Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

MODELLING OF AIR QUALITY IN AND AROUND RAILWAY STATION ENVIRONMENTS

A dissertation submitted to the University of Dublin for the requirements for the Degree of Doctor of Philosophy

June 2026

YUXUAN GUO

Supervisor: Prof. Brian Caulfield

Co-Supervisor: Dr. John Gallagher & Prof. Margaret O'Mahony

DEPT OF CIVIL, STRUCTURAL & ENVIRONMENTAL ENGINEERING

DECLARATION

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

I confirm that the work has been completed in full compliance with relevant university policies, including the Academic Integrity policy, the Trinity Policy on Good Research Practice, and the guidance on AI and GenAI in Teaching, Learning, Assessment and Research. I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

Yuxuan Guo

Yuxuan Guo

02/June/2026

SUMMARY

Air pollution is a major health concern, contributing to an estimated 1,300 premature deaths annually in Ireland. Most Irish rail services still use diesel traction, and over 97% of the network is not electrified. Busy, semi-enclosed termini can therefore become local hotspots of particulate matter and nitrogen oxides. This thesis examines air quality in and around two major diesel railway stations in Dublin: Heuston and Connolly. It addresses the lack of detailed emission tools and dispersion models for rail hubs, and the limited evidence on how train idling and operations affect pollution exposure for passengers, staff, and nearby residents. A systematic literature review of air quality modelling at transportation hubs is first conducted to identify research gaps and to guide the choice of statistical, emission and dispersion methods used in the case studies.

Field campaigns then measure $PM_{2.5}$, PM_{10} and NO_2 on platforms and concourses, alongside background air quality and local meteorology. Timetables and on-train recorder data are used to derive hourly train movements and cumulative idling time in different platform zones. Random Forest models show that urban background concentrations set the main baseline, but that idling in the enclosed platform regions is the strongest operational factor driving station increments.

A detailed idling analysis and Bayesian modelling link the distribution of idling time to timetable constraints and driver behaviour. The results show that small changes in layover times or shutdown practices could reduce unnecessary idling. To convert train operational activities into emissions, a physics-based tool is developed that combines fleet information, notch-based emission factors and movement data to produce time- and space-resolved NOx and $PM_{2.5}$ fields for the main diesel fleets. These inventories reveal that station-area emissions form only a small share of national rail totals but are highly concentrated near platforms, depots and approach tracks, with Class 201 locomotives dominating per-train emissions. The same tool is also applied to quantify the emission reduction achieved when Class 201 locomotives operate on Hydrotreated Vegetable Oil instead of diesel.

Finally, an ADMS dispersion model uses the rail and road emission inventories to map pollutant concentrations around Connolly station and to test separate rail and road contributions through scenario runs. Results indicate that station-related increments can be comparable in magnitude to nearby road-traffic increments at the most exposed receptors along the station–road corridor, but are highly localised and strongly dependent on wind direction and distance from rail stations. Together, the thesis provides an integrated framework that links monitoring, statistical modelling, emissions calculation, and dispersion modelling, and it offers practical measures to reduce air pollution from diesel train operations in semi-enclosed urban stations.

ACKNOWLEDGMENTS

Many people have supported and helped me over the past three years, and I would like to express my sincere gratitude to them.

First, I would like to thank my supervisor, Professor Brian Caufield, for his outstanding guidance and support. He is highly experienced in both his research field and PhD supervision. Whenever I felt anxious or hesitant, his encouragement restored my confidence, and our meetings always helped to resolve problems quickly and keep me motivated. I would not have had the confidence to reach this point without him.

I am also very grateful to my co-supervisor and PI of my PhD project, Associate Professor John Gallagher, for his patience, insight and constant support. He not only guided my research but also helped me adjust to life in Dublin. As PI, he ensured I had the technical and financial support needed at every stage of my work. I would also like to thank my co-supervisor, Professor Margaret O'Mahony, especially for her guidance in developing my professional academic writing.

I would also like to sincerely thank my examiners, Associate Professor James Tate and Dr Rui Teixeira, for their time, careful reading of my thesis, and constructive feedback during the examination process.

I gratefully acknowledge the Irish Environmental Protection Agency (EPA) and the National Transport Authority (NTA) for funding my research under the Station_Air project and for providing valuable data. I would like to thank Heidi Hopper Duffy from the Irish Railways and David Kelleghan and Patrick Malon from the EPA for their on-site support and expert guidance.

Thanks to my colleague, Shanmuga Priyan, who collaborated with me and supported me throughout the project. I also thank my office mates, Mengyi Jin and Kiran Apsunde, for making work an enjoyable experience.

I am deeply grateful to my parents, Jufang Yan (闫菊芳) and Jiandong Guo (郭建东), for their unconditional love and support throughout this journey.

Finally, I would like to thank my wife, Qianwen Duan, whose love and support have carried me forward throughout these years—steady and constant.

Contents

Abbreviations

ADMS	Atmospheric Dispersion Modelling System	**ADMS-Airport**	ADMS model configuration for airport-source applications
ADMS-Urban	ADMS model configuration for urban (street-scale) applications	**AERMOD**	American Meteorological Society & U.S. EPA Regulatory Model
AIC	Akaike Information Criterion	**API**	Application Programming Interface
AQ	Air quality	**APU**	Auxiliary power unit
As	Arsenic	**Ba**	Barium
BC	Black carbon	**BIOCHEM**	Bologna limited area model for meteorology and chemistry
CALPUFF	California Puff Modelling System	**CAMx**	Comprehensive Air Quality Model
Cd	Cadmium	**CFD**	Computational Fluid Dynamics
CHM	Canopy height model	**Cl^-**	Chloride ion
CMAQ	Community Multiscale Air Quality	**CMEM**	Comprehensive Modal Emission Model
CO	Carbon monoxide	**CO_2**	Carbon dioxide
Co	Cobalt	**Cr**	Chromium
CR	Cancer risk	**CrI**	Credible interval
Cu	Copper	**C-PORT**	Community model for near-port applications
DMU	Diesel multiple unit	**DOAS**	Differential optical absorption spectroscopy
DSM	Digital surface model	**DTM**	Digital terrain model
EC	Elemental carbon	**ECDF**	Empirical cumulative distribution function
EDMS	Emissions and Dispersion Modelling System	**EEA**	European Environment Agency
EPA	Environmental Protection Agency (Ireland)	**EU**	European Union
FAC2	Fraction of predictions within a factor of two of observations	**Fe**	Iron
FTIR	Fourier-transform infrared	**FUER**	Fuel use and emission rate
GAM	Generalised additive model	**GIS**	Geographic information system
GPS	Global Positioning System	**GSV**	Ground support vehicles
HC	Hydrocarbon	**HEI**	Health Effects Institute
HI	Hazard index	**HVAC**	Heating, ventilation, and air conditioning
HVO	Hydrotreated vegetable oil	**ICR**	InterCity Railcar (Class 22000 DMU)
ITM	Irish Transverse Mercator	**LNG**	Liquefied natural gas
LTO	Landing and take-off	**LUR**	Land use regression
MB	Mean bias	**Mn**	Manganese
MOVES	Motor vehicle emission simulator	**NH_3**	Ammonia
NH_4^+	Ammonium ion	**Ni**	Nickel

NMB	Normalised mean bias	**NME**	Normalised mean error
NO	Nitric oxide (nitrogen monoxide)	**NO_2**	Nitrogen dioxide
NO_3^-	Nitrate ion	**NOx**	Oxides of nitrogen
NTA	National Transport Authority (Ireland)	**O_3**	Ozone
OC	Organic carbon	**OSM**	OpenStreetMap
OTMR	On-Train Monitoring Recorder	**PAHs**	Polycyclic aromatic hydrocarbons
PCA	Principal component analysis	**PEMS**	Portable Emissions Measurement System
Pb	Lead	**PM**	Particulate matter
$PM_{2.5}$	Particulate matter with aerodynamic diameter ≤ 2.5 µm	**PM_{10}**	Particulate matter with aerodynamic diameter ≤ 10 µm
PMF	Positive matrix factorisation	**PNC**	Particle number concentration
R	Correlation coefficient	**R^2**	Coefficient of determination
RANS	Reynolds-averaged Navier–Stokes	**RF**	Random forest
RMSE	Root mean square error	**RPM**	Revolutions per minute
RSSB	Rail Safety and Standards Board	**Sb**	Antimony
SCATS	Sydney Coordinated Adaptive Traffic System	**SO_2**	Sulphur dioxide
SO_4^{2-}	Sulfate ion	**TAPM**	The Air Pollution Model
TC	Timetable constraint	**TEF**	Toxic equivalency factor
TH	Transport hub	**TRAP**	Traffic-related air pollution
UFP	Ultrafine particles	**V**	Vanadium
VOCs	Volatile organic compounds	**WHO**	World Health Organization
WRF	Weather Research and Forecasting model	**WRF-Chem**	WRF coupled with Chemistry
WRF-CMAQ	WRF coupled with CMAQ	**Zn**	Zinc

1. Introduction

1.1 Background

1.1.1 Rail emissions and air pollution

The World Health Organization (WHO) estimates that the combined effects of ambient and household air pollution cause around 6.7–7 million premature deaths every year worldwide—more than one in eight deaths—making air pollution one of the largest environmental health risks globally (WHO, 2024.; WHO Ambient (Outdoor) Air Pollution, 2024). In Europe alone, air pollution caused 518,700 premature deaths in 41 countries due to exposure to $PM_{2.5}$, NO_2, and O_3 in 2018 (EEA, 2020). Additionally, air pollution has been linked to several health issues, including stroke, heart disease, lung cancer, and both chronic and acute respiratory diseases like asthma (WHO, 2022).

Transport emissions significantly contribute to ambient air pollution. In the European Union (EU), the transport sector is responsible for 55.38% of NO_X emissions and 19.85% of $PM_{2.5}$ emissions (EEA, 2019). Within the transport sector, rail accounts for only a small fraction of these pollutants, contributing roughly 3% of transport-related NOx and about 4.5% of particulate matter emissions in Europe (Railways, 2024). While electrification has reduced emissions from rail transportation, 43.9% of Europe's rail lines still depend on diesel, which results in the release of pollutant gases and particulate matter (European Commission, 2021). Diesel-powered trains account for 2.0%, 2.8%, and 2.5% of mobile sources of NO_X, $PM_{2.5}$, and BC in the EU, respectively (Borken-Kleefeld & Ntziachristos, 2012).

Diesel trains release pollutants through exhaust and non-exhaust emissions (Delgado-Saborit et al., 2022). The exhaust emissions come from the diesel engines, which power and provide electricity to trains. The main pollutants from the combustion process can be classified into three groups: nitrogen oxides (NO_X), hydrocarbons and CO, and particulate matter (Yamada et al., 2011). Particulate matter is generated from diesel fuel through pyrolysis, nucleation, surface growth, agglomeration, and attachment (Tree & Svensson, 2007). The particles released in exhaust gases contain carbon and metals, and most of these particles are fine and ultrafine, below 2.5 micrometres (Lim et al., 2021). It is important to note that emissions of air pollutants do not directly increase with the train's speed or distance travelled (Kim et al., 2020; Rahman et al., 2013). In fact, on a g/kWh basis, NO_X and PM levels are typically highest when the engine is idling, meaning more NO_X is produced per unit of power output or fuel consumption at idle than at higher engine power levels (Bohac et al., 2012; Kim et al., 2020).

Diesel exhaust has been classified by WHO as "carcinogenic to humans" (IARC, 2012). Research across various occupations reveals that lung cancer is the most common cancer associated with exposure to diesel exhaust (Silverman et al., 2012; Villeneuve et al., 2011). Beyond carcinogenic effects, diesel exhaust exposure is also linked to non-cancerous outcomes, including neurobehavioral disorders among railroad workers and electrical technicians (Kilburn, 2000). Hart et al. (2006) found that locomotive drivers and conductors exposed to diesel exhaust showed increased mortality rates due to chronic obstructive pulmonary disease. Furthermore, exposure to ultrafine particles has been associated with DNA damage (Liu et al., 2020).

In addition to exhaust emissions, rail transport contributes to particulate matter production through non-exhaust emissions, primarily generated from the friction between wheel and rail, as well as brake pad and disc. Studies have demonstrated that particles from brake wear cause a pro-inflammatory response in human lung epithelial cells (Gasser et al., 2009). Particles from these sources are different from combustion emission particles in terms of size distribution and composition, since they contain a significant amount of metals and minerals (Abbasi et al., 2012). Fridell et al. (2010) calculated the average non-exhaust PM_{10} emission factors for the three types of freight, commuter, and feeder trains, which were 2.9, 0.48 and 0.24 g/train-km, respectively.

1.1.2 Air quality at rail stations

When diesel trains stop or shunt at enclosed or semi-enclosed stations, the dispersion of pollutants from idling and low-speed engines is limited, leading to high pollutant concentrations at stations. In addition to emissions from trains, various sources contribute to the pollution burden at train stations. These include emissions from packaged food stores (Chong et al., 2015; Font et al., 2020) and human activities, such as the resuspension of deposited black carbon particles by passenger movement (Qian et al., 2020). Furthermore, the air quality at train stations is influenced by external pollution from road traffic and other urban and regional sources that infiltrate the stations.

In the United Kingdom, air quality assessments have been conducted at several train stations, including London Paddington (Chong et al., 2015), Birmingham New Street (Hickman et al., 2018; Thornes et al., 2017), Edinburgh Waverley Station, and London King's Cross Station (Font et al., 2020). At Paddington, measurements indicated five hourly NO_2 concentrations exceeding the EU outdoor ambient limit of 200 µg/m³, with generally higher concentrations at the train station compared to a nearby busy roadside (Chong et al., 2015). At Birmingham New Street, NO_2 levels exceeded the 200 μg/m³ threshold in 33-66% of the measurements during the campaign period (Hickman et al., 2018). Human health is at risk at such levels, especially for those working in stations for long periods.

In addition to the human health impacts of air pollution in stations, rail hubs often represent local areas with high levels of air pollution (Munir et al., 2020). Stations with predominantly diesel train services potentially influence local air quality around the station. Although there are no studies specifically focused on rail hubs, there is extensive research on the impacts of other THs (airports, ports and bus stations) on local air quality (Buccolieri et al., 2016; G. Y. Park, 2022; Vichi et al., 2016; Wang et al., 2019; Xu et al., 2023).

1.1.3 Air quality modelling for rail stations

Modelling approaches are typically used for air quality studies at various transport hubs (THs) (Arunachalam et al., 2017; Priyan et al., 2025). Deterministic dispersion models are more commonly employed for large-scale THs, such as airports and ports, where the main emitting vehicles operate outside rather than within the passenger buildings. Used in conjunction with meteorological models and emission inventories, these models compute the spatial and temporal distribution of pollutants emitted in proximity to these THs, facilitating the exploration of the impacts of transport emissions on local air quality. Computational fluid dynamics (CFD) models were used for dispersion modelling of air pollution within bus stations; the enclosed nature of bus terminals reduces computational demands compared to other TH (Yoo et al., 2022a, 2024).

By contrast, there are relatively few air quality modelling studies focused specifically on railway stations (Priyan et al., 2025). Existing work has mainly considered underground or partially enclosed stations, using simplified mass-balance models, box models, or statistical approaches to relate PM and NO_2 concentrations to train traffic, ventilation, and background conditions (Font et al., 2020a; Walther & Bogdan, 2017)(Font et al., 2020; Walther & Bogdan, 2018; Faugier et al., 2023).

A further challenge is that rail emissions data are often reported at spatial and temporal scales that are too coarse for station-level air quality assessment. National or regional emission inventories typically use route-level activity data and fleet-average emission factors, which limit the ability to attribute emissions to specific station areas or operating modes (CLEAR, 2022.; Rastogi & Frey, 2021). This motivates the development of dedicated tools to estimate emissions from diesel trains in station areas at the spatial and temporal resolution required for local air quality and exposure assessments, and to couple these with suitable dispersion models for rail stations.

1.1.4 Railway stations in Ireland

Ireland has electrified 2.75% of its rail network. This compares to an EU average of 53.71%, ranking Ireland last in the league table for countries with rail (Irish Rail Strategy, 2021). In 2022, Ireland used a total of 40.3 million litres of diesel to run the rail network. Diesel fuel accounts for 82% of Ireland's total rail emissions (Irish Rail Annual Report, 2022).

Connolly and Heuston stations rank among Dublin's most frequented rail hubs, each servicing an average of over 500 passenger trains daily. Except for the Dublin Area Rapid Transit, all trains operating at these stations are diesel-powered. The body structures of Connolly and Heuston stations feature semi-enclosed architecture without active ventilation systems. It has been reported that NO_2 levels in Connolly station surpass the European Union's regulatory limit (Government of Ireland, 2021).

In addition, both locations are major THs where public transportation services such as trams, buses, cabs, and private vehicles intersect. In 2019, exceedances of NO_2 standards were recorded at EPA continuous monitoring stations for the local environment in the vicinity of both stations (UTRAP, 2021). Consequently, it is imperative to examine the air quality dynamics in these railway stations and establish models to assess both the internal air quality and its effects on the adjacent environment.

1.2 Research Objectives

This thesis explores air quality in stations and the impacts on the surrounding environment through modelling and provides systematic mitigation measures based on two Dublin stations (Connolly and Heuston) as case studies. The following specific objectives are put forward to achieve the aim of the present research work:

1. Undertake a systematic literature review of modelling approaches to assess air quality at THs and identify literature gaps.

2. Collect air quality and traffic data at Heuston and Connolly stations.

3. Apply statistical modelling methods to evaluate the effect of diesel trains and their idling time on the air quality inside train stations.

4. Develop a diesel train emissions calculation tool that uses train movement recorder data and train emissions factors (based on different engine states) to calculate total emissions for a given area based on train operations.

5. To quantify the spatial distribution of transport-related air pollution (TRAP) around stations, and the relative contributions of station and nearby road emissions, by coupling the diesel train emissions tool with the Atmospheric Dispersion Modelling System (ADMS-Urban).

6. Investigate different technological or policy measures to reduce the impact of air pollution from train stations on local air quality.

1.3 Overall Research Structure

Figure 1.1 shows the overall structure of this research. The work starts from several data sources: air quality monitoring inside the stations, background and meteorological data, train timetables and On-Train Monitoring Recorder (OTMR) records, train emission factors, and nearby road traffic. These datasets provide the information needed on both exposure and activity at the two Dublin terminals, Heuston and Connolly. Together they implement Objective 2, which is to monitor and collect air quality and traffic data at the study sites, and they also supply the input data required for the later modelling objectives (Objectives 3–5). Before the modelling work, Chapter 2 conducts a literature review of air quality modelling at THs and identifies gaps relevant to rail stations.

The central "Modelling/Analysis" column in Figure 1.1 groups the main analytical tasks, which are developed in Chapters 3–6. Statistical air quality models inside stations (Chapter 3) are used to describe pollutant levels and to relate them to background conditions, train movements and idling. This is the main step in addressing Objective 3, which is to quantify the effect of diesel trains and their idling time on air quality inside stations. The idling time analysis (Chapter 4) examines in more detail how long trains remain stationary with engines running and how this behaviour depends on timetable constraints and operations at Connolly and Heuston. This analysis also contributes to Objective 3 by explaining the operational patterns behind the statistical relationships found in Chapter 3.

The high-resolution emission calculation for diesel trains in stations (Chapter 5) converts train movement data and engine-state-specific emission factors into time- and space-resolved emissions in the station area. This work fulfils Objective 4, which is to develop a diesel train emission calculation tool suitable for station-scale applications. Finally, the air quality dispersion models around stations (Chapter 6) combine these rail emissions with road traffic and meteorology to estimate pollutant concentrations in the area surrounding the stations and to separate the contributions from rail and road sources. This step addresses Objective 5, which concerns the spatial distribution of transport-related air pollution around stations.

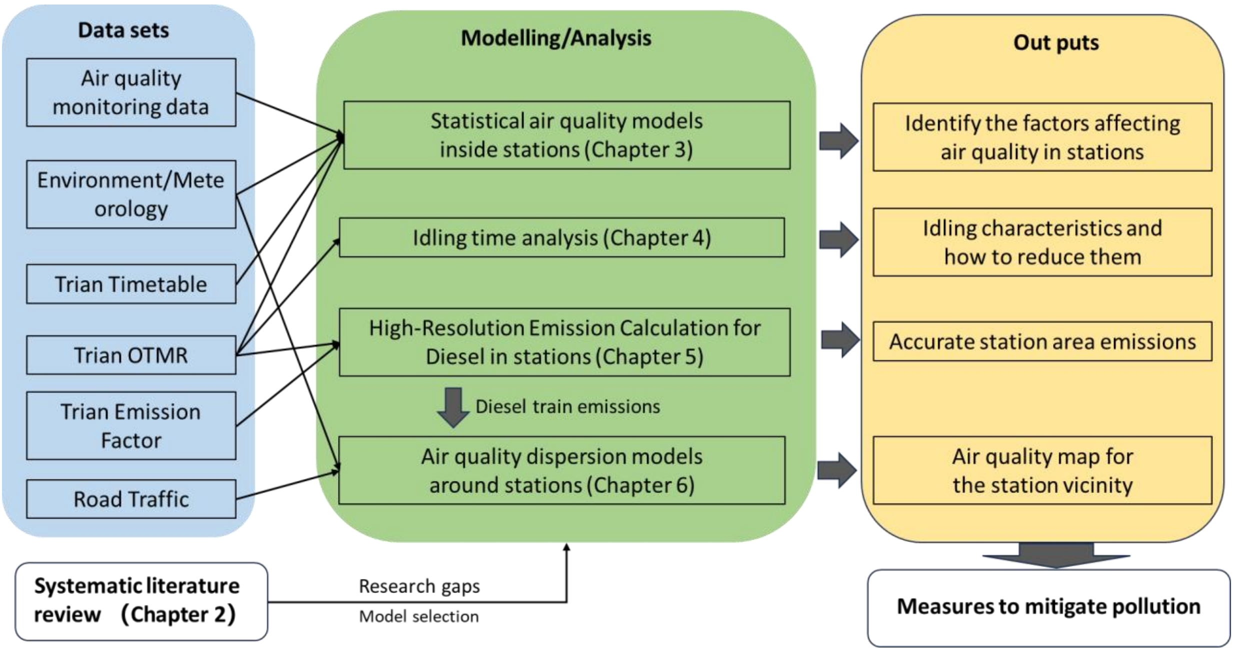


Figure 1.1 Overall research structure

The right-hand "Outputs" column summarises the main outcomes and shows how they feed into the final objective. Identifying the factors that affect air quality in stations, understanding idling characteristics and options to reduce them, and producing accurate station-area emission inventories and air quality maps all follow from the analyses in Chapters 3–6 and correspond to Objectives 3–5. These outputs then provide the evidence base for Objective 6, which is to investigate technological and policy measures to reduce the impact of air pollution from train stations on local air quality.

1.4 Thesis Layout

The thesis is organised into seven chapters, as follows.

Chapter 1 – Introduction

Provides the background to rail emissions and air quality in railway stations, outlines the situation in Ireland, and states the research aim, objectives and overall structure of the study.

Chapter 2 – Systematic Literature Review

Reviews air quality modelling studies for transport hubs such as airports, ports, underground systems, bus terminals and train stations. It summarises statistical, dispersion and traffic emission models, and identifies gaps related to rail stations and combined emissions–dispersion frameworks.

Chapter 3 – Air Quality Modelling in Stations

Describes the monitoring campaigns and train activity datasets for Heuston and Connolly, presents an overview of $PM_{2.5}$, PM_{10} and NO_2 levels, and uses Random Forest models to relate station air quality to background pollution, meteorology, train frequency and idling time.

Chapter 4 – Operational Drivers of Diesel Train Idling in Stations: Timetable Constraints, Driver Behaviour, and Mitigation

Defines idling and timetable constraints, analyses the distribution of idling events for diesel multiple units at both stations, and applies Bayesian regression models to link idling to timetable design and driver behaviour. The chapter then discusses practical options to reduce unnecessary idling.

Chapter 5 – High-Resolution Emission Calculation for Diesel Trains in Stations

Develops a physics-based exhaust calculation tool that uses OTMR data, timetables and emission factors to estimate NO_x and $PM_{2.5}$ emissions in and around the stations. It reports zoned and temporal emission patterns, compares fleets, and assesses emission reductions from switching to HVO fuel.

Chapter 6 – Air Quality Dispersion Modelling Around the Station Area

Uses the emission outputs from Chapter 5, together with road traffic and meteorological inputs, in an ADMS dispersion model to map pollutant concentrations around the stations and evaluate current impacts on nearby streets.

Chapter 7 – Conclusion

Summarises the main findings of the thesis, highlights the contributions to knowledge, discusses implications for rail and air quality policy, and sets out recommendations for further research and for wider application of the methods to the Irish rail system.

