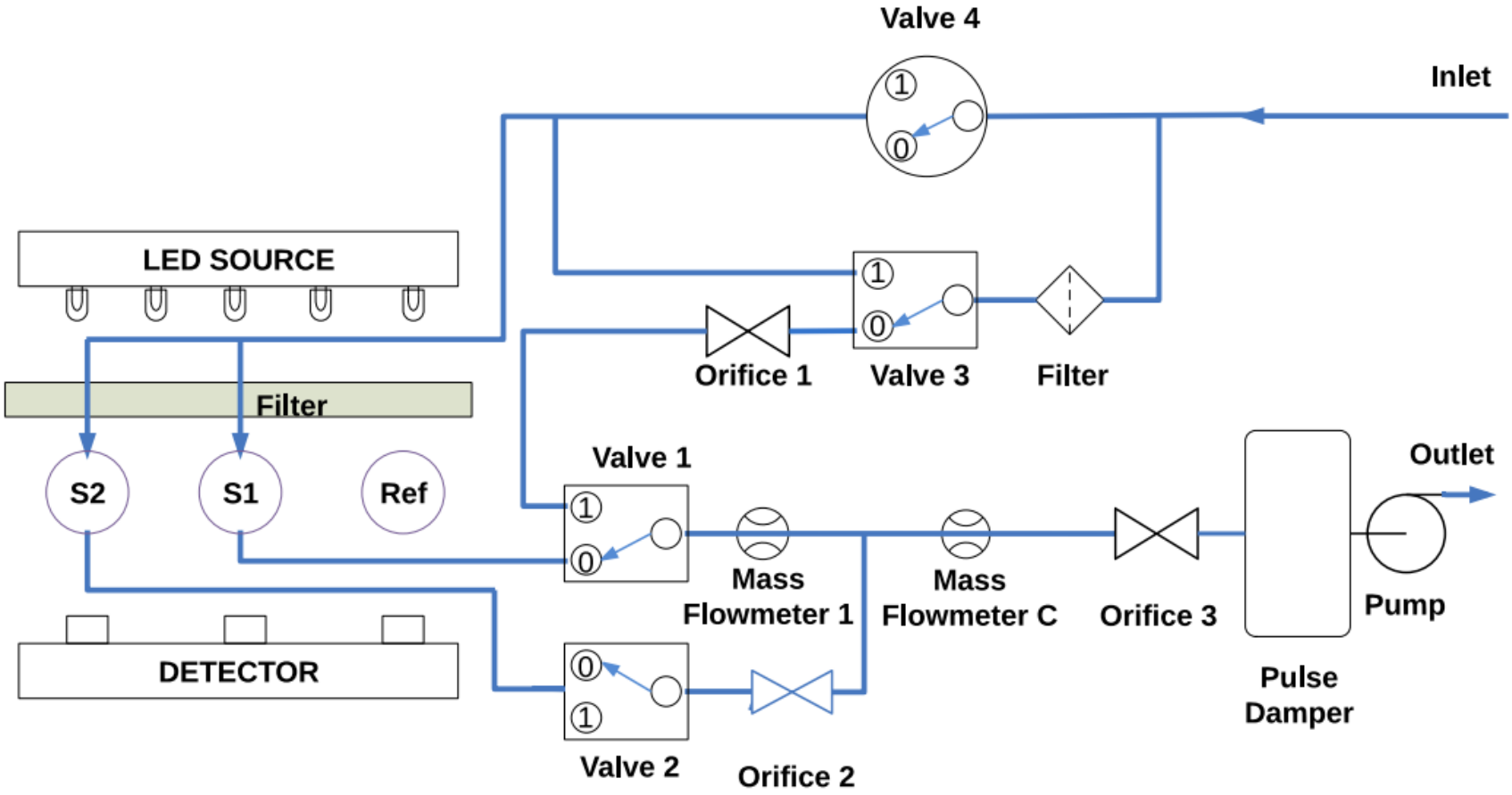


Figure 2.8: *Flow diagram of the AE33 aethalometer (Drinovec et al., 2015).*

2.2.5.1 Mass Concentration Calculation

The operating principles of the AE33 optical chamber are summarised in the schematic diagram in Figure 2.9. The attenuation (*ATN*) of light by the filter is calculated using the intensity of light that passes through the sampling spots and the reference spot using the following equation:

$$ATN \equiv 100(\ln\frac{I_0}{I_x}) \quad \textbf{(2.3)}$$

where I_x corresponds to the light intensity passing through either sample spot 1 (I_1) or sample spot 2 (I_2), and I_0 corresponds to the light intensity passing through the reference spot (Sandradewi et al., 2008a). The change in light attenuation as a function of time, ΔATN, is then used to derive an aerosol attenuation coefficient, b_{ATN}.

$$b_{ATN} \equiv \frac{S(\Delta ATN/100)}{F_{in}\Delta t} \quad \textbf{(2.4)}$$

where S is the tape spot area, ΔATN is calculated over time period, t, and F_{in} is the flow. F_{in} is calculated using the measured flow, F_{out}, and the leakage factor, ζ, as defined by the user:

$$F_{in} = F_{out}(1 - \zeta) \quad \textbf{(2.5)}$$

The attenuation coefficient and the multiple scattering parameter, C, defined by the user, are used to calculate the absorption coefficient, b_{abs}, according to the following equation:

$$b_{abs} = \frac{b_{ATN}}{C} \quad \textbf{(2.6)}$$

The eBC mass concentration is calculated as follows:

$$eBC = \frac{b_{abs}}{\sigma_{air}} \quad \textbf{(2.7)}$$

where σ_{air} is the mass absorption cross section (MAC). The MAC value is set by the manufacturer for each wavelength (Table 2.1) and relates the absorption efficiency of a particle per unit mass ($m^2\ g^{-1}$). Uncoated, freshly emitted BC has a MAC value of $7.5 \pm 1.2\ m^2\ g^{-1}$ at 550 nm (Bond et al., 2006).

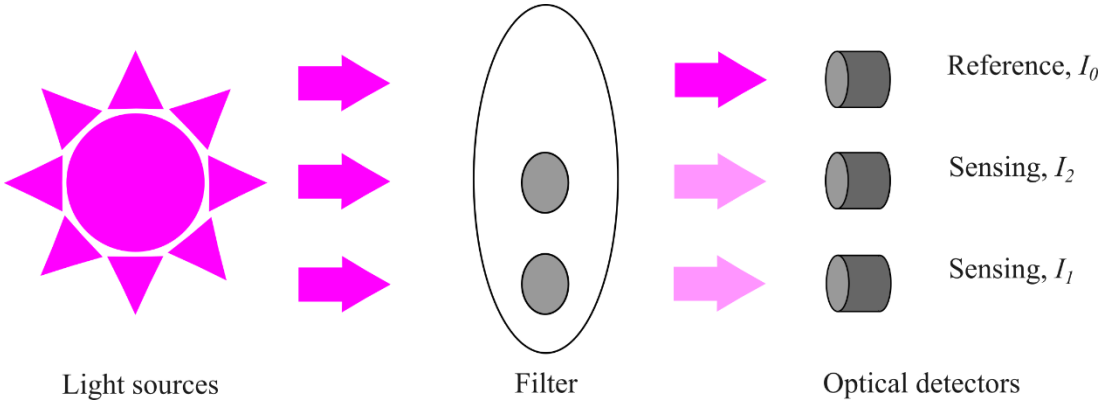


Figure 2.9: *Schematic diagram of the AE33 optical chamber light attenuation. Adapted from Magee Scientific (2016).*

The error associated with the filter loading is accounted for using a real-time compensation algorithm (Drinovec et al., 2015). In summary, the correction method eliminates the loading effect by employing the aethalometer's two simultaneous optical absorption measurements, BC_1 and BC_2, that are conducted with different degrees of particle accumulation, ATN_1 and ATN_2, respectively. The loading compensation parameter, k, is calculated following Equations 2.8 and 2.9:

$$BC_1 = BC(1 - k(ATN_1)) \quad \textbf{(2.8)}$$

$$BC_2 = BC(1 - k(ATN_2)) \quad \textbf{(2.9)}$$

A final eBC concentration can be calculated with the combination of the above equations:

$$eBC = \frac{S(\Delta ATN/100)}{F_1(1 - \zeta)(\sigma_{air})(C)(1 - k(ATN_\lambda)(\Delta t))} \quad \textbf{(2.10)}$$

Table 2.1: *Wavelengths and corresponding mass absorption cross section values for the AE33 aethalometer and related calculations.*

Measurement channel	**Wavelength**	**Mass absorption cross section**
	nm	**$m^2\ g^{-1}$**
1	370	18.47
2	470	14.54
3	520	13.14
4	590	11.58
5	660	10.35
6	880	7.77
7	950	7.19

2.3 Data Analysis Techniques

The analysis of timeseries involves the systematic examination of sequential data points collected over time to identify patterns, trends, and anomalies in air pollutant concentration levels. In terms of AQS network data, appropriate analytical methods are crucial for understanding both the spatial variations and the temporal dynamics of air quality, such as diurnal cycles, seasonal variations, and long-term trends. Additional source apportionment techniques are achievable with eBC data. The data analysis and visualisations presented in this thesis were carried out using the Julia Programming Language (Bezanson et al., 2017) and the R openair package (Carslaw & Ropkins, 2012). Maps were produced using ArcGIS (ESRI, 2011).

2.3.1 Multiple Linear Regression

MLR is a statistical technique used to model the relationship between a dependant variable and multiple independent variables. In comparison to a simple or bivariate linear regression model, which only involves one independent variable, MLR allows for the simultaneous inclusion of multiple predictors, allowing for the analysis of complex datasets where the outcome is influenced by numerous factors. The general form of the model is given in Equation 1.2 (Section 1.4.3). With MLR models, each coefficient c_i represents the expected change in the dependent variable $\hat{Y}$ for every unit change in the corresponding independent variable X_i when all other variables are constant.

For a MLR model to produce valid results, certain conditions must be met. The relationship between the dependant variable and each independent variable is assumed

to be linear and observations are assumed to be independent of each other (the error terms are not correlated) (Tranmer et al., 2020). Additionally, the error terms must be normally distributed, and the independent variables should not be highly correlated with one another, as multicollinearity can make it difficult to assess the effect of each predictor (Tranmer et al., 2020). The coefficient, c_0, c_1,…, c_n in the MLR model are often estimated using the ordinary least squares (OLS) method, whereby the sum of the squared differences between the observed and predicted values are minimised (Witte & Witte, 2017).

2.3.2 Cluster Analysis

Cluster analysis is a valuable unsupervised statistical technique used to identify natural groupings in a dataset by classifying the data into distinct groups, or clusters, without needing pre-classified or labelled data to train the algorithm. It systematically works to separate the data by minimising within group variation and maximizing between group variation. Cluster analysis is often used in air quality analysis, including describing pollution diurnal variation and identifying distinct diurnal patterns, pollution source identification, and identifying spatial patterns in particle compositions (Austin et al., 2012, 2013; Flemming et al., 2005). Hierarchical clustering (HC) and fuzzy c-means (FCM) clustering are utilised in this work. HC offers the hierarchical representation of clusters, aiding in the identification of nested relationships, while FCM allows for a more flexible approach to the cluster assignments.

For cluster analysis, the data is typically structured as an $n \times m$ matrix, with n observations or data points (rows), and m features or variables describing the

observations (columns). Missing values should be handled prior to cluster analysis as most clustering algorithms require complete data.

2.3.2.1 Hierarchical Clustering

HC is a connectivity-based clustering algorithm, it does not separate the data into a defined number of clusters in a single step, but rather consists of a series of separations or merges of the data (Everitt et al., 2011). This type of clustering gives an advantage over partition-based algorithms, whereby the user is not required to specify the number of clusters prior to beginning the algorithm. There are two main types of approaches for HC – agglomerative clustering (bottom-up) and divisive clustering (top-down). Agglomerative clustering treats the individual data points as numerous singleton clusters from the beginning, and the recursive algorithm successively agglomerates pairs of clusters until the merging finalises with a single cluster containing all of the data points. In contrast, divisive clustering begins by treating all data points as a single cluster and recursively splits the clusters until individual data points are contained in singleton clusters. Figure 2.10 illustrates the distinction between both approaches. The agglomerative approach is used in this thesis.

The HC algorithm typically includes the production of a dendrogram illustrating the tree-like structure of the nested clusters, i.e. it depicts the hierarchical relationships between the groups (Figure 2.11). For agglomerative HC, the dendrogram is created through the iterative process of merging clusters, and the process is based on the measure of similarity or distance between data points. Individual observations are situated at the bottom of the diagram, while the largest possible cluster (that containing

all observations) is located at the top. The dendrogram can be sliced at certain heights, representing the linkage distance, to generate certain numbers of clusters.

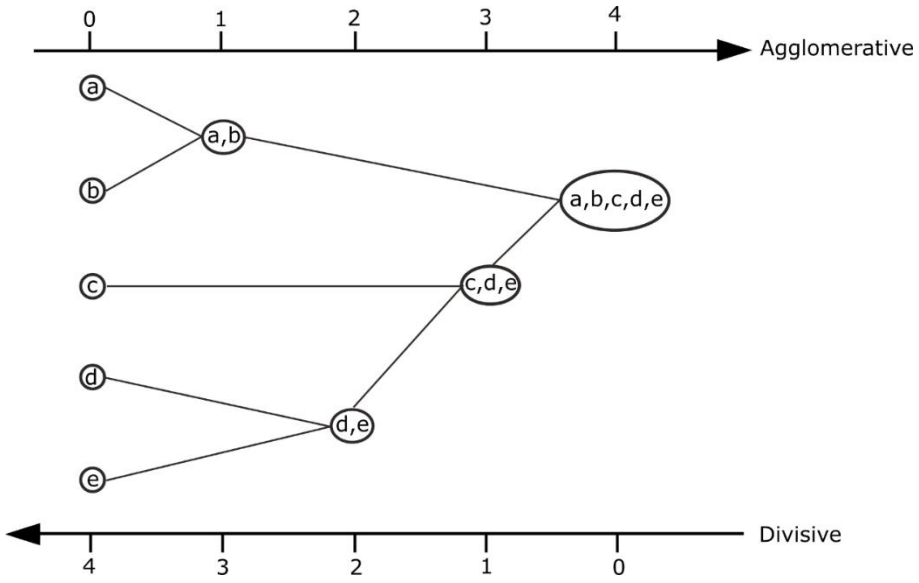


Figure 2.10: *Distinction between agglomerative and divisive HC algorithms. Adapted from Kaufman & Rousseeuw (1990).*

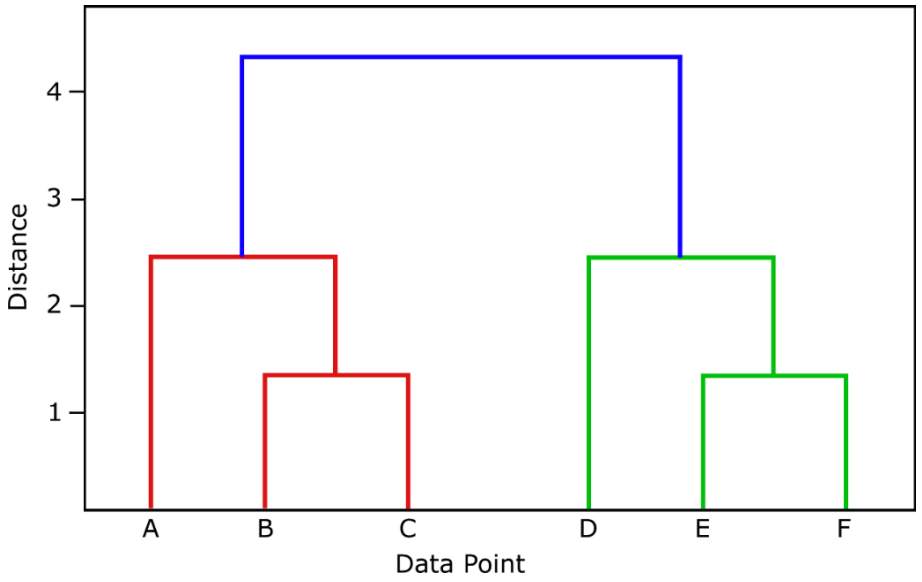


Figure 2.11: *Example of a HC dendrogram where slicing the dendrogram at different distance levels can produce different degrees of clustering. Slicing along the blue lines will give two clusters, cluster 1 (A, B, C) and cluster 2 (D, E, F). Slicing at lower distance levels will produce increasing numbers of clusters, with the maximum being when there is one data point per cluster.*

The merging of two clusters involves computing the distance between each pair and merging the pair with the least distance, i.e. the most similar clusters. A proximity matrix represents the distance between each pair of data points quantified with a distance function. The Euclidean distance is a commonly used distance function representing the straight-line distance between two points in space (Everitt et al., 2011; Murtagh & Contreras, 2012), and is generalised with the following equation:

$$d(i,j) = \sqrt{\sum_{k=1}^{p}(x_{ik} - x_{jk})^2} \tag{2.11}$$

where x_{ik} and x_{jk} are the kth variable value of the p-dimensional observations for points i and j, respectively. In order to merge clusters, a linkage function must be selected that defines the distance between clusters and decides which clusters are merged at each step of the algorithm. The Ward's method, used in this work, is a common linkage function with which the similarity of two clusters is based on the increase in squared error when two clusters are merged (Kaufman & Rousseeuw, 1990).

2.3.2.2 Fuzzy c-Means Clustering

FCM is a partitional, centroid-based clustering technique, where the number of clusters must be predefined. In the traditional hard, or non-fuzzy, clustering such as k-means clustering, each point belongs exclusively to a single cluster based on its distance from the centroid of that cluster (Figure 2.12). However, with soft clustering, the output is a membership score or probability likelihood of a data point belonging to each of the pre-defined clusters (Gentle et al., 1991). A unique mapping of each data

point to a specific cluster does not necessarily exist, as data points may belong to multiple clusters with varying degrees of membership.

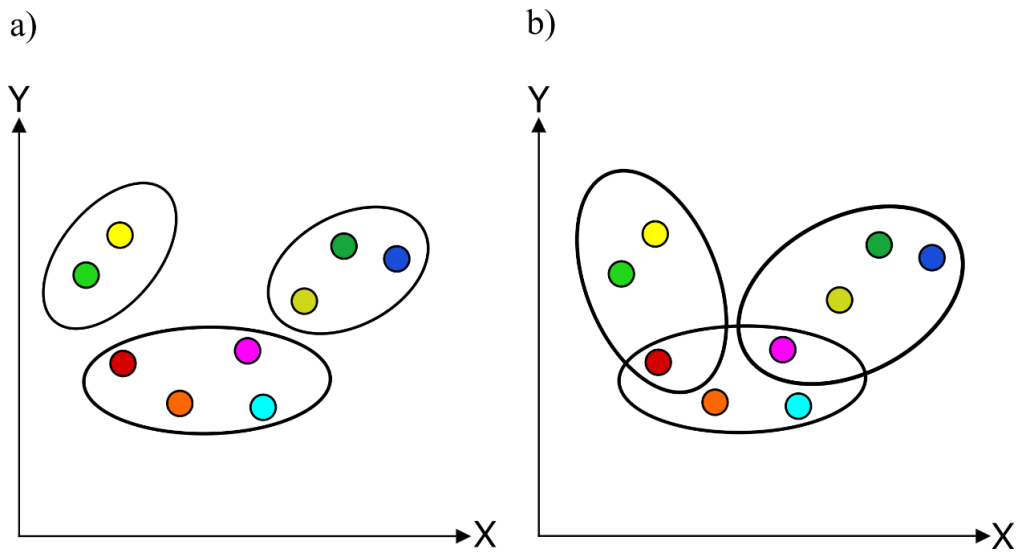


Figure 2.12: *Graphical distinction of cluster assignments between a) non-fuzzy and b) fuzzy clustering approaches.*

The FCM algorithm works iteratively to optimise a cost function that balances the degree of membership of each data point in each cluster with the distance between the data point and the centroid of that cluster and works to minimise intra-cluster variance. Figure 2.13 illustrates the iterative calculation process of the algorithm. To begin, an initial number of clusters, c, is defined and the membership matrix, U, is randomly initialised, where the sum of membership values for each data point across all clusters equals 1. The number of clusters must satisfy the rule: $2 \leq c \leq n$, for an $n \times m$ matrix (Wang et al., 2015). The initial cluster centre matrix, V, is calculated based on the weighted average of all data points, where the membership values act as the weights for each cluster. The membership matrix is then updated with the calculation of the Euclidean distance between every data point and the cluster centres. For a given data

point, the membership value will be closer to one when the data point is closer to the cluster centroid. These steps are repeated through each iteration step, *s*, until the changes in the membership matrix between iterations falls below a predefined threshold or fuzziness parameter, *e* (Julia Programming Language, 2024). This parameter is a weighting exponent that controls the extent of sharing among fuzzy clusters (Bezdek, 1981). The lower the value of *e*, the harder the clusters are, meaning that data points can only belong to one cluster with a high membership degree. If the threshold is not met, a new cluster centre matrix is calculated, resulting in the repositioning of centroids closer to the points with higher membership and the process continues. The final output once *e* is satisfied is a membership matrix reflecting the assignment of every data point to a cluster as a distribution over all available clusters, with each data point having varying degrees of membership toward each cluster.

Although the FCM algorithm tends to be more computationally complex (Ghosh & Dubey, 2023), the information achieved with soft clustering can be particularly useful when dealing with datasets exhibiting overlapping patterns or uncertainties in classification as opposed to initially partitioning into hard clusters with no fuzzy method applied (Gentle et al., 1991).

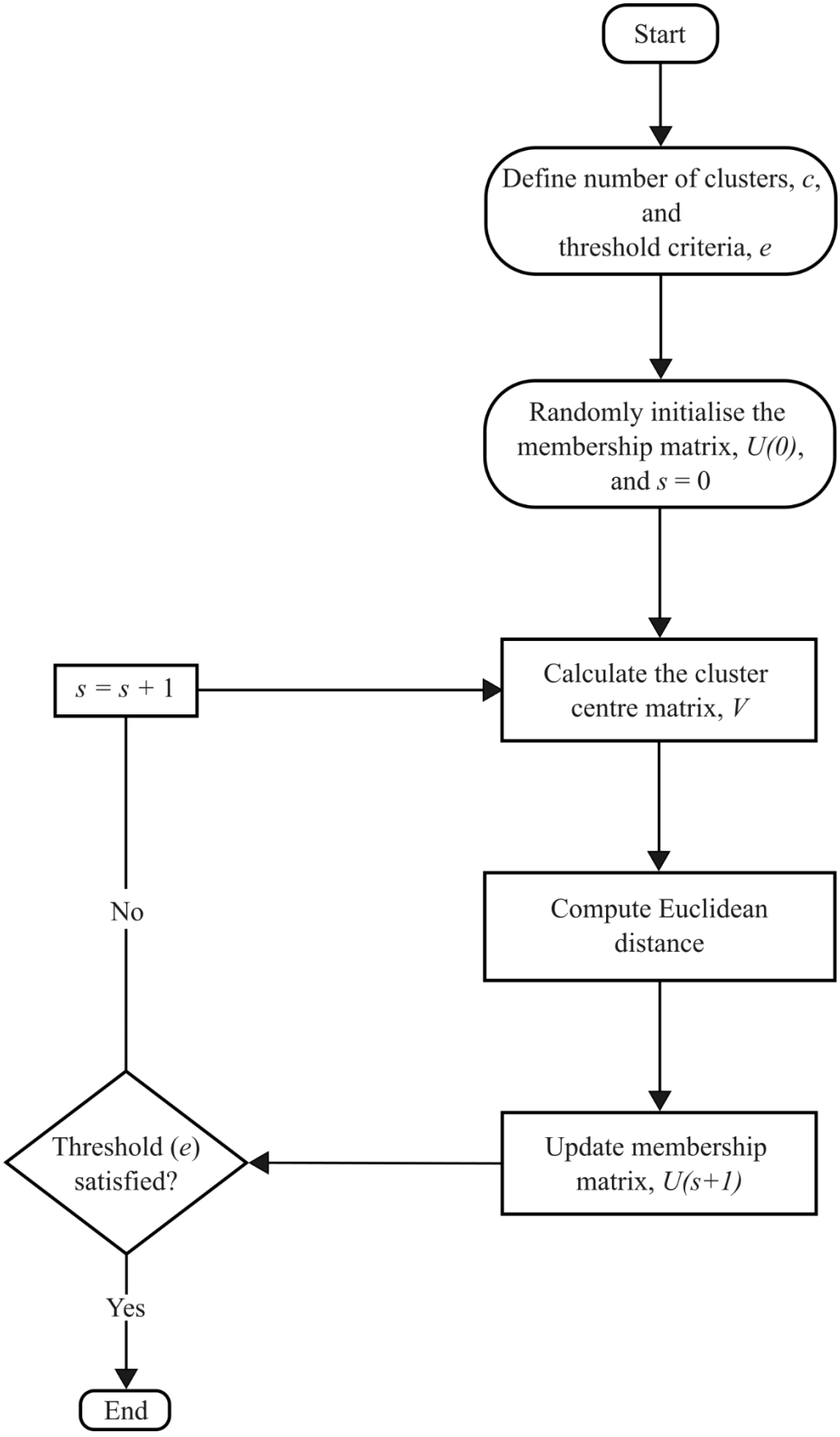


Figure 2.13: *Calculation flowchart of the FCM clustering algorithm. Adapted from Wang et al. (2015).*

2.3.2.3 Cluster Evaluation

With both HC and FCM clustering, the quality of cluster assignments can be assessed with various evaluation metrics to choose the optimal number of clusters. As the "true" cluster classifications are not known here, validation must be performed using the clustering algorithm itself. To assess the quality of the hierarchical clustering assignments, the silhouette metric was used along with the Calinski-Harabasz index to assess the FCM assignments (Caliñski & Harabasz, 1974; Rousseeuw, 1987).

The silhouette score quantifies how similar a data point is to its own cluster compared to other clusters (Rousseeuw, 1987). For each data point, the average distance, *a(i)*, between the data point, *i*, and all other data points in the same cluster represents the cohesion within that cluster. The average distance, *b(i)*, between the data point, *i*, and all data points in the nearest neighbouring cluster (i.e. the closest cluster to which *i* does not belong) represents the separation from the nearest cluster (Arbelaitz et al., 2013). Using these values, the silhouette score of a given data point, *s(i)*, is found with the following calculation:

$$s(i) = \frac{b(i) - a(i)}{maximum(a(i), b(i))} \quad \textbf{(2.12)}$$

The score ranges from −1 to +1, and when calculated for each member of a cluster then the mean silhouette score from all members indicates an overall assignment quality. A mean score closer to one indicates higher quality clusters where the data point is far from neighbouring clusters. A score closer to zero is indicative of a data point being near the decision boundary of two neighbouring clusters, and a negative score indicates poorer cluster assignments where the data point may have been assigned to the wrong cluster.

The Calinski-Harabasz index can be applied to soft clustering techniques and measures the ratio of the sum of between-cluster dispersion to within-cluster dispersion, with a higher index indicating better cluster quality (Caliñski & Harabasz, 1974). For a set of data points, x_1, x_2,…,x_n, spread across clusters, C_1, C_2,…,C_k, a given cluster will have a cluster centre, c_j, and the data will have a global data centre, $\bar{x}$ (the mean of all data points). The Euclidian distance function, d, can be applied to find the distance between cluster centres and the global data centre (between cluster dispersion), and distance between a data point and the cluster centre (within cluster dispersion) (Arbelaitz et al., 2013). The Calinski-Harabasz index, CH, is calculated as follows:

$$CH = \left(\frac{n-k}{k-1}\right)\left(\frac{\sum_{c_j}|C_j|\, d(c_j,\bar{x})}{\sum_{c_j}\sum_{x_i \in C_j} d(x_i c_j)}\right) \quad \textbf{(2.13)}$$

2.3.3 Spectral Analysis

2.3.3.1 Data Transformation

In a continuous timeseries of measured pollutant concentrations, it is commonly observed that the values fluctuate in a repeatable and cyclical manner. The duration of fluctuations is often associated with regular events in the diurnal cycle, e.g. rush hour traffic and smoke emissions from residential heating in the evenings, i.e. there are events that occur with certain frequencies, e.g. once per day or twice per day. Information about the data can be gained by extracting these frequencies from the data series, identifying which, if any, periodic components explain a significant portion of the data variance and describing the characteristics of these cycles (Warner, 1998). While certain data fluctuations relate to regular, predictable events, many of the cycle

lengths fitting the data are often not known a priori because certain cycles are not regular enough to be easily and accurately identified with visual examination. Spectral analysis is an exploratory method that enables the spectral decomposition of a timeseries, i.e. the deconstruction of a timeseries into a set of cyclic components (Figure 2.14).

The Fourier Transform (FT) is a mathematical operation which transforms the data from the time domain into the frequency domain by approximating the signal with a variety of sinusoidal waves, or cyclic components, allowing one to delve deeper into the data series to understand the underlying periodic patterns (Percival & Walden, 1993). Any periodic signal can be approximated as a sum of sinusoidal waves, and the higher the number of waves used to approximate the signal, the closer the approximation is to the original (Figure 2.15). More specifically, the FT involves producing estimates of complex exponential functions, or spectral components, of a signal over a frequency range (Rapuano & Harris, 2007).

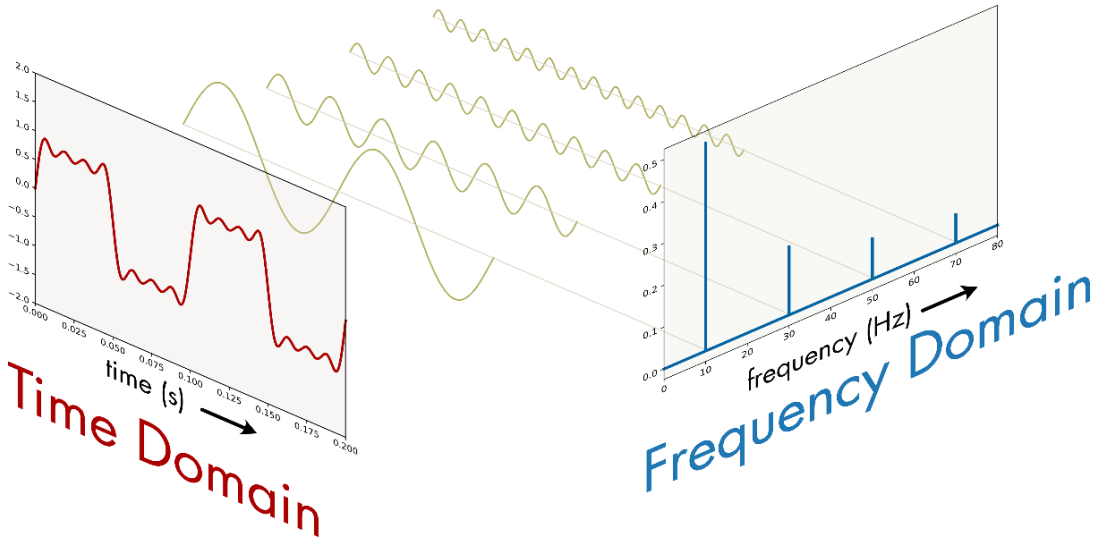


***Figure 2.14**: A signal in the time domain is broken down into its consitutent sine waves. The frequency domain represents an alternative view of the signal, with the frequency axis (x-axis) and amplitude (y-axis) of each component shown. Adapted from McKinnon & Ribas (2024).*

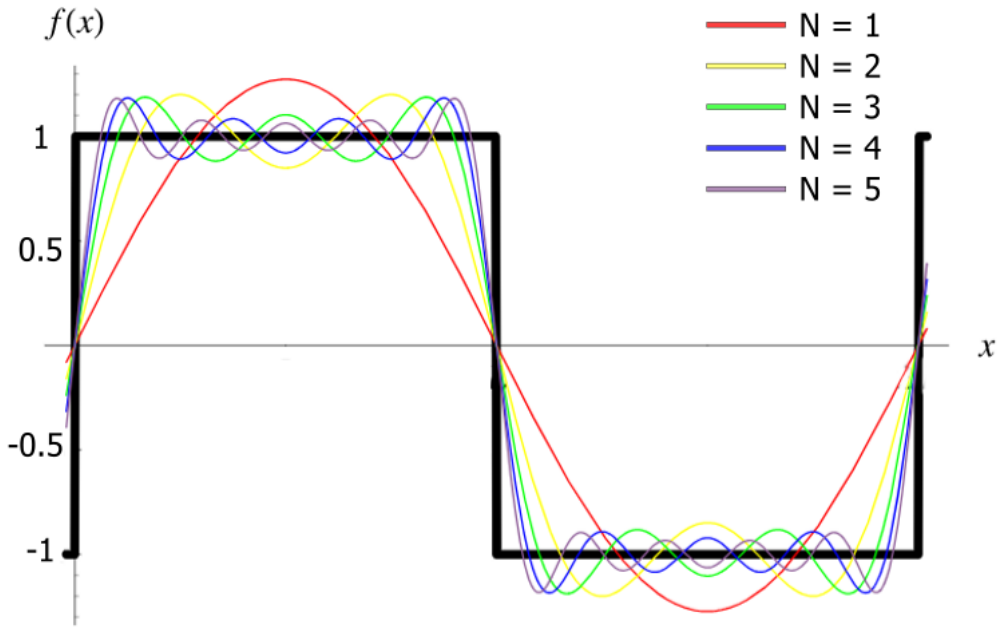


Figure 2.15: *The Fourier series approximation of a square signal (black) by summing an increasing number (N) of sinusoidal waves with decreasing amplitude. The figure illustrates how as an increasing number of sinusoidal waves are added, the approximation becomes progressively closer to the original signal. Adapted from Ma & Sun (2020).*

The Discrete Fourier Transform (DFT) is a form of the FT used when investigating discrete data, i.e. data measured successively over a finite period, as the FT assumes the data to be continuous. The frequencies associated with a transformation are subject to limits on the maximum frequency and frequency resolution. The maximum frequency is dependent on the sampling intervals of the original signal and frequency resolution is determined by the total length of the signal (Sevgi, 2007). A data series spanning a longer time period will have a higher frequency resolution, which is given by the following equation:

$$\Delta f = \frac{1}{T} \tag{2.14}$$

where Δf is the interval between successive frequencies, and T is the duration of the finite signal.

When t is the interval between consecutive measurements (for hourly data $t = 1$), then the highest frequency achievable is:

$$f_{max} = \frac{1}{2\Delta t} \tag{2.15}$$

The periodogram, calculated as the square of the absolute of the complex DFT output as a function of frequency, is a typical representation of the data transformation. As an estimate of the spectral density function, the periodogram is a measure of the relative strength of each cyclic component and can be used to identify dominant frequencies within a timeseries. With a finite data series of length N, the periodogram, P, for a finite timeseries is calculated as follows:

$$P(v_k) = |f(k)|^2 = \left| \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{(-2\pi i v_k n)} \right|^2 \tag{2.16}$$

where x_n, $k = 0, 1, \ldots, N - 1$, and $v_k = k/N$.

2.3.3.2 Frequency Interpretation

The power of a signal is conserved during the DFT (Percival & Walden, 1993), and the integral of the spectral density function with respect to frequency is equal to the variance of the original data series (Marr & Harley, 2002). Therefore, the contribution of specific cyclic components can be quantified by calculating the proportion of variance within specific frequency bins relative to the total variance of the periodogram. This can be achieved by integrating under the specific frequency bins and comparing to the integrated area under the entire periodogram, where specific frequency components are associated with distinct pollution sources, or source

categories. In urban environments and smaller population centres a variety of pollutant origins generate a heterogenous aerosol profile due to varied chemical compositions, and the chemical composition and hourly concentration of PM can vary from location to location within the one town or city (Ceburnis et al., 2006; Cesari et al., 2016; Kim Oanh et al., 2009; Tunno et al., 2012). Locally sourced $PM_{2.5}$ will exhibit rapid, short-term fluctuations in concentration measurements due to the heterogeneity of the aerosol which is influenced by local dispersion conditions and anthropogenic patterns in the local emissions (Seinfeld & Pandis, 2016). These rapid concentration fluctuations are approximated with high frequency components in the frequency domain. In contrast, regionally transported emissions result in more homogeneous $PM_{2.5}$ chemical profiles and slowly elevated baselines in the concentration measurements due to the transport of aerosols over a wide area contributing to the overall background concentration of PM in the atmosphere. The regional or seasonal changes in $PM_{2.5}$ correspond to low frequency cycles.

2.3.3.3 DFT Pre-Processing

A key processing step must be carried out prior to transforming the data which is related to a basic property of the DFT – the value output as the zero-frequency component is related to the mean of all samples in the data series (a signal component with 0 Hz frequency is a constant component), known as the DC offset. In order to centre the data series around zero and avoid this large zero-frequency component which may obfuscate the other frequencies, the overall mean of the given data series was subtracted from that entire data series. This does not affect any other resulting frequency components as the relative change between consecutive measurements

remains the same. Additionally, the DFT is applied with the assumption that the data is present at equally spaced intervals. Any missing data points in AQS and reference datasets were interpolated to achieve the necessary equally spaced measurements. If the first or final data point in the series was missing, it was replaced with the overall mean of the data series. Other missing data points throughout the series were replaced with the mean of the two nearest neighbouring data points (i.e. the pair of nearest data points before and after the missing data point). In the final processing step, the overall mean was subtracted from each data series, thereby centring them around zero.

2.3.4 Source Apportionment of Equivalent Black Carbon

The principal data parameters exported during data acquisition from the AE33 aethalometer are the attenuation measurements for the three spots, the calculated eBC mass concentration for each of the spots, and the filter loading correction factor, *k*. The AE33 measurements facilitate source apportionment analysis on the estimated eBC mass concentration, based on the wavelength dependent absorption efficiency. The source apportionment model describes the carbonaceous material with respect to the absorption coefficients of particles originating from distinct sources (biomass burning (eBC_{BB}) and fossil fuel (eBC_{FF}) related emissions) measured at various wavelengths (Sandradewi et al., 2008b). Using the Ångström absorption exponent (AAE), or alpha (α) values, the model differentiates between the various sources of BC in line with the following relationship describing the spectral dependence of absorption:

$$b_{abs} \propto \lambda^{-\alpha} \quad \textbf{(2.17)}$$

where b_{abs} is the absorption coefficient and λ is the wavelength. According to this relationship, α is the exponential slope in the regression model of absorption as a function of wavelength (Grange et al., 2020; Moosmüller et al., 2011). The model is based on the different absorption of BC particles originating from different sources. Organic components present in wood smoke tends to result in strong absorption in the UV region, while traffic-related emissions tend to absorb strongly in the IR region (Sandradewi et al., 2008a). The AE33 software provides an on-line analysis using a basic implementation of the source apportionment model employing an AAE pair of $\alpha_{FF} = 1$ and $\alpha_{BB} = 2$ for fossil fuel and biomass burning related emissions, respectively.

When using the information relating to two wavelengths, one high and one low, the following relationships are found:

$$\frac{b_{abs}(470\ nm)_{FF}}{b_{abs}(950\ nm)_{FF}} = \left(\frac{470}{950}\right)^{-\alpha_{FF}} \quad \textbf{(2.18)}$$

$$\frac{b_{abs}(470\ nm)_{BB}}{b_{abs}(950\ nm)_{BB}} = \left(\frac{470}{950}\right)^{-\alpha_{BB}} \quad \textbf{(2.19)}$$

$$b_{abs}(470\ nm) = b_{abs}(470\ nm)_{FF} + b_{abs}(470\ nm)_{BB} \quad \textbf{(2.20)}$$

$$b_{abs}(950\ nm) = b_{abs}(950\ nm)_{FF} + b_{abs}(950\ nm)_{BB} \quad \textbf{(2.21)}$$

where $b_{abs}(\lambda)$ corresponds to the absorption coefficient relating to the wavelength, λ, $b_{abs}(\lambda)_{FF}$ and $b_{abs}(\lambda)_{BB}$ are the fossil fuel fraction the biomass burning fraction of the absorption coefficient, respectively. Following this, the portion of biomass burning, *BB*(%), is calculated as:

$$BB(\%) = \frac{b_{abs}(950\ nm)_{BB}}{b_{abs}(950\ nm)} \quad \textbf{(2.22)}$$

The biomass burning and fossil fuel related eBC fractions are calculated as:

$$eBC_{BB} = BB(eBC) \quad \textbf{(2.23)}$$

$$eBC_{FF} = (1 - BB)(eBC) \quad \textbf{(2.24)}$$

The source apportionment model was applied off-line to the eBC measurements. The nomenclature was, however, adjusted in part from those described above. Solid fuel related (eBC_{SF}) emissions are used in placed of wood burning, the phrase used by studies conducted in continental Europe (Sandradewi et al., 2008a; Zotter et al., 2017), to reflect the variety of types typically used in Ireland. Traffic related (eBC_{TR}) emissions are used in place of fossil fuel, to incorporate typical road traffic and shipping traffic related emissions. Emissions relating to the burning of coal have been shown to have variable characterisations when using AAE values, are often subsumed into the solid fuel portion with the aethalometer (Buckley, 2019). The two source aethalometer model has limitations when applied to a multi-fuel environment as opposed to one where wood burning is the main solid fuel type used. Therefore, the suitability of various AAE pairs when applying the model in different time periods is investigated.

3 Highly Local Sources and Large Spatial Variations in $PM_{2.5}$ Across a City: Evidence from a City-Wide Sensor Network in Cork, Ireland

Reproduced from the published article:

Byrne, R., Ryan, K., Venables, D. S., Wenger, J. C., Hellebust, S. (2023). Highly local sources and large spatial variations in $PM_{2.5}$ across a city: Evidence from a city-wide sensor network in Cork, Ireland. *Environmental Science: Atmospheres*, *3*, 919–930.

https://doi.org/10.1039/D2EA00177B

3.1 Abstract

When dominated by local emissions, levels of ambient particulate matter (PM) can vary appreciably within a city. In Ireland, domestic solid fuel burning is the main $PM_{2.5}$ emission source. As a result, smoke from a small number of chimneys can have a major impact on the air quality in the surrounding area. The emergence of low-cost air quality sensors has enabled researchers to quantify this variation with high spatial density and high temporal resolution. In this study, a network of 16 PurpleAir (PA) devices measured $PM_{2.5}$ over a one-year period in Cork City, Ireland. Raw measurements were corrected based on the relationship between a co-located PA device and a reference instrument, and the influence of temperature and relative humidity were also accounted for. Hourly corrected data showed significant spatial variation across Cork City, with some areas experiencing daily average peaks approximately double that of others. Significant pollution periods occurred across the

different location types (city centre, areas surrounding the city centre, and commuter towns) causing infrequent exposure to high $PM_{2.5}$ concentrations. The seasonal variation was consistent across all locations, with higher variation seen in winter compared to summer diurnal profiles. The winter months displayed a characteristic evening peak in $PM_{2.5}$ associated with use of solid fuel for home heating. High frequency (two-minute timescale) data was analysed in a novel way to classify and quantify local sources and non-local sources of particulate emissions. Short-lived spikes in $PM_{2.5}$ were found to likely be from highly local sources, such as individual chimneys in the vicinity of the sensor. These characteristic spikes indicated the impact of highly local sources and short episodes of extreme air pollution. The classification estimates corroborated the spatial variation found and showed some devices were exposed to local emissions for 40–50% of the period analysed, while for others this contribution was below 20%.

3.2 Introduction

Particulate matter with an aerodynamic diameter below 2.5 μm ($PM_{2.5}$) is the air pollutant of greatest concern. Exposure to $PM_{2.5}$ affects the respiratory and cardiovascular systems, causes a range of other detrimental health effects and is responsible for several million global deaths every year (Kim et al., 2015; Waidyatillake et al., 2021). Despite greater knowledge about the impacts of air pollution on morbidity and mortality, the number of global deaths attributable to ambient $PM_{2.5}$ has risen by 102% from 1990 to 2019 (Sang et al., 2022). In 2020, the Health Effects Institute estimated that air pollution was the 4th leading risk factor for all deaths across the globe, making air pollution the leading environmental risk factor for premature death (Health Effects Institute, 2020). $PM_{2.5}$ is also the dominant air pollutant in Ireland (EPA, 2020). Air quality in Ireland is generally good during the warmer months of the year owing to Ireland's modest population density and the prevailing air masses from the Atlantic Ocean. On winter evenings, in contrast, Irish cities and towns frequently experience elevated levels of PM pollution from burning solid fuels for domestic heating (Kourtchev et al., 2011; Lin et al., 2018, 2019b; Wenger et al., 2020). As a result, the WHO Air Quality Guideline (AQG) values for $PM_{2.5}$ are frequently exceeded. For example, in 2020, the annual WHO (2005) AQG value was breached at 9 of the 64 air quality monitoring stations and the 24-hour mean value was exceeded at 34 stations (EPA, 2020).

Numerous studies have shown that several different solid fuels (coal, peat, and wood) contribute to PM pollution in Ireland (Dall'Osto et al., 2013a; Lin et al., 2017, 2018; Wenger et al., 2020), and in many homes across the country, solid fuel burning is a daily occurrence. A recent Irish Environmental Protection Agency (EPA) survey of

1823 households found that most (1043) were regular users of solid fuel for domestic heating (Eakins et al., 2022). Solid fuel was the primary source of heating in only 16% of households, but 38% of households used solid fuel for supplementary heating. A wide range of solid fuel types were used. Low-smoke coal and peat were prevalent fuels for primary heating, while low-smoke coal and wood logs or peat briquettes were the most common fuels for supplementary heating purposes.

Several strategies have been adopted to improve air quality in Ireland. The approaches include the introduction and expansion of "low-smoke zones" which forbid the sale of bituminous or “smoky” coal in towns and cities with populations over 10,000 people. New legislation to restrict the sale, distribution, and use of smoky fuels across the country was enacted on 31 October 2022. Irish air quality monitoring infrastructure has also been greatly expanded over the last few years. When the expansion programme is completed in 2023, the network will consist of 116 regulatory air quality monitoring stations across Ireland (EPA, 2022). As in other countries, however, the cost of equipping and operating each air quality monitoring station means that the number of stations in Irish cities is relatively small, and many towns have no monitoring infrastructure at all. As a result, regulatory air quality monitoring is sparsely distributed and provides little insight into the spatial variation of air pollution at the neighbourhood or town scale. Moreover, regulatory measurements of $PM_{2.5}$ are typically collected and reported on an hourly basis. Short-term changes (on the order of minutes) in local air pollution are poorly understood.

The capital and operational costs of regulatory air quality monitoring are a global challenge to measuring air quality. To address some of the limitations of regulatory air quality monitoring, so-called low-cost air quality sensors (AQSs) have attracted much recent attention (Alfano et al., 2020; Bousiotis et al., 2021; Brzozowski et al.,

2019; Castell et al., 2017; De Vito et al., 2020; Kosmopoulos et al., 2020). The cost of most AQSs is modest and they are easy to install and use, although expertise is needed in siting sensors, assessing measurement accuracy, and interpreting data. The accuracy of AQS measurements is lower than official reference and equivalent measurements and the quality of data can be variable. It is therefore important to carefully assess the accuracy of AQSs and, if necessary, calibrate the measurements against co-located reference instruments. Such calibrations can depend on the climatological and air pollution characteristics of specific locations (De Vito et al., 2020).

AQS measurements have been used around the world, but their application to building an air quality monitoring network was not explored in Ireland until 2019 when Cork City Council deployed PurpleAir (PA) sensors for monitoring PM across the city (Cork City Council & University College Cork, 2022). In addition to being the first AQS air quality network in Ireland, this was also one of the first citywide PA networks in Europe. The high frequency measurements and increased spatial resolution provided by the sensor network allows deeper insight into the spatial and temporal patterns of PM in the city. Such measurements therefore provide information that is complementary to the existing regulatory monitoring network.

This study aims to explore the use of PA $PM_{2.5}$ measurements in a citywide air quality monitoring network in order to:

1. Assess the spatial differences in $PM_{2.5}$ concentrations across different areas within the city.
2. Study the temporal characteristics of air pollution across the city. By exploiting the fast time response and wide distribution of the sensors,

intense local pollution sources such as chimneys in residential areas can be identified and discriminated from more distant sources.

The purpose of this work is therefore to serve as a case study to show how a AQS network can complement regulatory monitoring networks, provide deeper insight into $PM_{2.5}$ exposure across a city and support more strategic, targeted approaches for improving local air quality.

3.3 Materials and Methods

3.3.1 Air Quality Sensors

The AQS $PM_{2.5}$ network operated by Cork City Council is made up of 16 PurpleAir PA-II SD units distributed across the city (PurpleAir, Inc.), shown in Figure 3.1. Each PA unit consists of two Plantower PMS5003 sensors to measure $PM_{2.5}$ and PM_{10} mass concentrations ($\mu g\ m^{-3}$) based on light scattering. PA units and other systems based on the PMS5003 sensors are widely used for monitoring PM concentrations and, in general, their measurements are highly correlated with PM levels measured by gold-standard instrumentation (Ardon-Dryer et al., 2020; Sayahi et al., 2019). The measurement alternates between the two sensors every 5 seconds and the average for each sensor is obtained over a 120 second period. Data is uploaded to the PA online platform via a Wi-Fi connection and the data is also recorded on an SD card. The PA devices contain a Bosch BME280 sensor for pressure, temperature, and humidity readings. Raw sensor data is available via the PA API.



Figure 3.1: *Map of Cork City with PurpleAir locations.*

3.3.2 Data Specification and Preparation

3.3.2.1 Hourly $PM_{2.5}$ Measurements

The data collected by the PA units was processed and analysed using the Julia programming language and the openair package, which is based on the R programming language (Bezanson et al., 2017; Carslaw & Ropkins, 2012). For each PA unit in the sensor network, data from both PM sensors (sensor A and sensor B) was used. As each sensor measures data on individual 120 second measurement cycles, the timestamps were floored to the next lowest minute to extract common timestamps between the channels. Data between sensors A and B were usually highly correlated; however, some data points were excluded where there was a significant difference between the two channels. In situations where the value of one sensor measurement was above 15 $\mu g\ m^{-3}$, data were excluded if the relative difference between sensor A and B was greater than 5%. If a sensor measurement was below

15 μg m^{-3} and the relative difference between measurements from sensor A and B was greater than 50%, these data were also excluded. This procedure was carried out in order to omit any data points that would be considered an anomaly, as the two internal sensors would have grossly disagreed on the value measured at that time. The two-step removal approach was used to ensure that smaller PM values were not removed owing to the large relative differences between small values, which would unnecessarily reduce the dataset size. On average, this quality assurance procedure reduced the number of data points by 17% per PA unit. Data points above 1000 μg m^{-3} exceeded the operational range of the sensors and were also removed (PurpleAir, 2022). The mean of the remaining data points from the two channels produced a single stream of data per unit, with measurements approximately every 2 minutes. Hourly averages were calculated from this data.

The same steps were applied to all data retrieved from the PA units in the sensor network to create a dataset of hourly values between 01 January 2021 and 31 December 2021. The PA network dataset comprised a total of 228,371 individual hourly measurements. Although these dates incorporated some periods of COVID-19 pandemic restrictions, such restrictions mostly affected NO_2 concentrations and have not been found to have an unequivocal impact on PM levels in Ireland (EPA, 2020).

3.3.2.2 Two-minute $PM_{2.5}$ Measurements

Uncorrected, high frequency data were used to investigate the use of PA data for classifying the proportion of emissions from the immediate vicinity. This analysis was based on the relative gradients between successive data points, and it was deemed appropriate to use the uncorrected data for this purpose. The two sensors in each unit

were averaged across common timestamps, similarly to how the hourly data were processed. The analysis was applied to a winter period (December 2021) and a summer period (June 2021) with the chosen months corresponding to periods when the maximum number of devices were working. At other times in the year, some of the PA units were not functioning or transmitting data.

3.3.2.3 Meteorological Measurements

Weather data collected by Ireland's National Meteorological Service, Met Éireann, at Cork Airport (located 5.5 km from the city centre), was used for investigating the association between local $PM_{2.5}$ levels and meteorological factors. This data was accessed from the Met Éireann website (www.met.ie).

3.3.3 Air Quality Sensor Performance Evaluation

Several PA devices (Units 1, 2, 3, 4 and 5) were deployed at the University College Cork (UCC) air quality monitoring station (51.900111, -8.486329) for co-location with a beta-attenuation monitor (BAM-1020, Met One Instruments) and inter-unit comparison. The units and co-location periods are listed in Table A1.1. These additional units did not form part of the city-wide network.

3.3.3.1 Sensor Calibration

The BAM-1020 is classified as a reference equivalent method for measuring $PM_{2.5}$ mass concentration and has a measurement cycle time of 1 hour, a measurement resolution of ± 0.1 μg m^{-3}, and an operational range of 0–1000 μg m^{-3}. Raw, hourly

PA $PM_{2.5}$ values were corrected using a calibration factor derived from comparing data collected by Unit 1 to that obtained by the BAM-1020 over a period of seven months. A multiple linear regression (MLR) model, as described in Section 2.3.1, was used to correct the hourly PA data. An MLR is used to assess the association between two or more independent variables and a single continuous dependant variable and can be expressed using the general equation:

$$\hat{Y} = c_0 + c_1X_1 + c_2X_2 + \ldots c_nX_n \quad \textbf{(3.1)}$$

where $\hat{Y}$ is the predicted or the dependant variable, in this case the "true" $PM_{2.5}$ values, X_i through X_n, are the independent or predictor variables and c_0 through c_n are the estimated regression coefficients. Each regression coefficient represents the change in $\hat{Y}$ relative to a one-unit change in the respective independent variable. The model was trained by setting $\hat{Y}$ as the reference $PM_{2.5}$ value and the independent variables used were from meteorological variables along with the sensor data. The data from Unit 1 was randomly split into a training set and a testing set, with 70% of data in the training set. Data where the reference value was more than five times the PA value were deemed outliers and were removed; these instances were minimal when considering the trend between the PA values. A total of 3,981 data points were used for the calibration.

The MLR model incorporated the known effects of relative humidity and temperature on PM measurement (Malings, et al., 2020a). Multiplicative terms were also included to capture the full relationship between PM and the meteorological variables. However, the coefficients of multiplicative terms were small, which has also been found in previous studies (Barkjohn et al., 2021), and these terms were ignored in

applying the calibration equation to the data. The full calibration equation is shown in Equation 3.2.

$$\hat{Y} = -5.63 + 0.80(X) + 0.19(T) + 0.06(RH) - 0.00002(X * T) \quad \textbf{(3.2)}$$
$$-0.003(X * RH) - 0.001(T * RH) - 0.0002(X * RH)$$

where $\hat{Y}$ is the calibrated sensor value, X is the raw sensor value, T is the temperature in degrees Celsius, and RH is the percentage relative humidity.

The PA and BAM measurements were well correlated (R^2 = 0.9288) with a small offset (0.93 μg m^{-3}). The PA measurements were consistently higher than the reference measurements (slope = 0.58), as has been found in previous studies with PA devices and other air quality sensors (Datta et al., 2020; Karaoghlanian et al., 2022). After applying the correction, good agreement between the reference and PA measurements is observed, although the corrected PA values are somewhat lower than the reference data under high aerosol loading (Figure A1.1, Figure A1.2).

3.3.3.2 Inter-unit variability

It is important for this analysis that the devices exhibit low inter-unit variability for the purpose of meaningfully comparing different locations. Units 2, 3, 4, 5 were simultaneously co-located at the UCC monitoring station and the hourly averaged uncorrected data for a two-week period were analysed to establish the variability between different units. The time series and box plots for the four units during this period are shown in Figure 3.2, with the mean and maximum values and R^2 values between the devices recorded in Table 3.1. The agreement between the mean and maximum concentrations determined by the different units was excellent. The units

also showed a strong correlation with each other, even at very low $PM_{2.5}$ values. These results suggest that different PA units reliably produce similar measurements and that variations in unit-to-unit response are therefore not primarily responsible for the significantly differences in $PM_{2.5}$ concentrations measured across the Cork PM sensor network.

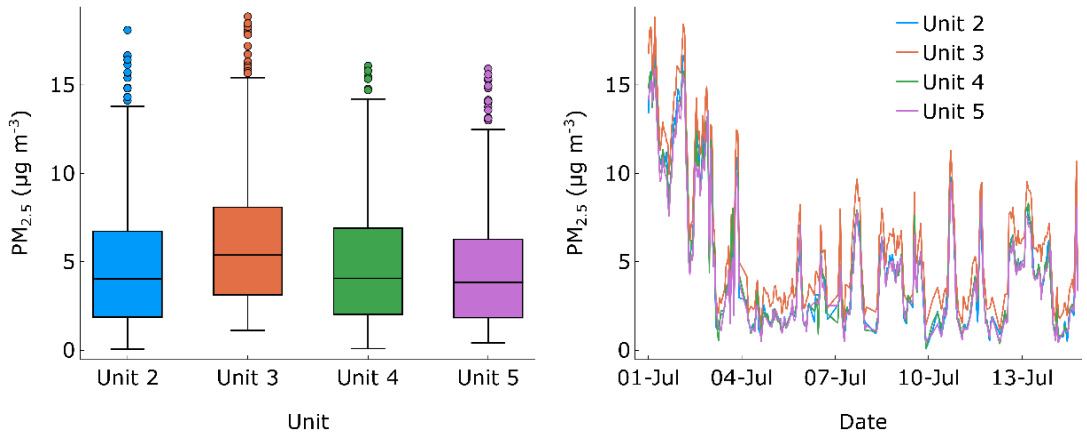


Figure 3.2: *Box plots and time series for uncorrected data from Units 2, 3, 4, and 5 obtained during the inter-unit sensor comparison period (n = 312 per device).*

The correlation between the two internal sensors in each device was investigated using the two data channels with common timestamps after quality assurance steps were carried out. Both sensors being highly correlated ($R^2 > 0.98$) and agreed closely with each other (Table A1.2). The excellent correlation between the two sensors in each device demonstrates that sharp changes in the high frequency data are linked to changes in atmospheric composition and do not arise from measurement noise.

***Table 3.1**: R^2 values for the inter-unit comparison and mean and max values for each unit.*

	Unit 2	**Unit 3**	**Unit 4**	**Unit 5**	**Mean**	**Maximum**
					µg m^{-3}	**µg m^{-3}**
Unit 2	1				5.05	18.11
Unit 3	0.989	1			6.37	18.87
Unit 4	0.989	0.992	1		4.96	16.09
Unit 5	0.989	0.988	0.991	1	4.75	15.93

3.4 Results and Discussion

3.4.1 Spatial and Temporal Variation in $PM_{2.5}$ Pollution

The calibrated network of PA units was used to analyse the spatial and temporal variations in $PM_{2.5}$ across Cork City. The units were categorised by their location in the city (Table 3.2). Note that the devices located in the commuter towns are within the official city boundary.

The winter (01 October 2021–31 December 2021) diurnal profiles for $PM_{2.5}$ (Figure 3.3a) show the expected pattern for locations in Ireland where residential solid-fuel burning emissions dominate. The air temperature in Ireland generally starts to drop in October so to capture the effect of the full heating season we take October as the beginning of winter. In contrast, the diurnal profiles for the summer months (01 June 2021–31 August 2021) shows lower levels of $PM_{2.5}$ and minimal variation across the day (Figure 3.3b). Winter $PM_{2.5}$ concentrations around noon are similar to those in summer. However, from about 15:00 on winter days, there is a sharp increase in $PM_{2.5}$ as households start burning solid fuel for home heating.

Although the same diurnal pattern is observed across the city, marked differences are seen in the maximum $PM_{2.5}$ levels in different regions of the city, indicating that residents in some parts of the city are exposed to appreciably higher PM concentrations than those in other areas. The devices categorised as City West showed diurnal profiles with peak $PM_{2.5}$ concentrations about half of those in other areas. Surprisingly, the outlying commuter towns (Blarney and Ballincollig) experienced the highest daily average for $PM_{2.5}$ concentrations, even above those of neighbourhoods within the contiguous built-up areas of Cork City. While the values of $PM_{2.5}$ are indicative due

to the nature of the sensors, the data provides clear evidence of the relative variation in $PM_{2.5}$ concentrations between location types and seasons.

Table 3.2: *City area allocation and maximum diurnal average $PM_{2.5}$ for each device with sufficient yearly data in 2021.*

Location type	**Device ID**	**Maximum diurnal value µg m^{-3}**
North	CCC4	13.4
	CCC6	9.0
	CCC14	18.9
South	CCC1	10.1
	CCC5	11.2
	CCC9	14.1
East	CCC2	11.4
	CCC12	14.8
	CCC13	14.5
West	MTU	7.5
	CCC8	8.9
City Centre	UCC	11.3
	CCC3	9.1
	CCC11	11.5
Commuter Towns	CCC15	19.4

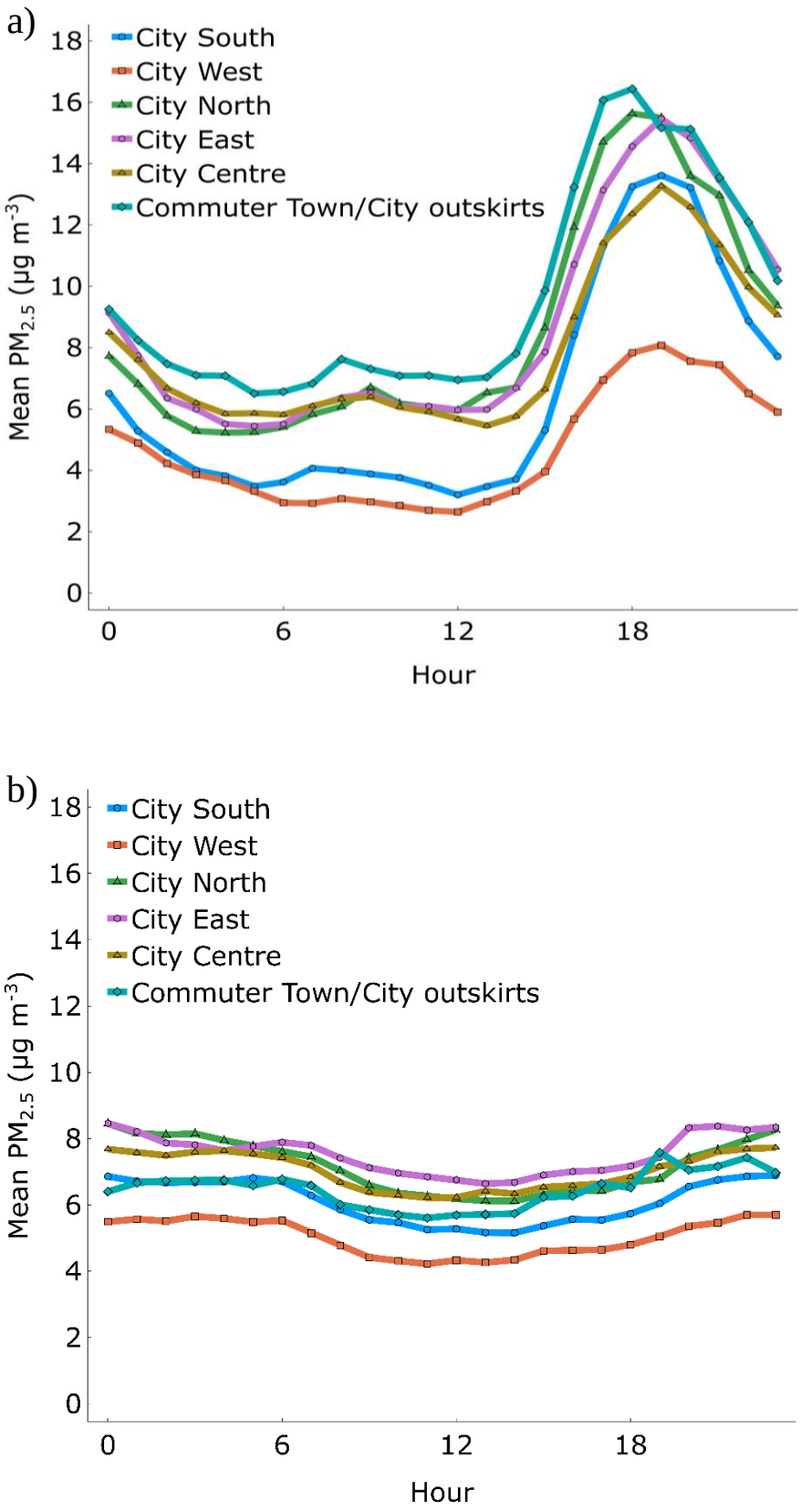


Figure 3.3: *2021 $PM_{2.5}$ diurnal averages per location type during a) winter and b) summer.*

The comparison of overall average $PM_{2.5}$ concentrations for the summer and winter periods is shown in Figure 3.4. The commuter towns experienced the largest increase in PM levels from summer to winter. Based on the diurnal profiles, the significant seasonal variation in $PM_{2.5}$ across different location types suggests a major impact from solid fuel burning.

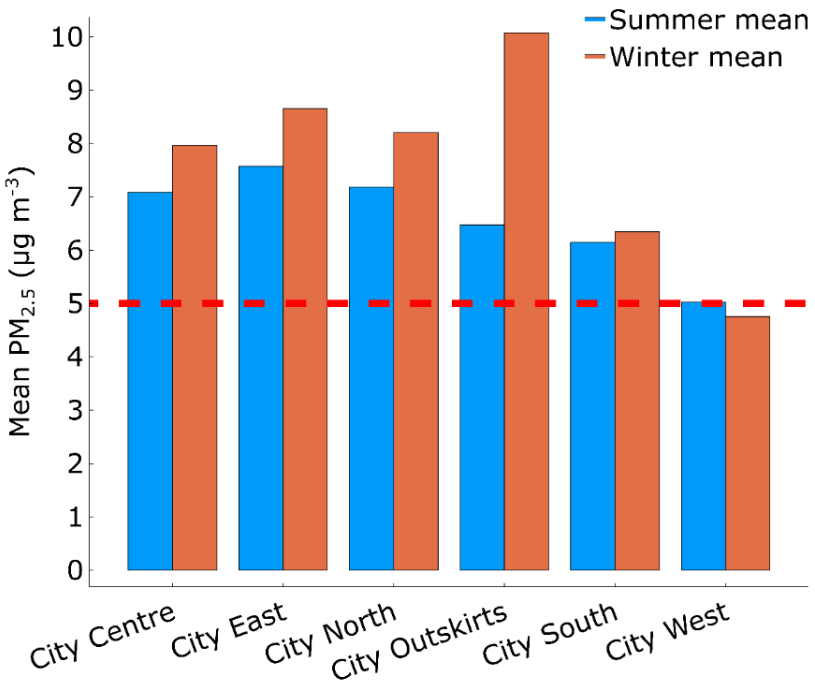


Figure 3.4: *2021 $PM_{2.5}$ averages over summer and winter periods per city district with WHO AQG annual mean (5 µg m^{-3}) indicated.*

The diurnal profiles show the total hourly average $PM_{2.5}$ levels in a location across a given period, however it is also important to consider the effect of pollution events which may not be apparent when looking at the average values. The maximum hourly average from the winter dataset was found for each location type. Table 3.3 compares these values to the maximum values seen in the diurnal profiles along with the mean value for the winter period. The commuter town locations showed a diurnal average

peak around double that of the city west devices, while they both exhibited significant maximum hourly averages. This suggests that while a location type may experience lower PM levels on average relative to elsewhere in the city, there can still be significant exposure to high levels of $PM_{2.5}$ during pollution events.