2. Literature Review

2.1 Introduction

Ambient air pollution is a complex mixture of particulate and gaseous pollutants from multiple sources. $PM_{2.5}$ and PM_{10} are the main regulatory particle metrics and are linked to a wide range of adverse health outcomes, with $PM_{2.5}$ able to penetrate deep into the lungs and partly enter the bloodstream (Health Effects Institute, 2024a; World Health Organization, 2024a). Primary particles are emitted directly from combustion in transport, power generation, industry and household fuel use, while secondary particles form in the atmosphere from gaseous precursors such as NO_x, SO_2 and VOCs (European Environment Agency, 2023a).

Key gaseous pollutants include NO_2, O_3, SO_2 and CO, all of which have well-documented health impacts. Epidemiological and toxicological evidence links NO_2 exposure to asthma onset and exacerbation, reduced lung function in children, and increased emergency visits and hospital admissions for asthma, with high-certainty evidence for 24-hour NO_2 in recent systematic reviews (Health Effects Institute, 2024a; World Health Organization, 2024a; Zheng et al., 2021). O_3, formed secondarily from NO_x and VOCs in the presence of sunlight, is associated with increased risks of respiratory and cardiovascular mortality and hospital admissions even at relatively low concentrations, as shown by a global time-series analysis in 406 locations and subsequent mortality studies (European Environment Agency, 2023a; Health Effects Institute, 2024a; Vicedo-Cabrera et al., 2020). SO_2 mainly arises from combustion of sulphur-containing fuels and specific industrial processes; short-term rises in SO_2 are consistently associated with increased all-cause and respiratory mortality, and with asthma exacerbations in children and adults, with high or moderate certainty of evidence in recent meta-analyses (Orellano et al., 2021; Zheng et al., 2021). CO, produced by incomplete combustion, reduces the oxygen-carrying capacity of blood via carboxyhaemoglobin formation, and short-term exposure to ambient CO has been linked to increased risk of myocardial infarction and cardiovascular hospital admissions in quantitative reviews (Lee et al., 2020; World Health Organization, 2024a).

TRAP refers to the mixture of pollutants emitted by road traffic, rail, aviation and shipping. It is typically characterised by elevated concentrations of NO_2, ultrafine particles, black carbon and traffic-related $PM_{2.5}$ in near-road environments and THs. Because transport emissions are released close to where people live, work and travel, TRAP often dominates local variation in urban air quality. Systematic reviews and meta-analyses report consistent associations between long-term exposure to TRAP and all-cause and cardiovascular mortality, lung cancer, asthma onset and other respiratory outcomes, with growing evidence for cardiometabolic effects and adverse birth outcomes (Boogaard et al., 2022; Khreis et al., 2024).

The spatial distribution of TRAP is strongly influenced by the built environment and transport system design. High exposures occur along busy roads, at intersections and in street canyons, as well as in enclosed or semi-enclosed microenvironments such as tunnels, underground stations and covered platforms (Khreis et al., 2024). Rail transport generally has lower emissions per passenger-kilometre than private cars or aviation when averaged over whole networks, but diesel-powered rail operations can create localised hotspots in and around stations and depots. In semi-enclosed stations, exhaust from diesel locomotives and multiple units can accumulate on platforms and concourses, especially when trains idle for extended periods, while emissions from adjacent roads add to the local TRAP mixture (Font et al., 2020a; Walther & Bogdan, 2017). These characteristics make rail stations an important, but comparatively understudied, setting for understanding TRAP and its mitigation.

Against this background, the remainder of Chapter 2 reviews air quality studies at different types of THs and summarises the modelling approaches that have been applied to quantify pollutant concentrations and exposure in and around these environments. This provides the conceptual and methodological context for the station-focused monitoring and modelling work presented in later chapters. This chapter is part of a submitted journal paper.

Part of this chapter has been published in Atmospheric Pollution Research (DOI: 10.1016/j.apr.2025.102709). The candidate is a co-first author, and content related to the air quality modelling in this paper is conceived, analysed and written by the candidate.

2.2 Method

The systematic literature review (SLR) was performed to synthesise the research evidence and identify the novelty associated with THs (airports, ports, underground, bus terminals and train stations) pollution. It focuses on modelling approaches for evaluating air quality in THs to identify transferable methods and spatiotemporal considerations relating to similarities and differences for these hubs.

The methodology for this systematic literature review was divided into four steps: literature search, literature screening, literature feature marking, and literature analysis.

Literature search: A literature search was conducted on the Web of Science based on the research topic. Four types of characters as keywords were used in the search: 1) air pollution*, particulate, air quality; 2) rail, train, bus, underground, metro, subway, tram, transport*; 3) station, stop, terminal, port, airport, hub; 4) model*, effect*, assessment, prediction, study, investigate, explore. The literature search was completed in October 2022.

Literature screening: Titles and abstracts are screened according to the following criteria： 1) The subject of the study must be a TH environment, including the area in or around the TH building; 2) The research method must be to build or use air quality or air pollution models. Articles that only monitor air quality or provide a descriptive analysis of air quality were excluded; 3) Models must incorporate environmental factors or traffic-related factors, excluding modelling articles that predict air pollution only from historical data; 4) Only research published in English in peer-reviewed journals was included. Conference abstracts, book chapters and reviews were excluded; 5) Articles published before 2010 were excluded. Sixty-one articles were filtered from 4331 search results to be included in the literature library.

Literature feature marking: Key information was extracted from each article, such as type of TH, study area, time of the study, type of air pollutant, and type of model. All articles were analysed based on the specific model types used in each study. The types of models are divided into the following major categories: Deterministic Mathematical Dispersion Model, Statistical Mode, and Traffic Emission Model.

2.3 Overview of the results

In the realm of TH air quality modelling, a comprehensive literature review, distilled from 4331 search results down to 61 related articles, evidences the field’s dynamic evolution. As detailed in Figure 2.1, there is a clear increased research focus on this subject beginning in 2014, with a significant uptick in the most recent two years. An analysis of the literature reveals ports as the preeminent focus, comprising 38% (23 out of 61) of the selected articles. Airports are a close second in terms of research attention. In contrast, railway stations feature minimally, with two articles - both published post-2020.

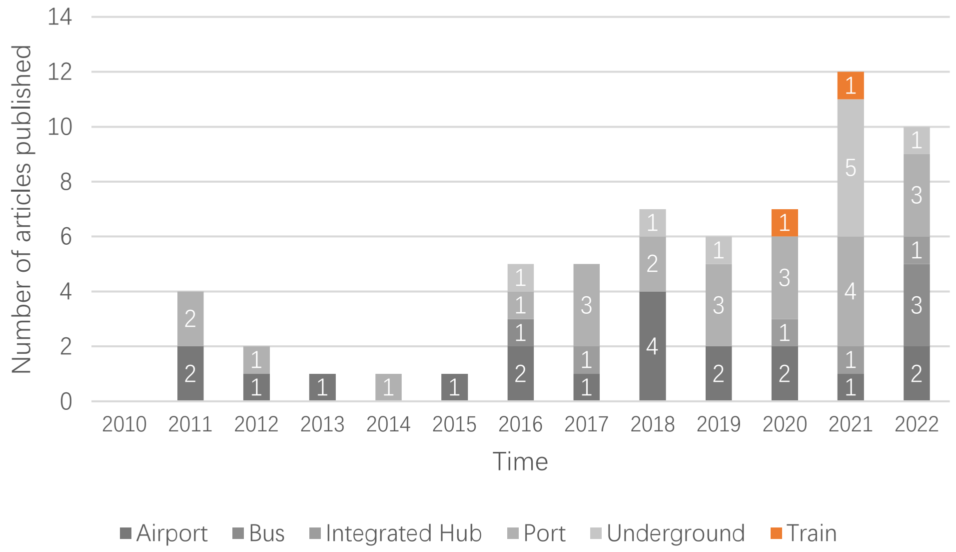


Figure 2.1 Publication dates of the TH air quality modelling literature

The literature can be categorised by model type: statistical models, deterministic dispersion models, and traffic emission models. Figure 2.2 shows which model types are used in the reviewed studies and the transport hubs (THs) to which they have been applied. Deterministic dispersion modelling, predominantly applied to airports and ports, is primarily focused on investigating the dispersion of transport pollutants within the vicinity of these hubs. Statistical models have been instrumental in modelling non-exhaust air pollution at underground stations and airports. Traffic emission models play a crucial role in air quality modelling at bus and train stations.

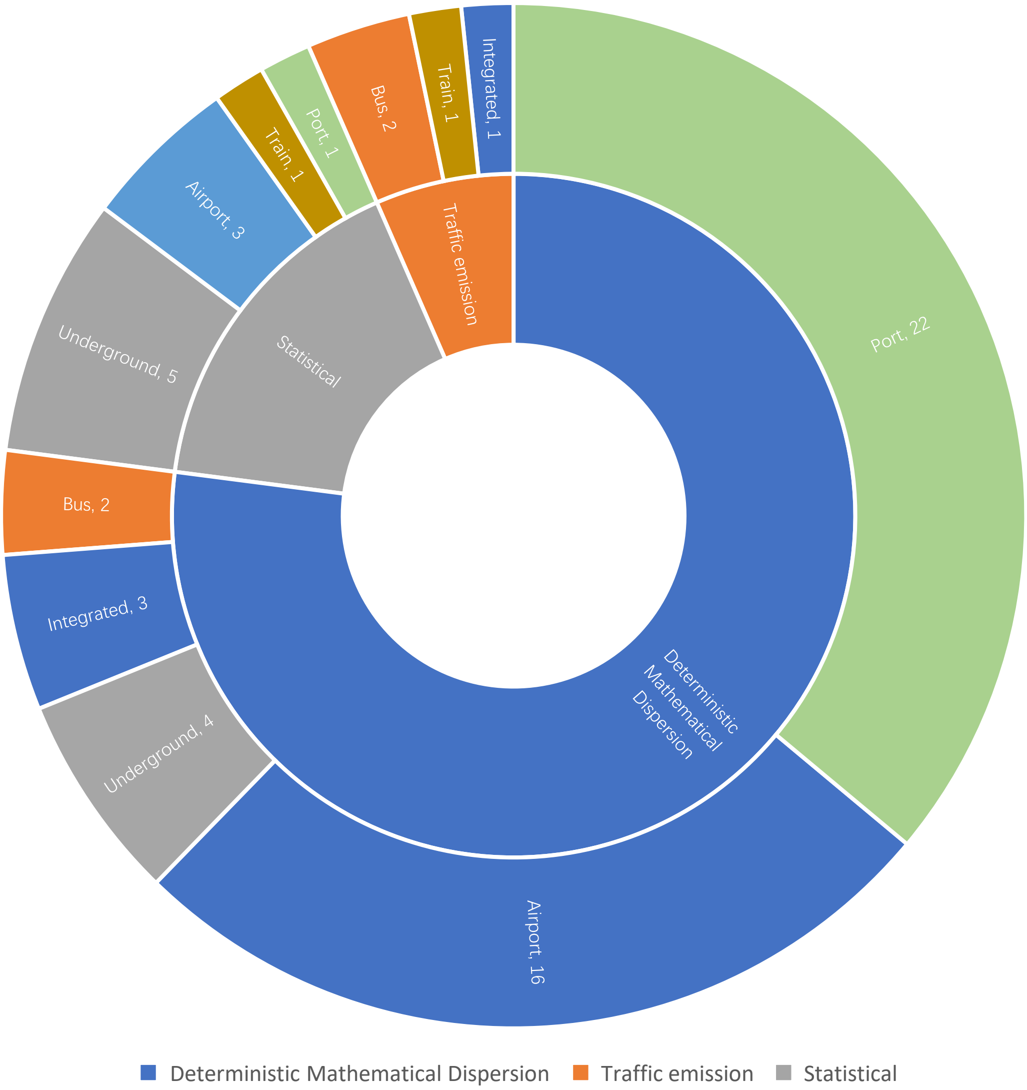


Figure 2.2 Air quality modelling pie charts

2.4 Statistical models

Statistical models play a prominent role in air pollution research at THs, both indoors and in the surrounding outdoor environment. They are used to identify key predictors and to quantify relationships between pollutant concentrations and variables such as background levels, meteorology, traffic activity, and station characteristics, based on historical data. The models applied in the literature range from relatively simple linear and linear mixed models to more flexible approaches, including GAMs, machine-learning methods, and LUR. These approaches differ in their data requirements, ability to capture non-linear relationships and interactions, and the extent to which they provide interpretable effect estimates that can inform mitigation.

Table 2.1 summarises the statistical models that have been used to characterise air quality at THs, distinguishing between applications inside hubs (for example, platforms and concourses) and those in adjacent outdoor areas. The following subsections discuss these model families, drawing out common findings across different hub types and highlighting strengths and limitations that are most relevant for this thesis.

Table 2.1 Statistical models reported in the literature to characterise air quality at THs, either inside the station or adjacent at the TH

Model	Hub Type	Pollutant Type	Data Input	Advantages	Disadvantages
Inside the station					
Multiple linear	Underground Station	Non-Exhaust	Environment, transport, background concentration	Assessment of the degree of influence of each influencing factor on the pollutant using a slope	Only linear relationships can be interpreted; Need to validate and transform variables to satisfy a linear relationship with air quality; Requires a priori knowledge
Random-forest regression	Railway Station	Exhaust	Environment, transport, background concentration	PM explanation Nonlinear relationship	NO explanation; PM Incremental concentration explanation; Unexplained impact factors
Boosted regression trees	Airport	Exhaust	Environment, transport, background concentration	Dealing with continuous and categorical variables Dealing with missing data; No need to transform variables Incorporating non-linear variables	Long-term projections
Artificial neural network	Underground Station	Non-Exhaust	Environment, transport, background concentration, historical pollutant concentration	Short-term prediction of PM concentrations	Unexplained impact factors
Space adjacent to the transport hub					
Land Use Regression	Airport	Exhaust	Environment, transport, background concentration	Higher spatial resolution Low-cost passive; sampling techniques Incorporation of a large number of explanatory variables	Poor performance in the formaldehyde concentration
Deep collaborative learning	Port	Exhaust	Historical pollutant concentration, transport, climate	High accuracy of SO_x and NO_x emissions from ships and O_3 is indirectly affected by ships	Poor performance on particulate concentration from multiple sources
Generalised additive	Airport	Exhaust	Environment, transport, climate	No specific restrictions on the form of distribution of the dependent variable;	Requires independence between

				Better describe the nonlinear influence of independent variables on dependent variables	variables

2.4.1 *Linear regression models*

Linear and linear mixed models have been widely used to analyse non-exhaust PM in underground systems, where the environment is relatively controlled, and a limited set of dominant sources can be identified. Tu et al. (2019) introduced a linear model with terms representing braking effects and cumulative train activity to explore the relationship between train movements and PM concentrations in Stockholm underground stations. The model showed a clear positive association between train frequency and PM_{10} concentrations, with mean prediction errors of around 12% of observed values, but found no linear relationship between $PM_{2.5}$ and the cumulative effect term, indicating that a simple linear structure was not suitable for $PM_{2.5}$.

Subsequent work by Tu and Olofsson (2021a) advanced this line of research by developing a linear mixed “train-type frequency factor” model. This model incorporated fixed effects for train type and hourly equivalent frequency while treating station, date and hour as random effects to capture temporal and spatial autocorrelation. This allowed the authors to isolate the influence of particular train types (for example, CX and C20) on platform PM_{10}, while making efficient use of repeated measurements. Environmental factors were still excluded at this stage.

To address this, Tu & Olofsson (2021b) developed a multiple linear model incorporating fixed effect terms for non-traffic factors, including urban background air quality, passenger flow, ventilation, and night maintenance. This linear mixed model analysed the combined effects of train activity and environmental factors on PM_{10}, $PM_{2.5}$, and PM_1 hourly concentrations. The results demonstrated the model's predictive solid performance for these pollutants. Notably, train frequency exerted a more substantial influence on PM_{10} concentrations compared to the other pollutants examined. Tan et al. (2022) developed a multiple linear regression model to investigate the variation of $PM_{2.5}$, PM_{10}, CO and CO_2 concentrations with underground station characteristics.

Overall, linear and linear mixed models offer an advantage in underground settings because they provide interpretable coefficients that support comparison of the relative importance of predictors such as train frequency, ventilation and background concentrations. Their limitations are also consistent across studies: they require careful variable selection and pre-processing to meet linearity assumptions, and they have limited ability to capture non-linear responses and interactions (for example, threshold effects under stagnant ventilation conditions). These limitations become more acute in airports and ports, where meteorology and multiple sources drive non-linear patterns, motivating the use of GAMs and machine-learning methods in the next subsections.

2.4.2 Non-linear statistical models (GAM, machine learning and spatial statistical mapping)

Linear and mixed models are limited in their ability to represent non-linear responses, interactions, and threshold effects, all of which are common in air-quality data at THs. This limitation becomes particularly important in open-air hubs such as airports and ports, where meteorology and dispersion can dominate short-term variability, and in semi-enclosed environments where accumulation and lagged effects can occur. As a result, the TH literature increasingly applies non-linear statistical approaches—most commonly GAMs and machine-learning models—to improve prediction and attribution when relationships are not well approximated by linear functions.

GAMs extend regression modelling by replacing fixed linear slopes with smooth functions, enabling flexible relationships for key predictors such as wind speed, wind direction, temperature and background concentration while retaining an additive structure that supports interpretation (Hastie & Tibshirani, 1986). Han et al. (2022), for example, used GAM to assess airport influences on NOx while controlling for background concentrations and meteorological variables. Their results suggested that airport activity contributed a measurable share of NOx variability (reported as ~13%), but that background and meteorology accounted for a larger proportion. This reflects a broader pattern across outdoor TH studies: local operational signals are often superimposed on strong regional background and weather-driven dispersion effects, so capturing the non-linear influence of meteorology is essential for isolating hub contributions (Han et al., 2022).

Machine-learning models are adopted for similar reasons, but are typically preferred when predictor sets are large, relationships are highly complex and correlated, or the aim is counterfactual evaluation rather than estimating interpretable coefficients. Carslaw et al. (2012) used boosted regression trees to quantify the effect of a short-term flight ban on NOx concentrations near an airport. By training the model under routine operations and then comparing observed concentrations during the ban to predicted “business-as-usual” concentrations under comparable meteorological conditions, the study demonstrated an attribution strategy that is highly relevant for TH environments, where controlled experiments are rarely feasible (Carslaw et al., 2012).

Park et al. (2018) developed an artificial neural network to predict PM_{10} concentrations in an underground station using both concurrent predictors (outdoor background concentration, train frequency and ventilation rate) and lagged predictors (including the previous hour's station and background concentrations). The improved performance (with 67 ~ 80% of accuracy) when lagged predictors were included is consistent with accumulation, whereby concentrations depend not only on current activity and ventilation but also on prior conditions (Park et al., 2018).

For rail stations, Font et al. (2020) applied Random Forest models to quantify the effects of diesel train numbers and rolling stock type on air quality in two enclosed UK stations. A consistent finding was that background concentration remained the primary explanatory variable for hourly $PM_{2.5}$ and NO_2, with train activity contributing additional variation. Model performance was stronger for $PM_{2.5}$ than for NO_2, and the model struggled to predict incremental $PM_{2.5}$ concentrations relative to ambient levels. The authors attributed part of this limitation to the lack of detailed operational information on idling, highlighting a broader issue for TH attribution: even flexible non-linear models can have limited explanatory power when operational activity is represented only by coarse proxies (e.g., train counts) rather than behavioural metrics (e.g., idling time and engine state) (Font et al., 2020a).

Port applications demonstrate similar dependence on the quality and spatial scale of activity data. Sim et al. (2022) proposed a deep collaborative learning approach based on ConvLSTM to predict air pollution influenced by ship emissions using historical pollutant concentrations, meteorology, and ship movement/emission characteristics. The model performed better for combustion-related pollutants (NOx and SOx) than for $PM_{2.5}$ and PM_{10}, which often have multiple sources and formation pathways. Prediction accuracy also depended on the spatial definition of relevant activity, with best performance reported when ships within approximately 3 km were considered (Sim et al., 2022).

A complementary, but conceptually distinct, statistical approach in the TH literature is land use regression (LUR), which is primarily designed to explain spatial variation in long-term mean concentrations rather than short-term operational effects. Whereas GAM and machine-learning models in TH studies are commonly applied to explain temporal variability at fixed monitoring locations, LUR relates annual or seasonal mean concentrations measured across multiple sites to GIS-derived predictors such as road density, land use and proximity to the hub (Hoek et al., 2008). Gaeta et al. (2016) applied LUR around an airport and found that nearby road traffic, rather than airport activity, predominantly explained spatial patterns for NO_2 and several VOCs, while acrolein showed a clearer airport-related signal. This illustrates a recurring challenge in TH environments:

for common TRAP indicators such as NO_2, strong near-road gradients can dominate spatial variability, limiting the ability of purely spatial statistical models to isolate a hub contribution (Gaeta et al., 2016).

Overall, the non-linear modelling literature indicates that GAMs and machine-learning models are increasingly used across THs to account for complex non-linear dependence on meteorology, background, and operational activity. GAMs offer a balance between flexibility and interpretability, while machine-learning methods often provide stronger predictive performance with reduced transparency in mechanistic interpretation. Across these approaches, a consistent theme is that robust quantification of TH contributions requires careful control for background and meteorology, and that attribution improves substantially when operational activity is represented at sufficient temporal and spatial resolution.

2.5 Deterministic mathematical dispersion model

Deterministic mathematical dispersion modelling has been the most frequently adopted tool employed in atmospheric sciences to support decision making (Genikhovich, 2004). Dispersion models use mathematical expressions to describe the physics and chemistry involved in the ambient (Johansson et al., 2022).

Dispersion modelling is primarily used in studies to evaluate model results against measurements, forest air quality, apportion TH emissions to local air quality, and assess the reduction in pollutant concentrations achieved through the implementation of policy and technological measures (Gao & Zhou, 2024). The commonly reported dispersion models in the THs include Gaussian models, Eulerian & Lagrangian models, and CFD (Gao & Zhou, 2024; Thunis et al., 2016). Table 2.2 provides the deterministic mathematical dispersion models reported in the literature to characterise air quality at THs.

2.5.1 Gaussian models

The Gaussian plume model assumes that pollutants are continuously emitted, representing them as a single expanding plume over time (Johnson, 2022). The pollutant species concentration is directly proportional to the emission rate of the source, diluted by the wind velocity at the source (Gibson et al., 2013). The dispersion of the pollutant is calculated using the standard deviation associated with the Gaussian distribution (Holmes & Morawska, 2006). Gaussian-derived models, which include the American Meteorological Society and U.S. Environmental Protection Agency

Regulatory Model (AERMOD), the Community model for near-port applications (C-Port) and the Atmospheric Dispersion Modelling System (ADMS), are typically employed for AQ research at airports and ports.

AERMOD is a grid-less model that has the potential of predicting sources impact from 1 to 50km away. The model has been used to understand the impact of point and volume pollution sources at a local scale, including the sources which emit benzene, hydrogen cyanide, SO_2 and PM (Gibson et al., 2013; Kalhor & Bajoghli, 2017). Similarly, ADMS (Viz. ADMS-airport & ADMS-urban) can model pollutant concentrations across a range of scales, from street level to city-wide level. ADMS-airport is the most used model in the airport environment, while ADMS-urban is the model mostly employed in the port environment to model the pollutant concentrations in its vicinity (Cesari et al., 2018, 2018; Mokalled et al., 2022). ADMS-airport includes all the features of ADMS-urban, additionally it incorporates the airport specific sources such as air traffic, ground support vehicles, auxiliary power units and other less-defined airport sources. Additionally, ADMS-airport incorporates a jet model which captures the near-field dispersion of the jet exhaust emission during take-off and landing activities (Voogt et al., 2023). Both AERMOD and ADMS incorporate advanced modules such as sulphate chemistry, NO_x – NO_2 chemistry (Holmes & Morawska, 2006).