Table 3.3: *2021 Winter diurnal and, hourly maxima and overall mean values across location types.*

Location type	**Maximum diurnal value**	**Maximum hourly value**	**Mean**
	µg m^{-3}	**µg m^{-3}**	**µg m^{-3}**
North	15.6	84.2	8.2
South	13.6	99.8	6.3
East	15.5	57.6	8.7
West	8.1	49.2	4.8
Centre	13.3	73.2	8
Commuter towns	16.4	56.7	10.1

3.4.2 Identification and Estimation of Local Emissions

High frequency time series data from the network units exhibited intense, short-lived spikes in $PM_{2.5}$ concentrations (Figure 3.5). These spikes predominantly occurred on winter evenings. Across the network, both sensors A and B in the different PA units consistently showed the same behaviour with a high correlation coefficient between the two sensors, including during the spikes (Table A1.2). We conclude therefore that the spikes are not measurement noise but indicate rapid variations in the PM concentrations measured by the different sensor units.

The high-frequency (two-minute interval) data was analysed to investigate the contribution of local area emissions of $PM_{2.5}$ to the total $PM_{2.5}$ burden across the city. Using uncorrected $PM_{2.5}$ data obtained by the PA devices for June and December (to represent summer and winter and to ensure sufficient sensor data was available during analysis), the difference between two consecutive measurements (*ΔPM)* was calculated:

$$\Delta PM_t = PM_t - PM_{t-1} \quad \textbf{(3.3)}$$

where ΔPM_t is the *ΔPM* value at time, *t*, and PM_t and PM_{t-1} are the corresponding PM values at times *t* and *t-1*. Large *ΔPM* values indicate substantial changes in $PM_{2.5}$ concentration over a very short time frame. Such changes likely arise from local emission sources rather than $PM_{2.5}$ in transported air masses. Figure 3.5a shows the *ΔPM* for unit CCC9 during the first week of December 2021 and Figure 3.5b shows the corresponding plot for the first week of June 2021.

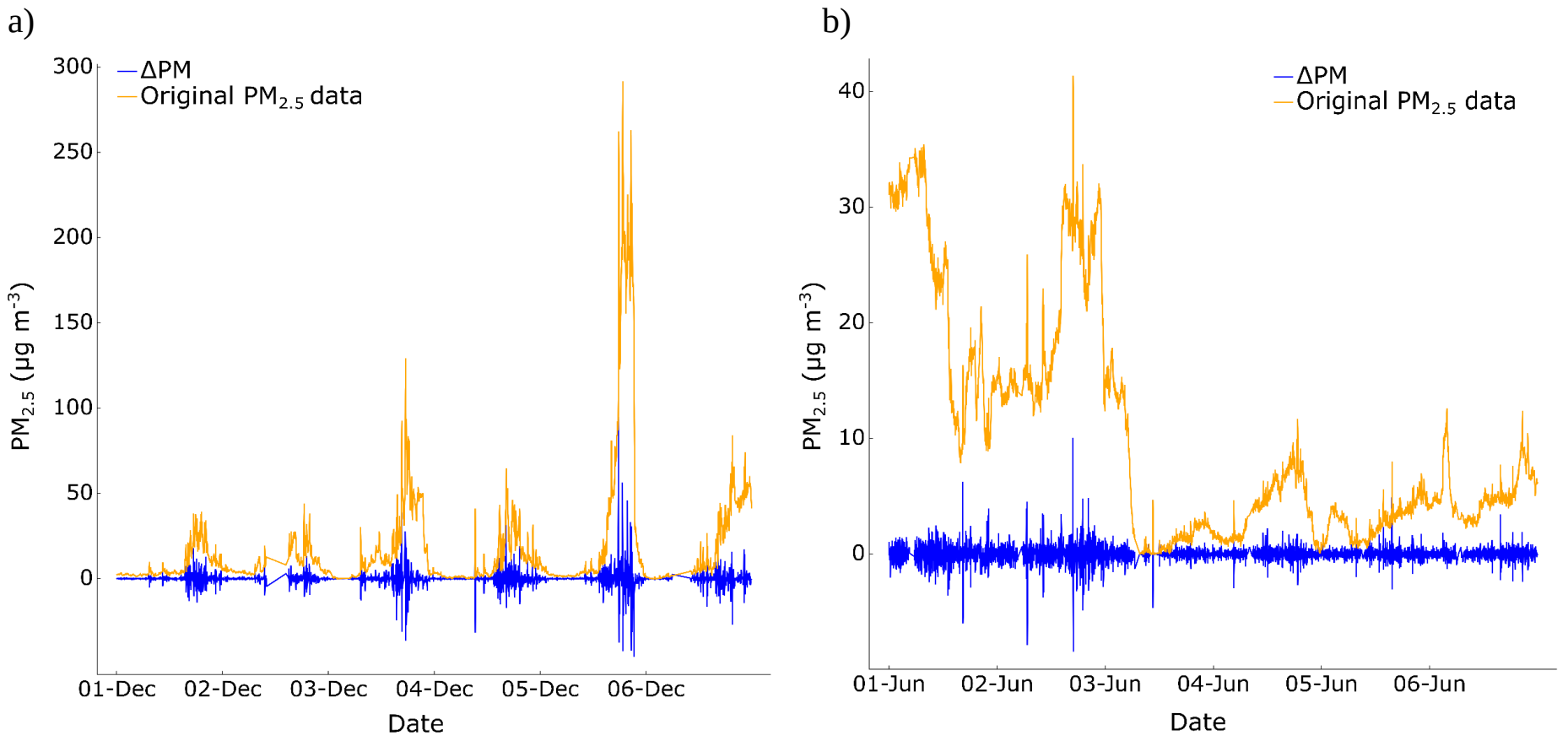


Figure 3.5: *Example time series for unit CCC9 with ΔPM during a) the first week in December 2021, and b) the first week in June 2021.*

Hourly standard deviations of ΔPM ($\sigma_{\Delta PM}$) were calculated to investigate the variability in PM concentrations at different times:

$$(\sigma_{\Delta PM})_T = \sqrt{\frac{\sum(\Delta PM_i - \bar{x}_T)^2}{n-1}} \quad \textbf{(3.4)}$$

where $(\sigma\Delta PM)_T$ is the standard deviation of ΔPM over hour T, ΔPM_i is the ΔPM value at time i (within the hour T), $\bar{x}_T$ is the mean of ΔPM over hour T, and n is the total number of ΔPM values in hour T. This approach was selected to identify periods of rapid fluctuations in measurements and to ensure that the selection process was not influenced by the background concentration or more gradual changes in PM levels over time. To identify periods of rapid fluctuations, periods where the hourly standard deviation was greater than the mean hourly standard deviation for a given dataset were highlighted, as shown on the unit CCC9 data in December (Figure 3.6a) and June (Figure 3.6b). In these figures, the data points in blue correspond to periods of low deviation from the hourly mean, and the periods in yellow correspond to the instances highlighted by the $\sigma_{\Delta PM}$ filtering process where there is higher deviation from the hourly mean value, indicating rapid fluctuations in $PM_{2.5}$ levels.

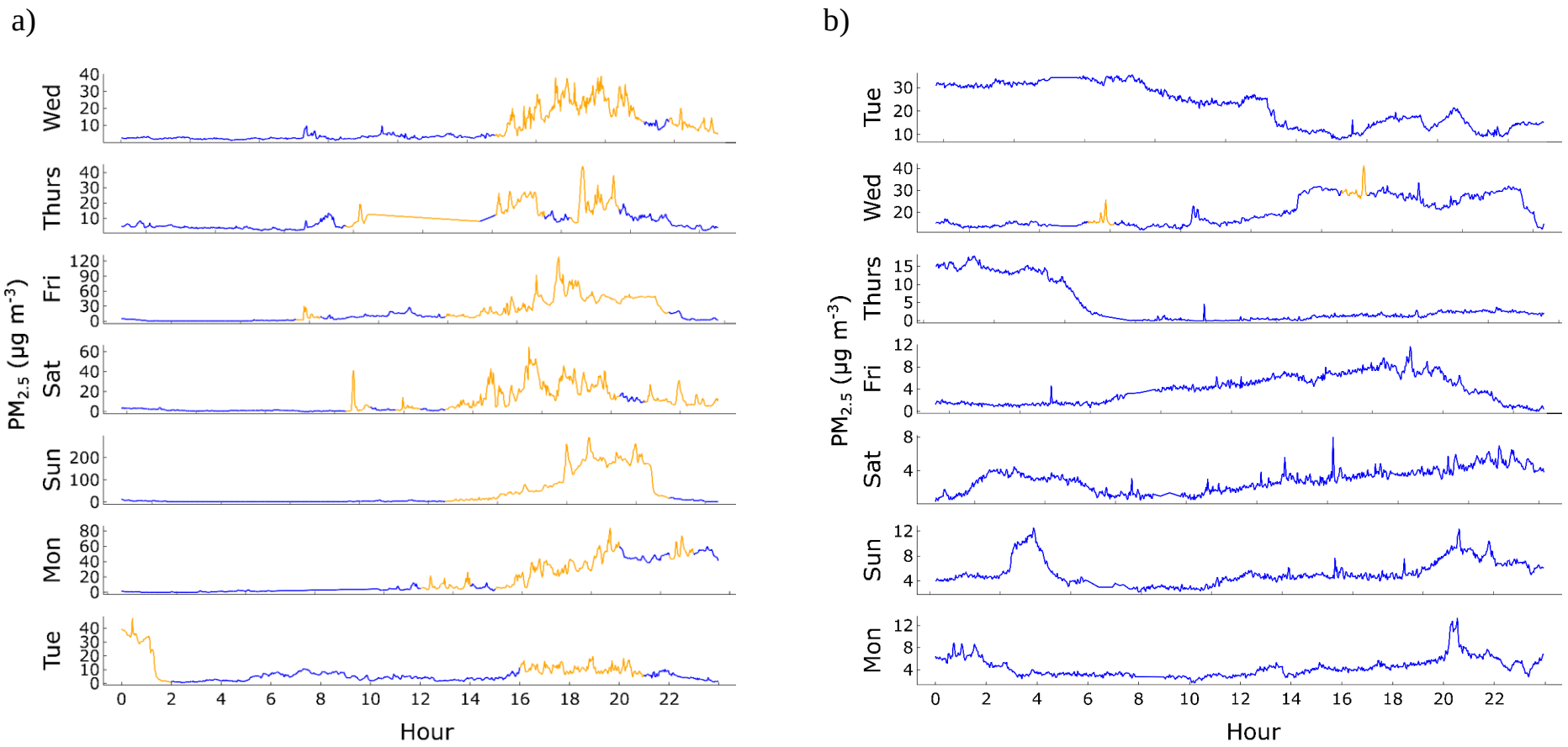


Figure 3.6: *σΔPM filter applied to unit ccc9 for a) the first week in December 2021 and b) the first week in June 2021. Blue data points correspond to periods of low σΔPM, yellow data points correspond to periods of high σΔPM.*

Periods of large fluctuations in PM concentrations (where the data deviated appreciably from the hourly mean) were much more frequent in December than in June. Even though some periods with high $PM_{2.5}$ concentrations were observed in June, these instances were generally not highlighted by the filtering process. The rapid fluctuations typically occurred at times when solid fuels are burnt for home heating. Large fluctuations in $PM_{2.5}$ levels are sometimes also a feature of sites impacted by intermittent, high traffic emissions such as at busy traffic junctions. While this method cannot distinguish between traffic and solid fuel burning related emissions, traffic emissions are unlikely to be a major contributor to the majority of sensors within the Cork AQS network as the rapid fluctuations were a common characteristic across all PA units, despite the fact that most units were not immediately adjacent to busy roads and intersections.

The average $\sigma_{\Delta PM}$ across each daily hour for June and December was calculated. On average, hour 18 (18:00–19:00) has the highest hourly standard deviation in December, with the lowest deviation seen in the early morning hours (despite relatively high PM levels). The winter month shows higher hourly deviation of ΔPM compared to the warmer summer month by a factor of 2.5 on average. The percentage of times with high hourly fluctuations was calculated (Table 3.4). $PM_{2.5}$ concentrations in the various locations had markedly different profiles for the winter period, with the proportion of time with high fluctuation periods ranging from 8% to 55%.

Table 3.4: *Estimated $PM_{2.5}$ contributions during times dominated by local sources for June and December 2021.*

Device ID	Location type	Percent of time identified as having local influence		Percent of total $PM_{2.5}$ observed during locally influenced periods		Percent of additional $PM_{2.5}$ observed attributable to local emissions	
		%	%	%	%	%	%
		June	December	June	December	June	December
CCC4	North	5	-	7	-	7	-
CCC5	South	5	30	7	56	7	55
CCC9	South	4	36	7	75	6	74
CCC2	East	8	34	8	56	7	55
CCC12	East	4	41	5	57	6	57
CCC13	East	18	55	28	87	27	87
MTU	West	5	16	7	38	7	37

Table 3.4 continued.

Device ID	Location type	Percent of time identified as having local influence %		Percent of total $PM_{2.5}$ observed during locally influenced periods %		Percent of additional $PM_{2.5}$ observed attributable to local emissions %	
		June	**December**	**June**	**December**	**June**	**December**
CCC8	West	2	11	5	25	5	24
UCC	City Centre	1	8	3	24	3	23
CCC11	City Centre	14	24	17	41	17	40
CCC7	Commuter Town	22	41	21	61	19	60
CCC15	Commuter Town	-	36	-	61	-	60
Average		**8**	**30**	**10**	**53**	**10**	**52**

To estimate the difference in measured $PM_{2.5}$ levels between periods of high frequency fluctuations and periods of more stable $PM_{2.5}$ levels, the relative differences were assessed using integration of the concentration-time profile. The total area under the $PM_{2.5}$ concentration-time profile (A_T) represents 100% of the concentration. The total integrated area under the profile of all high fluctuation periods for a given device (A_F) represents the $PM_{2.5}$ concentration measurements associated with high frequency periods and the fraction A_F/A_T represents the percentage of $PM_{2.5}$ exposure experienced during high frequency periods. These values are reported for both June and December in Table 3.4. During December, the A_F/A_T values for each location show considerable variation, with estimated percentages ranging from 24% to 87%. This variation is less apparent in June, with consistently low values measured by most devices. These figures, when compared with the percentage of time that locations experienced high frequency variations in $PM_{2.5}$ levels, suggest that on average, 53% of the $PM_{2.5}$ was measured during 30% of the time across the network locations during December. In other words, for a resident population, over half of their $PM_{2.5}$ exposure occurred during a third of the time, and based on the $\sigma_{\Delta PM}$ analysis above it is argued that the $PM_{2.5}$ emissions during those periods originated in the local area. One further estimate can be developed, which is the portion of $PM_{2.5}$ that is specifically attributable to local emissions above the background levels during the periods identified by the high frequency filter used in this study. If the integrated area A_f represents the area under the curve above a threshold defined as the background level, then A_f/A_T represents an estimate of the additional $PM_{2.5}$ in the locality that is due to emissions in the vicinity of the sensor. The background level here is taken as the level interpolated between the start time and the end time of each high frequency period. Estimates of both A_F/A_T and A_f/A_T are included in Table 3.4. Interestingly, the two

estimates are almost identical numerically, which suggests that without local emissions the $PM_{2.5}$ levels would be considerably lower than what was measured in this study.

This observation suggests that when the *ΔPM* filter is applied across a network of PA sensors, it can be used to distinguish between areas which experience greater levels of highly local PM pollution from those that experience city-wide PM. The high hourly standard deviation of *ΔPM* indicates periods with short-lived pollution events in a specific area. Based on previous work, such intense pollution is associated with smoke plumes from nearby chimneys (Dall'Osto et al., 2013a; Healy et al., 2010; Kourtchev et al., 2011). Of course, changes in wind speed and direction may also affect the air pollution in an area. However sustained changes in air pollution are associated with transported air masses from different directions while rapid fluctuations are associated with local emissions. The method employed here allows discrimination between these pollution sources. Furthermore, the high temporal resolution of the PA sensor measurements shows that the rapid fluctuations in PM levels are a characteristic feature of wintertime air pollution in Ireland. Such short-lived plumes would be missed with a lower temporal resolution, such as with hourly data.

Our A_f/A_T estimate assumes that the baseline level is not significantly changed and follows the same rationale as the Lenschow, or incremental, approach and is subject to similar dependencies (Lenschow et al., 2001; Pey et al., 2010). In the incremental approach, the urban impact is taken as the sum of the following terms: the Lenschow urban increment, the city spread and the background deviation (Thunis, 2018). In the approach presented here, local impact is confounded with the city spread, while the background is considered as the baseline level when local sources are non-dominant.

Notably, our A_F/A_T estimate makes no such assumption, but supports the same conclusion.

Analysis of the spatial and temporal variations in $PM_{2.5}$ across Cork City showed that the impact of local PM emissions varied from location to location, as shown by the differences in the diurnal averages across location types and in the winter and summer averages (Figures 3.3). The results showed that the devices categorised as City West showed the lowest PM concentrations during winter, with the sensors in the commuter towns showing the highest diurnal peak of all locations (Table 3.3). However, all location types experienced high maxima. This was in stark contrast to the average air quality being experienced for the overall period. The results show that while certain areas may experience lower daily average $PM_{2.5}$ overall during winter, they can at times experience strong maximum peaks resulting from more severe pollution events over the winter season. The results from this analysis provide support for the emission classifications given in Section 3.2. The estimated local contributions varied considerably across the devices, with some locations experiencing over double the local estimated contributions of others. The difference between the winter and summer results in the emission classification correlates well with the spatial and temporal variation analysis. By looking at the average $PM_{2.5}$ levels across the winter and summer months, it was shown that there are significant differences in $PM_{2.5}$ levels between the two periods and between locations. This is also shown with the emission classification estimates. During the colder period in December, there were elevated levels of estimated local $PM_{2.5}$ contributions, with some locations experiencing higher levels than others. In particular, the devices located in smaller suburban areas further from the city centre showed significant local $PM_{2.5}$ contribution estimates, comparable to or even higher than the more urban counterparts. For example, the device CCC15,

located in Ballincollig, a commuter town approximately 9 km outside of the city centre, experienced one of the highest local $PM_{2.5}$ contribution estimates during the winter period, which relates to the high winter averages seen in the spatial and temporal analysis. During the month of June, the estimated contribution of local $PM_{2.5}$ for all locations remained low and stable. This relates to the similar summer averages shown across the location types previously.

3.4.3 Meteorological Factors

It is a common observation that an important factor determining $PM_{2.5}$ levels in a locality is the weather. Emissions generally follow economic cycles and tend to exhibit regular patterns, whereas the weather is highly variable. So, the weather influences the $PM_{2.5}$ levels by influencing emissions, such as causing increased home heating and energy use during cold periods, and through the dispersion of emissions, such as effective transport of pollution by wind, with generally better mixing and dispersion at higher wind speeds. The weather is not generally responsible for directly generating $PM_{2.5}$ emissions, although PM_{10} levels can increase at higher wind speeds due to increased resuspension by wind.

The argument that is developed here is that during periods of less effective atmospheric mixing, i.e. at low wind speeds, the emissions are not dispersed effectively, and they build up in the local area where they are emitted. The prevailing wind direction in Cork City is from the South-West. However, as shown in the wind rose in Figure A1.3, during the period covered in this study, this pattern was not as evident, and the weather was less dominated by the prevailing wind direction than is normally the case.

Figure 3.7 shows polar plots for each device location indicating the wind direction and wind speed associated with higher $PM_{2.5}$ concentrations in each location. For some locations situated on the periphery of the city, such as the devices labelled MTU and CCC8, located at the western edge of the city, the higher $PM_{2.5}$ levels are clearly associated with low wind speeds from the easterly direction. For the most part, however, there is no consistent directionality to the $PM_{2.5}$ concentrations. Higher $PM_{2.5}$ is always associated with low wind speeds. These observations are again consistent with the suggestion that $PM_{2.5}$ levels are dominated by nearby emissions.

The meteorological data obtained from Met Éireann reflect the weather on a synoptic scale while the meteorological field across the city will also be influenced by local factors, street canyons and local topography. It therefore cannot be assumed that the wind direction measured at the airport (approximately 4–11 km from individual devices) is identical for all devices in the network. Nevertheless, the regional scale wind direction will dominate the general meteorological field across the city, and it can be reasonably supposed that transport of ambient $PM_{2.5}$ would be observable as a concentration gradient across the area if such transport was a significant factor in observed $PM_{2.5}$ levels. Wind speeds measured at the airport, which is situated in an elevated location above the city, generally exceed those experienced in the city. In other words, while the wind direction measured at the airport may have deviated from the wind direction at a given device, the fact that the polar plots imply no particular directionality suggests that wind direction is not an important factor. Moreover, the fact that higher $PM_{2.5}$ levels are associated with low wind speeds in general supports the hypothesis of very local origin.

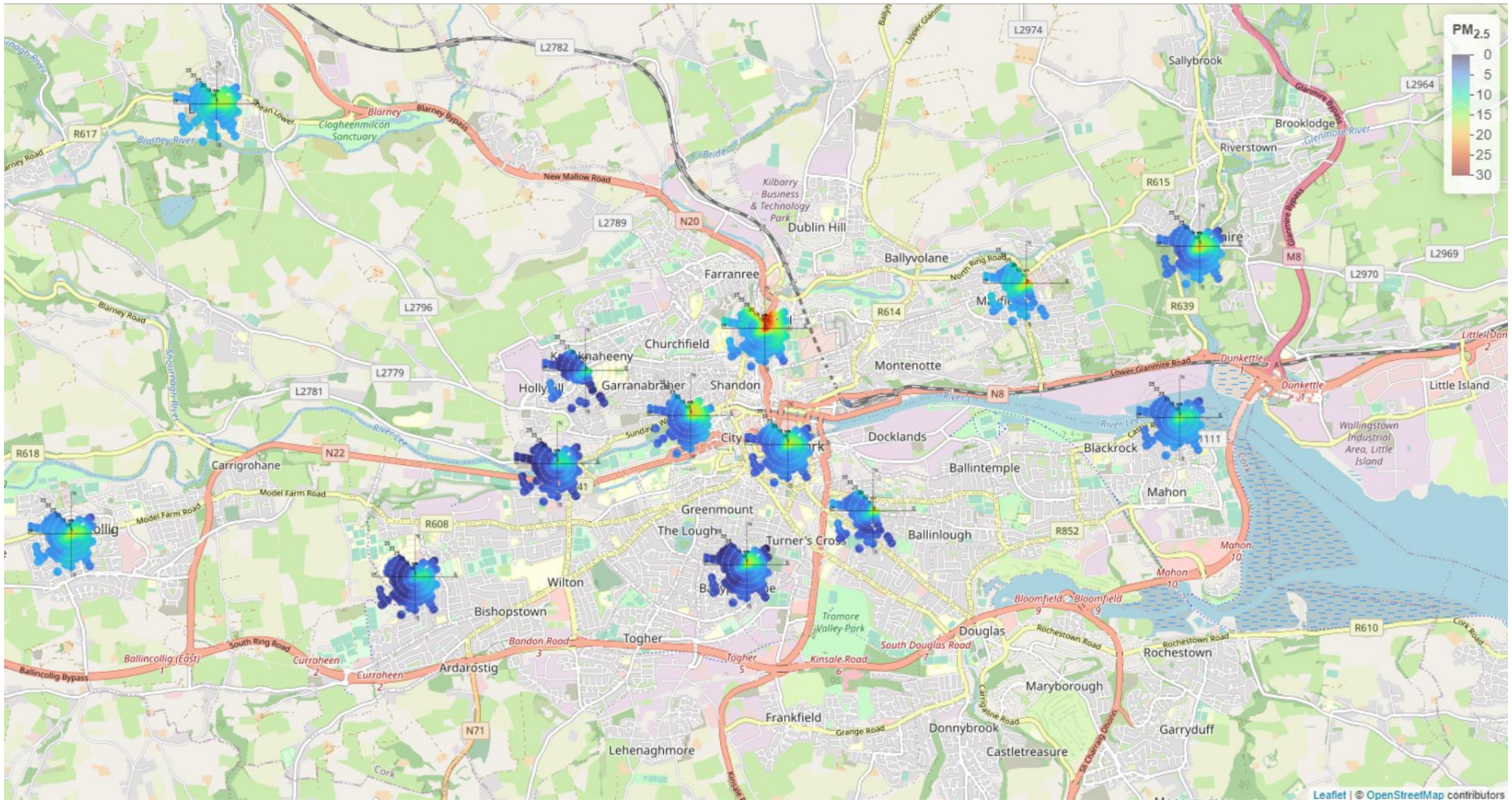


Figure 3.7: *Polar plots showing associations between wind speed and direction measured at Cork Airport, and local $PM_{2.5}$ levels at the locations monitored in the sensor network. Note the common scale for $PM_{2.5}$ concentration. High $PM_{2.5}$ at low wind speeds was exhibited in the majority of locations within the sensor network.*

3.4.4 Comparison with Source Apportionment Studies

The measurements and estimates reported in Table 3.4, particularly for December, suggest that measured $PM_{2.5}$ levels could mainly be attributed to residential solid fuel burning. A comparison with previous source apportionment studies in Cork and other places in Ireland, may therefore be instructive. Detailed source apportionment studies have been previously carried out in the Port of Cork, but they are separated from this study by over a decade. Using chemical markers associated with the combustion of peat, coal and wood, the estimated contribution of local solid fuel burning to $PM_{2.5}$ concentrations measured at the Port of Cork in August 2008 was 5% (Healy et al., 2010). Using a different approach, Kourtchev et al. (2011), obtained a value of 6% for the solid fuel burning contribution during the same period. These values are in line with the local $PM_{2.5}$ contribution estimates reported for many locations during June (Table 3.4), including CCC2, which is the PA unit closest to the Port of Cork measurement site. Kourtchev et al. (2011), also obtained a value of 28% for the solid fuel burning contribution to $PM_{2.5}$ measured during February 2009, while a more comprehensive source apportionment method was used by Dall'Osto et al. (2013a) to deliver an estimate of 46–50% for the same period. The results of this latter study are more consistent with the local $PM_{2.5}$ contribution estimates obtained for December 2021 in this work. A more recent study employing detailed in-situ chemical characterisation of ambient wintertime $PM_{2.5}$ and source apportionment modelling has been carried out at multiple diverse locations in Ireland, including the capital city, Dublin, a small midland town (Birr) and costal locations (Carnsore Point and Mace Head). Average estimates of local solid fuel contributions were 50%, 25%, 48%, and 39%, for Dublin, Carnsore Point, Birr, and Mace Head, respectively (Lin et al., 2019b).

The results obtained for the urban areas of Dublin and Birr are in line with local $PM_{2.5}$ contribution estimates obtained for December 2021 in this work.

While other studies have utilised air quality sensors in similar ways to estimate source contributions, these are based on measurements in different environments where the resulting estimates would not be comparable to the specific results presented here (Gameli Hodoli et al., 2023; Heimann et al., 2015; Lung et al., 2020). They do, however, showcase the usefulness and feasibility of using these kinds of devices in new and meaningful ways to attain more in-depth information about an area's air quality without using traditional and more costly facilities and instrumentation.

3.5 Conclusion

The PA sensor network in Cork City has provided information on the spatial and temporal variations in $PM_{2.5}$ across the urban area that were previously unachievable with traditional air quality monitoring methods. The accuracy of the measurements was improved by calibrating the sensor network through a co-location study with a reference monitor. While the PA data is considered indicative and not of regulatory monitoring standard, this study showcases the opportunities available with an AQS network. In particular, a carefully planned and implemented AQS network delivers complementary information to that from a regulatory grade air quality monitoring network. This includes expanding monitoring across cities where regulatory monitoring is already in place. Such AQS networks can provide valuable information on local air quality in a context where regulatory monitoring network is unaffordable, such as in developing country cities and in smaller towns in wealthy nations.

This study also shows it is possible to distinguish between local and background/other emissions using a PM sensor network. Such information is vital for developing targeted and effective interventions to improve urban air quality. We applied a novel method using high frequency data to highlight time periods where the PM measured by the AQS is assumed to be dominated by an intense local pollution source as opposed to a background/other source in the city. Significant differences were found between the different locations, in particular between devices west of Cork City and in the other areas of the city. The western locations had higher $PM_{2.5}$ levels associated with easterly winds of low wind speed and were among the lowest in terms of percent of time with high fluctuations and percent of $PM_{2.5}$ measured during these periods. However, the majority of devices showed higher $PM_{2.5}$ levels with low wind speeds,

but no consistent directionality, suggesting that local emission sources dominate the $PM_{2.5}$ levels.

The considerable difference in estimated local contributions highlights the importance of measuring in multiple locations within an urban area. Surprisingly, the surrounding commuter towns were found to have local PM contribution estimates that were on par with or above those of other location types. This shows that less populated locations and small towns can experience higher $PM_{2.5}$ levels than areas closer to the city centre. Nevertheless, all location types experienced high daily maxima, indicating that while daily average values may differ appreciably at locations across the city owing to higher instances of residential solid fuel burning or favourable meteorological conditions, all areas experience pollution episodes with high $PM_{2.5}$ concentrations.

3.6 Appendix 1

Table A1.1: *PurpleAir devices and periods when co-located at the UCC air quality monitoring station.*

Device ID	Operational dates
Unit 1	01 January 2021–01 June 2021
Unit 2	01 July 2021–14 July 2021
Unit 3	01 July 2021–14 July 2021
Unit 4	01 July 2021–14 July 2021
Unit 5	01 July 2021–14 July 2021

Table A1.2: *Correlations between sensor A and sensor B (validated data) within each PurpleAir device unit for the total dataset.*

Device ID	**R^2**	**Slope**	**y-intercept**
		µg m^{-3}	**µg m^{-3}**
CCC1	0.996	1.003	0.4
CCC2	0.998	0.999	−0.4
CCC3	0.999	1.032	0.0
CCC4	0.999	1.010	0.2
CCC5	0.999	0.961	−0.3
CCC6	0.999	1.027	0.1
CCC7	0.993	0.975	−0.5
CCC8	0.995	1.013	0.8
CCC9	1.000	1.006	−0.1
CCC11	0.999	1.013	0.1
CCC12	0.986	0.985	−0.2
CCC13	0.999	1.008	0.3
CCC14	0.999	1.044	0.4
CCC15	0.999	1.021	0.2
MTU	0.996	1.029	0.2
UCC	0.998	0.993	1.1

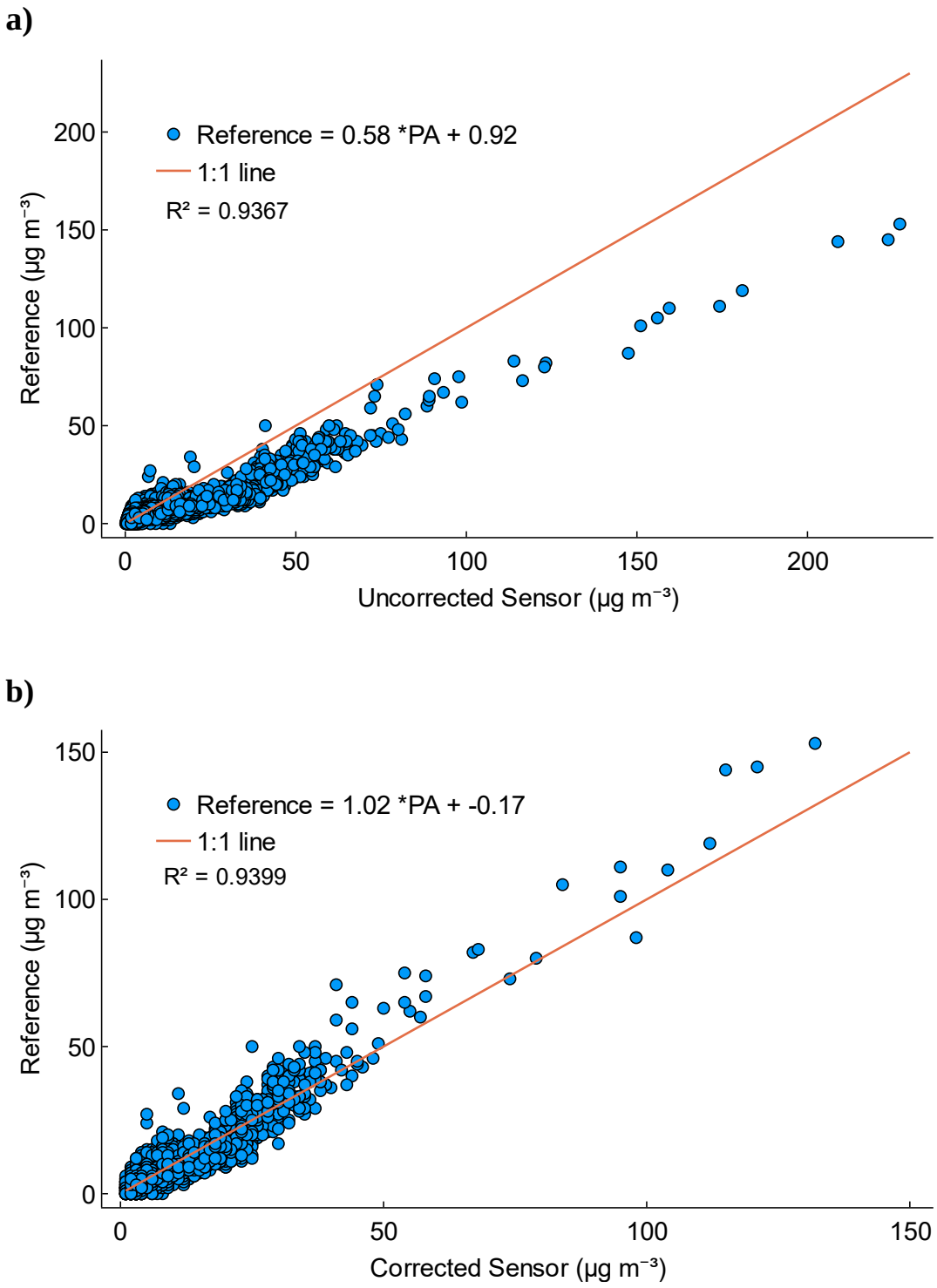


Figure A1.1: *Linear fit of hourly reference test data with a) uncorrected and b) corrected Unit 1 data.*

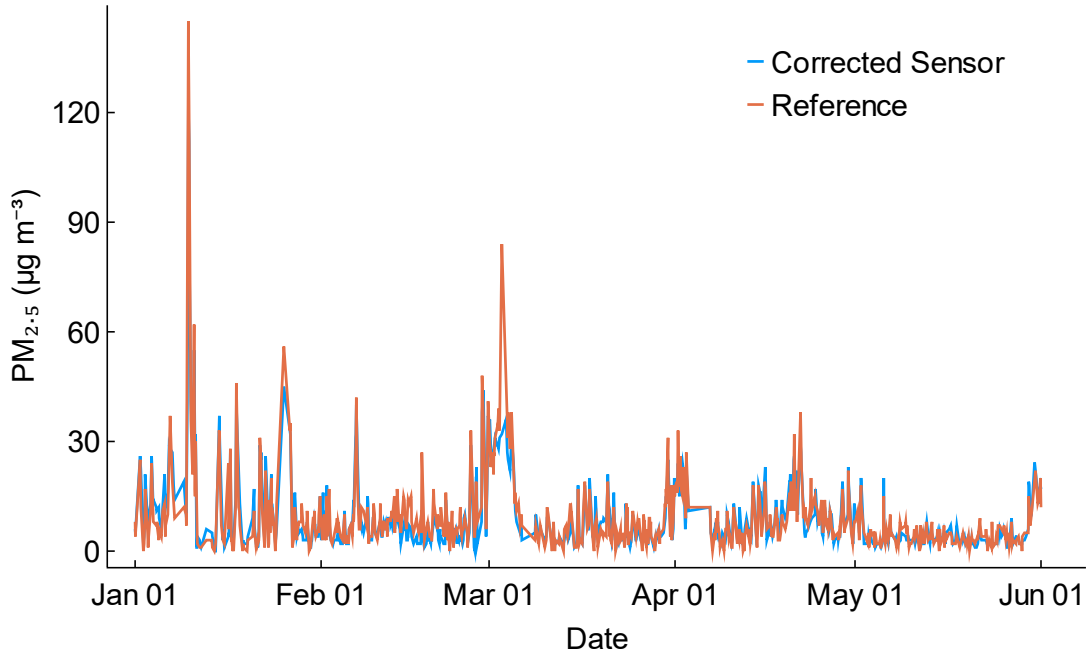
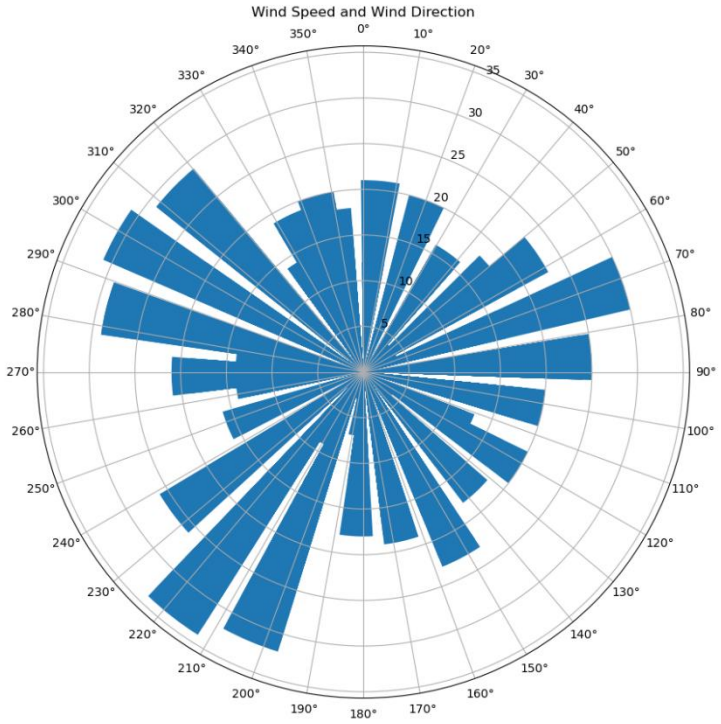


Figure A1.2: *Time series of hourly corrected Unit 1 and hourly reference model testing data.*

a)



b)

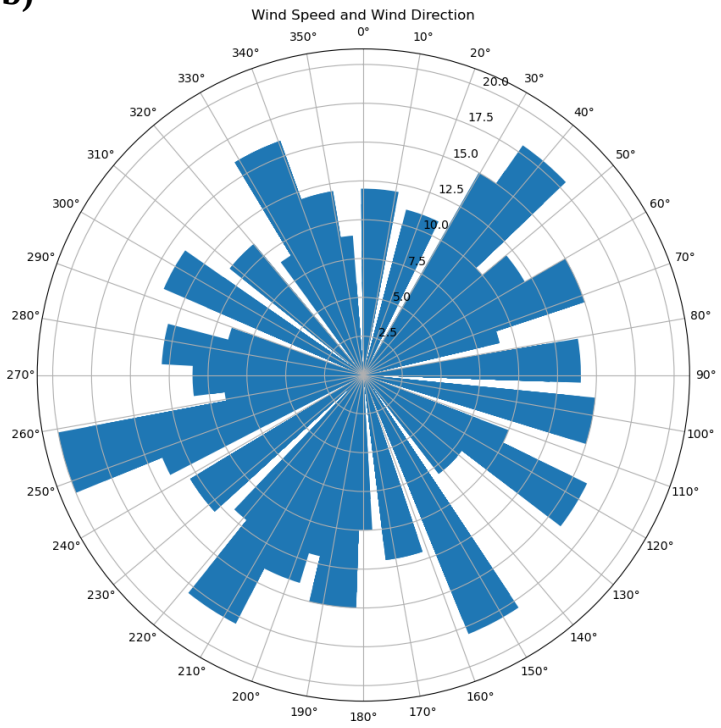


***Figure A1.3**: Wind roses showing the recorded wind speeds and direction at Cork City Airport during a) December 2021 and b) June 2021.*

4 Spatial Analysis of $PM_{2.5}$ Using a Concentration Similarity Index Applied to Air Quality Sensor Networks

Reproduced from the published article:

Byrne, R., Wenger, J. C., & Hellebust, S. (2024). Spatial analysis of $PM_{2.5}$ using a Concentration Similarity Index applied to air quality sensor networks. *Atmospheric Measurement Techniques, 17(17),* 5129–5146.

https://doi.org/10.5194/amt-17-5129-2024

4.1 Abstract

AQS networks are useful for mapping $PM_{2.5}$ in urban environments, but quantitative assessment of the observed spatial and temporal variation is currently under-developed. This study introduces a new metric – the concentration similarity index (CSI) – to facilitate a quantitative and time-averaged comparison of the concentration-time profiles of $PM_{2.5}$ measured by each sensor within an air quality sensor network. Following development on a dataset with minimal unexplained variation and robust tests, the CSI function is used to represent an unbiased and fair depiction of the air quality variation within an area covered by a monitoring network. The measurement data is used to derive a CSI value for every combination of sensor pairs in the network, yielding valuable information on spatial variation in $PM_{2.5}$. This new method is applied to two separate AQS networks, in Dungarvan and in Cork City, Ireland. In Dungarvan there was a lower mean CSI value ($\bar{x}_{CSI,\ Dungarvan} = 0.61$, $\bar{x}_{CSI,\ Cork} = 0.71$), indicating lower overall similarity between locations in the network.

In both networks, the average diurnal plots for each sensor exhibit an evening peak in $PM_{2.5}$ concentration due to emissions from residential solid fuel burning, however, there is considerable variation in the size of this peak. Clustering techniques applied to the CSI matrices identify two different location types in each network; locations in central or residential areas which experience more pollution from solid fuel burning and locations on the edge of the urban areas which experience cleaner air. The difference in mean $PM_{2.5}$ between these two location types was 6 μg m^{-3} in Dungarvan, and 2 μg m^{-3} in Cork. Furthermore, the examination of winter and summer months (January and May) indicates higher $PM_{2.5}$ levels during periods of increased residential solid fuel burning act as a major driver for greater differences (lower similarity indices) between locations in both networks, with differences in mean seasonal CSI values exceeding 0.25 and differences in mean seasonal $PM_{2.5}$ exceeding 7 μg m^{-3}. These findings underscore the importance of including wintertime PM data in analyses as the differences between locations is enhanced during periods when solid fuel burning activities are at a peak. Additionally, the CSI method facilitates the assessment of the representativeness of the $PM_{2.5}$ measured at regulatory air quality monitoring locations with respect to population exposure, showing here that location type is more important than physical proximity in terms of similarity and spatial representativeness assessments. Applying the CSI in this manner can allow for the placement of monitoring infrastructure to be optimised**.** The results indicate that the population exposure to $PM_{2.5}$ in Dungarvan is moderately represented ($\bar{x}_{CSI}$ = 0.63) by the current regulatory monitoring location, and the regulatory monitoring location assessed in Cork represented the city-wide $PM_{2.5}$ levels well ($\bar{x}_{CSI}$ = 0.76).

4.2 Introduction

Air pollution affects the environment, quality of life, and is a major cause of premature death and disease (Cesaroni et al., 2013; Lelieveld et al., 2015; Pedersen et al., 2013; Raaschou-Nielsen et al., 2013b). The category of air pollutant with the largest impact on human mortality and health is fine particulate matter, i.e. atmospheric particles with an aerodynamic diameter of 2.5 micrometres or less ($PM_{2.5}$) (Pope et al., 2020; Pope & Dockery, 2012; Samoli et al., 2013). In many regions around the world, air quality monitoring and management have become critical endeavours to mitigate the detrimental health effects of air pollution, especially $PM_{2.5}$.

Over the years, technological advances have provided valuable tools to enhance our understanding of air pollution, and low-cost AQSs are emerging as promising instruments for collecting real-time air quality data at improved spatial and temporal resolutions (Kumar et al., 2015; Munir et al., 2019). When used in networks, air quality sensors have the potential to enhance and supplement regulatory monitoring and assessment (Malings, et al., 2020a). However, further work needs to be carried out to assess the effectiveness of sensor networks and how to make best use of the data for gaining further insights into air pollution within a locality, because the data quality obtained with such low-cost devices does not meet the standards for regulatory monitoring. Careful consideration must be given to the quality of the data provided by sensors and the requirement for calibration must be assessed (Diez et al., 2022). Recent studies have shown that the performance and calibration of a $PM_{2.5}$ sensor is dependent on the type of sensor and often on the measurement location, suggesting the need for site-specific and individual calibrations to correct for the absolute level of $PM_{2.5}$ (Kaur & Kelly, 2023b; Sayahi et al., 2019; Wang et al., 2015; Zamora et al., 2020). When

these factors are considered and accounted for, AQS networks offer an unprecedented opportunity to gain further insights into the complex dynamics of air pollution in localised areas, such as urban environments, industrial zones, and residential neighbourhoods (Crawford et al., 2021; Frederickson et al., 2022; Gameli Hodoli et al., 2023; Heimann et al., 2015; O'Regan et al., 2022).

Information on the spatial variation of air quality is important because air pollution is not homogenous and can exhibit significant variations across different areas even on a local scale (Frederickson et al., 2022, 2023; Kassomenos et al., 2014; Wang et al., 2018). The variability of air pollution can be influenced by a multitude of factors such as traffic patterns, industrial activities, meteorological conditions, and local topography. Consequently, relying on single monitoring locations or limited data resolution can provide an incomplete picture and inadequate understanding of local air quality in a certain area (Li et al., 2019). Understanding these variations is crucial for targeted interventions and policy decisions aimed at improving air quality and safeguarding public health. Spatial analysis, facilitated by sensor networks allows for a more accurate and nuanced understanding of how air quality, and therefore exposure to pollution, varies across a population centre.

In a recent study, we used data collected by a $PM_{2.5}$ sensor network in Cork City, Ireland, to estimate the contribution of local pollution sources as separate and distinct from regional or transported air pollution (Byrne et al., 2023, see Chapter 3). The results highlighted the very localised nature of $PM_{2.5}$ caused by residential solid fuel burning during winter, which is a notable problem in many towns and cities in Ireland and elsewhere (Dall'Osto et al., 2013a; Kourtchev et al., 2011; Lin et al., 2019b; Lin et al., 2018; Ovadnevaite et al., 2021; Wenger et al., 2020; Zhang et al., 2021).

In this work, we propose a new approach for assessing the spatial profile of air quality using an AQS network. The method yields a time-averaged concentration similarity index (CSI) for quantitative assessment of the similarity between the complete data series produced by different sensors within a network. The CSI is built on the premise that sensors exposed to similar ambient conditions and pollutant sources will produce comparable $PM_{2.5}$ temporal trends. Conversely, sensors subject to different conditions might display divergent $PM_{2.5}$ concentration trends. The motivation for the development of an assessment method based on the temporal variation over an extended period is the realisation that the annual average is often an incomplete representation of true population exposure, which is experienced from hour to hour and day to day. If hourly or daily $PM_{2.5}$ variability is high, it is not always adequate to merely compare annual averages of $PM_{2.5}$ in different locations when comparing $PM_{2.5}$ exposure experienced by local populations. While the annual average and hourly/daily values are often well correlated, numerous studies have found positive associations between short-term exposure to particulate matter and increased morbidity and mortality due to respiratory and cardiovascular diseases (Fajersztajn et al., 2017; Orellano et al., 2020; Weinmayr et al., 2010). This method aims to translate this idea into a quantifiable metric by calculating the time-averaged degree of similarity between two sensor datasets. After development and testing, the CSI analysis is applied to an AQS network in the town of Dungarvan in Ireland to identify areas that may be experiencing persistently elevated or very localised $PM_{2.5}$ pollution compared to others. Clustering techniques are used to group sensors based on the similarity of their $PM_{2.5}$ measurements. The CSI method is also retrospectively applied to the data collected in the Cork City network to investigate the transferability of the method between sensor networks and to explore any differences between the locations.

4.3 Methodology

4.3.1 Data Collection, Preprocessing, and Calibration

The collection, preprocessing, and calibration of the data collected by the $PM_{2.5}$ sensor networks in Dungarvan and Cork City was carried out using the Julia programming language (Bezanson et al., 2017). Since low-cost AQSs are not of regulatory standard, care needs to be taken with quality assessment and quality control of the data. In particular, the degree to which changes or differences in $PM_{2.5}$ measurements between devices can be trusted needs to be considered. The methodology proposed here addresses these inherent issues to deliver an approach for assessing the spatial representativeness of any monitoring location and to facilitate comparison of different environment types, regardless of geographical distance to the location.

4.3.2 Dungarvan $PM_{2.5}$ Sensor Network

The Dungarvan sensor network consisted of 18 solar powered Clarity Node-S devices (Clarity Movement Co., USA) which utilise the Plantower PM6003 sensor to measure $PM_{2.5}$ within the range 1–1000 $\mu g\ m^{-3}$ and at a resolution of 1 $\mu g\ m^{-3}$ (Clarity Movement Co., 2023). By default, the Node-S devices take measurements every 15 minutes, allowing sufficient data upload and battery sleep time in between sampling periods, however this can be adjusted to higher or lower frequencies. A sampling interval of 8 minutes was used, providing efficient battery performance.

The Clarity Node-S devices were typically attached to street light poles approximately 2–4 m above the ground. The sensors were positioned in a range of different environments including urban background, residential, coastal, and roadside locations

(Figure A2.1). Many of these locations were a mix of the different environments. The majority of devices were operational from 01 November 2022 to 31 May 2023, however three devices (AP7, AY9N, AY93) with the Clarity Wind Module were only deployed from 12 January 2023. Measurements were taken over a continuous period covering different meteorological seasons (Winter, Spring, and early Summer), thus ensuring temporal variations in $PM_{2.5}$ concentrations were captured comprehensively.

Prior to and after deployment in Dungarvan, the Clarity Node-S devices were co-located on the roof of the Ellen Hutchins Building, UCC (51.895136, -8.516146) to compare their performance. Details of the three co-location periods are outlined in Table A2.1. Although some devices were not available for all three co-location periods, the three periods combined provide a comparison between the sensors across different seasons. This co-location dataset enabled the CSI method to be developed on measurements that in theory should be equal and the function could then be modified, if necessary, to allow for sensor behaviour/limitations, uncertainties, and errors.

The raw sensor data from the co-location periods and field deployment, underwent a series of preprocessing steps to mitigate potential sources of error in the measurement and ensure data quality and consistency. Data points outside of the operational range of the sensors ($> 1000\ \mu g\ m^{-3}$) were removed, although instances of these were minimal. The eight-minute data were averaged to produce hourly measurements. Missing data points could potentially affect the temporal continuity of the data; however, the data coverage was overall very good for the co-location and measurement campaign periods. On average, the devices had an hourly measurement coverage of 87% for the field measurement campaign, corresponding to an average of 4443 hourly measurements per device for the campaign period.

Assessing the consistency of measurements across the sensor network was paramount. Although the $PM_{2.5}$ readings were very well correlated when the devices were co-located (Table A2.2), a data harmonisation procedure was performed to ensure the uniformity of sensor measurements, a prerequisite for the subsequent development of the CSI. As there was no reference-grade $PM_{2.5}$ data available during the co-location periods, the $PM_{2.5}$ concentrations from each sensor were scaled to a common reference point, represented with the mean of all data points across the whole co-location dataset (Figure A2.2). The data series for each sensor was then individually compared with the calculated mean dataset and subsequently harmonised to the common reference point using simple linear regression. The equations resulting from this harmonisation procedure were applied to the measurements collected from all devices during the subsequent field measurement campaign. While this procedure did not convert the measured $PM_{2.5}$ to reference-equivalent concentrations, it minimised sensor output variability and facilitated a more equitable comparison between sensor measurements (Table A2.2).

4.3.3 Cork City $PM_{2.5}$ Sensor Network

The Cork City sensor network consisted of 16 PurpleAir PA-II-SD units which each contain two Plantower PMS5003 sensors to measure $PM_{2.5}$ within the effective range 0–500 μg m^{-3}, with a maximum range of 1000 μg m^{-3}, and at a resolution of 1 μg m^{-3} (Yong, 2016). In this study, data recorded by the devices in the network for the periods 01 January 2021 to 31 May 2021 and 01 September 2021 to 31 December 2021 were collated and analysed. However, four devices were found to have limited data capture for the specified periods (< 60%) and were therefore omitted from the analysis. The

12 sensors used in this analysis had an average data capture of 85% for the specified periods; their locations are shown in Figure A2.3.

Due to logistical constraints, it was not possible to co-locate all of the PurpleAir devices together to assess variability in $PM_{2.5}$ concentrations. However, low inter-sensor and inter-unit variability was exhibited by four co-located PurpleAir devices in our previous study on the Cork City network, where all inter-sensor and inter-unit comparisons yielded R^2 values greater than 0.98 (Byrne et al., 2023, See Chapter 3). Moreover, PurpleAir $PM_{2.5}$ measurements were highly correlated (R^2 = 0.92) with hourly values of $PM_{2.5}$ concentrations obtained using a Beta-Attenuation Monitor (BAM-1020, Met One Instruments). The comparison yielded a low offset (0.93 $\mu g\ m^{-3}$), although reference measurements tended to be lower than the sensor measurements (slope = 0.58). A co-location dataset was then used to derive calibration factors incorporating the effects of temperature and relative humidity. The data processing procedures for obtaining the $PM_{2.5}$ concentrations reported here are identical to those reported by Byrne et al. (2023) (See Chapter 3).

The Cork City dataset spans a similar measurement period to the Dungarvan dataset to allow for comparable results due to the known seasonality of $PM_{2.5}$ pollution in Ireland (Ovadnevaite et al., 2021). Although the year 2021 included some periods of COVID-19 pandemic restrictions, such measures mainly affected NO_2 concentrations and were not shown to have a significant impact on PM levels in Ireland (EPA, 2020).

4.3.4 Meteorological Measurements

Meteorological data was analysed in each location. For Cork City, data collected at Cork Airport by Ireland's National Meteorological Service, Met Éireann, was

accessed from the website (www.met.ie). The airport weather station is located approximately 5.5 km from Cork City centre.

There is no weather station located nearby Dungarvan that provides hourly measurements, however three of the Clarity Node-S devices were fitted with Clarity Wind Modules (AP7, AY93, AY9N), which provide high time-resolution measurements of wind speed and direction (Clarity Movement Co., USA.). Due to technical difficulties, device AY9N did not capture wind direction measurements, however its wind speed is included. The Wind Module contains a solid-state two-axis ultrasonic anemometer which provides wind speed measurements with a range of 0−60.00 m s^{-1}, and a resolution of 0.01 m s^{-1} along with wind direction at a resolution of 0.1°, over a range of 0–359.9° (Wind Module technical sheet, 2024). These measurements have not been validated against reference meteorological data; however, they are included for indicative purposes.

4.4 Development of the Concentration Similarity Index

The CSI derived here quantifies the degree of likeness between $PM_{2.5}$ concentration profiles from two sensors for a defined period of time and forms the basis for assessing the spatial disparities in $PM_{2.5}$ measurements within sensor networks. The methodology proposed was developed through multiple iterations in order to adjust and improve the procedure. An overview of the development is described, showing the evolution towards the final method.

4.4.1 Original Function Application

The first phase of development was based directly on the work carried out by Piersanti et al. (2015), who used a concentration similarity function to assess the spatial representativeness of $PM_{2.5}$ and O_3 monitoring stations in the Italian air quality monitoring network. Using modelled hourly air pollutant data covering Italy with a 4×4 km^2 grid cell resolution, Piersanti et al. (2015) produced maps showing how representative certain sites in the Italian monitoring infrastructure were. The application proposed here compares point measurement to point measurement as opposed to comparing modelled grid cell data, however the underlying principle of comparing two concentration-time profiles to produce a single indication of similarity between them still applies. The function value $f_{site}(x, y)$ used by Piersanti et al. (2015) to assess the spatial coverage of point measurements is given in Equation 4.1.

$$f_{site}(x,y) = \frac{\sum_{i=1}^{N_t} flag}{N_t}, \quad \textbf{(4.1)}$$

$$where\ flag = \begin{cases} 1, & \frac{|C(X_{site},Y_{site},t_i) - C(x,y,t_i)|}{C(X_{site},Y_{site},t_i)} < 0.2 \\ 0, & \frac{|C(X_{site},Y_{site},t_i) - C(x,y,t_i)|}{C(X_{site},Y_{site},t_i)} > 0.2 \end{cases}$$

where, $C(x, y, t_i)$ represents the surface concentration from the modelled data in a grid point at time t_i, $C(X_{site}, Y_{site}, t_i)$ represents the modelled data of a specific site of interest at time t_i, and N_t is the total number of time steps. The study defined a modelled grid cell at the site of interest as representative of a surrounding grid cell area if the condition $f_{site}(x, y) > 0.9$ is true.

In the first step of our approach, this function was applied to the hourly average $PM_{2.5}$ data obtained from the co-located Clarity Node-S units by comparing two sensor data series at a time. The concentration at the point of interest and surrounding grid cell concentration inputs were substituted for sensor PM concentration values from any given sensor A and sensor B pair, $C(A, t_i)$ and $C(B, t_i)$. Over a total of 1565 co-located hours, the mean number of comparable data points per $C(A, t_i)$, $C(B, t_i)$ pair was 654, due to devices being present at different stages during the co-location periods (Table A2.1). In theory, the function value comparing two sensor data series would be 1, given that the measurements were collected in the same location and were known to all represent the same air parcel at each point in time. However, it was found that the function was not comprehensive enough to allow for an acceptable comparison of the sensor data. The results showed discrepancies between some device pairs because the function value deviated significantly from 1 in many cases (Table 4.1) and was as low as 0.51 in some cases, with an overall mean of 0.82.

Table 4.1: *Function values, fsite(x, y), for hourly averaged $PM_{2.5}$ measured by a range of co-located Clarity Node-S devices. Device labels in the columns were set as C(Xsite, Ysite, t) and device labels in the rows were set at C(x, y, ti).*

	A3	**A4**	**A8H**	**A8Z**	**A9**	**AQ**	**AZ**	**A7**	**A6P**	**AJ3**	**AP7**	**AQV**	**ARF**	**AW6**	**AWF**	**AY9N**	**AY93**	**AYG**
A3	1	0.83	0.88	0.80	0.80	0.86	0.88	0.87	0.87	0.99	0.67	0.90	0.80	0.99	0.80	0.61	0.80	0.98
A4	0.84	1	0.81	0.89	0.83	0.87	0.89	0.87	0.87	0.98	0.80	0.90	0.90	0.97	0.85	0.72	0.79	0.99
A8H	0.86	0.79	1	0.79	0.78	0.87	0.89	0.83	0.78	0.76	0.66	0.85	0.72	0.92	0.71	0.62	0.72	0.73
A8Z	0.82	0.91	0.81	1	0.78	0.88	0.89	0.85	0.84	0.97	0.87	0.88	0.90	0.97	0.86	0.81	0.73	0.98
A9	0.76	0.82	0.80	0.76	1	0.75	0.80	0.78	0.78	0.78	0.69	0.77	0.74	0.67	0.78	0.62	0.76	0.69
AQ	0.88	0.87	0.90	0.87	0.77	1	0.94	0.88	0.86	0.97	0.87	0.97	0.84	1	0.81	0.81	0.75	0.94
AZ	0.89	0.88	0.89	0.88	0.80	0.94	1	0.91	0.90	0.97	0.80	0.97	0.82	0.99	0.82	0.75	0.84	0.93
A7	0.87	0.87	0.86	0.83	0.82	0.88	0.91	1	0.89	0.98	0.68	0.91	0.81	0.99	0.80	0.63	0.74	0.96
A6P	0.87	0.86	0.79	0.82	0.77	0.87	0.92	0.89	1	0.89	0.77	0.90	0.83	0.75	0.82	0.80	0.78	0.89

Table 4.1 continued.

	A3	**A4**	**A8H**	**A8Z**	**A9**	**AQ**	**AZ**	**A7**	**A6P**	**AJ3**	**AP7**	**AQV**	**ARF**	**AW6**	**AWF**	**AY9N**	**AY93**	**AYG**
AJ3	0.99	0.97	0.76	0.97	0.75	0.97	0.98	0.98	0.89	1	0.72	0.88	0.91	0.76	0.82	0.77	0.81	0.89
AP7	0.71	0.85	0.63	0.88	0.65	0.86	0.86	0.74	0.77	0.71	1	0.75	0.81	0.54	0.80	0.83	0.75	0.80
AQV	0.90	0.89	0.86	0.86	0.75	0.96	0.96	0.91	0.88	0.88	0.77	1	0.85	0.77	0.80	0.76	0.82	0.85
ARF	0.82	0.90	0.72	0.91	0.74	0.86	0.85	0.83	0.85	0.90	0.80	0.84	1	0.74	0.82	0.81	0.81	0.92
AW6	0.96	0.94	0.92	0.94	0.65	1	0.99	0.98	0.72	0.72	0.50	0.73	0.70	1	0.69	0.51	0.49	0.70
AWF	0.80	0.88	0.70	0.86	0.76	0.82	0.82	0.81	0.82	0.82	0.81	0.80	0.83	0.73	1	0.81	0.77	0.86
AY9N	0.62	0.76	0.59	0.81	0.60	0.79	0.75	0.65	0.78	0.77	0.81	0.74	0.80	0.55	0.80	1	0.72	0.80
AY93	0.76	0.76	0.67	0.65	0.72	0.69	0.74	0.65	0.78	0.84	0.75	0.80	0.81	0.55	0.77	0.75	1	0.82
AYG	0.99	0.99	0.72	0.99	0.67	0.95	0.95	0.98	0.87	0.87	0.78	0.83	0.90	0.75	0.84	0.79	0.76	1

4.4.2 Function Parameter Optimisation and Introduction of PM Limit

Analysis of the results obtained from direct application of the original function showed that the conditions set out by it were too strict to apply to the sensor data given the variations that can occur in AQS measurements. The areas of the entire sensor networks discussed here could be within the original single grid cell size analysed by Piersanti et al. (2015). Therefore, overall pollution dynamics would vary significantly, in part because of hyper-local effects, and pollution averaging effects would be more pronounced when assessing larger areas. Moreover, the high hourly $PM_{2.5}$ variation and very localised effects exhibited in a typical Irish winter $PM_{2.5}$ profile is not suited to the original function (Byrne et al., 2023, see Chapter 3). While the original application contains a mathematical function examining the difference between two pollutant concentrations and is independent of specifications regarding area size and pollution dynamics, the threshold values can be adapted to reflect the specific application of the function. A second threshold value, a PM mass concentration limit, PM_{lim}, was introduced to the function, with different relative concentration limits for the upper and lower PM values, $C_{lim,\ upper}$ and $C_{lim,\ lower}$, respectively. Treating larger and smaller $PM_{2.5}$ values differently when assessing the similarity between two data series is useful in capturing the nuanced relationships and patterns in the data. It allows for the real-world significance of the data to be reflected, acknowledging the varying implications of $PM_{2.5}$ measurements based on the magnitude. Higher $PM_{2.5}$ values can indicate a pollution episode or specific local pollution sources, while lower values can represent background levels. Therefore, treating lower $PM_{2.5}$ values with more leniency in the similarity assessment recognises that minor fluctuations in low hourly concentrations might not be as concerning as similar deviations in higher

concentrations and the health-related considerations associated with these high concentrations.

Another potential advantage of the PM limit concerns the varying degrees of accuracy of the AQS measurements. Allowing the leeway introduced here in assessing the similarity of lesser measurement values considers potential measurement uncertainties with these devices. However, it is important to note that this approach is not accommodating sensor limitations at the expense of accuracy but rather it is a strategy to ensure that the assessment remains faithful to the underlying air quality dynamics while accounting for the potential deficiencies in measurement equipment.

When the function is applied to a pair of sensors, the resulting CSI can differ slightly depending on which sensor was classified as $C(x, y, t_i)$ or $C(X_{site}, Y_{site}, t_i)$, or sensor A or sensor B, in Equation. 4.1 when computing the difference at each timestep. Due to the nature of the function, the denominator value of the relative difference calculation, the concentration of sensor A at a given timestep, is what makes the difference. To counteract this and to avoid the possibility of large discrepancies between the CSI values for a sensor pair depending on which sensor is taken as A or B, the function was modified to have the geometric mean, or the square root of the product, of $C(A, t_i)$ and $C(B, t_i)$ used as the denominator. This ensured symmetry in the function so that the CSI values were identical regardless of which sensor was classified as A or B in a sensor pair.

Equation 4.2 shows the next form of the concentration similarity function (function notation has been modified to be more suitable for this application).

$$CSI_{A,B} = \frac{\sum_{i=1}^{N_t} f}{N_t}, where\ f = \tag{4.2}$$

$$\begin{cases} 1\ if\ \dfrac{|C(B,t_i) - C(A,t_i)|}{\sqrt{C(A,t_i) \times C(B,t_i)}} < C_{lim,upper}\ and\ C(A,t_i)\ or\ C(B,t_i) > PM_{lim} \\ 0\ if\ \dfrac{|C(B,t_i) - C(A,t_i)|}{\sqrt{C(A,t_i) \times C(B,t_i)}} > C_{lim,upper}\ and\ C(A,t_i)\ or\ C(B,t_i) > PM_{lim} \\ 1\ if\ \dfrac{|C(B,t_i) - C(A,t_i)|}{\sqrt{C(A,t_i) \times C(B,t_i)}} < C_{lim,lower}\ and\ C(A,t_i)\ or\ C(B,t_i) < PM_{lim} \\ 0\ if\ \dfrac{|C(B,t_i) - C(A,t_i)|}{\sqrt{C(A,t_i) \times C(B,t_i)}} > C_{lim,lower}\ and\ C(A,t_i)\ or\ C(B,t_i) < PM_{lim} \end{cases}$$

where $C(A, t_i)$ and $C(B, t_i)$ are the $PM_{2.5}$ measurements from devices *A* and *B* at time, t_i. $C_{lim,\ lower}$ and $C_{lim,\ upper}$ are the threshold values defining the acceptable level of difference between two concentrations, and PM_{lim} is the PM mass concentration threshold value. In order to retain the operation of the function if either $C(A, t_i)$ or $C(B, t_i)$ equals exactly 0 μg m^{-3}, the value is changed to 0.1 μg m^{-3} to carry out the calculations. Additionally, if both concentrations are equal, *f* is automatically assigned as 1 to represent similar concentrations.

4.4.3 Development and Testing of the Modified Equation

The PM limit and associated concentration similarity limits introduced were chosen by iteratively testing the similarity function on the co-location data using different limits. Each co-located sensor pair was tested with different PM_{lim} values (5 μg m^{-3}, 10 μg m^{-3}, 15 μg m^{-3}, 20 μg m^{-3}) and with C_{lim} values ranging from 0.1 to 2.0 in steps of 0.1 for both the upper and lower limits. This produced a C_{lim} vs CSI comparison for each A-B pair for data above and below the corresponding PM_{lim} value. The C_{lim} value for each sensor comparison that gave a minimum CSI value of 0.95 was recorded with the overall mean of these C_{lim} values above and below each PM_{lim} value taken forward.

The mean C_{lim} pair values were then applied to the co-location measurements with the respective PM_{lim} values to give final CSI values for each sensor pair, highlighting how the $PM_{2.5}$ concentration profile of each sensor compares to that of all the other sensors. The highest mean CSI value for all co-located A-B pairs was found for PM_{lim} = 15 μg m^{-3}, $C_{lim,\ upper}$ = 0.2, and $C_{lim,\ lower}$ = 0.7. When applying these new limits, all sensor pairs gave a CSI above 0.85, with 99% of pairs above 0.90, and an overall CSI mean of 0.98. These final limits enabled a good comparison for the hourly co-located AQS measurements (Table 4.2).