Table 2.2 Deterministic mathematical dispersion models reported in the literature to characterise air quality at THs

Model Type	Model	Study number	Domain	Spatial resolution	Time resolution	TH	Pollutant	Performance Indicator (NMB)	References
Gaussian	AERMOD	14	Local scale (up to 40km)	100m-200m	Hourly-Annual	Airport	PM_{10}	-2.31 to 11%, R^2: 0.62-0.78	(Abrutytė et al., 2014; Buccolieri et al., 2016; Cesari et al., 2018, 2021; Cohan et al., 2011; Ekmekçioğlu et al., 2020; Gibson et al., 2013; Groma et al., 2018; Henry-Lheureux et al., 2021; Isakov et al., 2017; Kose et al., 2022; Koulidis et al., 2020; S. Kuzu, 2018; S. L. Kuzu et al., 2021; Makridis & Lazaridis, 2019; Mandal et al., 2021; Merico et al., 2019; Mokalled et al., 2022; Popoola et al., 2018; Sabatino et al., 2011; Song & Shon, 2012; Sorte et al., 2019)
							$PM_{2.5}$	-2.2 to 10.4%, R^2: 0.68-0.81	
							NO_X	5 to 13%	
							CO	0.29 to 6.6%, R^2: 0.52-0.72	
						Port	NO_X	-11 to -65%, R^2: 0.4-0.45	
							$PM_{2.5}$	24 to 36%	
	ADMS	6	Local scale (up to 13km)	45m-230m	Hourly	Airport	NO_2	5 to 20.4%, R^2: 0.33-0.65	
							$PM_{2.5}$	-4 to 16.1%	
							O_3	7 to 10%, R^2: 0.65-0.79	
	C-PORT	2	Local scale (up to 10km)	10m-40m	Hourly	-	-	-	
Eulerian	TAMP	3	Local scale (up to 28km)	250m-500m	Hourly-Annual	Airport	NO_2	-11 to -38%, R^2: 0.16-0.25	(Ramacher et al., 2019, 2020; Tang et al., 2020)
							$PM_{2.5}$	-7 to -54%, R^2: 0.25- 0.71	
						Port	NO_2	5 To 15%, R^2:0.73- 0.79	
	CMAQ	5	Regional scale	1km-27km	Hourly-Monthly	Airport	$PM_{2.5}$	-53.9 to 82%	(Rissman et al., 2013; Song et al., 2015; Yang et al., 2018; J. Zhang et al., 2021; Zhou et al., 2020)
							SO_4	-57.4 to -57.6%	
							NO_3	5.4 to 6.1%	
							O_3	3.1 to 20.6%	
	CAMx	2		6km-36km	Monthly	-	-	-	(Bo et al., 2019; Merico et al., 2017)
	WRF-Chem	1		3km	Monthly	Port	$PM_{2.5}$	-14.4 to -21.3%, R^2: 0.59-0.66	(D. Chen et al., 2017)
Lagrangian	SPRAY	2	Local scale (up to 50km)	100m-1km	Hourly	Airport	NO_X	-36 to -89%	(Gaeta et al., 2016; Kontos et al., 2017; Murena et al., 2018; Pecorari et al., 2016; Poplawski et al., 2011)
							CO	-35 to -60.8%	
							HC	-44 to -134%	
	CALPUFF	3	Local scale (up to 30km)	100m-200m	Hourly	Port	SO_2	30 to 52%	
							$PM_{2.5}$	36 to 69%	
							NO_2	-44 to 14%	
CFD		2	Microscale (up to 10m)	Centimetre	Per second	-	-	-	(Yoo et al., 2022b; Yu et al., 2022a)

The plume transport model, AERMOD and ADMS cannot process the variable calm wind condition (Cohan et al., 2011; Makridis & Lazaridis, 2019). A common way to deal with this problem is to replace the meteorological condition with an allowable wind speed that is representative of the physical processes. Wind speed is generally chosen as the smallest non-zero velocity which can be handled by the model and wind directions are equally chosen for all directions for calm winds. If the calm wind condition periods are of low percentage, these periods can be excluded in the modelling to prevent overestimation of results. For instance, Cohan et al. (2011) selected the smallest non-zero velocity of 1 knot as a cut-off value and assumed wind direction equally probable in all directions to account for calm and variable winds in AERMOD. The study highlighted roadway emissions as the major local pollution compared to ship emissions in the ports of Los Angeles and Long Beach.

C-port is another dispersion modelling system which can estimate emissions from the port terminals and helps in assessing the AQ impacts from ports (Isakov et al., 2017). The C -port models the primary pollutants such as SO_2, NO_X, $PM_{2.5}$, CO, formaldehyde, benzene, acrolein, and acetaldehyde (CMAS, 2016). The model represents multiple activities in the port including rail, road, and terminal activities. The model's algorithm for the point and area sources is similar to AERMOD, and the line-source algorithm is consistent with the US EPA C-line model. However, this algorithm is simplified for computational efficiency, as C-port is a web-based application. The C-port model currently uses data from 21 seaports in the U.S. and incorporates the local meteorology of these locations (C-PORT, 2016). Hence, its applicability is limited for use in other port locations worldwide, and the model is not further reviewed in the article.

The study domains used in the studies performed in the port and airport ranged from 6 × 9 km to 40 × 40 km, with grid resolutions between 100 m and 200 m. Both AERMOD and ADMS models can predict hourly pollution levels; however, if only annual or monthly emissions are available, emissions can be parameterised based on traffic data and assume a steady release of emissions. Vutukuru & Dabdub (2008) parameterised emissions such that 70% occur between 8 AM and 8 PM and 30% between 8 PM and 8 AM at the port. Weekend emissions are set at 50% of the weekly average.

Estimating hourly pollution levels is computationally intensive for a larger domain, some studies have restricted the simulation to shorter periods rather than for the whole year (Cohan et al., 2011;

Vutukuru & Dabdub, 2008). All the emission sources in both the port and airport are treated as an area source (Gibson et al., 2013). However, there is a limitation in AERMOD in treating sources like landing and take-off (LTO) of aircraft as area sources. The emissions during LTO are assumed to be uniformly disturbed over a defined area with a specific vertical spread (plume rise). In AERMOD, area sources do not allow the treatment of plume rise, which restricts the accurate prediction of buoyant plume rise of jet exhaust. In contrast, ADMS-airport incorporates the treatment of plume rise, accounting for both buoyancy of jet exhaust as well as momentum (Popoola et al., 2018). Despite its limitation, the studies conducted at airports using AERMOD have shown good agreement between modelled and measured values, with mean biases of 5 to 13% for NO_x, -0.29 to 6.6% for CO, -2.2 to -10.4% for $PM_{2.5}$ and -2.31 to -11% for PM_{10}. Similarly, the airport also showed good agreement with measured concentrations, with mean bias ranging from 5.0% to 20.4% for NO_2, −4.0% to 16.1% for $PM_{2.5}$, and 7.0% to 10.0% for O_3. AERMOD is widely employed in port environments, with the mean bias error of -11 to -65% for $PM_{2.5}$ and 24 to 36% for NO_X. The model tends to exhibit lower errors (-11 to 24%) in areas closer to the port compared to farther away locations.

Overestimation of NO_X near the port and airport is observed from both AERMOD and ADMS-airport, respectively. This may partly be due to the absence of chemical reactions that act as sinks for NO_x in the model. Similarly, underestimation of PM is observed, which could be attributed to the underrepresentation of the secondary organic aerosol component of PM, likely caused by larger uncertainties in the PM emission inventory. Both models showed improved results with the incorporation of the photochemistry module. Additionally, the inclusion of background concentrations improved the models' results.

Overestimation of NO_x near ports and airports has been observed in both AERMOD and ADMS-Airport applications. This may partly be due to simplified treatment of NO_x chemistry, especially reactions that remove NO_x from the atmosphere. Key sink pathways include the daytime reaction of NO_2 with hydroxyl radicals to form nitric acid, and nighttime conversion through nitrate radical and N_2O_5 chemistry, which can also form nitric acid or particulate nitrate. These processes typically occur over timescales of several hours to about one day, depending on sunlight, oxidant levels, aerosol loading and humidity (Brown & Stutz, 2012). Similarly, underestimation of PM may be attributed to the underrepresentation of secondary aerosol formation and uncertainty in PM emission inventories. Both models showed improved performance when photochemistry modules and background concentrations were included.

Similar to the aircraft plume, ship manoeuvring in ports also requires plume rise treatment. However, manoeuvring emissions are relatively lower compared to hoteling emissions, which are the most significant among port emissions. As a result, manoeuvring emissions have a lesser impact on model results and are often ignored during modelling.

2.5.2 Eulerian & Lagrangian models

Eulerian dispersion modelling uses the conservation of mass equation for a pollutant species. A stationary grid of reference is presumed, with the dispersion phenomena being computed as a concentration field with respect to the reference grid (Collett & Oduyemi, 1997). The physical conditions in the reference grid are treated as turbulent, so that dependant variables are composed of an average component (Collett & Oduyemi, 1997). Lagrangian dispersion modelling differs from Eulerian counterparts in that pollutant dispersal is stimulated relative to the shifting reference grid (Collett & Oduyemi, 1997; Rissman et al., 2013). Eulerian & Lagrangian dispersion models identified in the airport and port studies include TAPM, CALPUFF, SPRAY, CAMx, WRF-Chem and CMAQ (D. Chen et al., 2017; Gaeta et al., 2016; Merico et al., 2017; Ramacher, Tang, et al., 2020). These models are typically employed to understand the process of air pollution including the transport of pollutants, deposition, and chemical transformation in regional scale (Gao & Zhou, 2024; Ramacher, Tang, et al., 2020).

CAMx, CMAQ, TAMP and WRF-Chem are three-dimensional Eulerian multi-scale models that support grid nesting and multi-model integration, capable of simulating complex physical and chemical interactions among atmospheric pollutants (Gao & Zhou, 2024). TAPM uses grid based Eulerian dispersion model for regional scale with Lagrangian particle mode to near source concentrations (Ramacher et al., 2019; Ramacher, Matthias, et al., 2020). While these models are primarily used on the regional scale, they can be configured with smaller grid sizes (e.g. 250 x 250 m) in nested domains to focus on specific areas such as ports or airports. However, these finer-resolution nested domains require boundary conditions provided by a larger, coarser surrounding domain. Sub-grid analysis is performed to understand the airport or port contribution to the pollutant concentrations (Ramacher, Tang, et al., 2020). For instance, Ramacher et al. (2019) used TAPM to evaluate the effects of local shipping emissions on AQ with five nested grid system ranging from 10 x 10km to 250 x 250m. The chemical boundary conditions were taken from the CMAQ simulation on a 4km x 4km grid. The study reveals a significant impact of local shipping on local NO2 concentrations, contributing between 5-15% of the annual grid mean NO2 concentration.

CALPUFF is a Lagrangian puff multi-species model capable of simulating the effect of varying meteorological conditions and longer-range transport effects on pollutants transformation and removal (Kontos et al., 2017; Poplawski et al., 2011). SPRAY is a Lagrangian particle mode model used in studying dispersion and deposition of the non-reactive emissions of passive pollutants (Gaeta et al., 2016; Pecorari et al., 2016). Unlike other regional-scale models, CALPUFF and SPRAY can calculate dispersion from multiple sources at the microscale with a smaller grid size (100 m × 100 m) without requiring a coarser surrounding domain. Murena et al. (2018) assessed the impact of cruise ship emissions on Naples, Italy, using CALPUFF with a 36 × 24 cell grid and 200m cell spacing to model cruise ship emissions. The results show a low correlation between the fixed monitoring stations' data and the simulations, suggesting that the contribution of NO2 emissions from cruise ships is insignificant within the Naples metropolitan area. Pecorari et al. (2016) analysed the effects of meteorology on aircraft exhaust dispersion at Marco Polo Airport in Venice employing SPRAY with a 500 × 500 cell grid and 100 m spacing to assess spatial dispersion of HC and CO during LTO cycles. The study highlighted aircraft taxiing as the most impactful during operations, with specific meteorological events (low wind speed and inversion conditions) causing increased HC and CO concentrations in the airport's neighbourhoods.

Among Eulerian models, TAPM and WRF-Chem tend to predict values close to measured concentrations in the port vicinity (TAPM, mean bias: 5 to 15% for NO2; WRF-Chem, mean bias: -14.3 to -19.8% for PM2.5). However, Eulerian models tend to produce larger mean bias errors near airport vicinities (TAPM, -7 to -54% for $PM_{2.5}$ & -11 to 38% for NO_2; CMAQ, 53.9 to 82% for $PM_{2.5}$). The large mean bias near airports can be attributed to several factors, including the lack of precise information on aircraft timing and trajectories, the limited number of emitters used to represent the various modes of aircraft emissions, uncertainties in aircraft engine emission estimates, and inadequate representation of plume characteristics. Few studies have incorporated advanced plume treatments in their models, where emissions from aircraft sources are represented as Gaussian puffs (Bo et al., 2019; Yang et al., 2018). A chemical mechanism "CMAQ-APT" is used to simulate reactions both within the puff as well as within the grid cells. However, the inclusion of this process did not result in a significant improvement in the model's performance. The insufficient availability of measurements of aircraft plumes makes it challenging to accurately represent real-world plume characteristics in the model.

The WRF-CMAQ model showed reasonable agreement with measured ozone concentrations in the airport environment, attributed to the substantial emissions of VOCs and NO_x (Song et al., 2015). The mean bias error between modelled and observed values ranged from -3.1% to 20.6%.

Like Eulerian models, Lagrangian models also exhibit larger mean bias error for airport vicinities than port vicinities. SPRAY is commonly used to understand the impact of airport emissions on local NOx, HC, and CO concentrations in the airport's neighbourhood. The mean bias error for pollutants modelled using SPRAY in the airport vicinity ranges from -36% to 89.6% for NO_x, -35% to -60.8% for CO, and -44% to -134% for HC. The model performance slightly improved when considering the background concentration. The SPRAY model is effective in modelling the physical dispersion of pollutants but does not account for chemical transformations. However, the atmosphere around airports is a complex environment with hundreds of chemical reactions occurring, which could contribute to significant errors in the model results. While CALPUFF is commonly employed in port to understand the effect on local SO2, NO2, and PM2.5. The mean bias error for pollutants modelled using CALPUFF in the port vicinity ranges from 30% to 52% for SO_2, 36% to 69% for $PM_{2.5}$, and - 44% to 14% for NO_2. Inadequate representation of plume characteristics in the CALPUFF model, particularly building downwash of the plume, is attributed to underestimation of concentration (Poplawski et al., 2011). Also, representing ship emissions as a line source is not the best choice; the ship emissions should be represented as a series of volume sources.

Gaussian model results showed reasonable agreement with ground-level measurements compared to Eulerian and Lagrangian regional models in airport vicinities. However, Eulerian models (TAPM and WRF-Chem) performed slightly better in predicting pollutants near ports. Additionally, the Eulerian model exhibited better agreement with ground-level measurements at locations farther from the port than the Gaussian model, as it is better equipped to handle complex terrain and non-steady-state meteorological conditions typical of coastal areas.

For regulatory compliance purposes, when the focus is on near-field impacts, Gaussian models are more suitable due to their computational efficiency. ADMS-airport is more appropriate for the airport environment than the AERMOD. However, if the study involves modelling secondary pollutants or long-range transport, Eulerian models are more appropriate, as they can account for chemical reactions, temporal variability, and complex dispersion dynamics.

2.5.3 Computational fluid dynamics models

CFD tools have become integral to environmental modelling and health protection. These models, grounded in the Navier-Stokes equations, are frequently employed across complex geometries at local scales. CFD stands as the preeminent methodology in the analysis and design of building safety and indoor environments. Additionally, CFD models have been effectively used to simulate pollutant dispersion in both enclosed and semi-enclosed spaces like tunnels. Despite its benefits, CFD usage is often characterised by substantial computational demands.

Yoo et al. (2022) employed CFD to simulate the dispersion of NO_2 originating from a semi-enclosed bus station terminal to optimise the placement of an air purification system. The study involved the construction of a comprehensive model that incorporated the geometric features of the bus station and used the RNG k-epsilon turbulence model to predict the distribution of pollutants. The findings revealed that the effectiveness of the air cleaning system depends on the specific combination of inlet and outlet locations. Similarly, Yu et al. (2022a) employed CFD to investigate the dispersion of CO from a bus exiting a station. The study focused on assessing the impact of bus stop platform type and passenger location on individual exposure to pollution. Three studies adopted CFD to model PM concentrations and other fluid properties in the underground stations; however, this setting was excluded from the review.

Numerical investigations using CFD have been conducted to explore the impact of train piston effects on air pollutant dispersion in underground stations. Bolourchi et al. (2018) presented results of CFD modelling concerning $PM_{2.5}$ and PM_{10} concentrations in underground stations, revealing that the predicted concentrations were approximately 8% lower than the measured concentrations. The numerical simulations focused on air velocity, pressure, kinetic energy, and PM concentrations as the train entered the station. Durand et al. (2021) developed a CFD model to understand particle-flow interaction during low-speed underground train passage, considering particle sizes ranging from 0.02µm to 3µm. Coherent structures within the train wake, including counter-rotating streamwise vortices and a recirculation region, are identified as key mechanisms contributing to lateral and vertical particle dispersion, respectively. Wind tunnel experiments with reduced-scale models confirm the presence of these structures and provide validation for the numerical simulations. The findings suggest that the side extremities of central platform stations could be favourable locations for air purification devices and the concentration rates between the train and the station wall. In the work conducted by Durand et al. (2021), the focus was on the train wake and the region between the train and the wall. Subsequently, Izadi et al. (2021) expanded earlier piston-effect CFD set-ups by modelling particle dispersion across multiple connected station spaces (platforms, stairwells and ticket halls). The simulation of train-induced turbulence was conducted using Reynolds Averaged Navier-Stokes (RANS) equations. The findings indicate that PM in rail stations can originate from primary non-exhaust sources, including brake wear and wheel–rail contact, while previously deposited rail dust may also be re-suspended and dispersed through station spaces by train-induced airflow and passenger movement.

2.5.4 Other dispersion models

Photochemical models are widely used for simulating the impacts of transport emissions on local air quality, such as Community Multiscale Air Quality (CMAQ) and Comprehensive Air Quality Model (CAMx). Rissman et al. (2013) used a plume-in-grid process in the CMAQ model to investigate the impacts of aircraft emissions on $PM_{2.5}$ concentrations around Hartsfield-Jackson Atlanta International Airport. This approach enabled the simulation of emissions sources at a finer scale plume-in-grid system than the traditional CMAQ grid system. The plume-in grids treat the dynamic and chemical processes of secondary pollutants more realistically than traditional grids. Zhou et al. (2020) used CMAQ to examine the impacts of port-related source emissions on the main port areas of Shanghai. The study found that port-related sources contributed to an average of 37% and 3% of NO_X and $PM_{2.5}$ concentrations respectively, towards all anthropogenic sources in Shanghai's main areas. Yang et al. (2018) used WRF-CMAQ to assess the impact of $PM_{2.5}$ from airport emissions over the domain surrounding the Beijing Capital International Airport. The finding indicated that the impact was maximum within a radius of 1km around the airport. Chen et al. (2017) developed a high-resolution ship emission inventory for Qingdao Port, China, and employed CAMx to model the impacts on the surrounding ambient air quality. The results reveal that the contribution of ship emissions to annual ambient $PM_{2.5}$ concentration ranges between 1.5-13% and could be over 20% near the densely populated urban areas located close to the docks during certain periods of the year. A similar approach was adopted by Merico et al. (2017), where the CAMx model and measurements were used to assess shipping impact on air quality along the coast of four Adriatic-Ionian port cities. Results revealed that the shipping impact on gaseous pollutants is greater than that of PM.

2.6 Traffic emission models

In the domain of transport-related pollution assessment, traffic emission models serve as a valuable tool for calculating emissions associated with specific modes of transportation. Multiple models have been developed to evaluate pollutant emissions across different TH. Direct measurements are often required to determine emission factors within these traffic emission models accurately (Lyu et al., 2021). Additionally, some traffic emission models integrate with deterministic dispersion models to forecast air quality in the immediate vicinity of TH. Table 2.3 provides the traffic emissions models reported in the literature to characterise air quality at THs.

2.6.1 Motor vehicle emission simulator

The motor vehicle emission simulator (MOVES) is a comprehensive tool for modelling motor vehicle emissions at various scales, enabling the assessment of emissions for individual vehicle activities (US EPA, 2016c). MOVES was employed to evaluate the emission inventory of a bus terminus in Xi'an, China. The data inputs for this assessment included bus characteristics, operating cycles, and meteorological data (Qiu et al., 2016). Notably, the study demonstrated that substituting diesel fuel with natural gas significantly reduced $PM_{2.5}$ and CO emissions from buses. Park (2022) used MOVES to assess the emission reduction impact of a port consignment truck replacement programme on local air pollution in the ports of New York and New Jersey. The assessment considered both driving and idling emissions of trucks. The results indicated that replacing trucks with higher emission standards led to a substantial decrease in NO_X and $PM_{2.5}$ emissions from trucks. Moreover, the study evaluated the impact of the consignment truck replacement programme on local air quality using C-PORT.

2.6.2 Comprehensive modal emission models

Yu et al. (2022) applied a Comprehensive Modal Emission Model (CMEM) to precisely estimate the emissions of CO_2, CO, NO_X, and HC from buses in Nanjing, China. The model achieved an error margin of less than 10% when comparing these estimates with Portable Emissions Measurement System (PEMS) records. To simulate bus emissions in different bus station types (kerbside and bus bay), the study combined CMEM with multimodal traffic simulation software VISSIM and further used a CFD model to simulate the dispersion of pollutants at the bus stations. The findings of the study indicate that bus station types do not exert a substantial influence on bus emissions.

2.6.3 Fuel use and emission rate models

Rastogi & Frey (2021) developed a Fuel Use and Emission Rates (FUERs) model to segment and accurately quantify diesel locomotives' emission intensity and map railway emission hotspots. This model incorporates data sources, including 1Hz emission data obtained from PEMS, engine activity data, and locomotive activity records. The study also used Classification and Regression Trees (CART) to identify the variables influencing the hotspot areas for different pollutants. The results show that the intensity of emissions from hotspots for NO_X and PM is much higher than from non-hotspots. Most of the hotspot areas were within 1.25 miles of the stations.

Table 2.3 Traffic emissions models reported in the literature to characterise air quality at THs

Model	Hub Type	Study Area	Inference	References
Motor vehicle emission simulator (2)	Bus terminal and Port	Semi-enclosed space, Vicinity of the hub	Assessment of individual vehicle emissions; Need for accurate emission factors and engine states	**(G. Y. Park, 2022; Qiu et al., 2016)**
Comprehensive modal emission (1)	Bus terminal	Semi-enclosed space	High precision Second-by-second data; High computational volume; Requires PEMS validation	(Yu et al., 2022b)
Fuel use and emission rates (1)	Railway Station	Vicinity of the hub	Accurate emission estimates; Spatial variability; High-cost emission measurements Based on PEMS	(Rastogi & Frey, 2021)

2.7 Discussion

Air quality modelling methodologies in the existing the TH literature vary according to (i) whether emissions and exposure occur primarily outdoors or in enclosed/semi-enclosed spaces, (ii) whether the modelling aim is near-field concentration mapping, source contribution/apportionment, or operational attribution, and (iii) the availability of activity and emissions data at an appropriate temporal and spatial resolution. Across hubs, these choices create a trade-off between physical realism (dispersion/CFD), statistical attribution and forecasting (regression, GAM, machine learning), and operational emissions representation (traffic emissions models). For airports and ports—where the dominant emission sources are generally outside terminal buildings—deterministic dispersion models are most commonly applied to quantify the spatial footprint of transport emissions and to evaluate policy scenarios. Gaussian-based systems such as AERMOD and ADMS are widely used at local scales where detailed meteorology and emissions inventories can be specified, while Eulerian chemical transport models (e.g., CMAQ) are applied when regional chemistry and secondary formation are central to the research question. These applications show that, when emissions inputs are credible and meteorology is well constrained, deterministic models can reproduce observed patterns with useful skill, enabling comparisons between scenarios and supporting impact assessment. However, the relevant literature for airports and harbours focuses mainly on outdoor concentration fields. There are no studies modelling air quality inside terminal or port buildings.

Underground metro systems represent enclosed environments where non-exhaust particulate sources (e.g., braking, wheel–rail contact, resuspension) can dominate platform exposures, and where meteorology influences concentrations mainly through ventilation and exchange with outdoor air. Consequently, statistical models are frequently used to attribute concentration variability to train frequency, station characteristics, ventilation, passenger flow, maintenance plans and urban background, while CFD is used to resolve the within-station dispersion and the influence of geometry on pollutant transport. The underground literature illustrates two important points that generalise to other semi-enclosed transport environments. First, background concentrations remain a key determinant of observed levels, setting the baseline upon which station processes add an increment. Second, operational metrics that more directly represent emission-generating behaviours (e.g., braking intensity, ventilation mode, dwell characteristics) tend to be more informative than coarse proxies such as simple train counts. These findings matter for rail stations because many mainline platforms share partial enclosure and constrained ventilation, but with the additional complexity of diesel exhaust sources that vary strongly with engine state.

Bus terminals and stops provide a further intermediate case, where exposure occurs in semi-enclosed microenvironments, and the dominant sources are road vehicles with highly transient operating modes. Here, studies commonly combine a traffic emissions model (e.g., MOVES, CMEM) with a high-resolution dispersion component (often CFD) to capture the effects of acceleration, queuing and idling on local concentration gradients. This paired approach highlights a recurring limitation across TH studies: dispersion skill is strongly constrained by the realism of emissions inputs. When second-by-second emissions or engine-state information is available, near-field predictions become more credible; when emissions are specified using generic factors or averaged inventories, uncertainty increases substantially—particularly for particulate matter, where contributions can include primary exhaust, non-exhaust, and secondary components.

Railway stations are comparatively under-represented in the modelling literature, despite their potential for high exposure due to the proximity of passengers and staff to active diesel engines and the presence of semi-enclosed platform structures. The limited rail-station studies identified in this review tend to fall into two groups. The first focuses on operational attribution using statistical or machine-learning models with variables such as train frequency and rolling stock type. These studies indicate that background pollution often remains the dominant predictor of station concentrations, while train activity explains additional variability; they also show that model explanatory power can be limited when key behaviours (notably idling and engine state) are not measured and must be represented using coarse proxies (e.g., train counts). The second group focuses on emissions characterisation, such as fuel-use and emission-rate modelling approaches that use high-frequency operational and measurement data to map emission intensity and identify hotspots around rail infrastructure, but these approaches are typically data-intensive and not readily transferable without comparable measurement and activity datasets. Across all TH types, several cross-cutting themes emerge that inform the design of robust station-scale studies. First, credible attribution of a hub contribution requires careful control for background concentrations and meteorological/ventilation conditions, because these factors can explain a large fraction of observed variance and can confound operational signals if not modelled explicitly. Second, the temporal resolution and behavioural specificity of activity data strongly determine how well any model class can isolate transport contributions: emissions and dispersion processes respond to engine state, idling duration, and operating mode, which are often not captured by schedule-level or count-based proxies. Third, the choice between statistical and deterministic approaches is not purely methodological; it reflects the question being asked. Statistical models are well suited to identifying influential variables and quantifying associations in measured datasets, but they do not by themselves provide a spatially resolved emissions field or a mechanistic basis for scenario testing. Deterministic dispersion models can map spatial patterns and compare mitigation scenarios, but only when the emissions specification is sufficiently detailed and consistent with real operations. This motivates integrated workflows where statistical attribution and operational analysis inform the emissions representation that is then used in dispersion modelling.

2.8 Research gaps from the literature

1. Modelling air quality at airports and ports frequently uses verified transportation emission models or emission inventories to evaluate the pollution produced by traffic tools comprehensively. Conversely, in railway stations, there is an evident deficiency in such emission inventories or tools. This could be attributed to the lack of accessible data regarding train services and fleet specifics and the lack of direct emission measurements from diesel railway locomotives.

2. Railway stations/hubs lack an all-encompassing model, integrating both a traffic emissions model and a deterministic dispersion model, essential for determining the role of railway systems in local air pollution. This comprehensive model should also account for road traffic

emissions in the vicinity, enabling an exploration of the various transport modes' contribution to the surrounding air pollution.

3. Modelling air quality in railway stations is a challenge, as emission sources are more intricate compared to the well-defined sources in underground stations, particularly for particulate matter. Emissions of such matter in railway stations originate from diesel vehicle exhaust and wheel-rail friction across all vehicles. Current research cannot pinpoint these particulate matter sources at railway stations through chemical particulate analysis or statistical modelling. In addition, the idling time of trains in stations was not included in the model.

4. Despite the success of CFD modelling in underground stations, adopting a similar modelling technique to emulate the spread of particulate matter in railway stations is not straightforward. Railway stations are not wholly enclosing environments and exhibit complex architectural structures, and this adds significantly to the computationally demanding nature of the modelling process. Although a statistical model has investigated the impact of the number and type of diesel vehicles on railway station air quality, they still require additional data, such as locomotive features and operational conditions, for better prediction of pollutant concentrations.

These gaps map directly onto the objectives of this thesis. The lack of station-scale rail emissions inputs motivates development of a diesel train emissions calculation tool based on operational data and emission factors (Objective 4) and supports attribution of in-station concentrations to diesel activity and idling (Objective 3). The lack of integrated rail–road neighbourhood assessment motivates coupling rail emissions with deterministic dispersion modelling, alongside road-traffic inputs, to quantify spatial impacts around stations (Objective 5). The limited representation of PM sources and omission of idling behaviour motivates targeted analysis of operational drivers, including idling-time assessment and its influence on exposure (Objectives 3 and 4). Finally, the absence of robust mitigation evaluation frameworks motivates scenario-based assessment of technological and policy measures (Objective 6), grounded in the literature synthesis presented in this chapter (Objective 1) and supported by the monitoring datasets compiled for the Dublin case studies (Objective 2).

3. Air quality monitoring and statistical modelling in Dublin rail stations

3.1 Introduction

This chapter investigates air quality in two major Dublin railway stations—Heuston and Connolly—with the aim of characterising in-station exposure conditions and identifying the key determinants of pollutant variability. The analysis focuses on three pollutants: $PM_{2.5}$, PM_{10} and NO_2. These pollutants are prioritised because they are widely used indicators in urban air-quality assessment, have well-established relevance for human health, and are directly linked to transport-related emissions and exposure in commuter microenvironments (European Environment Agency, 2023; Health Effects Institute, 2024; World Health Organization, 2024).

$PM_{2.5}$ and PM_{10} are included to capture both fine and coarse particle behaviour in stations. Fine particles are influenced by combustion sources (including diesel exhaust) and can remain suspended for longer periods, while coarse particles are more sensitive to mechanical processes such as resuspension and abrasion and can exhibit strong spatial gradients near sources (Harrison, 2020; World Health Organization, 2021a). In mainline rail stations, PM may reflect a combination of diesel exhaust (where diesel traction or idling occurs), non-exhaust rail processes (e.g., wheel–rail and braking wear) and contributions from adjacent road traffic (Abbasi et al., 2012; Fussell et al., 2022; Loxham & Nieuwenhuijsen, 2019).

NO_2 is included as a marker of combustion-related emissions, particularly diesel exhaust, and is commonly used to represent the gaseous component of transport-related air pollution in urban environments (European Environment Agency, 2023; World Health Organization, 2024). Considering PM fractions alongside NO_2 therefore supports a more complete assessment of station air quality, where both particulate and gaseous pollutants may contribute to exposure.

From a health and regulatory perspective, $PM_{2.5}$ and PM_{10} are also key because guideline and limit values are expressed directly in terms of these metrics. The 2021 WHO Global Air Quality Guidelines recommend annual mean concentrations not exceeding 5 $\mu g/m^3$ for $PM_{2.5}$ and 15 $\mu g/m^3$ for PM_{10}, with corresponding 24-h guideline values of 15 and 45 $\mu g/m^3$, respectively (World Health Organization, 2021). For NO_2, the WHO guidelines recommend an annual mean of 10 $\mu g/m^3$ and a 24-h mean of 25 $\mu g/m^3$ (World Health Organization, 2021). In the European Union, Directive 2008/50/EC sets the legal limit values currently in force, including annual limits of 40 $\mu g/m^3$ for both PM_{10} and NO_2 and 25 $\mu g/m^3$ for $PM_{2.5}$ (European Commission, 2008).

The chapter addresses two connected questions. First, how do $PM_{2.5}$, PM_{10} and NO_2 vary over time within each station, and how do levels differ between stations and monitoring locations? Second, to what extent can observed variability be explained by train activity, operational conditions and external influences such as meteorology and background urban concentrations? The analysis is designed to support later chapters by identifying which operational behaviours are most plausibly linked to in-station concentrations and by clarifying the extent to which in-station air quality reflects local station processes versus regional background.

To answer these questions, the chapter combines descriptive analysis with statistical modelling. Descriptive results summarise concentration distributions and temporal patterns at each station. Inferential screening is then used to test whether pollutant levels differ systematically across operational categories and conditions. Finally, multivariate non-linear modelling (Random Forest regression) is applied to quantify the relative explanatory importance of train-activity indicators, background concentrations and meteorological variables, and to assess predictive skill in each station context. This staged approach is adopted because station air quality is influenced by multiple correlated drivers and may exhibit non-linear responses and interactions that are not well represented by simple linear models.

The remainder of the chapter is organised as follows. Section 3.2 provides a case-study context for Irish Rail and summarises the operational and physical characteristics of Heuston and Connolly, and describes the monitoring design, instrumentation, datasets and pre-processing steps, including background and meteorological inputs and the construction of train-activity variables. Section 3.3 presents results from descriptive analysis and statistical modelling. Sections 3.4 and 3.5 synthesise the main findings and set out the implications for the operational analysis and emissions–dispersion modelling undertaken in subsequent chapters.

3.2 Methods

3.2.1 Case-study context: Irish Rail, Heuston and Connolly

Irish Rail (Iarnród Éireann) is the national rail operator and infrastructure manager for Ireland, providing intercity, commuter and freight rail services. The operational rail network extends to approximately 2,400 km of track and includes 144 stations.

This thesis focuses on Dublin's two principal city-centre rail stations, Heuston Station and Connolly Station (Figure 3.1). Heuston is located to the west of the city core and primarily serves as a terminal for intercity and commuter services on the Cork, Galway, Limerick and Waterford corridors. Connolly is located to the north-east of the city core and functions for intercity and commuter services while also acting as a key interchange for Dublin's electrified suburban rail services (DART).

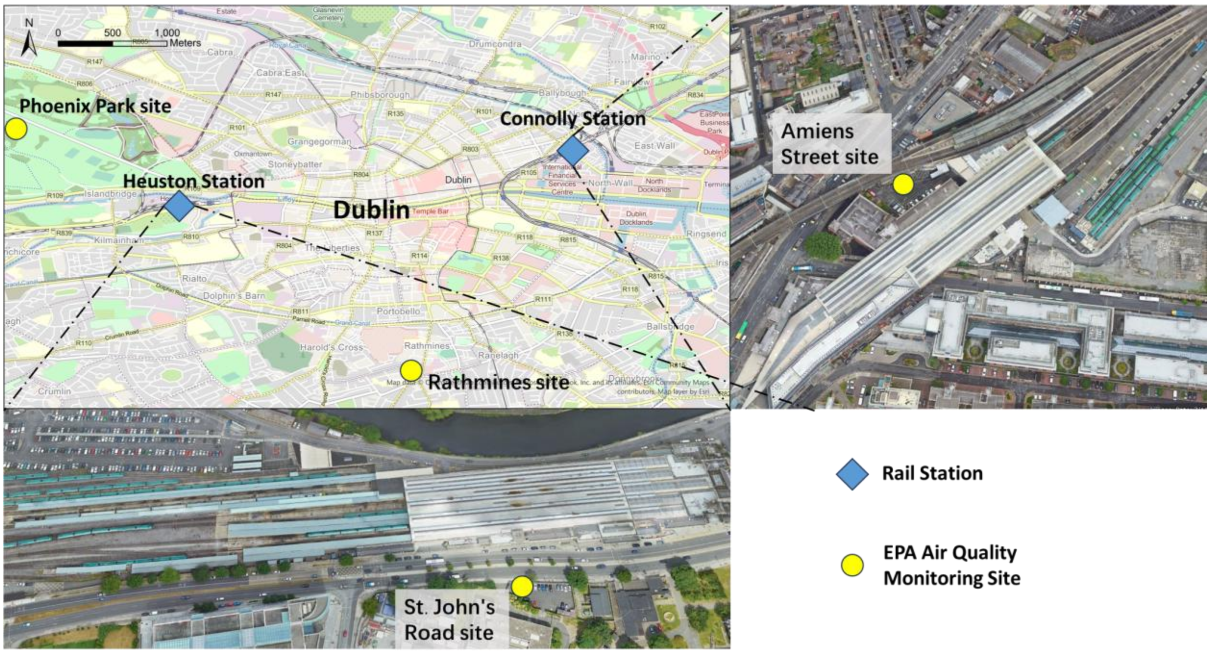


Figure 3.1 Location of the two case-study stations (Heuston and Connolly) and external reference (EPA) monitoring sites in Dublin

Figure 3.1 maps the locations of both stations within the Dublin urban area and shows the Environmental Protection Agency (EPA) monitoring station sites used in this chapter to characterise near-station urban concentrations, including Amiens Street (20 m from Connolly) and St John's Road (37 m from Heuston). In addition, the EPA sites at Phoenix Park (6.3 km from Connolly) and Rathmines (4.2 km from Connolly), which were used as urban background inputs for the dispersion model in Chapter 6, are also indicated.

Heuston Station

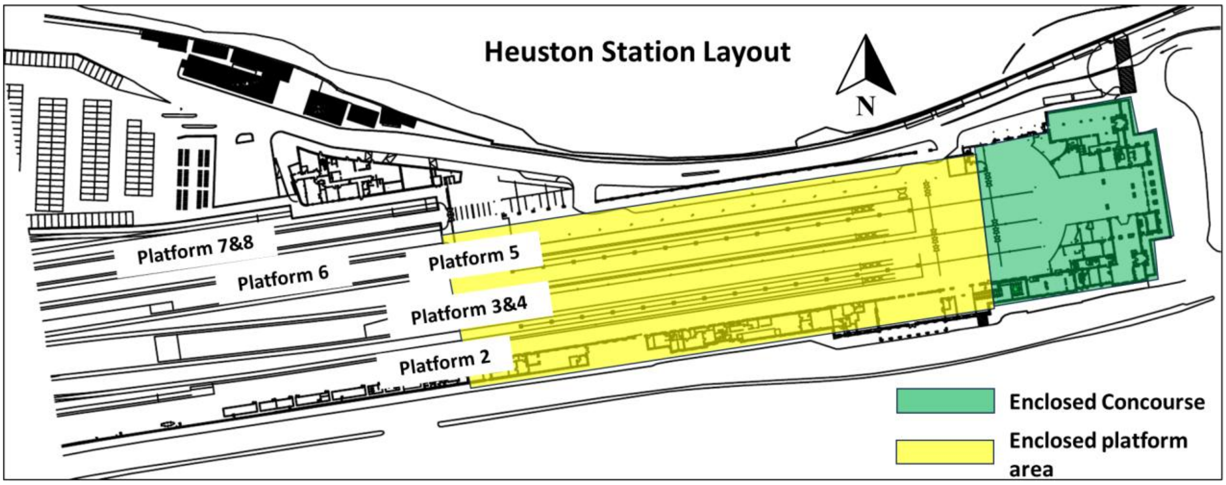


Figure 3.2 Heuston Station layout

Heuston is a terminal station with tracks oriented broadly east–west (Figure 3.2). The station is semi-enclosed: platforms 2 –5 are located under a canopy (the enclosed platform, 7357 m^2), while platforms 6–8 are unenclosed. The principal external opening to the covered platform hall is at the west end, where trains enter and exit. The track gradient at this entrance is approximately 13‰, with trains approaching the station downhill and departing uphill. Additional lateral openings provide pedestrian and vehicle access on the north and south sides. The main concourse (2866 m^2) is situated at the east end of the station and is separated from the platforms by a glass wall. In the covered hall, longitudinal roof vents are the primary outlet points for exhaust air.

All passenger services operating at Heuston are diesel-powered. Over the 2023–2024 study period, the average daily number of planned diesel services at Heuston was 168 on weekdays, 123 on Saturdays, and 68 on Sundays. On weekdays, these services were predominantly operated by Class 22000 diesel multiple unit (DMU) sets (154 services; 91.7%), with the remainder operated by Class 201 locomotive-hauled MkIV sets (14 services; 8.3%). Appendix 1 shows photographs of trains serving at Connolly and Heuston.

DMU is the train that comprises self-propelled carriages rather than a single locomotive hauling unpowered coaches. The Class 22000 DMU fleet can be flexibly marshalled into InterCity Railcar (ICR) sets. In this thesis, the notation of ICR + n is used to indicate the number of carriages in the set. For example, ICR4 denotes an InterCity Railcar set consisting of four Class 22000 DMU carriages.

Passenger throughput at Heuston is high at the national scale. On the 2024 National Rail Census day, Heuston recorded 14,326 boardings and 16,199 alightings (National Transport Authority, 2025).

Connolly Station

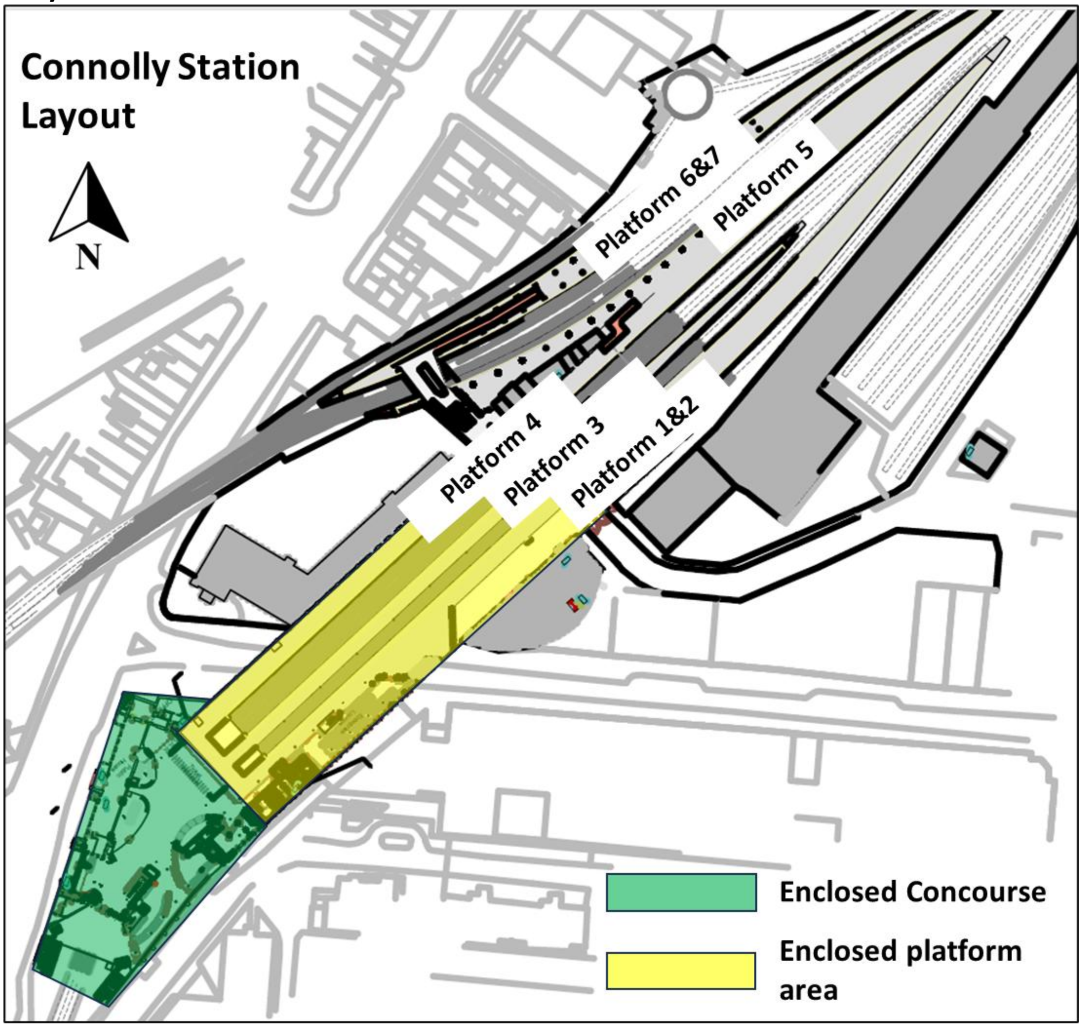


Figure 3.3 Connolly Station plan layout

Connolly has two functionally distinct rail areas (Figure 3.3). Platforms 1–4 serve as terminus platforms for diesel services and are located under a roof (the enclosed platform, 6230 m^2). The principal external opening for this enclosed hall is at the north-east end, where trains enter and exit. The track gradient at this entrance is approximately 8.3‰, with trains approaching the station uphill and departing downhill. Additional openings include pedestrian access points on the eastern side near Platform 2 (connecting towards the station parking area) and internal connections between the terminus platforms and the through-platform area. Connolly's main concourse is located at the southern end of the station (888 m^2) and serves as the primary passenger circulation space linking the station's main entrances with the platform areas.

Platforms 5–7 operate as open (unenclosed) through platforms and accommodate both electric and diesel services, although they predominantly serve electric suburban DART movements. In particular, Platform 5 is a key platform for southbound DART services, while Platforms 6–7 primarily serve northbound DART services. The DART is Dublin's electrified suburban rail system running along the coastal corridor from the northside (Howth/Malahide) through the city towards the southside to Greystones, and it forms a major component of commuter rail activity in the station environs. The main station concourse is located to the south-west of the platforms and is separated from the platform areas by a curtain wall.

Over the 2023–2024 study period, Connolly handled approximately 269 diesel services per weekday, with 153 on Saturday and 118 on Sunday. On weekdays, diesel services comprised 105 Class 22000 DMU services (39.0%), 145 Class 29000 DMU services (53.9%), and 18 De Dietrich services hauled by Class 201 traction (6.7%). In addition, the station handled approximately 203 electric services per weekday as a through station via the outdoor platforms (electric share of total weekday services: 43.0%; 203 out of 472 services).

Connolly also experiences high passenger throughput. On the 2024 National Rail Census day, Connolly recorded 22,356 boardings and 20,097 alightings, reflecting its combined role as a terminus and an interchange node within the Dublin rail network (National Transport Authority, 2025).

In summary, both stations combine substantial passenger throughput with partially enclosed platform environments, but they differ in operational configuration and traction mix. Heuston is a totally diesel terminal with a semi-enclosed main platform hall and no electrified through-platform interface. Connolly combines an enclosed diesel terminus hall (Platforms 1–4) with an open, through-running platform area (Platforms 5–7) that is strongly influenced by frequent electric DART movements. These differences in traction mix and enclosure are expected to affect ventilation and pollutant accumulation and are considered in the analysis.

3.2.2 Air quality data

3.2.2.1 Monitoring instrumentation

Three measurement approaches were used to characterise pollutant levels inside both stations: (i) compact low-cost multipollutant stations providing high-frequency data at multiple locations; (ii) Palmes-type passive diffusion tubes providing integrated NO_2 concentrations over several weeks;

and (iii) a high-resolution fixed NO_2 monitor based on the Sonitus DM30 Dustsens platform. Photographs of the instruments shown in Figure 3.4.

Figure 3.4 Photographs of monitoring equipment

Low-cost sensors (eLichens eLos)

For continuous monitoring of particulate matter and nitrogen dioxide, eLos outdoor air-quality stations (eLichens, 2023) were deployed at multiple positions within each station. Each eLos unit is an autonomous, solar-powered outdoor station that houses optical and gas sensors in a compact weatherproof enclosure, together with on-board data logging and communication electronics. The eLos main body size is 18cm*15cm*6cm, equipped with a solar panel with a size of 20*12cm. The instrument measures several pollutants relevant for this study – PM_1, $PM_{2.5}$, PM_{10} and NO_2 – alongside temperature, relative humidity and pressure, with data recorded at resolution 1 hour. PM was measured via light scattering and NO_2 via electrochemical sensors.

The PM channels use a light-scattering particle counter, which has been evaluated in manufacturer and shown to agree closely with reference analysers (reported coefficients of determination up to $R^2 \approx 0.98$ for $PM_{2.5}$) (eLichens, 2020).

In this study, the eLos units were installed at approximately 2–3 m above floor level on existing structural elements (e.g. columns or railings), so that the inlets sampled air within the breathing zone but remained clear of passenger movement. In this study, instrument performance was assessed and, where necessary, corrected by co-location with EPA monitors, as described in Section 3.2.2.2.

Passive NO_2 diffusion tubes (Gradko Palmes tubes)

To obtain longer-term integrated NO_2 concentrations and to complement the high-frequency sensor data, Palmes-type diffusion tubes supplied and analysed by Gradko International were installed at selected locations (Gradko International Ltd, 2024). These passive samplers consist of a small clear plastic tube, approximately 7 cm in length, open at one end and closed at the other, with stainless-steel grids at the closed end coated in triethanolamine solution to absorb NO_2 from the surrounding air. The tubes were deployed in protective shelters at selected locations inside each station and exposed for four weeks, following guidance for Palmes tubes.

Palmes-type NO_2 tubes are widely used as an indicative method for ambient monitoring in the UK and Europe (Butterfield et al., 2021; Department for Environment, 2024; Targa et al., 2008). When prepared, exposed and analysed in line with Defra guidance, typical expanded uncertainty for annual mean NO_2 is of order ±25% relative to chemiluminescent reference analysers, meeting the data-quality objective for indicative measurements under European legislation (Butterfield et al., 2021; Heal et al., 2018; Targa et al., 2008). In this study, Gradko's UKAS-accredited laboratory undertook all chemical analysis.

High-resolution NO_2 monitor (Sonitus DM30-based)

To provide a higher-resolution reference for NO_2 within the station environment, continuous fixed NO_2 monitoring was undertaken using a customised station based on the Sonitus DM30 Dustsens platform (Sonitus Systems Ltd, 2022). The custom station employed the same processing hardware as the DM30 but was equipped with an electrochemical NO_2 sensor rather than the standard particulate configuration, enabling high-resolution (60 seconds) NO_2 measurements within Connolly Station. The instrument is housed in a rugged weatherproof enclosure with approximate external dimensions of 430 × 410 × 185 mm and a 185 mm inlet, with a mass of around 12 kg (Sonitus Systems Ltd, 2022). The DM30 Dustsens is MCERTS-certified as an indicative ambient monitor.

Although electrochemical NO_2 sensors are useful for high-temporal-resolution monitoring, they are not equivalent to reference-grade analysers. Their measurements can be affected by temperature, relative humidity, sensor drift, and cross-sensitivity to other oxidising gases, especially ozone. Previous field evaluations have shown that site-specific calibration and careful interpretation are needed when using low-cost electrochemical NO_2 sensors for ambient air-quality studies (Mead et al., 2013; Papaconstantinou et al., 2023).