The CSI function was also applied to data obtained from the four co-located PurpleAir devices in order to make sure that the function was applicable across the two AQS types. The data was harmonised by following the same procedure as the Clarity data, through scaling each data from each sensor to the mean data series of all four sensors. Although this co-location period was shorter than that of the Clarity dataset used for the function development, it still allowed for the CSI to be calculated from around 250 common data points per sensor pair. All device pairs reported a CSI close to 1.0, with a mean CSI of 0.99 (Table A2.3).

Table 4.2: *Concentration similarity indices for hourly averaged $PM_{2.5}$ measured by a range of co-located Clarity Node-S devices. PM_{lim} = 15 μg m^{-3}, $C_{lim,\ upper}$ = 0.2, $C_{lim,\ lower}$ = 0.7.*

	A3	**A4**	**A8H**	**A8Z**	**A9**	**AQ**	**AZ**	**A7**	**A6P**	**AJ3**	**AP7**	**AQV**	**ARF**	**AW6**	**AWF**	**AY9N**	**AY93**	**AYG**
A3	1	0.97	0.99	0.96	0.92	0.99	1	0.99	0.99	1	0.96	1	0.97	1	0.96	0.94	0.99	1
A4		1	0.97	0.99	0.95	0.98	0.99	0.99	0.99	1	0.99	0.99	0.99	1	0.98	0.97	0.97	1
A8H			1	0.96	0.92	0.99	1	0.99	0.99	0.97	0.94	0.99	0.96	0.98	0.96	0.94	0.97	0.98
A8Z				1	0.94	0.97	0.99	0.98	0.98	1	0.98	0.98	1	1	0.98	0.96	0.97	1
A9					1	0.92	0.92	0.94	0.95	0.97	0.96	0.95	0.97	0.92	0.97	0.95	0.96	0.96
AQ						1	1	1	0.99	1	0.97	1	0.97	1	0.96	0.97	1	1
AZ							1	0.99	0.99	1	0.97	1	0.98	1	0.96	0.95	1	1
A7								1	0.99	1	0.95	0.99	0.98	1	0.96	0.94	0.97	1
A6P									1	0.99	0.97	0.99	0.99	0.96	0.97	0.97	0.99	1

Table 4.2 continued.

	A3	**A4**	**A8H**	**A8Z**	**A9**	**AQ**	**AZ**	**A7**	**A6P**	**AJ3**	**AP7**	**AQV**	**ARF**	**AW6**	**AWF**	**AY9N**	**AY93**	**AYG**
AJ3										1	0.98	0.98	0.99	0.95	0.99	0.98	1	0.99
AP7											1	0.96	0.99	0.85	0.99	0.99	0.97	0.98
AQV												1	0.98	0.98	0.97	0.96	0.99	0.99
ARF													1	0.96	0.98	0.99	0.99	1
AW6														1	0.92	0.87	0.9	0.97
AWF															1	0.98	0.97	0.98
AY9N																1	0.97	0.99
AY93																	1	0.99

The function described in Equation 4.2 was further tested by comparing the sensors to numerous sets of synthetic data created from each sensor's measurements to assess the impact of a range of scenarios. Comparing a sensor dataset to itself establishes a baseline for the comparison where the CSI is 1 and any subsequent adjustments to the data to create the synthetic data can be explored, resulting in a new CSI. The first scenario investigated changes in CSI when outliers are present in the data. To explore this, the sensor data was changed so a certain number of data points could be considered outliers (n = 1, 10, 500, 1000). An outlier data point was created by increasing a value by 100 μg m^{-3} in order to ensure discrepancy between it and the original value. The function was then tested in a scenario where the data was scaled linearly so the mean remained constant, but the variance of the data was increased, and it was also tested in a scenario where the entire data was offset by 5 μg m^{-3}, 10 μg m^{-3}, 15 μg m^{-3}, and 20 μg m^{-3}. The final test scenario involved the introduction of noise to the dataset, representing impactful variations in that data. Gaussian noise with various values of standard deviation was added to the data. The CSI results for the synthetic data tests were also compared to the results when the R^2 was found between any two given datasets. Low variations were found during all synthetic data analyses with the resulting CSI values having standard deviations ≤ 0.05 across the individual devices for each test. As an example, the effects of these tests on the CSI results for A6P are shown in Table 4.3, where 4406 data points were included in the calculations.

It is clear that in the case of the linearly scaled data with higher variability but the same overall mean, the CSI is impacted (CSI = 0.52) when the standard deviation is increased by just a factor of 1.5, indicating that such a dataset is dissimilar to the original. In comparison, the R^2 is not an accurate reflection of the change, as it does

not deviate from 1. Offsetting the data by different degrees also shows a major change in the CSI (CSI < 0.55). However, this is not reflected well in the R^2 values, which do not deviate from 1. The CSI method is quite robust with respect to outliers, whereas the R^2 is more sensitive (R^2 = 0.94) when 10 outliers are introduced to the dataset, which is approximately 0.2% of the total data points. The R^2 is significantly reduced (R^2 = 0.45) with 500 outliers (~11% of the total data points), whereas the CSI is only slightly impacted (CSI = 0.89). As the method yields a time-averaged result, low numbers of outliers do not hugely affect the index for a given sensor pair. So, two datasets that are generally similar, but where one experiences some outliers, will be deemed similar by the method. The R^2 also shows a more limited response when larger amounts of Gaussian noise are added, resulting in a value of 0.96 when the standard deviation of the noise is 4 $\mu g\ m^{-3}$, while the CSI is adjusted to 0.7. From a health-impact and exposure point of view, increased variation and higher offset represent very different exposure scenarios, whereas a large difference in the occasional hourly average in an otherwise similar exposure regime does not. The CSI offers a more appropriate comparison between hourly measurements collected at two locations.

Table 4.3: *Influence of data outliers and other factors on CSI determined in test scenarios with device A6P.*

Number of outliers	**CSI**	**R^2**	**Standard deviation factor increase**	**CSI**	**R^2**	**$PM_{2.5}$ positive offset** $\mu g\ m^{-3}$	**CSI**	**R^2**	**$PM_{2.5}$ negative offset** $\mu g\ m^{-3}$	**CSI**	**R^2**	**Standard deviation of added noise** $\mu g\ m^{-3}$	**CSI**	**R^2**
0	1	1	0	1	1	0	1	1	0	1	1	0	1	1
1	1	1	1.5	0.52	1	5	0.54	1	5	0.34	1	1	0.96	1
10	1	0.94	2	0.29	1	10	0.05	1	10	0.02	1	4	0.70	0.96
500	0.89	0.45	4	0.10	1	15	0.01	1	15	0.01	1	10	0.36	0.77
1000	0.77	0.38				20	0.01	1	20	0.004	1			

4.4.4 Application to Sensor Networks and Analysis of Spatial Trends

The CSI methodology developed above was subsequently applied to the Dungarvan and Cork City sensor networks to evaluate the similarity and spatial variations in $PM_{2.5}$. A systematic pairwise comparison approach was employed, wherein each sensor was individually compared to every other sensor within the network. Hierarchical clustering and fuzzy *c*-means (FCM) clustering were both performed on the CSI results, to identify groupings based on each sensors' relationship to other sensors in the network which can then be reflected spatially.

With both clustering techniques, the quality of cluster assignments can be assessed with various evaluation metrics to choose the optimal number of clusters. As the "true" cluster classifications are not known here, validation must be performed using the clustering algorithm itself. To assess the quality of the hierarchical clustering assignments and an optimal cluster number, the silhouette metric was used along with the Calinski-Harabasz index to assess the FCM assignments (Caliñski & Harabasz, 1974; Rousseeuw, 1987). The silhouette score, ranging from −1 to +1, can be calculated for each member of a cluster and then the mean silhouette score from all members indicates an overall assignment quality for members of that cluster, with a high score closer to 1 indicating higher quality clusters, and a low or negative score indicating poorer cluster assignments. The Calinski-Harabasz index also quantifies the quality of cluster assignments with higher scores indicating better quality.

4.5 Results and Discussion

4.5.1 Dungarvan $PM_{2.5}$ Sensor Network

Analysis of the harmonised data obtained from the sensors in the Dungarvan $PM_{2.5}$ network was conducted to determine CSI values and assess the spatial variation of air pollution across the town. Although the $PM_{2.5}$ concentrations are not as accurate as those collected by reference instrumentation, any relative differences between the sensors and between individual sensor data trends can be regarded as genuine due to the low inter-sensor variation observed after data harmonisation procedures, where the standard deviation of the mean $PM_{2.5}$ co-located measurements was 1.7 μg m^{-3}.

The temporal and spatial trends of $PM_{2.5}$ across the Dungarvan sensor network are reflected in the average diurnal plots obtained for each sensor, Figure 4.1. These diurnal profiles all show large evening peaks in $PM_{2.5}$, which are typical for towns and cities in Ireland affected by residential solid fuel burning during winter evenings (Dall'Osto et al., 2014; Healy et al., 2010; Wenger et al., 2020). However, there are clear disparities in some of the average evening peak values between the sensors. One group of sensors has maximum values above 35 μg m^{-3} (A3, A4, A8H, A9, AQ, A7, AW6, AQV), while the sensors with maxima below 35 μg m^{-3} can be further divided into three smaller groups. Sensors labelled AJ3, AWF, and AZ all have a maximum diurnal concentration around 30 μg m^{-3}; sensors AY9N, AY93, ARF, A8Z, AYG, and A6P all have maxima in the 20–26 μg m^{-3} range; sensor AP7 has a significantly lower evening peak compared to the other devices.

Most sensors exhibited the diurnal maximum around the same time of day, between 18:00 and 20:00, however AP7 and ARF, showed a slightly delayed peak from 20:00 to 22:00. AP7 had the lowest peak concentration and did not exhibit the sharp rise and

subsequent decrease associated with evening solid fuel burning that the other sensors showed. AP7 was located on the south-western edge of the town, and since the predominant wind direction is south westerly, did not measure as much local pollution as other locations in the eastern part of the network.

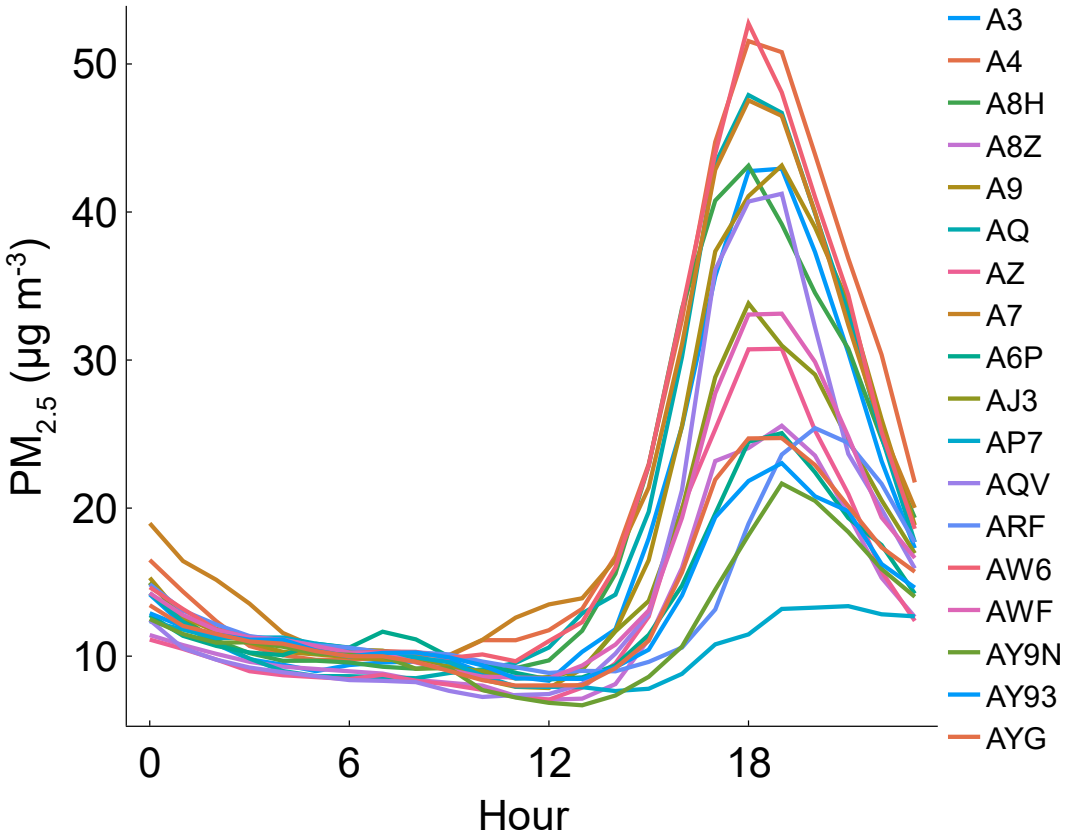


Figure 4.1: *Diurnal profiles for hourly averaged measurements of $PM_{2.5}$ in the Dungarvan sensor network (September 2022 to May 2023). The device AP7 reported the lowest diurnal maximum $PM_{2.5}$ mass concentration.*

Summary statistics obtained for the 18 sensors in the Dungarvan network are listed in Table 4.4. Unsurprisingly, most of the devices with diurnal maxima above 35 µg m^{-3} have the highest mean, median, and maximum values. Out of this subset of devices, AQV has the lowest overall mean (15 µg m^{-3}), but still has a relatively high standard deviation (22 µg m^{-3}), indicating the $PM_{2.5}$ values tend to vary widely but are lower on

average. This could be indicative of fluctuating particle concentrations, consistent with intermittent pollution sources such as residential solid fuel burning.

The wind speed and direction recorded at sites AP7 and AY93 showed some variation (Figure A2.4), however wind speeds measured at all three sites showed a moderate correlation with all R^2 values above 0.65. The measured wind direction at the AP7 and AY93 sites reported a moderate correlation (R^2 = 0.63). Both sites measured winds emanating from a broad range of directions. Both locations reported, generally southerly winds 53% of the time, and south westerly winds 30% of the time. The temporal variations of wind speed measured at the three sites are detailed in Figure A2.5. Little diurnal variation is seen between devices AY93 and AY9N, however it is clear that AP7 tended to experience slightly lower wind speeds than AY93 and AY9N during the measurement campaign. Nevertheless, this difference did not exceed 1 m s^{-1} in any of the temporal variation assessments and all three sites reported the same overall trends in wind speed. The variations in wind measurements between the sites indicate some slight local meteorological differences; however, the overall meteorological field is not likely to differ greatly between the three sites.

Table 4.4: *Summary statistics of hourly average $PM_{2.5}$ concentrations obtained for all sensors in the Dungarvan sensor network (September 2022 to May 2023).*

Device ID	**Mean (σ)**	**Median**	**Maximum hourly value**	**Maximum diurnal value**	**Hour of maximum diurnal value**
	µg m^{-3}	**µg m^{-3}**	**µg m^{-3}**	**µg m^{-3}**	
AP7	11 (12)	7	153	12	21
AY9N	12 (14)	7	136	23	19
A8Z	13 (16)	7	274	25	19
A6P	13 (18)	8	311	26	19
AY93	13 (15)	8	176	22	19
AZ	14 (18)	7	286	30	19
ARF	14 (18)	8	243	26	20
AYG	14 (16)	8	259	24	19
AQV	15 (22)	8	281	44	19
AJ3	16 (19)	9	270	33	18
AWF	16 (17)	9	189	31	19
A3	18 (27)	9	412	45	19
A9	18 (27)	9	409	45	19
A8H	19 (28)	9	482	40	18
AQ	19 (28)	10	480	48	18
A4	21 (27)	12	370	52	18
A7	21 (27)	12	361	45	18
AW6	21 (26)	11	319	51	18

4.5.1.1 Concentration Similarity Index

The matrix of CSI values obtained for the Dungarvan sensor network is shown in Table 4.5. The results can be analysed in a number of ways. Firstly, the indices for one sensor can be used to assess how similar or dissimilar the measurements are to all other sensors in the network, thus providing information on the spatial representativeness of that particular location. Secondly, the indices of all sensors can be looked at together to elucidate any potential relationship between sensor measurement locations.

The minimum CSI value (0.85) determined during the co-location deployment can act as the lower limit for when two sensor locations can be considered very similar. The reported CSI values for Dungarvan sensors ranged from 0.48 (ARF vs A7) to 0.79 (AYG vs AWF) with a mean of 0.61, indicating a significant difference in air quality representation between locations across the town. The device with the lowest mean of its CSI values with respect to the other locations was A4 (0.55), and although device ARF was only slightly above this (0.57), it reported a larger range of CSI values, including the lowest of the entire dataset. AJ3, AQV, and AYG all shared the highest mean CSI values (0.66).

Table 4.5: *Concentration similarity indices for the hourly averaged PM2.5 concentrations measured by Clarity Node-S devices in the Dungarvan sensor network.*

	A3	**A4**	**A8H**	**A8Z**	**A9**	**AQ**	**AZ**	**A7**	**A6P**	**AJ3**	**AP7**	**AQV**	**ARF**	**AW6**	**AWF**	**AY9N**	**AY93**	**AYG**
A3	1	0.56	0.64	0.56	0.58	0.67	0.58	0.62	0.64	0.59	0.53	0.61	0.50	0.63	0.55	0.59	0.60	0.57
A4		1	0.53	0.52	0.55	0.53	0.56	0.58	0.53	0.57	0.55	0.58	0.49	0.54	0.55	0.56	0.55	0.55
A8H			1	0.58	0.60	0.61	0.59	0.57	0.61	0.61	0.57	0.64	0.54	0.61	0.56	0.61	0.62	0.61
A8Z				1	0.62	0.55	0.69	0.54	0.62	0.65	0.71	0.66	0.60	0.56	0.62	0.66	0.62	0.67
A9					1	0.53	0.60	0.57	0.58	0.63	0.60	0.63	0.52	0.58	0.60	0.61	0.58	0.60
AQ						1	0.58	0.57	0.60	0.59	0.50	0.62	0.49	0.63	0.53	0.56	0.58	0.56
AZ							1	0.58	0.63	0.68	0.68	0.71	0.60	0.62	0.66	0.62	0.61	0.74
A7								1	0.57	0.58	0.53	0.60	0.48	0.62	0.55	0.56	0.57	0.55
A6P									1	0.72	0.62	0.66	0.58	0.63	0.64	0.65	0.68	0.71
AJ3										1	0.64	0.76	0.60	0.64	0.76	0.67	0.70	0.78

Table 4.5 continued.

	A3	**A4**	**A8H**	**A8Z**	**A9**	**AQ**	**AZ**	**A7**	**A6P**	**AJ3**	**AP7**	**AQV**	**ARF**	**AW6**	**AWF**	**AY9N**	**AY93**	**AYG**
AP7											1	0.64	0.67	0.52	0.64	0.65	0.63	0.69
AQV												1	0.59	0.65	0.72	0.65	0.68	0.77
ARF													1	0.56	0.59	0.62	0.61	0.63
AW6														1	0.61	0.60	0.62	0.62
AWF															1	0.66	0.68	0.79
AY9N																1	0.59	0.67
AY93																	1	0.72
AYG																		1

To further investigate the effect of solid fuel burning on local air quality, the CSI function was applied to data from two isolated months – January and May 2023. The purpose of this assessment was to evaluate the extent to which residential solid fuel burning dictates the CSI between two sensors, given that one month (January) will have higher $PM_{2.5}$ levels with measurements heavily influenced by solid fuel burning, and the other will not (May). For both months, all sensors had data capture above 65% and the mean capture was 94% for January and 92% for May. The January mean CSI from all comparisons was 0.51, and the May mean CSI was 0.84 (Table A2.4, Table A2.5). The large discrepancy between the mean CSI for January and May is most likely due to the higher variation typically seen in wintertime $PM_{2.5}$ due to residential solid fuel burning ($\sigma_{January}$ = 25 μg m^{-3}, σ_{May} = 9 μg m^{-3}) (Figure 4.1). This highlights the importance of seasonality when assessing the spatial representativeness of monitoring network locations.

4.5.1.2 Clustering

Clustering techniques were employed on the CSI matrix to uncover any inherent spatial relationships between different locations in the network. Hierarchical clustering produced a dendrogram showing the hierarchical relationship between the sensor locations and was used to identify clusters (Figure 4.2). The highest mean silhouette score was found with two clusters (Figure A2.6). However, it was not a high silhouette score (0.19), indicating that the quality of the cluster assignments was low. The highest Calinski-Harabasz index corresponded to the assignment of members to two clusters when applying the FCM clustering (Figure A2.7).

From both the dendrogram (Figure 4.2) and the FCM membership weights (Figure 4.3), it is clear that devices A4 through to AQ are grouped together in one cluster (Cluster 1), and devices AQV to AP7 are grouped in another cluster (Cluster 2). This split is very similar to the easily visualised groupings shown in the diurnal profile maxima (Figure 4.1), with the only difference being device AQV. The devices in Cluster 1 are also those with the highest mean $PM_{2.5}$ for the measurement period. The mean CSI for each sensor mostly corresponds to the cluster assignments, with Cluster 1 devices having a mean CSI equal to or below 0.6, and all devices in Cluster 2 have a mean CSI above 0.6, except for device ARF. Interestingly, this grouping also appears to have spatial importance too, as shown in Figure 4.4. Cluster 2 devices are mainly located around the edge of the town and generally experience cleaner air ($\bar{x}_{PM2.5}$ = 13 μg m^{-3}, $s_{PM2.5}$ = 27 μg m^{-3}), while Cluster 1 devices are located in central and residential areas ($\bar{x}_{PM2.5}$ = 19 μg m^{-3}, $s_{PM2.5}$ = 17 μg m^{-3}), which are more polluted during winter months.

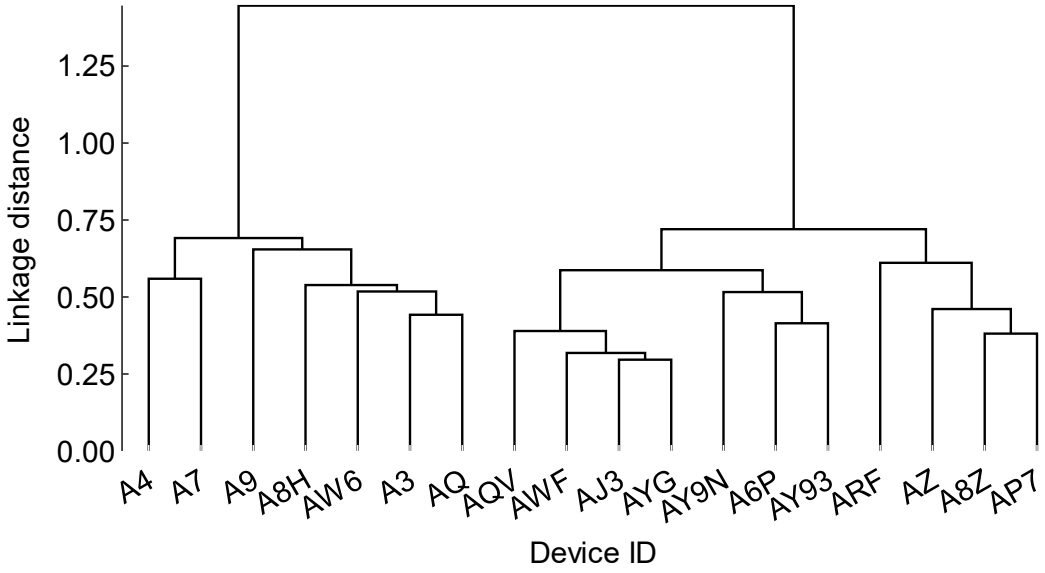


Figure 4.2: *Dendrogram output from hierarchical clustering of the CSI data from the Dungarvan sensor network.*

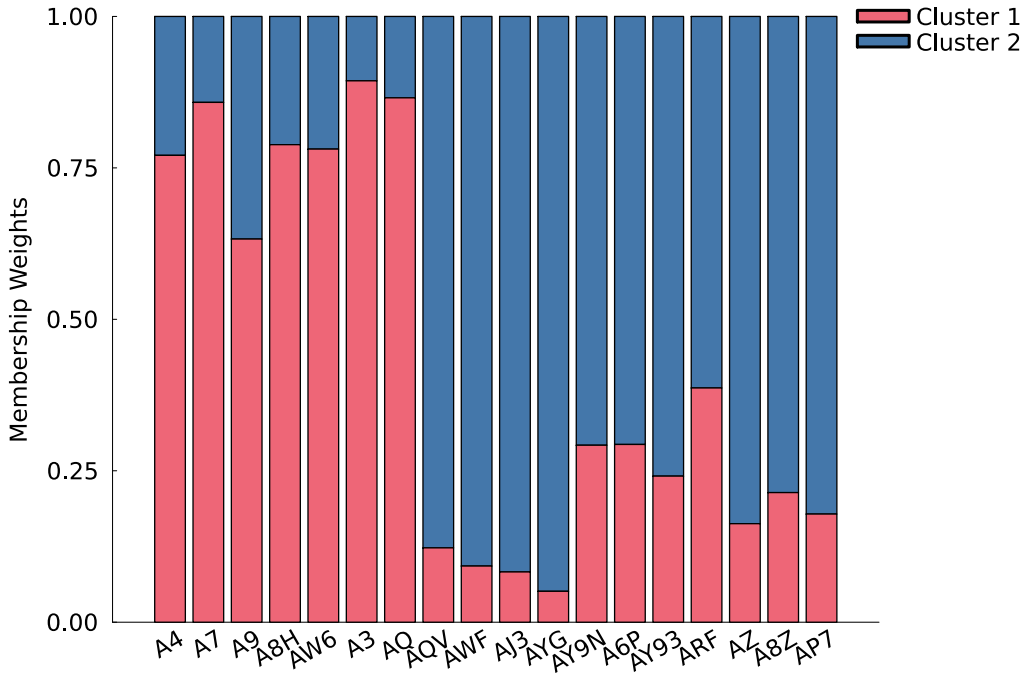


Figure 4.3: *Membership weights from FCM clustering of the CSI data from the Dungarvan sensor network.*



Figure 4.4: *Dungarvan AQS locations with two cluster groups indicated. Cluster 1 devices (red triangle markers) are mainly located in central and residential areas, while cluster 2 devices (blue cross markers) are mainly located on the edge of the town.*

4.5.2 Cork City $PM_{2.5}$ Sensor Network

The same approach as above was used to analyse the data collected by the Cork City AQS network. In this case, the corrected measurements are indicative of the actual $PM_{2.5}$ experienced in each location. The diurnal plots for each sensor in the Cork City network are similar to those observed in Dungarvan, with a sizeable evening peak in $PM_{2.5}$ concentrations (19:00–21:00) due to emissions from residential solid fuel burning. Again, there is considerable variation in the peak concentration of $PM_{2.5}$ (Figure 4.5). Device MTU showed the lowest diurnal average maximum of 9 μg m^{-3}. This device is located on the western side of the city and has few upwind pollution sources contributing to air pollution at the location as the prevailing wind direction is from the South-West. Devices CCC12 and CCC9 both showed the highest diurnal average maximum, 17 μg m^{-3}. CCC12 is located northeast of the city, and so likely experiences urban $PM_{2.5}$ sources up-wind from it or has strong localised sources. Similarly, CCC9 is located to the east of the city, in a residential area. Table 4.6 contains summary statistics for each of the sensors in the Cork City network. Some devices had very high $PM_{2.5}$ maxima, e.g. 201 μg m^{-3} for CCC11, which were more than double the maxima of other devices, e.g., CCC8 which had the lowest overall maximum of 47 μg m^{-3}. Device MTU had the lowest diurnal maximum value, indicating that this location is the least affected by local emissions from solid fuel burning. However, it measured a significant overall $PM_{2.5}$ maximum of 99 μg m^{-3} and significant spikes in pollution were occasionally observed, likely due to meteorological conditions or specific localised effects. When looking at all of the parameters listed in Table 4.6, CCC11 stands out. This sensor has the highest maximum hourly average $PM_{2.5}$ concentration in the network, but the standard deviation (8 μg m^{-3}) is in the middle of the range, indicating that the location had

relatively stable $PM_{2.5}$ levels throughout the measurement period with less variation than other devices but was still susceptible to occasional spikes in $PM_{2.5}$.

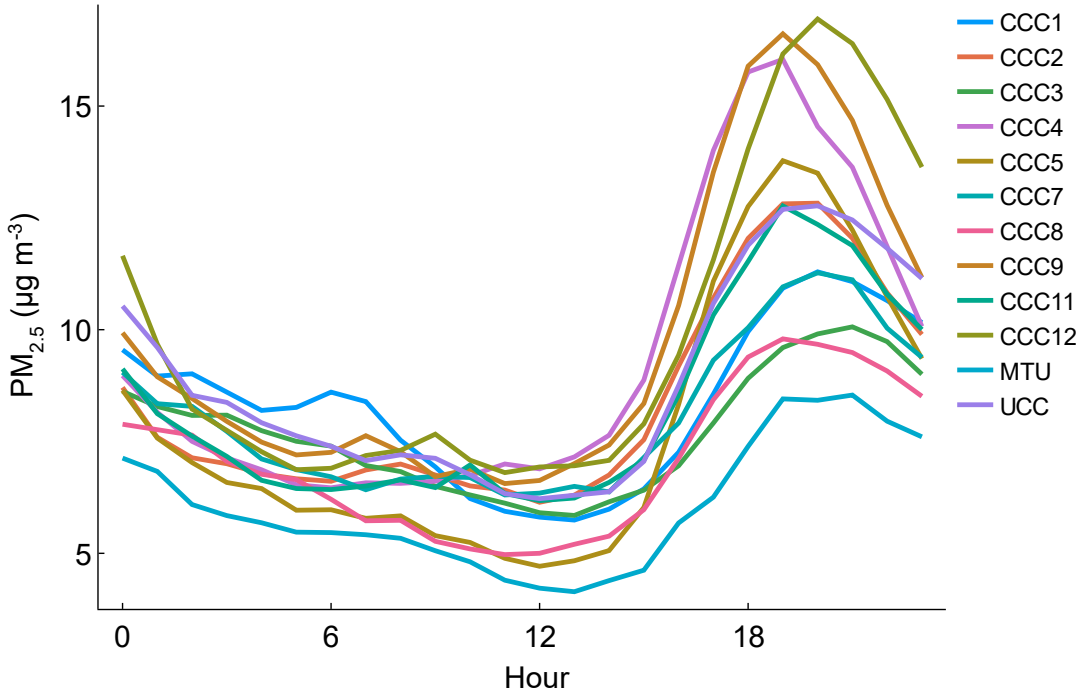


Figure 4.5: *Diurnal $PM_{2.5}$ profiles for all AQSs in the Cork City network (January to May 2021 and September to December 2021).*

The meteorological data retrieved from Cork Airport, which is approximately 4−11 km from each device in the Cork City sensor network, was investigated for the measurement period in 2021. While data obtained from the airport site indicates the meteorological conditions on a synoptic scale, the local weather experienced at individual locations within the city are additionally shaped by factors such as street canyon effects and local topography. Consequently, the wind direction measured at the airport site cannot be assumed to mirror that of all devices in the network. Wind speeds measured at the airport generally surpass those within the city as it is situated at a higher elevation than the city. However, the broader regional wind patterns are

expected to exert a predominant influence on the overall meteorological conditions across the city and therefore the relationship with meteorological conditions and local $PM_{2.5}$ levels can be investigated. The Cork Airport site recorded southerly winds 59% of the time, and south westerly winds 39% of the time (Figure A2.8).

Table 4.6: *Summary statistics of hourly averaged $PM_{2.5}$ obtained for all sensors in the Cork City network (January to May and September to December 2021).*

Device ID	**Mean (*s*)**	**Median**	**Maximum**	**Maximum diurnal value**	**Hour of maximum diurnal value**
	µg m^{-3}	**µg m^{-3}**	**µg m^{-3}**	**µg m^{-3}**	
MTU	6 (6)	4	99	9	21
CCC8	7 (6)	5	47	10	19
CCC3	8 (7)	5	61	10	21
CCC5	8 (10)	5	181	14	19
CCC	8 (7)	6	71	11	20
CCC11	8 (8)	6	201	13	19
CCC1	8 (8)	6	92	11	20
CCC2	8 (8)	6	122	13	20
UCC	9 (8)	6	108	13	20
CCC4	9 (8)	7	97	16	19
CCC9	10 (10)	7	158	17	19
CCC12	10 (10)	7	117	17	20

4.5.2.1 Concentration Similarity Index

The matrix of CSI values obtained for the Cork City sensor network is shown in Table 4.7. The values range from 0.52 (CCC12 vs MTU and CCC9 vs MTU) to 0.85 (CCC2 vs CCC11) with a mean of 0.71. The high maximum CSI indicates a high degree of similarity between those locations in the network, and overall, the Cork City locations show a higher degree of similarity compared to those in Dungarvan.

The isolated CSI results for the months of January and May 2021 were also assessed for Cork City. The average data coverage during both periods was 92%. The mean CSI value in January (0.55) was considerably lower than that observed in May (0.82), Table A2.6, Table A2.7. This result is similar to that found for the Dungarvan network, again indicating that the large difference in mean scores between the two months can be attributed to higher wintertime $PM_{2.5}$ variation by residential solid fuel burning ($\sigma_{January}$ = 15 μg m^{-3}, σ_{May} = 3 μg m^{-3}).

Table 4.7: *Concentration similarity indices for the hourly averaged $PM_{2.5}$ concentrations measured by PurpleAir devices in the Cork City AQS network.*

	CCC1	**CCC2**	**CCC3**	**CCC4**	**CCC5**	**CCC7**	**CCC8**	**CCC9**	**CCC11**	**CCC12**	**MTU**	**UCC**
CCC1	1	0.73	0.76	0.68	0.65	0.71	0.66	0.64	0.76	0.67	0.66	0.76
CCC2		1	0.79	0.82	0.65	0.77	0.68	0.73	0.85	0.78	0.61	0.79
CCC3			1	0.73	0.76	0.80	0.80	0.65	0.82	0.70	0.76	0.80
CCC4				1	0.63	0.73	0.64	0.76	0.82	0.78	0.56	0.77
CCC5					1	0.65	0.74	0.70	0.66	0.60	0.69	0.71
CCC7						1	0.68	0.65	0.78	0.70	0.66	0.73
CCC8							1	0.70	0.67	0.60	0.67	0.74
CCC9								1	0.72	0.74	0.52	0.80
CCC11									1	0.79	0.61	0.84
CCC12										1	0.52	0.77
MTU											1	0.62
UCC												1

4.5.2.2 Clustering

The two clustering algorithms were applied to investigate the CSI results of the Cork City network. The silhouette scores for each number of assigned clusters (two to five) were low, with two clusters showing the highest mean score (Figure A2.9). Similarly, with the FCM analysis, two clusters showed the highest score with the Calinski-Harabasz indices (Figure A2.10).

The dendrogram produced from the hierarchical clustering and the membership weights for two clusters from FCM clustering are shown in Figure 4.6 and Figure 4.7, respectively. It is clear that devices MTU, CCC5, and CCC8 are all grouped together in one branch, Cluster 2, with the remainder of the devices in the other branch. The one assignment difference between the two clustering methods is CCC3, which has a higher membership weight towards Cluster 2 with the FCM method but does not branch with that cluster in the dendrogram. However, its membership weight is close to 0.5. CCC1 also shows a membership weight close to 0.5, however it is showing a higher weight towards Cluster 1, as per the hierarchical clustering results. Devices in Cluster 2, except for CCC3, all have the lowest mean CSI value.

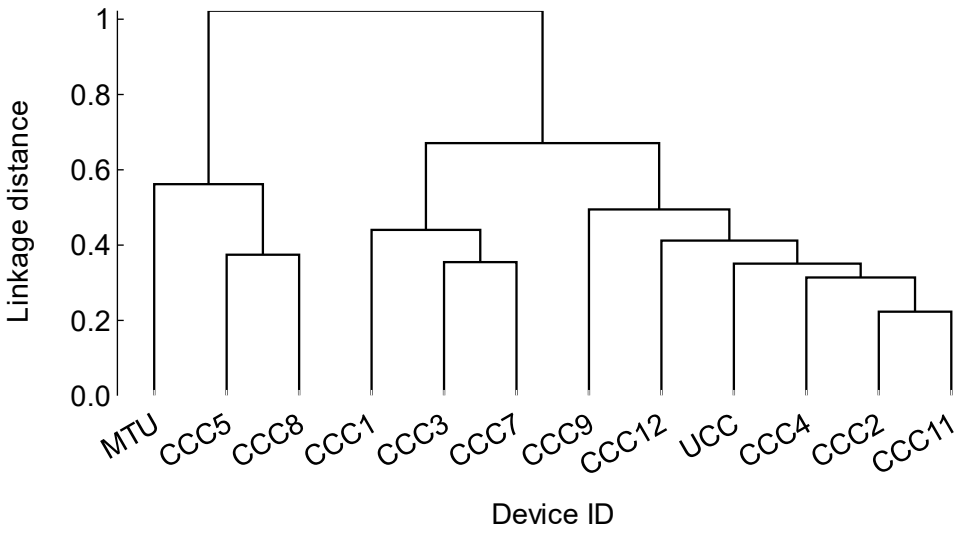


Figure 4.6: *Dendrogram output from hierarchical clustering of the CSI data from the Cork sensor network.*

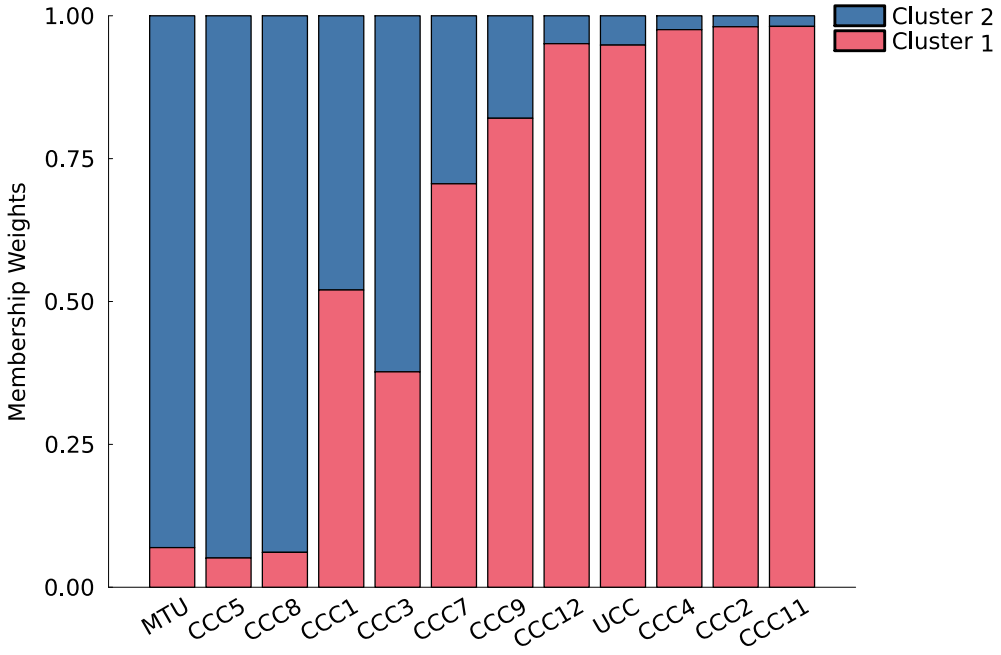


Figure 4.7: *Membership weights from FCM clustering of the CSI data from the Cork sensor network.*

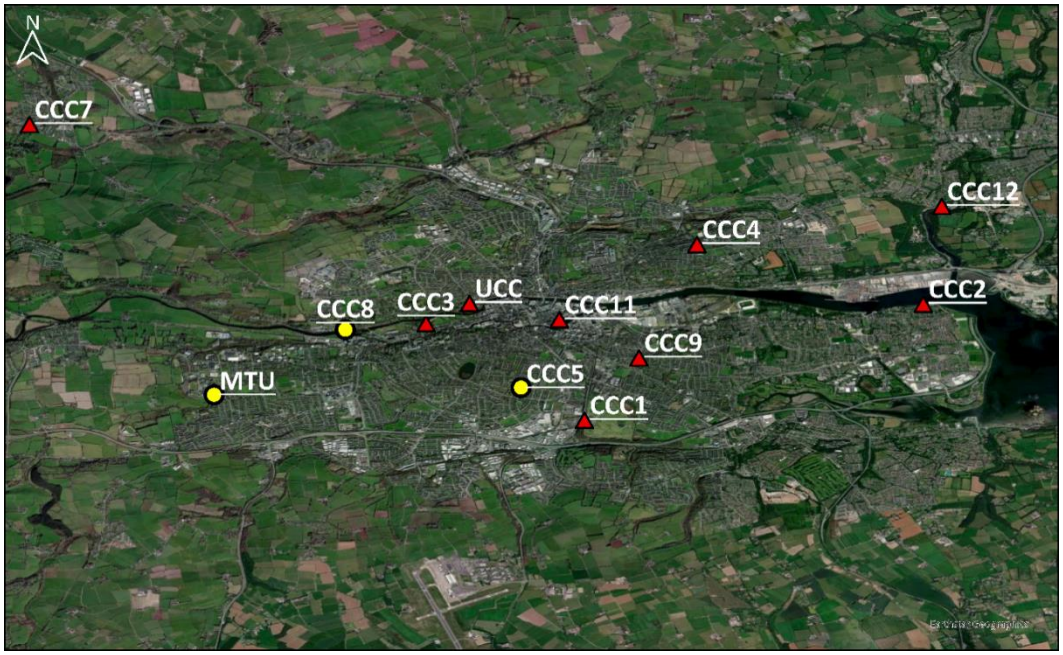


Figure 4.8: *Cork City AQS locations with two cluster groups indicated. Cluster 1 devices (red triangle markers) are located in the city centre and east/northeast, while Cluster 2 devices (yellow circle markers) are mainly located on the western side of the city.*

4.5.3 Assessment of the Spatial Representativeness of Air Quality Monitoring Locations

A key benefit of the CSI metric for AQS networks is that one sensor can be singled out and analysis can assess the spatial representativeness of a given location in the AQS network by quantitively exploring how similar its $PM_{2.5}$ profile is to other locations. If a network sensor is co-located with a reference instrument, then the CSI values for that sensor can be used to provide a measure of the representativeness of the designated monitoring location and how well it informs the assessment of population exposure to air pollution.

In Dungarvan, the device A6P was co-located with a $PM_{2.5}$ instrument (Osiris, Turnkey) deployed as part of the national air quality monitoring network. The

instrument is not a reference instrument but is certified to provide indicative measurements of $PM_{2.5}$ (EPA), 2023; Turnkey, 2024). A6P had a mean CSI of 0.63, the fifth highest of the mean CSI values across all devices. The similarity indices for A6P are included in Table 4.5 and represented spatially in Figure 4.9. All CSI values are below the minimum similarity threshold of 0.85 for two Clarity S-node devices in the Dungarvan network. The most similar devices are found to the north-east of this location, AJ3 and AYG. Interestingly, the similarity of PM profiles does not decrease with increasing distance from A6P. Devices on the furthest western (AZ, A8Z, AP7) and eastern (AWF, AY9N) edges of the town have a CSI within 0.6 to 0.7, yet devices A4, A7, AQ, and A9 are all at or below 0.6 despite being physically closer to A6P. This suggests that the location type is more important when it comes to assessing the similarity of locations within Irish towns as opposed to physical proximity, as A4, A7, AQ, and A9 are all fully surrounded by residential areas, whereas the other devices are in more open areas or on the edge of the town.



Figure 4.9: *Dungarvan AQS locations with CSI results indicated in coloured circles (blue = lowest CSI, yellow = highest CSI), and A6P location indicated with red pin marker.*

One of the devices in the Cork City senor network, UCC, was co-located alongside a reference instrument (BAM-1020) at the national air quality monitoring location on UCC campus. The CSI values for device labelled UCC are shown in Figure 4.10, showing how similar the measurements at this site are compared to the rest of the locations in the sensor network. The CSI scale on the map has been adjusted for these values. Similar to the Dungarvan case, there are devices which show high similarity (CCC4, CCC2, CCC12, CCC1) with UCC which are not located nearby.

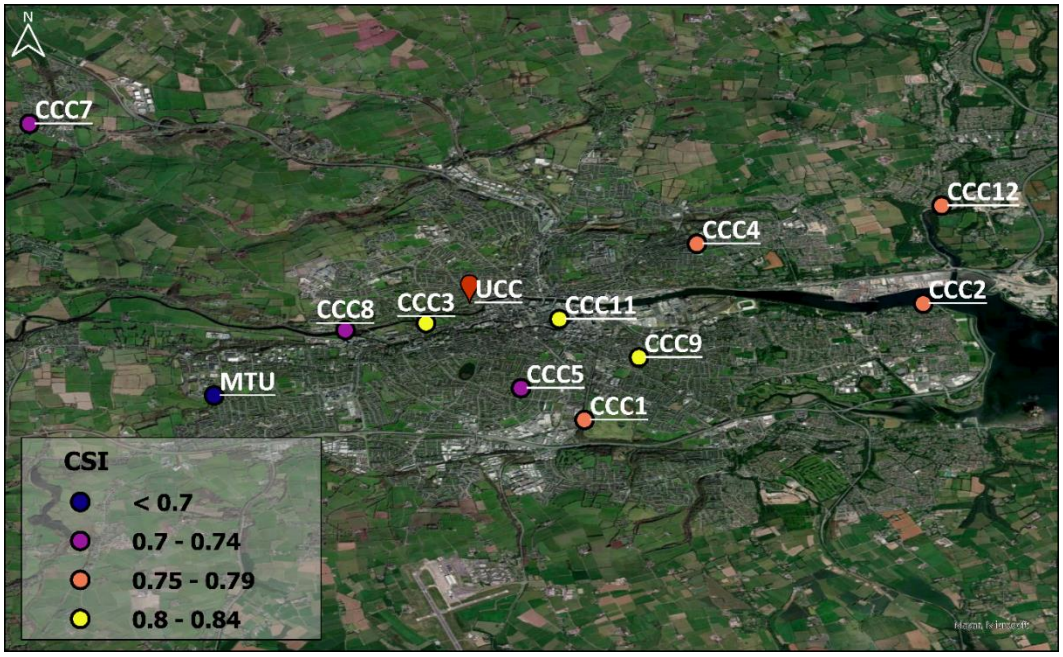


Figure 4.10: *Cork AQS locations with UCC CSI results indicated in coloured circles (blue = lowest CSI, yellow = highest CSI), and UCC location indicated with red pin marker.*

4.6 Conclusion

A robust framework for comparing data series from individual AQSs within a network has been established with the development of a new metric, the CSI, after optimisation and testing on a co-location dataset. The CSI method considers the monitoring network in terms of the similarity of the concentration-time profile of $PM_{2.5}$ at one location to those at the other locations within the network. A harmonised dataset with minimal unexplained inter-sensor variation underpins the development of the CSI method, along with robust tests to ensure that the function represents an unbiased and fair depiction of the inter-sensor relationships after deployment in a monitoring network.

The CSI method has been used to analyse data in two $PM_{2.5}$ sensor networks in Ireland – Dungarvan and Cork City. Clustering techniques are applied to the CSI matrix and comparable similarity trends between locations drives the distinctions made with the clustering algorithm. The resulting groupings can provide several insights into the $PM_{2.5}$ profile at each location, including the likelihood of similarity in pollution sources, spatial patterns, and temporal trends. An interesting contrast in the CSI results from the two monitoring networks was obtained from the clustering analysis. In Dungarvan, the locations generated clusters that were well reflected when comparing the individual diurnal profiles and specifically the diurnal maximum values, indicating that this factor has a major influence when relating the concentration-time profiles at each location to one another in this network. However, for the Cork City network results, this was not as apparent. The clusters were not aligned based on diurnal peaks but rather the differentiating factor was more nuanced. Both clusters contained locations with a mix of higher and lower diurnal maxima and overall maxima.

However, both network groupings reflect that devices may report dissimilar CSI results to other devices located nearby, and that considering location specifications or types, such as residential areas, is more important than physical proximity when it comes to understanding and quantifying the similarities between locations.

The CSI function was also applied to two separate months in the network datasets, with January chosen to represent a period of higher $PM_{2.5}$ levels due to solid fuel burning emissions, and May chosen to represent a period with lower $PM_{2.5}$ concentrations due to reduced solid fuel burning. In both locations, the mean CSI for the network comparisons was higher in May than in January, indicating that higher $PM_{2.5}$ levels is a major driver for lower similarity indices between sensor locations. Combining this with the findings of our previous study, we provide further evidence that high levels of localised $PM_{2.5}$ cause distinct disparities in exposure to poor air quality in different locations. Furthermore, to properly assess the burden of $PM_{2.5}$ experienced by a population and to accurately compare the measurements at two locations, the wintertime PM data must be included in the assessment.

The similarity of $PM_{2.5}$ measured at designated sites in the national air quality monitoring network compared to the rest of locations in the sensor networks was analysed to give an estimation of the representativeness of the air pollution measured at the designated monitoring site. The national monitoring site location in Dungarvan was shown to be moderately representative of the other AQS network locations in the town, with CSI values ranging from 0.53 to 0.72. The CSI values for the Cork City comparison ranged from 0.62 to 0.84, also showing a fair representation of the air pollution experienced in the rest of the network. The CSI function was also tested via synthetic datasets which showed that a positive offset of just 5 $\mu g\, m^{-3}$ resulted in almost halving the CSI, which was a lower CSI than most of the sensor comparisons

in both network locations. So, while a CSI of 0.85 was used as a limit for two sensor measurement sets being very similar, CSI values between 0.6 and 0.7 are still moderately similar. In general, the CSI values in Cork City for the reference site comparison were higher (mean = 0.75) than that of Dungarvan (mean = 0.63), indicating less similarity between the reference site and devices in the Dungarvan network relative to Cork City.

While the function was developed and tested on multiple sensor pairs, and further validated with additional co-located pairs, validation with co-located $PM_{2.5}$ measurements of the PM_{lim}, $C_{lim,upper}$, and $C_{lim,lower}$ parameters for specific applications is recommended to ensure the index represents the dataset accurately. Co-location assessments are also recommended to ensure minimal inter-sensor variation. Nonetheless, the differentiation between higher and lower PM values in the concentration similarity assessment is a strategic choice which acknowledges the complexity of $PM_{2.5}$ data, the varying significance of concentration levels, and the limitations of sensors. It allows for a more accurate representation of similarities while considering real world implications and measurement uncertainties and minimises the potential biases that could arise from an indiscriminate approach, thus ensuring an impartial and unbiased evaluation.

The CSI function displays the potential for AQS networks to be used in conjunction with a regulatory monitoring system. This study has shown the potential for sensor networks to assess the need for additional regulatory monitoring in an area, identifying locations that are poorly represented by the current system. Furthermore, the CSI method can be used to optimise a sensor network by carrying out short term sensor deployments, identifying areas of similarity or dissimilarity, and thus assessing where the best locations for sensors are based on the similarity in exposure to air pollution.

4.7 Appendix 2

Table A2.1: *Co-location periods for Clarity Node-S devices used in the Dungarvan sensor network. Grey cells indicate a device was not present for the co-location, blue cells indicate a device was present for part of the co-location period, and green cells indicate a device was present for the entire co-location period.*

	Co-location deployment period		
Device ID	**19/01/2022–28/01/2022**	**29/10/2022–22/11/2022**	**16/06/2023–17/07/2023**
A3			
A4			
A8H			
A8Z			
A9			
AQ			
AZ			
A7			
A6P			
AJ3			
AP7			
AQV			
ARF			
AW6			
AWF			
AY9N			
AY93			
AYG			

Table A2.2: *Linear regression results for co-located Clarity Node-S devices when compared to the mean values of all co-located data, before and after data harmonisation procedures were carried out.*

	Pre-harmonisation			**Post-harmonisation**		
Device ID	**Slope**	**Intercept**	**R^2**	**Slope**	**Intercept**	R^2
	µg m^{-3}	**µg m^{-3}**		**µg m^{-3}**	**µg m^{-3}**	
A3	1.036	0.933	0.992	1	0	0.992
A4	1.012	−0.278	0.996	1	0	0.996
A8H	0.905	0.979	0.995	1	0	0.995
A8Z	1.003	−0.302	0.997	1	0	0.997
A9	1.06	−0.535	0.979	1	0	0.979
AQ	0.989	0.859	0.997	1	0	0.997
AZ	0.979	0.079	0.997	1	0	0.997
A7	0.947	0.047	0.992	1	0	0.992
A6P	1.027	0.632	0.995	1	0	0.995
AJ3	0.989	−0.415	0.993	1	0	0.993
AP7	0.964	−0.477	0.992	1	0	0.992
AQV	1.042	0.797	0.997	1	0	0.997
ARF	1.016	−0.276	0.996	1	0	0.996
AW6	0.958	0.291	0.984	1	0	0.984
AWF	0.965	−0.554	0.989	1	0	0.989
AY9N	1.065	−0.364	0.988	1	0	0.988
AY93	0.955	0.298	0.995	1	0	0.995
AYG	0.976	−0.168	0.994	1	0	0.994

Table A2.3: *Concentration similarity indices for hourly averaged $PM_{2.5}$ measured by a range of co-located PurpleAir devices. PM_{lim} = 15 μg m^{-3}, $C_{lim,\ upper}$ = 0.2, $C_{lim,\ lower}$ = 0.7.*

	Unit 2	**Unit 3**	**Unit 4**	**Unit 5**
Unit 2	1	0.98	0.99	1
Unit 3		1	0.99	0.99
Unit 4			1	0.99
Unit 5				1

Table A2.4: *Concentration similarity indices for the hourly averaged $PM_{2.5}$ measured by Clarity Node-S devices in the Dungarvan AQS network for January 2023.*

	A3	**A4**	**A8H**	**A8Z**	**A9**	**AQ**	**AZ**	**A7**	**A6P**	**AJ3**	**AP7**	**AQV**	**ARF**	**AW6**	**AWF**	**AY9N**	**AY93**	**AYG**
A3	1	0.48	0.60	0.46	0.50	0.63	0.45	0.51	0.50	0.46	0.26	0.46	0.40	0.58	0.42	0.41	0.38	0.46
A4		1	0.47	0.46	0.55	0.48	0.52	0.51	0.43	0.54	0.38	0.51	0.45	0.51	0.51	0.50	0.45	0.49
A8H			1	0.49	0.53	0.54	0.46	0.56	0.49	0.53	0.37	0.55	0.44	0.59	0.47	0.57	0.41	0.50
A8Z				1	0.58	0.42	0.62	0.46	0.49	0.57	0.57	0.57	0.52	0.52	0.55	0.61	0.49	0.60
A9					1	0.48	0.57	0.53	0.46	0.54	0.47	0.56	0.45	0.54	0.50	0.51	0.44	0.56
AQ						1	0.43	0.50	0.47	0.45	0.26	0.45	0.39	0.56	0.37	0.45	0.35	0.42
AZ							1	0.52	0.52	0.61	0.52	0.61	0.51	0.53	0.56	0.53	0.44	0.67
A7								1	0.46	0.51	0.34	0.53	0.41	0.56	0.47	0.46	0.37	0.48
A6P									1	0.61	0.42	0.50	0.49	0.57	0.55	0.51	0.49	0.61
AJ3										1	0.49	0.67	0.54	0.59	0.72	0.53	0.63	0.71

Table A2.4 continued.

	A3	**A4**	**A8H**	**A8Z**	**A9**	**AQ**	**AZ**	**A7**	**A6P**	**AJ3**	**AP7**	**AQV**	**ARF**	**AW6**	**AWF**	**AY9N**	**AY93**	**AYG**
AP7											1	0.45	0.50	0.37	0.49	0.51	0.44	0.54
AQV												1	0.50	0.58	0.63	0.55	0.52	0.66
ARF													1	0.52	0.54	0.59	0.46	0.58
AW6														1	0.54	0.59	0.49	0.58
AWF															1	0.49	0.64	0.73
AY9N																1	0.50	0.54
AY93																	1	0.63
AYG																		1

Table A2.5: *Concentration similarity indices for the hourly averaged $PM_{2.5}$ measured by Clarity Node-S devices in the Dungarvan AQS network for May 2023.*

	A3	**A4**	**A8H**	**A8Z**	**A9**	**AQ**	**AZ**	**A7**	**A6P**	**AJ3**	**AP7**	**AQV**	**ARF**	**AW6**	**AWF**	**AY9N**	**AY93**	**AYG**
A3	1	0.85	0.90	0.83	0.84	0.89	0.87	0.88	0.90	0.88	0.80	0.90	0.81	0.88	0.82	0.81	0.86	0.86
A4		1	0.84	0.82	0.87	0.80	0.82	0.89	0.82	0.87	0.81	0.85	0.75	0.81	0.81	0.79	0.79	0.81
A8H			1	0.83	0.84	0.86	0.83	0.87	0.88	0.86	0.79	0.86	0.75	0.82	0.79	0.80	0.85	0.83
A8Z				1	0.87	0.84	0.91	0.85	0.85	0.87	0.89	0.88	0.78	0.76	0.82	0.82	0.82	0.90
A9					1	0.79	0.86	0.89	0.84	0.91	0.85	0.88	0.78	0.81	0.89	0.85	0.78	0.84
AQ						1	0.82	0.83	0.84	0.83	0.76	0.87	0.74	0.80	0.73	0.77	0.83	0.84
AZ							1	0.82	0.88	0.90	0.83	0.93	0.84	0.84	0.86	0.80	0.81	0.94
A7								1	0.86	0.88	0.83	0.86	0.77	0.86	0.82	0.80	0.82	0.82
A6P									1	0.92	0.84	0.93	0.84	0.90	0.90	0.87	0.86	0.93
AJ3										1	0.86	0.93	0.85	0.89	0.94	0.91	0.84	0.94

Table A2.5 continued.