The customised NO_2 monitor was the inlet positioned to sample the mixed air in the station rather than being directly in a plume from a specific train. Data were logged at 1 min and aggregated to hourly means for comparison with both the eLos network and external reference stations. Before installation at the station, the instrument was co-located with the Pearse Street EPA air-quality station for 10 days to derive an adjustment factor to the reference scale.

3.2.2.2 Monitoring campaigns

Air-quality monitoring was conducted in three phases (Monitoring Activities 1–3), summarised in Table 3.1.

Table 3.1 Description of air quality monitoring activities in train stations

<table>
<tr><th>Monitoring Activity</th><th>Devices type</th><th>instrument model</th><th>Pollutants</th><th>Time resolution</th><th>Location</th><th>Data from</th><th>Data to</th></tr>
<tr><td rowspan="3">1</td><td rowspan="3">Low-cost monitoring sensor</td><td rowspan="3">eLichens eLos station</td><td rowspan="3">$PM_{2.5}$, PM_{10}, NO_2</td><td rowspan="3">1 Hour</td><td>Heuston</td><td>26-May-23</td><td>19-Jun-23</td></tr>
<tr><td rowspan="2">Connolly</td><td>8-Aug-23</td><td>17-Aug-23</td></tr>
<tr><td>28-Sep-23</td><td>7-Oct-23</td></tr>
</table>

2	Palmes-type diffusion tube	Gradko Nitrogen Dioxide Diffusion Tube	NO_2	4 Weeks	Heuston	25-May-23	20-Jun-23
				4 Weeks	Connolly	8-Aug-23	7-Sep-23
3	High-resolution monitoring station	Sonitus DM30 Dustsen	NO_2	1 Min	Connolly	1-Mar-24	31-May-24

Monitoring Campaign 1 – Low-cost sensor deployment and calibration

In the first stage, low-cost monitoring (LCM) devices (eLos) were deployed in Heuston and Connolly to measure $PM_{2.5}$, PM_{10} and NO_2. The devices simultaneously recorded temperature, humidity, pressure. An anemometer at each train station provided local meteorological data. All parameters were measured at hourly for evaluation and calibration.

Prior to in-station deployment, five LCMs (IDs 12, 35, 36, 43 and 54) were co-located at the Ringsend EPA air-quality monitoring station (reference site 1) for two weeks (4–17 May 2023) to enable calibration and performance assessment (Figure 3.5). The co-location provided paired hourly time series of pollutants and facilitated:

- Screening of units based on consistency and agreement with the reference analyser.
- Estimation of calibration relationships between LCM outputs and reference concentrations.

Figure 3.5 Low-cost monitors co-located with the Ringsend EPA station (reference site 1)

Following co-location, a subset of validated LCMs was deployed inside the stations. At Heuston, validated units were installed on the main platform, concourse and roof vent area for 25 days, at the locations shown in Figure 3.6. At Connolly, validated units were installed on Platforms 2, 3, 4, Platform 6/7 and the concourse for 20 days (Figure 3.7). These deployments produced spatially distributed time series of $PM_{2.5}$ and PM_{10} within each station, suitable for comparing enclosed and open areas and for relating pollutant levels to train operations. During Heuston monitoring, an LCM was installed at the St John’s Road (reference site 2) EPA monitoring site near Heuston (Location shown at Figure 3.1) to test the performance of the calibration algorithm. The same calibration algorithm was applied to other LCMs.

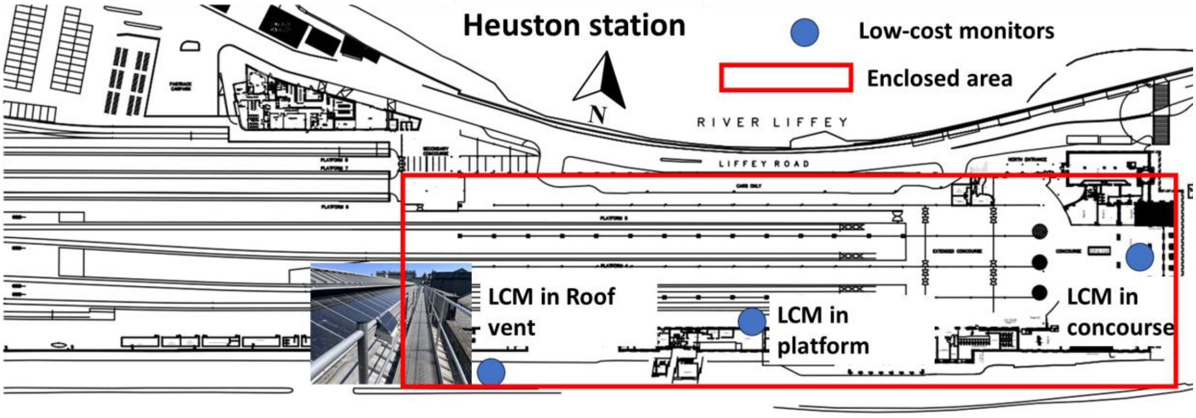


Figure 3.6 Locations of low-cost monitors at Heuston Station

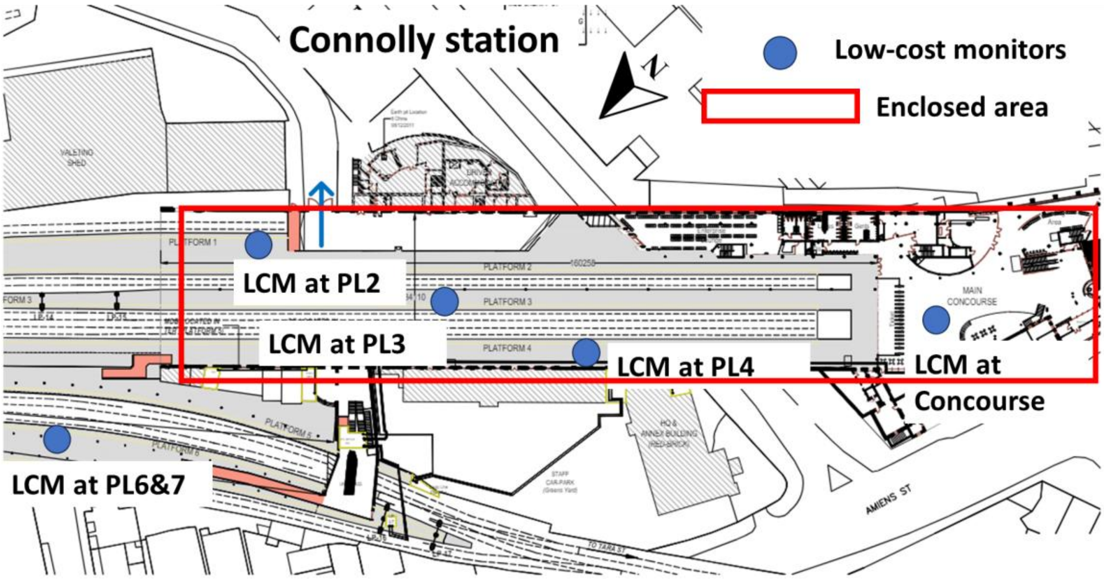


Figure 3.7 Locations of low-cost monitors at Connolly Station

Calibration of low-cost particulate sensors

Calibration of the low-cost particulate sensors was carried out using co-location data from the Ringsend EPA air-quality station (reference site 1) and validated using the St John's Road EPA station (reference site 2). Hourly PM measurements from each low-cost monitor (LCM 12, 36, 43 and 54) were paired with the corresponding hourly reference analyser values at Ringsend. On this training dataset, several calibration algorithms were fitted to map raw low-cost measurements to reference-equivalent concentrations. The candidate models comprised ordinary least-squares linear regression, ridge regression, least absolute shrinkage and selection operator (LASSO) regression, decision tree regression and Random Forest regression. Each model produced a calibration function for each sensor. These functions were then applied to LCM 36 during the Heuston monitoring phase, when it was co-located with the St John's Road EPA station, to evaluate out-of-sample performance. Performance was quantified using the Pearson correlation coefficient and the root mean squared error (RMSE) between the calibrated sensor output and the reference analyser. The comparative results of these calibration methods and the final sensor-specific calibration equations adopted for $PM_{2.5}$ and PM_{10} are reported in Section 3.3.1.

Monitoring campaign 2 – Diffusion tubes

In the second stage, Palmes-type diffusion tubes were used to estimate daily average NO_2 concentrations within stations. Tubes were deployed at multiple platform and concourse positions and exposed for a 5-week period. Deployment locations are shown in Figure 3.12 (Heuston) and Figure 3.13 (Connolly). Two NO_2 diffusion tubes were co-located with the EPA reference station at St John's Road (Figure 3.1). The diffusion tube results were compared with measurements from the EPA NO_2 analyser over the same exposure period. A Local Bias Adjustment Factor was then calculated using the Defra diffusion tube bias adjustment approach (Defra, 2011). All diffusion tube concentrations were multiplied by this local adjustment factor (0.89) to obtain the final adjusted NO_2 concentrations.

In this chapter, the diffusion tubes provide spatial information on NO_2 levels across the platform and concourse locations and help identify high-concentration areas for Monitoring campaign 3.

Monitoring campaign 3 – Continuous NO_2 and meteorology at Connolly

In the third stage, continuous NO_2 monitoring was carried out at the centre of Platform 4 in Connolly for 90 days (1 March–31 May 2024) using the customised Sonitus-based station described in Section 3.2.2.1. This location was selected on the basis of diffusion-tube results, which indicated high NO_2 levels in this area. NO_2 was recorded at 1-minute resolution and aggregated to hourly means for analysis.

A weather station (HOBO U30, Onset Computer Corporation, MA, USA) was installed initially at the open end of Platform 2, where it recorded outdoor meteorological parameters for one month (1–31 March 2024). The station was then relocated to Platform 4 and operated alongside the continuous NO_2 monitor from 1 April to 31 May 2024. Meteorological data (air temperature, relative humidity, pressure, wind speed and wind direction) were logged at 5-minute resolution and aggregated to hourly means to match the NO_2 data.

3.2.3 Train activity data

3.2.3.1 Train frequency and type

Irish Rail's published passenger timetables provide scheduled frequencies and timings of services at Heuston and Connolly (Irish Rail, 2023). In addition, a public Application Programming Interface (API) supplies real-time information on train running, including delays and service status. A Python-based application was developed to harvest data from this API and store them in a database by station and date.

The system has run continuously since April 2023 on a high-capacity computer, creating a record of actual train arrivals, departures and pass-through movements, together with unit-level information on fleet type and consist length. These data were used to compute hourly counts of arrivals, departures and passes, disaggregated by train class and set length (e.g. 3-, 4-, 5- and 6-car DMUs, MkIV locomotive-hauled sets). Figure 3.8 summarises the workflow for obtaining real train frequencies and types.

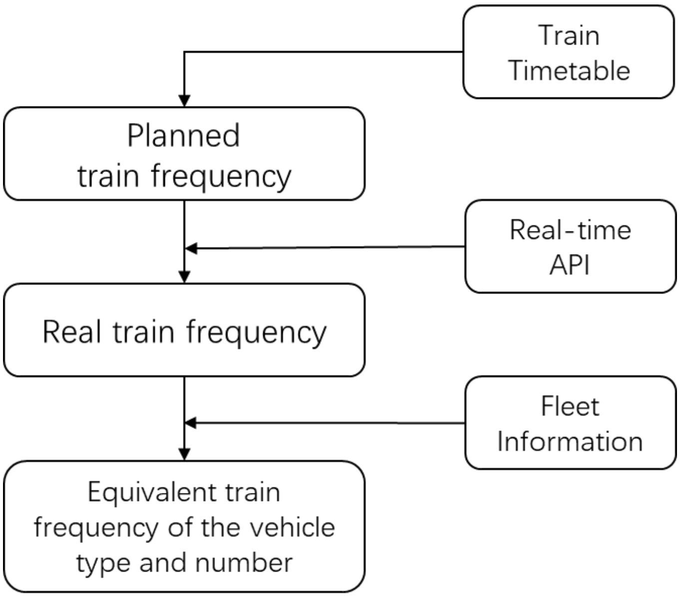


Figure 3.8 Workflow for obtaining real train frequencies and types

3.2.3.2 On-train monitoring recorder (OTMR) data and idling

OTMR data were obtained from two on-board systems used by Irish Rail: the TELOC event recorder and the Train Management and Diagnostic System (TMDS). Both systems are standard equipment fitted to the fleet and together provide a complete OTMR record for each train, including time, speed, location and engine-related diagnostics. OTMR data are collected routinely by Irish Rail for safety, performance monitoring and incident investigation.

For this study, Irish Rail exported archived OTMR files from TELOC and TMDS for trains serving Heuston and Connolly over the air-quality monitoring periods. These data were used to reconstruct train movements in and around the stations and to derive detailed measures of idling behaviour at the platforms.

For Heuston, OTMR data were available from 25 May to 25 June 2023, covering Monitoring campaign 1. For Connolly, OTMR data were available from 1 March to 31 May 2024, covering Monitoring campaign 3.

3.2.3.3 Definition and extraction of idling

Idling time statistics

Idling time is defined as the period when a train is stationary at a platform (speed = 0), but the engine remains running, so exhaust emissions continue. To extract idling episodes from OTMR data, the processing workflow proceeds as follows:

1. OTMR records are sorted by unit number and timestamp.

2. Time differences between consecutive records are calculated.

3. A new idling period is initiated whenever the time gap between records exceeds a threshold (10 seconds), ensuring that data gaps do not fragment single idling episodes.

4. Within each candidate episode, records with speed = 0 and non-zero engine speed (or equivalent traction status) are retained and used to characterise idling.

For each identified idling period, the start time, end time and total duration were calculated. The most frequently occurring spatial region during that period was then determined from the train's recorded positions (Figure 3.9 and Figure 3.10). All train identifiers appearing within the same idling period were concatenated into a single combined identifier to account for overlaps or consist changes. Finally, idling periods that overlapped for the same unit were checked, and any overlaps were resolved by retaining only the longer of the overlapping periods.

Spatial regions for idling

To link idling to the physical layout of each station, three spatial regions were defined using polygons in a geographic information system (Figure 3.9 and Figure 3.10).

At Heuston, Region 1 is the area of the station platform under the station roof. Region 2 is the area of the platform that is not under the station roof. Region 3 is the extended area of the platform under the roof, which is entirely open.

Figure 3.9 Idling regions at Heuston Station

At Connolly, Region 1 is the enclosed terminal platform area under the roof (Platforms 1–4). Region 2 is an extended open area of the terminal platform. Region 3 is the open through-platform area (Platforms 5–7).

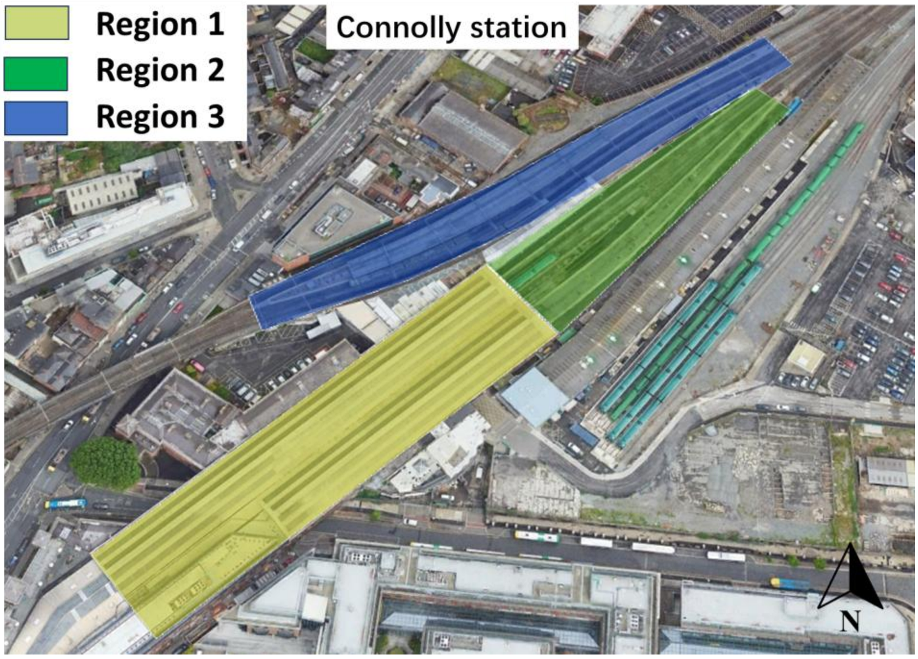


Figure 3.10 Idling regions at Connolly Station

Each OTMR position record is mapped into one of these polygons, and each idling episode is assigned to the region in which it spends most of its duration.

Carriage-adjusted idling

Because emissions during idling scale with the number of powered vehicles in a formation, idling durations were adjusted by the number of powered vehicles. For a DMU, each powered coach contributes to emissions; for locomotive-hauled services, the locomotive is treated as the powered unit.

For each idling episode, the basic idling duration in minutes was first multiplied by the number of powered vehicles in the consist to give a value in carriage-minutes. These carriage-minutes were then summed across all idling episodes occurring within each hour and within each spatial region.

This produces hourly cumulative carriage-adjusted idling times for Regions 1–3 at each station. These hourly series are then merged with the hourly pollutant, background and meteorological data and used in subsequent regression and Random Forest analyses.

3.2.4 Statistical analysis and modelling

The statistical analysis was organised in a sequence of stages so that simple exploratory techniques built towards more complex multivariate models. The first stage used descriptive statistics and visualisations to characterise pollutant levels and temporal patterns at each station. The second stage related these patterns to straightforward indicators of train activity using simple regressions and grouped plots. A third stage applied non-parametric tests and correlation analysis to screen a broader set of candidate predictors and to reduce redundancy among them. The final stage employed Random Forest regression models to quantify how background conditions, meteorology and operational variables jointly influence pollutant levels at the platforms and to rank the relative importance of different drivers (Breiman, 2001; Hastie et al., 2009).

3.2.4.1 Descriptive and exploratory analyses

Descriptive analysis provided the initial characterisation of air quality inside the stations and at the nearby background sites. For each pollutant and station, daily and hourly summary statistics were calculated, including means, medians, selected percentiles and ranges. These summaries were produced separately for each monitoring location (platforms, concourse, roof vents) and for the relevant EPA background sites. Time-series plots and mean diurnal profiles were then constructed by averaging pollutant concentrations over the 24 hours of the day. These plots highlight typical daily patterns, differences between enclosed and open parts of the stations, and periods of systematically higher concentrations. $PM_{2.5}/PM_{10}$ ratios were also calculated to provide a simple indicator of particle size distribution, helping to distinguish fine exhaust-dominated environments from locations with a greater contribution from coarse dust and outdoor particles.

The descriptive stage was complemented by simple exploratory analyses that linked observed pollutant levels to basic operational indicators. Hourly mean pollutant concentrations at the platforms were paired with hourly train activity measures, such as the total number of trains, and the separate numbers of arrivals, departures and pass-through movements. In addition, hourly pollutant data were merged with the carriage-adjusted idling times in each of the defined station regions. For some analyses, 24 hour-of-day mean values were used instead of all hourly observations so that each clock hour is represented by a single mean pollutant level and a single mean activity measure. Simple linear regressions were fitted between pollutant concentrations and each operational metric. These regressions summarise average relationships between station operations and air quality, provide interpretable slopes in units such as micrograms per cubic metre per additional train or per minute of idling, and indicate, through the coefficient of determination, how much of the systematic variation in pollution can be attributed to the operational variable alone.

3.2.4.2 Variable screening and selection

To extend the analysis beyond simple bivariate relationships and to prepare for multivariate modelling, a structured variable screening procedure was applied. The aim was to identify which operational and environmental variables have statistically detectable associations with pollutant levels and to reduce redundancy among predictors before fitting Random Forest models.

Each continuous candidate predictor, including background pollutant concentrations, temperature, pressure, wind speed, relative humidity and carriage-adjusted idling times in each region, was discretised into ordered categories by dividing its distribution into approximately equal-frequency bins, typically deciles. Discrete predictors such as hourly train counts were grouped by their observed integer values or into small ranges where appropriate.

For each pollutant–station combination, the Kruskal–Wallis one-way analysis of variance on ranks was then used to test whether the distribution of hourly pollutant concentrations differed across the categories of each predictor (Kruskal & Wallis, 1952). Under the null hypothesis, pollutant concentrations in all categories arise from the same population; a small p-value indicates that pollutant levels vary systematically across the range of that variable. A significance threshold of $p < 0.01$ was adopted to identify predictors with a clear association with pollutant concentrations. This non-parametric screening is robust to non-normality and outliers and provides a transparent ranking of predictors according to how strongly they stratify the pollutant distributions.

Because many operational and meteorological variables are correlated with one another, an additional correlation-based selection step was applied to avoid redundancy and potential instability in the multivariate models. For all predictors retained after the Kruskal–Wallis screening, pairwise Pearson correlation coefficients were calculated. Where two variables displayed a high degree of linear association, with a correlation greater than 0.7, only one was retained for modelling. In such cases, preference was given to the variable with the stronger Kruskal–Wallis statistic or to the metric with clearer operational interpretation.

3.2.4.3 Random Forest modelling

Random Forest regression was used to quantify how the selected predictors jointly influence platform pollutant levels and to provide a ranked assessment of the relative importance of different drivers (Breiman, 2001; Hastie et al., 2009). Three hourly models were developed within a single framework: platform $PM_{2.5}$ and PM_{10} at Heuston, and NO_2 at Connolly Platform 4. For Heuston, the period 26 May–19 June 2023 provides 425 hours of valid PM and auxiliary data from the main platform, concourse and urban background site. For Connolly, the hourly NO_2 model uses the Platform 4 fixed monitor and the corresponding background and operational data from 1 March to 31 May 2024.

All models use the same basic approach for variable selection, model design and diagnostic analysis, and differ only in the response pollutant and the station data set used. The predictor variable set was constructed to be as consistent as data availability allowed across models and includes urban background concentration, local meteorological conditions at the station (air temperature, relative humidity, atmospheric pressure, wind speed and wind direction), hourly carriage-adjusted idling times in each of the three spatial regions, and hourly counts of arrivals, departures and pass-through trains disaggregated by train type and set length. All data were first aligned to a common hourly time base, and hours with missing key variables were excluded from the modelling.

Random Forest regression was chosen because it can capture non-linear relationships and interactions between predictors, is relatively robust to outliers and collinearity, and provides internal measures of variable importance (Breiman, 2001; Hastie et al., 2009). Each Random Forest model consists of an group of decision trees, each fitted to a bootstrap resample of the training data and using a random subset of predictors at each split (Breiman, 2001; Hastie et al., 2009). For each pollutant–station combination, the available hourly observations were randomly split into a training set comprising 80 % of the data and a test set comprising the remaining 20 %. Hyperparameters such as the number of trees in the ensemble, the maximum tree depth and the minimum number of observations in terminal nodes were tuned using 10-fold cross-validation on the training set. Model performance was evaluated on the independent test set using the coefficient of determination (R^2) and the root mean squared error (RMSE).

To interpret the fitted Random Forest models, two main outputs were examined: variable importance scores and partial dependence plots. Variable importance, computed using the mean decrease in impurity, indicates how much each predictor contributes to reducing prediction error across the ensemble and thus provides a ranked list of the most influential drivers of platform pollutant levels (Breiman, 2001). Variable importance scores are dimensionless and are normalised so that they sum to one across all predictors; larger values indicate a greater contribution to reducing prediction error in the Random Forest. Partial dependence plots show the marginal relationship between an individual predictor and the predicted pollutant concentration, averaged over the joint distribution of all other predictors (Friedman, 2001). These plots help to identify whether effects are approximately linear or strongly non-linear, whether they saturate at high or low values, and over what ranges changes in a predictor have the greatest impact on predicted concentrations.

3.3 Results

3.3.1 Low-cost sensor performance and calibration

The Ringsend co-location (reference site 1) provided an initial assessment of LCM performance relative to an EPA air quality station. Table 3.2 summarises performance metrics for each LCM against the reference at Ringsend. LCMs 12, 36, 43 and 54 showed robust mutual consistency, with inter-sensor correlations exceeding 0.94 for $PM_{2.5}$ and PM_{10}. Against the EPA reference, these sensors captured temporal variability well but exhibited a tendency to underestimate absolute $PM_{2.5}$ and PM_{10} concentrations (negative bias). In contrast, sensor 35 displayed substantial differences in both $PM_{2.5}$ and PM_{10}, with very low correlations and high RMSEs, and was therefore excluded from subsequent analysis.