	A3	A4	A8H	A8Z	A9	AQ	AZ	A7	A6P	AJ3	AP7	AQV	ARF	AW6	AWF	AY9N	AY93	AYG
AP7											1	0.82	0.81	0.72	0.87	0.79	0.80	0.82
AQV												1	0.84	0.89	0.90	0.86	0.84	0.94
ARF													1	0.77	0.84	0.72	0.76	0.87
AW6														1	0.83	0.76	0.81	0.84
AWF															1	0.91	0.76	0.92
AY9N																1	0.77	0.88
AY93																	1	0.83
AYG																		1

***Table A2.6**: Concentration similarity indices for hourly averaged $PM_{2.5}$ measured by PurpleAir devices in the Cork City AQS network for January 2021.*

	CCC1	**CCC2**	**CCC3**	**CCC4**	**CCC5**	**CCC7**	**CCC8**	**CCC9**	**CCC11**	**CCC12**	**MTU**	**UCC**
CCC1	1	0.55	0.63	0.52	0.72	0.61	0.60	0.42	0.70	0.54	0.55	0.59
CCC2		1	0.54	0.58	0.63	0.48	0.59	0.56	0.69	0.60	0.40	0.59
CCC3			1	0.48	0.58	0.67	0.65	0.35	0.67	0.49	0.63	0.56
CCC4				1	0.59	0.45	0.51	0.54	0.60	0.58	0.42	0.50
CCC5					1	0.54	0.62	0.65	0.70	0.60	0.40	0.66
CCC7						1	0.54	0.33	0.59	0.42	0.60	0.46
CCC8							1	0.65	0.66	0.57	0.40	0.75
CCC9								1	0.50	0.57	0.18	0.70
CCC11									1	0.59	0.46	0.65
CCC12										1	0.33	0.55
MTU											1	0.31
UCC												1

Table A2.7: *Concentration similarity indices for hourly averaged $PM_{2.5}$ measured by PurpleAir devices in the Cork City AQS network for May 2021.*

	CCC1	**CCC2**	**CCC3**	**CCC4**	**CCC5**	**CCC7**	**CCC8**	**CCC9**	**CCC11**	**CCC12**	**MTU**	**UCC**
CCC1	1	0.58	0.70	0.51	0.63	0.63	0.58	0.45	0.50	0.39	0.74	0.58
CCC2		1	0.91	0.94	0.93	0.92	0.94	0.89	0.95	0.91	0.81	0.95
CCC3			1	0.85	0.93	0.95	0.92	0.84	0.88	0.79	0.90	0.92
CCC4				1	0.91	0.86	0.90	0.94	0.93	0.91	0.74	0.92
CCC5					1	0.92	0.90	0.83	0.94	0.83	0.86	0.96
CCC7						1	0.93	0.86	0.90	0.80	0.86	0.91
CCC8							1	0.89	0.92	0.84	0.79	0.90
CCC9								1	0.92	0.91	0.63	0.91
CCC11									1	0.91	0.75	0.96
CCC12										1	0.60	0.87
MTU											1	0.79
UCC												1



Figure A2.1: *Map of Dungarvan with Clarity Node-S sensor locations and labels.*

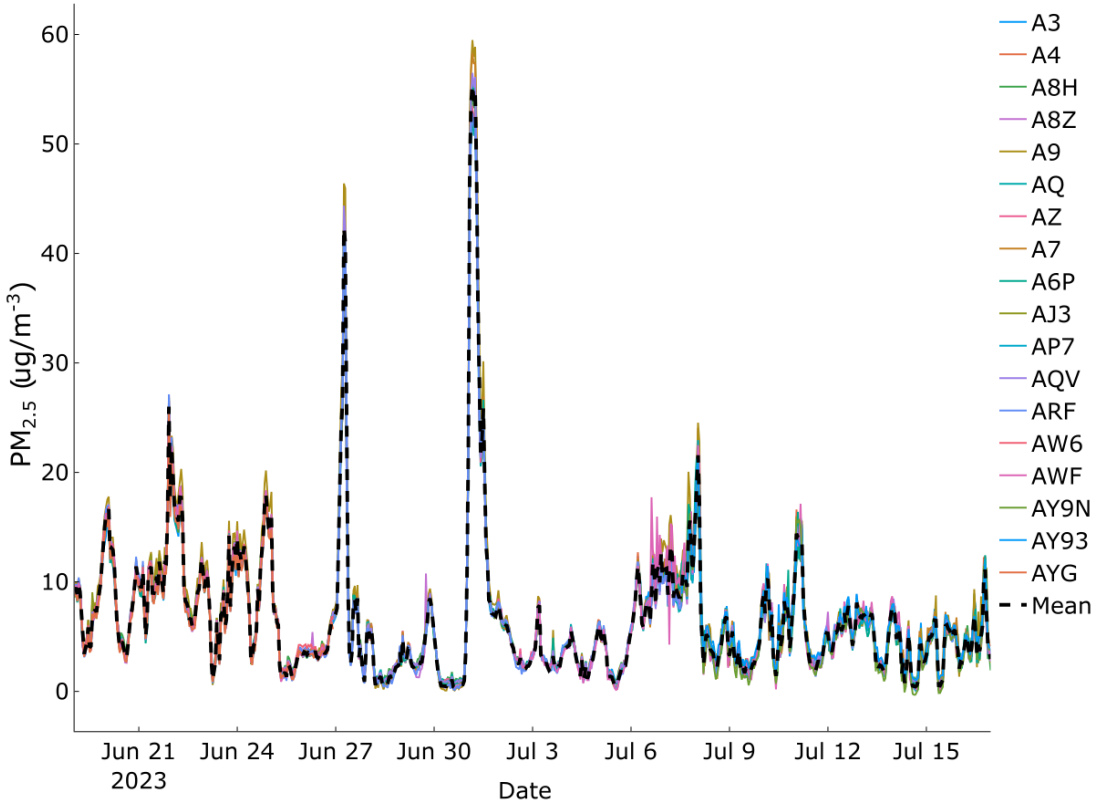


Figure A2.2: *Hourly averaged $PM_{2.5}$ from co-located and harmonised Clarity Node-S devices including the mean $PM_{2.5}$ data series used for the harmonisation (19 June 2023–18 July 2023).*

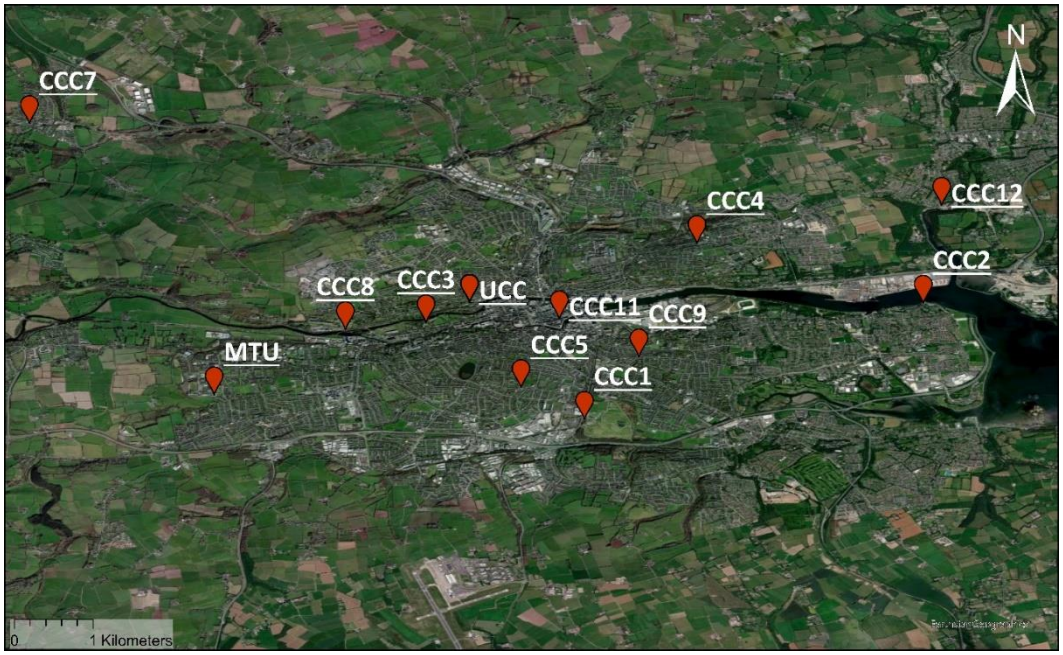


Figure A2.3: *Map of Cork City with PurpleAir PA-II-SD sensor locations and labels.*

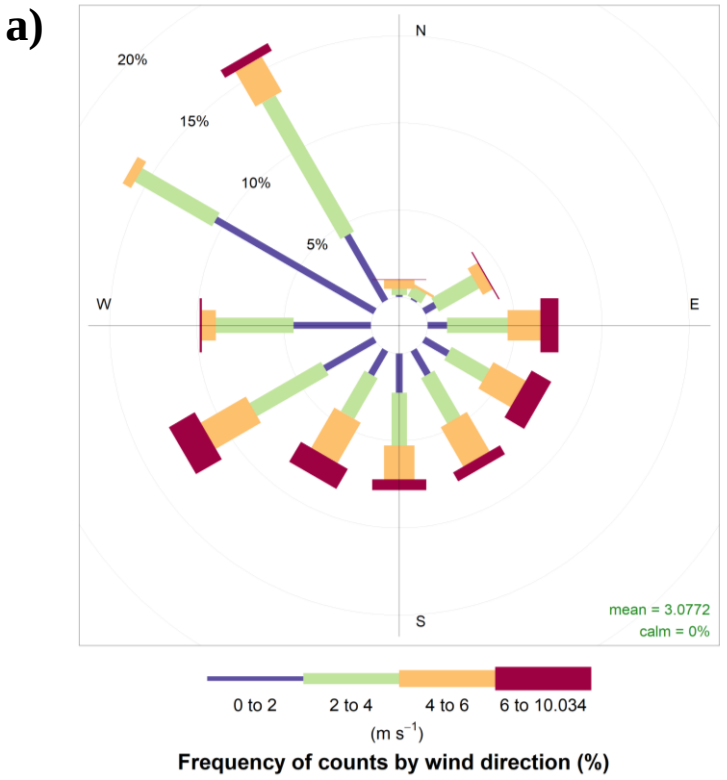


Figure A2.4: *Wind rose plots for devices a) AP7 and b) AY93 in the Dungarvan sensor network (January 2023–May 2023).*

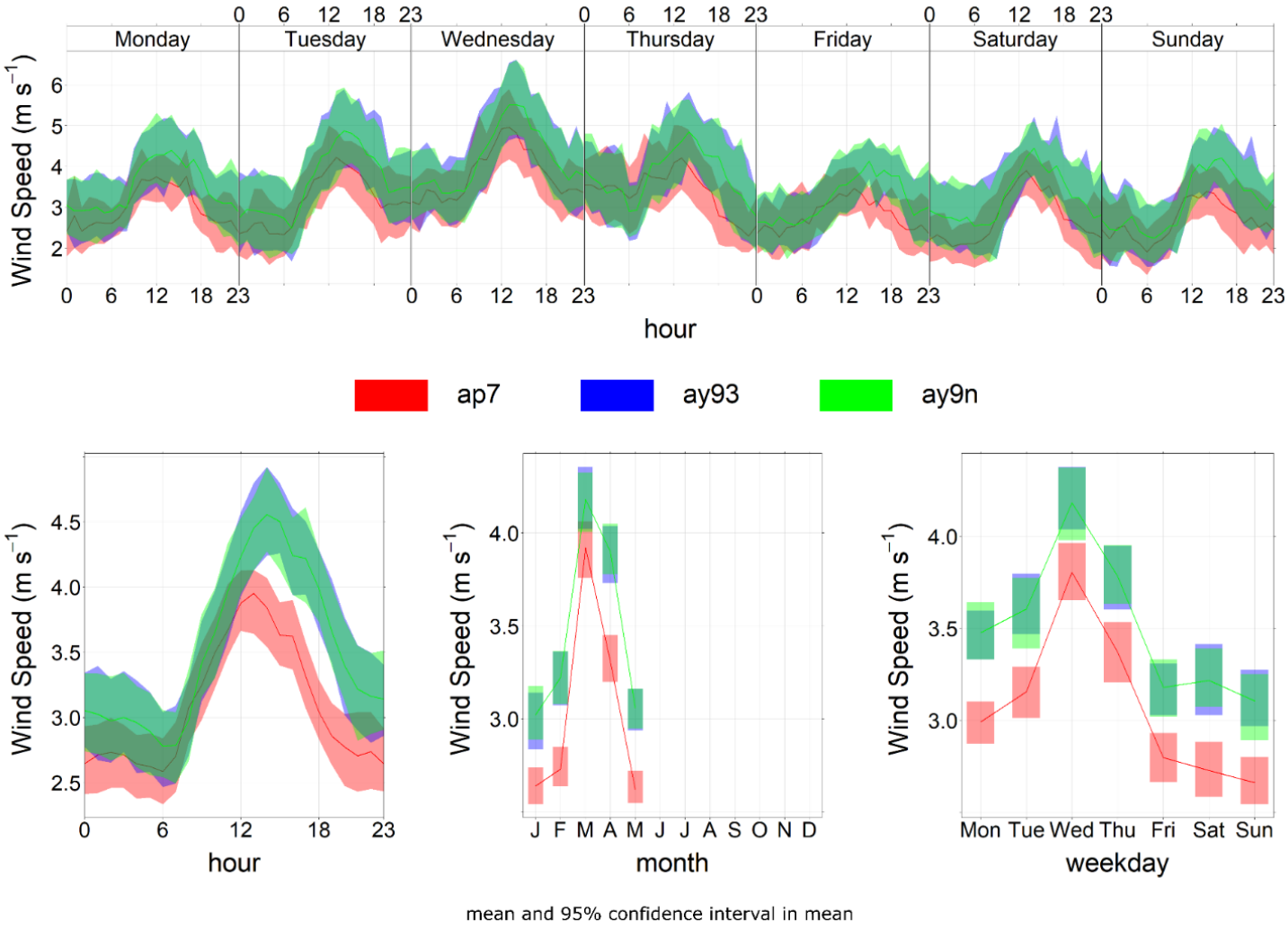


Figure A2.5: *Temporal variation (hour of day - day of week variation, mean diurnal variation, monthly variation, and day of the week variation) in the wind speed recorded by devices AP7, AY93, and AY9N for the January 2023–May 2023 measurement period, showing that each wind sensor reports similar wind speeds.*

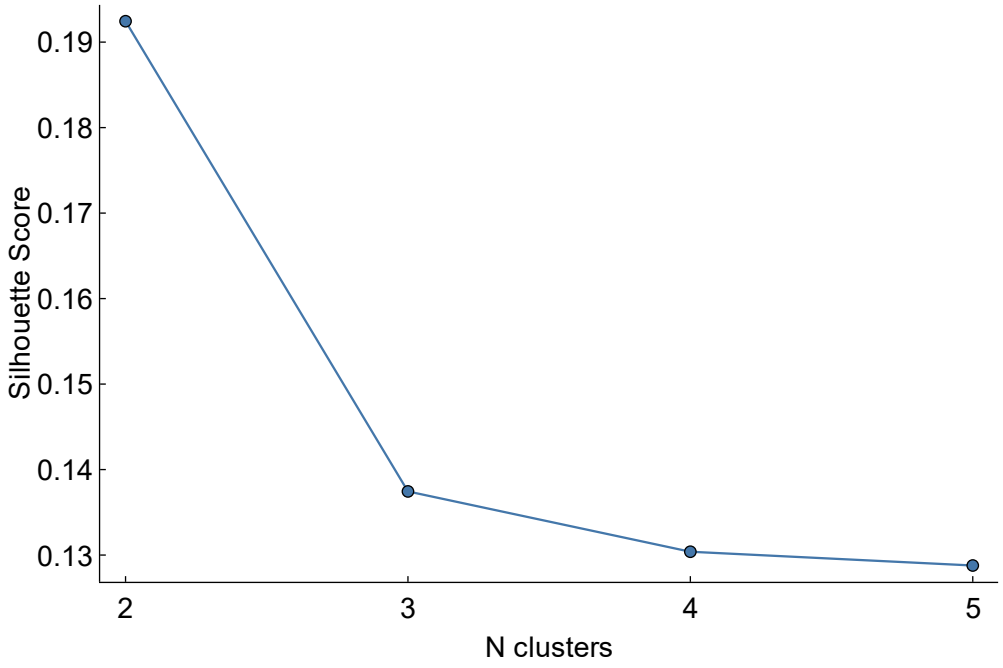


Figure A2.6: *Silhouette scores for hierarchical clustering on the Dungarvan sensor network CSI results. The silhouette score can be used to assess the overall quality of the cluster assignments, where a score close to one indicates high quality clusters and a negative score is indicative of low-quality clusters.*

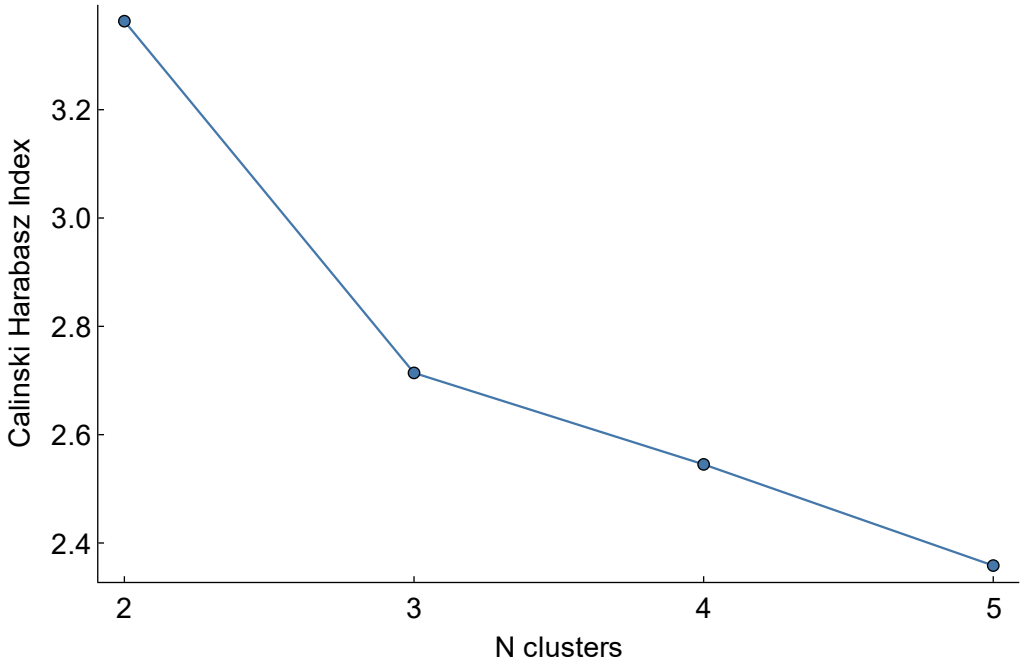


Figure A2.7: *Calinski-Harabasz indices for FCM clustering on the Dungarvan sensor network CSI results. The Calinski-Harabasz index can be used to assess the quality of soft cluster assignments. A higher score indicates high-quality clusters.*

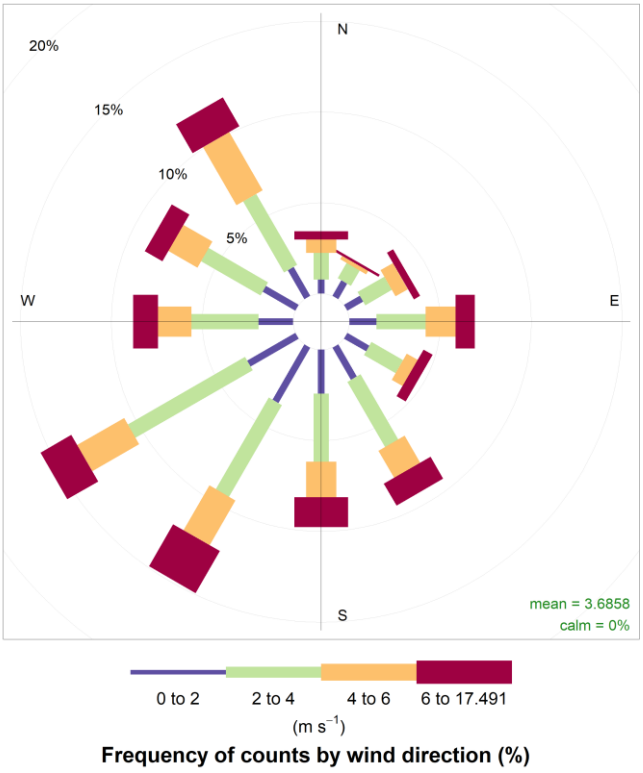


Figure A2.8: *Wind rose plot for Cork Airport (January to May 2021 and September to December 2021).*

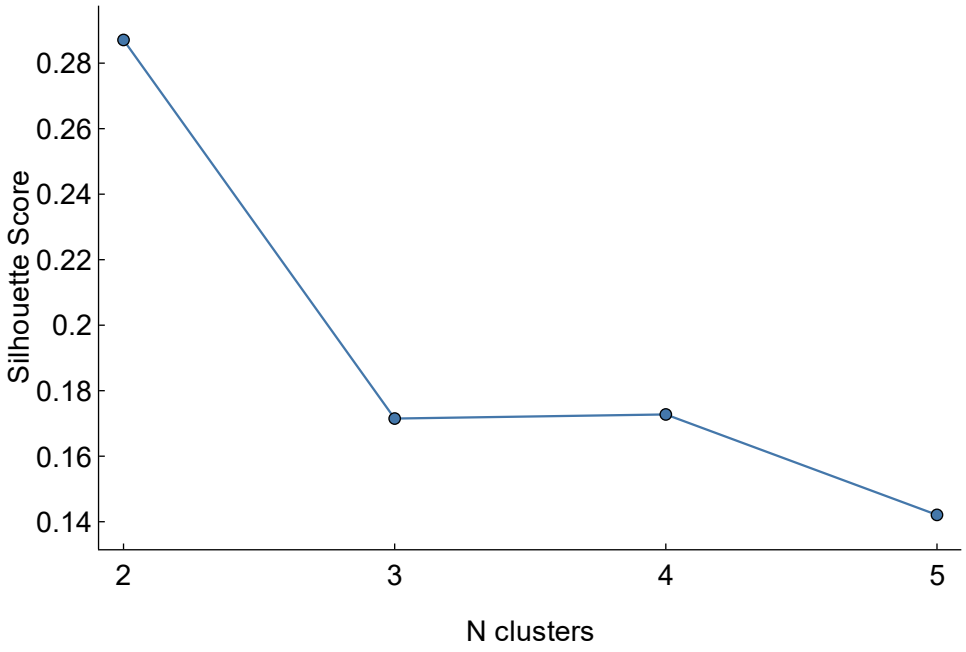


Figure A2.9: *Silhouette scores for hierarchical clustering on the Cork sensor network CSI results. The silhouette score can be used to assess the overall quality of the cluster assignments, where a score close to one indicates high quality clusters and a negative score is indicative of low-quality clusters.*

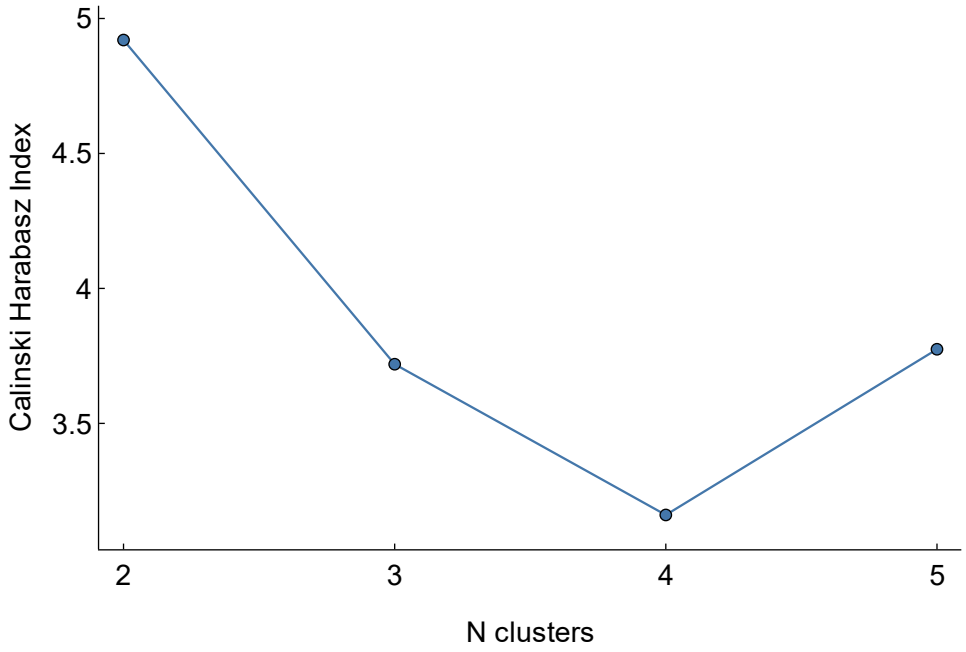


Figure A2.10: *Calinski-Harabasz indices for FCM clustering on the Cork sensor network CSI results. The Calinski-Harabasz index can be used to assess the quality of soft cluster assignments. A higher score indicates high-quality clusters.*

5 Spectral Analysis of Air Quality Sensor Network Data for Local Source Contribution Estimates of $PM_{2.5}$

Reproduced from the manuscript prepared for submission:

Byrne, R., Wenger, J. C., Hellebust, S. Spectral analysis of air quality sensor network data for local source contribution estimates of $PM_{2.5}$.

5.1 Abstract

Spectral analysis by Discrete Fourier Transform is used in conjunction with networks of low-cost air quality sensors (AQSs) to estimate the local source contributions to $PM_{2.5}$ levels. The methodology is supported by the low inter-sensor variation observed for both PurpleAir (PA) and Clarity Node-S devices ($R^2 > 0.8$ for all inter-sensor comparison) and good correlations with reference instruments ($R^2 \geq 0.75$ for all sensor-reference comparisons). A comparison of data from both sensor types and reference instruments produced similar local source contribution estimates (LSCEs), 61%, 68%, and 62% for the reference, PA, and Clarity Node-S, respectively, indicating that sensor errors do not bias the results. Comparisons between uncorrected and corrected sensor data showed negligible differences in the LSCEs (< 3 pp (pp = percentage points)), confirming the robustness of the validated but uncorrected AQS data. The spectral analysis approach was applied to three validated sensor network datasets obtained in Ireland – Cork City 2021, Cork City 2023, and Dungarvan – without calibration against reference data. A strong seasonal variation was observed in each network, with LSCEs varying between 44–71% (Cork City 2021), 26–73% (Cork City 2023) and 49–75% (Dungarvan) during winter months due

to emissions from residential solid fuel burning. The LSCEs for the Cork City network in 2021 were generally similar to those previously obtained using a previously validated method, however, spectral analysis yielded higher estimates in the summer, likely due to the focus on high-frequency components across the entire signal, capturing more local traffic emissions. An unexpectedly low contribution of 17% from local sources was identified in Cork City during March 2021 and subsequently attributed to a regional pollution event. Analysis of the Cork City network in 2023 indicated higher spatial variation in LSCEs in December 2023 (range = 45 pp) compared to December 2021 (range = 35 pp). A reduction in the LSCEs in December 2023 compared to December 2021 was also found for all common device locations with the decreases ranging 28 pp, however this was not reflected in the yearly estimates, again highlighting the strong seasonality of local air pollution in Ireland. Extending the spectral analysis to the Dungarvan network in 2023 revealed less variation in local sources across the network (range = 26 pp) compared to the winter month estimate for 2023 in Cork City. Overall, this study demonstrates that spectral analysis of validated AQS network data can provide local source contribution estimates of $PM_{2.5}$ without the need for spatially dense deployment of reference instrumentation. Moreover, this approach has the potential to supplement regulatory monitoring activities and also offers a cost-effective alternative for determining $PM_{2.5}$ sources, especially in areas with limited access to reference grade measurements.

5.2 Introduction

As urbanisation and industrialisation continue to expand globally, the need for effective air quality monitoring tools becomes increasingly vital to measure and understand the health and societal impacts of air pollution (Cesaroni et al., 2013; Lelieveld et al., 2015; Pedersen et al., 2013). Fine particulate matter, comprising particles with a diameter of 2.5 micrometres or less ($PM_{2.5}$), poses significant risks to human health. $PM_{2.5}$ can penetrate deep into the respiratory system, cause respiratory and cardiovascular diseases (Pope et al., 2020; Pope & Dockery, 2012; Samoli et al., 2013), and has also been suggested as a risk factor for developing vascular dementia and Alzheimer's disease (Jung et al., 2015; Oudin et al., 2016). $PM_{2.5}$ is both a regional and local issue, with contributions from long-range transport, urban sources, and hyper-local emissions in cities. Given the prevalence of $PM_{2.5}$, it is crucial to extend air quality monitoring efforts to deliver high-spatial and temporal resolution measurements to understand local variations in the concentrations and exposure impacts of $PM_{2.5}$.

Traditional air quality monitoring tools offer precise, quality assured data. However, high data quality comes with high costs for equipment and personnel, often restricting the availability of such tools in expansive monitoring networks. Air quality sensors (AQSs), on the other hand, offer a means to measure pollutant concentrations at increased spatial and temporal resolutions at a fraction of the cost (Kumar et al., 2015; Munir et al., 2019). Although AQSs only provide indicative measurements of pollutant concentrations, the growing number of reports on the reliability and quality of $PM_{2.5}$ sensors in particular (Liu et al., 2019; Liu et al., 2020; Malings, et al., 2020a), demonstrates their potential for future widespread deployment. When incorporated

into a well-designed network and when data is appropriately assessed and interpreted, $PM_{2.5}$ sensors enhance the spatial and temporal granularity of ambient air pollutant measurements, thereby supplementing existing air quality monitoring infrastructure. They can also provide measurements of $PM_{2.5}$ in places where air pollution data has previously been unattainable, thus tackling the global challenge of delivering effective air quality monitoring in low- and middle-income countries (Coker et al., 2021; McFarlane et al., 2021; Okure et al., 2022; Raheja et al., 2021).

The growing trend of deploying $PM_{2.5}$ sensors in networks has created a need for methodologies that can maximise the benefits of these systems while addressing their limitations. One such method is spectral analysis, specifically the Discrete Fourier Transform (DFT), which allows for the mathematical analysis of air pollutant concentrations measured over extended periods. The measured concentrations often exhibit cyclical fluctuations due to events like rush hour traffic and residential heating, and by decomposing a timeseries into its frequency components, spectral analysis helps identify significant periodic patterns in the timeseries and understand their contributions to overall data variance (Akdi et al., 2021; Choi et al., 2008; Frederickson et al., 2022; Hogrefe et al., 2000, 2006; Lazić et al., 2016; Marr & Harley, 2002; Warner, 1998). This method can reveal valuable insights into temporal patterns and sources of $PM_{2.5}$ in areas where local emissions cause rapid, short-term fluctuations, or high frequency cycles in concentration, due to the heterogenous aerosol profile (Ceburnis et al., 2006; Cesari et al., 2016; Kim Oanh et al., 2009; Seinfeld & Pandis, 2016; Tunno et al., 2012). Conversely, strong regional transport of air pollutants results in slowly elevating concentration baselines. While specific sources cannot be identified with this method or with $PM_{2.5}$ mass concentration values

alone, important distinctions in source types can be made without requiring in-depth chemical analysis of the ambient $PM_{2.5}$.

In a previous study, we used data collected by a $PM_{2.5}$ sensor network in Cork City, Ireland, to estimate the contribution of local pollution sources in a summer and winter period in 2021 (Byrne et al., 2023, See Chapter 3). This method, the rapid fluctuation method, relied on the high frequency resolution of the two-minute timeseries data, to pinpoint periods of highly fluctuating $PM_{2.5}$ concentrations that are dominated by local $PM_{2.5}$ emission sources. The results highlighted the very localised nature of $PM_{2.5}$ caused by residential solid fuel burning during winter periods, which has previously been shown to be a significant problem in many towns and cities in Ireland (Dall'Osto et al., 2013a; Kourtchev et al., 2011; Lin et al., 2019b; Lin et al., 2018; Ovadnevaite et al., 2021; Wenger et al., 2020).

This study aims to explore the application of spectral analysis to discriminate between local and non-local sources contributing to $PM_{2.5}$ measurements across a sensor network. Using hourly $PM_{2.5}$ measurements, we analyse the frequencies within the data and compare the results to those previously obtained with the rapid fluctuation method for Cork City in 2021. The spectral analysis method is also used to provide an updated assessment of $PM_{2.5}$ in the Cork City network during 2023, after the Irish Government introduced new legislation regarding the sale and distribution of the most polluting solid fuels in October 2022, the Air Pollution Act 1987 (Solid Fuels) Regulations 2022. We also apply this method to another $PM_{2.5}$ sensor network in Dungarvan, a town in the south-east of Ireland.

5.3 Methodology

5.3.1 Data Collection and Preprocessing

All data collection, preprocessing, and subsequent data transformations and analyses were carried out using the Julia Programming language and the openair package written for the R programming language. (Bezanson et al., 2017; Carslaw & Ropkins, 2012). Calibration of the $PM_{2.5}$ data obtained from the sensor networks was not needed because the spectral analysis method relies on the relative changes in the measured signal, rather than the absolute concentrations.

5.3.1.1 Cork City Sensor Network

The Cork City sensor network consisted of 16 locations monitored with PurpleAir PA-II SD devices (PurpleAir, USA.). Each of these devices contains two Plantower PMS5003 sensors to measure $PM_{2.5}$ via light scattering within an effective range of 0−500 μg m^{-3}, a maximum range of 0−1000 μg m^{-3}, and at a resolution of 1 μg m^{-3} (Yong, 2016). The PurpleAir (PA) devices obtain measurements by alternating between the two Plantower sensors every 5 seconds, to produce an averaged measurement every 120 seconds.

The Cork City network is operated and maintained by Cork City Council. Two periods are extracted from the PA online platform for the Cork City network, one spanning the calendar year 2021 (01 January 2021–31 December 2021) and the other spanning the year 2023 (01 January 2023–31 December 2023). Units placed at specific locations are identified with unique names, although the specific sensor in place at a given location may have changed between 2021 and 2023 (Figure 5.1). For the locations to

be addressed consistently, we maintain the same identification per location between both years. The data processing procedure followed in this study to produce filtered hourly averaged measurements for both Cork City sensor network datasets is identical to that reported for the 2021 data (Byrne et al. (2023) (See Chapter 3), omitting the calibration via a multiple linear regression model, unless explicitly used for comparative purposes.

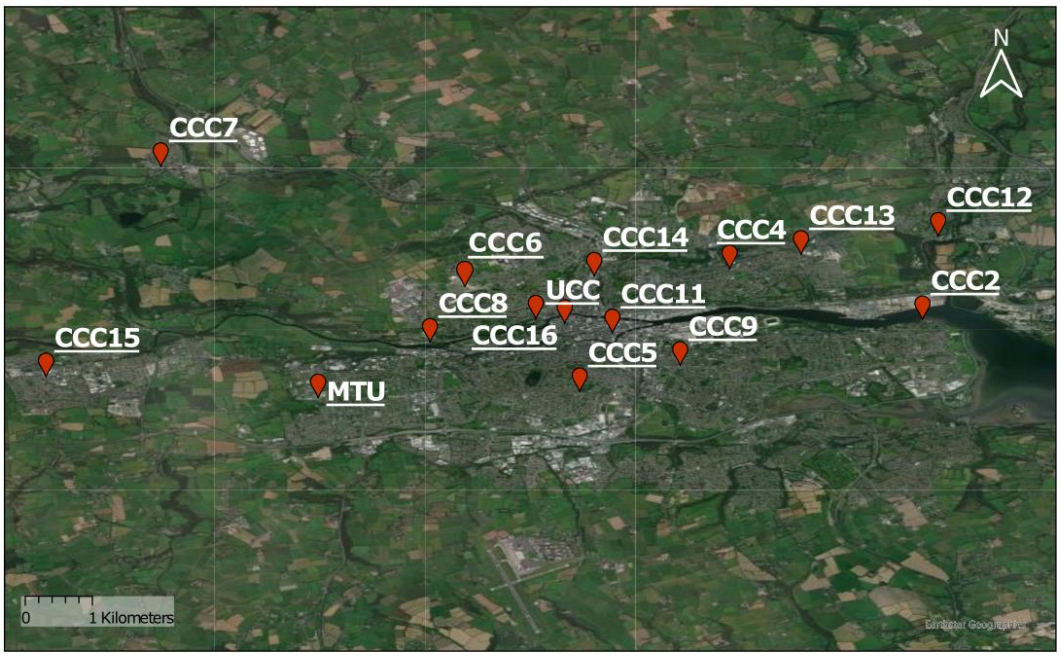


Figure 5.1: *Map of Cork City with PurpleAir sensor locations and labels.*

5.3.1.2 Dungarvan Sensor Network

The Dungarvan sensor network consisted of 18 solar powered Clarity Node-S devices (Clarity Movement Co., USA). These AQSs use the Plantower PMS6003 sensor to measure $PM_{2.5}$ via light scattering within a range of 1–1000 μg m^{-3} and at a resolution of 1 μg m^{-3} (Clarity Movement Co., 2023b). The Clarity Node-S devices measured the ambient $PM_{2.5}$ mass concentration every 8 minutes. The devices were deployed across

the town of Dungarvan, encompassing a range of different environments including residential, coastal, and roadside locations (Figure 5.2). Many of the locations were a mix of these different environments. The measurement period in Dungarvan was 01 November 2022–31 May 2023, however the deployment of three devices (AP7, AY9N, AY93) was delayed until 12 January 2023 due to logistical constraints. The data processing procedure to produce hourly averaged measurements is identical to that reported on the same data by Byrne et al. (2024) (See Chapter 4), omitting the data harmonisation step.



Figure 5.2: *Map of Dungarvan with Clarity Node-S sensor locations and labels.*

5.3.1.3 Co-location assessments

Comparisons of the AQS data with reference-grade measurements during multiple co-location assessments were carried out to support the methodology in this study. Reference data for the relevant location and period was downloaded in hourly format from the website for the national ambient air quality monitoring network

(www.airquality.ie), managed by the EPA, Ireland. This data is presented on the website as automatic data collections from the monitoring locations and as such is not officially validated by the EPA. A PA sensor was co-located with a BAM (BAM-1020, Met One Instruments) in 2021 at the UCC air quality monitoring site (51.900111, -8.486329). The reference instrument reported data for all of 2021 except for 18 May 2021−25 June 2021. A PA and a Clarity Node-S were both co-located with a Palas Fidas 200 instrument at the EPA air quality monitoring site in Edenderry, Co. Offaly (53.342666, -7.047655) from 01 February 2022 to 01 March 2022.

Additional assessments were carried out using a number of co-located AQSs for inter-sensor variation analysis. Three PA units were co-located for the period 01 October 2023–02 February 2024, and the 18 Clarity units underwent a full co-location assessment prior to and after deployment in Dungarvan on the roof of the Ellen Hutchins Building, UCC (51.895136, -8.516146). All co-located sensor data was processed following the same procedures described in the previous sections.

5.3.1.4 Meteorological measurements

Meteorological data was used for investigating the association between local $PM_{2.5}$ levels and atmospheric conditions. For Cork City, data recorded at Cork Airport by Ireland's National Meteorological Service, Met Éireann, was accessed from the website (www.met.ie). The airport weather station is located approximately 5.5 km from Cork city centre.

While there are no official hourly meteorological measurements collected nearby Dungarvan, three of the Clarity Node-S devices were fitted with Clarity Wind Modules (AP7, AY93, AY9N), providing measurements of wind speed and direction (Clarity

Movement Co., 2024). For technical reasons, device AY9N did not capture wind direction measurements, however wind speed and other meteorological variables were recorded. The Wind Module contains a solid-state two-axis ultrasonic anemometer, providing wind speed measurements with a range of 0–60.00 m s^{-1} and a resolution of 0.01 m s^{-1}, and wind direction at a resolution of 0.1°, over a range of 0–359.9°. The module measures ambient temperature, relative humidity, and atmospheric pressure with ranges of −20–75 °C, 10–99%, and 300–1100 hPa, respectively. The corresponding resolutions are 0.1 °C, 0.1%, and 0.02 hPa, respectively (Clarity Movement Co., 2024). Wind speed, wind direction, and temperature measurements were investigated in relation to the local source contributions estimated using spectral analysis of data collected by the Dungarvan sensor network.

5.3.2 Air Quality Sensor Performance Evaluation

The performance characteristics of both AQS types were determined from data collected during multiple co-location tests to ensure homogeneity in the reported $PM_{2.5}$ measurements. Due to logistical constraints, it was not possible to co-locate all the PA devices in the Cork City network to assess specific sensor variability. However, low inter-sensor and inter-unit variability was shown with co-located devices in our previous study on the 2021 Cork City network, where a comparison of four PA units yielded R^2 values consistently exceeding 0.98 (Byrne et al., 2023, See Chapter 3). The PA $PM_{2.5}$ measurements were also highly correlated with hourly reference measurements (R^2 = 0.92). We also obtained a more recent co-location dataset for three PA units (Section 5.4.1) to further emphasise the low variability reported between devices prior to data correction procedures.

The procedure for the co-location assessment of the Clarity units is described in Byrne et al. (2024) (See Chapter 4). When measurements from each sensor were compared to the mean of all sensor measurements, R^2 values above 0.97 were obtained for each of the 18 sensors across the co-location periods prior to applying data-harmonisation procedures. This provided confirmation of low inter-sensor variation, and data harmonisation procedures have not been applied to the Clarity Node-S dataset presented here.

When comparing the sensor datasets during co-location exercises, the R^2 value is used to assess the correlation between two datasets and the Mean absolute error (MAE) and Root mean square error (RMSE) are used to assess the average differences between the two datasets. The equations for the MAE and RMSE calculations are as follows:

$$MAE = \frac{\sum_{i=1}^{N} |(y_i - x_i)|}{N} \quad \textbf{(5.1)}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y_i - x_i)^2}{N}} \quad \textbf{(5.2)}$$

where N is the total number of observations, and y and x are the two concentration variables being compared.

5.3.3 Fourier Transform and Periodogram Representation

DFT spectral analysis is applied with the assumption that the data is present at equally spaced intervals. Any missing data points in AQS and reference datasets were interpolated to achieve the necessary equally spaced measurements. If the first or final data point in the series was missing, it was replaced with the overall mean of the data series. Other missing data points throughout the series were replaced with the mean of

the two nearest neighbouring data points (i.e. the pair of nearest data points before and after the missing data point). The final processing step involved subtraction of the overall mean in each data series to centre each data series around zero.

Most sensors exhibited good data capture rates over the period in question for the respective sensor network, however if a sensor reported low coverage (< 60%) across an entire measurement period or during a specific monthly assessment, it was not included for the source contribution analysis. Sensors with adequate data coverage only across a single month were excluded from analysis of the overall measurement period. The average data capture rates for sensors included in spectral analysis of the entire measurement period in Cork City 2021, Cork City 2023, and Dungarvan were 94%, 92%, and 92%, respectively.

The periodogram was determined for each sensor and used to identify dominant frequencies within the timeseries. The periodogram, *P*, for each timeseries was calculated as follows:

$$P(v_k) = |f(k)|^2 = \left| \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{(-2\pi i v_k n)} \right|^2 \quad \textbf{(5.4)}$$

where N is the total length of the finite data series, x_n, $k = 0, 1, \ldots, N-1$, and $v_k = k/N$. P was calculated for the entire data transformation. When comparing co-located sensor data to reference data, the periodograms were normalised against the highest periodogram peak between the co-located sensor and the reference datasets for a given comparison. Similarly, the periodograms for AQS data series within a given network were normalised against the highest peak from all sensor periodograms within that network for a given comparison.

5.3.4 Source Contribution Estimates

The contributions of local and non-local $PM_{2.5}$ can be distinguished by relating the frequencies in the periodograms to the various $PM_{2.5}$ sources (Frederickson et al., 2022; Hogrefe et al., 2000, 2006). While previous studies typically divided the pollutant measurements into three or more categories usually following a local, city-wide, and regional-background or seasonal split, here we focus on discriminating between just local and non-local emissions. The local emissions of $PM_{2.5}$ in Irish towns and cities are primarily road traffic (typically not seasonal) and residential solid fuel burning, which is dominant in wintertime (Ovadnevaite et al., 2021; Wenger et al., 2020). Since these emission sources follow a regular diurnal pattern, we classify all frequency components above 0.0313 h^{-1} (32 h) as high frequency, local contributions. Previous studies have employed an iterative moving-average based low-pass filter (Kolmogorov-Zurbenko filter) to remove high-frequency (local emission related) variations from a timeseries (Hogrefe et al., 2000; Rao et al., 1997), with the use of study-specific parameters relating to the assignment of frequency components recommended to provide optimal results (Rao et al., 1997). While we do not have chemical profile data relating to the study domain, an empirical methodology has been utilised to estimate an appropriate cut-off frequency to separate local and non-local contributions based on acknowledgement of the anthropogenic activity linked cycles in the data. As Figure 5.3 shows, the time of the maximum $PM_{2.5}$ level is variable, with 18:00 being the most common time, but it is often observed at different hours, meaning that the duration between maximum levels per day can be more than 24 hours. The 0.0417 h^{-1} (24 h) component was significantly higher in most cases than surrounding frequency components in the periodograms (Figure 5.4a), reflecting the characteristic rise in $PM_{2.5}$ during wintertime evenings in Ireland, however, when

decomposing isolated months of timeseries data for comparison, the frequency resolution is naturally reduced (Sevgi, 2007) and contribution information was being lost with a limit of 0.0417 h^{-1} (Figure 5.4b).

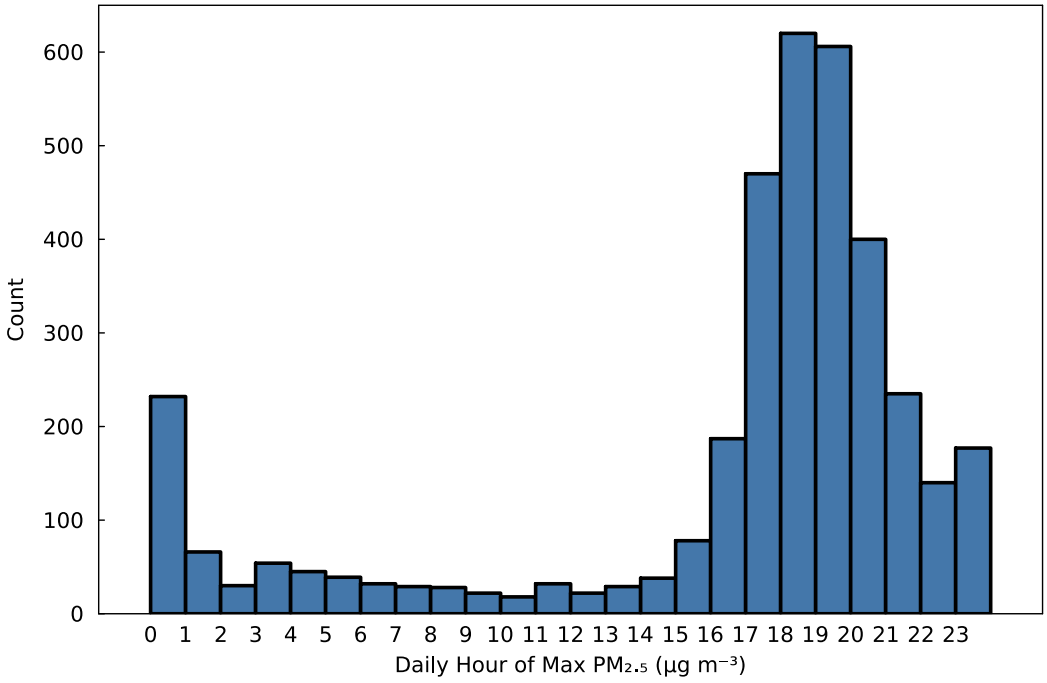


Figure 5.3: *Distribution (by hour) of daily maximum $PM_{2.5}$ measurement from all sensors in the Dungarvan sensor network.*

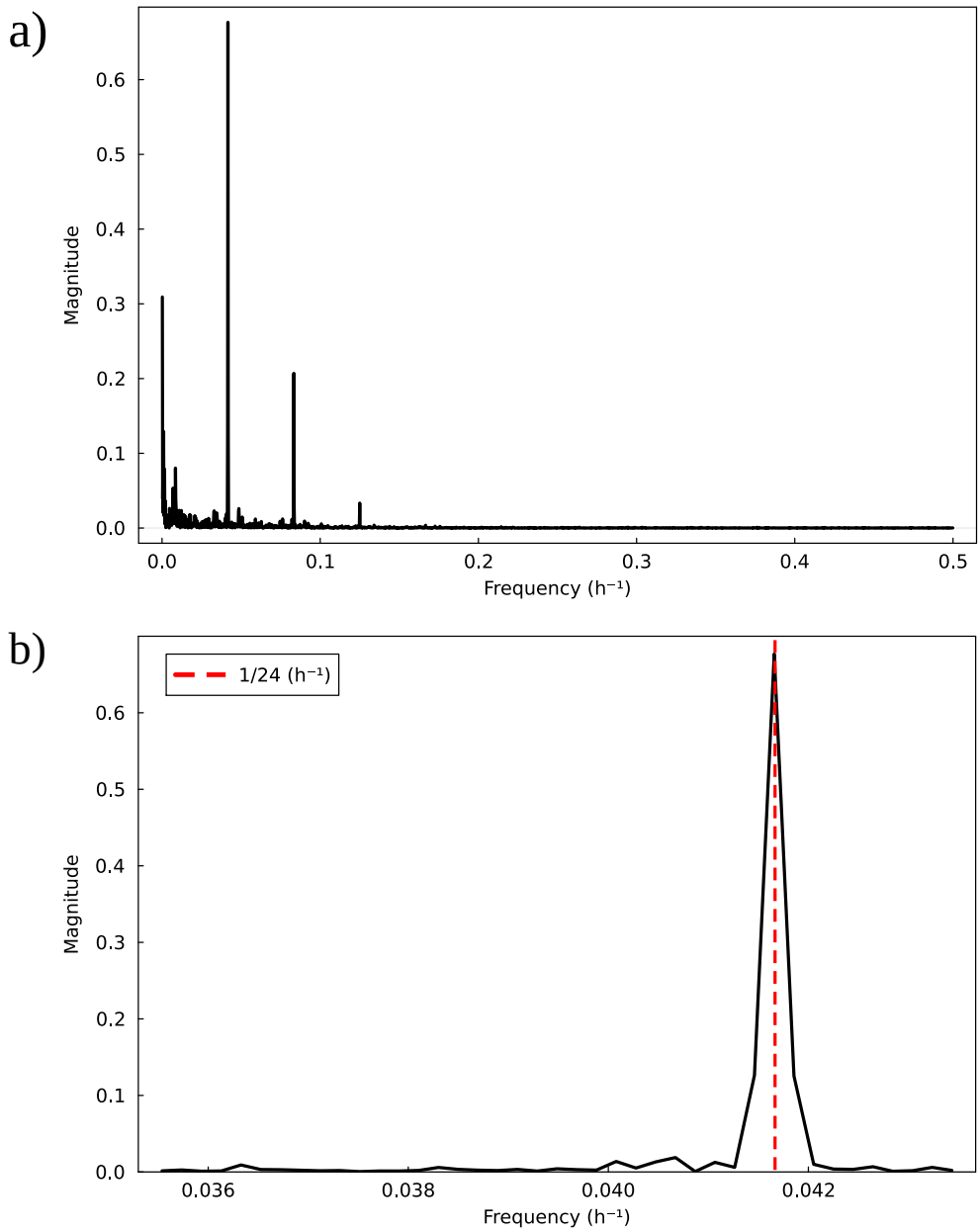


Figure 5.4: *a) Full periodogram of device A3 in the Dungarvan network and b) cropped section of the periodogram of device A3, focussing on the frequency range 0.036 to 0.042 h^{-1} with the loss of component information at the 0.0417 h^{-1} frequency highlighted.*

The autocorrelation function of a timeseries is the correlation between two values in the timeseries, at a time *t* and time *t + 1*. This can be calculated for any number of subsequent time steps, or lags, where a lag *k* autocorrelation refers to the correlation between values that are *k* time periods apart. The lag autocorrelation plots for the Dungarvan $PM_{2.5}$ measurements (Figure 5.5) shows that the autocorrelation falls

immediately after $k = 1$. There is then a local peak when $k = 25$ for most data series, however some data series peak when $k = 24$. This indicates there are relatively strong periodic variations occurring in the data series at a frequency of around 0.04−0.0417 h^{-1}, which, again, is consistent with the typical diurnal profile seen in the data series. After the local maximum is reached, the autocorrelation values begin to decrease. However, on average, the lag autocorrelation begins to become insignificant (using approximate significance bounds given by $\pm 1.96/\sqrt{n}$, where n is the sample size (Shumway & Stoffer, 2017)) when $k > 32$, suggesting there may be meaningful variations in the data at frequencies beyond the traditional 24-hour cycle that are still significant and related to that cycle. This is especially important to consider when splitting $PM_{2.5}$ into just two fractions – local and non-local. It is imperative that the non-local component, relating to regional and seasonal variations, does not encompass contributions from the local emissions. The lag autocorrelations also demonstrate how $PM_{2.5}$ can vary within a town. Different locations show varying levels of autocorrelation with lagged values, with some showing higher consecutive correlations than others, when $1 < k < 25$. Namely, AP7 shows the smallest decrease in autocorrelation between those lag values, indicating less variation and relatively high correlation between lagged data points.

Similar trends in the hourly distribution of daily maximum $PM_{2.5}$ and lag autocorrelations are seen with data from the Cork City 2021 network (Figure A3.1, Figure A3.2); however, the autocorrelation value becomes insignificant (using the same significance bounds, adjusted for the different sample size) when $k > 35$. We expand beyond the 24-hour cycle exhibited by the typical wintertime $PM_{2.5}$ profile to include cycles with periods up to 32 hours to capture the slightly longer-term variations in the local emissions and reduce loss of important contributing components

in the 24-hour cycle, ensuring the analysis encompasses the full range of temporal dynamics in the $PM_{2.5}$ data. All frequency components below the location specific limits are classified as the non-local changes in $PM_{2.5}$ which may be attributable to seasonal changes or regional/transported emissions.

The local frequency limit is applied and the periodogram corresponding to every AQS data series in each network or reference data series are integrated across the entire periodogram and in the local frequency bin. The integrals are compared to give a percentage estimate, the local source contribution estimate (LSCE). Note that varying the cut-off frequency between 0.0313 h^{-1} (32 h) and 0.0286 h^{-1} (35 h) resulted in a LSCE difference no greater than 1% for all applications in this study. Both the long-term trends and shorter-term LSCE trends are investigated by decomposing yearly and monthly AQS data, respectively.

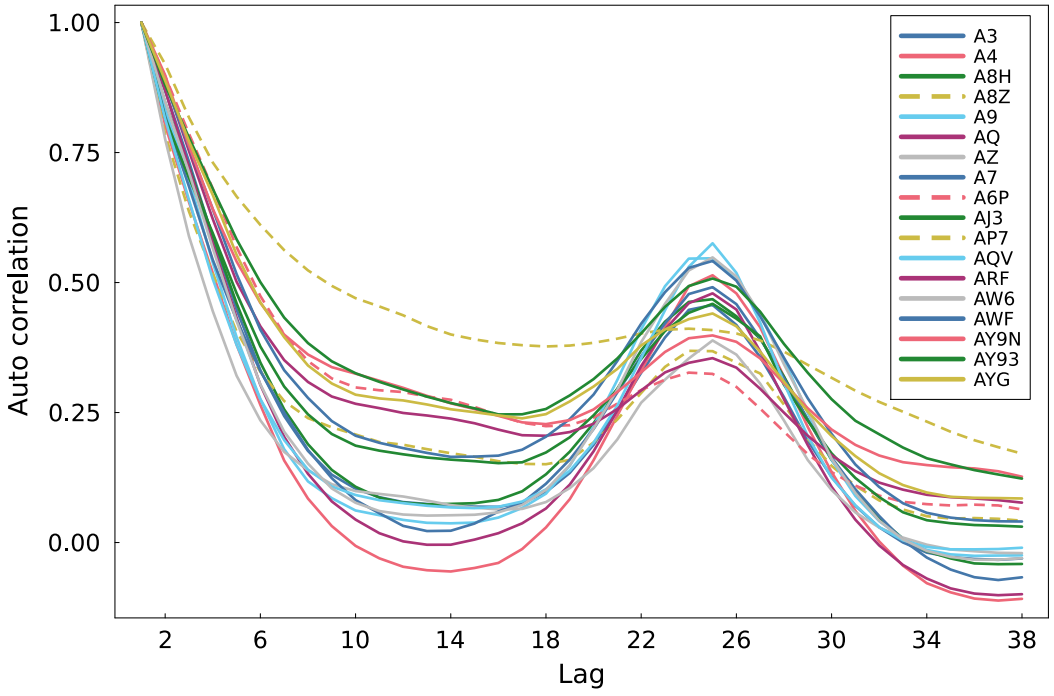


Figure 5.5: *Lag autocorrelation for the $PM_{2.5}$ data series from the Dungarvan sensor network. Solid and dashed lines indicate data series with a local maximum at k = 24 and k = 25, respectively.*

5.4 Results and Discussion

5.4.1 Sensor Performance Evaluation

The sensor-sensor co-location datasets are used to assess the inter-sensor variation, while data from sensor-reference co-location pairs are compared to understand the level of uncertainty in the LSCEs obtained from the sensor network datasets. To facilitate comparison between the different co-location assessments, February is taken as a typical winter month and July as a typical summer month.

5.4.1.1 Co-location of PurpleAir and Clarity Sensors

The co-located PA unit PA_epa and Clarity Node-S A8H_epa (uncorrected data) were both compared to each other and to the Palas Fidas 200 instrument for the month of February 2022. There was little variability observed between the measurements of PA_epa ($\bar{x}$ = 11 μg m^{-3}) and A8H_epa ($\bar{x}$ = 12 μg m^{-3}). $PM_{2.5}$ readings from the two sensors were highly correlated with each other (R^2 = 0.91) and both showed good correlation (R^2 values of 0.75) with the reference data ($\bar{x}$ = 9.46 μg m^{-3}). The sensor measurements also yielded low MAE (< 3 μg m^{-3}) and RMSE values (< 5 μg m^{-3}) with each other and with the reference measurements (Table 5.1). In terms of the data transformation and subsequent periodogram representation (Figure 5.6), the same dominant peaks are present in all three periodograms; however, the relative magnitudes are different. The sensor periodogram peaks are generally higher than that of the reference periodogram, due to the higher mass concentrations reported by the sensors compared to the reference measurements, i.e. the inherent bias in the sensor measurements. However, when comparing the relative areas under the total frequency

curve and the high frequency portion (> 0.0313 h^{-1}), the relative proportions of high frequency area to total area are very similar, and thus the LSCEs are also similar, at 61%, 68%, and 62% for the reference, PA_epa, and A8H_epa, respectively.

Table 5.1: *Statistical comparison of $PM_{2.5}$ signals from the co-located sensors, PA_epa and A8H_epa, and the Palas Fidas 200 reference instrument (01 February 2022–01 March 2022).*

Comparison	**R^2**	**Mean absolute error** **µg m^{-3}**	**Root mean square error** **µg m^{-3}**
Reference and PA_epa	0.75	2.5	4.7
Reference and A8H_epa	0.75	2.7	4.6
A8H_epa and PA_epa	0.91	1.8	2.7

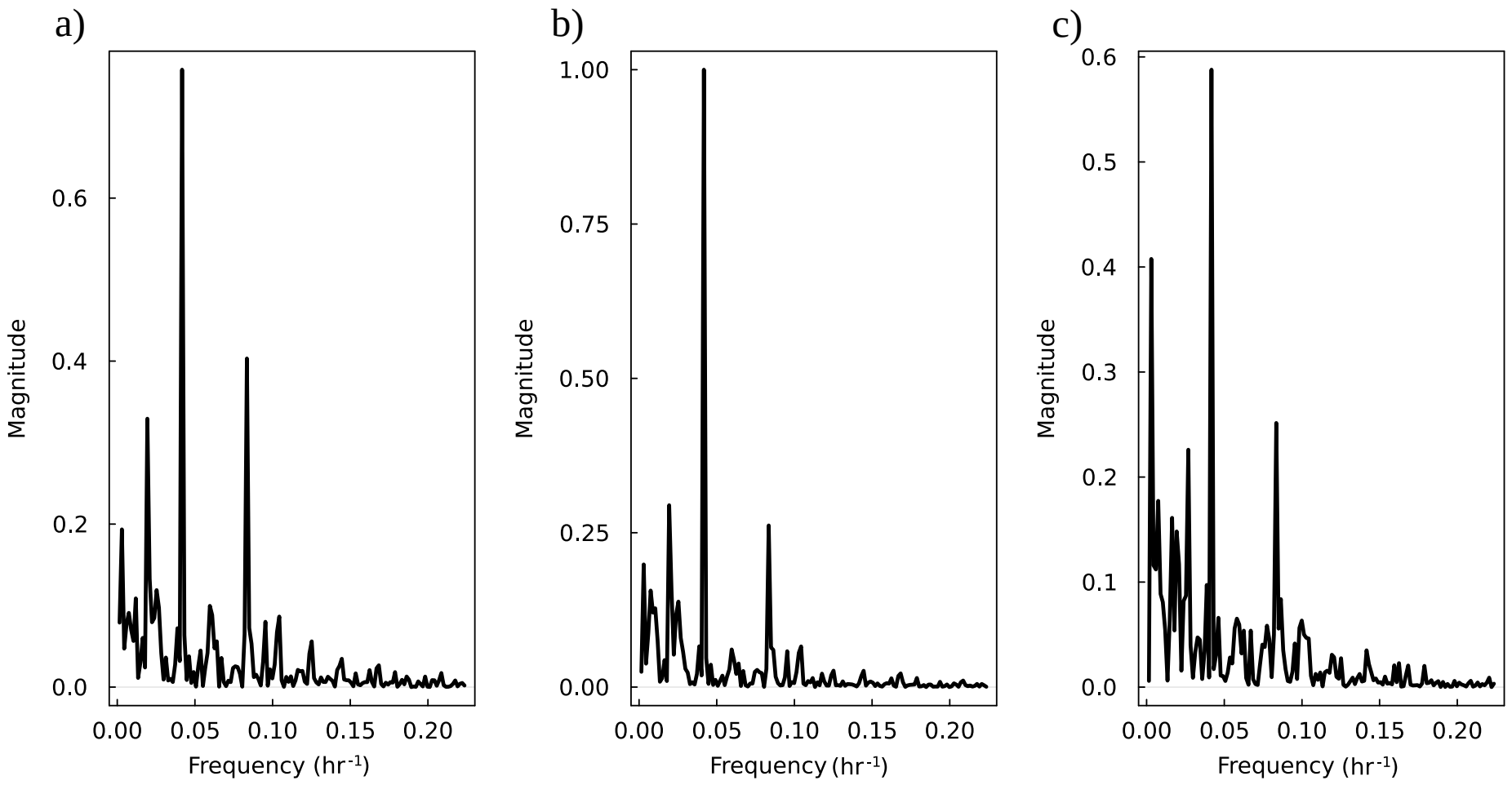


Figure 5.6: *Periodogram of $PM_{2.5}$ signals from a) the co-located PA sensor, b) co-located Clarity sensor, and c) the reference instrument in Edenderry in February 2022. Peaks are normalised against the maximum peak value.*

5.4.1.2 Co-location of PurpleAir Sensors

Reference Data Comparison

The yearly periodograms for the co-located PA device (uncorrected data) and BAM-1020 reference instrument in the Cork City 2021 network exhibit similar frequency profiles (Figure A3.3) with resulting local contribution estimates for the reference and the sensor periodograms of 38% and 36%, respectively (Table 5.2). Although the $PM_{2.5}$ measurements recorded by the sensor ($\bar{x}$ = 11 μg m^{-3}) were again higher than those measured by the reference instrument ($\bar{x}$ = 6.38 μg m^{-3}), the relative changes and the proportion of high frequency components remain similar, and thus the percentage of $PM_{2.5}$ due to local sources is comparable, albeit with a certain degree of uncertainty which must always be considered with AQS data.

Table 5.2: *Local source contribution estimates for co-located PurpleAir device and reference instrument across the entire year 2021 and February and July 2021 seasonal estimates.*

Data source	**Local source contribution estimate %**		
	Year	**July**	**February**
Uncorrected PurpleAir	36	28	49
Corrected PurpleAir	35	29	47
Reference	38	36	58

The same analysis was also conducted on the individual monthly data (February and July) to understand how the seasonal variations in LSCEs compare between the reference and the co-located sensor data. In both months, similar to the yearly assessment, the sensor reported higher area under both the entire periodogram and the high frequency portion, with mean $PM_{2.5}$ measurements of 11 μg m^{-3} and 9 μg m^{-3} for February and July, respectively. The reference instrument reported mean $PM_{2.5}$ measurements of 8.96 μg m^{-3} and 7.22 μg m^{-3} for February and July, respectively. The uncorrected sensor data reported good correlations with the reference data ($R^2 > 0.75$) for both months. The resulting percentages of the total signal attributable to local sources, were in good agreement with a difference of less than 10 percentage points (pp) in both cases. The reference and sensor reported LSCEs of 36% and 28% in July, respectively, and 58% and 49% in February, respectively.

Uncorrected vs Corrected Sensor Data

The LSCEs obtained for the uncorrected PA co-location data (Table 5.2) were compared with those derived from a calibrated version of the same data, as reported by Byrne et al. (2023) (see Chapter 3). The corrected sensor data yielded an overall mean $PM_{2.5}$ concentration for 2021 of 8 μg m^{-3}, with February and July mean values of 8 μg m^{-3}, and 7 μg m^{-3}, respectively, a decrease from the uncorrected mean values. The yearly LSCE for the calibrated sensor was 35%, with July and February values of 29% and 47%, respectively. Both the yearly and the monthly estimates showed minor differences (< 3 pp) between the corrected and uncorrected sensor data. Additionally, with both the uncorrected and corrected sensor datasets, the LSCE deviated further from the estimate derived from reference data during winter, with differences of 9 pp

and 11 pp for the uncorrected and corrected data, respectively. This is potentially due to increased uncertainties in the sensor data under high $PM_{2.5}$ loading during high pollution episodes caused by residential solid fuel burning. Nonetheless, the results show the ability for appropriately processed but uncorrected AQS data to effectively capture the real changes in $PM_{2.5}$ that occur, with the calibration resulting in negligible differences in the source contribution estimates.

Inter-sensor variation assessment

Additional PA devices (PA_1, PA_2, PA_3) were co-located for the period 01 October 2023–02 February 2024 to facilitate inter-sensor variation analysis of the uncorrected data. The devices reported similar mean values for the measurement period ($\bar{x}_{PA\_1}$ = 6 μg m^{-3}, $\bar{x}_{PA\_2}$ = 4 μg m^{-3}, $\bar{x}_{PA\_3}$ = 5 μg m^{-3}) and exhibited good correlations ($R^2 > 0.8$), with low MAE (< 2 μg m^{-3}) and RMSE (< 4 μg m^{-3}) values between one another (Table 5.3). Furthermore, the spectral analysis produced very similar LSCEs (68%, 66%, 70%), confirming minimal inter-sensor variations in the fluctuations for uncorrected $PM_{2.5}$ measurements.

Table 5.3: *Statistical comparison of hourly $PM_{2.5}$ signals from the co-located PurpleAir sensors in Cork City (01 October 2023–02 February 2024).*

Comparison	R^2	Mean absolute error µg m^{-3}	Root mean square error µg m^{-3}
PA_1 and PA_2	0.86	1.4	2.6
PA_1 and PA_3	0.84	1.2	2.4
PA_2 and PA_3	0.83	0.9	2.5

5.4.1.3 Co-location of Clarity Sensors

The 18 Clarity Node-S devices were co-located across multiple periods (Byrne et al., 2024, See Chapter 4), and 11 of them were operational for sufficient time to support reliable DFT analysis for producing LCSEs. Data from the 11 devices were initially compared and the uncorrected mean MAE (Table A3.1), RMSE (Table A3.2), and R^2 (Table A3.3) values from every sensor-sensor pair were 1 µg m^{-3}, 1.3 µg m^{-3}, and 0.97, respectively, showing very low inter-sensor variation between the devices. Additionally, LSCEs obtained from the sensor timeseries ranged from 30–36%, averaging at 33% (σ_{LSCE} = 2%).

5.4.2 Cork City 2021 Sensor Network

The uncorrected $PM_{2.5}$ measurements for the Cork City 2021 sensor network (summarised in Table A3.4) were subjected to DFT spectral analysis to determine

LSCEs on a monthly and annual basis. The yearly source contribution estimates, shown in Table 5.4, illustrate the extent to which the contribution of local sources can vary within the city, with 16 pp separating the highest and lowest values.

The monthly estimates for one device, CCC2, are shown in Figure 5.7. There is a strong seasonal variation, with LSCEs highest during winter (64% in November) and lowest during summer **(**15% in August). Interestingly, the second lowest monthly estimate of 17% was observed in March, which agreed very well with the value of 18% derived from the reference instrument data. The mean temperature recorded at Cork Airport in March was 7.19 °C, lower than that in November (8.12 °C) and similar to December (7.88 °C), so it is reasonable to assume that residential solid fuel burning was still taking place in March. Upon further investigation of the timeseries for this month, a strong elevation in the baseline $PM_{2.5}$ measurements was recorded by all AQSs in the city network for approximately 5 days between 28th February and 5th March 2021 (Figure A3.4), indicative of a regional or transported pollution event. The minimum uncorrected $PM_{2.5}$ recorded during this period was 21 μg m^{-3}, measured by both CCC13 and CCC4 at 02:00 on 1st March, although levels throughout the network did not fall below 34 μg m^{-3} for the remainder of the pollution event. The mean uncorrected $PM_{2.5}$ across the network for this period was 41 μg m^{-3}, compared to 12 μg m^{-3} for the rest of March. Meteorological analysis showed that the higher levels of $PM_{2.5}$ in March 2021 were associated with easterly winds (Figure 5.8), instead of the more localised winter $PM_{2.5}$ typically observed in this network (Figure 5.11). A Saharan dust storm affected the PM concentrations over Europe during this time (ECMWF, 2021), and while Ireland is generally not heavily influenced by dust storms, PM levels can be elevated when the storm is associated with easterly winds that transport polluted air from mainland Europe to Ireland (Hermes et al., 2024; Wang et

al., 2020). The unexpectedly low LSCE for March 2021 is therefore attributed to this regional pollution event which dominated the high $PM_{2.5}$ levels in March and significantly influenced the monthly average value.

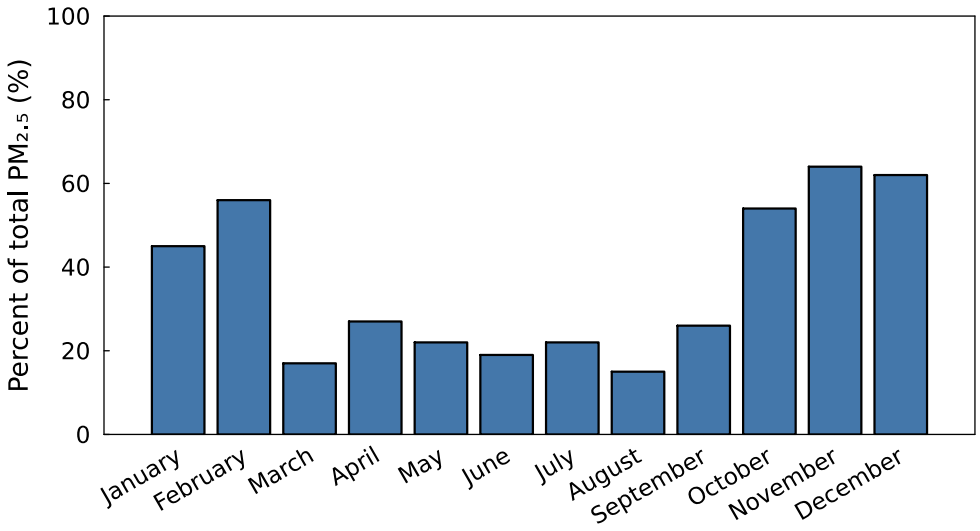


Figure 5.7: *Monthly local source contribution estimates for device CCC2 in the Cork City 2021 sensor network.*

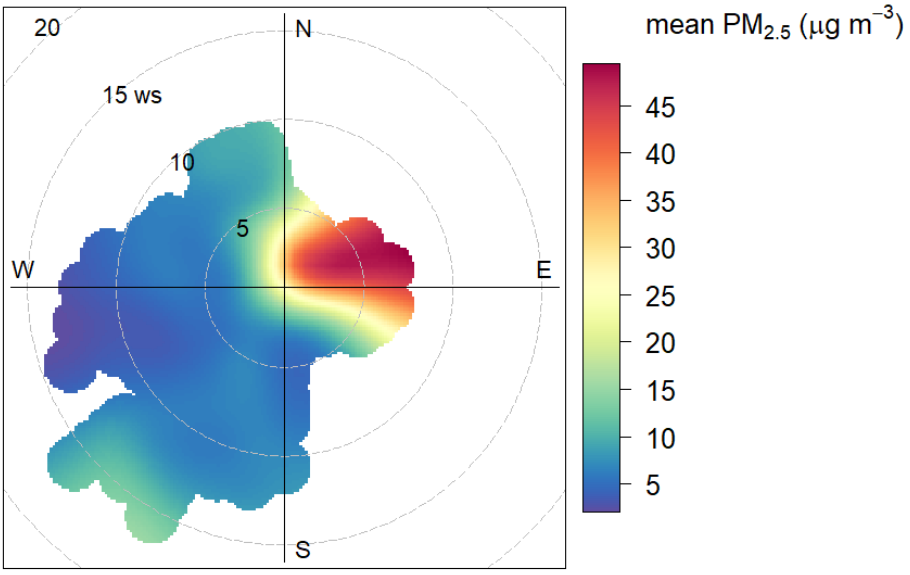


Figure 5.8: *Polar plot showing the average $PM_{2.5}$ measured across all sensors in Cork City as a function of wind speed and direction in March 2021, highlighting the increased pollution associated with slack easterly winds.*

Table 5.4: *Local source contribution estimates for the Cork City sensor network (01 January 2021 to 31 December 2021, and 01 January 2023 to 31st December 2023). Values are provided for devices that met the data coverage criterion.*

Device ID	Local source contribution estimate %	
	2021	**2023**
CCC2	36	47
CCC4	42	-
CCC5	40	-
CCC7	30	34
CCC6	-	29
CCC8	26	28
CCC9	37	42
CCC11	36	-
CCC12	42	31
CCC13	42	-
CCC14	-	55
CCC15	-	41
CCC16	-	39
MTU	31	-
UCC	34	26

5.4.2.1 Comparison of Local Source Contribution Estimates from the Spectral Analysis and Rapid Fluctuation Methods

A comparison of the LSCEs determined by the spectral analysis (Table A3.6) and rapid fluctuation methods in Cork City for June and December 2021 is shown in Figure 5.9. The same general trends were observed across the network for both analytical approaches, with more variation between devices during December than in June, highlighting the importance of considering wintertime $PM_{2.5}$ separately, as well as in relation to the annual $PM_{2.5}$ profile. The difference between the two methods was generally lower (< 10 pp) for December than in June.

The rapid fluctuation method focuses on periods in the data series dominated by rapidly fluctuating $PM_{2.5}$ levels, which are most likely caused by local solid fuel burning emissions (Byrne et al., 2023, see Chapter 3). In contrast, the spectral analysis method is based on the relative proportion of high-frequency (period ≤ 32 h) data fluctuations across the *entire* signal. Due to the difference in the approaches, a higher LSCE is expected from the spectral analysis method as any data variations in the signal are considered, not just those that occur during the periods characterised by very rapid fluctuations. For instance, local traffic sources may not lead to part of the signal being entirely dominated by rapidly fluctuating $PM_{2.5}$ levels and would not be captured as local emissions by the rapid fluctuation method, but with spectral analysis, any resulting signal variation would appear in the periodogram, thereby contributing to the local contribution estimate. This is particularly relevant for June where larger differences are observed between the methods. During this month, residential solid fuel burning would not dominate the $PM_{2.5}$ profile and less periods dominated by local emissions would be identified with the rapid fluctuation method, leading to a lower source contribution estimate than spectral analysis. In December, when the $PM_{2.5}$

levels measured are known to be dominated by the influence of solid fuel burning, this difference is less apparent at most locations, suggesting that the majority of local $PM_{2.5}$ is occurring during the periods captured by the rapid fluctuation method.

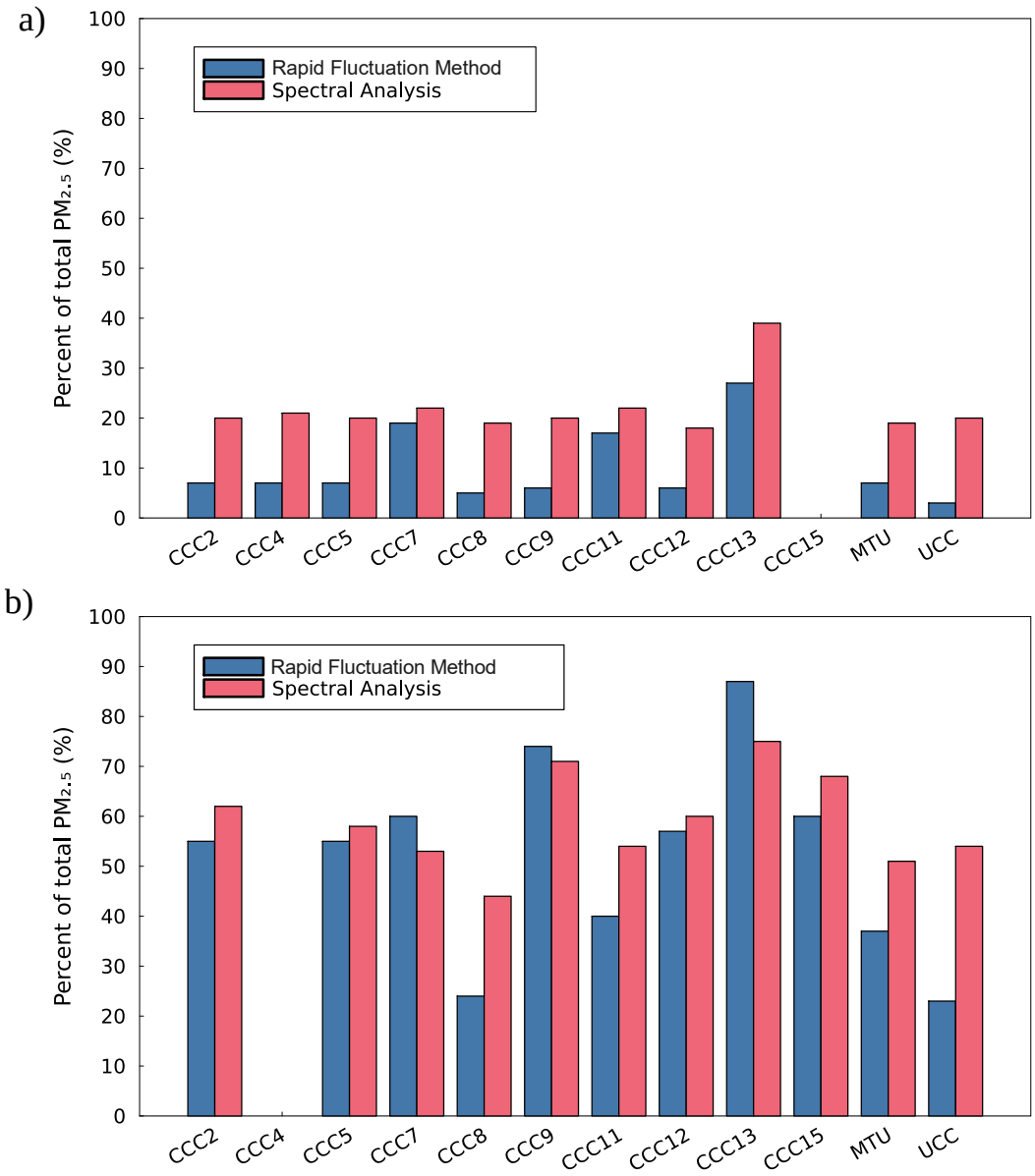


Figure 5.9: *Comparison of a) June 2021, and b) December 2021 local source contribution estimates between both methods for the Cork City sensor network. Values are provided for devices that met the data coverage criteria in December 2021. CCC15 did not record sufficient data in June, and CCC4 did not record sufficient data in December.*

5.4.3 Cork City 2023 Sensor Network

Spectral analysis of the Cork City 2023 sensor network data was conducted to produce average monthly and yearly contribution estimates of $PM_{2.5}$ from local sources. Larger average variation in mean $PM_{2.5}$ levels (Table A3.5) was observed in the winter months ($\sigma_{PM2.5}$ = 7 μg m^{-3}) compared to the summer month ($\sigma_{PM2.5}$ = 2 μg m^{-3}), reaffirming the increased variability experienced across a single town or city during winter. The degree of variation in the average $PM_{2.5}$ levels was more pronounced for 2023 than 2021, indicating more dissimilarities between locations in the later period. The seasonal variation in LSCEs in 2023 is highlighted in Figure 5.10 and Table A3.6. For June, all devices in the network reported LSCEs below 20%, with little variation between locations (σ_{LSCE} = 2%). For the six devices (UCC, CCC12, CCC8, CCC9, CCC7, CCC2) with sufficient data for comparison between June 2021 and 2023, all showed small decreases (< 10 pp). For December, the variation in the LSCEs (σ_{LSCE} = 17%) was much higher than in June, with a range of 45 pp. CCC6 had the highest contribution (73%), and although the same location was not included in the 2021 analysis, it is similar to the maximum LSCE observed in 2021 (75%, CCC13). The five devices common between December 2021 and December 2023 (UCC, CCC12, CCC8, CCC9, CCC2) all reported a decrease in LSCEs from 2021 to 2023, with a minimum decrease of 4 pp (CCC2) and a maximum decrease of 37 pp (CCC9).

The yearly source estimates for the 2023 sensor network (Table 5.4) show a yearly range of 29 pp in the LSCEs across the city. Of the six devices with sufficient data for both 2021 and 2023 yearly analyses, two (CCC12, UCC) reported a decrease in LSCE from 2021 to 2023, while four (CCC2, CCC7, CCC8, CCC9) reported an increase. However, in all cases, the changes were less than 10 pp.

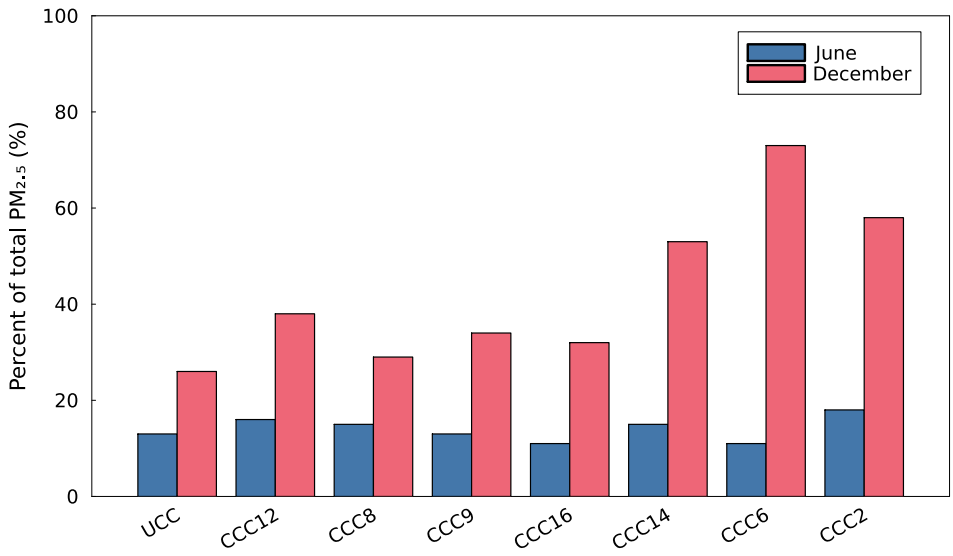


Figure 5.10: *Local source contribution estimates for June and December 2023 for the Cork City sensor network. Values are provided for devices that met the data coverage criterion for both months.*

The reductions in winter LSCEs observed at the five locations could possibly be linked to lower local emissions in December 2023 resulting from the new solid fuel regulations introduced in October 2022. However, there are other factors influencing ambient $PM_{2.5}$, such as meteorology, which need to be considered. Byrne et al. (2023) (See Chapter 3) showed that the highest $PM_{2.5}$ levels in Cork City during 2021 were largely associated with low wind speeds which limit atmospheric mixing and dispersion, thereby leading to a build-up of pollutants in the local area where they are emitted. The same observation was again evident in 2023, as shown in Figure 5.11, where polar plots for December 2021 and December 2023 are compared. Although higher mean $PM_{2.5}$ values were recorded in December 2023, the elevated levels remain concentrated around periods when wind speeds were low. The mean wind speeds recorded in the two months were similar (11.4 m s^{-1} for December 2021 and 12.4 m s^{-1} for December 2023), indicating that the average effect of atmospheric mixing and dispersion, would likely be similar for the two periods. Additionally, the diurnal wind

speed patterns are consistent between both years with an absolute average diurnal difference less than 1 m s^{-1} for all hours (Figure A3.5).

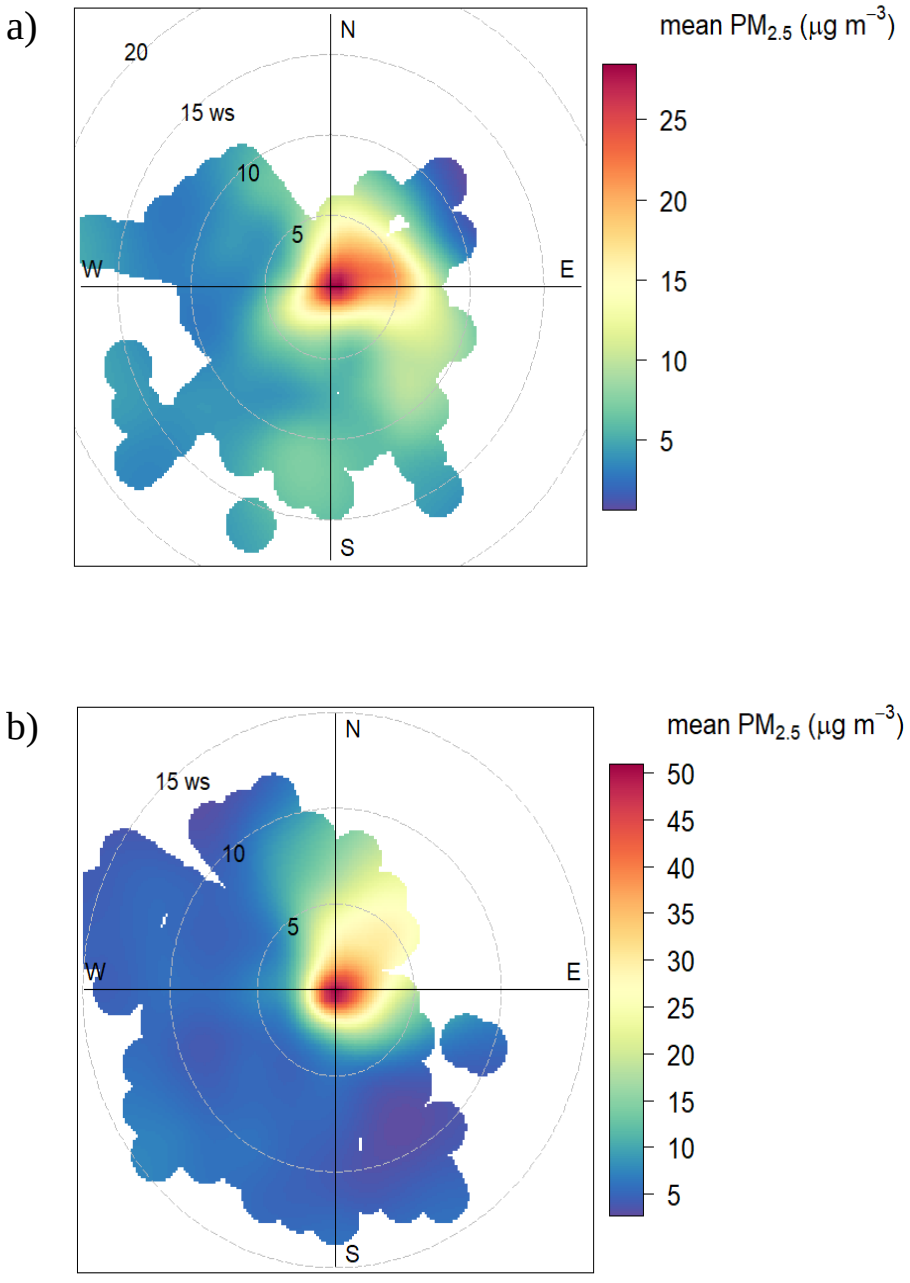


Figure 5.11: *Polar plot showing the average $PM_{2.5}$ measured across all sensors in Cork City as a function of wind speed and direction in a) December 2021 and b) December 2023. Note that the mean $PM_{2.5}$ scales are different on each figure and represent uncalibrated sensor values.*

A comparison of the maximum (13 °C, 12.9 °C), minimum (1.1 °C, −1.1 °C) and mean (7.9 °C, 7.7 °C) temperatures recorded at Cork Airport in December 2021 and 2023 shows little variation between the two periods. A temperature below 5 °C was measured in 14% and 17% of the hours in December 2021 and 2023 respectively. There was more variation between June 2021 and 2023, with maximum (21.8 °C, 22.4 °C), minimum (6.9 °C, 9.2 °C), and mean (13.5 °C, 15.7 °C) values showing slightly warmer conditions during 2023. $PM_{2.5}$ generally increases as the temperature decreases (Megaritis et al., 2014), and while the difference in temperature may, at least in part, contribute to the observed reduction in $PM_{2.5}$ between June 2021 and 2023, the overall effect is modest. Moreover, since the temperature profiles for December 2021 and 2023 were almost identical, temperature is highly unlikely to account for any observed difference in $PM_{2.5}$ between these years.

Since both wind and temperature conditions were very similar during December 2021 and 2023, it is tentatively proposed that reduced local emissions from residential solid fuel burning could be an important factor in the observed drop in $PM_{2.5}$ at the five locations in Cork City. While this could potentially be linked to the new legislation restricting the sale and use of the most highly polluting solid fuels across the country, it is noted that the reductions are not uniform across the sensor network and the LSCEs show a large degree of variation. Clearly, further work, such as detailed source apportionment studies (Ovadnevaite et al., 2021; Wenger et al., 2020), is required to fully assess the impact of the new solid fuel regulations on air quality in Ireland.

5.4.4 Dungarvan Sensor Network

As the measurements only spanned the period November 2022–May 2023 (summarised in Table A3.7), yearly LSCEs could not be obtained for Dungarvan, and May 2023 was taken to represent a summer month. February 2023 was selected to represent a typical winter month to maximise data coverage across all sensors as some devices (AP7, AY93, AY9N) were not deployed until mid-January. Similar to the Cork City results, a greater degree of variation in the uncorrected mean values was observed during the winter month ($\sigma_{PM2.5}$ = 4 μg m^{-3}) compared to the summer month ($\sigma_{PM2.5}$ = 1 μg m^{-3}). While the uncorrected data does not reflect the absolute mass concentrations of $PM_{2.5}$ in the area, the winter variability does highlight the importance of capturing high spatial resolution measurements, even in a small town, to accurately reflect the relative exposure differences.

The LSCEs for each sensor in the Dungarvan network for February and May 2023 are shown in Figure 5.12. The range of LSCEs is lower in May (17 pp) compared to February (26 pp). The maximum estimate in May (A4, 41%) was also lower than that in February (AW6 and A9, 75%). In comparison to the same year in Cork City, the summer month mean LSCE is higher in Dungarvan ($\bar{x}_{LSCE,\ Dungarvan}$ = 31%, $\bar{x}_{LSCE,\ Cork}$ = 14%). The mean temperature recorded in Dungarvan in May with AY93 was 13.89 °C, while the mean temperature at Cork Airport in June was 14.10 °C, which does not represent a substantial difference in terms of the atmospheric effects on the local ambient $PM_{2.5}$. However, home heating habits are often influenced by ingrained habits and behaviours, such as the "weekend effect" of increased wood burning as secondary, or decorative, heating sources on the weekend observed in London (Fuller et al., 2014). Therefore, it is possible that local heating habits differ between any two locations regardless of the similar meteorological profiles.

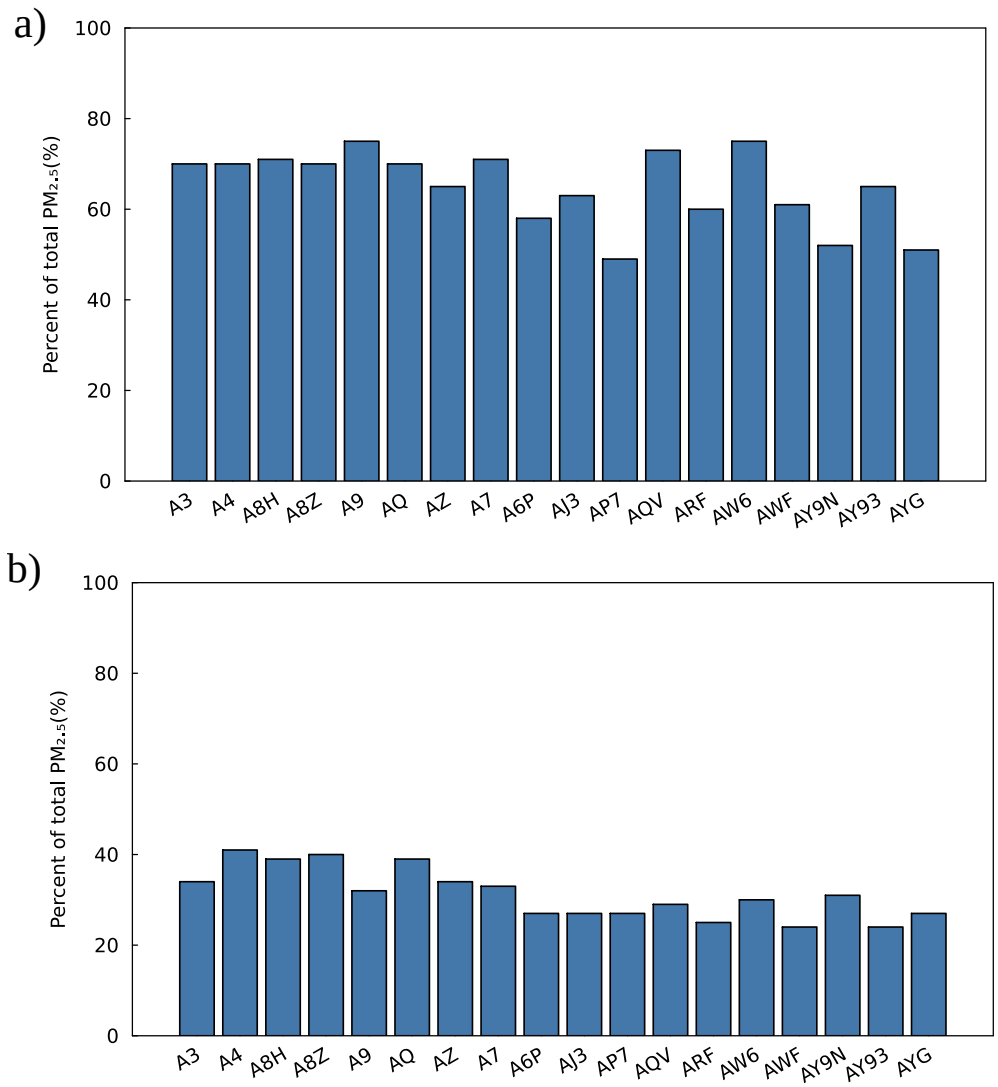


Figure 5.12: *Local source contribution estimates during a) February 2023 and b) May 2023 for the Dungarvan sensor network.*

While both Dungarvan and Cork City had similar maximum LSCEs during the winter month, the Dungarvan LSCEs are more uniform across the network ($\sigma_{LSCE,\ Dungarvan}$ = 8%, $\sigma_{LSCE,\ Cork}$ = 17%) with a higher mean estimate than in Cork City ($\bar{x}_{LSCE,\ Dungarvan}$ = 65%, $\bar{x}_{LSCE,\ Cork}$ = 43%). The uncorrected mean $PM_{2.5}$ during February (16 μg m^{-3}) is higher than that during December in Cork City (10 μg m^{-3}). The mean temperature recorded by AY93 in February was 8.64 °C, while the mean temperature recorded at Cork Airport in December was 6.82 °C. So, while the temperature was

slightly lower in Cork City in December, the average proportion of local $PM_{2.5}$ across the city was lower than that in Dungarvan in February, indicating potential differences in home heating habits or fuel choice between the two locations.

Another factor to consider is that Cork City is connected to the Irish gas pipeline network, while Dungarvan is not, resulting in higher availability of less polluting heating options for residents in Cork City. In 2022, 60% of private households in Cork City, used gas for central heating compared to just 0.6% in Dungarvan (CSO, 2022). Data from the Census of Population in Ireland shows an increase in the number of dwellings using natural gas and electricity for space heating between the 2016 and 2022, and a reduction in the number of houses relying on coal and peat/turf (CSO, 2022), indicating a general shift towards cleaner space heating options. However, a greater proportion of households in Dungarvan listed coal (10%) and oil (63%) as the fuel for central heating in 2022, compared to 1.6% (coal) and 16% (oil) in Cork City (CSO, 2022). These figures also translate to higher fuel use density in Dungarvan for coal (79 households per km^2) and oil (492 houses per km^2) compared to Cork City (7 households per km^2 using coal, 70 households per km^2 using oil). Given the significantly higher density of polluting fuel use in Dungarvan, it is unsurprising that the proportion of $PM_{2.5}$ originating from local sources is higher than in Cork City. Similarly high contributions of solid fuel burning to ambient $PM_{2.5}$ levels have been previously reported for other small towns in Ireland (Wenger et al., 2020). In addition, Cork City has had regulations to reduce the use of bituminous (or smoky) coal since 1995, while similar regulations only included Dungarvan from October 2022.

The lower variation in the LSCEs for Dungarvan implies that the local sources are picked up by the sensors relatively uniformly, suggesting more pronounced local contributions overall than in Cork City. Additionally, the disparity in the recorded

$PM_{2.5}$ concentrations, coupled with the more consistent local source influence found here could suggest that while local sources are similarly impactful across the network in terms of the proportion of total $PM_{2.5}$, there are differences in the absolute levels of these local emissions.

The meteorological conditions during the measurement period in Dungarvan showed little variation across the town, with maximum, minimum and mean temperatures of around 22.66 °C, −1.77 °C, and 9.54 °C, respectively, as previously reported (Byrne et al., 2024, See Chapter 4). The wind speeds and direction were also similar, indicating that the overall meteorological field did not differ greatly between the three sites. The relationship between the wind measurements (using AY93 as the reference) and the $PM_{2.5}$ measured by each sensor in the Dungarvan network in February 2023 is represented by the polar plots in Figure 5.13. Unsurprisingly, most devices show significant levels of $PM_{2.5}$ associated with low wind speeds due to accumulation of local pollution from solid fuel burning. Those with the strongest association between wind speed and $PM_{2.5}$ (A4, A7, A9, and AW6) had among the highest LSCEs. The devices with the lowest levels of $PM_{2.5}$ attributable to local sources (AP7, AY9N, AYG), show relatively lower levels of $PM_{2.5}$ at low wind speeds, indicating a weaker influence of local emissions from solid fuel burning.

Interestingly, all devices in the network also show higher mean $PM_{2.5}$ values associated with strong south-easterly winds. Upon further investigation, the meteorological data from AY93 yielded additional information. Prior to 11th February, multiple high $PM_{2.5}$ episodes (> 50 μg m^{-3}) coincided with low wind speeds (< 3.6 m s^{-1}), contributing to the localised accumulation of pollution shown in Figure 5.13. For approximately 2.5 days (12 February 2023–15 February 2023), consistently high $PM_{2.5}$ levels (> 18 μg m^{-3}, uncorrected) were recorded (Figure A3.6), along with

higher wind speeds and a change from north westerly to south easterly winds. During this period of increased $PM_{2.5}$, the pressure dropped from higher, stable levels (> 1030 hPa) to a lower pressure system indicative of more dynamic weather conditions. High wind speeds can increase horizontal and vertical mixing, resulting in more ideal conditions for both the transport of pollutants from other areas and the resuspension of local particles.

While the adverse weather conditions likely caused particles to be transported to Ireland, the duration of the event was relatively short-lived compared to the 5-day pollution event experienced by Cork City in March 2021. Moreover, the mean $PM_{2.5}$ (28 μg m^{-3}, AY9N uncorrected) recorded during this period was much lower than that recorded during the regional pollution event in Cork City (44 μg m^{-3}, CCC2 uncorrected). In addition, the $PM_{2.5}$ levels during the event in Dungarvan were more variable ($\sigma_{Dungarvan}$ = 17 μg m^{-3}, σ_{Cork} = 11 μg m^{-3}) indicating an overall lower and less consistent elevation in $PM_{2.5}$ levels, with multiple strong evening $PM_{2.5}$ peaks contributing to the local frequencies. The event in Cork City also dominated the high $PM_{2.5}$ measurements in March 2021, whereas the event in February 2023 was not responsible for the majority of high $PM_{2.5}$ measurements during that month in Dungarvan as higher $PM_{2.5}$ levels were recorded previous to this event at the beginning of February and again towards the end of the month. This is also evident from the $PM_{2.5}$ polar plots (Figure 5.8, Figure 5.13). During March 2021 in Cork City, there is relatively low levels of $PM_{2.5}$ associated with low wind speeds. However, in February 2023 in Dungarvan, the majority of locations still experienced significantly elevated $PM_{2.5}$ levels at low winds speeds in addition to the $PM_{2.5}$ associated with south-easterly winds.

.

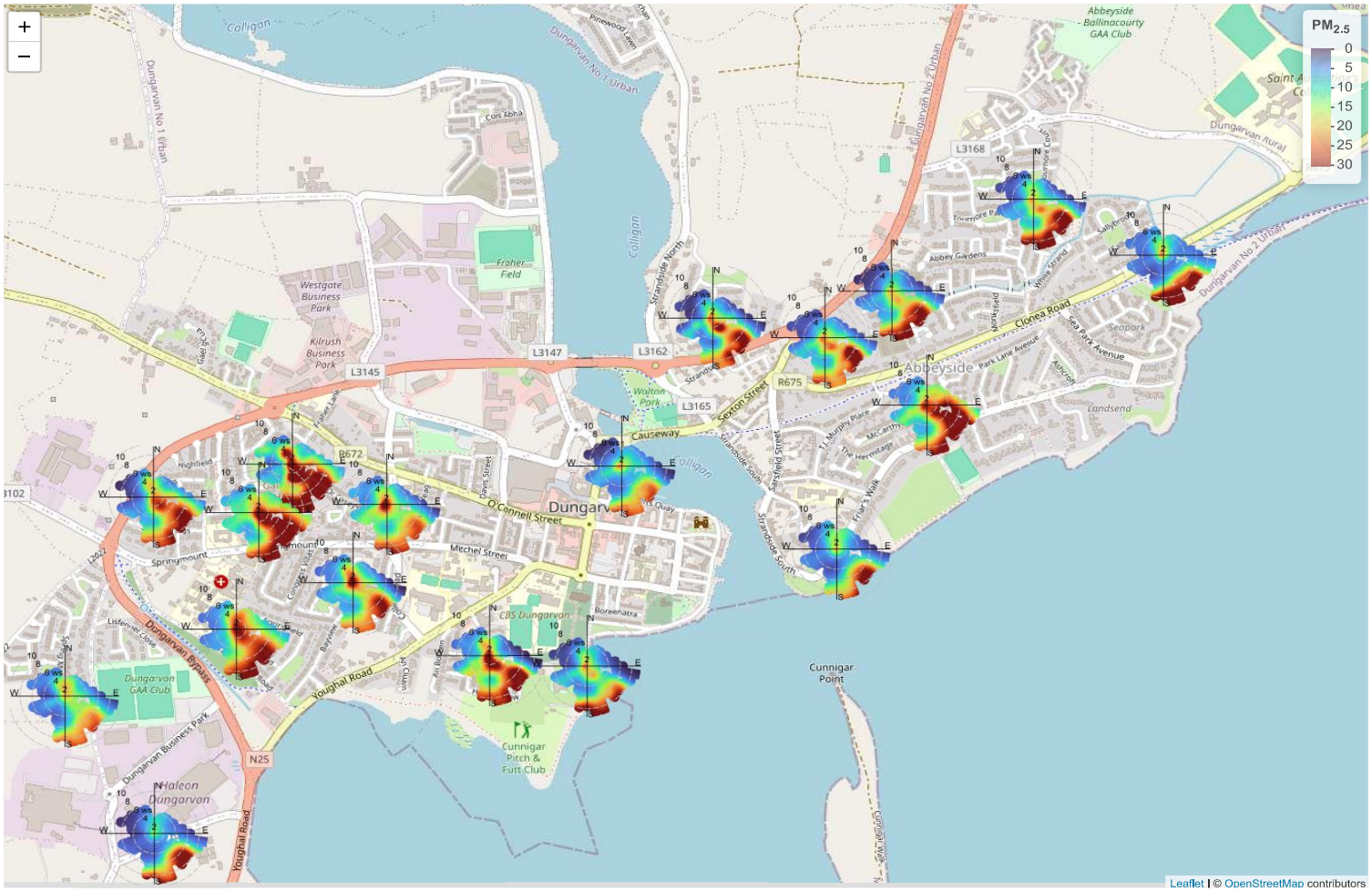


Figure 5.13: *Polar plots showing the average $PM_{2.5}$ measured by sensors in the Dungarvan sensor network as a function of wind speed and direction (as measured by device AY93) for February 2023.*

5.5 Limitations

The main limitation of this study is the lack of relevant chemical information about the ambient $PM_{2.5}$ to further validate the source contribution findings. Although $PM_{2.5}$ composition and source apportionment studies have previously been carried out in Cork City (Dall'Osto et al., 2013a; Healy et al., 2010; Hellebust et al., 2010; Kourtchev et al., 2011), they were conducted in the period 2008–2010. Despite this long gap between measurements, Byrne et al. (2023) (See Chapter 3) showed that the results of the rapid fluctuation method applied to the Cork City 2021 sensor data were generally in line with the reported $PM_{2.5}$ source apportionment estimates for Cork during similar summer and winter periods. This suggests that the principal $PM_{2.5}$ sources have not largely changed, even if they have reduced in intensity. Another limitation is the selection of the frequency cut-off to partition the local emission related frequencies. While the threshold proposed here reflects the characteristics and cycles of the data, which are in line with those of reference measurements, it would be instructive to repeat the analysis in different locations, ideally with additional source apportionment data, to test the generality of the limit.

5.6 Conclusion

The spectral analysis method was used to determine local source contribution estimates (LSCEs) of $PM_{2.5}$ using uncorrected data obtained from three sensor network datasets (Cork City 2021, Cork City 2023, and Dungarvan). The approach taken assesses the potential for the use of PM sensors, with a focus on $PM_{2.5}$ facilitated by access to $PM_{2.5}$ reference instruments. The study aims to support the broader adoption of scalable, cost-effective air quality monitoring initiatives, especially in underserved areas where reference-grade data is limited or not readily available. The low inter-sensor variation observed among the sensors, coupled with good correlations with co-located reference instruments underpins the proposed methodology. Moreover, the comparison of data transformations, frequency profiles, and LSCEs derived from both reference data and sensor data demonstrated that the estimates produced from both data types are, in general, quite similar (< 10 pp). Crucially, comparison of LSCEs from uncorrected and corrected sensor data revealed little variation in the outputs (< 3 pp), indicating the direct sensor measurements do not introduce significant biases into the LSCEs, thus affirming the robustness of the methodology.

The 2021 Cork City LSCEs were compared to estimates previously determined for the same network and measurement period using a different method (the rapid fluctuation method), offering valuable insights into the method outputs. Both method estimates were found to agree well in terms of the trends across the city network and relative differences between sensors. However, the spectral analysis method tended to give higher estimates during the summer, likely due to the spectral method capturing more of the local traffic emissions than the rapid fluctuation method. There is potential for both methods to be further developed to provide estimates for traffic-related emissions

based on the differences between both methods and to validate these estimates with chemical analysis of the ambient $PM_{2.5}$ or integrated with co-located black carbon source apportionment results. Additionally, the spectral analysis method does not rely on high-temporal resolution data and can be applied to hourly data obtained from a regulatory monitoring network for investigating the spatial differences in the contribution of local sources over a larger area, if resources for in depth chemical analysis of the ambient aerosol are limited.

To provide an updated network wide $PM_{2.5}$ source analysis for Cork City, the sensor network data was analysed for the calendar year 2023, providing an assessment of the change in local pollution experienced in common locations between 2021 and 2023. Higher variation was found in December 2023 between all locations in the network, with a range of 45 pp in reported LSCEs, compared to a range of 35 pp for 2021. All common devices showed a decrease in the LSCEs between December 2021 and December 2023, with a range in decrease of 33 pp, suggesting a possible disparity in the reduction of proportion of local sources across the city. Meteorological patterns during the two winter periods were not found to vary so significantly as to cause a substantial change in the local air quality, with higher levels of $PM_{2.5}$ associated with low wind speeds during both periods. Most devices available for both yearly analyses showed little change (< 10 pp decrease) in the percent of total $PM_{2.5}$ originating from local sources, while two devices (CCC12, UCC) showed slight increases. The larger difference in the reduction of local source contributions in December compared to that across the year once again highlights the seasonality associated with local emissions which arise mainly from residential solid fuel burning during the colder months and offers an early indication that the updated legislation on solid fuels may be helping to mitigate the impact of local wintertime $PM_{2.5}$ emissions, however the effect may not

be uniform across the city. Follow-up studies are required to investigate whether this represents a real and consistent change.

The analysis extended beyond Cork City to the Dungarvan network, which offered a spatial contrast to the Cork City network. Dungarvan is a small town on the south-east coast of Ireland. Higher levels of local pollution were reported for February than in May. The Dungarvan network had a lower range in winter LSCEs compared to Cork City in both 2021 and 2023, indicating there is less variation in the impact of local sources across Dungarvan than in Cork City. However, it is also a smaller town, so less local source variation may be expected despite varying levels of $PM_{2.5}$ reported.

With spectral analysis of the Cork City network, a regional pollution event due to the transport of Saharan dust in March 2021 was identified. A similar, but less impactful event was observed in Dungarvan in February 2023, where high $PM_{2.5}$ levels were associated with strong south-easterly winds, however, unlike the event in Cork City, it did not represent a significant increase in $PM_{2.5}$ compared to the rest of the month. So, while the method can respond to impactful reductions in the emissions contributions from local sources, it is robust to changes that do not significantly alter the relative proportion of local sources overall in a given period.

5.7 Appendix 3

Table A3.1: *Mean absolute error values (μg m^{-3}) for co-located Clarity Node-S devices (16 June 2023−17 July 2023).*

	A9	**AZ**	**A8H**	**A3**	**A7**	**AQ**	**A4**	**ARF**	**A6P**	**AWF**	**AQV**
A9	0	0.8	1	1.9	0.7	1.6	0.5	0.8	1.6	0.6	1.6
AZ		0	0.7	1.3	0.5	1	0.6	0.4	1	0.9	1
A8H			0	1.2	0.8	0.9	0.9	0.8	0.9	1.2	0.9
A3				0	1.5	0.7	1.8	1.4	0.6	2.1	0.6
A7					0	1.2	0.6	0.5	1.2	0.9	1.3
AQ						0	1.5	1.1	0.6	1.9	0.3
A4							0	0.6	1.5	0.6	1.6
ARF								0	1.1	0.9	1.1
A6P									0	1.8	0.5
AWF										0	1.9
AQV											0

Table A3.2: *Root mean square error values (μg m^{-3}) for co-located Clarity Node-S devices (16 June 2023–17 July 2023).*

	A9	**AZ**	**A8H**	**A3**	**A7**	**AQ**	**A4**	**ARF**	**A6P**	**AWF**	**AQV**
A9	0	1.1	1.2	2.1	1.1	1.9	0.8	1	1.8	1.1	1.9
AZ		0	0.9	1.5	0.7	1.1	0.8	0.5	1.1	1.4	1.1
A8H			0	1.6	1	1.2	1.1	1	1.2	1.6	1.3
A3				0	1.7	0.9	2.1	1.6	0.9	2.5	0.8
A7					0	1.4	0.8	0.7	1.4	1.3	1.5
AQ						0	1.7	1.2	0.9	2.2	0.5
A4							0	0.7	1.7	1	1.7
ARF								0	1.3	1.3	1.2
A6P									0	2.2	0.8
AWF										0	2.2
AQV											0

Table A3.3: *R^2 values for co-located Clarity Node-S devices (16 June 2023–17 July 2023).*

	A9	**AZ**	**A8H**	**A3**	**A7**	**AQ**	**A4**	**ARF**	**A6P**	**AWF**	**AQV**
A9	1	0.98	0.97	0.98	0.97	0.97	0.98	0.97	0.97	0.96	0.98
AZ		1	0.98	0.97	0.98	0.99	0.99	0.99	0.97	0.97	0.99
A8H			1	0.97	0.98	0.98	0.98	0.97	0.97	0.95	0.98
A3				1	0.97	0.97	0.98	0.97	0.97	0.96	0.97
A7					1	0.98	0.98	0.98	0.97	0.96	0.98
AQ						1	0.99	0.99	0.97	0.96	0.99
A4							1	0.99	0.97	0.97	0.99
ARF								1	0.97	0.97	0.99
A6P									1	0.95	0.97
AWF										1	0.97
AQV											1

Table A3.4: *Mean values (uncorrected) from the Cork City 2021 sensor network for the entire year period and for the summer and winter months.*

Device ID	Year	December	June
	µg m^{-3}	µg m^{-3}	µg m^{-3}
CCC15	-	10	-
MTU	6	6	5
CCC5	7	7	5
CCC8	7	6	7
CCC2	8	9	7
CCC4	8	-	7
CCC7	8	10	7
CCC11	8	9	8
UCC	8	8	7
CCC9	9	8	7
CCC12	9	10	8
CCC13	14	11	8

Table A3.5: *Mean values (uncorrected) from the Cork City 2023 sensor network for the entire year period and for the summer and winter months.*

Device ID	Year µg m^{-3}	December µg m^{-3}	June µg m^{-3}
CCC6	-	5	9
CCC7	-	16	-
CCC15	-	4	-
CCC8	5	4	9
CCC12	5	5	8
CCC2	6	5	9
UCC	7	6	12
CCC9	9	17	11
CCC16	9	14	11
CCC14	13	22	13

Table A3.6: *Local source contribution estimates for the Cork City sensor network (June 2021, June 2023 December 2021, and December 2023). Values are provided for devices that met the data coverage criterion.*

Device ID	Local source contribution estimate %			
	June 2021	**June 2023**	**December 2021**	**December 2023**
CCC16	-	11	-	32
CCC14	-	15	-	53
CCC6	-	11	-	73
CCC4	21	-	-	-
CCC8	19	15	44	29
MTU	19	-	51	-
CCC7	22	14	53	-
UCC	20	13	54	26
CCC11	22	-	54	-
CCC5	20	-	58	-
CCC12	18	16	60	38
CCC2	20	18	62	58
CCC15	-	12	68	-
CCC9	20	13	71	34
CCC13	39	-	75	-

Table A3.7: *Mean values (uncorrected) from the Dungarvan sensor network for the entire measurement period and for the summer and winter months.*

Device ID	Full period µg m^{-3}	February µg m^{-3}	May µg m^{-3}
AP7	11	9	11
AY9N	12	12	10
A8Z	13	12	10
A6P	13	12	9
AY93	13	13	11
AYG	13	12	9
AZ	14	14	9
ARF	14	12	10
AQV	15	15	9
AJ3	16	14	11
AWF	16	14	10
A3	18	18	9
A9	18	20	11
A8H	19	20	11
AQ	19	19	10
A4	21	22	12
A7	21	21	12
AW6	21	20	11

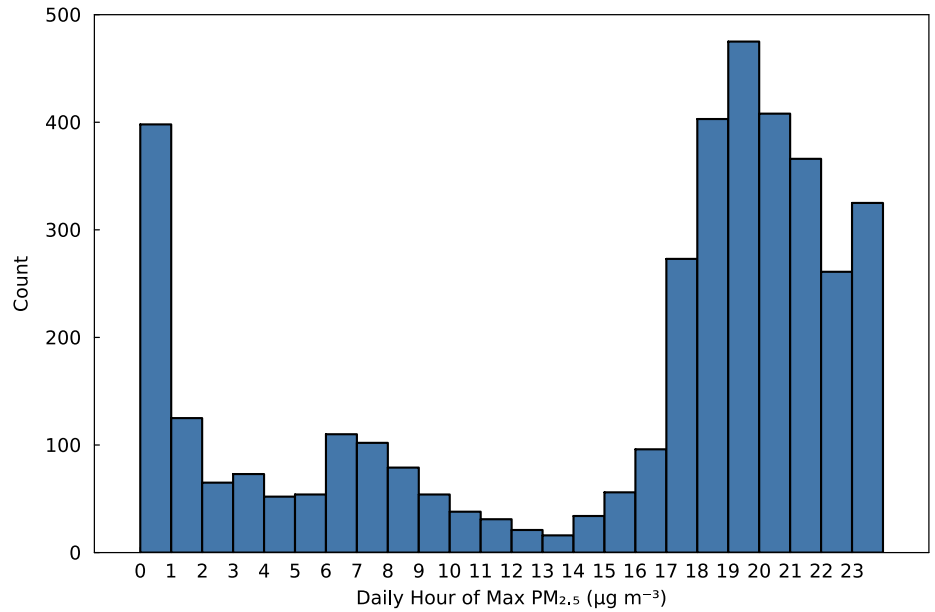


Figure A3.1: *Distribution of hour of daily maximum $PM_{2.5}$ measurement from all sensors in the Cork City 2021 sensor network.*

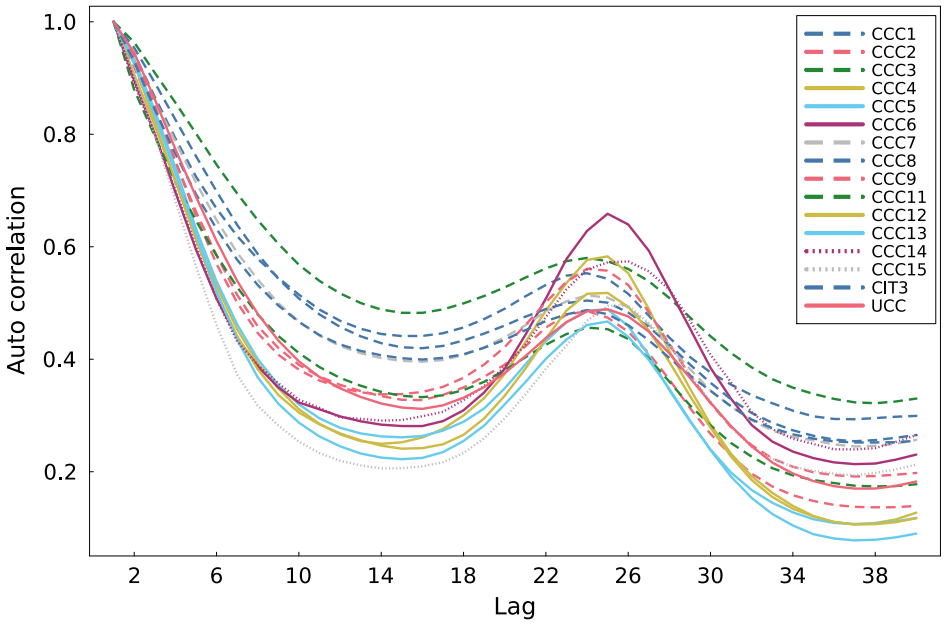


Figure A3.2: *Lag autocorrelation for $PM_{2.5}$ data series from the Cork City 2021 sensor network. Solid and dashed lines indicate data series with a local maximum at k = 24 and k = 25, respectively.*

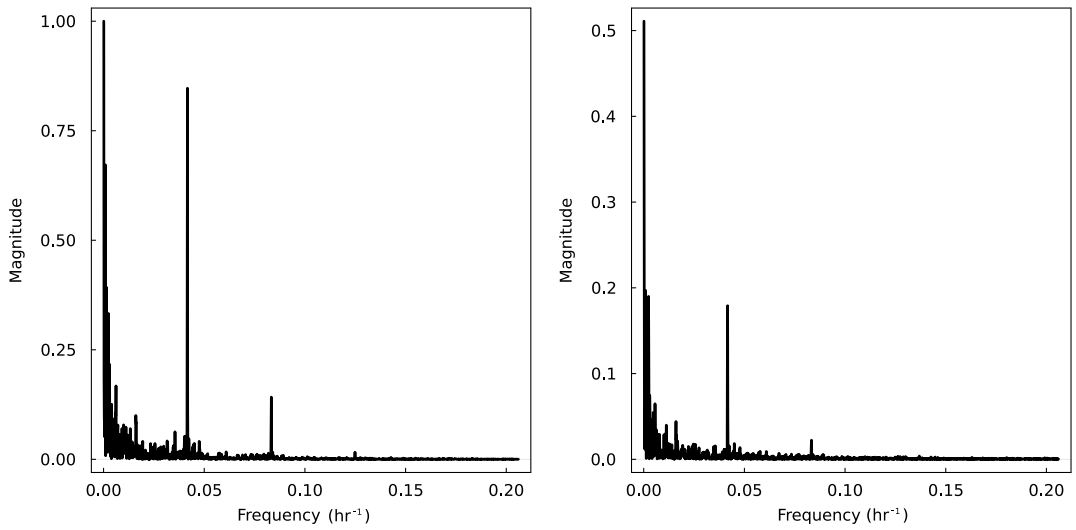


Figure A3.3: *Periodogram of $PM_{2.5}$ signals from a) the co-located sensor and b) the reference monitor in the Cork City network across the year 2021. Both periodograms are normalised against the overall maximum peak value.*

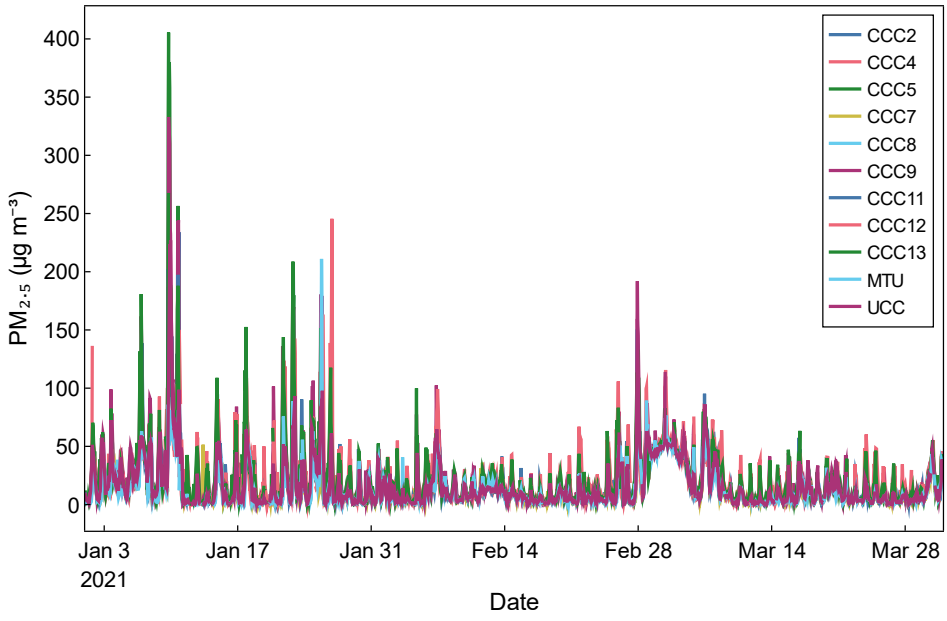


Figure A3.4: *Isolated concentration-time profiles from devices in the Cork City 2021 network from 01 January 2021 to 31 March 2021. An elevated baseline is noted during the first week in March 2021 for all devices.*

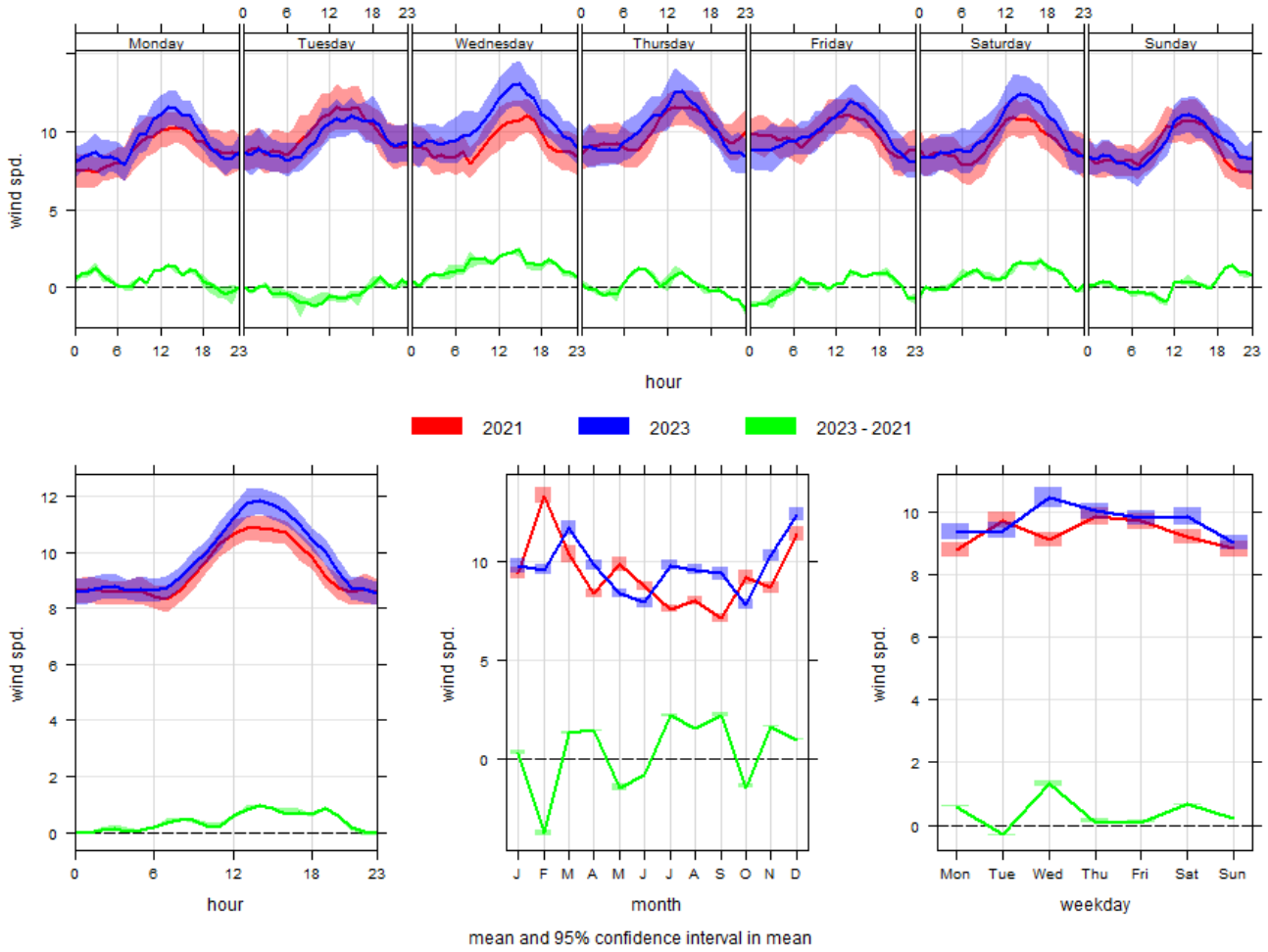


Figure A3.5: *Temporal variation (hour of day – day of week variation, mean diurnal variation, monthly variation, and day of the week variation) in the wind speed recorded at Cork Airport in 2021 and 2023. The red and blue lines correspond to mean wind speed measurements from 2021 and 2023, respectively, while the green lines indicate the mean differences between the two years.*

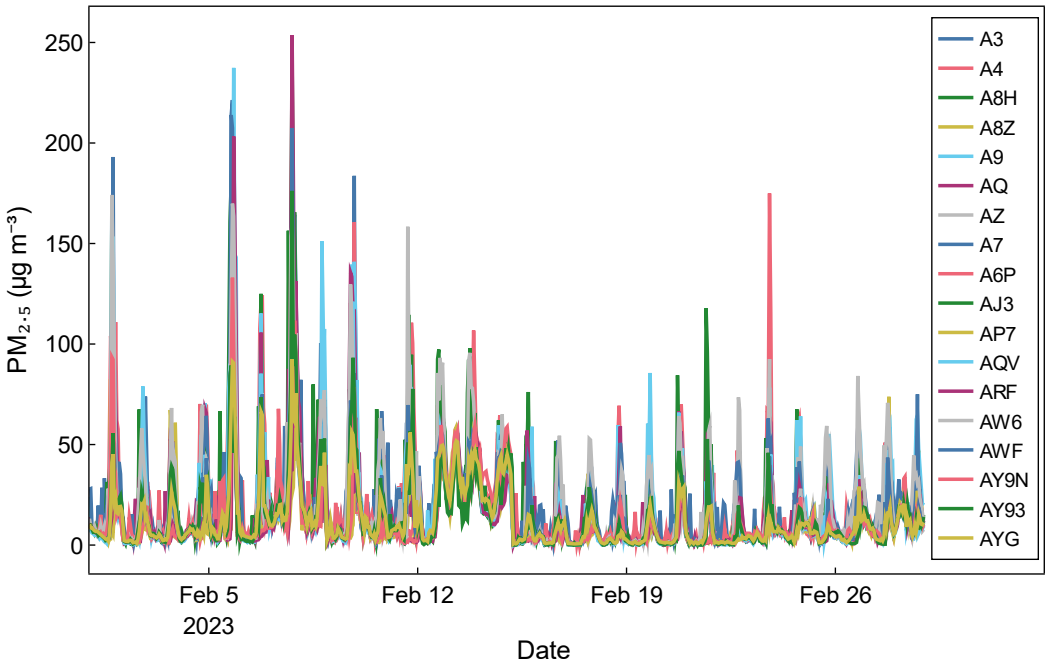


Figure A3.6: *Isolated concentration-time profiles from devices in the Dungarvan network from 01 February 2023 to 28 February 2023. An elevated baseline is noted during the first week in February 2023 for all devices.*

6 The Changing Nature of Black Carbon and $PM_{2.5}$ Emissions over Multiple Winters Characterised by Spectral Analysis and Aethalometer Data

Reproduced from the manuscript prepared for submission:

Byrne, R., Heffernan, E., Wenger, J. C., Hellebust, S. The Changing Nature of Black Carbon and $PM_{2.5}$ Emissions over Multiple Winters Characterised by Spectral Analysis and Aethalometer Data.

6.1 Abstract

This study assesses the temporal variability of wintertime particulate pollution in a small town in Ireland with a history of poor air quality due to residential solid fuel burning. Yearly winter $PM_{2.5}$ data from 2018 to 2024 and co-located equivalent black carbon (eBC) data from three winter periods were analysed to evaluate the impact of solid fuel regulations introduced in 2020. The relative changes in local and regional/background emissions were investigated by integrating temporal analysis of $PM_{2.5}$ and eBC with local source contribution estimates (LSCEs) derived from spectral analysis. A considerable reduction in mean wintertime eBC and $PM_{2.5}$ levels was observed between 2019 and 2024, with a decrease of 1.61 $\mu g\ m^{-3}$ for eBC and 8.27 $\mu g\ m^{-3}$ for $PM_{2.5}$. However, continued high maxima (> 170 $\mu g\ m^{-3}$) and periodic evening elevations in $PM_{2.5}$ concentrations during the later winters indicate continued reliance on solid fuel burning for heating. While the mean Ångström absorption

exponent (AAE) increased over three winter periods between 2019 and 2024, suggesting a shift away from fossil fuels like coal towards biomass-based fuel burning as a key emission source, the wide range of AAE values suggested contributions from multiple different emission sources. All periods showed a high proportion of $PM_{2.5}$ and eBC originating from local sources (> 50%). The 2023/2024 winter showed similar LSCE levels (52%) to previous years, suggesting that a general reduction in both local and regional pollution sources has occurred between 2019 and 2024. This study underscores the evolving nature of pollution in Irish towns, highlighting the influence of external factors and changing legislation on air quality trends. Despite overall reductions in ambient pollution, local sources continue to play a substantial role in $PM_{2.5}$ levels, emphasising the need for ongoing, detailed pollution monitoring and tailored mitigation strategies.

6.2 Introduction

Fine particulate matter ($PM_{2.5}$) continues to be recognised as the atmospheric pollutant with the greatest impact on human mortality and health (Pope et al., 2020; Pope & Dockery, 2012; Samoli et al., 2013). Black carbon (BC), which represents a fraction of $PM_{2.5}$ in the atmosphere, is primarily formed during the incomplete combustion of fossil fuels and biomass. These particles are characterised by distinct optical properties and their resistance to chemical transformation, and therefore, BC is often a good indicator of primary emissions (Bond et al., 2013).

Equivalent black carbon (eBC) refers to BC concentrations measured via optical methods such as with a multi-wavelength aethalometer. The difference in optical properties of eBC emissions from combustion of fossil fuels and biomass is utilized to apportion BC to these two sources (Herich et al., 2011; Sandradewi et al., 2008a, 2008b; Zotter et al., 2017). This source apportionment model makes use of the absorption Ångström exponent (AAE), or α, which describes the wavelength dependence of aerosol optical absorption by eBC (Bond et al., 2013). The AAE for BC depends on the nature and the origin of the particles and has been found to vary significantly between locations (Helin et al., 2021a). Fossil fuel combustion particles tend to have an AAE ≈ 1.0, i.e. no wavelength dependence, while the AAE of biomass combustion particles tends towards higher values (> 1.5), due to enhanced absorption in the UV range.

The result of the aethalometer source apportionment model is highly dependent on the selection of an appropriate AAE pair, α_{TR} for traffic-related fossil fuel emissions and α_{SF} for biomass or solid fuel emissions, and the assumption of fixed values for these parameters is a major limitation of the model. While some studies have used AAE

values cited in literature, others have determined these site-specific values by comparing the aethalometer data to collocated measurements such as ^{14}C or by using levoglucosan markers (Aas et al., 2018; Crilley et al., 2015; Sandradewi et al., 2008b; Zotter et al., 2017). It is often assumed that both contributions are predominantly local, but the model does not explicitly distinguish between local and non-local BC emissions.

In pollutant timeseries, concentration fluctuations often align with daily events like rush hour traffic and home heating emissions. Frequency analysis is useful for identifying these periodic components, decomposing the timeseries into cyclic frequencies that explain a significant portion of the variance of the signal (Warner, 1998). The Fourier Transform (FT) is used to transform data from the time domain to the frequency domain, allowing for the extraction of underlying patterns and helping to understand temporal variations (Percival & Walden, 1993). In urban areas, local $PM_{2.5}$ sources cause rapid, high-frequency fluctuations in measured PM levels, while regional sources and seasonal changes lead to less regular, gradual, and low-frequency changes. The Fourier transform of the timeseries signal can therefore be employed to distinguish between local and regional pollution sources based on the frequency characteristics, aiding in exposure assessment without the need for detailed chemical analysis (Frederickson et al., 2022).

In Ireland, solid fuel burning remains a part of the daily wintertime routine in many homes and studies have consistently identified these activities as a dominant contributor to particulate pollution across the country with a variety of solid fuels (wood, peat/turf, coal) being identified (Dall'Osto et al., 2013b, 2014; Eakins et al., 2022; Healy et al., 2010; Hellebust et al., 2010a, 2010b; Kourtchev et al., 2011; Lin et al., 2017, 2018, 2019b; Ovadnevaite et al., 2021; Wenger et al., 2020). High levels of

wintertime $PM_{2.5}$ are experienced in most towns and cities, with the mean annual (5 μg m^{-3}) and mean 24-hour (15 μg m^{-3}) WHO AQG values both exceeded at over 90% of the 79 $PM_{2.5}$ monitoring stations in the national air quality monitoring network in 2022 (EPA, 2023). To tackle the ongoing pollution caused by solid fuel burning, the Irish government introduced new regulations in October 2022 to restrict the sale, distribution, and use of certain solid fuels extending beyond larger towns and cities to the entire country.

This study aims to analyse $PM_{2.5}$ and eBC data collected during three distinct winter periods in Enniscorthy, Wexford, a small rural town in Ireland: December 2019–January 2020, November 2020–January 2021 and November 2023–January 2024. The first period was before the solid fuel regulations were in place in Enniscorthy, while restrictions to reduce the transmission of COVID-19 were in place intermittently during the 2020/2021 winter. The winter periods are compared via temporal trend analysis of the eBC and $PM_{2.5}$ concentrations with investigation into the contribution of different sources to these major pollutants. Source apportionment of eBC is assessed using the aethalometer model and local source contributions of $PM_{2.5}$ and eBC are estimated using spectral analysis of the timeseries.

6.3 Methodology

6.3.1 Campaign and Data Specification

Pollutant measurements were taken at the EPA air quality monitoring site on the grounds of the public library near the town centre of Enniscorthy (52.50035, -6.57054). Hourly $PM_{2.5}$ concentrations have been measured continuously by EPA Ireland with a FIDAS 200 since 2018 (Palas, 2024). Winter $PM_{2.5}$ data was obtained for the period of 01 December to 31 January for each year between 2018 and 2024. An AE33 multi-wavelength dual spot aethalometer (Magee Scientific, 2016) was used to measure BC over three winter periods – 01 December 2019 to 31 January 2020, 01 November 2020 to 21 January 2021, and 01 November 2023 to 31 January 2024.

The AE33 continuously measures the light attenuation of BC across 7 wavelengths (Table 6.1) every 60 s. The eBC measurements were averaged to achieve an hourly resolution. BC absorbs light strongly across the entire visible wavelength region; however, previous studies have shown the light absorption of other absorbing aerosols, such as brown carbon and mineral dust, is negligible at 880 nm, so the concentrations determined at this wavelength represent the reported eBC concentrations (Sandradewi et al., 2008a). The mass absorption cross-section (MAC) values relate the absorption efficiency of a particle sample per unit mass ($m^2\ g^{-1}$). The MAC values employed by the AE33 aethalometer for the evaluation of eBC concentrations are shown in Table 6.1. The wavelength specific attenuation of BC as measured by the aethalometer is converted into an absorption coefficient, b_{abs} (m^{-1}), which is then, using the pre-defined MAC values, converted into eBC mass concentrations at the various wavelengths.

Table 6.1: *Wavelengths and mass absorption cross section values used for source apportionment calculations.*

Channel	Wavelength nm	Mass absorption cross section (MAC) $m^2\ g^{-1}$
1	370	18.47
2	470	14.54
3	520	13.14
4	590	11.58
5	660	10.35
6	880	7.77
7	950	7.19

Meteorological data was measured by the Irish national meteorological service, Met Éireann, at Johnstown Castle, Co. Wexford. Historical hourly data is available publicly via www.met.ie. The weather station is located approximately 23.5 km southeast of the monitoring site in Enniscorthy.

6.3.2 Diurnal Profile Assessment Metrics

The Relative Peak Increase (RPI) metric (Equation 6.1) quantifies the relative magnitude of $PM_{2.5}$ concentration peaks in diurnal profiles. $PM_{2.5}$ often exhibits concentration peaks at specific times of the day, typically influenced by local emission sources (mainly solid fuel burning in Ireland) and atmospheric conditions (e.g. cold temperatures at night leading to more pollutant accumulation). Understanding the relative increases of the seasonal diurnal peaks is important for assessing the change

of impact of emission sources, and therefore local air quality, over time. The RPI metric allows for a standardised comparison across different periods, providing a clearer picture of the temporal variability and potential changes in source influence in the $PM_{2.5}$ data.

$$RPI = \frac{C_{peak} - C_{baseline}}{C_{baseline}} \quad \textbf{(6.1)}$$

where C_{peak} is the maximum diurnal concentration during peak hours (16:00–20:59), and $C_{baseline}$ is the average diurnal concentration during periods of low activity, or non-peak hours, (03:00–07:59) which represents the underlying concentration level.

6.3.3 Source Apportionment

Based on the wavelength dependent absorption efficiency of BC, the aethalometer source apportionment model uses linear regression to relate carbonaceous material originating from distinct sources to the absorption coefficients at different wavelengths (Sandradewi et al., 2008b). Sources of BC, traffic related (eBC_{TR}) and solid fuel (biomass) burning related emissions (eBC_{SF}), are differentiated using AAE values. The AAEs describe the spectral dependence of absorption – the exponent in the regression model of BC absorption as a function of wavelength (Grange et al., 2020). The aethalometer model is based on the principle that woodburning related BC emissions absorb well in the UV and lower visible wavelength range compared to fossil fuel combustion related BC particles which absorb strongly in the IR region (Sandradewi et al., 2008a; Wang et al., 2011a, 2011b).

A range of AAEs associated with specific source types have been observed, typically ranging between 0.8–1.1 and 1.6–2.5 for traffic and wood combustion aerosols, respectively, for absorption measurements at wavelengths of 470 nm and 950 nm, respectively (Kirchstetter et al., 2004; Schnaiter et al., 2003, 2005). Previous source apportionment studies in Ireland employing the aethalometer model have used values of $\alpha_{TR} = 0.9$ and $\alpha_{SF} = 1.68$, as proposed by Zotter et al. (2017) (Wenger et al., 2020). Sandradewi et al. (2008b) proposed values of 1.1 for traffic and 1.86 for wood burning following measurements in an Alpine valley in winter, while the real-time source apportionment analysis provided by the AE33 built-in software is based on values of $\alpha_{SF} = 1$ and $\alpha_{TR} = 2$. The two source aethalometer model has certain limitations in Ireland's multi-fuel environment, where a mix of solid fuels – wood, coal, and peat/turf – are commonly used. Although coal is a fossil fuel and its emissions are typically classified as such, in terms of source control it is often grouped with solid fuel and biomass combustion. Peat combustion produces a highly oxidised material which absorbs efficiently in the UV region, similar to wood burning emissions. Additionally, the AAE value for solid fuel emissions is less well-defined than for traffic related aerosols as it depends on the efficiency of combustion.

6.3.4 Spectral Analysis

6.3.4.1 Discrete Fourier Transform

The contribution of specific cyclic components to a temporal signal (i.e. measurements at regular time intervals) can be quantified with the transformation of the signal from the time domain into the frequency domain. The Discrete Fourier Transform (DFT) was applied under the assumption of equally spaced time intervals. Missing data were

interpolated; series endpoints were replaced with the overall mean, while internal gaps were filled using the mean of adjacent data points. Each data series was centred around zero by subtracting the overall mean. The periodogram, $P(v_k)$, was computed for a given timeseries (Equation 6.2).

$$P(v_k) = |f(k)|^2 = \left| \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{(-2\pi i v_k n)} \right|^2 \quad \textbf{(6.2)}$$

where N is the total length of the finite data series, x_n, $k = 0, 1, \ldots, N-1$, and $v_k = k/N$. The magnitude of peaks in the periodogram reflect the importance of the corresponding frequencies in the original signal. A larger peak indicates that the associated cyclic component accounts for a greater proportion of the signal's variance.

6.3.4.2 Local Source Contribution Estimates

The high frequency components in the periodogram can be integrated and compared to the total integrated signal. In other words, local and non-local $PM_{2.5}$ contributions were differentiated using periodogram frequencies, focusing on local emissions (e.g., traffic and residential heating). High-frequency components above 0.0313 h^{-1} (32 h) were classified as local contributions, accounting for diurnal patterns in emissions. Autocorrelation analysis of $PM_{2.5}$ concentrations (the correlation between two values in the timeseries, at a time *t* and time *t + 1*) (Figure A4.1) supported this frequency cut-off, revealing strong periodic variations around 0.04–0.0417 h^{-1} (24–25 h), consistent with diurnal patterns and a lag autocorrelation peak with the 24-hour cycle (Figure A4.1). Beyond 32-hour cycles, the lag autocorrelation becomes insignificant (assuming approximate significance bounds given by $\pm\ 1.96/\sqrt{n}$, where *n* is the

sample size (Shumway & Stoffer, 2017)), suggesting the importance of the frequencies below this limit in terms of the daily local emission cycle.