Table 3.2 Performance comparison of LCMs with reference sites

Sensor	Pollutant	Sensor Mean(ug/m^3)	Ref site 1 Mean	Sensor STD	Ref site 1 STD	R^2	RMSE (ug/m^3)
12	$PM_{2.5}$	3.04	6.68	5.41	4.83	0.7	4.7
	PM_{10}	5.9	15.34	5.68	13.86	0.39	14.67
	NO_2	11.91	19.7	4.83	12.05	0.12	16.4
35	$PM_{2.5}$	35.51	6.68	26.76	4.83	0.02	39.21
	PM_{10}	40.73	15.34	27.37	13.86	0.01	38.92
	NO_2	13.22	19.7	4.67	12.05	0.06	15.41
36	$PM_{2.5}$	2.78	6.68	4.83	4.83	0.73	4.71
	PM_{10}	4.41	15.34	5.26	13.86	0.38	15.78
	NO_2	12.01	19.7	4.9	12.05	0.08	16.17
43	$PM_{2.5}$	2.54	6.68	4.51	4.83	0.71	4.92
	PM_{10}	4.96	15.34	4.91	13.86	0.39	15.47
	NO_2	13.9	19.7	6.22	12.05	0.16	16.66
54	$PM_{2.5}$	3.84	6.68	6.14	4.83	0.68	4.5
	PM_{10}	6.05	15.34	6.22	13.86	0.34	14.72
	NO_2	12.02	19.7	4.57	12.05	0.05	15.83

LCM NO_2 readings showed weak correlations with the reference analyser and large RMSEs at Ringsend. This behaviour is consistent with known limitations of low-cost electrochemical NO_2 sensors in complex ambient environments and indicates that the eLos NO_2 channel was not reliable enough for quantitative analysis. As a result, this chapter uses low-cost sensors only for $PM_{2.5}$ and PM_{10}; quantitative NO_2 analysis relies on diffusion tubes and the customised high-resolution monitor.

Table 3.3 Performance of calibration algorithms for $PM_{2.5}$ at St John's Road (LCM 36 vs

reference)

Model	Correlation (R)	RMSE (μg/m³)
Raw LCM 36 (uncalibrated)	0.85	4.45
Linear regression	0.85	3.02
Ridge regression	0.85	3.02
LASSO regression	0.85	2.93
Decision-tree regression	0.80	3.54
Random Forest regression	0.82	3.20

To correct the underestimation of PM and to obtain station-ready concentration series, a set of calibration models, trained at Ringsend, was evaluated out-of-sample using LCM 36 at the St John's Road EPA station (reference site 2) near Heuston. At this site, raw LCM 36 $PM_{2.5}$ data reproduced temporal variability but underestimated peaks, with a correlation coefficient of 0.85 and an RMSE of 4.45 μg/m³ relative to the reference analyser. Applying the candidate calibration models preserved the correlation but reduced the RMSE to varying degrees (Table 3.3). The least absolute shrinkage and selection operator (LASSO) regression gave the lowest RMSE of 2.93 μg/m³, while maintaining R = 0.85. The time-series comparison (Figure 3.11) shows that the LASSO-calibrated series closely tracks both the magnitude and temporal variability of the reference $PM_{2.5}$, with the underestimation of high episodes largely removed. On this basis, LASSO regression was selected as the preferred calibration approach for $PM_{2.5}$ and PM_{10} across all retained sensors.

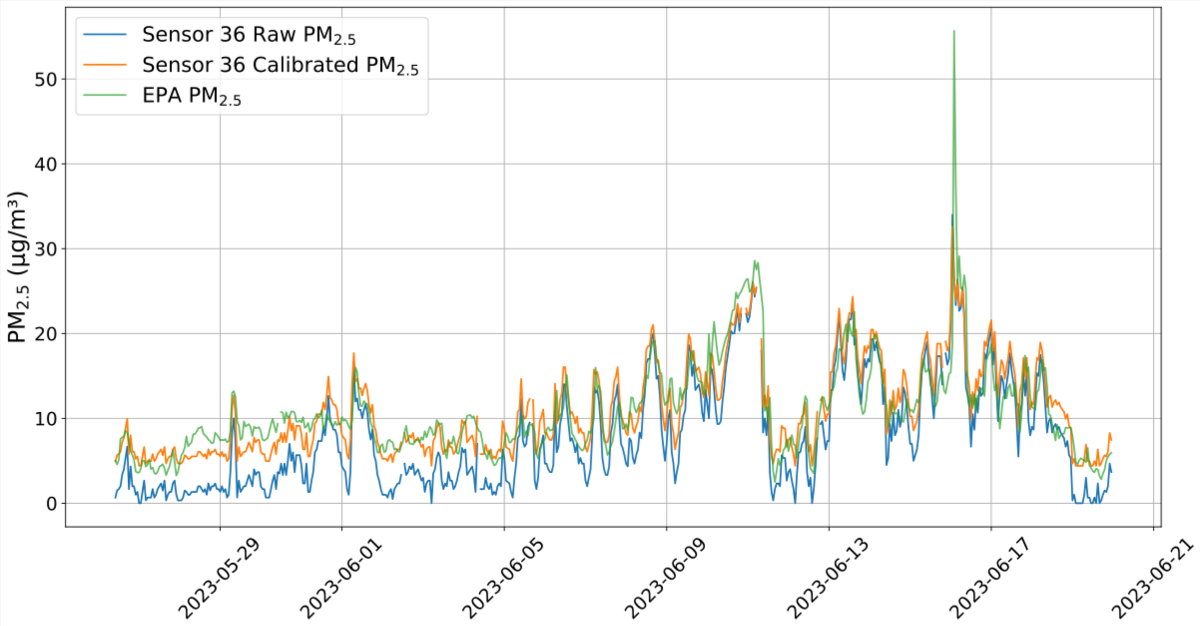


Figure 3.11 $PM_{2.5}$ concentrations at LCM 36 and St John's Road (reference site 2) before and after Lasso calibration

For each of the four retained sensors (LCM 12, 36, 43 and 54), the final LASSO models for $PM_{2.5}$ and PM_{10} reduced to simple linear correction equations of the form "calibrated concentration = a × raw concentration + b". The fitted coefficients are summarised in Table 3.4 and Table 3.5. For $PM_{2.5}$, slopes range from 0.64 to 0.88 and intercepts from about 4.2 to 4.5 µg/m³, which confirms the need for a positive offset and a modest multiplicative adjustment. For PM_{10}, slopes range from 1.29 to 1.75 and intercepts from around 6.6 to 8.1 µg/m³, reflecting a larger multiplicative correction and a substantial additive term. Together, these coefficients imply that raw LCM readings underestimate both fine and coarse particle concentrations, especially at higher levels, and that the LASSO calibration restores the magnitude of peaks while preserving temporal patterns. All subsequent analyses of $PM_{2.5}$ and PM_{10} in this chapter use concentrations corrected with these equations.

Table 3.4 LASSO calibration equations for $PM_{2.5}$

Sensor ID	Slope *a*	Intercept *b*	Calibration equation ($PM_{2.5}$)
12	0.7139	4.5135	Calibrated $PM_{2.5}$ = 0.7139 × $PM_{2.5}$_raw + 4.5135
36	0.8304	4.4095	Calibrated $PM_{2.5}$ = 0.8304 × $PM_{2.5}$_raw + 4.4095
43	0.8826	4.4786	Calibrated $PM_{2.5}$ = 0.8826 × $PM_{2.5}$_raw + 4.4786
54	0.6418	4.2268	Calibrated $PM_{2.5}$ = 0.6418 × $PM_{2.5}$_raw + 4.2268

Table 3.5 LASSO calibration equations for PM_{10}

Sensor ID	Slope *a*	Intercept *b*	Calibration equation (PM_{10})
12	1.4880	6.5563	Calibrated PM_{10} = 1.4880 × PM_{10}_raw + 6.5563
36	1.6273	8.0605	Calibrated PM_{10} = 1.6273 × PM_{10}_raw + 8.0605
43	1.7544	6.6909	Calibrated PM_{10} = 1.7544 × PM_{10}_raw + 6.6909
54	1.2926	7.3374	Calibrated PM_{10} = 1.2926 × PM_{10}_raw + 7.3374

In summary, the co-location experiments show that, once calibrated using the LASSO-derived equations, the selected low-cost monitors provide reliable hourly $PM_{2.5}$ and PM_{10} series suitable for assessing in-station air quality. The following subsections use these calibrated particulate measurements to characterise spatial and temporal patterns of pollution at Heuston and Connolly.

3.3.2 Overview of air quality levels

3.3.2.1 Variation of NO2 across Station

Passive Palmes-type diffusion tubes were deployed at multiple locations within Heuston and Connolly to provide spatially integrated NO_2 concentrations over the four-week exposure periods. The resulting daily-mean concentrations are plotted on station layout maps in Figure 3.12 and Figure 3.13, where each coloured point corresponds to an individual tube and the colour scale denotes concentration classes. The red boxes indicate the roofed or largely enclosed area at each station, allowing comparison between enclosed and more open areas.

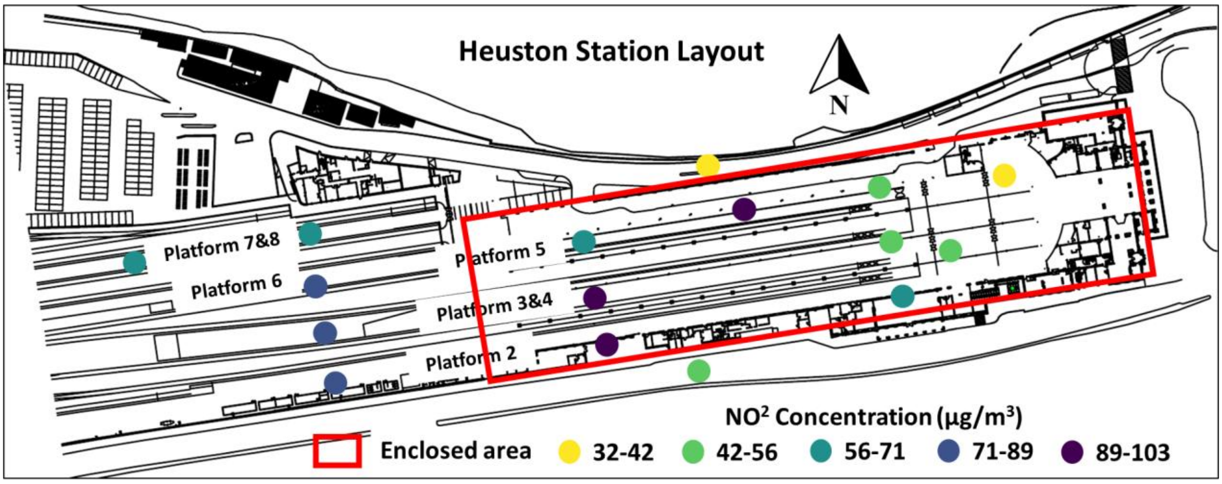


Figure 3.12 NO_2 diffusion-tube locations and daily mean concentrations at Heuston Station

At Heuston, the highest daily mean NO_2 concentration was 103 µg/m³ in the central section of the enclosed Platforms 2–5, more than twice the EU annual limit of 40 µg/m³. Moving towards the uncovered extensions of these platforms beyond the roof, concentrations fall slightly but remain high, in the range 71–89 µg/m³. Near the concourse end of the enclosed platforms, NO_2 levels drop further to about 42–56 µg/m³, and the tube in the concourse itself records a mean of 38 µg/m³, just below the EU limit. The open Platforms 6–7 exhibit NO_2 concentrations slightly lower than those at the centre of the enclosed platforms, at 56–71 µg/m³. The lowest value in the station, 32.0 µg/m³, is observed at the north pedestrian crossing outside the station.

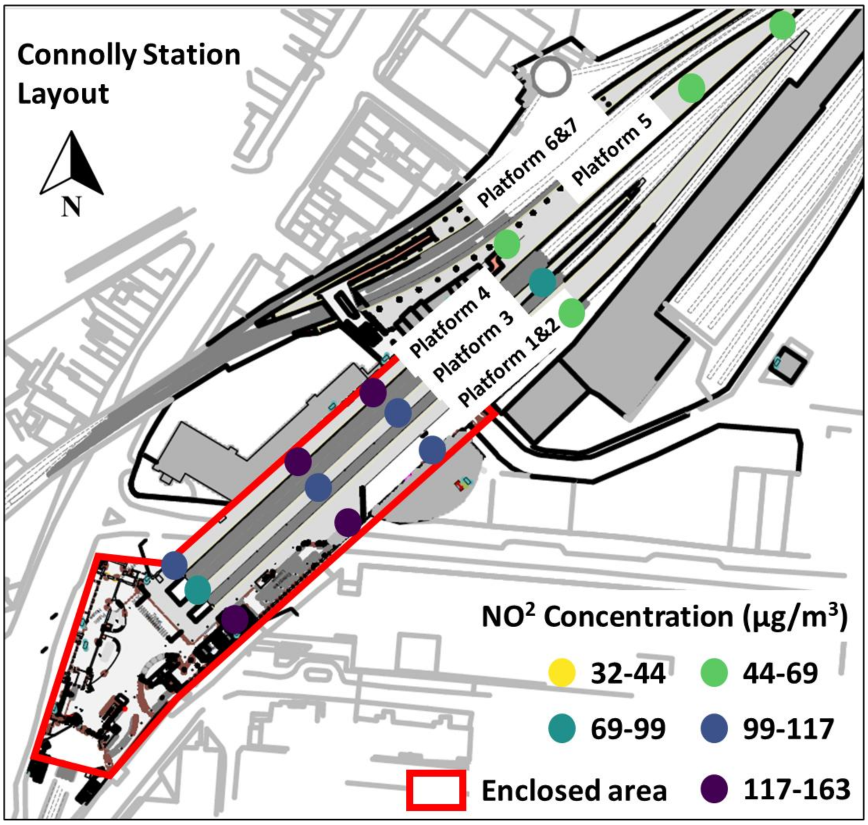


Figure 3.13 NO_2 diffusion-tube locations and daily mean concentrations at Connolly Station

At Connolly, average NO_2 concentrations differed between the enclosed and outdoor platforms. On the enclosed platforms, values during the campaign ranged from 83.0 to 162.7 µg/m³, with especially high levels along the full length of Platform 4. In contrast, NO_2 on the open platforms ranged from 40.5 to 71.5 µg/m³, with higher values near the open end of Platform 3. Concentrations in the concourse and station entrance ranged from 63 to 68 µg/m³ and 56 to 59 µg/m³, respectively. All fixed monitoring locations exceeded the EU annual NO_2 limit of 40 µg/m³. These spatial patterns motivated the selection of Platform 4 as the continuous NO_2 monitoring location in Monitoring Campaign 3.

3.3.2.2 PM2.5 and PM10 levels at Heuston

Three LCM locations inside Heuston Station are located at the platform, concourse and roof vent. Data from the St. John's Road EPA air monitoring station located 37 meters south of the station Figure 3.1 were used as background concentrations to compare pollutant concentrations inside and outside the station.

Table 3.6 Summary of 24-hour average $PM_{2.5}$ and PM_{10} concentrations at Heuston Station by

location

Location	**N days**	**$PM_{2.5}$ min (µg/m³)**	**$PM_{2.5}$ max (µg/m³)**	**$PM_{2.5}$ mean (µg/m³)**	**PM_{10} min (µg/m³)**	**PM_{10} max (µg/m³)**	**PM_{10} mean (µg/m³)**	**% days $PM_{2.5}$ > 15**	**% days PM_{10} > 45**	**Mean $PM_{2.5}/PM_{10}$**
Platform	26	6.5	21.6	12.7	18.5	41.7	27.8	30.8	0.0	0.45
Roof	26	5.8	24.1	12.2	18.6	45.8	27.8	30.8	3.8	0.43
Concourse	26	5.4	22.3	11.6	20.2	46.2	27.0	23.1	3.8	0.42
Background	26	4.1	22.8	9.7	7.4	30.5	15.4	11.5	0	0.62

At Heuston Station, 24-hour average $PM_{2.5}$ and PM_{10} levels were homogeneous between the platform, roof and concourse locations but still showed clear increments above background (Table 3.6 and Figure 3.14). During the monitoring period (26 May – 20 June 2023), daily mean $PM_{2.5}$ on the main platform ranged from 6.5 to 21.6 µg/m³ and PM_{10} from 18.5 to 41.7 µg/m³. On the roof, $PM_{2.5}$ ranged from 5.8 to 24.1 µg/m³ and PM_{10} from 18.6 to 45.8 µg/m³. In the concourse, $PM_{2.5}$ ranged from 5.4 to 22.3 µg/m³ and PM_{10} from 20.2 to 46.2 µg/m³. The WHO $PM_{2.5}$ guideline (15 µg/m³) was exceeded on about 31% of days on the platform and roof, and on 23% of days in the concourse. The PM_{10} guideline (45 µg/m³) was exceeded only on a small number of days (about 4% at the roof and concourse, and not exceeded on the platform). Figure 3.14 shows the distribution of hourly PM monitoring data.

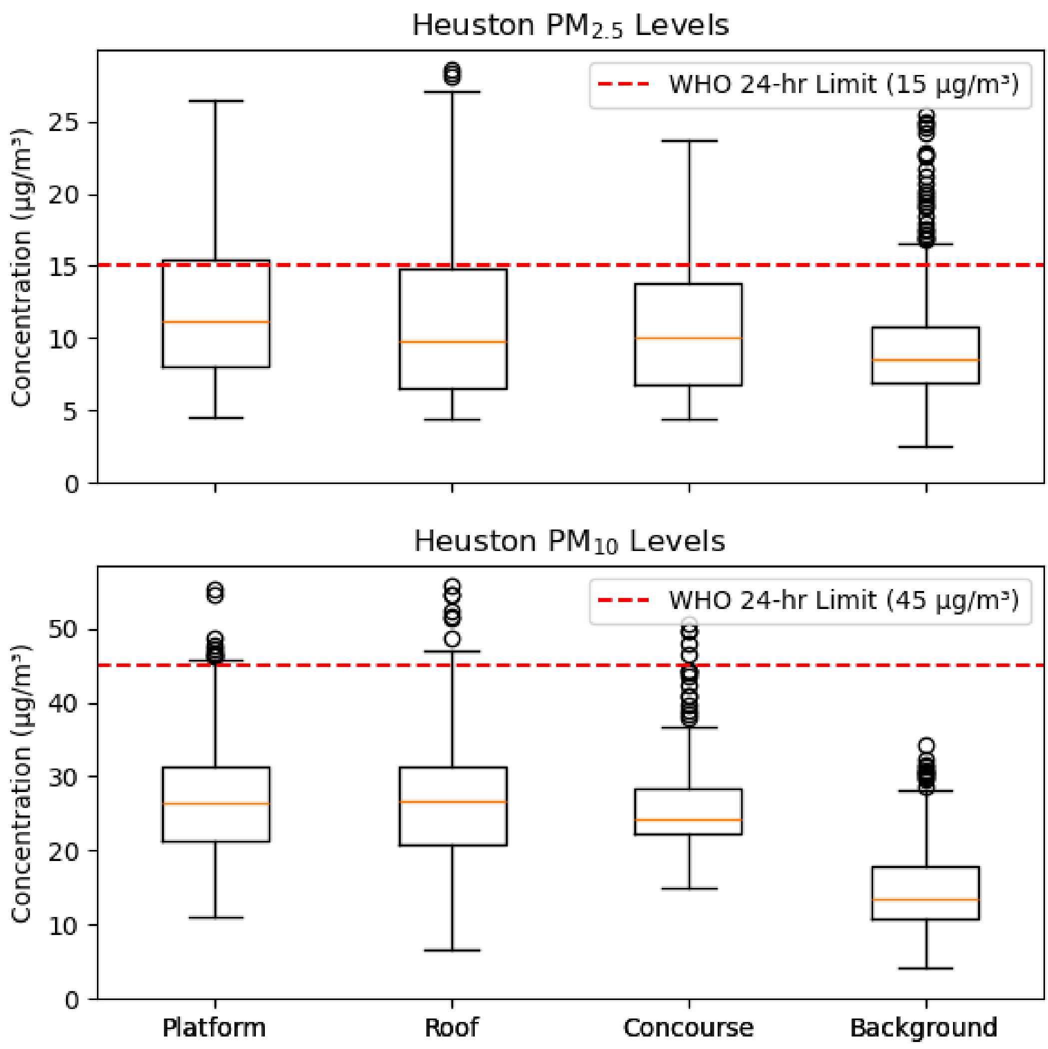


Figure 3.14 Heuston hourly $PM_{2.5}$ and PM_{10} boxplots including WHO 24-hour limit values

The diurnal pattern (Figure 3.15) at Heuston shows lower $PM_{2.5}$ and PM_{10} concentrations during the night and early morning, with an increase after 06:00 when train services begin. Concentrations remain high throughout the daytime, with peaks of $PM_{2.5}$ between late morning and early evening when train frequency is highest, and then decline again during the late evening.

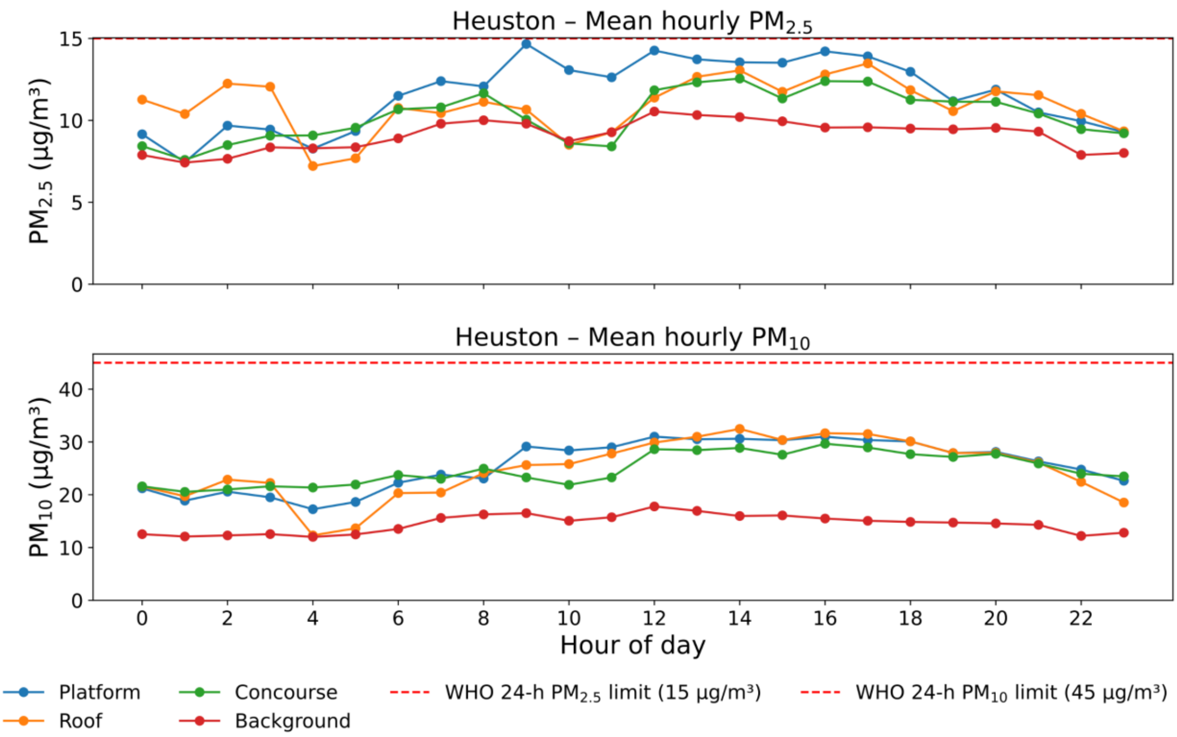


Figure 3.15 Mean diurnal $PM_{2.5}$ and PM_{10} at Heuston including WHO 24-hour limit values

The $PM_{2.5}/PM_{10}$ ratios at Heuston are lower than at the enclosed Connolly platforms (Table 3.6). Mean ratios are about 0.45 on the platform, 0.43 on the roof, and 0.42 in the concourse. This indicates a larger coarse fraction and suggests an important role for coarse outdoor particles and resuspended dust.

3.3.2.3 PM2.5 and PM10 levels at Connolly

Five LCM locations inside Heuston Station are located at the platform and concourse. Data from the EPA air monitoring station (Amiens Street) located 20 meters west of Connolly station were used as background concentrations to compare pollutant concentrations inside and outside the station. Figure 3.16 shows the distribution of hourly PM monitoring data at Connolly.

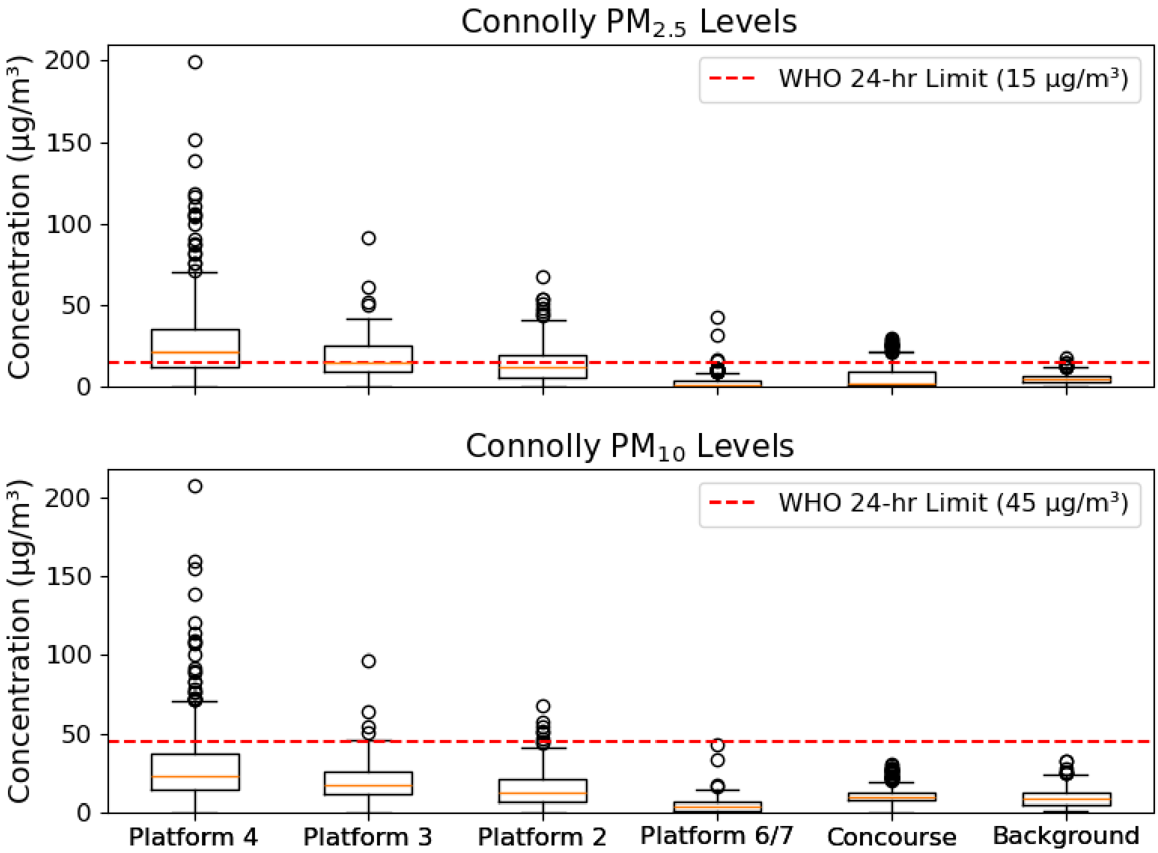


Figure 3.16 Connolly hourly $PM_{2.5}$ and PM_{10} boxplots including WHO 24-hour limit values

The 24-hour average $PM_{2.5}$ and PM_{10} levels in Connolly Station showed clear differences between enclosed and open areas (Table 3.7, Figure 3.16). On enclosed Platforms 2, 3 and 4, daily mean $PM_{2.5}$ ranged from 7.6 to 20.8 μg/m³, 11.6 to 25.0 μg/m³ and 14.2 to 57.5 μg/m³, respectively. Corresponding PM_{10} averages ranged from 7.9 to 22.7 μg/m³ on Platform 2, 14.5 to 28.3 μg/m³ on Platform 3 and 14.6 to 59.9 μg/m³ on Platform 4. The WHO 24-hour $PM_{2.5}$ guideline of 15 μg/m³ was exceeded on about 44%, 73% and 94% of monitored days on Platforms 2, 3 and 4, respectively. The PM_{10} guideline of 45 μg/m³ was exceeded on about 11% of days on Platform 4 and was not exceeded on Platforms 2 and 3. These high concentrations on the enclosed platforms are consistent with frequent diesel operations.

Table 3.7 Summary of 24-hour average $PM_{2.5}$ and PM_{10} concentrations at Connolly Station

Location	N days	$PM_{2.5}$ min (µg/m³)	$PM_{2.5}$ max (µg/m³)	$PM_{2.5}$ mean (µg/m³)	PM_{10} min (µg/m³)	PM_{10} max (µg/m³)	PM_{10} mean (µg/m³)	% days $PM_{2.5}$ > 15	% days PM_{10} > 45	Mean $PM_{2.5}/PM_{10}$
Platform 2	16	7.6	20.8	14.3	7.9	22.7	15.0	43.8	0.0	0.96
Platform 3	11	11.6	25.0	17.3	14.5	28.3	19.0	72.7	0.0	0.91
Platform 4	18	14.2	57.5	28.3	14.6	59.9	30.5	94.4	11.1	0.93
Platform 6/7	20	0.7	12.2	2.7	1.0	13.1	4.2	0.0	0.0	0.68
Concourse	20	0.5	16.7	5.9	3.1	17.4	10.2	5.0	0.0	0.52
Background	20	2.3	11.5	5.1	3.9	18.0	9.4	0	0	0.55

In contrast, much lower daily averages were observed on Platform 6/7 and in the Concourse (Table 3.7). On Platform 6/7, which mainly serves pass through diesel and electric trains and is exposed to outdoor air, 24-hour mean $PM_{2.5}$ ranged from 0.7 to 12.2 µg/m³ and PM_{10} from 1.0 to 13.1 µg/m³. In the Concourse, $PM_{2.5}$ averages ranged from 0.5 to 16.7 µg/m³ and PM_{10} from 3.1 to 17.4 µg/m³. At these locations, the $PM_{2.5}$ guideline was rarely exceeded (0% of days on Platform 6/7 and 5% in the Concourse), and the PM_{10} guideline was not exceeded. The boxplots in Figure 3.16 show that the median and upper quartile $PM_{2.5}$ values on Platforms 3 and 4 lie close to or above the WHO limit, while the distributions for Platform 6/7 and the Concourse remain well below both $PM_{2.5}$ and PM_{10} limits.

At Connolly, mean hourly $PM_{2.5}$ and PM_{10} show a diurnal cycle with strong spatial contrasts across platforms (Figure 3.17). Concentrations are highest at Platform 4 throughout the day, with elevated levels overnight (00:00–02:00), a marked early-morning minimum around 03:00–05:00, and an increase through late morning. Both pollutants exhibit a midday peak at approximately 13:00, followed by a decline in mid-afternoon and a renewed rise during the evening period. The concourse and background series remain comparatively low and stable across all hours, and Platforms 6/7 consistently show the lowest platform concentrations, which is related to its good ventilation and the large number of electric trains passing by.

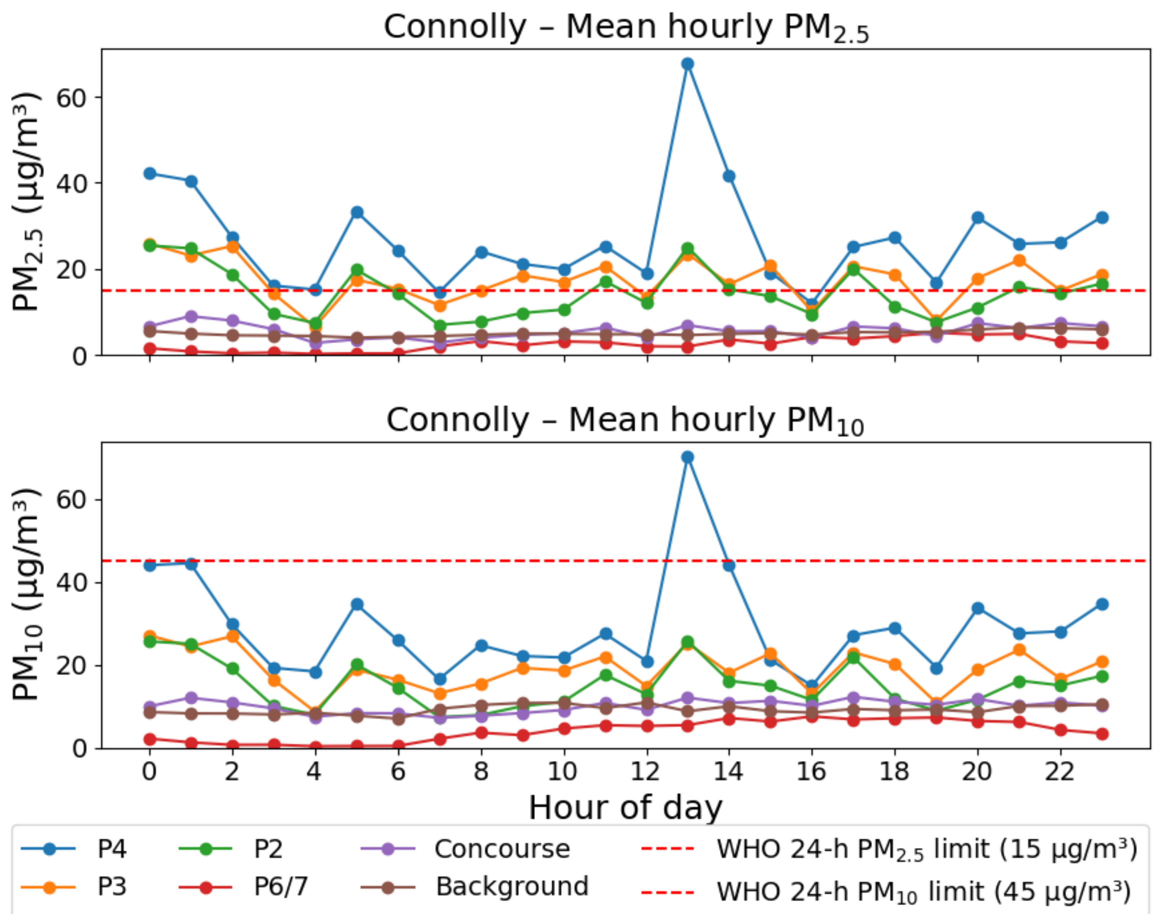


Figure 3.17 Mean diurnal $PM_{2.5}$ and PM_{10} at Connolly including WHO 24-hour limit values

The particle size fractions also differ between enclosed and open areas. On Platforms 2, 3 and 4, the mean $PM_{2.5}/PM_{10}$ ratios are 0.96, 0.91 and 0.93, indicating that the fine fraction dominates and that diesel exhaust is an important source. On Platform 6/7 and in the Concourse, the mean ratios are lower, 0.68 and 0.52, which suggests a larger contribution from coarse particles such as resuspended dust and outdoor road dust (Table 3.7).

3.3.2.4 NO_2 levels at Connolly (continuous monitoring)

Continuous NO_2 measurements on closed Platform 4 at Connolly showed very high levels compared with the nearby urban background (Table 3.8). Daily mean NO_2 concentrations on the platform ranged from 52.0 to 242.0 µg/m³, with an average of 137.9 µg/m³. Over the same period, background NO_2 ranged from 15.0 to 85.0 µg/m³, with an average of 43.1 µg/m³. The daily mean station–background increment was therefore large, with an average of about 94.8 µg/m³ and a median of 97.2 µg/m³. Even on the cleanest day, the daily mean on Platform 4 remained more than double the WHO 24-hour guideline of 25 µg/m³ for NO_2, and all 74 days in the dataset exceeded this guideline.

Table 3.8 Summary of daily mean NO_2 at Connolly Platform 4 and background NO_2 (daily means over 74 days), with exceedances of the WHO 24-hour NO_2 guideline

Location	N days	Daily mean NO_2 min (μg/m³)	Daily mean NO_2 max (μg/m³)	Mean daily NO_2 (μg/m³)	Median daily NO_2 (μg/m³)	% days with daily mean > 25 μg/m³
Platform 4 (NO_2)	74	52.0	242.0	137.9	141.9	100.0
Background NO_2	74	15.0	85.0	43.1	44.0	75.7

The contrast is also clear in the hourly data (Table 3.9, Figure 3.18). Hourly NO_2 on the platform ranged from 8.1 to 622.2 μg/m³, with a mean of 138.2 μg/m³ and a 95th percentile of 307.2 μg/m³. Background hourly NO_2 ranged from 3.2 to 126.9 μg/m³, with a mean of 42.5 μg/m³. The hourly station–background increment had a mean of 95.5 μg/m³ and a 75th percentile of about 134.9 μg/m³, showing that three-quarters of hours had NO_2 at least 35 μg/m³ above background and many hours far higher. Compared with the WHO 1-hour NO_2 guideline of 200 μg/m³, about 18.4% of all hours on Platform 4 exceeded this level, and roughly 1.9% of hours were above 400 μg/m³. On 78.4% of days, the daily maximum hourly NO_2 was above 200 μg/m³, indicating frequent short-term peaks that are likely to affect passengers and staff during busy periods.

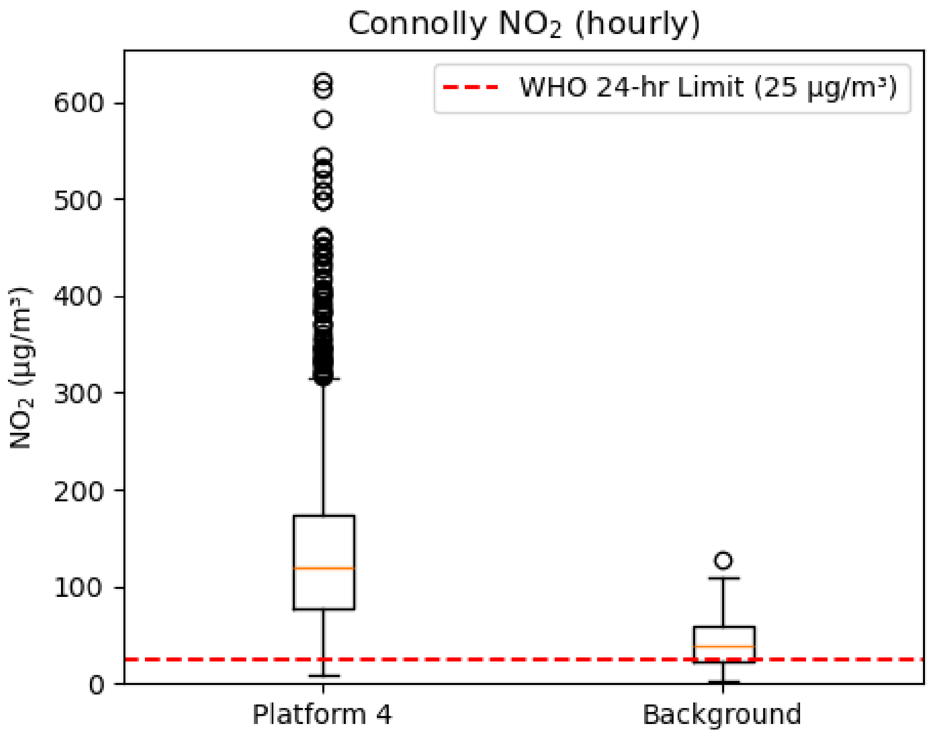


Figure 3.18 Boxplots of hourly NO_2 concentrations at Connolly Platform 4 and at the background site including WHO 24-hour limit values

Table 3.9 Summary of hourly Connolly NO_2 concentrations on Platform 4, background NO_2, and station–background increment

Statistic	Connolly NO_2 (µg/m³)	Background NO_2 (µg/m³)	Station − background (µg/m³)
N (hours)	1 735	1 710	1 710
Mean	138.24	42.48	95.50
Std dev	86.85	23.83	92.53
Min	8.07	3.15	−64.06
5th percentile	39.18	9.37	−9.53
25th percentile	77.89	22.24	34.27
Median	119.39	39.73	72.47
75th percentile	173.45	59.00	134.86
95th percentile	307.18	85.50	277.39
Max	622.20	126.92	600.78

The diurnal mean NO_2 (Figure 3.19) is lowest in the late evening and early night, around 90–150 µg/m³ between 19:00 and 03:00, and increases sharply during the morning and midday. A first peak appears at 05:00 (about 177 µg/m³), associated with the early outbound diesel services. Concentrations then rise further, with mean values of about 204 µg/m³ at 11:00 and a maximum of around 229 µg/m³ at 13:00. After the early afternoon, NO_2 gradually declines but remains above 120 µg/m³ for much of the day. Weekday mean NO_2 (142.4 µg/m³) is slightly higher than at weekends (127.5 µg/m³), which is consistent with the denser weekday diesel timetable and shorter gaps between services.

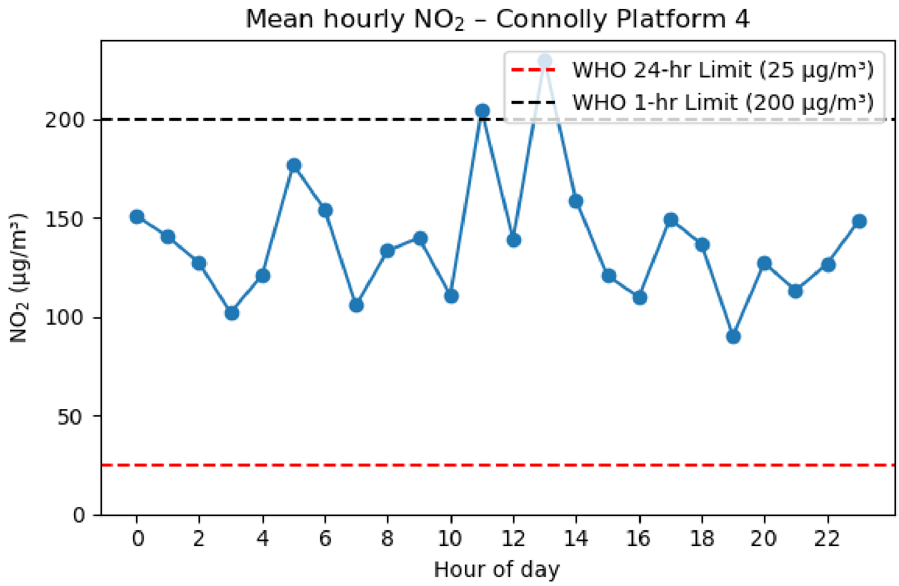


Figure 3.19 Connolly mean diurnal NO_2 concentrations and comparison including WHO 24-hour and 1-hour limit values

Overall, the NO_2 results at Connolly confirm and strengthen the conclusions from the PM analysis. While $PM_{2.5}$ and PM_{10} are high on the enclosed diesel platforms and show clear increments above background, NO_2 levels are much more extreme, with daily means always above the WHO 24-hour guideline and frequent exceedances of the 1-hour limit. This indicates that diesel exhaust is the dominant source of NO_2 on Platform 4 and that ventilation in the enclosed platform area is not sufficient to prevent high short-term exposures during busy periods.

3.3.3 Operational drivers of in-station pollution

This section examines how train operations influence air pollution inside the stations. The analysis focuses on $PM_{2.5}$ and PM_{10} at Heuston and on NO_2 at Connolly. For each station, pollutant levels are related to hourly train frequency and to the cumulative time that trains spend idling at the platforms. Because on-train monitoring (OTMR) data at Connolly do not cover the same period as the low-cost monitoring (LCM) campaign, the operational analysis is restricted to PM at Heuston and NO_2 at Connolly.

3.3.3.1 Train frequency vs PM at Heuston

The average hourly train frequency at Heuston is shown in Figure 3.20. Over the monitoring period, the mean number of trains is about 4.7 trains/hour, with a clear weekday pattern. From Monday to

Friday, the mean hourly frequency is between 5.5 and 6.6 trains/hour, while on Saturdays and Sundays it falls to 4.4 and 3.4 trains/hour, respectively. Almost no trains operate between 01:00 and 05:00. The maximum mean frequency occurs between 09:00 and 10:00, when about 9–10 trains per hour use the station, followed by a broad plateau of 6–9 trains/hour until the early evening. After 20:00, the frequency declines steadily and drops below 4 trains/hour by 23:00.

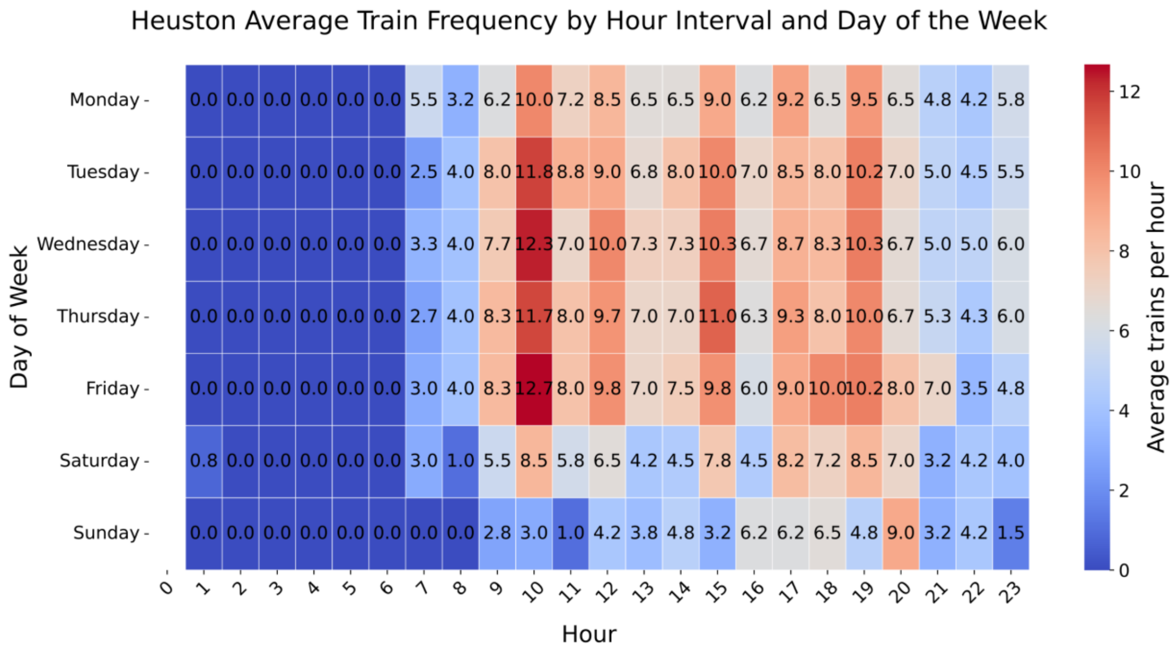


Figure 3.20 Heuston train frequency heat map

The diurnal pattern of platform PM follows the train timetable. Mean hourly platform $PM_{2.5}$ increases from about 8–9 μg/m³ during the night to about 14–15 μg/m³ during the daytime, and declines again in the late evening. Platform PM_{10} shows a similar rise, from around 18–19 μg/m³ at night to more than 30 μg/m³ during the busiest hours. A regression using the 24-hour mean data shows that platform $PM_{2.5}$ is strongly related to the average hourly number of trains, with $R^2 = 0.75$ and a slope of 0.56 μg/m³ per additional train/hour. Platform PM_{10} is even more sensitive, with $R^2 = 0.83$ and a slope of 1.28 μg/m³ per train/hour (Table 3.10, Figure 3.21).

Table 3.10 Heuston Linear regression between hourly mean $PM_{2.5}$/PM_{10} and total train frequency at Heuston

Location	Pollutant	R^2 vs total trains	Slope (μg/m³ per train/hour)
Platform	$PM_{2.5}$	0.75	0.56
Platform	PM_{10}	0.83	1.28

Location	Pollutant	R^2 vs total trains	Slope (µg/m³ per train/hour)
Concourse	$PM_{2.5}$	0.41	0.29
Concourse	PM_{10}	0.53	0.67

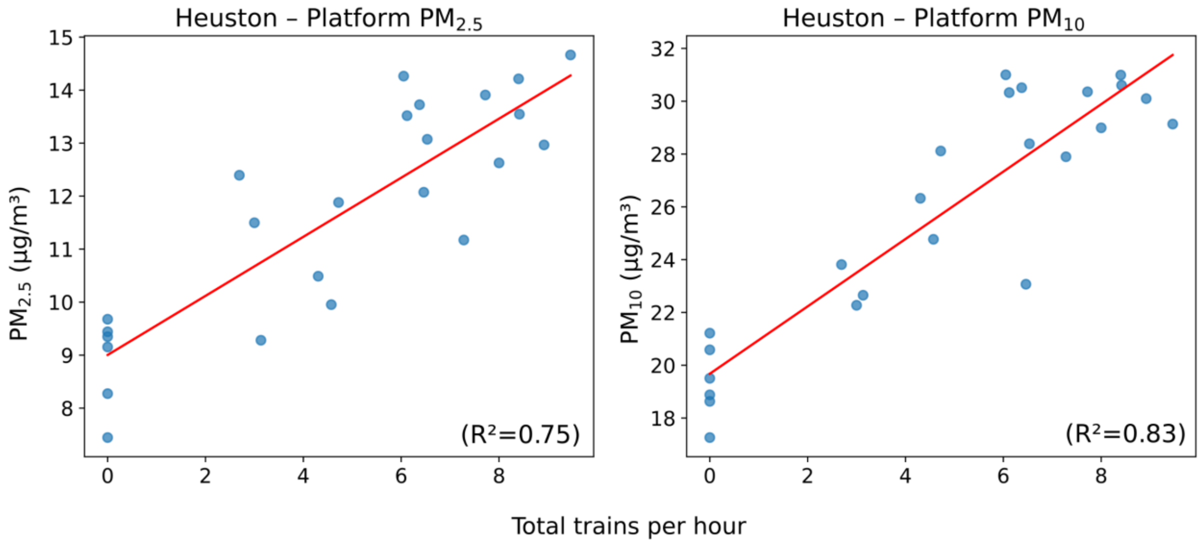


Figure 3.21 Relationship between train frequency and platform PM at Heuston station

At the concourse in Heuston station, $PM_{2.5}$ and PM_{10} show only a moderate association with train frequency. For concourse $PM_{2.5}$, R^2 is 0.41 with a slope of 0.29 µg/m³ per train/hour, and for concourse PM_{10}, R^2 is 0.53 with a slope of 0.67 µg/m³ per train/hour (Table 3.10). This weaker relationship reflects the shielding effect of the glass curtain wall, which reduces direct exchange between the platform and concourse.