Following the above analysis of the high-frequency components, local source contribution estimates (LSCEs) were calculated from the periodogram by expressing the integral of values within the local frequency range (> 0.0313 h^{-1}) as a fraction of the total integrated signal. This approach captures the temporal dynamics of $PM_{2.5}$, distinguishing local from non-local sources, however it is recommended that the frequency cut-off is investigated further in future applications of the method. The length of a timeseries affects the DFT output as more long-term fluctuations can be captured with a timeseries covering a longer time period (Sevgi, 2007). To ensure comparable cycle lengths between data series, the DFT was carried out on data between 01 December and 21 January for each of the winter periods.

6.4 Results and Discussion

6.4.1 Meteorological Variations

The meteorological conditions did not exhibit notable variation between the three winter periods (Table 6.2). Similar mean wind speeds and variations were observed in each period, with a difference of 0.86 m s^{-1} between the highest (winter 23/24) and lowest (winter 21/22) mean values. Similarly, comparable mean temperatures and levels of rainfall were also recorded. The predominant wind directions were similar (Figure 6.1), with south-westerly winds occurring 50% or more of the time in all periods.

Table 6.2: *Summary of the mean and standard deviation of hourly meteorological parameters measured at Johnstown Castle, Co. Wexford.*

	Mean wind speed (σ) m s^{-1}	**Mean temperature (σ) °C**	**Mean precipitation (per hour) (σ) mm**
Winter 2018/2019	4.75 (2.52)	8.16 (2.52)	0.19 (0.63)
Winter 2019/2020	4.81 (2.45)	7.04 (3.58)	0.16 (0.53)
Winter 2020/2021	4.92 (2.24)	6.62 (3.58)	0.20 (0.62)
Winter 2021/2022	4.24 (2.34)	7.73 (2.71)	0.11 (0.52)
Winter 2022/2023	4.90 (2.33)	7.40 (3.66)	0.17 (0.61)
Winter 2023/2024	5.1 (2.31)	7.58 (3.19)	0.16 (0.54)

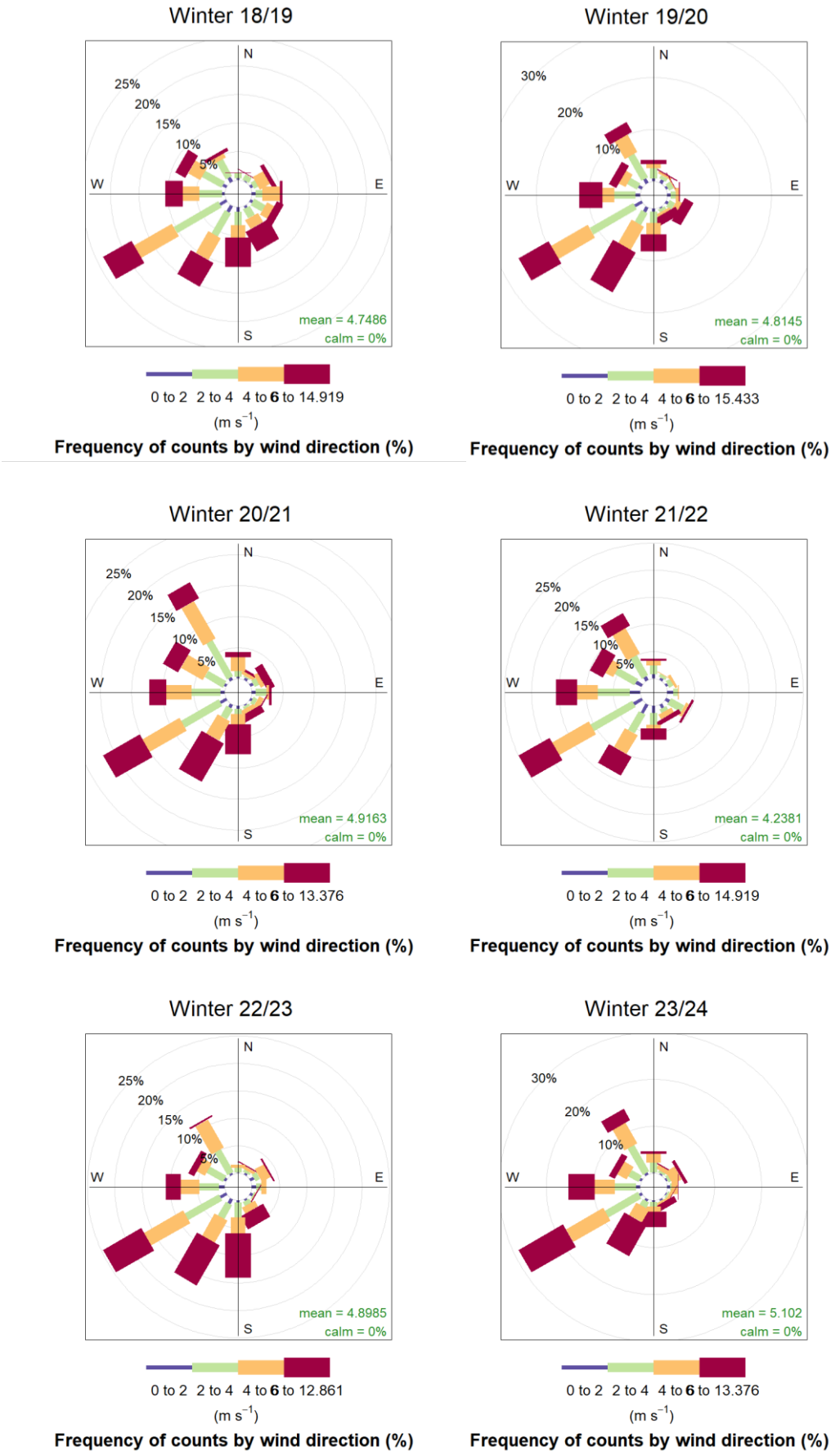


Figure 6.1: *Seasonal wind rose plots from wind measurements at Johnstown Castle, Co. Wexford.*

6.4.2 Temporal Variations in the Ambient Aerosol

6.4.2.1 $PM_{2.5}$

The temporal profiles of $PM_{2.5}$ in each winter between 2018 and 2024 are shown in Figure 6.2. In general, there appears to be a downward trend in the levels of $PM_{2.5}$ after winter 19/20 ($\bar{x}$ = 21.09 μg m^{-3}). There was a difference of 8.27 μg m^{-3} in the mean values between winter 19/20 and winter 23/24 ($\bar{x}$ = 12.82 μg m^{-3}). However, the mean $PM_{2.5}$ concentrations fluctuated during the intermediate years, although not exceeding the levels of those in the earlier winters, with mean values of 13.61 μg m^{-3}, 15.36 μg m^{-3}, and 14.30 μg m^{-3}, in winter 20/21, 21/22, and 22/23, respectively (Table 6.3). The highest maximum $PM_{2.5}$ recorded overall was in winter 19/20 (350.3 μg m^{-3}) and is almost double that of winter 23/24 (175.47 μg m^{-3}). However, a maximum of 311 μg m^{-3} was recorded in winter 22/23, indicating large, episodic $PM_{2.5}$ elevations continue to occur in the area, despite a reduction in mean values and updated legislation. During a 2015 measurement campaign in the same location, a maximum of 236.6 μg m^{-3} was recorded, although the measurement periods differed (08 January 2015–14 February 2015). The mean during that period in 2015 was higher than all periods reported in this work ($\bar{x}$ = 29.2 μg m^{-3}) (Buckley, 2019).

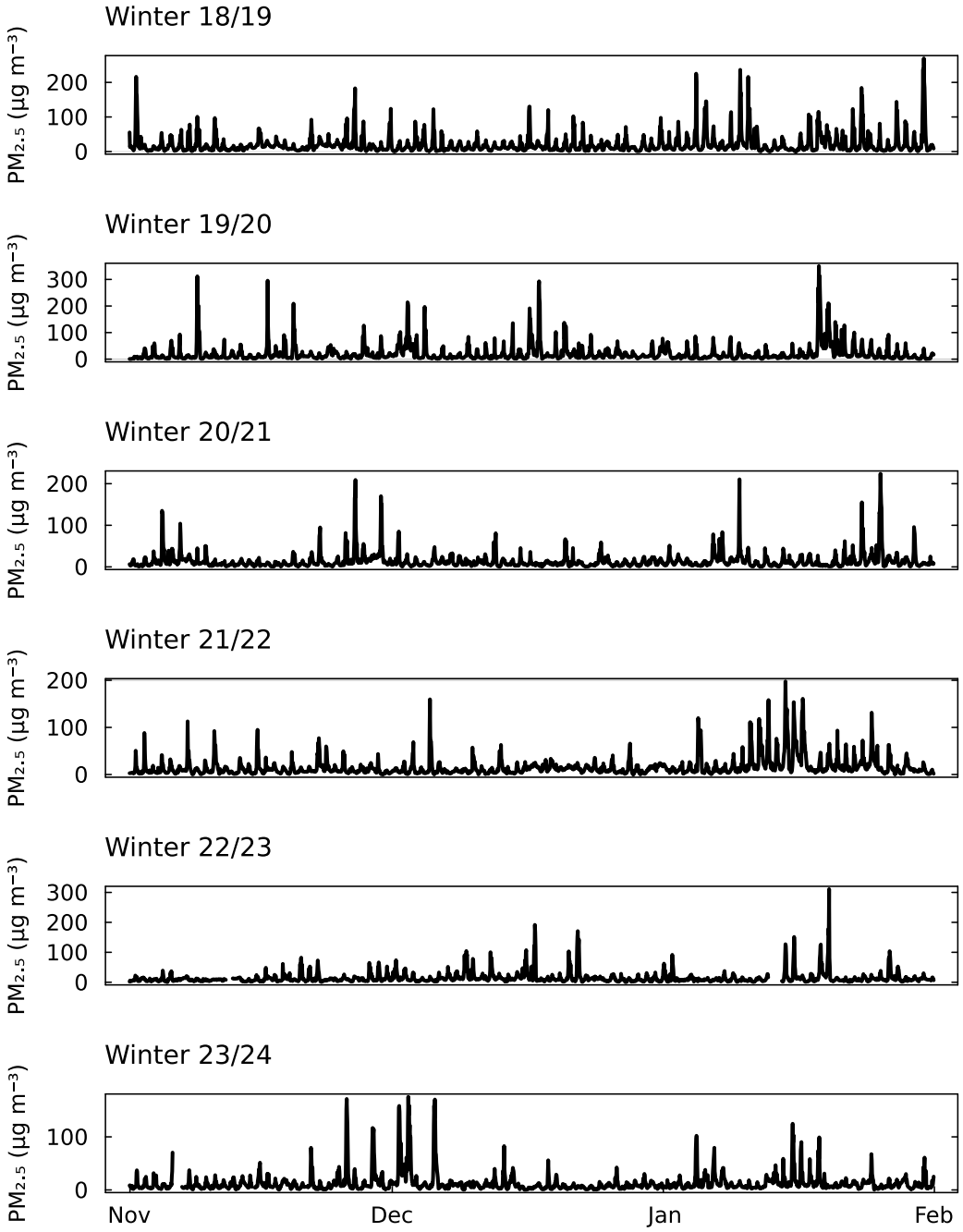


Figure 6.2: *Seasonal temporal profiles of $PM_{2.5}$. Note the $PM_{2.5}$ mass concentration axes scales vary in each yearly plot.*

Due to the large difference in mean levels between the first period and the later winter periods, the coefficient of variation (CV) is a more appropriate metric compared to the absolute variability reported by the standard deviation (Table 6.3). The CV for all winters (127–155%) shows all periods had high variability relative to the mean values,

suggestive of high episodic pollution events in addition to lower levels occurring at other times, which is indicative of emissions dominated with solid fuel burning activities which peak in the evening time.

While there is little variation in the minimum hourly $PM_{2.5}$ measured during each period and no minimum value exceeded 0.50 μg m^{-3}, a 27% decrease in the 25th percentile values for $PM_{2.5}$ from winter 18/19 (6.4 μg m^{-3}) to winter 23/24 (4.67 μg m^{-3}) was observed, indicating a decrease in baseline concentrations between those periods (Table 6.3). However, similar to the mean concentrations, the 25th percentiles fluctuated slightly during the intermediate years (4.90–5.70 μg m^{-3}) although they did not exceed the winter 18/19 value. The changes suggest the baseline reduction is not consistent and can vary from year to year, while the change in the 75th percentile values from the first winter to the latest (Δ = 7.61 μg m^{-3}) represents a strong decrease (36%) in the frequency of high concentrations, with a clearer decreasing trend presented in the intermediate years compared to the 25th percentiles. As the meteorological parameters measured during each period are generally similar leading to similar atmospheric conditions for pollution accumulation, the reduction in the extremity and frequency of severe pollution episodes is suggestive of effective implementation of fuel-related legislation leading to a drop in local emission related $PM_{2.5}$ concentrations during peak activity times. Additionally, this reduction has likely led to an improvement in regional particulate concentrations from the reduction of local emissions in different areas across the country leading to a generalised reduction in $PM_{2.5}$ levels over time.

The $PM_{2.5}$ measured in winter 21/22 and winter 22/23 were overall slightly higher and more variable than in winter 20/21 which incorporated periods of COVID-19 related lockdown measures. However, lockdown measures were still ongoing intermittently

during winter 21/22 and the mean $PM_{2.5}$ concentration and variation were more similar to the later periods. Additionally, the mean $PM_{2.5}$ concentration was lower in winter 23/24 than any other winter period, therefore the effect of the COVID-19 restrictions is uncertain. Studies have shown NO_2 concentrations were most severely impacted due to reduced numbers of vehicles on the roads, while the impact on $PM_{2.5}$ levels has been found to be less certain (Perillo et al., 2022; Spohn et al., 2022). Similar to the results suggested here for the 20/21 winter, Evangeliou et al. (2021) found reduced variability in $PM_{2.5}$ concentrations during lockdown periods in Ireland. Using a meteorological normalisation model, Heffernan (2022) estimated an average reduction of 37% in ambient $PM_{2.5}$ levels in Enniscorthy due to reduced activity as a consequence of the COVID-19 lockdown measures. However, as the ambient levels were similarly reduced in winter 23/24, it is likely that general legislative mitigation strategies, i.e. the ban on bituminous coal and other solid fuel restrictions being enacted in Enniscorthy in September 2020, also played a role in the reduction of $PM_{2.5}$ in winter 20/21.

All winter periods exhibit a rise in $PM_{2.5}$ diurnal concentrations in the evening time, indicative of increasing solid fuel burning emissions (Figure 6.3). However, the absolute diurnal peak in winter 18/19 (53.14 $\mu g\ m^{-3}$) and winter 19/20 (54.61 $\mu g\ m^{-3}$) is much larger than the peaks in all of the later winters, where the overall maximum was 34.06 $\mu g\ m^{-3}$ in winter 21/22. The RPI in winter 18/19 (6.20) and winter 19/20 (5.42) indicates a larger relative increase in the evening peak concentration compared to the later winters, where the maximum RPI was 4.21 in winter 20/21. Buckley (2019) reported a diurnal maximum of approximately 63 $\mu g\ m^{-3}$ in 2015, indicating a substantial reduction in the average $PM_{2.5}$ exposure has occurred over time.

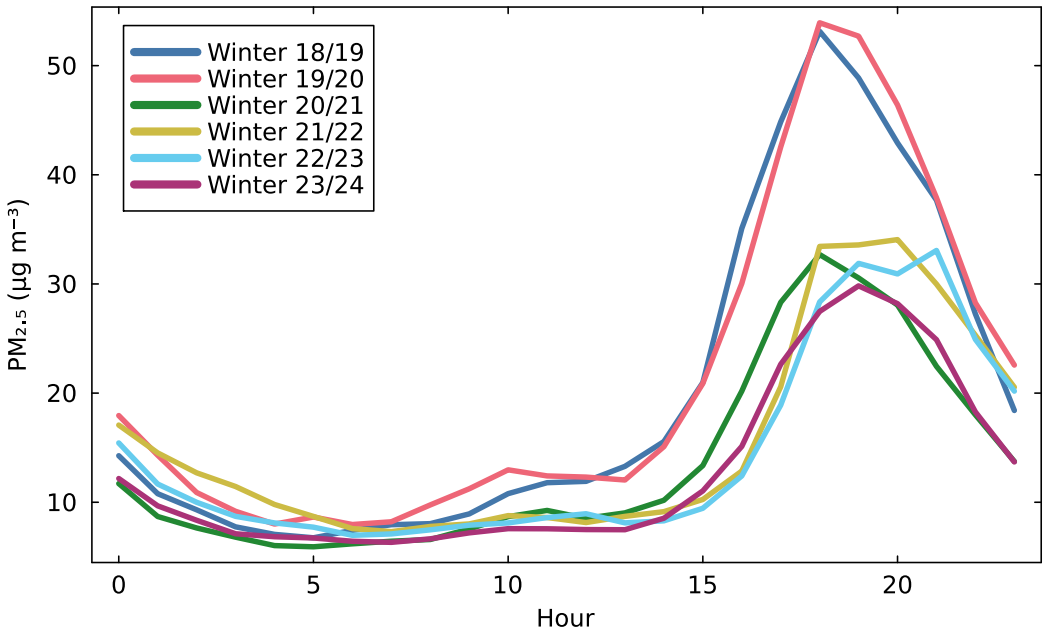


Figure 6.3: *Seasonal diurnal profiles of $PM_{2.5}$.*

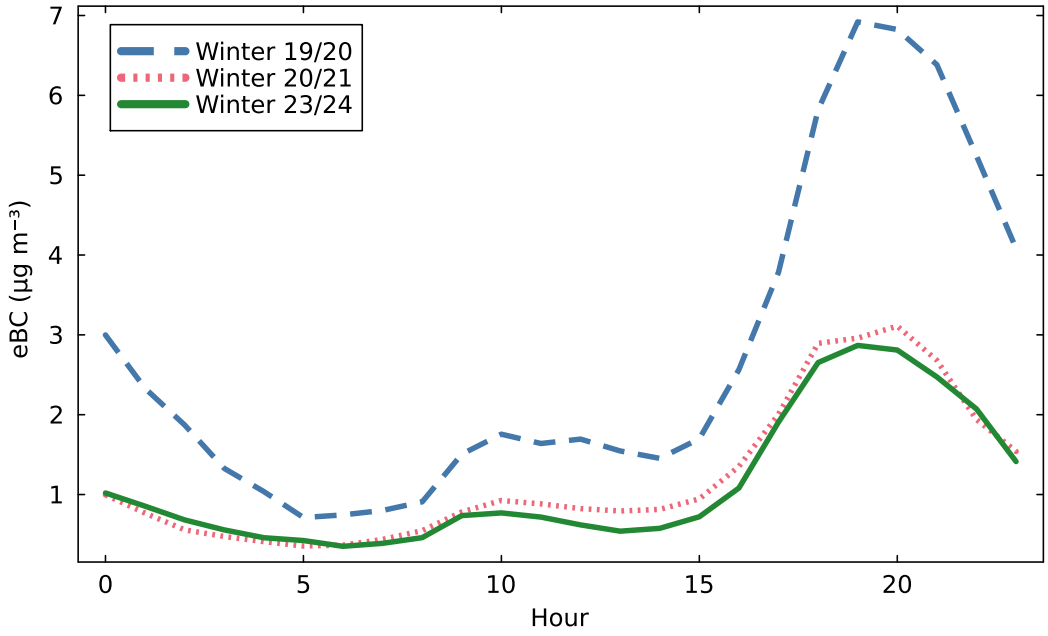


Figure 6.4: *Seasonal diurnal profiles of eBC from AE33 measurements.*

Higher $PM_{2.5}$ concentrations were primarily associated with relatively low wind speeds (2–5 m s^{-1}) in each period, when pollution dispersion is less effective, and

accumulation is likely to occur (Figure 6.5). In general, wind direction had less impact on the $PM_{2.5}$ concentrations compared to wind speed. Although concentrations of approximately 10–15 μg m^{-3} were observed at higher wind speeds too with $PM_{2.5}$ emissions originating from several directions relative to the sampling location, likely due to the influence of surrounding residential areas.

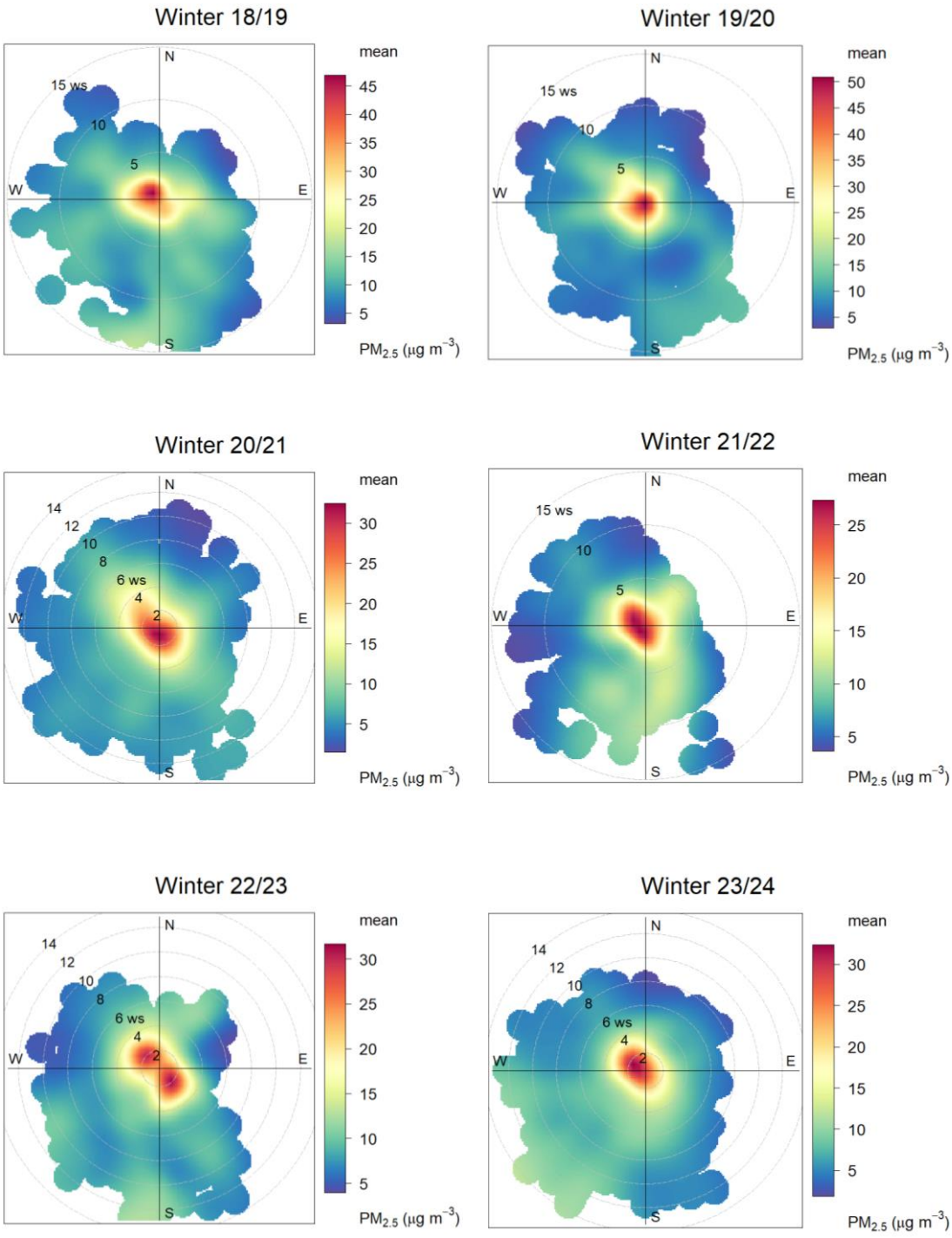


Figure 6.5: *Seasonal mean $PM_{2.5}$ concentration as a function of wind speed and wind direction.*

Table 6.3: *Summary statistics of hourly $PM_{2.5}$ concentrations for each winter period.*

	Mean (σ)	**Coefficient of variation**	**Maximum**	**Minimum**	**25th percentile**	**75th percentile**	**RPI**
	µg m^{-3}	**%**	**µg m^{-3}**	**µg m^{-3}**	**µg m^{-3}**	**µg m^{-3}**	
Winter 2018/2019	20.04 (26.61)	133	269.30	0.50	6.40	21.30	6.20
Winter 2019/2020	21.09 (32.69)	155	350.30	0.30	5.70	21.73	5.42
Winter 2020/2021	13.61 (18.76)	138	223.90	0.30	4.90	15.10	4.21
Winter 2021/2022	15.36 (19.55)	127	197.60	0.18	5.55	16.70	2.80
Winter 2022/2023	14.30 (19.18)	134	311.00	0.45	5.70	14.60	3.13
Winter 2023/2024	12.82 (18.3)	143	175.47	0.27	4.67	13.69	3.46

6.4.2.2 Equivalent Black Carbon

In general, a reduction in eBC concentrations (measured at 880 nm) occurred between 2019 and 2024, with an overall reduction in the mean eBC values of 1.61 μg m^{-3}, from 2.75 μg m^{-3} in winter 19/20 to 1.14 μg m^{-3} in winter 23/24 (Table 6.4). The variation in hourly concentrations was also found to decrease overall from a standard deviation of 4.74 μg m^{-3} in winter 19/20 to 2.02 μg m^{-3} in winter 23/24. A lower standard deviation was observed in winter 20/21 ($\sigma_{PM2.5}$ = 1.84 μg m^{-3}) compared to that in winter 23/24, similar to the $PM_{2.5}$ trends observed. The highest maximum eBC concentration was found in winter 19/20 (41.08 μg m^{-3}), with the other two periods reporting maximum concentrations around half of that at 23.1 μg m^{-3} and 20.38 μg m^{-3} for winters 20/21 and 23/24, respectively. The maximum eBC reported by Buckley (2019) was 35.2 μg m^{-3}, lower than winter 19/20, however the mean eBC reported by Buckley (2019) ($\bar{x}$ = 3.6 μg m^{-3}) points towards a general decreasing trend in ambient aerosol concentrations over time at the measurement site. The 25th percentile values for the eBC concentrations reduced from winter 19/20 (0.26 μg m^{-3}) to winter 23/24 (0.21 μg m^{-3}), representing a 19% reduction, while the 75th percentile values decreased by 66%, from 3.17 μg m^{-3} to 1.08 μg m^{-3} between winters 19/20 and 23/24, respectively.

Similar to the $PM_{2.5}$ measurements, the eBC diurnal concentrations from the three periods demonstrated similar patterns, with a rise in concentrations in the evening time exhibited in each winter period (Figure 6.4). Buckley (2019) observed a diurnal maximum of approximately 9 μg m^{-3} in 2015, higher than that found for winter 19/20 (~7 μg m^{-3}). Of the three winter periods, winter 19/20 exhibited the highest mean eBC concentrations for every hour of the day, including a more pronounced morning peak and a much higher evening peak. All three periods showed a similar relative increase

in diurnal evening concentrations, with the winter 19/20 RPI (6.49) slightly higher than winter 23/24, which had the lowest RPI (5.58).

The ΔC parameter ($BC_{370\ nm}$–$BC_{880\ nm}$) can offer an indication of the presence of specific solid fuel emissions due to the organic components of wood/peat burning particles absorbing light at 370 nm more effectively than at 880 nm (Wang et al., 2011a, 2011b, 2012). However, it is only an indicator of the presence of wood smoke and is not a quantitative measurement (Allen et al., 2004; Harrison et al., 2013). Figure 6.6 demonstrates a rise in ΔC concentration values in the evening time for each period, with winter 19/20 exhibiting the highest absolute increase. However, the RPI values during each period provide additional insights. The RPI for winter 19/20 (9.12) was lower than that measured for both of the later winters (10.68 and 13.27 for winters 20/21 and 23/24, respectively). The gradual increase in the relative size of the evening ΔC peak is indicative of a relative importance of specific emissions, namely wood/peat burning, to the overall BC profile.

The slight ΔC increase between approximately 10:00 and 13:00 in winter 19/20 suggests potential changes in heating habits from the first winter to the former two, where perhaps people are not relying on solid fuel burning for morning heating to the same extent as in the earlier winter. This was also observed in the $PM_{2.5}$ diurnal profiles, where a relatively larger morning peak was seen only in winter 19/20. Additionally, the ΔC diurnal pattern is similar to that observed with the eBC diurnal concentrations (Figure 6.4), suggesting that wood burning strongly influenced total eBC levels in all periods.

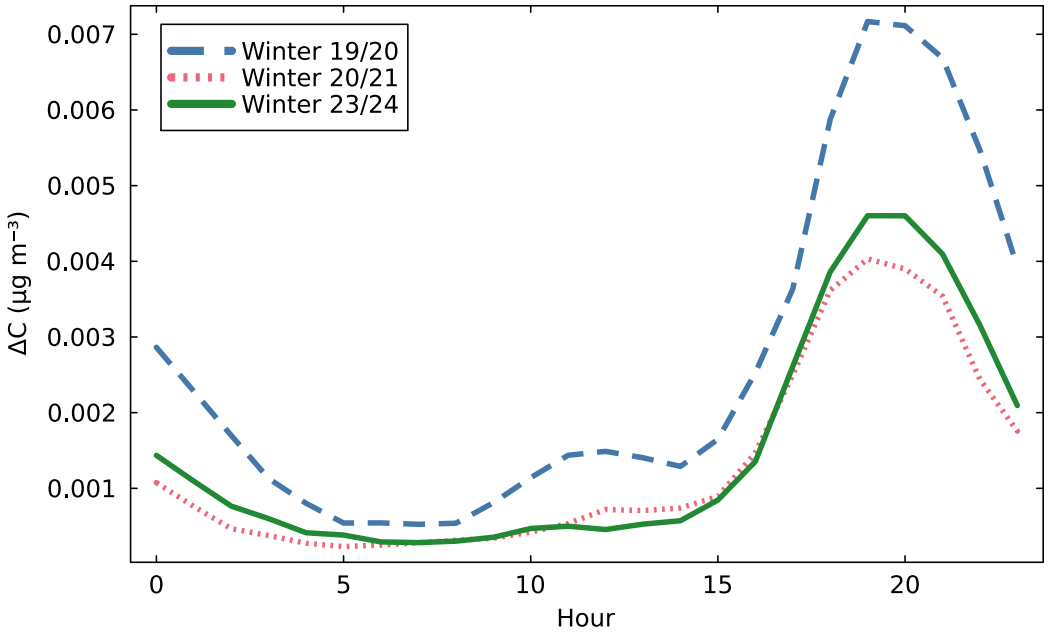


Figure 6.6: *Seasonal diurnal profiles of ΔC (BC1 – BC6), calculated from AE33 measurements.*

Table 6.4: *Summary statistics of hourly averaged eBC concentrations for each winter period.*

	Mean (σ)	**Maximum**	**Minimum**	**RPI**	**Contribution to total $PM_{2.5}$**
	μg m^{-3}	**μg m^{-3}**	**μg m^{-3}**		**%**
Winter 2019/2020	2.75 (4.74)	40.43	0.01	6.49	10
Winter 2020/2021	1.23 (1.84)	23.1	0.01	6.61	8.8
Winter 2023/2024	1.14 (2.02)	20.38	0.01	5.58	7.7

6.4.3 Source Apportionment of eBC

Apportionment of eBC to traffic and solid fuel burning is highly dependent on the choice of an appropriate AAE value for each category, solid fuel (α_{SF}) or traffic (α_{TR}), because incorrect values can result in significantly altered results, particularly the relative magnitudes of the diurnal profiles and the overall diurnal patterns (Harrison et al., 2013). Unfortunately, the spectral dependencies are somewhat site-specific due to the varying sources of BC emissions, namely the different fuel types used for home heating (Harrison et al., 2013; Helin et al., 2021b; Zotter et al., 2017). Without additional chemical data to determine the optimal site-specific pair, the applicability of the AAE pair cannot be directly assessed. However, other criteria can give an indication of the suitability of a specific pair of values to eBC concentrations, such as the number of negative concentrations in the eBC_{SF}/eBC_{TR} concentration split, which has been shown to correspond to the incorrect assignment of the AAE pair (Zotter et al., 2017), and an increase in the ΔC parameter is indicative of a relative increase in woodburning related emissions.

Analysis of the distribution of AAE values for each seasonal period can also indicate the dominant sources of BC during each measurement period and how these might vary between periods. For instance, AAE $\approx$ 1 suggests a dominant contribution from fossil fuel combustion, i.e. traffic related emissions, while a predominance of higher AAE values around 2 is indicative of a significant contribution from biomass or solid fuel burning. A wider distribution of AAE values within a single period can suggest a mix of sources contributing to BC. Similar distributions of AAE values across different periods suggest stable source contributions over time, whereas significant

changes in the distributions indicate shifts in the dominant sources of BC, such as a shift in the type of fuel dominating the combustion emissions.

6.4.3.1 Distributions of the Ångström Absorption Exponent

Figure 6.7 shows the distribution of AAE values during each winter period. While the mean and median AAE values for all three periods are generally similar ($1.53 \leq \bar{x} \leq 1.6$ μg m^{-3}), both mean and median values increase between each winter season. The same increasing trend is observed when the mean AAE from 2015 (1.45) is considered (Buckley, 2019). The majority of values in all cases are above 1, suggesting the BC in the ambient atmosphere was dominated by non-traffic related emissions during each period. However, as the AAE do not tend towards excessively high values, it is likely that traffic related emissions also contributed to the BC concentrations during each period, which is in line with previous source apportionment carried out at the same location (Wenger et al., 2020).

The AAE range occurring most frequently, as per the histogram counts, was 1.62–1.64 in winter 19/20, 1.58–1.6 in winter 20/21, and 1.7–1.72 in winter 23/24. The distributions corresponding to the first two winters are slightly skewed left (skewness = −0.56, −0.55 for winter 19/20 and winter 20/21, respectively) with a low kurtosis (< 3), indicating a tendency for the AAE values to be concentrated slightly towards higher values with extreme outliers being uncommon. The similarity in the variation of the distributions, σ_{AAE} = 0.33 μg m^{-3}, 0.35 μg m^{-3} for winter 19/20 and winter 20/21, respectively, suggests relatively stable source contributions and atmospheric conditions throughout each respective period. The final period has a similar level of skewness (−0.41); however, a higher variation is observed

(σ_{AAE} = 0.43 µg m^{-3}). The distribution also appears to exhibit a tri-modal distribution, indicating a wider range of sources contributing to increased absorption at lower wavelengths. The increased variation and tri-modality is suggestive of more variability in the emissions affecting BC measurements in winter 23/24.

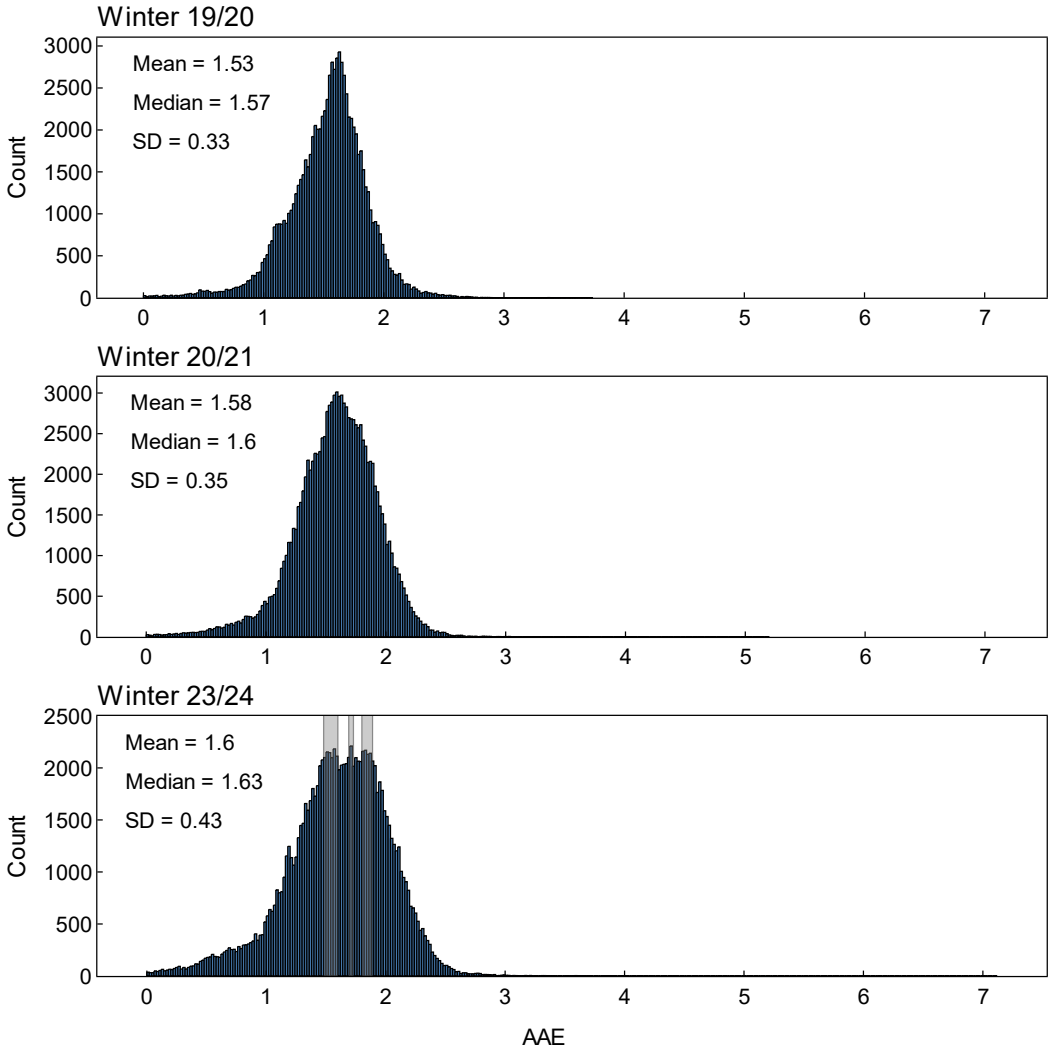


Figure 6.7: *Frequency distribution and summary information of the absorption Ångström exponent measured in Enniscorthy with the AE33 during each winter period. The tri-modality of winter 23/24 is highlighted.*

The diurnal profile of AAE values during each period (Figure 6.8) also indicates a rise in solid fuel burning related emissions in the evening time, when AAE > 1.7. This peak is larger in winter 23/24, indicative of a shift in the dominant emission sources towards more UV absorbing emissions, which is consistent with the RPI in the ΔC diurnal profiles being highest in winter 23/24. The observed trend in the increasing maximum diurnal AAE is also consistent with the lower value reported in 2015 (~1.55) (Buckley, 2019).

Interestingly, the AAE peak between approximately 05:00 and 09:00 in winter 20/21 is not observed during the other periods. As this winter involved periods of COVID-19 government restrictions and traffic counts during the typical morning rush hour time were significantly lower, it is possible that a compositional change in the morning ambient carbonaceous aerosol occurred. This, in conjunction with a potential increase in solid fuel burning due to people being at home more, would lead to a dominance of UV absorbing species in the morning. BC trends were investigated in other European locations during the COVID-19 lockdowns, where a general decrease in BC levels was observed across Europe, however the extent of the decrease was variable across locations and in certain regions it remained unchanged or exhibited an increase due to increased residential combustion-related emissions (Evangeliou et al., 2021).

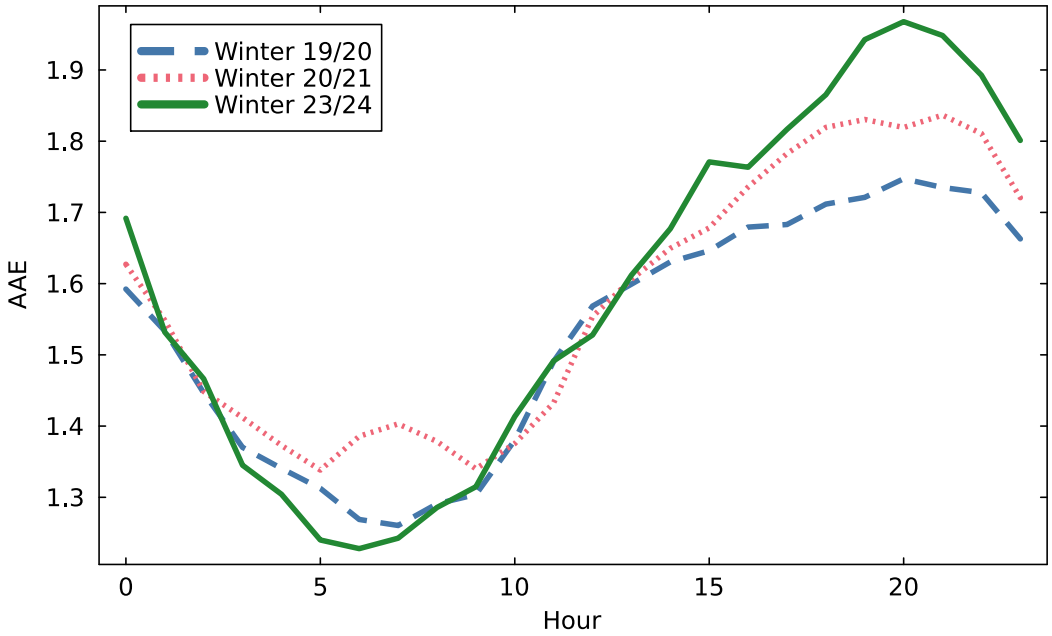


Figure 6.8: *Seasonal diurnal profiles of absorption Ångström exponent values.*

6.4.3.2 Aethalometer Model Parameter Selection

The source apportionment results of eBC using values of α_{SF} = 1.68 and α_{TR} = 0.9 (Zotter et al., 2017) shows the relative amount of eBC_{SF} increased between 2019 and 2024 by 7.94% (Table A4.1). Employing this parameter pair results in an increase in the proportion of negative results between each consecutive period under investigation (1.76–5.29% between winters 20/21 and 23/24). The diurnal profiles of the apportioned eBC (Figure A4.2) also show the unsuitability of the exponents in the later periods due to the negative diurnal values of the traffic contribution. Although the proportion of negative eBC_{SF} and eBC_{TR} with values of α_{SF} = 2 and α_{TR} = 1 applied (the AE33 on-line source apportionment values) are relatively low (Table A4.2), the diurnal concentration profiles (Figure A4.3) show the model is mainly unsuitable in the first two periods due to the unrealistically high evening eBC_{TR} peak, a result also shown in Enniscorthy in 2015 (Buckley, 2019).

Due to the change in the nature of the emissions indicated by the observed change in AAE distributions and the lack of suitability for a single one of the aforementioned source apportionment model parameter pairs across every year in this analysis, different α_{SF} values were used for each year, based on optimising the ranges of values verified by the apportioned eBC concentrations. α_{TR} was varied between 0.9 and 1.1 in steps of 0.1, and α_{SF} was varied between 1.5 and 2.2 in steps of 0.1. An AAE value was deemed optimal by minimising the number of negative eBC_{SF} and eBC_{TR} concentration values and maximising the number of hourly TR and SF percentage splits between 0-100%. A maximum of 5% of concentrations were allowed to be negative, and a 95% of percentage splits had to fall within the stated range. Additionally, the difference in the morning and evening traffic-related eBC peaks were kept minimal to reflect the realistic similarity between emissions during periods, which was informed by the RPI criterion. While this optimisation procedure cannot be validated with external data and is a relatively simple fitting approach, it provides an AAE pair on the basis that the resulting concentrations and percentage splits are as realistic as possible. A similar procedure was suggested by Buckley (2019), where an AAE pair of $\alpha_{TR} = 0.9$ and $\alpha_{SF} = 1.65$ was found to be most suitable for Enniscorthy in 2015, resulting in an eBC_{SF} of 84%.

Table 6.5 details the optimal α_{SF} values for each period and the resulting source apportionment results, with an overall increase in the optimal α_{SF} observed over time. The source apportionment results for eBC_{SF} in winter 19/20 (59.29%), indicates a greater overall reliance on solid fuel burning in this period, and then falls to approximately 54% for both of the subsequent winter periods. Additionally, the diurnal profiles (Figure 6.9) show realistic hourly concentrations, with a rise in solid fuel related concentrations in the evening time (which is most notable for the first

winter), and a fossil fuel morning and evening peak of similar magnitude for each period, however the difference was relatively larger in winter 2. The eBC_{SF} peak observed in winter 19/20 (~5 μg m^{-3}) is over double that observed in the latter two winters (~2 μg m^{-3}). The peak of the traffic-related portion of eBC, eBC_{TR}, decreased by approximately 1 μg m^{-3}, with the lowest diurnal peak observed in winter 23/24.

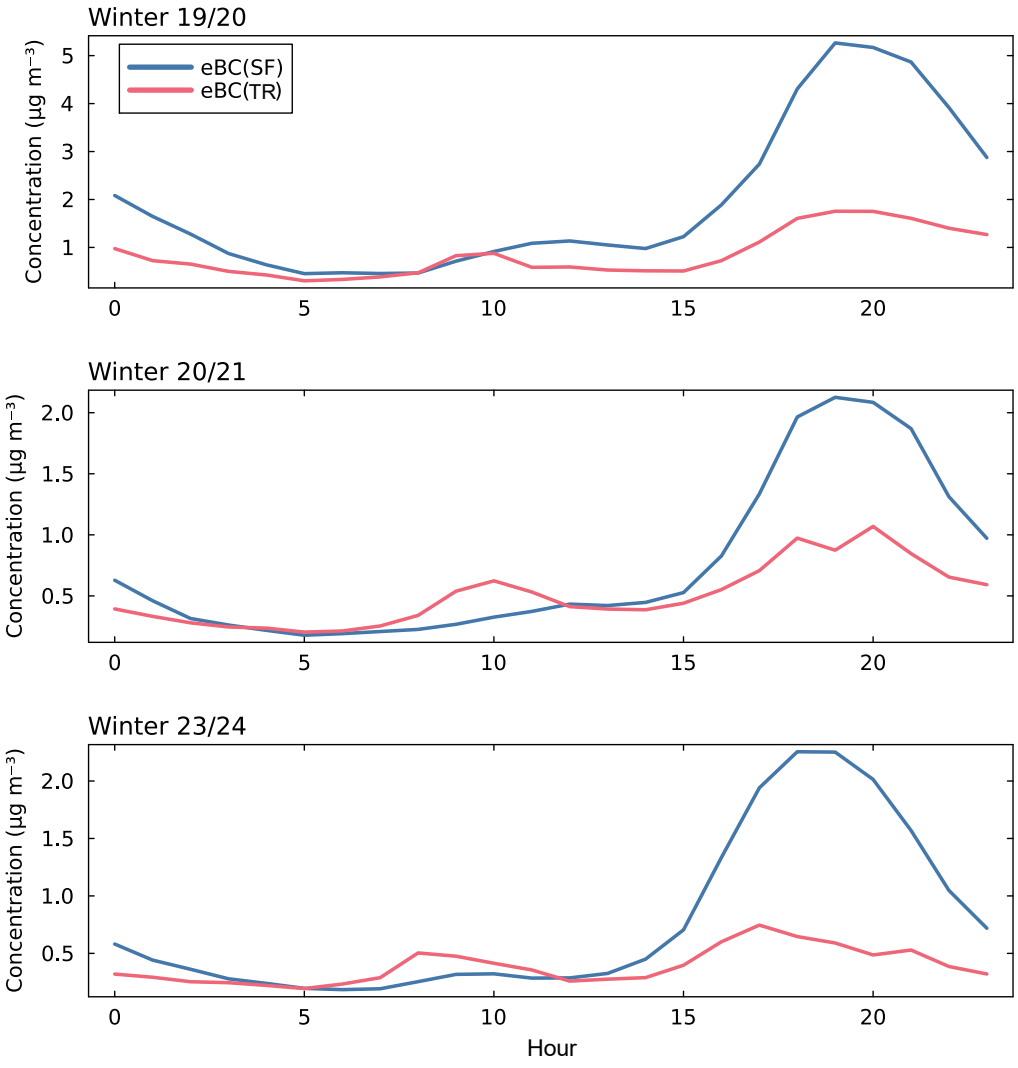


Figure 6.9: *Seasonal diurnal profiles of eBC_{SF} and eBC_{TR}, calculated from AE33 measurements using optimal AAE values.*

Table 6.5: *Optimised AAE values for each winter period and summary of source apportionment of hourly averaged eBC using the optimised AAE values.*

	α_{SF}	α_{TR}	eBC_{SF}	eBC_{TR}	Mean eBC_{SF}	Mean eBC_{TR}
			%	%	µg m^{-3}	µg m^{-3}
Winter 2019/2020	1.82	0.9	58.24	41.75	1.60	1.15
Winter 2020/2021	1.91	0.9	52.00	48.00	0.64	0.59
Winter 2023/2024	1.99	0.9	51.01	48.98	0.58	0.56

The Census of Population results from 2016 and 2022 can provide insight into the changes in fuel use for residential space heating (CSO, 2016; CSO, 2022). From 2016 to 2022, the proportion of houses using coal for central heating in Enniscorthy decreased by 9%. Over an area of approximately 8 km^2, coal represented a fuel use density of 146 houses per km^2 in 2016 and 105 houses per km^2 in 2022. The proportion of electricity and oil use increased by 4.4% (additional 29 houses per km^2) and 2.3% (additional 37 houses per km^2), respectively, while the use of wood officially reduced from 1.8% to 1.2%, although with wood there is the potential for some non-traded, and therefore unrecorded, fuel volumes. While the use of coal is decreasing over time, the use of oil has increased which may contribute to the eBC_{TR} portion of measured BC. Although emissions are typically not as high with residential oil burners as with the combustion of coal, oil burners can produce notable SO_2 emissions which may contribute to secondary particle formation in the atmosphere (Hays et al., 2008; Sippula et al., 2009).

6.4.4 Variations in the Relationship of eBC and $PM_{2.5}$

Figure 6.10 depicts the linear relationship between the eBC and $PM_{2.5}$ hourly concentrations. The strong linear relationship exhibited in winter 19/20 ($R^2 = 0.94$) is indicative of a shared common source between both pollutants and is comparable to the correlation reported for 2015 (0.92) (Buckley, 2019). However, as meteorological factors were relatively constant during each winter period (Section 3.1), the reduction in correlation (0.94 to 0.86) and the slope (0.133 to 0.097) over time is interesting. As the observed distribution of measured AAE values (Figure 6.7) changes from winter 19/20 to winter 23/24, there is a clear indication that the fuel mixture is changing over

time. This would be expected due to legislative changes. The gradual reduction in the average eBC to $PM_{2.5}$ ratio (slope) and the reduction in the goodness-of-fit (R^2 values) both indicate that the relationship between eBC and $PM_{2.5}$ is changing, which is also indicative of a change in fuel mixture.

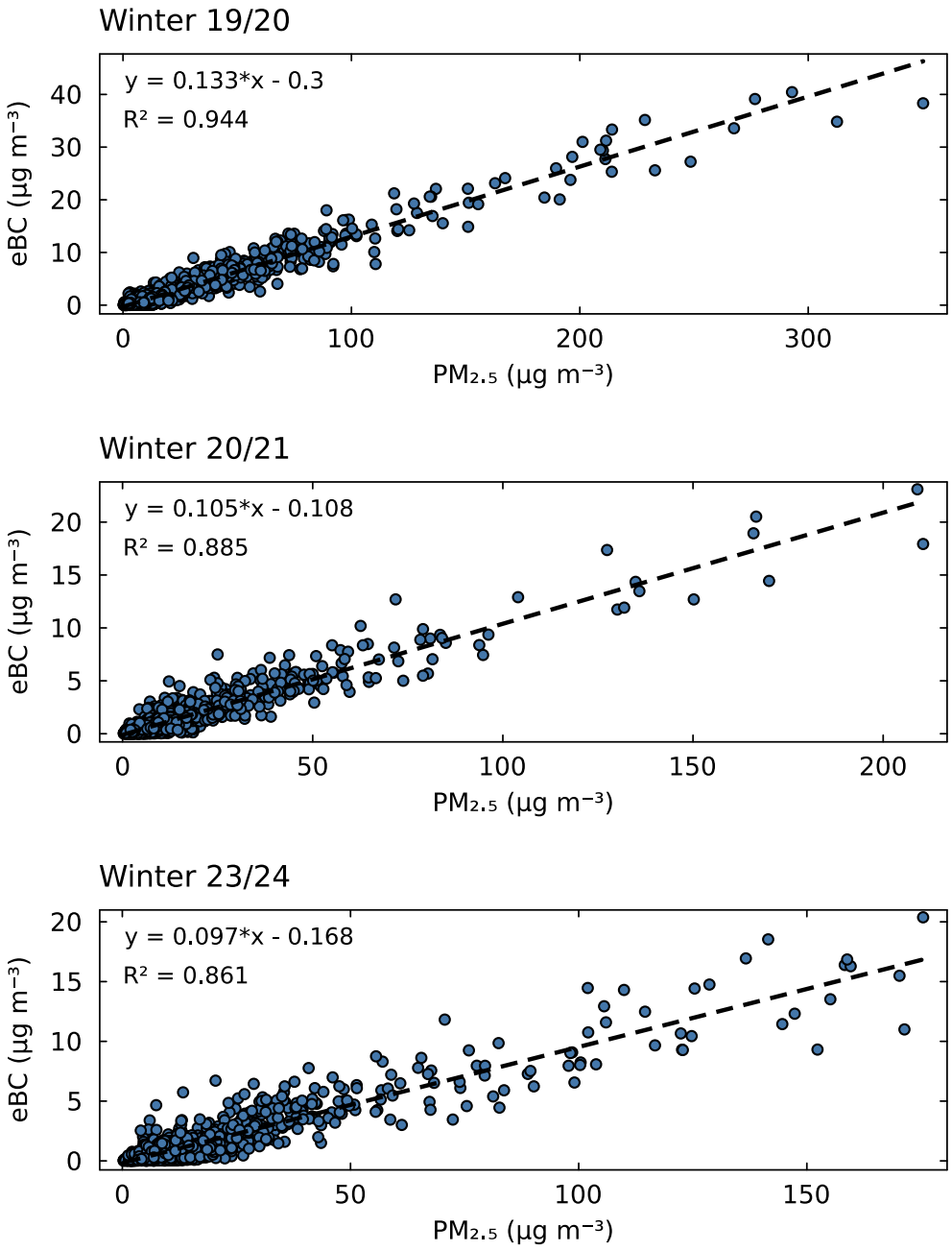


Figure 6.10: *Seasonal linear comparison of $PM_{2.5}$ and eBC.*

The contribution of eBC to total $PM_{2.5}$ appears to be decreasing in each consecutive winter period, from 10% in winter 19/20 to 7.7% in winter 23/24 (Table 6.4), including from that reported in 2015 (13.2%) (Buckley, 2019). An uncertainty propagation method was used to assess the uncertainty in the relative change in percentage contribution of eBC to $PM_{2.5}$ between each period. To investigate the significance of the change in contribution between different periods, a test statistic was then calculated based on the change in contribution relative to the uncertainty in that change. The associated uncertainties for individual measurements (obtained with the Fidas 200 and the AE33) were used to calculate the overall uncertainty of each dataset (Equation 6.3). The $PM_{2.5}$ measurement uncertainty is assumed to be 9.7% (Palas, 2024), which can be expressed in μg m^{-3} by calculating the mean relative uncertainty across all measurements in a given period, i.e., the mean of the ratios of the uncertainty value (0.097) against each individual measurement. The uncertainties associated with aethalometer measurements can vary depending on various factors (Cuesta-Mosquera et al., 2021). Taking this into account, the AE33 uncertainty was estimated based on the concentration outputs during the "clean air tests" carried out automatically by the AE33, where a bypass filter is utilised to remove particles from the ambient air stream. The mean eBC concentration measured during these "clean air" periods, and thus the estimated uncertainty of the measurement sets, for winters 19/20, 20/21, and 23/24 were 0.90 μg m^{-3}, 0.03 μg m^{-3}, and 0.39 μg m^{-3}, respectively. Using the assumed uncertainties for both the $PM_{2.5}$ and the eBC datasets, the overall uncertainty for the comparison was found with the following equation:

$$U_{(n)} = \sqrt{\left(\frac{u_{eBC(n)}}{\bar{x}_{eBC(n)}}\right)^2 + \left(\frac{u_{PM2.5(n)}}{\bar{x}_{PM2.5(n)}}\right)^2} \quad \textbf{(6.3)}$$

where $U_{(n)}$ is the uncertainty associated with the contribution of eBC to $PM_{2.5}$ in the period, n. u_{eBC} and $u_{PM2.5}$ are the mean measurement uncertainties for the eBC and $PM_{2.5}$ measurements, expressed in μg m^{-3}, respectively, and $\bar{x}_{eBC}$ and $\bar{x}_{PM2.5}$ are the mean eBC and $PM_{2.5}$ measurements, respectively.

The uncertainty associated with each contribution was then carried through to determine the uncertainty in the relative change in the contribution of eBC to $PM_{2.5}$ between two periods, i and j (Equation 6.4).

$$\Delta U_{i,j} = \sqrt{U_{(i)}{}^2 + U_{(j)}{}^2} \quad \textbf{(6.4)}$$

where $\Delta U_{i,j}$ is the uncertainty associated with the relative change in contribution of eBC to $PM_{2.5}$ between two periods, i and j.

The test statistic was carried out with a 90% confidence level (critical value = ±1.645). The score for the relative change from winter 19/20 to winter 23/24 was −1.78, and −0.13 when comparing winter 20/21 to winter 23/24. The results indicate that the relative decrease in the contribution of eBC to $PM_{2.5}$ between winters 19/20 and 23/24 was statistically significant; however, the change from winter 20/21 to winter 23/24 was found to be within the uncertainty limits, i.e. statistically insignificant and attributed to measurement uncertainty.

One possibility for the change in the eBC to $PM_{2.5}$ relationship over time, particularly after the introduction of new legislation in Enniscorthy, is that the composition of $PM_{2.5}$ is changing due to a change in the secondary organic aerosol formation resulting from a change in combustion-related volatile organic compound (VOC) emissions. With a shift from fuels that produce more directly emitted carbonaceous aerosol towards fuels that produce more UV-absorbing organic aerosol, some of which may

originate from transformation of emitted VOCs, the relationship between BC and $PM_{2.5}$ becomes less direct. In other words, the changing relationship between BC and $PM_{2.5}$ could indicate an increase in the significance of secondary formation from gaseous precursors (sulfur dioxide, nitrogen oxides, and VOCs) of $PM_{2.5}$. The aging of pollutants, where primary emissions undergo chemical transformations, might have affected the physical and chemical properties of $PM_{2.5}$ in the later periods, leading to a more complex relationship with eBC.

6.4.5 Spectral analysis of $PM_{2.5}$

$PM_{2.5}$ is comprised of a mix of primary and secondary particles from various sources (traffic, industrial emissions, natural sources, domestic heating), each with its own temporal pattern. BC, on the other hand, is primarily from combustion processes (traffic and residential heating). Spectral analysis of the temporal profiles of both concentration-time profiles can offer additional insights into the relative proportion of locally emitted $PM_{2.5}$ and eBC compared to the regional or background proportion. $PM_{2.5}$ and eBC concentration timeseries from each winter period were transformed using the DFT to investigate the underlying frequency components in each dataset. The LSCE for each dataset was found by expressing the integral of the high frequency ($> 32\ hr^{-1}$) portion of the periodogram as a fraction of the total integrated area.

The polar plots indicated high $PM_{2.5}$ concentrations were mainly associated with low wind speeds, and therefore local emissions, with some additional influences from further afield, which is in line with all LSCEs for $PM_{2.5}$ being over 50% (Table 6.6). The largest LSCE was observed in winter 18/19 (77%), although this was similar to the LSCE observed in winter 22/23 (71%). Despite high concentrations overall, a

lower LSCE was found for winter 19/20 (56%), indicating a larger proportion of the ambient particulate pollution was background or regional in nature.

The winter 20/21 periodograms for both $PM_{2.5}$ (Figure 6.11) and eBC (Figure 6.12) are interesting as COVID-19 restrictions were ongoing during this period. There is a notable lack of low-frequency, or long-term, cycles present for both datasets. While this did not result in the highest overall LSCEs, the same lack of low frequency components was not noted in the periodograms for periods where a considerably high LSCE (> 70%) was found for $PM_{2.5}$ (Table 6.6). The full effect of the lockdown period is difficult to ascertain due to the introduction of legislation in Enniscorthy in September 2020, however it is possible that regional or long-range pollution sources (such as transported industrial or shipping emissions) were diminished, which could explain the low magnitudes of the low frequency components.

The changes in the LSCE for both eBC and $PM_{2.5}$ between 2018 and 2024 indicates that the proportion of regional to local pollution is subject to change from year to year. However, the considerable reduction in overall concentrations and continually high LSCEs through the different periods suggests a proportional decrease in both the background/regional and local portions of the particulate concentrations over time. BC is often considered to be primarily a local pollutant, however the results proposed here would suggest a significant portion is not. Regional eBC has been observed and measured in various locations across Europe and globally (Ahlberg et al., 2023; Healy et al., 2017; Mbengue et al., 2020; Prasad et al., 2018; Wang et al., 2017), however the extent of the regional impact can vary between locations and is dependent on various factors such as topography, atmospheric conditions, and emission sources.

While long-term BC studies are lacking in comparison to brief, intensive measurement campaigns, previous European-based studies have found decreasing eBC trends; however, consistent levels of BC originating from certain emission sources were also observed. Savadkoohi et al. (2023) observed a general decreasing trend in eBC at both traffic sites and urban background sites across Europe. However, while traffic related eBC was found to be decreasing over the decade in the study, the eBC originating from residential and commercial sources remained relatively constant, or even slightly increased in some regions. Kutzner et al. (2018) found that background BC in Germany was decreasing slowly compared to urban BC.

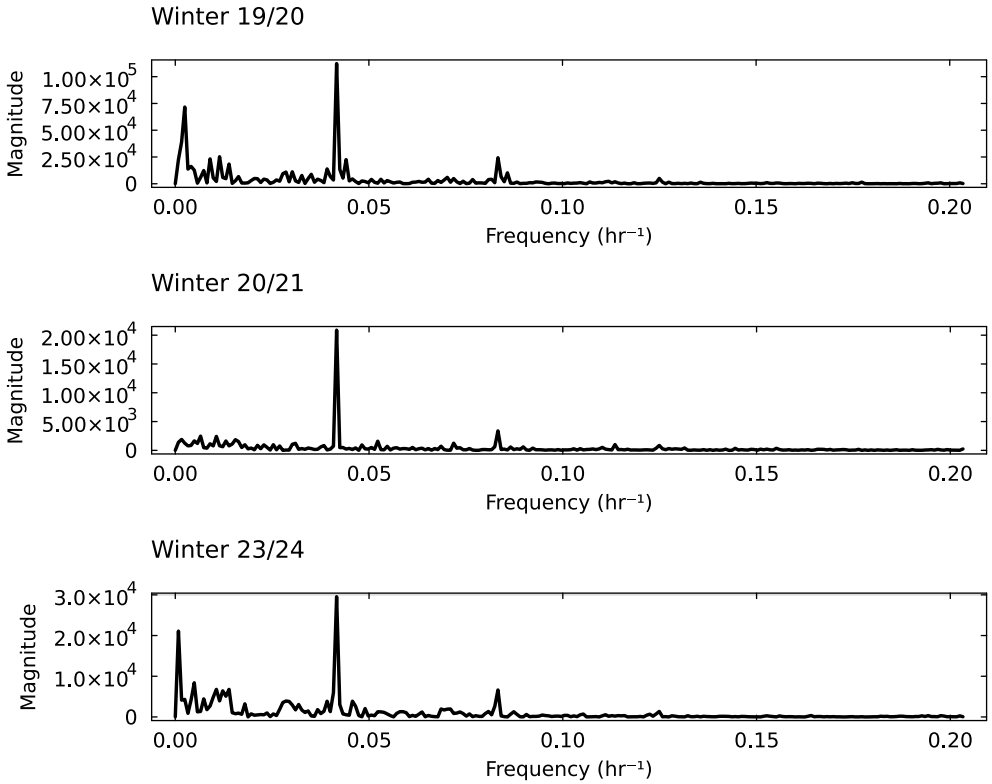


Figure 6.11: *Seasonal periodogram representations of the discrete Fourier transform of the $PM_{2.5}$ signals.*

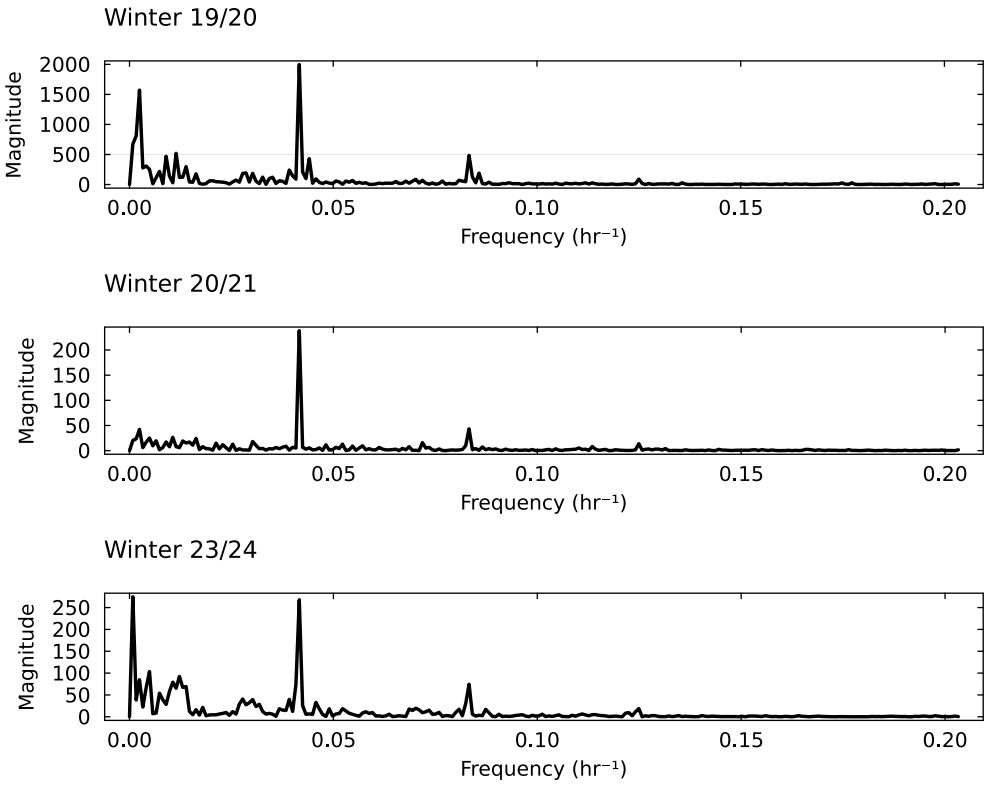


Figure 6.12: *Seasonal periodogram representations of the discrete Fourier transform of the eBC signals.*

Table 6.6: *Seasonal local source contribution estimates of $PM_{2.5}$ and eBC.*

	$PM_{2.5}$	**eBC**
	%	**%**
Winter 2018/2019	77	–
Winter 2019/2020	56	52
Winter 2020/2021	66	63
Winter 2021/2022	56	–
Winter 2022/2023	71	–
Winter 2023/2024	52	50

6.5 Limitations

Both the standard on-line AAE pair employed by the AE33 software, and the values proposed by Zotter et al. (2017), which have been used in numerous studies in different environments, were shown to be unsuitable to the data included in this study. However, even with an optimally fitted AAE pair the aethalometer model has significant limitations in environments with more than two dominant BC sources, because it is a two-parameter model and it was developed in an environment dominated by wood smoke (Sandradewi et al., 2008b). While the values selected provided more appropriate diurnal patterns and reduced the quantity of negative values compared to both other options assessed, the periodograms produced provide further insight into the reliability of the model (Figure A4.4, Figure A4.5). The transformed eBC_{SF} signal contains strong peaks corresponding to the 24-hour and 12-hour cycles, and smaller magnitude peaks in winters 2 and 3 corresponding to the 8-hourcycle, typically observed with traffic emissions. Additionally, the eBC_{TR} periodograms contain peaks corresponding to the 24-hour and 12-hour cycles in winters 1 and 2, but these are not apparent in winter 23/24. An 8-hour cycle is not clear in either winter period for the eBC_{TR} transformations. The periodograms suggest the AAE pair applied to the data, in particular the winter 23/24 data, did not produce reliable eBC_{SF} and eBC_{TR} timeseries, based on assumptions of daily emission profiles. However, similar issues were also found when the values proposed by Zotter et al. (2017) were applied (Figure A6.4, Figure A4.7).

The values proposed by Zotter et al. (2017) have previously been found to have questionable results in other European locations (Savadkoohi et al., 2023). Previous studies have shown that the AAE values can vary significantly depending on various

factors such as the specific type of fuel used, the burning conditions, and the microphysical properties of the BC particles (Helin et al., 2021b; Li et al., 2016; Liu et al., 2018). While an investigation into optimal AAE values was carried out, no additional chemical data was used to corroborate these values which would have likely provided additional insights into the optimal AAE pair for each period.

In addition to the necessity of selecting a correct AAE pair, the aethalometer model generally assumes that the MAC values, which describe the BC absorption per mass, remain constant at a given wavelength and that this is independent of the observational site (Srivastava et al., 2022). However, a wide range of MAC values have been reported at certain wavelengths in different environments, depending on the physiochemical properties of the BC particles (Bond & Bergstrom, 2006; Liousse et al., 1993; Nordmann et al., 2013; Savadkoohi et al., 2024; Schwarz et al., 2008). The MAC values are influenced by a coating that can build up on BC which may act as a lens and focus light into the core of the BC particle (Knox et al., 2009), and the coating can vary with source composition. The MAC values are highly dependent on the mixing state of BC and other chemical processes and can influence the light absorption of BC particles and can therefore affect the apportioned concentrations of eBC.

6.6 Conclusion

A combined approach to assess the temporal variability in winter air quality in Enniscorthy was employed in this study. Temporal variation analysis of the ambient aerosol and local source contribution estimates (LSCEs) obtained via spectral analysis of concentration data were used in conjunction to compare the local emissions across winter periods between 2018 and 2024.

A considerable reduction in the mean eBC and $PM_{2.5}$ between winter 19/20 and winter 23/24 was observed (1.61 μg m^{-3} and 8.27 μg m^{-3}, for eBC and $PM_{2.5}$, respectively). The change in the contribution of eBC to overall $PM_{2.5}$ over time was statistically significant, suggesting a compositional change in emissions over time, potentially due to an increased significance of secondary $PM_{2.5}$ due to a changing fuel mixture. Additionally, the Ångström absorption exponent (AAE) values suggest the profile of emissions in Enniscorthy have changed, with the mean AAE increasing from 1.53 in winter 19/20 to 1.63 in winter 23/24. The shift in the AAE distribution is indicative of more UV absorbing emissions, such as wood or peat/turf burning, becoming more important factors in the ambient aerosol.

High local contributions from both $PM_{2.5}$ and eBC (> 50%) were observed in all periods, likely driven by residential solid fuel burning. Winter 18/19 saw the highest LSCE for both $PM_{2.5}$ (77%) and the highest mean $PM_{2.5}$ concentration (20.04 μg m^{-3}), highlighting the localised nature of PM emissions in Ireland. The LSCEs for $PM_{2.5}$ and eBC data were similar, again suggesting that the $PM_{2.5}$ emissions are mainly local in nature and strongly linked to combustion (solid fuel burning and traffic), which can be both locally emitted and regionally transported. While the emissions in general have reduced over time, the consistently high LSCEs suggest a reduction in both

regional/background and local emissions, though a considerable relative proportion is still originating from local emission sources in the area. Additionally, the lack of a consistent downward trend in LSCEs over time suggests the relative importance of regional vs local emissions can change from year to year. The high evening concentrations observed for $PM_{2.5}$ in winter 23/24, coupled with the consistently high LSCE, indicates the use of solid fuels for home heating remains strong.

Although the impact of the COVID-19 restrictions was not entirely apparent through temporal analysis of $PM_{2.5}$ and eBC, a suggested decrease in the relative importance of regional emissions compared to local emissions was observed, as evidenced by the lack of high magnitude low frequency components in the winter 20/21 periodogram. However, the LSCE in winter 22/23 was higher than in winter 20/21, indicating the relative importance of regional vs local emissions to the overall $PM_{2.5}$ signal can shift without extreme external factors such as the COVID-19 restrictions.

The two-source aethalometer model has limitations in a multi-fuel source environment, even when the AAE parameters are optimised to the specific conditions observed in a given period as both the standard AAE pair and the optimised values were found to be limited for correctly apportioning the eBC concentrations. While the nomenclature employed here refers to a traffic and a solid fuel burning portion of the total measured eBC in order to reflect the multi-fuel environment in Ireland, the aethalometer model cannot distinguish between more than two main emission sources. The wide range of AAE values in each winter period is suggestive of multiple contributors to light absorbing particles beyond just traffic aerosol and woodsmoke aerosol. Additionally, the lack of a single pair of AAE values which provided realistic outputs in all three winter periods is indicative of the need for spatially and temporally specific model parameters. However, the optimisation procedure did not adjust the

AAE value related to the traffic emissions, suggesting that this fuel type does not change considerably over time, or indeed, between different environments in Europe.

The complex and dynamic situation observed in Enniscorthy across the three winter periods highlights the need for effective pollution monitoring and assessment, in particular, long-term studies of the ambient aerosol are required to assess the evolution over time. $PM_{2.5}$ and eBC measurements were used in tandem with more in-depth timeseries analysis to provide insights into the changing nature of pollution, while also highlighting the ongoing contribution of local emissions to the ambient $PM_{2.5}$ levels in the area. The effects of legislative changes were evident, making ongoing, long-term assessments crucial to understanding pollution dynamics over time.

6.7 Appendix 4

Table A4.1: *Summary of source apportionment of hourly averaged eBC and percent of eBC_{SF} and eBC_{TR} with negative values using AAE values of α_{SF} = 1.68 and α_{Tr} = 0.9 for each winter period.*

	eBC_{SF}	**eBC_{TR}**	**Negative eBC_{SF} values**	**Negative eBC_{TR} values**
	%	**%**	**%**	**%**
Winter 2019/2020	72.52	27.48	1.76	16.89
Winter 2020/2021	78.81	21.19	1.91	26.65
Winter 2023/2024	80.46	19.54	5.29	34.6

Table A4.2: *Summary of source apportionment of hourly averaged eBC and percent of eBC_{SF} and eBC_{TR} with negative values using AAE values of α_{SF} = 2 and α_{TR} = 1 for each winter period.*

	eBC_{SF}	**eBC_{TR}**	**Negative eBC_{SF} values**	**Negative eBC_{TR} values**
	%	**%**	**%**	**%**
Winter 2019/2020	41.75	58.25	3.19	0.47
Winter 2020/2021	45.95	54.05	3.00	1.19
Winter 2023/2024	47.05	52.95	6.93	3.17

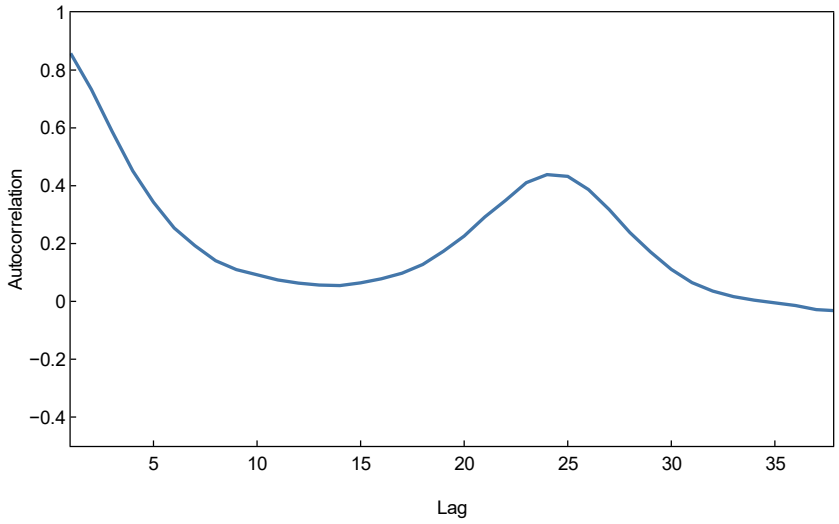


Figure A4.1: *Lag autocorrelation for hourly $PM_{2.5}$ data from all winter periods for k = 1 to k = 38.*

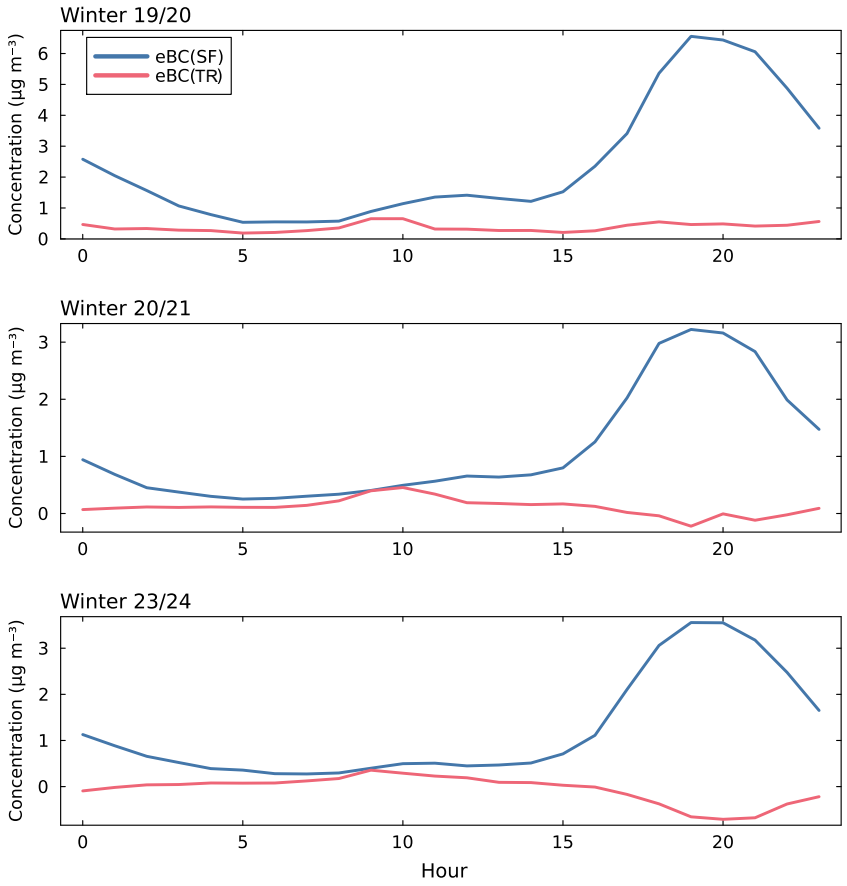


Figure A4.2: *Seasonal diurnal profiles of eBC_{SF} and eBC_{TR}, calculated from AE33 measurements using AAE values of α_{SF} = 1.68 and α_{TR} = 0.9.*

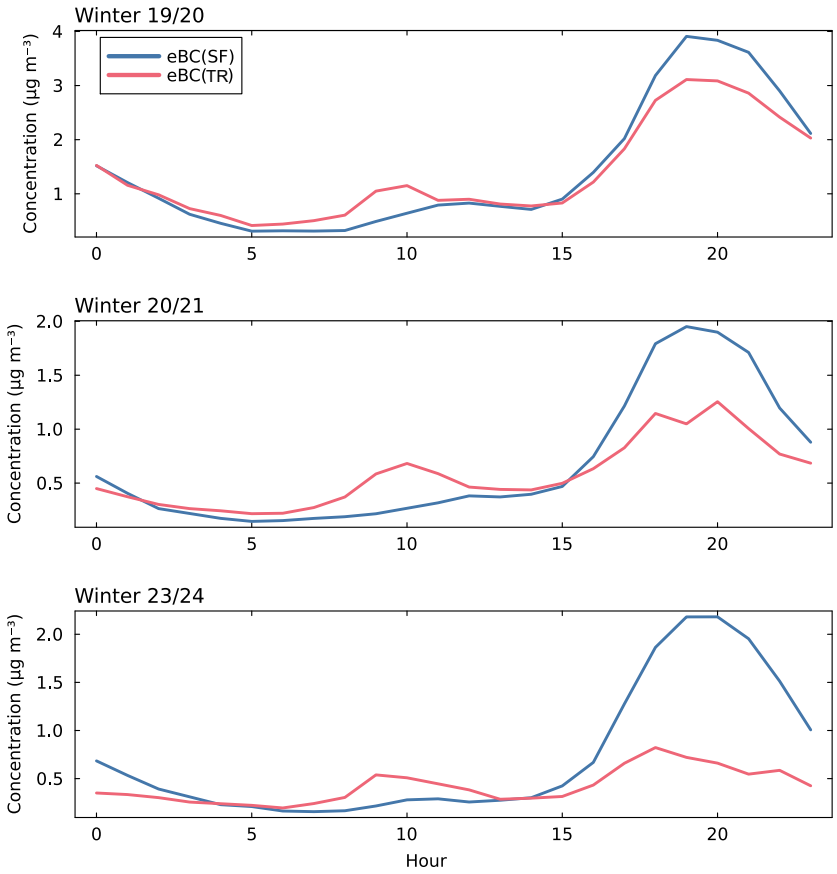


Figure A4.3: *Seasonal diurnal profiles of eBC_{SF} and $eBC_{TR,}$ calculated from AE33 measurements using AAE values of α_{SF} = 2 and α_{TR} = 1.*

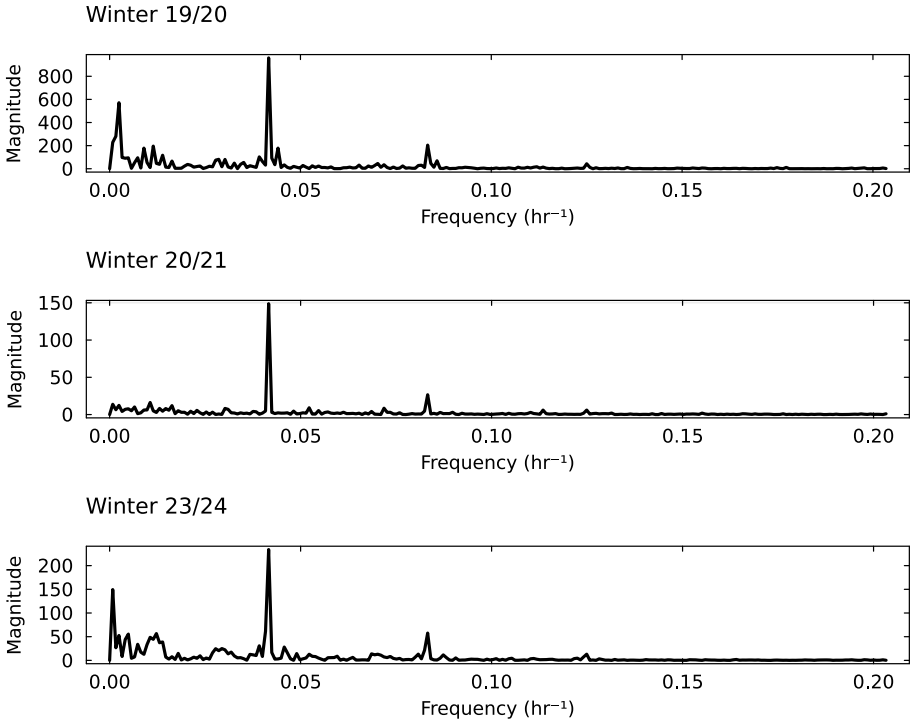


Figure A4.4: *Seasonal periodogram representations of the discrete Fourier transform of the eBC_{SF} signals, using optimised AAE values.*

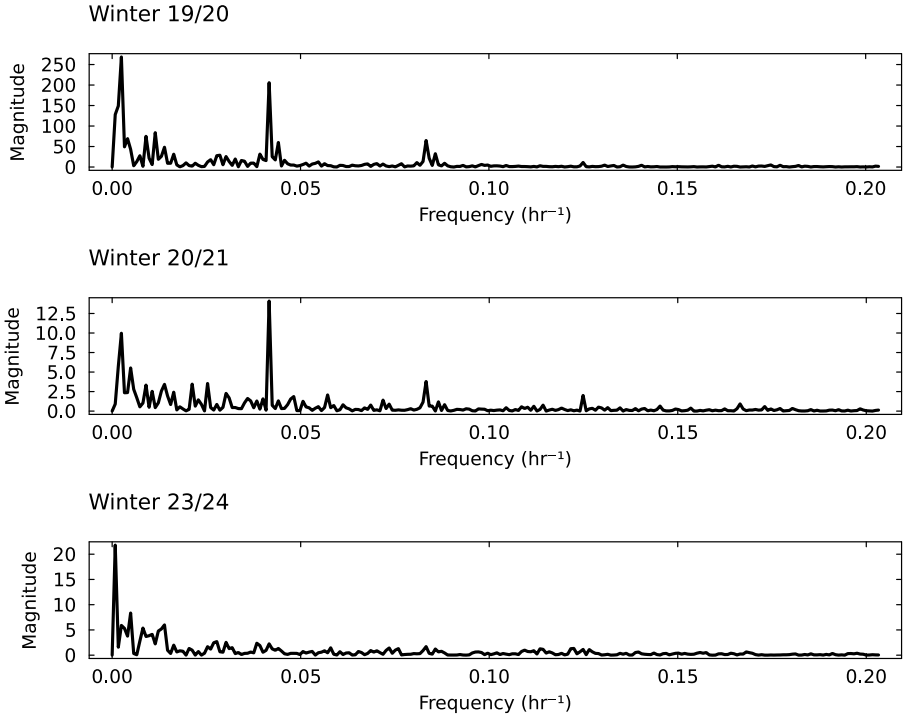


Figure A4.5: *Seasonal periodogram representations of the discrete Fourier transform of the eBC_{TR} signals, using optimised AAE values.*

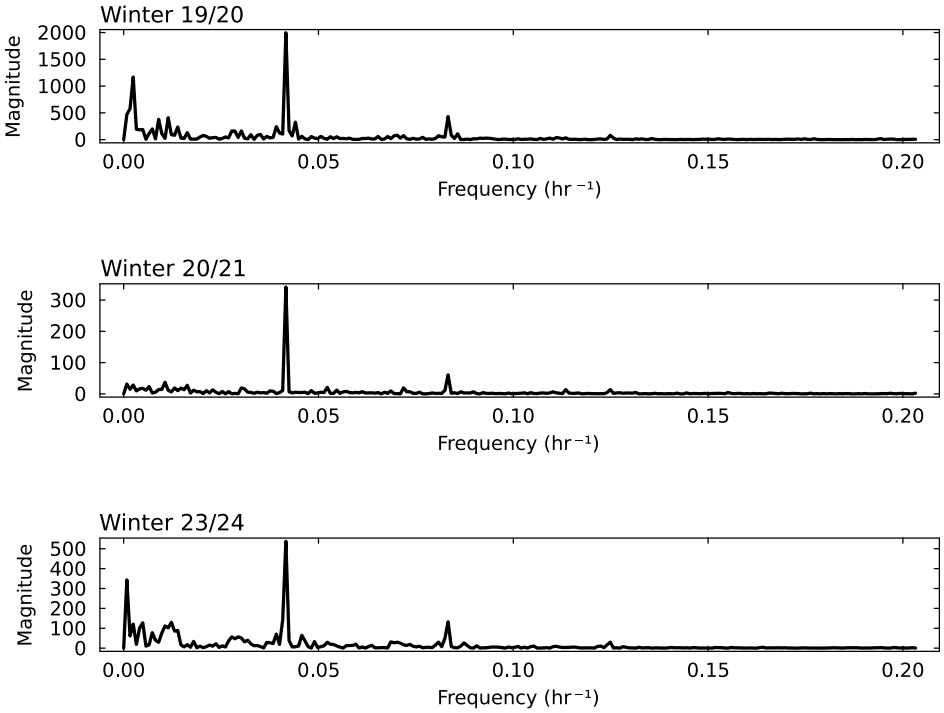


Figure A4.6: *Seasonal periodogram representations of the discrete Fourier transform of the eBC_{SF} signals, using AAE values of $\alpha_{SF} = 1.68$ and $\alpha_{TR} = 0.9$.*

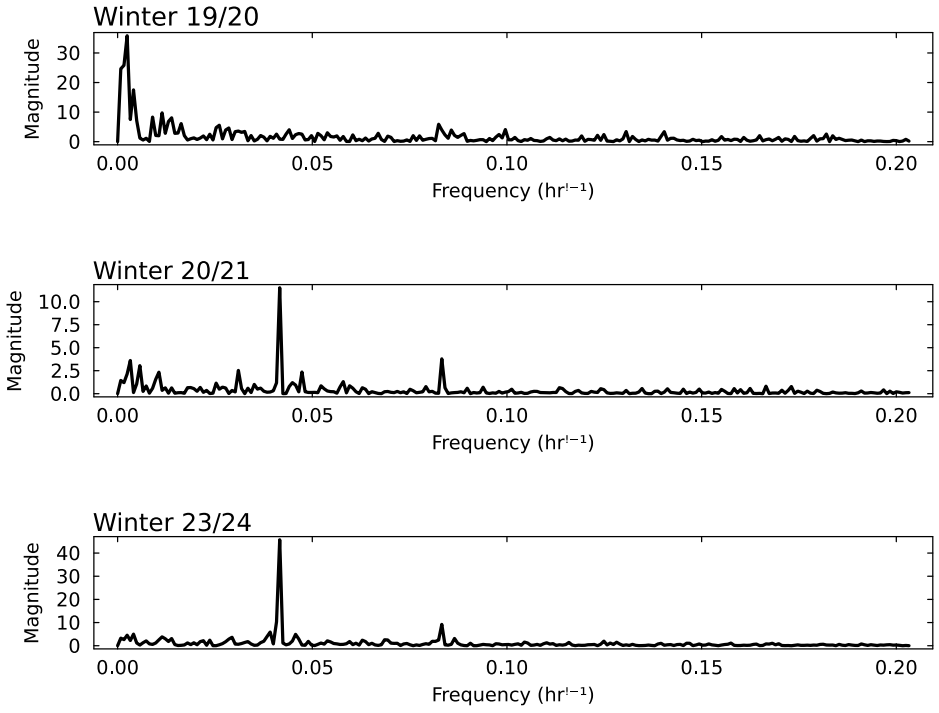


Figure A4.7: *Seasonal periodogram representations of the discrete Fourier transform of the eBC_{TR} signals, using AAE values of $\alpha_{SF} = 1.68$ and $\alpha_{TR} = 0.9$.*

7 Conclusion

7.1 Summary

This research explored the potential for AQS networks to monitor $PM_{2.5}$ and provide an affordable and flexible supplement to traditional air quality monitoring infrastructure. Data generated by the first AQS sensor networks in Ireland (Cork City and Dungarvan) has been analysed using several methods and demonstrates the importance of low inter-sensor variation and appropriate sensor calibration through co-location studies. Exploratory data analysis beyond calibration exercises also enabled investigation into the spatial variability of $PM_{2.5}$ and the influence of local sources within both urban and semi-urban settings. In addition, the analysis carried out on BC data in Enniscorthy showed how advanced data analysis techniques can be integrated with regulatory monitoring to further our understanding of the nature and origins of particulate pollution and how these change over time. The results also emphasise that absolute population size is not always a driving factor for air pollution levels. Cork City has a population over ten times that of Dungarvan and Enniscorthy, yet the air pollution does not increase proportionally, and in fact, Cork City has relatively low levels of $PM_{2.5}$ pollution compared to Enniscorthy.

Leveraging the high temporal resolution of two-minute PurpleAir (PA) data in Cork City led to the development of a procedure to investigate the concentration fluctuations and further understand the spatial variability in local $PM_{2.5}$ pollution. By linking rapid, short-lived concentration changes with locally emitted pollution, high levels of $PM_{2.5}$ were linked to local sources. The impact of local sources was particularly evident in smaller towns/commuter areas surrounding the larger urban area of Cork city centre,

where emissions were found to be on par with or exceeding those closer to the city centre. Valuable insights into the spatial heterogeneity of $PM_{2.5}$ across various locations in the city were also provided, information that was previously unattainable through traditional air quality monitoring infrastructure, with significant differences observed in the $PM_{2.5}$ concentrations across Cork City.

The considerable differences in the estimated local contributions across a relatively small city highlighted the importance of achieving high spatial density with monitoring systems. However, this methodology has certain limitations. Mainly, that the very high temporal resolution achieved with the PA sensor network is not possible with all AQSs, particularly when relying on solar powered options for increased spatial coverage, such as with the Clarity devices. Considerations must be given to balancing the ease of sensor deployment, in terms of power source (for example, solar powered vs mains powered), appropriate siting within the network area, and temporal resolution. The spectral analysis method overcame this limitation and was developed to estimate the local source contributions of $PM_{2.5}$ data based on hourly data. However, the temporal information achieved with the rapid fluctuation method was not available via spectral analysis, i.e. pinpointing where in the concentration-time profiles the local sources were dominating. The Fourier transform does not provide high temporal resolution in tandem with high frequency resolution.

The spectral analysis method did, however, reaffirm the utility in using uncorrected AQS data, with low variation when comparing the results from uncorrected data to both corrected sensor data and reference data. By employing the spectral analysis method across two separate years in Cork City, 2021 and 2023, seasonal trends and reductions in the proportion of winter emissions from local, residential solid fuel burning were observed, suggesting legislative-linked improvements. However, this

effect was unevenly distributed across the city and did not translate to yearly reductions in local source contributions. The application of the spectral analysis method also extended beyond a temporal variation analysis to capture the influence of regional pollution events, such as Saharan dust transport.

A new metric for AQS network data comparison and spatial representativeness assessment, the concentration similarity index (CSI), was developed and extensively tested on colocation data. The CSI method, applied in both Cork City and Dungarvan, enabled a nuanced analysis of how $PM_{2.5}$ concentration-time profiles compared between various locations. The clustering of CSI matrices revealed that the similarity in $PM_{2.5}$ profiles was more influenced by local pollution sources and site characteristics than by geographical proximity, as residential/central areas and locations at the edge of urban areas displayed dissimilar $PM_{2.5}$ concentration-time profiles. The CSI analysis also demonstrated the importance of including seasonal data, particularly from winter months, when evaluating the representativeness of sensor networks and regulatory monitoring stations as solid fuel burning related emissions are a main driver for location dissimilarity. Additionally, the development of the CSI enabled investigation into the spatial representativeness of the EPA monitoring locations within Cork City and Dungarvan. In Cork City, the CSI results indicated moderate representativeness of the EPA monitoring site relative to the broader sensor network. Similarly, in Dungarvan, the EPA site was found to fairly represent the air quality throughout the network, although to a lesser extent than in Cork City.

The analysis in Enniscorthy offered an opportunity to integrate the spectral analysis method, after development on AQS data, with regulatory $PM_{2.5}$ data and additional information from the measurement of equivalent black carbon (eBC). The results in

Enniscorthy showed significant wintertime reductions in both $PM_{2.5}$ and eBC between 2019 and 2024, suggesting that updated solid fuel legislation is successfully helping to mitigate local emissions, however the presence of high evening maxima highlighted the continued reliance on solid fuel burning. Source apportionment of BC using the two-source aethalometer model was carried out, however limitations of applying this model in a mixed fuel environment were shown and spectral analysis of the traffic and solid fuel burning portions of eBC found they were likely unreliable. Additionally, while the AE33 measures BC with high temporal resolution, the aethalometer model does not distinguish between local and non-local emission sources. Using eBC data in conjunction with the spectral analysis method allowed for additional insight into the changing profile of the ambient aerosol in Enniscorthy and the distinction in local and non-local emissions. It is suggested a reduction in both regional and local particulate emissions.

Overall, the research presented underscores the transformative potential of both AQS networks and data-driven approaches to air quality monitoring and assessment. AQS networks can provide critical, location-specific information on air quality variation in urban and semi-urban environments, while new methodologies for $PM_{2.5}$ assessment enables the identification of pollution sources and can support the development of targeted mitigation strategies. In particular, the development of innovative tools like the CSI and the application of spectral analysis further enhances the capabilities of novel air quality assessment techniques to complement and optimise traditional air quality monitoring systems. As demonstrated, sensor networks can fill crucial gaps in air quality monitoring, particularly in areas underserved by regulatory systems, and offer valuable data to inform policy and improve public health outcomes.

7.2 Perspectives

This research marks a significant step forward in the deployment, and notably the advanced application and interpretation, of AQS networks for the monitoring and assessment of $PM_{2.5}$ concentrations. Traditional air quality monitoring systems are highly accurate and important for regulatory purposes; however, they are limited by high costs and often sparse spatial coverage, which frequently leaves both urban and rural areas inadequately represented. The work conducted in Cork City and Dungarvan demonstrates that AQS networks offer an affordable, scalable solution to fill this gap, once the inherent associated limitations are addressed. Specifically, the strong influence of anthropogenic activities was demonstrated, including the spatial disparity in the effects of local pollution sources across both urban areas and small, rural towns.

As sensor technologies continue to improve, becoming more affordable and accurate, we can expect their adoption to grow, particularly in underserved regions. There is considerable scope for integrating sensor networks with developing technologies like machine learning and real-time data analytics to forecast local pollution episodes and mitigate health risks. Alert systems could be established to inform cohorts particularly vulnerable to episodes of elevated air pollutant levels, allowing them to take precautionary measures. Furthermore, advancements in sensor durability and accuracy will likely enable long-term and large-scale deployments. The continuous, reliable data across large regions can provide a foundation for evidence-based policy making, allowing governments to track their progress towards emission reduction targets.

Additional monitoring campaigns incorporating sensor networks, both within Ireland and internationally in more contrasting environments, will help to build upon the tools and methodologies in this work. With further application and testing, these tools offer

the potential for advanced insights into local pollution dynamics, particularly in under-represented regions where it can be done at a fraction of the cost of traditional approaches, enabling data-driven, equitable solutions. Moreover, additional monitoring for longitudinal studies on emission profiles will help to understand how these profiles change, particularly in relation to changing fuel use patterns. Particularly, long-term measurements for BC are required to develop targeted mitigation strategies to limit emissions. Currently, the national air quality monitoring network in Ireland does not include extensive BC data, and while this is likely to be rectified in the incoming updates in European legislation, the analysis in Enniscorthy highlights the changing nature of carbonaceous aerosol emissions and the need for additional monitoring. While this research suggests legislative changes may be positively affecting local $PM_{2.5}$ concentrations, particularly in winter months, future work could build upon this and systematically investigate the long-term effects of socio-economic development and policy interventions in Ireland. Using the fine-scale spatiotemporal data gathered from AQS networks, future research could also focus on linking air pollution levels to health outcomes in specific population groups, including vulnerable communities and different socio-economic groups, to assess the differences, or lack thereof, in the impact of localised $PM_{2.5}$ exposure and determine if these impacts change over time.

As AQS technology allows for the collection of a large amount of data, it raises the question of how many sensors are necessary for optimal spatial representation. Future work can focus on applying the CSI during sensor deployment in order to refine sensor deployment strategies by reducing redundancies and filling spatial gaps in population exposure to pollutants. Similarly, the CSI has potential for optimising regulatory air quality networks, ensuring broader and more accurate population exposure

representation. Both applications would enable more efficient, representative monitoring infrastructure, allowing city planners and policymakers to identify pollution hotspots or underrepresented areas and prioritise interventions while making effective pollution exposure estimates. Additionally, the analysis carried out in Enniscorthy reaffirms the potential for integrating these methods into regulatory monitoring. There is potential for the spectral analysis method to be applied in various regulatory monitoring locations in Ireland to compare the temporal variation in $PM_{2.5}$ and as a broader assessment, to investigate the differences in the impact of local emission sources on local ambient pollution.

The spectral analysis, in particular, opens the door for deploying sensor networks in regions where resources for air quality management are scarce, such as developing countries or smaller, rural towns. With the ability to capture both local emissions and regional pollution events, such as Saharan dust transport, using low-cost AQS networks adds a new dimension to air quality management strategies, particularly in regions susceptible to transboundary pollution. Comparison of the spectral analysis method and the rapid fluctuation method in Cork City showcased the potential for both methods to be developed in tandem to provide estimates for specific traffic-related emissions. It is suggested that differences in the source contribution estimates between both methods is linked to traffic pollution, which is not captured with the rapid fluctuation method. A limitation of the current state of the spectral analysis method was noted with regards to the unavailability of temporal information along with frequency information. This, however, could be overcome with investigation into the use of wavelet transforms instead of the Fourier transform, whereby increased temporal and frequency resolution can be obtained in tandem. The spectral analysis methodology could also be extended to include cluster analysis, whereby clustering

would be applied to periodograms produced in a variety of location types, potentially allowing for a location classification system based on emission profiles, taking the entire concentration-time profile into account. For example, this could identify characteristic patterns in urban, industrial, and rural areas, allowing for more dynamic location-based air quality management.

At the time of this work, the reliability of $PM_{2.5}$ sensors is far beyond that of low-cost gas sensing technologies. While electrochemical gas sensors typically show a real response, they are subject to more cross-sensitivities and therefore currently require more resource intensive data calibration procedures to produce meaningful results. However, gas sensors are likely to improve and will then allow for the incorporation of other pollutants, such as NO_x and O_3, and enable more effective multi-pollutant analysis. Both the CSI method and the spectral analysis can be adapted for other pollutants and can provide a more comprehensive picture of air quality and the interplay between different emission sources.

Challenges remain, however, with the use of AQSs, particularly in ensuring the reliability and accuracy of sensor networks compared to reference-grade monitors. The methods developed in this research are underpinned by co-location assessments demonstrating low inter-sensor variation and good correlations with reference data and have addressed many of these issues. However, continuous validation and improvement of these techniques are necessary. Further research into hybrid systems, where sensor networks and regulatory monitors work in tandem, could help enhance data accuracy while maintaining cost-efficiency. Additionally, as AQS networks expand, issues related to data management, privacy, and the integration of these networks into existing environmental frameworks must be addressed.

In conclusion, the work presented here not only highlights the immense potential of AQS networks, and the tools developed, for understanding and managing air quality but also lays the foundation for their broader adoption. As cities worldwide grapple with the health impacts of air pollution, sensor networks offer a scalable, cost-effective solution for monitoring and mitigating these effects. With continued research and innovation, these networks could become a cornerstone of global efforts to ensure cleaner, healthier air for all.

References

Aarhaug, T. A., Bacquart, T., Boyd, R., & Daniels, C. (2024). Review of sampling and analysis of particulate matter in hydrogen fuel. *International Journal of Hydrogen Energy*, *49*, 1293–1305. https://doi.org/10.1016/J.IJHYDENE.2023.09.225

Aas, W., Platt, S., & Yttri, K. E. (2018). *EMEP/ACTRIS intensive measurement period-Winter 2018*.

Ahlberg, E., Ausmeel, S., Nilsson, L., Spanne, M., Pauraite, J., Klenø Nøjgaard, J., Bertò, M., Skov, H., Roldin, P., Kristensson, A., Swietlicki, E., & Eriksson, A. (2023). Measurement report: Black carbon properties and concentrations in southern Sweden urban and rural air-the importance of long-range transport. *Atmospheric Chemistry and Physics*, *23*(5), 3051–3064. https://doi.org/10.5194/ACP-23-3051-2023

Air Pollution Act, Pub. L. No. S.I. No. 6/1987 (1987). https://www.irishstatutebook.ie/eli/1987/act/6/enacted/en/print#sec46

Air Pollution Act (Marketing, Sale and Distribution of Fuels) Regulations, 1990, Pub. L. No. S.I. No. 123/1990 (1990). https://www.irishstatutebook.ie/eli/1990/si/123/made/en/print?q=air+pollution+act

Air Pollution Act (Retail Sale of Fuels) Regulations 1989, Pub. L. No. S.I. No. 333/1989 (1989). https://www.irishstatutebook.ie/eli/1989/si/333/made/en/print?q=retail+sale

Air Quality Standards Regulations 2011, Pub. L. No. S.I. No. 180/2011 (2011). https://www.irishstatutebook.ie/eli/2011/si/180/made/en/print?q=+Air+Quality+Standards+Regulations+2011

Aitken, J. (1889). I.—On the Number of Dust Particles in the Atmosphere. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*, *35*(1), 1–19. https://doi.org/10.1017/S0080456800017592

Aitken, J. (1895). III—On the Number of Dust Particles in the Atmosphere of certain Places in Great Britain and on the Continent, with Remarks on the Relation between the Amount of Dust and Meteorological Phenomena. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*, *37*(1), 17–49. https://doi.org/10.1017/S008045680003249X

Akdi, Y., Gölveren, E., Ünlü, K. D., & Yücel, M. E. (2021). Modelling and forecasting of monthly PM2.5 emission of Paris by periodogram-based time series methodology. *Environmental Monitoring and Assessment*, *193*(10), 1–15. https://doi.org/10.1007/S10661-021-09399-Y/TABLES/9

Alastuey, A., Querol, X., Aas, W., Lucarelli, F., Pérez, N., Moreno, T., Cavalli, F., Areskoug, H., Balan, V., Catrambone, M., Ceburnis, D., Cerro, J. C., Conil, S., Gevorgyan, L., Hueglin, C., Imre, K., Jaffrezo, J.-L., Leeson, S. R., Mihalopoulos, N., … Yttri, K. E. (2016). Geochemistry of PM 10 over Europe during the EMEP intensive measurement periods in summer 2012 and winter 2013. *Atmos. Chem. Phys*, *16*, 6107–6129. https://doi.org/10.5194/acp-16-6107-2016

Alfano, B., Barretta, L., Giudice, A. del, Vito, S. de, Francia, G. di, Esposito, E., Formisano, F., Massera, E., Miglietta, M. L., & Polichetti, T. (2020). A Review

of Low-Cost Particulate Matter Sensors from the Developers' Perspectives. *Sensors*, *20*(23), 6819. https://doi.org/https://doi.org/10.3390/s20236819

Allen, G. A., Babich, P., & Poirot, R. L. (2004). Evaluation of a New Approach for Real Time Assessment of Wood Smoke PM. *Air & Waste Management Association Visibility Specialty Conference on Regional and Global Perspectives on Haze: Causes, Consequences and Controversies*.

Ammann, M., Kalberer, M., Jost, D. T., Tobler, L., Rössler, E., Piguet, D., Gäggeler, H. W., & Baltensperger, U. (1998). Heterogeneous production of nitrous acid on soot in polluted air masses. *Nature 1998 395:6698*, *395*(6698), 157–160. https://doi.org/10.1038/25965

Anderson, J. O., Thundiyil, J. G., & Stolbach, A. (2012). Clearing the Air: A Review of the Effects of Particulate Matter Air Pollution on Human Health. *Journal of Medical Toxicology*, *8*(2), 166–175. https://doi.org/10.1007/S13181-011-0203-1/TABLES/5

Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., & Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognition*, *46*(1), 243–256. https://doi.org/10.1016/J.PATCOG.2012.07.021

Ardon-Dryer, K., Dryer, Y., Williams, J. N., & Moghimi, N. (2020). Measurements of PM2.5 with PurpleAir under atmospheric conditions. *Atmospheric Measurement Techniques*, *13*(10), 5441–5458. https://doi.org/10.5194/AMT-13-5441-2020

Austin, E., Coull, B. A., Zanobetti, A., & Koutrakis, P. (2013). A framework to spatially cluster air pollution monitoring sites in US based on the PM2.5 composition. *Environment International*, *59*, 244–254. https://doi.org/10.1016/J.ENVINT.2013.06.003

Austin, E., Coull, B., Thomas, D., & Koutrakis, P. (2012). A Framework for Identifying Distinct Multipollutant Profiles in Air Pollution Data. *Environment International*, *45*(1), 112. https://doi.org/10.1016/J.ENVINT.2012.04.003

Awokola, B. I., Okello, G., Mortimer, K. J., Jewell, C. P., Erhart, A., & Semple, S. (2020). Measuring Air Quality for Advocacy in Africa (MA3): Feasibility and Practicality of Longitudinal Ambient PM2.5 Measurement Using Low-Cost Sensors. *International Journal of Environmental Research and Public Health 2020, Vol. 17, Page 7243*, *17*(19), 7243. https://doi.org/10.3390/IJERPH17197243

Bachwenkizi, J., Liu, C., Meng, X., Zhang, L., Wang, W., van Donkelaar, A., Martin, R. V., Hammer, M. S., Chen, R., & Kan, H. (2022). Maternal exposure to fine particulate matter and preterm birth and low birth weight in Africa. *Environment International*, *160*, 107053. https://doi.org/10.1016/J.ENVINT.2021.107053

Bai, H., Yan, R., Gao, W., Wei, J., & Seong, M. (2022). Spatial representativeness of PM2.5 monitoring stations and its implication for health assessment. *Air Quality, Atmosphere and Health*, *15*(9), 1571–1581. https://doi.org/10.1007/S11869-022-01202-2/FIGURES/9

Barkjohn, K. K., Gantt, B., & Clements, A. L. (2021). Development and application of a United States-wide correction for PM2.5 data collected with the PurpleAir sensor. *Atmospheric Measurement Techniques*, *14*(6), 4617–4637. https://doi.org/10.5194/amt-14-4617-2021

Barkjohn, K. K., Holder, A. L., Frederick, S. G., & Clements, A. L. (2022). Correction and Accuracy of PurpleAir PM2.5 Measurements for Extreme Wildfire Smoke. *Sensors*, *22*(24), 9669. https://doi.org/10.3390/S22249669/S1

Beamish, F. E. (1949). Inorganic gravimetric analysis. *Analytical Chemistry*, *21*(1), 144–160. https://doi.org/10.1021/AC60025A027

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A. L., Dufresne, J. L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., … Stevens, B. (2020). Bounding Global Aerosol Radiative Forcing of Climate Change. *Reviews of Geophysics*, *58*(1), e2019RG000660. https://doi.org/10.1029/2019RG000660

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, *59*(1), 65–98. https://doi.org/10.1137/141000671

Bezdek, J. C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. In *Pattern Recognition with Fuzzy Objective Function Algorithms* (1st ed.). Springer US. https://doi.org/10.1007/978-1-4757-0450-1

Blifford, I. H., & Meeker, G. O. (1967). A factor analysis model of large scale pollution. *Atmospheric Environment (1967)*, *1*(2), 147–157. https://doi.org/10.1016/0004-6981(67)90042-X

Bond, T. C., & Bergstrom, R. W. (2006). Light Absorption by Carbonaceous Particles: An Investigative Review. *Aerosol Science and Technology*, *40*(1), 27–67. https://doi.org/10.1080/02786820500421521

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., Deangelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D., Kinne, S., Kondo, Y., Quinn, P. K., Sarofim, M. C., Schultz, M. G., Schulz, M., Venkataraman, C., Zhang, H., Zhang, S., … Zender, C. S. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, *118*(11), 5380–5552. https://doi.org/10.1002/JGRD.50171

Bond, T. C., Habib, G., & Bergstrom, R. W. (2006). Limitations in the enhancement of visible light absorption due to mixing state. *Journal of Geophysical Research: Atmospheres*, *111*(D20), 20211. https://doi.org/10.1029/2006JD007315

Bousiotis, D., Beddows, D. C. S., Singh, A., Haugen, M., Diez, S., Edwards, P. M., Boies, A., Harrison, R. M., & Pope, F. D. (2022). A study on the performance of low-cost sensors for source apportionment at an urban background site. *Atmospheric Measurement Techniques*, *15*(13), 4047–4061. https://doi.org/10.5194/AMT-15-4047-2022

Bousiotis, D., Singh, A., Haugen, M., Beddows, D. C. S., Diez, S., Murphy, K. L., Edwards, P. M., Boies, A., Harrison, R. M., & Pope, F. D. (2021). Assessing the sources of particles at an urban background site using both regulatory instruments and low-cost sensors-a comparative study. *Atmos. Meas. Tech*, *14*(6), 4139–4155. https://doi.org/10.5194/amt-14-4139-2021

Bové, H., Bongaerts, E., Slenders, E., Bijnens, E. M., Saenen, N. D., Gyselaers, W., Van Eyken, P., Plusquin, M., Roeffaers, M. B. J., Ameloot, M., & Nawrot, T. S. (2019). Ambient black carbon particles reach the fetal side of human placenta. *Nature Communications 2019 10:1*, *10*(1), 1–7. https://doi.org/10.1038/s41467-019-11654-3

Bradford Royal Infirmary. (2024). *Born In Bradford*. https://borninbradford.nhs.uk/news-events/news/new-born-in-bradford-project-is-a-breath-of-fresh-air-for-bradford-school-kids/ (Accessed on 02/09/2024)

Bravo, M. A., & Miranda, M. L. (2022). A longitudinal study of exposure to fine particulate matter during pregnancy, small-for-gestational age births, and birthweight percentile for gestational age in a statewide birth cohort.

Environmental Health : A Global Access Science Source, *21*(1). https://doi.org/10.1186/S12940-021-00823-X

Brimblecombe, P. (2012). *The big smoke: A history of air pollution in London since medieval times* (1st ed.). Taylor and Francis. https://doi.org/10.4324/9780203813898

Brzozowski, K., Ryguła, A., & Maczyński, A. (2019). The use of low-cost sensors for air quality analysis in road intersections. *Transportation Research Part D: Transport and Environment*, *77*, 198–211. https://doi.org/10.1016/j.trd.2019.10.019

Buckley, P. (2019). *Nature and origin of black carbon in Ireland*. University College Cork.

Byrne, R., Ryan, K., Venables, D. S., Wenger, J. C., & Hellebust, S. (2023). Highly local sources and large spatial variations in PM2.5 across a city: evidence from a city-wide sensor network in Cork, Ireland. *Environmental Science: Atmospheres*, *3*(5), 919–930. https://doi.org/10.1039/D2EA00177B

Byrne, R., Wenger, J. C., & Hellebust, S. (2024). Spatial analysis of PM2.5 using a Concentration Similarity Index applied to air quality sensor networks. *Atmospheric Measurement Techniques*, *17*(17), 5129–5146. https://doi.org/https://doi.org/10.5194/amt-17-5129-2024

Cabada, J. C., Khlystov, A., Wittig, A. E., Pilinis, C., & Pandis, S. N. (2004). Light scattering by fine particles during the Pittsburgh Air Quality Study: Measurements and modeling. *Journal of Geophysical Research: Atmospheres*, *109*(D16), 16–19. https://doi.org/10.1029/2003JD004155

Caliński, T., & Harabasz, J. (1974). A Dendrite Method For Cluster Analysis. *Communications in Statistics*, *3*(1), 1–27. https://doi.org/10.1080/03610927408827101

Carotenuto, F., Bisignano, A., Brilli, L., Gualtieri, G., & Giovannini, L. (2023). Low-cost air quality monitoring networks for long-term field campaigns: A review. *Meteorological Applications*, *30*(6), e2161. https://doi.org/10.1002/MET.2161

Carslaw, D. C., & Ropkins, K. (2012). openair — An R package for air quality data analysis. *Environmental Modelling & Software*, *27–28*, 52–61. https://doi.org/10.1016/J.ENVSOFT.2011.09.008

Castell, N., Dauge, F. R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., Broday, D., & Bartonova, A. (2017). Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environment International*, *99*, 293–302. https://doi.org/10.1016/J.ENVINT.2016.12.007

Ceburnis, D., Yin, J., Allen, A. G., Jennings, S. G., Harrison, R. M., Wright, E., Fitzpatrick, M., Healy, T., Barry, E., Yin, J., & McCusker, D. (2006). Local and regional air pollution in Ireland during an intensive aerosol measurement campaign. *Journal of Environmental Monitoring*, *8*(4), 479–487. https://doi.org/10.1039/b516029d

Central Statistics Office (CSO). (2016). *2016 Census*. https://www.cso.ie/en/census/census2016reports/

Central Statistics Office (CSO). (2022). *2022 Census*. https://www.cso.ie/en/statistics/population/censusofpopulation2022/

Cesari, D., Donateo, A., Conte, M., Merico, E., Giangreco, A., Giangreco, F., & Contini, D. (2016). An inter-comparison of PM2.5 at urban and urban background sites: Chemical characterization and source apportionment. *Atmospheric Research*, *174–175*, 106–119. https://doi.org/10.1016/J.ATMOSRES.2016.02.004

Cesaroni, G., Badaloni, C., Gariazzo, C., Stafoggia, M., Sozzi, R., Davoli, M., & Forastiere, F. (2013). Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. *Environmental Health Perspectives*, *121*(3), 324–331. https://doi.org/10.1289/EHP.1205862

Cesaroni, G., Forastiere, F., Stafoggia, M., Andersen, Z. J., Badaloni, C., Beelen, R., Caracciolo, B., De Faire, U., Erbel, R., Eriksen, K. T., Fratiglioni, L., Galassi, C., Hampel, R., Heier, M., Hennig, F., Hilding, A., Hoffmann, B., Houthuijs, D., Jöckel, K. H., … Peters, A. (2014). Long term exposure to ambient air pollution and incidence of acute coronary events: Prospective cohort study and meta-analysis in 11 European cohorts from the escape project. *BMJ (Online)*, *348*. https://doi.org/10.1136/BMJ.F7412

Chatoutsidou, S. E., & Lazaridis, M. (2022). Mass concentrations and elemental analysis of PM2.5 and PM10 in a coastal Mediterranean site: A holistic approach to identify contributing sources and varying factors. *Science of The Total Environment*, *838*, 155980. https://doi.org/10.1016/J.SCITOTENV.2022.155980

Chen, D., Liu, X., Lang, J., Zhou, Y., Wei, L., Wang, X., & Guo, X. (2017). Estimating the contribution of regional transport to PM2.5 air pollution in a rural area on the North China Plain. *Science of The Total Environment*, *583*, 280–291. https://doi.org/10.1016/J.SCITOTENV.2017.01.066

Choi, Y. S., Ho, C. H., Chen, D., Noh, Y. H., & Song, C. K. (2008). Spectral analysis of weekly variation in PM10 mass concentration and meteorological conditions over China. *Atmospheric Environment*, *42*(4), 655–666. https://doi.org/10.1016/J.ATMOSENV.2007.09.075

Chow, J. C. (1995). Measurement Methods to Determine Compliance with Ambient Air Quality Standards for Suspended Particles. *Journal of the Air & Waste Management Association*, *45*(5), 320–382. https://doi.org/10.1080/10473289.1995.10467369

Clarity Movement Co. (2023a). *Clarity Movement Co.* https://www.clarity.io/

Clarity Movement Co. (2023b). *Node-S technical sheet*. https://click.clarity.io/hubfs/Marketing%20Assets%20-%20PDFs/Product%20and%20Specification%20Sheets/Node-S%20Specifications%20Sheet.pdf

Clarity Movement Co. (2023c). *Wind module technical sheet*. https://click.clarity.io/hubfs/Marketing%20Assets%20-%20PDFs/Product%20and%20Specification%20Sheets/2%20Pager%20Flyer%20%2B%20Specifications%20%E2%80%94%C2%A0Wind%20&%20Met%20Module.pdf

Clements, A. L., Griswold, W. G., Abhijit, R. S., Johnston, J. E., Herting, M. M., Thorson, J., Collier-Oxandale, A., & Hannigan, M. (2017). Low-Cost Air Quality Monitoring Tools: From Research to Practice (A Workshop Summary). *Sensors 2017, Vol. 17, Page 2478*, *17*(11), 2478. https://doi.org/10.3390/S17112478

Coker, E. S., Amegah, A. K., Mwebaze, E., Ssematimba, J., & Bainomugisha, E. (2021). A land use regression model using machine learning and locally developed low cost particulate matter sensors in Uganda. *Environmental Research*, *199*, 111352. https://doi.org/10.1016/j.envres.2021.111352

Connerton, P., Nogueira, T., Kumar, P., & Ribeiro, H. (2023). Use of Low-Cost Sensors for Environmental Health Surveillance: Wildfire-Related Particulate Matter Detection in Brasília, Brazil. *Atmosphere*, *14*(12), 1796. https://doi.org/10.3390/ATMOS14121796/S1

Cork City Council, & University College Cork. (2022, September 14). *Cork City Council Air Quality Dashboard*. https://corkairquality.ie (Accessed on 18/04/2023)

Crawford, B., Hagan, D. H., Grossman, I., Cole, E., Holland, L., Heald, C. L., & Kroll, J. H. (2021). Mapping pollution exposure and chemistry during an extreme air quality event (the 2018 Kīlauea eruption) using a low-cost sensor network. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(27), e2025540118. https://doi.org/https://doi.org/10.1073/pnas.2025540118

Crilley, L. R., Bloss, W. J., Yin, J., Beddows, D. C. S., Harrison, R. M., Allan, J. D., Young, D. E., Flynn, M., Williams, P., Zotter, P., Prevot, A. S. H., Heal, M. R., Barlow, J. F., Halios, C. H., Lee, J. D., Szidat, S., & Mohr, C. (2015). Sources and contributions of wood smoke during winter in London: assessing local and regional influences. *Atmos. Chem. Phys*, *15*, 3149–3171. https://doi.org/10.5194/acp-15-3149-2015

Crilley, L. R., Shaw, M., Pound, R., Kramer, L. J., Price, R., Young, S., Lewis, A. C., & Pope, F. D. (2018). Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring. *Atmospheric Measurement Techniques*, *11*(2), 709–720. https://doi.org/10.5194/AMT-11-709-2018

Crilley, L. R., Singh, A., Kramer, L. J., Shaw, M. D., Alam, M. S., Apte, J. S., Bloss, W. J., Hildebrandt Ruiz, L., Fu, P., Fu, W., Gani, S., Gatari, M., Ilyinskaya, E., Lewis, A. C., Ng'ang'a, D., Sun, Y., Whitty, R. C. W., Yue, S., Young, S., & Pope, F. D. (2020). Effect of aerosol composition on the performance of low-cost optical particle counter correction factors. *Atmospheric Measurement Techniques*, *13*(3), 1181–1193. https://doi.org/10.5194/AMT-13-1181-2020

Cuesta-Mosquera, A., Mocnik, G., Drinovec, L., Müller, T., Pfeifer, S., Minguillon, M. C., Briel, B., Buckley, P., Dudoitis, V., Fernández-García, J., Fernandez-Amado, M., De Brito, J. F., Riffault, V., Flentje, H., Heffernan, E., Kalivitis, N., Kalogridis, A. C., Keernik, H., Marmureanu, L., … Wiedensohler, A. (2021). Intercomparison and characterization of 23 Aethalometers under laboratory and ambient air conditions: Procedures and unit-to-unit variabilities. *Atmospheric Measurement Techniques*, *14*(4), 3195–3216. https://doi.org/10.5194/AMT-14-3195-2021

Dall'Osto, M., Hellebust, S., Healy, R. M., Connor, I. P., Kourtchev, I., Sodeau, J. R., Ovadnevaite, J., Ceburnis, D., O'Dowd, C. D., & Wenger, J. C. (2014). Apportionment of urban aerosol sources in Cork (Ireland) by synergistic measurement techniques. *Science of The Total Environment*, *493*, 197–208. https://doi.org/10.1016/J.SCITOTENV.2014.05.027

Dall'Osto, M., Ovadnevaite, J., Ceburnis, D., Martin, D., Healy, R. M., O'Connor, I. P., Kourtchev, I., Sodeau, J. R., Wenger, J. C., & O'Dowd, C. (2013a). Characterization of urban aerosol in Cork city (Ireland) using aerosol mass spectrometry. *Atmospheric Chemistry and Physics*, *13*(9), 4997–5015. https://doi.org/10.5194/ACP-13-4997-2013

Dall'Osto, M., Ovadnevaite, J., Ceburnis, D., Martin, D., Healy, R. M., O'Connor, I. P., Kourtchev, I., Sodeau, J. R., Wenger, J. C., & O'Dowd, C. (2013b). Characterization of urban aerosol in Cork city (Ireland) using aerosol mass spectrometry. *Atmos. Chem. Phys*, *13*, 4997–5015. https://doi.org/10.5194/acp-13-4997-2013

Datta, A., Saha, A., Zamora, M. L., Buehler, C., Hao, L., Xiong, F., Gentner, D. R., & Koehler, K. (2020). Statistical field calibration of a low-cost PM2.5 monitoring

network in Baltimore. *Atmospheric Environment*, *242*, 117761. https://doi.org/10.1016/J.ATMOSENV.2020.117761

de Preux, L., Rizmie, D., Fecht, D., Gulliver, J., & Wang, W. (2023). Does It Measure Up? A Comparison of Pollution Exposure Assessment Techniques Applied across Hospitals in England. *International Journal of Environmental Research and Public Health*, *20*(5), 3852. https://doi.org/10.3390/IJERPH20053852

De Vito, S., Esposito, E., Castell, N., Schneider, P., & Bartonova, A. (2020). On the robustness of field calibration for smart air quality monitors. *Sensors and Actuators, B: Chemical*, *310*, 127869. https://doi.org/10.1016/J.SNB.2020.127869

deSouza, P., Karoline Barkjohn, ab, Clements, A., Lee, J., Kahn, R., Crawford, B., & Kinney, P. (2023). An analysis of degradation in low-cost particulate matter sensors †. *Environ. Sci.: Atmos*, *3*, 521–536. https://doi.org/10.1039/d2ea00142j

Di Antonio, A., Popoola, O. A. M., Ouyang, B., Saffell, J., & Jones, R. L. (2018). Developing a relative humidity correction for low-cost sensors measuring ambient particulate matter. *Sensors (Switzerland)*, *18*(9), 2790. https://doi.org/10.3390/s18092790

Diez, S., Lacy, S., Bannan, T. J., Flynn, M., Gardiner, T., Marsden, N., Martin, N., Read, K., & Edwards, P. M. (2022). Air pollution measurement errors: is your data fit for purpose? *Atmospheric Measurement Techniques*, *15*(13), 4091–4105. https://doi.org/https://doi.org/10.5194/amt-15-4091-2022

Donaldson, K., Stone, V., Seaton, A., & MacNee, W. (2001). Ambient particle inhalation and the cardiovascular system: potential mechanisms. *Environmental Health Perspectives*, *109*, 523–527. https://doi.org/10.1289/EHP.01109S4523

Drinovec, L., Močnik, G., Zotter, P., Prévôt, A. S. H., Ruckstuhl, C., Coz, E., Rupakheti, M., Sciare, J., Müller, T., Wiedensohler, A., & Hansen, A. D. A. (2015). The "dual-spot" Aethalometer: An improved measurement of aerosol black carbon with real-time loading compensation. *Atmospheric Measurement Techniques*, *8*(5), 1965–1979. https://doi.org/10.5194/AMT-8-1965-2015

Dún Laoghaire-Rathdown County Council. (2024). *Air Quality Monitoring*. https://www.dlrcoco.ie/environment/environmental-health/air-quality-monitoring (Accessed on 17/2024)

Dunmore, J. H., Hamilton, R. J., & Smith, D. S. G. (1964). An instrument for the sampling of respirable dust for subsequent gravimetric assessment. *Journal of Scientific Instruments*, *41*(11), 669. https://doi.org/10.1088/0950-7671/41/11/304

Duyzer, J., van den Hout, D., Zandveld, P., & van Ratingen, S. (2015). Representativeness of air quality monitoring networks. *Atmospheric Environment*, *104*, 88–101. https://doi.org/10.1016/J.ATMOSENV.2014.12.067

Eakins, J., Power, B., Dunphy, N., & Sirr, G. (2022). *Residential Solid Fuel Use in Ireland and the Transition Away from Solid Fuels*. ISBN 978-1-80009-035-4

Encyclopaedia of the Environment. (2019). *Airborne particulate matter and their health effects*. https://www.encyclopedie-environnement.org/en/health/airborne-particulate-health-effects/ (Accessed on 12/12/2024)

Environmental Protection Agency. (2015). *National Ambient Air Quality Monitoring Programme 2017-2022*. www.epa.ie (Accessed on 16/08/2024)

Environmental Protection Agency. (2022). *LIFE EMERALD*. https://www.epa.ie/environment-and-you/air/life-emerald/ (Accessed on 12/10/2023)

Environmental Protection Agency. (2023a). *Air Quality in Ireland 2022*. ISBN 978-1-80009-070-5
Environmental Protection Agency. (2023b). *National Ambient Air Quality Monitoring Network*. https://airquality.ie/ (Accessed on 26/10/2023)
Environmental Protection Agency. (2024). *EPA Maps*. 2024. https://gis.epa.ie/EPAMaps (Accessed on 14/08/2024)
Environmental Protection Agency (EPA). (2020). *Air Quality in Ireland 2020*. ISBN 978-1-80009-014-9
Environmental Protection Agency (EPA). (2022). *Air Quality in Ireland 2021*. ISBN 978-1-80009-070-5
Environmental Research Group, & Imperial College London. (2024). *Breathe London*. https://www.breathelondon.org/about (Accessed on 02/09/2024)
Environmental Systems Research Institute, Inc. (2011). *ArcGIS Desktop*. Environmental Systems Research Institute.
European Centre for Medium-Range Weather Forecasts. (2021). *Saharan dust events in the spring of 2021*. https://www.ecmwf.int/en/newsletter/168/news/saharan-dust-events-spring-2021 (Accessed on 17/12/2024)
European Commission. (2018). *Spatial representativeness*. https://environment.ec.europa.eu/topics/air/air-quality/spatial-representativeness_en (Accessed on 30/04/2024)
European Commission. (2019). *Fitness Check of the Ambient Air Quality Directives*. https://commission.europa.eu/publications/fitness-check-eu-ambient-air-quality-directives_en (Accessed on 19/08/2024)
European Commission. (2022). *Proposal for a Directive of the European Parliament and of the Council on ambient air quality and cleaner air for Europe*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0542&qid=1668771169224 (Accessed on 16/08/2024)
European Environment Agency. (2024a). *Europe's air quality status 2024*. https://www.eea.europa.eu/publications/europes-air-quality-status-2024 (Accessed on 19/08/2024)
European Environment Agency. (2024b). *Exceedance of air quality standards in Europe*. https://www.eea.europa.eu/en/analysis/indicators/exceedance-of-air-quality-standards (Accessed on 19/08/2024)
European Environment Agency, & Joint Research Centre. (2007). *FAIRMODE - Forum for Air quality Modeling*. https://fairmode.jrc.ec.europa.eu/activity/ct8 (Accessed on 30/07/2024)
Evangeliou, N., Platt, S. M., Eckhardt, S., Lund Myhre, C., Laj, P., Alados-Arboledas, L., Backman, J., Brem, B. T., Fiebig, M., Flentje, H., Marinoni, A., Pandolfi, M., Yus-Dìez, J., Prats, N., Putaud, J. P., Sellegri, K., Eleftheriadis, K., Vratolis, S., Wiedensohler, A., & Stohl, A. (2021). Changes in black carbon emissions over Europe due to COVID-19 lockdowns. *Atmos. Chem. Phys*, *21*, 2675–2692. https://doi.org/10.5194/acp-21-2675-2021
Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). Cluster analysis: Fifth edition. *Cluster Analysis: Fifth Edition*, 1–330. https://doi.org/10.1002/9780470977811
Fajersztajn, L., Saldiva, P., Pereira, L. A. A., Leite, V. F., & Buehler, A. M. (2017). Short-term effects of fine particulate matter pollution on daily health events in Latin America: a systematic review and meta-analysis. *International Journal of Public Health*, *62*(7), 729–738. https://doi.org/10.1007/S00038-017-0960-Y

Feeney, P., Cahill, T., Olivera, J., & Guidara, R. (1984). Gravimetric Determination of Mass on Lightly Loaded Membrane Filters. *Journal of the Air Pollution Control Association*, *34*(4), 376–378. https://doi.org/10.1080/00022470.1984.10465760/ASSET//CMS/ASSET/A932153B-9158-45DC-8968-360263C03D83/00022470.1984.10465760.FP.PNG

Finlayson-Pitts, B. J., & Pitts, J. N. (2000). *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Applications* (First Ed.). Academic Press. http://www.sciencedirect.com/science/book/9780122570605

Fiordelisi, A., Piscitelli, P., Trimarco, B., Coscioni, E., Iaccarino, G., & Sorriento, D. (2017). The mechanisms of air pollution and particulate matter in cardiovascular diseases. *Heart Failure Reviews*, *22*(3), 337–347. https://doi.org/10.1007/S10741-017-9606-7/FIGURES/1

Flemming, J., Stern, R., & Yamartino, R. J. (2005). A new air quality regime classification scheme for O3, NO2, SO2 and PM10 observations sites. *Atmospheric Environment*, *39*(33), 6121–6129. https://doi.org/10.1016/J.ATMOSENV.2005.06.039

Fowler, D., Brimblecombe, P., Burrows, J., Heal, M. R., Grennfelt, P., Stevenson, D. S., Jowett, A., Nemitz, E., Coyle, M., Lui, X., Chang, Y., Fuller, G. W., Sutton, M. A., Klimont, Z., Unsworth, M. H., & Vieno, M. (2020). A chronology of global air quality. *Philosophical Transactions of the Royal Society A*, *378*(2183). https://doi.org/10.1098/RSTA.2019.0314

Frederickson, L. B., Russell, H. S., Fessa, D., Khan, J., Schmidt, J. A., Johnson, M. S., & Hertel, O. (2023). Hyperlocal air pollution in an urban environment - measured with low-cost sensors. *Urban Climate*, *52*, 101684. https://doi.org/10.1016/J.UCLIM.2023.101684

Frederickson, L. B., Russell, H. S., Raasch, S., Zhang, Z., Schmidt, J. A., Johnson, M. S., & Hertel, O. (2024). Urban vertical air pollution gradient and dynamics investigated with low-cost sensors and large-eddy simulations. *Atmospheric Environment*, *316*, 120162. https://doi.org/10.1016/j.atmosenv.2023.120162

Frederickson, L. B., Sidaraviciute, R., Schmidt, J. A., Hertel, O., & Johnson, M. S. (2022). Are dense networks of low-cost nodes really useful for monitoring air pollution? A case study in Staffordshire. *Atmospheric Chemistry and Physics*, *22*(21), 13949–13965. https://doi.org/10.5194/acp-22-13949-2022

Fuller, G. W., Tremper, A. H., Baker, T. D., Yttri, K. E., & Butterfield, D. (2014). Contribution of wood burning to PM10 in London. *Atmospheric Environment*, *87*, 87–94. https://doi.org/10.1016/J.ATMOSENV.2013.12.037

Fuller, R., Landrigan, P. J., Balakrishnan, K., Bathan, G., Bose-O'Reilly, S., Brauer, M., Caravanos, J., Chiles, T., Cohen, A., Corra, L., Cropper, M., Ferraro, G., Hanna, J., Hanrahan, D., Hu, H., Hunter, D., Janata, G., Kupka, R., Lanphear, B., … Yan, C. (2022). Pollution and health: a progress update. *The Lancet Planetary Health*, *6*(6), e535–e547. https://doi.org/10.1016/S2542-5196(22)00090-0

Gameli Hodoli, C., Coulon, F., & Mead, M. I. (2020). Applicability of factory calibrated optical particle counters for high-density air quality monitoring networks in Ghana. *Heliyon*, *6*(6), e04206. https://doi.org/10.1016/J.HELIYON.2020.E04206

Gameli Hodoli, C., Coulon, F., & Mead, M. I. (2023). Source identification with high-temporal resolution data from low-cost sensors using bivariate polar plots in urban areas of Ghana. *Environmental Pollution*, *317*, 120448. https://doi.org/10.1016/J.ENVPOL.2022.120448

Gao, M., Cao, J., & Seto, E. (2015). A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China. *Environmental Pollution*, *199*, 56–65. https://doi.org/10.1016/j.envpol.2015.01.013

Gentle, J. E., Kaufman, L., & Rousseuw, P. J. (1991). Finding Groups in Data: An Introduction to Cluster Analysis. In *Biometrics* (Vol. 47, Issue 2). John Wiley & Sons, Inc. https://doi.org/10.2307/2532178

Ghosh, S., & Dubey, S. K. (2023). Comparative Analysis of K-Means and Fuzzy C-Means Algorithms. *International Journal of Advanced Computer Science and Applications*, *4*(4), 35–39. https://doi.org/http://dx.doi.org/10.14569/IJACSA.2013.040406

Givati, A., & Rosenfield, D. (2004). Quantifying Precipitation Suppression Due to Air Pollution. *Journal of Applied Meteorology and Climatology*, *43*(7), 1038–1056. https://doi.org/DOI:10.1175/1520-0450(2004)043<1038:QPSDTA>2.0.CO;2

Goetz, A. (1953). Application of Molecular Filter Membranes to the Analysis of Aerosols. *Am J Public Health Nations Health*, *43*(2), 150–159. https://doi.org/10.2105/AJPH.43.2.150

Grahame, T. J., & Schlesinger, R. B. (2010). Cardiovascular health and particulate vehicular emissions: A critical evaluation of the evidence. *Air Quality, Atmosphere and Health*, *3*(1), 3–27. https://doi.org/10.1007/S11869-009-0047-X/TABLES/3

Grange, S. K., Lötscher, H., Fischer, A., Emmenegger, L., & Hueglin, C. (2020). Evaluation of equivalent black carbon source apportionment using observations from Switzerland between 2008 and 2018. *Atmospheric Measurement Techniques*, *13*(4), 1867–1885. https://doi.org/10.5194/AMT-13-1867-2020

Hafkenscheid, T. L., & Vonk, J. (2014). Evaluation of equivalence of the MetOne BAM-1020 for the measurement of PM2.5 in ambient air. *National Institute for Public Health and the Environment*.

Hagan, D. H., Gani, S., Bhandari, S., Patel, K., Habib, G., Apte, J. S., Hildebrandt Ruiz, L., & Kroll, J. H. (2019). Inferring Aerosol Sources from Low-Cost Air Quality Sensor Measurements: A Case Study in Delhi, India. *Environmental Science and Technology Letters*, *6*(8), 467–472. https://doi.org/10.1021/ACS.ESTLETT.9B00393

Hagan, D. H., & Kroll, J. H. (2020). Assessing the accuracy of low-cost optical particle sensors using a physics-based approach. *Atmospheric Measurement Techniques*, *13*(11), 6343–6355. https://doi.org/10.5194/AMT-13-6343-2020

Haldane, J. S. (1931). Atmospheric pollution and fogs. *British Medical Journal*, *1*, 366. https://doi.org/10.1136/BMJ.1.3660.366

Halleux, V. (2024). *Revision of the Ambient Air Quality Directives*. European Parliament - Legislative Train Schedule. https://www.europarl.europa.eu/legislative-train/theme-a-european-green-deal/file-revision-of-eu-ambient-air-quality-legislation (Accessed on 16/08/2024)

Han, W., Li, Z., Guo, J., Su, T., Chen, T., Wei, J., & Cribb, M. (2020). The Urban–Rural Heterogeneity of Air Pollution in 35 Metropolitan Regions across China. *Remote Sensing*, *12*(14), 2320. https://doi.org/10.3390/RS12142320

Hansen, A. D. A., Rosen, H., & Novakov, T. (1984). The aethalometer — An instrument for the real-time measurement of optical absorption by aerosol particles. *Science of The Total Environment*, *36*(C), 191–196. https://doi.org/10.1016/0048-9697(84)90265-1

Hao, X., Zhang, Y., Yu, G., He, B., Yang, F., Zou, Z., Zhang, C., Yang, X., Ouyang, B., & Chang, Y. (2022). Online vertical measurement of air pollutants: Development of a monitoring platform on a skyscraper and its application in Shanghai. *Atmospheric Pollution Research*, *13*(7), 101477. https://doi.org/10.1016/J.APR.2022.101477

Harrison, R. M., Beddows, D. C. S., Jones, A. M., Calvo, A., Alves, C., & Pio, C. (2013). An evaluation of some issues regarding the use of aethalometers to measure woodsmoke concentrations. *Atmospheric Environment*, *80*, 540–548. https://doi.org/10.1016/J.ATMOSENV.2013.08.026

Hays, M. D., Beck, L., Barfield, P., Lavrich, R. J., Dong, Y., & Vander Wal, R. L. (2008). Physical and chemical characterization of residential oil boiler emissions. *Environmental Science and Technology*, *42*(7), 2496–2502. https://doi.org/10.1021/ES071598E

He, M., Kuerbanjiang, N., & Dhaniyala, S. (2020). Performance characteristics of the low-cost Plantower PMS optical sensor. *Aerosol Science and Technology*, *54*(2), 232–241. https://doi.org/10.1080/02786826.2019.1696015

Health Effects Institute. (2020). *State of Global Air 2020*. ISSN 2578-6873, https://www.stateofglobalair.org/sites/default/files/documents/2020-10/soga-2020-report-10-26_0.pdf (Accessed on 17/08/2022)

Healy, R. M., Hellebust, S., Kourtchev, I., Allanic, A., O'Connor, I. P., Bell, J. M., Healy, D. A., Sodeau, J. R., & Wenger, J. C. (2010). Source apportionment of PM2.5 in Cork Harbour, Ireland using a combination of single particle mass spectrometry and quantitative semi-continuous measurements. *Atmospheric Chemistry and Physics*, *10*(19), 9593–9613. https://doi.org/10.5194/ACP-10-9593-2010

Healy, R. M., Sofowote, U., Su, Y., Debosz, J., Noble, M., Jeong, C. H., Wang, J. M., Hilker, N., Evans, G. J., Doerksen, G., Jones, K., & Munoz, A. (2017). Ambient measurements and source apportionment of fossil fuel and biomass burning black carbon in Ontario. *Atmospheric Environment*, *161*, 34–47. https://doi.org/10.1016/J.ATMOSENV.2017.04.034

Heffernan, E. (2022). *Spatial and temporal variation of ambient carbonaceous aerosol in Ireland and strategies for effective monitoring of source contributions*. University College Cork.

Heimann, I., Bright, V. B., McLeod, M. W., Mead, M. I., Popoola, O. A. M., Stewart, G. B., & Jones, R. L. (2015). Source attribution of air pollution by spatial scale separation using high spatial density networks of low cost air quality sensors. *Atmospheric Environment*, *113*, 10–19. https://doi.org/10.1016/J.ATMOSENV.2015.04.057

Helin, A., Virkkula, A., Backman, J., Pirjola, L., Sippula, O., Aakko-Saksa, P., Väätäinen, S., Mylläri, F., Järvinen, A., Bloss, M., Aurela, M., Jakobi, G., Karjalainen, P., Zimmermann, R., Jokiniemi, J., Saarikoski, S., Tissari, J., Rönkkö, T., Niemi, J. V., & Timonen, H. (2021a). Variation of Absorption Ångström Exponent in Aerosols From Different Emission Sources. *Journal of Geophysical Research: Atmospheres*, *126*(10), e2020JD034094. https://doi.org/10.1029/2020JD034094

Helin, A., Virkkula, A., Backman, J., Pirjola, L., Sippula, O., Aakko-Saksa, P., Väätäinen, S., Mylläri, F., Järvinen, A., Bloss, M., Aurela, M., Jakobi, G., Karjalainen, P., Zimmermann, R., Jokiniemi, J., Saarikoski, S., Tissari, J., Rönkkö, T., Niemi, J. V., & Timonen, H. (2021b). Variation of Absorption Ångström Exponent in Aerosols From Different Emission Sources. *Journal of*

Geophysical Research: Atmospheres, *126*(10), e2020JD034094. https://doi.org/10.1029/2020JD034094

Hellebust, S., Allanic, A., O'Connor, I. P., Jourdan, C., Healy, D., & Sodeau, J. R. (2010). Sources of ambient concentrations and chemical composition of PM2.5–0.1 in Cork Harbour, Ireland. *Atmospheric Research*, *95*(2–3), 136–149. https://doi.org/10.1016/J.ATMOSRES.2009.09.006

Hellebust, S., Allanic, A., O'Connor, I. P., Wenger, J. C., & Sodeau, J. R. (2010). The use of real-time monitoring data to evaluate major sources of airborne particulate matter. *Atmospheric Environment*, *44*(8), 1116–1125. https://doi.org/10.1016/j.atmosenv.2009.11.035

Henry, R. C., Lewis, C. W., Hopke, P. K., & Williamson, H. J. (1984). Review of receptor model fundamentals. *Atmospheric Environment*, *18*(8), 1507–1515. https://doi.org/10.1016/0004-6981(84)90375-5

Herich, H., Hueglin, C., & Buchmann, B. (2011). A 2.5 year's source apportionment study of black carbon from wood burning and fossil fuel combustion at urban and rural sites in Switzerland. *Atmospheric Measurement Techniques*, *4*(7), 1409–1420. https://doi.org/10.5194/AMT-4-1409-2011

Hermes, K., Quinting, J., Grams, C. M., Hoose, C., & Hoshyaripour, G. A. (2024). Impact of Saharan dust outbreaks on short-range weather forecast errors in Europe. *Quarterly Journal of the Royal Meteorological Society*, *150*(760), 1704–1723. https://doi.org/10.1002/QJ.4666

Hobbs, P. V. (2000). Introduction to Atmospheric Chemistry. In *Introduction to Atmospheric Chemistry*. Cambridge University Press. https://doi.org/10.1017/CBO9780511808913

Hogrefe, C., Porter, P. S., Gego, E., Gilliland, A., Gilliam, R., Swall, J., Irwin, J., & Rao, S. T. (2006). Temporal features in observed and simulated meteorology and air quality over the Eastern United States. *Atmospheric Environment*, *40*(26), 5041–5055. https://doi.org/10.1016/J.ATMOSENV.2005.12.056

Hogrefe, C., Rao, S. T., Zurbenko, I. G., & Porter, P. S. (2000). Interpreting the information in ozone observations and model predictions relevant to regulatory policies in the Eastern United States. *Bulletin of the American Meteorological Society*, *81*(9), 2083–2106. https://doi.org/10.1175/1520-0477(2000)081<2083:ITIIOO>2.3.CO;2

Holder, A. L., Mebust, A. K., Maghran, L. A., McGown, M. R., Stewart, K. E., Vallano, D. M., Elleman, R. A., & Baker, K. R. (2020). Field Evaluation of Low-Cost Particulate Matter Sensors for Measuring Wildfire Smoke. *Sensors*, *20*(17), 4796. https://doi.org/10.3390/S20174796

Hopke, P. K. (1988). Target transformation factor analysis as an aerosol mass apportionment method: A review and sensitivity study. *Atmospheric Environment*, *22*(9), 1777–1792. https://doi.org/10.1016/0004-6981(88)90066-2

Hopke, P. K. (2016). Review of receptor modeling methods for source apportionment. *Journal of the Air & Waste Management Association*, *66*(3), 237–259. https://doi.org/10.1080/10962247.2016.1140693

Hopke, P. K., Gladney, E. S., Gordon, G. E., Zoller, W. H., & Jones, A. G. (1976). The use of multivariate analysis to identify sources of selected elements in the Boston urban aerosol. *Atmospheric Environment (1967)*, *10*(11), 1015–1025. https://doi.org/10.1016/0004-6981(76)90211-0

Hyder, A., Lee, H. J., Ebisu, K., Koutrakis, P., Belanger, K., & Bell, M. L. (2014). PM2.5 exposure and birth outcomes: use of satellite- and monitor-based data.

Epidemiology (Cambridge, Mass.), *25*(1), 58–67. https://doi.org/10.1097/EDE.0000000000000027

IPCC. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II, and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.*

Janssen, N. A. H., Hoek, G., Simic-Lawson, M., Fischer, P., van Bree, L., Brink, H. Ten, Keuken, M., Atkinson, R. W., Ross Anderson, H., Brunekreef, B., & Cassee, F. R. (2011). Black carbon as an additional indicator of the adverse health effects of airborne particles compared with pm10 and pm2.5. *Environmental Health Perspectives*, *119*(12), 1691–1699. https://doi.org/10.1289/EHP.1003369/SUPPL_FILE/EHP.1003369.S001.PDF

Janssen, S., Tarrason, L., & Ross-Jones, M. (2023). *FAIRMODE WG8-Guidance Document on the estimation of Spatial Representativeness and of Exceedance Situation Indicators*.

Jayaratne, R., Liu, X., Thai, P., Dunbabin, M., & Morawska, L. (2018). The influence of humidity on the performance of a low-cost air particle mass sensor and the effect of atmospheric fog. *Atmospheric Measurement Techniques*, *11*(8), 4883–4890. https://doi.org/10.5194/AMT-11-4883-2018

Jennings, S. G., Ceburnis, D., Allen, A. G., Yin, J., Harrison, R. M., Fitzpatrick, M., Wright, E., Wenger, J., Moriarty, J., Sodeau, J. R., & Barry, E. (2000). *Nature and Origin of PM 10 and Smaller Particulate Matter in Urban Air*. ISBN 1 84095 191 5

Julia Programming Language. (2024). *Clustering.jl* (0.15.7). https://juliastats.org/Clustering.jl/stable/index.html

Jung, C. R., Lin, Y. T., & Hwang, B. F. (2015). Ozone, Particulate Matter, and Newly Diagnosed Alzheimer's Disease: A Population-Based Cohort Study in Taiwan. *Journal of Alzheimer's Disease*, *44*(2), 573–584. https://doi.org/10.3233/JAD-140855

Kahn, R. A., & Gaitley, B. J. (2015). An analysis of global aerosol type as retrieved by MISR. *Journal of Geophysical Research: Atmospheres*, *120*(9), 4248–4281. https://doi.org/10.1002/2015JD023322

Kahn, R. A., Gaitley, B. J., Martonchik, J. V, Diner, D. J., Crean, K. A., Holben, B., Kahn, C. :, Gaitley, B. J., Martonchik, J. V, Diner, D. J., Crean, K. A., & Holben, B. (2005). Multiangle Imaging Spectroradiometer (MISR) global aerosol optical depth validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations. *Journal of Geophysical Research: Atmospheres*, *110*(D10), 1–16. https://doi.org/10.1029/2004JD004706

Karagulian, F., Barbiere, M., Kotsev, A., Spinelle, L., Gerboles, M., Lagler, F., Redon, N., Crunaire, S., & Borowiak, A. (2019). Review of the Performance of Low-Cost Sensors for Air Quality Monitoring. *Atmosphere*, *10*(9), 506. https://doi.org/10.3390/atmos10090506

Karaoghlanian, N., Noureddine, B., Saliba, N., Shihadeh, A., & Lakkis, I. (2022). Low cost air quality sensors "PurpleAir" calibration and inter-calibration dataset in the context of Beirut, Lebanon. *Data in Brief*, *41*, 108008. https://doi.org/10.1016/J.DIB.2022.108008

Kassomenos, P. A., Vardoulakis, S., Chaloulakou, A., Paschalidou, A. K., Grivas, G., Borge, R., & Lumbreras, J. (2014). Study of PM10 and PM2.5 levels in three European cities: Analysis of intra and inter urban variations. *Atmospheric Environment*, *87*, 153–163. https://doi.org/10.1016/J.ATMOSENV.2014.01.004

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data*. https://doi.org/10.1002/9780470316801

Kaur, K., & Kelly, K. E. (2023a). Laboratory evaluation of the Alphasense OPC-N3, and the Plantower PMS5003 and PMS6003 sensors. *Journal of Aerosol Science*, *171*, 106181. https://doi.org/10.1016/J.JAEROSCI.2023.106181

Kaur, K., & Kelly, K. E. (2023b). Performance evaluation of the Alphasense OPC-N3 and Plantower PMS5003 sensor in measuring dust events in the Salt Lake Valley, Utah. *Atmos. Meas. Tech*, *16*(10), 2455–2470. https://doi.org/10.5194/amt-16-2455-2023

Kelleher, S., Quinn, C., Miller-Lionberg, D., & Volckens, J. (2018). A low-cost particulate matter (PM2.5) monitor for wildland fire smoke. *Atmospheric Measurement Techniques*, *11*(2), 1087–1097. https://doi.org/10.5194/AMT-11-1087-2018

Kelly, K. E., Whitaker, J., Petty, A., Widmer, C., Dybwad, A., Sleeth, D., Martin, R., & Butterfield, A. (2017). Ambient and laboratory evaluation of a low-cost particulate matter sensor. *Environmental Pollution*, *221*, 491–500. https://doi.org/10.1016/J.ENVPOL.2016.12.039

Kim, K.-H., Kabir, E., & Kabir, S. (2015). A review on the human health impact of airborne particulate matter. *Environment International*, *74*, 136–143. https://doi.org/https://doi.org/10.1016/j.envint.2014.10.005

Kim Oanh, N. T., Pongkiatkul, P., Upadhyay, N., & Hopke, P. P. (2009). Designing ambient particulate matter monitoring program for source apportionment study by receptor modeling. *Atmospheric Environment*, *43*(21), 3334–3344. https://doi.org/10.1016/J.ATMOSENV.2009.04.016

Kirchstetter, T. W., Novakov, T., & Hobbs, P. V. (2004). Evidence that the spectral dependence of light absorption by aerosols is affected by organic carbon. *Journal of Geophysical Research: Atmospheres*, *109*(D21), 21208. https://doi.org/10.1029/2004JD004999

Kloog, I., Melly, S. J., Ridgway, W. L., Coull, B. A., & Schwartz, J. (2012). Using new satellite based exposure methods to study the association between pregnancy pm2.5 exposure, premature birth and birth weight in Massachusetts. *Environmental Health: A Global Access Science Source*, *11*(1), 1–8. https://doi.org/10.1186/1476-069X-11-40/TABLES/3

Knox, A., Evans, G. J., Brook, J. R., Yao, X., Jeong, C. H., Godri, K. J., Sabaliauskas, K., & Slowik, J. G. (2009). Mass Absorption Cross-Section of Ambient Black Carbon Aerosol in Relation to Chemical Age. *Aerosol Science and Technology*, *43*(6), 522–532. https://doi.org/10.1080/02786820902777207

Kolb, C. E., & Worsnop, D. R. (2012). Chemistry and composition of atmospheric aerosol particles. *Annual Review of Physical Chemistry*, *63*(Volume 63, 2012), 471–491. https://doi.org/10.1146/ANNUREV-PHYSCHEM-032511-143706/

Kosmopoulos, G., Salamalikis, V., Pandis, S. N., Yannopoulos, P., Bloutsos, A. A., & Kazantzidis, A. (2020). Low-cost sensors for measuring airborne particulate matter: Field evaluation and calibration at a South-Eastern European site. *Science of the Total Environment*, *748*, 141396. https://doi.org/10.1016/J.SCITOTENV.2020.141396

Kourtchev, I., Hellebust, S., Bell, J. M., O'Connor, I. P., Healy, R. M., Allanic, A., Healy, D., Wenger, J. C., & Sodeau, J. R. (2011). The use of polar organic compounds to estimate the contribution of domestic solid fuel combustion and biogenic sources to ambient levels of organic carbon and PM2.5 in Cork Harbour,

Ireland. *Science of the Total Environment*, *409*(11), 2143–2155. https://doi.org/10.1016/J.SCITOTENV.2011.02.027

Kracht, O., Santiago, J. L., Martin, F., Piersanti, A., Cremona, G., Righini, G., Vitali, L., Delaney, K., Basu, B., & Ghosh, B. (2017). First outcomes of the fairmode & aquila intercomparison exercise on spatial representativeness. *International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes (HARMO 18)*, 355–359. ineris-01863209

Kracht, O., Santiago, J. L., Martin, F., Piersanti, A., Cremona, G., Righini, G., Vitali, L., Delaney, K., Basu, B., Ghosh, B., Spangl, W., Brendle, C., Latikka, J., Kousa, A., Pärjälä, E., Meretoja, M., Malherbe, L., Letinois, L., Beauchamp, M., … Gerboles, M. (2018). Spatial representativeness of air quality monitoring sites: Outcomes of the FAIRMODE/AQUILA intercomparison exercise. *Publications Office of the European Union*. https://doi.org/10.2760/60611

Krämer, U., Herder, C., Sugiri, D., Strassburger, K., Schikowski, T., Ranft, U., & Rathmann, W. (2010). Traffic-related air pollution and incident type 2 diabetes: results from the SALIA cohort study. *Environmental Health Perspectives*, *118*(9), 1273–1279. https://doi.org/10.1289/EHP.0901689

Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., & Britter, R. (2015). The rise of low-cost sensing for managing air pollution in cities. *Environment International*, *75*, 199–205. https://doi.org/10.1016/J.ENVINT.2014.11.019

Kutzner, R. D., von Schneidemesser, E., Kuik, F., Quedenau, J., Weatherhead, E. C., & Schmale, J. (2018). Long-term monitoring of black carbon across Germany. *Atmospheric Environment*, *185*, 41–52. https://doi.org/10.1016/J.ATMOSENV.2018.04.039

Laamanen, A., & Partanen, T. (1969). Factor Analysis of General and Special Components of Particulates in the Air of Four Finnish Cities on JSTOR. *Work, Environment, Health*, *6*(1), 50–54. https://www.jstor.org/stable/44376282

Lack, D. A., & Cappa, C. D. (2010). Impact of brown and clear carbon on light absorption enhancement, single scatter albedo and absorption wavelength dependence of black carbon. *Atmospheric Chemistry and Physics*, *10*(9), 4207–4220. https://doi.org/10.5194/ACP-10-4207-2010

Lazić, L., Urošević, M. A., Mijić, Z., Vuković, G., & Ilić, L. (2016). Traffic contribution to air pollution in urban street canyons: Integrated application of the OSPM, moss biomonitoring and spectral analysis. *Atmospheric Environment*, *141*, 347–360. https://doi.org/10.1016/J.ATMOSENV.2016.07.008

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., & Pozzer, A. (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature 2015 525:7569*, *525*(7569), 367–371. https://doi.org/10.1038/nature15371

Lenschow, P., Abraham, H. J., Kutzner, K., Lutz, M., Preuß, J. D., & Reichenbächer, W. (2001). Some ideas about the sources of PM10. *Atmospheric Environment*, *35*(SUPPL. 1), S23–S33. https://doi.org/10.1016/S1352-2310(01)00122-4

Li, H. Z., Gu, P., Ye, Q., Zimmerman, N., Robinson, E. S., Subramanian, R., Apte, J. S., Robinson, A. L., & Presto, A. A. (2019). Spatially dense air pollutant sampling: Implications of spatial variability on the representativeness of stationary air pollutant monitors. *Atmospheric Environment: X*, *2*, 100012. https://doi.org/10.1016/J.AEAOA.2019.100012

Li, J., Liu, C., Yin, Y., & Kumar, K. R. (2016). Numerical investigation on the Ångström exponent of black carbon aerosol. *Journal of Geophysical Research: Atmospheres*, *121*(7), 3506–3518. https://doi.org/10.1002/2015JD024718

Lin, C., Ceburnis, D., Hellebust, S., Buckley, P., Wenger, J., Canonaco, F., Prévôt, A. S. H., Huang, R. J., O'Dowd, C., & Ovadnevaite, J. (2017). Characterization of Primary Organic Aerosol from Domestic Wood, Peat, and Coal Burning in Ireland. *Environ Sci Technol*, *51*(18), 10624–10632. https://doi.org/DOI: 10.1021/acs.est.7b01926

Lin, C., Ceburnis, D., Huang, R. J., Canonaco, F., Prévôt, A. S. H., O'Dowd, C., & Ovadnevaite, J. (2019). Summertime Aerosol over the West of Ireland Dominated by Secondary Aerosol during Long-Range Transport. *Atmosphere 2019, Vol. 10, Page 59*, *10*(2), 59. https://doi.org/10.3390/ATMOS10020059

Lin, C., Ceburnis, D., Huang, R. J., Xu, W., Spohn, T., Martin, D., Buckley, P., Wenger, J., Hellebust, S., Rinaldi, M., Cristina Facchini, M., O'Dowd, C., & Ovadnevaite, J. (2019). Wintertime aerosol dominated by solid-fuel-burning emissions across Ireland: Insight into the spatial and chemical variation in submicron aerosol. *Atmospheric Chemistry and Physics*, *19*(22), 14091–14106. https://doi.org/10.5194/ACP-19-14091-2019

Lin, C., Ceburnis, D., Xu, W., Heffernan, E., Hellebust, S., Gallagher, J., Huang, R. J., O'Dowd, C., & Ovadnevaite, J. (2020). The impact of traffic on air quality in Ireland: Insights from the simultaneous kerbside and suburban monitoring of submicron aerosols. *Atmospheric Chemistry and Physics*, *20*(17), 10513–10529. https://doi.org/10.5194/ACP-20-10513-2020

Lin, C., Huang, R. J., Ceburnis, D., Buckley, P., Preissler, J., Wenger, J., Rinaldi, M., Facchini, M. C., O'Dowd, C., & Ovadnevaite, J. (2018). Extreme air pollution from residential solid fuel burning. *Nature Sustainability*, *1*(9), 512–517. https://doi.org/10.1038/s41893-018-0125-x

Liousse, C., Cachier, H., & Jennings, S. G. (1993). Optical and thermal measurements of black carbon aerosol content in different environments: Variation of the specific attenuation cross-section, sigma (σ). *Atmospheric Environment. Part A. General Topics*, *27*(8), 1203–1211. https://doi.org/10.1016/0960-1686(93)90246-U

Liu, C., Chung, C. E., Yin, Y., & Schnaiter, M. (2018). The absorption Ångström exponent of black carbon: From numerical aspects. *Atmospheric Chemistry and Physics*, *18*(9), 6259–6273. https://doi.org/10.5194/ACP-18-6259-2018

Liu, H. Y., Schneider, P., Haugen, R., & Vogt, M. (2019). Performance Assessment of a Low-Cost PM2.5 Sensor for a near Four-Month Period in Oslo, Norway. *Atmosphere 2019, Vol. 10, Page 41*, *10*(2), 41. https://doi.org/10.3390/ATMOS10020041

Liu, X., Jayaratne, R., Thai, P., Kuhn, T., Zing, I., Christensen, B., Lamont, R., Dunbabin, M., Zhu, S., Gao, J., Wainwright, D., Neale, D., Kan, R., Kirkwood, J., & Morawska, L. (2020). Low-cost sensors as an alternative for long-term air quality monitoring. *Environmental Research*, *185*, 109438. https://doi.org/10.1016/J.ENVRES.2020.109438

Local Government Act 2019, Pub. L. No. Act No. 1/2019 (2019). https://www.irishstatutebook.ie/eli/2019/act/1/enacted/en/pdf

Lohmann, U., & Feichter, J. (2005). Global indirect aerosol effects: a review. *Atmos. Chem. Phys*, *5*, 715–737. https://doi.org/https://doi.org/10.5194/acp-5-715-2005

Lung, S. C. C., Wang, W. C. V., Wen, T. Y. J., Liu, C. H., & Hu, S. C. (2020). A versatile low-cost sensing device for assessing PM2.5 spatiotemporal variation

and quantifying source contribution. *Science of The Total Environment*, *716*, 137145. https://doi.org/10.1016/J.SCITOTENV.2020.137145

Ma, R., & Sun, W. C. (2020). FFT-based solver for higher-order and multi-phase-field fracture models applied to strongly anisotropic brittle materials. *Computer Methods in Applied Mechanics and Engineering*, *362*, 112781. https://doi.org/10.1016/J.CMA.2019.112781

Magee Scientific. (2016). *AE33 User Manual*.

Magi, B. I., Cupini, C., Francis, J., Green, M., & Hauser, C. (2020). Evaluation of PM2.5 measured in an urban setting using a low-cost optical particle counter and a Federal Equivalent Method Beta Attenuation Monitor. *Aerosol Science and Technology*, *54*(2), 147–159. https://doi.org/10.1080/02786826.2019.1619915

Malings, C., Tanzer, R., Hauryliuk, A., Saha, P. K., Robinson, A. L., Presto, A. A., & Subramanian, R. (2020). Fine particle mass monitoring with low-cost sensors: Corrections and long-term performance evaluation. *Aerosol Science and Technology*, *54*(2), 160–174. https://doi.org/10.1080/02786826.2019.1623863

Malings, C., Westervelt, D. M., Hauryliuk, A., Presto, A. A., Grieshop, A., Bittner, A., Beekmann, M., & Subramanian, R. (2020). Application of low-cost fine particulate mass monitors to convert satellite aerosol optical depth to surface concentrations in North America and Africa. *Atmospheric Measurement Techniques*, *13*(7), 3873–3892. https://doi.org/10.5194/AMT-13-3873-2020

Marr, L. C., & Harley, R. A. (2002). Spectral analysis of weekday–weekend differences in ambient ozone, nitrogen oxide, and non-methane hydrocarbon time series in California. *Atmospheric Environment*, *36*(14), 2327–2335. https://doi.org/10.1016/S1352-2310(02)00188-7

Masic, A., Bibic, D., Pikula, B., Blazevic, A., Huremovic, J., & Zero, S. (2020). Evaluation of optical particulate matter sensors under realistic conditions of strong and mild urban pollution. *Atmospheric Measurement Techniques*, *13*(12), 6427–6443. https://doi.org/10.5194/AMT-13-6427-2020

Mbengue, S., Serfozo, N., Schwarz, J., Ziková, N., Šmejkalová, A. H., & Holoubek, I. (2020). Characterization of Equivalent Black Carbon at a regional background site in Central Europe: Variability and source apportionment☆. *Environmental Pollution*, *260*, 113771. https://doi.org/10.1016/J.ENVPOL.2019.113771

McFarlane, C., Isevulambire, P. K., Lumbuenamo, R. S., Ndinga, A. M. E., Dhammapala, R., Jin, X., McNeill, V. F., Malings, C., Subramanian, R., & Westervelt, D. M. (2021). First Measurements of Ambient PM2.5 in Kinshasa, Democratic Republic of Congo and Brazzaville, Republic of Congo Using Field-calibrated Low-cost Sensors. *Aerosol and Air Quality Research*, *21*(7), 200619. https://doi.org/10.4209/AAQR.200619

McKinnon, D., & Ribas, R. M. (2024). *Cellular Neurophysiology*. Department of Neurobiology and Behavior, Stony Brook University.

Megaritis, A. G., Fountoukis, C., Charalampidis, P. E., Denier Van Der Gon, H. A. C., Pilinis, C., & Pandis, S. N. (2014). Linking climate and air quality over Europe: effects of meteorology on PM2.5 concentrations. *Atmos. Chem. Phys*, *14*, 10283–10298. https://doi.org/10.5194/acp-14-10283-2014

Met One Instruments. (2008). *BAM-1020*. https://metone.com/products/bam-1020/ (Accessed on 28/02/2024)

Miller, M. S., Friedlander, S. K., & Hidy, G. M. (1972). A chemical element balance for the Pasadena aerosol. *Journal of Colloid and Interface Science*, *39*(1), 165–176. https://doi.org/10.1016/0021-9797(72)90152-X

Molina Rueda, E., Carter, E., L'Orange, C., Quinn, C., & Volckens, J. (2023). Size-Resolved Field Performance of Low-Cost Sensors for Particulate Matter Air Pollution. *Environmental Science and Technology Letters*, *10*(3), 247–253. https://doi.org/10.1021/ACS.ESTLETT.3C00030/ASSET/IMAGES/LARGE/EZ3C00030_0002.JPEG

Moosmüller, H., Chakrabarty, R. K., Ehlers, K. M., & Arnott, W. P. (2011). Absorption Ångström coefficient, brown carbon, and aerosols: Basic concepts, bulk matter, and spherical particles. *Atmospheric Chemistry and Physics*, *11*(3), 1217–1225. https://doi.org/10.5194/ACP-11-1217-2011

Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., Bedini, A., Chai, F., Christensen, B., Dunbabin, M., Gao, J., Hagler, G. S. W., Jayaratne, R., Kumar, P., Lau, A. K. H., Louie, P. K. K., Mazaheri, M., Ning, Z., Motta, N., … Williams, R. (2018). Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone? *Environment International*, *116*, 286–299. https://doi.org/10.1016/J.ENVINT.2018.04.018

Munir, S., Mayfield, M., Coca, D., Jubb, S. A., & Osammor, O. (2019). Analysing the performance of low-cost air quality sensors, their drivers, relative benefits and calibration in cities-a case study in Sheffield. *Environmental Monitoring and Assessment*, *191*(2), 94. https://doi.org/10.1007/S10661-019-7231-8

Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(1), 86–97. https://doi.org/10.1002/WIDM.53

Mustard, W. P., & Gummere, R. M. (1917). Seneca: ad Lucilium Epistulae Morales. *Harvard University Press and William Heinemann Ltd.* https://en.wikisource.org/wiki/Moral_letters_to_Lucilius/Letter_104 (Accessed on 04/07/2024)

Nguyen, N. H., Nguyen, H. X., Le, T. T. B., & Vu, C. D. (2021). Evaluating Low-Cost Commercially Available Sensors for Air Quality Monitoring and Application of Sensor Calibration Methods for Improving Accuracy. *Open Journal of Air Pollution*, *10*(1), 1–17. https://doi.org/10.4236/OJAP.2021.101001

Niedermeier, N., Held, A., Müller, T., Heinold, B., Schepanski, K., Tegen, I., Kandler, K., Ebert, M., Weinbruch, S., Read, K., Lee, J., Fomba, K. W., Müller, K., Herrmann, H., & Wiedensohler, A. (2014). Mass deposition fluxes of Saharan mineral dust to the tropical northeast Atlantic Ocean: An intercomparison of methods. *Atmospheric Chemistry and Physics*, *14*(5), 2245–2266. https://doi.org/10.5194/ACP-14-2245-2014

Nordmann, S., Birmili, W., Weinhold, K., Müller, K., Spindler, G., & Wiedensohler, A. (2013). Measurements of the mass absorption cross section of atmospheric soot particles using Raman spectroscopy. *Journal of Geophysical Research: Atmospheres*, *118*(21), 12,075-12,085. https://doi.org/10.1002/2013JD020021

Okure, D., Ssematimba, J., Sserunjogi, R., Gracia, N. L., Soppelsa, M. E., & Bainomugisha, E. (2022). Characterization of Ambient Air Quality in Selected Urban Areas in Uganda Using Low-Cost Sensing and Measurement Technologies. *Environmental Science and Technology*, *56*(6), 3324–3339. https://doi.org/DOI: 10.1021/acs.est.1c01443

O'Regan, A. C., Byrne, R., Hellebust, S., & Nyhan, M. M. (2022). Associations between Google Street View-derived urban greenspace metrics and air pollution

measured using a distributed sensor network. *Sustainable Cities and Society*, *87*, 104221. https://doi.org/10.1016/J.SCS.2022.104221

Orellano, P., Reynoso, J., Quaranta, N., Bardach, A., & Ciapponi, A. (2020). Short-term exposure to particulate matter (PM10 and PM2.5), nitrogen dioxide (NO2), and ozone (O3) and all-cause and cause-specific mortality: Systematic review and meta-analysis. *Environment International*, *142*, 105876. https://doi.org/10.1016/J.ENVINT.2020.105876

Oudin, A., Forsberg, B., Adolfsson, A. N., Lind, N., Modig, L., Nordin, M., Nordin, S., Adolfsson, R., & Nilsson, L. G. (2016). Traffic-related air pollution and dementia incidence in Northern Sweden: A longitudinal study. *Environmental Health Perspectives*, *124*(3), 306–312. https://doi.org/10.1289/ehp.1408322

Ouimette, J., Arnott, W. P., Laven, P., Whitwell, R., Radhakrishnan, N., Dhaniyala, S., Sandink, M., Tryner, J., Volckens, J., Patrick Arnott, W., & Rosa, S. (2023). Fundamentals of low-cost aerosol sensor design and operation. *Aerosol Science and Technology*, *58*(1), 1–15. https://doi.org/10.1080/02786826.2023.2285935

Ouimette, J., Malm, W. C., Schichtel, B. A., Sheridan, P. J., Andrews, E., Ogren, J. A., & Arnott, W. P. (2022). Evaluating the PurpleAir monitor as an aerosol light scattering instrument. *Atmos. Meas. Tech*, *15*, 655–676. https://doi.org/10.5194/amt-15-655-2022

Ovadnevaite, J., Lin, C., Rinaldi, M., Ceburnis, D., Buckley, P., Coleman, L., Facchini, M. C., Wenger, J., & O'Dowd, C. (2021). *Air Pollution Sources in Ireland*. ISBN 978-1-80009-007-1

Paatero, P., & Tapper, U. (1994). Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, *5*(2), 111–126. https://doi.org/10.1002/ENV.3170050203

Palas. (2024). *Fidas 200*. https://www.palas.de/en/product/fidas200?save=accept&cookie[sessiontracking]=1 (Accessed on 27/08/2024)

Pandis, S. N., Wexler, A. S., & Seinfeld, J. H. (1995). Dynamics of Tropospheric Aerosols. *J. Phys. Chem*, *99*(24), 9646–9659. https://doi.org/DOI: 10.1021/j100024a003

Pedersen, M., Giorgis-Allemand, L., Bernard, C., Aguilera, I., Andersen, A. M. N., Ballester, F., Beelen, R. M. J., Chatzi, L., Cirach, M., Danileviciute, A., Dedele, A., Eijsden, M. van, Estarlich, M., Fernández-Somoano, A., Fernández, M. F., Forastiere, F., Gehring, U., Grazuleviciene, R., Gruzieva, O., … Slama, R. (2013). Ambient air pollution and low birthweight: A European cohort study (ESCAPE). *The Lancet Respiratory Medicine*, *1*(9), 695–704. https://doi.org/10.1016/S2213-2600(13)70192-9

Percival, D. B., & Walden, A. T. (1993). *Spectral Analysis for Physical Applications*. Cambridge University Press. https://doi.org/10.1017/CBO9780511622762

Perillo, H. A., Broderick, B. M., Gill, L. W., McNabola, A., Kumar, P., & Gallagher, J. (2022). Spatiotemporal representativeness of air pollution monitoring in Dublin, Ireland. *Science of The Total Environment*, *827*, 154299. https://doi.org/10.1016/J.SCITOTENV.2022.154299

Petzold, A., Ogren, J. A., Fiebig, M., Laj, P., Li, S. M., Baltensperger, U., Holzer-Popp, T., Kinne, S., Pappalardo, G., Sugimoto, N., Wehrli, C., Wiedensohler, A., & Zhang, X. Y. (2013). Recommendations for reporting black carbon measurements. *Atmospheric Chemistry and Physics*, *13*(16), 8365–8379. https://doi.org/10.5194/ACP-13-8365-2013

Pey, J., Querol, X., & Alastuey, A. (2010). Discriminating the regional and urban contributions in the North-Western Mediterranean: PM levels and composition. *Atmospheric Environment*, *44*(13), 1587–1596. https://doi.org/10.1016/J.ATMOSENV.2010.02.005

Piersanti, A., Vitali, L., Righini, G., Cremona, G., & Ciancarella, L. (2015). Spatial representativeness of air quality monitoring stations: A grid model based approach. *Atmospheric Pollution Research*, *6*(6), 953–960. https://doi.org/10.1016/J.APR.2015.04.005

Plantower. (2022). *PMS5003*. https://www.plantower.com/en/products_33/74.html

Plantower. (2024). *PMS6003*. https://www.plantower.com/en/products_33/75.html (Accessed on 13/09/2024)

Pope, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., & Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA*, *287*(9), 1132–1141. https://doi.org/10.1001/JAMA.287.9.1132

Pope, C. A., Coleman, N., Pond, Z. A., & Burnett, R. T. (2020). Fine particulate air pollution and human mortality: 25+ years of cohort studies. *Environmental Research*, *183*, 108924. https://doi.org/10.1016/J.ENVRES.2019.108924

Pope, C. A., & Dockery, D. W. (2012). Health Effects of Fine Particulate Air Pollution: Lines that Connect. *Journal of the Air & Waste Management Association*, *56*(6), 709–742. https://doi.org/10.1080/10473289.2006.10464485

Pope, F. D., Gatari, M., Ng'ang'a, D., Poynter, A., & Blake, R. (2018). Airborne particulate matter monitoring in Kenya using calibrated low-cost sensors. *Atmospheric Chemistry and Physics*, *18*(20), 15403–15418. https://doi.org/10.5194/ACP-18-15403-2018

Pöschl, U. (2005). Atmospheric Aerosols: Composition, Transformation, Climate and Health Effects. *Angewandte Chemie International Edition*, *44*(46), 7520–7540. https://doi.org/10.1002/ANIE.200501122

Prasad, P., Roja Raman, M., Venkat Ratnam, M., Chen, W. N., Vijaya Bhaskara Rao, S., Gogoi, M. M., Kompalli, S. K., Sarat Kumar, K., & Suresh Babu, S. (2018). Characterization of atmospheric Black Carbon over a semi-urban site of Southeast India: Local sources and long-range transport. *Atmospheric Research*, *213*, 411–421. https://doi.org/10.1016/J.ATMOSRES.2018.06.024

PurpleAir. (2022). *PA-II SD Specs*. https://www2.purpleair.com/collections/air-quality-sensors/products/purpleair-pa-ii-sd

Putaud, J. P., Raes, F., Van Dingenen, R., Brüggemann, E., Facchini, M. C., Decesari, S., Fuzzi, S., Gehrig, R., Hüglin, C., Laj, P., Lorbeer, G., Maenhaut, W., Mihalopoulos, N., Müller, K., Querol, X., Rodriguez, S., Schneider, J., Spindler, G., Ten Brink, H., … Wiedensohler, A. (2004). A European aerosol phenomenology—2: chemical characteristics of particulate matter at kerbside, urban, rural and background sites in Europe. *Atmospheric Environment*, *38*(16), 2579–2595. https://doi.org/10.1016/J.ATMOSENV.2004.01.041

Raaschou-Nielsen, O., Andersen, Z. J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G., Hoffmann, B., Fischer, P., Nieuwenhuijsen, M. J., Brunekreef, B., Xun, W. W., Katsouyanni, K., Dimakopoulou, K., Sommar, J., Forsberg, B., Modig, L., Oudin, A., Oftedal, B., Schwarze, P. E., … Hoek, G. (2013a). Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). *The Lancet. Oncology*, *14*(9), 813–822. https://doi.org/10.1016/S1470-2045(13)70279-1

Raaschou-Nielsen, O., Andersen, Z. J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G., Hoffmann, B., Fischer, P., Nieuwenhuijsen, M. J., Brunekreef, B., Xun, W. W., Katsouyanni, K., Dimakopoulou, K., Sommar, J., Forsberg, B., Modig, L., Oudin, A., Oftedal, B., Schwarze, P. E., … Hoek, G. (2013b). Air pollution and lung cancer incidence in 17 European cohorts: Prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). *The Lancet Oncology*, *14*(9), 813–822. https://doi.org/10.1016/S1470-2045(13)70279-1

Raes, F., Van Dingenen, R., Vignati, E., Wilson, J., Putaud, J. P., Seinfeld, J. H., & Adams, P. (2000). Formation and cycling of aerosols in the global troposphere. *Atmospheric Environment*, *34*(25), 4215–4240. https://doi.org/10.1016/S1352-2310(00)00239-9

Raheja, G., Nimo, J., Appoh, E. K. E., Essien, B., Sunu, M., Nyante, J., Amegah, M., Quansah, R., Arku, R. E., Penn, S. L., Giordano, M. R., Zheng, Z., Jack, D., Chillrud, S., Amegah, K., Subramanian, R., Pinder, R., Appah-Sampong, E., Tetteh, E. N., … Westervelt, D. M. (2023). Low-Cost Sensor Performance Intercomparison, Correction Factor Development, and 2+ Years of Ambient PM2.5 Monitoring in Accra, Ghana. *Environmental Science and Technology*, *57*(29), 10708–10720. https://doi.org/10.1021/ACS.EST.2C09264

Raheja, G., Sabi, K., Sonla, H., Gbedjangni, E. K., McFarlane, C. M., Gameli Hodoli, C., & Westervelt, D. M. (2021). A Network of Field-Calibrated Low-Cost Sensor Measurements of PM2.5in Lomé, Togo, over One to Two Years. *ACS Earth and Space Chemistry*, *6*(4), 1011–1021. https://doi.org/10.1021/acsearthspacechem.1c00391

Rao, S. T., Zurbenko, I. G., Neagu, R., Porter, P. S., Ku, J. Y., & Henry, R. F. (1997). Space and Time Scales in Ambient Ozone Data. *Bulletin of the American Meteorological Society*, 2153–2166. https://doi.org/https://doi.org/10.1175/1520-0477(1997)078<2153:SATSIA>2.0.CO;2

Resch, F., Sunnu, A., & Afeti, G. (2008). Saharan dust flux and deposition rate near the Gulf of Guinea. *The Authors Journal Compilation C*, *60*, 98–105. https://doi.org/10.1111/j.1600-0889.2007.00286.x

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*(C), 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Samoli, E., Stafoggia, M., Rodopoulou, S., Ostro, B., Declercq, C., Alessandrini, E., Díaz, J., Karanasiou, A., Kelessis, A. G., Tertre, A. Le, Pandolfi, P., Randi, G., Scarinzi, C., Zauli-Sajani, S., Katsouyanni, K., Forastiere, F., Alessandrini, E., Angelini, P., Berti, G., … Pascal, M. (2013). Associations between fine and coarse particles and mortality in Mediterranean cities: Results from the MED-PARTICLES project. *Environmental Health Perspectives*, *121*(8), 932–938. https://doi.org/10.1289/EHP.1206124

Sandradewi, J., Prévôt, A. S. H., Szidat, S., Perron, N., Alfarra, M. R., Lanz, V. A., Weingartner, E., & Baltensperger, U. R. S. (2008). Using Aerosol Light Absorption Measurements for the Quantitative Determination of Wood Burning and Traffic Emission Contributions to Particulate Matter. *Environmental Science and Technology*, *42*(9), 3316–3323. https://doi.org/10.1021/ES702253M

Sandradewi, J., Prévôt, A. S. H., Weingartner, E., Schmidhauser, R., Gysel, M., & Baltensperger, U. (2008). A study of wood burning and traffic aerosols in an Alpine valley using a multi-wavelength Aethalometer. *Atmospheric*

Environment, *42*(1), 101–112. https://doi.org/10.1016/J.ATMOSENV.2007.09.034

Sang, S., Chu, C., Zhang, T., Chen, H., & Yang, X. (2022). The global burden of disease attributable to ambient fine particulate matter in 204 countries and territories, 1990–2019: A systematic analysis of the Global Burden of Disease Study 2019. *Ecotoxicology and Environmental Safety*, *238*, 113588. https://doi.org/10.1016/J.ECOENV.2022.113588

Sanyal, S., Rochereau, T., Maesano, C. N., Com-Ruelle, L., & Annesi-Maesano, I. (2018). Long-Term Effect of Outdoor Air Pollution on Mortality and Morbidity: A 12-Year Follow-Up Study for Metropolitan France. *International Journal of Environmental Research and Public Health 2018, Vol. 15, Page 2487*, *15*(11), 2487. https://doi.org/10.3390/IJERPH15112487

Savadkoohi, M., Pandolfi, M., Favez, O., Putaud, J. P., Eleftheriadis, K., Fiebig, M., Hopke, P. K., Laj, P., Wiedensohler, A., Alados-Arboledas, L., Bastian, S., Chazeau, B., María, Á. C., Colombi, C., Costabile, F., Green, D. C., Hueglin, C., Liakakou, E., Luoma, K., … Querol, X. (2024). Recommendations for reporting equivalent black carbon (eBC) mass concentrations based on long-term pan-European in-situ observations. *Environment International*, *185*, 108553. https://doi.org/10.1016/J.ENVINT.2024.108553

Savadkoohi, M., Pandolfi, M., Reche, C., Niemi, J. V., Mooibroek, D., Titos, G., Green, D. C., Tremper, A. H., Hueglin, C., Liakakou, E., Mihalopoulos, N., Stavroulas, I., Artiñano, B., Coz, E., Alados-Arboledas, L., Beddows, D., Riffault, V., De Brito, J. F., Bastian, S., … Querol, X. (2023). The variability of mass concentrations and source apportionment analysis of equivalent black carbon across urban Europe. *Environment International*, *178*, 108081. https://doi.org/10.1016/J.ENVINT.2023.108081

Sayahi, T., Butterfield, A., & Kelly, K. E. (2019). Long-term field evaluation of the Plantower PMS low-cost particulate matter sensors. *Environ Pollut*, *245*, 932–940. https://doi.org/https://doi.org/10.1016/j.envpol.2018.11.065

Schnaiter, M., Horvath, H., Möhler, O., Naumann, K. H., Saathoff, H., & Schöck, O. W. (2003). UV-VIS-NIR spectral optical properties of soot and soot-containing aerosols. *Journal of Aerosol Science*, *34*(10), 1421–1444. https://doi.org/10.1016/S0021-8502(03)00361-6

Schnaiter, M., Linke, C., Möhler, O., Naumann, K. H., Saathoff, H., Wagner, R., Schurath, U., & Wehner, B. (2005). Absorption amplification of black carbon internally mixed with secondary organic aerosol. *Journal of Geophysical Research: Atmospheres*, *110*(D19), 1–11. https://doi.org/10.1029/2005JD006046

Schwarz, J. P., Gao, R. S., Spackman, J. R., Watts, L. A., Thomson, D. S., Fahey, D. W., Ryerson, T. B., Peischl, J., Holloway, J. S., Trainer, M., Frost, G. J., Baynard, T., Lack, D. A., de Gouw, J. A., Warneke, C., & Del Negro, L. A. (2008). Measurement of the mixing state, mass, and optical size of individual black carbon particles in urban and biomass burning emissions. *Geophysical Research Letters*, *35*(13), 13810. https://doi.org/10.1029/2008GL033968

Seinfeld, J. H., & Pandis, S. N. (2016). *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change* (3rd ed.). John Wiley & Sons, Inc. ISBN 978-1-119-22117-3

Sevgi, L. (2007). Numerical fourier transforms: DFT and FFT. *IEEE Antennas and Propagation Magazine*, *49*(3), 238–243. https://doi.org/10.1109/MAP.2007.4293982

Shukla, K., & Aggarwal, S. G. (2022). Overview on Beta-Attenuation Method for the Monitoring of Particulate Matter in Ambient Air. *Aerosol and Air Quality Research*, *22*(12), 220195. https://doi.org/10.4209/aaqr.220195

Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and Its Applications* (4th ed.). Springer International Publishing. https://doi.org/10.1007/978-3-319-52452-8

Singh, A., Ng'ang'a, D., Gatari, M. J., Kidane, A. W., Alemu, Z. A., Derrick, N., Webster, M. J., Bartington, S. E., Thomas, G. N., Avis, W., & Pope, F. D. (2021). Air quality assessment in three East African cities using calibrated low-cost sensors with a focus on road-based hotspots. *Environmental Research Communications*, *3*(7), 075007. https://doi.org/10.1088/2515-7620/AC0E0A

Sippula, O., Hokkinen, J., Puustinen, H., Yli-Pirilä, P., & Jokiniemi, J. (2009). Comparison of particle emissions from small heavy fuel oil and wood-fired boilers. *Atmospheric Environment*, *43*(32), 4855–4864. https://doi.org/10.1016/J.ATMOSENV.2009.07.022

Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S. W., Shelow, D., Hindin, D. A., Kilaru, V. J., & Preuss, P. W. (2013). The changing paradigm of air pollution monitoring. *Environmental Science and Technology*, *47*(20), 11369–11377. https://doi.org/10.1021/ES4022602/ASSET/IMAGES/LARGE/ES-2013-022602_0002.JPEG

Sousa, G., Teixeira, J., Azevedo, R., Almeida, A., Delerue-Matos, C., Morais, S., Wang, X., Rodrigues, F., Freitas, M., & Oliveira, M. (2024). Gravimetric, morphological, and chemical characterization of fine and ultrafine particulate matter inside fire stations. *Building and Environment*, *255*, 111403. https://doi.org/10.1016/J.BUILDENV.2024.111403

Sousan, S., Koehler, K., Thomas, G., Park, J. H., Hillman, M., Halterman, A., & Peters, T. M. (2016). Inter-comparison of low-cost sensors for measuring the mass concentration of occupational aerosols. *Aerosol Science and Technology*, *50*(5), 462–473. https://doi.org/10.1080/02786826.2016.1162901

Spohn, T. K., Martin, D., Geever, M., & O'Dowd, C. (2022). Effect of COVID-19 lockdown on regional pollution in Ireland. *Air Quality, Atmosphere and Health*, *15*(2), 221–234. https://doi.org/10.1007/s11869-021-01098-4

Srivastava, P., Naja, M., Seshadri, T. R., Joshi, H., Dumka, U. C., Gogoi, M. M., & Babu, S. S. (2022). Implications of Site-specific Mass Absorption Cross-section (MAC) to Black Carbon Observations at a High-altitude Site in the Central Himalaya. *Asia-Pacific Journal of Atmospheric Sciences*, *58*(1), 83–96. https://doi.org/https://doi.org/10.1007/s13143-021-00241-6

The European Parliament and of the Council of the European Union. (2008). Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. *Official Journal of the European Union*.

The Government of Ireland. (2023). *Clean Air Strategy*. https://www.gov.ie/en/publication/927e0-clean-air-strategy/ (Accessed on 17/07/2024)

Thermo Fisher. (2018). *Model 1405 TEOM Ambient Air Monitor*. https://www.thermofisher.com/order/catalog/product/TEOM1405/tabs?defaultTab=0 (Accessed on 12/12/2024)

Thunis, P. (2018). On the validity of the incremental approach to estimate the impact of cities on air quality. *Atmospheric Environment*, *173*, 210–222. https://doi.org/10.1016/J.ATMOSENV.2017.11.012

Tomasi, C., & Lupi, A. (2017). Primary and Secondary Sources of Atmospheric Aerosol. In *Atmospheric Aerosols: Life Cycles and Effects on Air Quality and Climate*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9783527336449.CH1

Tranmer, M., Murphy, J., Elliot, M., & Pampaka, M. (2020). *Multiple Linear Regression (2nd Edition)* (1). https://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2020/multiple-linear-regression.pdf (Accessed on 05/09/2024)

Tryner, J., L'Orange, C., Mehaffy, J., Miller-Lionberg, D., Hofstetter, J. C., Wilson, A., & Volckens, J. (2020). Laboratory evaluation of low-cost PurpleAir PM monitors and in-field correction using co-located portable filter samplers. *Atmospheric Environment*, *220*, 117067. https://doi.org/10.1016/J.ATMOSENV.2019.117067

Tryner, J., Mehaffy, J., Miller-Lionberg, D., & Volckens, J. (2020). Effects of aerosol type and simulated aging on performance of low-cost PM sensors. *Journal of Aerosol Science*, *150*, 105654. https://doi.org/10.1016/j.jaerosci.2020.105654

Tunno, B. J., Shields, K. N., Lioy, P., Chu, N., Kadane, J. B., Parmanto, B., Pramana, G., Zora, J., Davidson, C., Holguin, F., & Clougherty, J. E. (2012). Understanding intra-neighborhood patterns in PM2.5 and PM 10 using mobile monitoring in Braddock, PA. *Environmental Health: A Global Access Science Source*, *11*(1), 1–15. https://doi.org/10.1186/1476-069X-11-76/TABLES/2

Turnkey. (2024). *Osiris*. https://turnkey-instruments.com/product/osiris/ (Accessed on 26/02/2024)

Vaughan, N. P., Milligan, B. D., & Ogden, T. L. (1989). FILTER WEIGHING REPRODUCIBILITY AND THE GRAVIMETRIC DETECTION LIMIT. *The Annals of Occupational Hygiene*, *33*(3), 331–337. https://doi.org/10.1093/ANNHYG/33.3.331

Waidyatillake, N. T., Campbell, P. T., Vicendese, D., Dharmage, S. C., Curto, A., & Stevenson, M. (2021). Particulate matter and premature mortality: A Bayesian meta-analysis. *International Journal of Environmental Research and Public Health*, *18*(14), 7655. https://doi.org/10.3390/IJERPH18147655/S1

Wallace, L. (2022). Intercomparison of PurpleAir Sensor Performance over Three Years Indoors and Outdoors at a Home: Bias, Precision, and Limit of Detection Using an Improved Algorithm for Calculating PM2.5. *Sensors*, *22*(7), 2755. https://doi.org/10.3390/S22072755

Wallace, L. (2023a). Cracking the code—Matching a proprietary algorithm for a low-cost sensor measuring PM1 and PM2.5. *Science of The Total Environment*, *893*, 164874. https://doi.org/10.1016/J.SCITOTENV.2023.164874

Wallace, L. (2023b). Testing a New "Decrypted" Algorithm for Plantower Sensors Measuring PM2.5: Comparison with an Alternative Algorithm. *Algorithms*, *16*(8), 392. https://doi.org/10.3390/A16080392

Wallace, L., Bi, J., Ott, W. R., Sarnat, J., & Liu, Y. (2021). Calibration of low-cost PurpleAir outdoor monitors using an improved method of calculating PM2.5. *Atmospheric Environment*, *256*, 118432. https://doi.org/10.1016/J.ATMOSENV.2021.118432

Wang, Q., Gu, J., & Wang, X. (2020). The impact of Sahara dust on air quality and public health in European countries. *Atmospheric Environment*, *241*, 117771. https://doi.org/10.1016/J.ATMOSENV.2020.117771

Wang, Y., de Foy, B., Schauer, J. J., Olson, M. R., Zhang, Y., Li, Z., & Zhang, Y. (2017). Impacts of regional transport on black carbon in Huairou, Beijing, China. *Environmental Pollution*, *221*, 75–84. https://doi.org/10.1016/J.ENVPOL.2016.11.006

Wang, Y., Hopke, P. K., Rattigan, O. V., Chalupa, D. C., & Utell, M. J. (2012). Multiple-year black carbon measurements and source apportionment using Delta-C in Rochester, New York. *Journal of the Air & Waste Management Association*, *62*(8), 880–887. https://doi.org/10.1080/10962247.2012.671792

Wang, Y., Hopke, P. K., Rattigan, O. V., Xia, X., Chalupa, D. C., & Utell, M. J. (2011). Characterization of residential wood combustion particles using the two-wavelength aethalometer. *Environmental Science and Technology*, *45*(17), 7387–7393. https://doi.org/10.1021/es2013984

Wang, Y., Hopke, P. K., Rattigan, O. V., & Zhu, Y. (2011). Characterization of ambient black carbon and wood burning particles in two urban areas. *Journal of Environmental Monitoring*, *13*(7), 1919–1926. https://doi.org/10.1039/C1EM10117J

Wang, Y., Li, J., Jing, H., Zhang, Q., Jiang, J., & Biswas, P. (2015). Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement. *Aerosol Science and Technology*, *49*(11), 1063–1077. https://doi.org/10.1080/02786826.2015.1100710

Wang, Z. T., Zhao, N. B., Wang, W. Y., Tang, R., & Li, S. Y. (2015). A fault diagnosis approach for gas turbine exhaust gas temperature based on fuzzy C-Means clustering and support vector machine. *Mathematical Problems in Engineering*, 240267. https://doi.org/10.1155/2015/240267

Wang, Z., Zhong, S., He, H. di, Peng, Z. R., & Cai, M. (2018). Fine-scale variations in PM2.5 and black carbon concentrations and corresponding influential factors at an urban road intersection. *Building and Environment*, *141*, 215–225. https://doi.org/10.1016/J.BUILDENV.2018.04.042

Warner, R. M. (1998). *Spectral analysis of time-series data* (T. D. Little, Ed.). Guilford Press. ISBN 9781572303386

Watson, J. G., Tropp, R. J., Kohl, S. D., Wang, X., & Chow, J. C. (2017). Filter Processing and Gravimetric Analysis for Suspended Particulate Matter Samples. *Aerosol Science and Engineering*, *1*(2), 93–105. https://doi.org/10.1007/S41810-017-0010-4/FIGURES/2

Weinmayr, G., Romeo, E., de Sario, M., Weiland, S. K., & Forastiere, F. (2010). Short-Term effects of PM10 and NO2 on respiratory health among children with asthma or asthma-like symptoms: A systematic review and Meta-Analysis. *Environmental Health Perspectives*, *118*(4), 449–457. https://doi.org/https://doi.org/10.1289/ehp.0900844

Wenger, J., Arndt, J., Buckley, P., Hellebust, S., Mcgillicuddy, E., O'Connor, I., Sodeau, J., & Wilson, E. (2020). *Source Apportionment of Particulate Matter in Urban and Rural Residential Areas of Ireland (SAPPHIRE)*. ISBN 978-1-84095-905-5

Wieringa, J. (1996). Does representative wind information exist? *Journal of Wind Engineering and Industrial Aerodynamics*, *65*(1–3), 1–12. https://doi.org/10.1016/S0167-6105(97)00017-2

Williams, J., De Reus, M., Krejci, R., Fischer, H., & Ström, J. (2002). Application of the variability-size relationship to atmospheric aerosol studies: Estimating aerosol lifetimes and ages. *Atmospheric Chemistry and Physics*, *2*(2), 133–145. https://doi.org/10.5194/ACP-2-133-2002

Winchester, J. W., & Nifong, G. D. (1971). Water pollution in Lake Michigan by trace elements from pollution aerosol fallout. *Water, Air, and Soil Pollution*, *1*(1), 50–64. https://doi.org/10.1007/BF00280779

Witte, R. S., & Witte, J. S. (2017). *Statistics* (11th ed.). John Wiley & Sons, Inc.

Wojtyla, C., Zielinska, K., Wojtyla-Buciora, P., & Panek, G. (2020). Prenatal Fine Particulate Matter (PM2.5) Exposure and Pregnancy Outcomes—Analysis of Term Pregnancies in Poland. *International Journal of Environmental Research and Public Health*, *17*(16), 5820. https://doi.org/10.3390/IJERPH17165820

World Health Organisation. (2005). *Air Quality Guidelines - Global update 2005*. https://www.who.int/publications/i/item/WHO-SDE-PHE-OEH-06.02 (Accessed on 16/12/2024)

World Health Organisation. (2024). *Ambient air pollution*. https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health (Accessed on 12/12/2024)

World Health Organisation International Agency for Research on Cancer (IARC). (2020). World Cancer Report: Cancer Research for Cancer Prevention. In *Cancer Control*. ISBN 978-92-832-0447-3

World Health Organisation Regional Office for Europe. (1987). *Air quality guidelines for Europe*.

Xue, T., Zhu, T., Geng, G., & Zhang, Q. (2019). Association between pregnancy loss and ambient PM2·5 using survey data in Africa: a longitudinal case-control study, 1998–2016. *The Lancet Planetary Health*, *3*(5), e219-ee225. https://doi.org/10.1016/S2542-5196(19)30047-6

Yang, L. H., Hagan, D. H., Rivera-Rios, J. C., Kelp, M. M., Cross, E. S., Peng, Y., Kaiser, J., Williams, L. R., Croteau, P. L., Jayne, J. T., & Ng, N. L. (2022). Investigating the Sources of Urban Air Pollution Using Low-Cost Air Quality Sensors at an Urban Atlanta Site. *Environmental Science & Technology*, *56*(11), 7063–7073. https://doi.org/10.1021/acs.est.1c07005

Yin, J., Allen, A. G., Harrison, R. M., Jennings, S. G., Wright, E., Fitzpatrick, M., Healy, T., Barry, E., Ceburnis, D., & McCusker, D. (2005). Major component composition of urban PM10 and PM2.5 in Ireland. *Atmospheric Research*, *78*(3–4), 149–165. https://doi.org/10.1016/J.ATMOSRES.2005.03.006

Yong, Z. (2016). *PMS5003 series data manual*. https://aqicn.org/air/sensor/spec/pms5003-english-v2.3.pdf

Zamora, M. L., Buehler, C., Datta, A., Gentner, D. R., & Koehler, K. (2023). Identifying optimal co-location calibration periods for low-cost sensors. *Atmospheric Measurement Techniques*, *16*(1), 169–179. https://doi.org/10.5194/AMT-16-169-2023

Zamora, M. L., Rice, J., & Koehler, K. (2020). One year evaluation of three low-cost PM2.5 monitors. *Atmospheric Environment*, *235*, 117615. https://doi.org/10.1016/j.atmosenv.2020.117615

Zhang, B., Liang, S., Zhao, J., Qian, Z., Bassig, B. A., Yang, R., Zhang, Y., Hu, K., Xu, S., Zheng, T., & Yang, S. (2016). Maternal exposure to air pollutant PM2.5 and PM10 during pregnancy and risk of congenital heart defects. *Journal of Exposure Science & Environmental Epidemiology 2016 26:4*, *26*(4), 422–427. https://doi.org/10.1038/jes.2016.1

Zhang, F., Wang, Y., Peng, J., Chen, L., Sun, Y., Duan, L., Ge, X., Li, Y., Zhao, J., Liu, C., Zhang, X., Zhang, G., Pan, Y., Wang, Y., Zhang, A. L., Ji, Y., Wang, G., Hu, M., Molina, M. J., & Zhang, R. (2020). An unexpected catalyst dominates formation and radiative forcing of regional haze. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(8), 3960–3966. https://doi.org/10.1073/PNAS.1919343117

Zhang, Y., Shi, Z., Wang, Y., Liu, L., Zhang, J., Li, J., Xia, Y., Ding, X., Liu, D., Kong, S., Niu, H., Fu, P., Zhang, X., & Li, W. (2021). Fine particles from village air in northern China in winter: Large contribution of primary organic aerosols from residential solid fuel burning. *Environmental Pollution*, *272*, 116420. https://doi.org/10.1016/J.ENVPOL.2020.116420

Zikova, N., Hopke, P. K., & Ferro, A. R. (2017). Evaluation of new low-cost particle monitors for PM2.5 concentrations measurements. *Journal of Aerosol Science*, *105*, 24–34. https://doi.org/10.1016/J.JAEROSCI.2016.11.010

Zotter, P., Herich, H., Gysel, M., El-Haddad, I., Zhang, Y., Močnik, G., Hüglin, C., Baltensperger, U., Szidat, S., & Prévôt, A. S. H. (2017). Evaluation of the absorption Ångström exponents for traffic and wood burning in the Aethalometer-based source apportionment using radiocarbon measurements of ambient aerosol. *Atmos. Chem. Phys*, *17*, 4229–4249. https://doi.org/10.5194/acp-17-4229-2017

Zou, Y., Clark, J. D., & May, A. A. (2021). Laboratory evaluation of the effects of particle size and composition on the performance of integrated devices containing Plantower particle sensors. *Aerosol Science and Technology*, *55*(7), 848–858. https://doi.org/10.1080/02786826.2021.1905148

Zusman, M., Schumacher, C. S., Gassett, A. J., Spalt, E. W., Austin, E., Larson, T. V., Carvlin, G., Seto, E., Kaufman, J. D., & Sheppard, L. (2020). Calibration of low-cost particulate matter sensors: Model development for a multi-city epidemiological study. *Environment International*, *134*, 105329. https://doi.org/10.1016/J.ENVINT.2019.105329