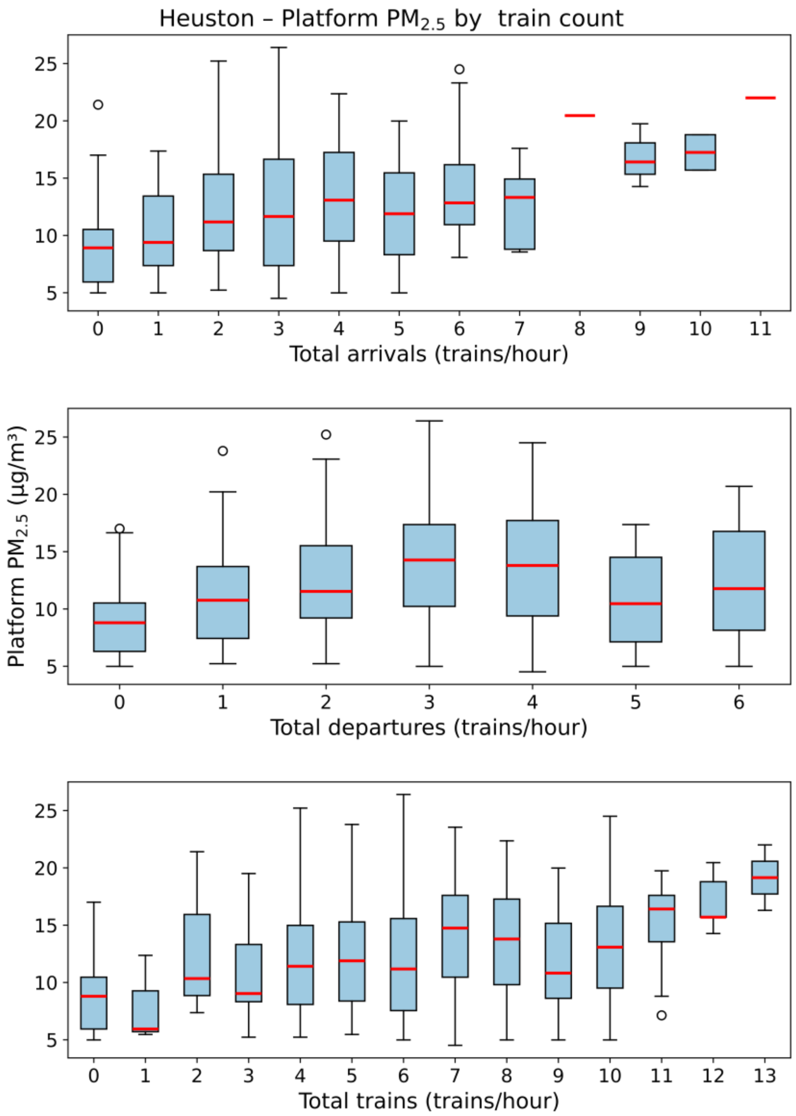


Figure 3.22 Boxplots of Platform PM$_{2.5}$ at Heuston station by train arrivals, departures and total trains

Box-and-whisker plots provide more detail on the relationship between train frequency and platform PM$_{2.5}$ (Figure 3.22). Hourly platform PM$_{2.5}$ is grouped by the number of arriving, departing, and total trains per hour. For all three metrics, broadly median PM$_{2.5}$ increases with train frequency. When there are 0–2 trains/hour, median platform PM$_{2.5}$ is around 9–10 µg/m^3, while for 1–12 trains/hour it rises to about 17–179 µg/m^3. Arriving and departing trains do not exhibit distinctly different patterns.PM$_{10}$ boxplots (Figure 3.23) show a similar pattern.

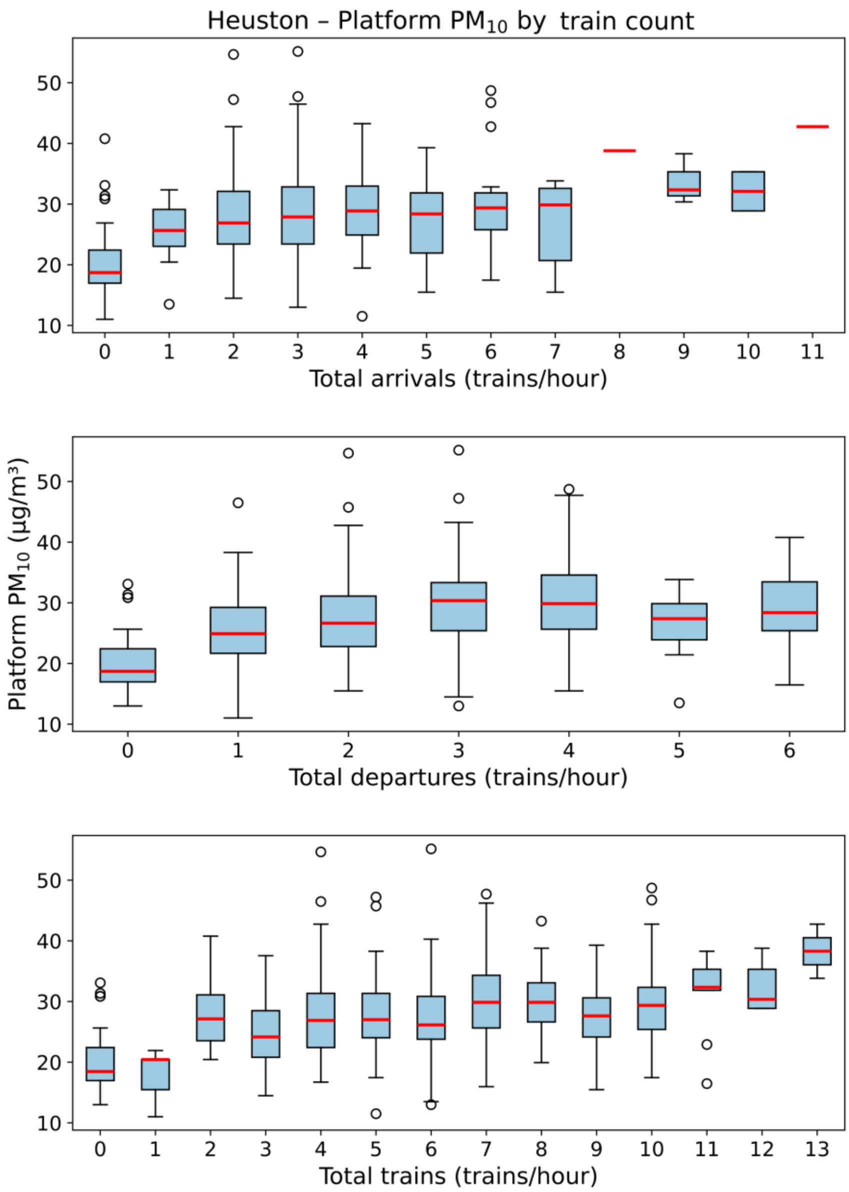


Figure 3.23 Boxplots of Platform PM_{10} at Heuston station by train arrivals, departures and total trains

3.3.3.2 Idling time vs PM pollution at Heuston

Heuston is a terminal station, so all services arrive, dwell, and then depart with a new train ID. During this layover, trains usually stay on the platform and engines are not always shut down. In this study, idling time is defined as the period when a train is stationary, but the engine remains on. Idling emissions can therefore make a large contribution to local air pollution.

Between 25 May and 26 June 2023, 4,243 idling episodes were identified at Heuston, with a total duration of 128,138 minutes. These idling events were assigned to three spatial regions along the platforms. Cumulative idling times were 50,643 minutes in Region 1, 57,431 minutes in Region 2, and 20,1630 minutes in Region 3. The distribution of individual idling periods is right-skewed (Figure 4.2). About 34.5% of idling episodes last 0–15 minutes and 22.8% last 15–30 minutes, while only around 11% last longer than 60 minutes.

Hourly mean platform $PM_{2.5}$ shows a strong dependence on cumulative idling time. Using the 24 hourly means, regressions between $PM_{2.5}$ and mean hourly idling time is $R^2 = 0.77$, 0.65 and 0.56 for Regions 1, 2 and 3, with slopes of about 0.012, 0.012 and 0.026 μg/m³ per idling minute, respectively (Table 3.11). Platform PM_{10} behaves similarly, with $R^2 = 0.64$, 0.57 and 0.45 and slopes of about 0.025, 0.025 and 0.050 μg/m³ per minute in Regions 1, 2 and 3. Thus, an additional 30 minutes of idling in Region 1 is associated with an increase of about 0.4 μg/m³ in mean hourly $PM_{2.5}$ and about 0.8 μg/m³ in PM_{10}.

Table 3.11 Regression between hourly mean platform $PM_{2.5}$/PM_{10} and hourly mean cumulative idling time in each region, and between idling time and total train frequency at Heuston.

Region	$PM_{2.5}$ R^2	$PM_{2.5}$ slope (μg/m³ per min)	PM_{10} R^2	PM_{10} slope (μg/m³ per min)	Train frequency R^2	Idling slope (min per train/hour)
Region 1	0.77	0.012	0.64	0.025	0.60	35.1
Region 2	0.65	0.012	0.57	0.025	0.57	32.2
Region 3	0.56	0.026	0.45	0.050	0.40	11.8

Hourly idling time is itself related to train frequency. The regression between mean hourly idling and total trains per hour gives R^2 values of 0.60, 0.57 and 0.40 in Regions 1, 2 and 3, with slopes between 11.8 and 35.1 minutes per additional train (Table 3.11). This indicates that hours with many train movements tend to have long cumulative idling times.

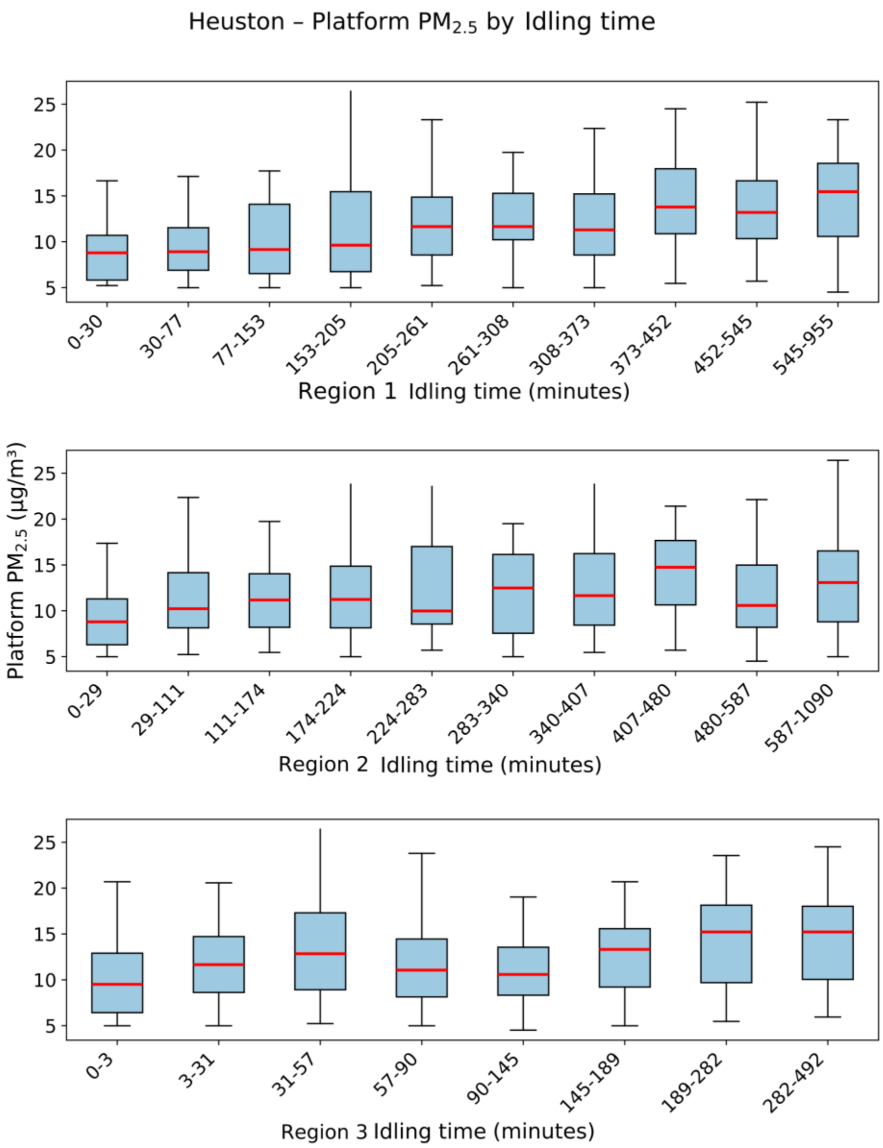


Figure 3.24 Boxplots of Platform $PM_{2.5}$ at Heuston station by idling time

To explore the shape of the relationship between idling and $PM_{2.5}$, hourly platform $PM_{2.5}$ concentrations were plotted against deciles of cumulative idling time in each region (Figure 3.24). In Region 1 there is a clear upward trend: median $PM_{2.5}$ increases from about 9 µg/m³ in the lowest idling decile to about 16 µg/m³ in the highest. Regions 2 and 3 show weaker and less stable trends, likely because these areas are more open and more strongly influenced by wind and outdoor air. PM_{10} boxplots (Figure 3.25) show a similar pattern, with increasing medians and a widespread at high idling levels.

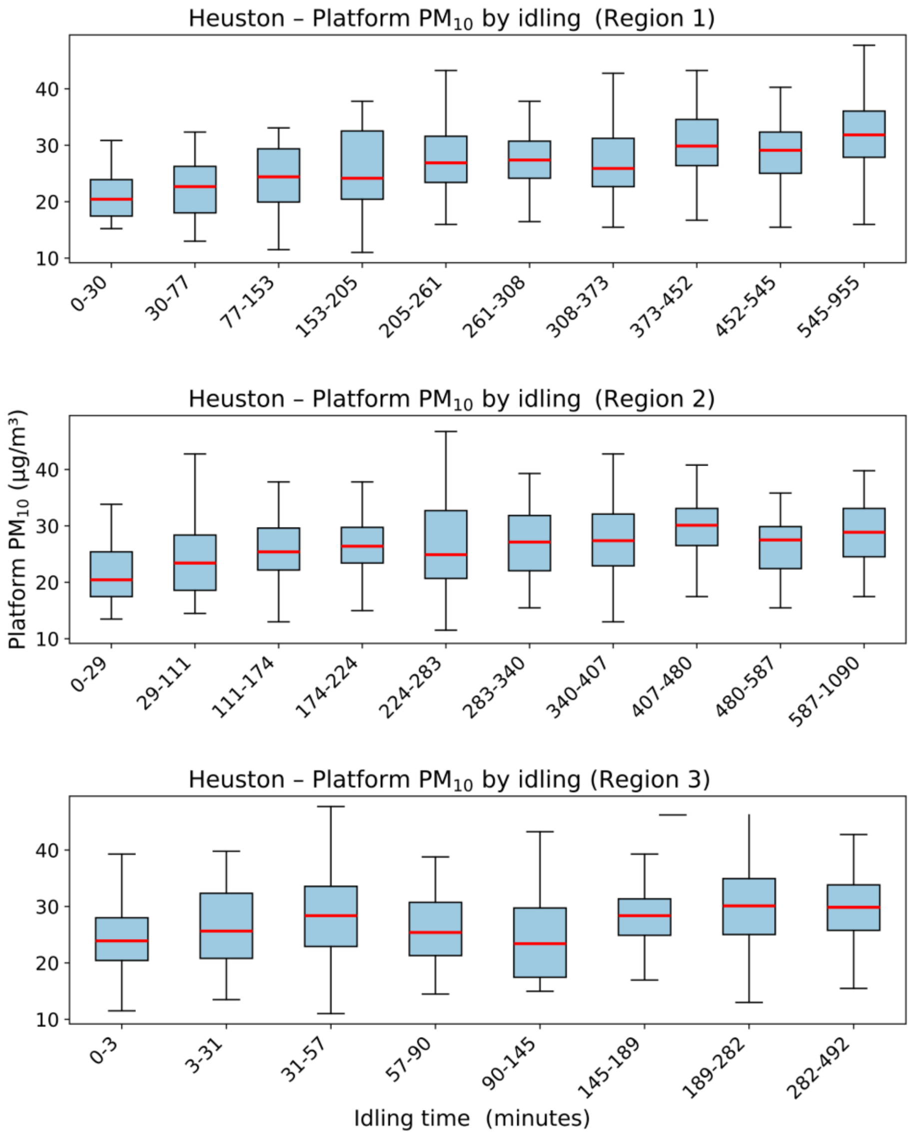


Figure 3.25 Boxplots of Platform PM10 at Heuston station by idling time

3.3.3.3 Train frequency vs NO_2 at Connolly

At Connolly, both electric and diesel services operate. Diesel terminal trains (Class 201 locomotive-hauled sets and DMU 29000/22000 units) mainly use the enclosed platforms 1-4, while pass and electric services use the open platforms 5-7. During the NO_2 monitoring period (March–May 2024), Platforms handled on average 9.5 train movements per hour, with higher frequencies on weekdays than weekends. The hourly timetable has very few movements between 01:00 and 04:00, with a sharp increase during the morning and peaks of 16–17 trains/hour in the late morning and early afternoon (Figure 3.26).

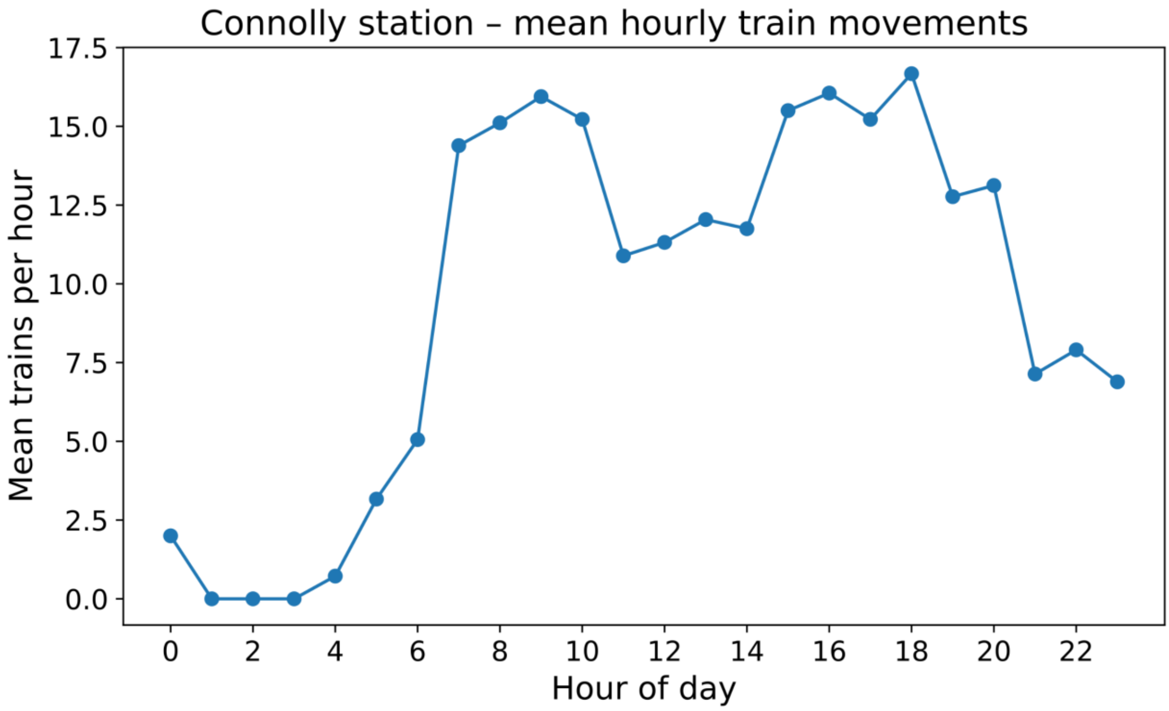


Figure 3.26 Mean hourly train movements at Connolly station

Mean hourly NO_2 at Platform 4 rises from about 100 µg/m³ at night to around 230 µg/m³ at 13:00. However, total train frequency does not explain much of this variation. A regression between the 24 hourly mean NO_2 values and the corresponding mean total train frequency gives $R^2 \approx 0.00$ (Table 3.12, Figure 3.27). This means that NO_2 is not simply a function of the total number of train movements.

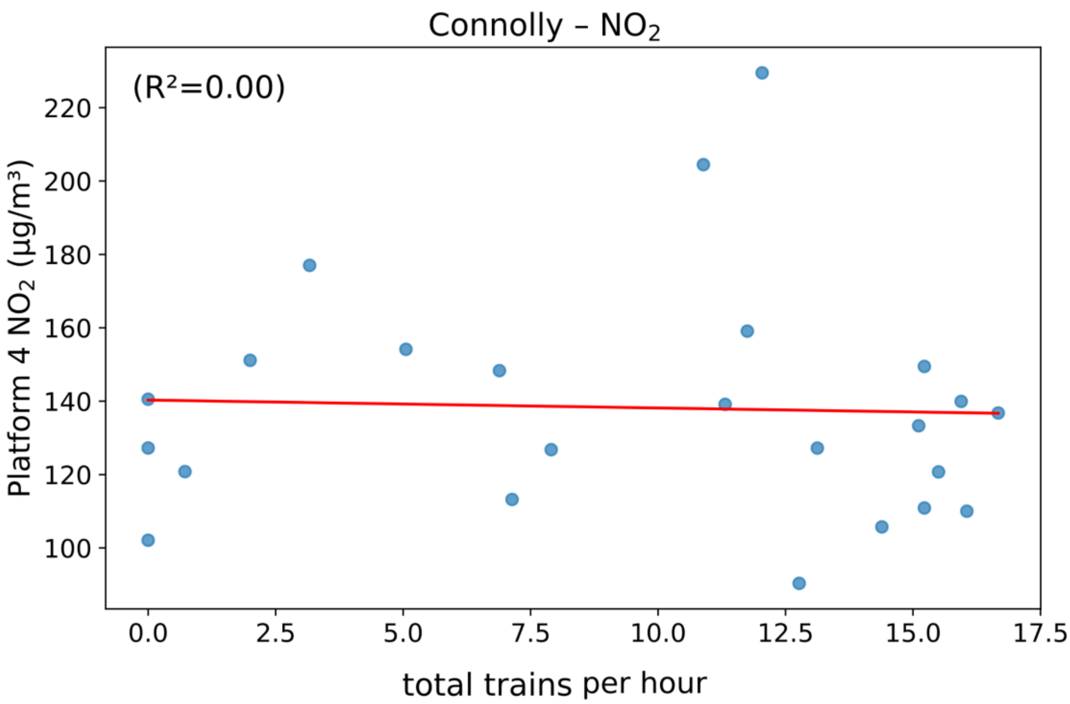


Figure 3.27 Relationship between NO_2 at Connolly Platform 4 and train activity

When arrivals, departures and passes are considered separately Table 3.12, total arrivals show almost no association with NO_2 ($R^2 \approx 0.00$). Total pass-through trains have a weak negative association ($R^2 \approx 0.13$, slope −3.51 µg/m³ per pass per hour), consistent with pass-through movements spending little time in the enclosed hall and often occurring under better-ventilated conditions. Total departures show a positive association ($R^2 \approx 0.16$, slope 5.77 µg/m³ per departure per hour), suggesting that departure events contribute more strongly to NO_2 than arrivals or passes, but still do not fully explain the variability.

Table 3.12 Regression between hourly mean NO_2 and train activity or idling at Connolly (24 hour-of-day means)

Predictor	R^2	Slope (μg/m³ per unit)
Total trains per hour	0.00	−0.22 per train/hour
Total arrivals per hour	0.00	0.54 per arrival/hour
Total departures per hour	0.16	5.77 per departure/hour
Total pass-through trains per hour	0.13	−3.51 per pass/hour
Idling Region 1 (minutes/hour)	0.56	0.40 per idling minute/hour
Idling Region 2 (minutes/hour)	0.02	0.10 per idling minute/hour
Idling Region 3 (minutes/hour)	0.08	−0.50 per idling minute/hour

Overall, these results show that, even when arrivals, departures and pass-through trains are considered separately, train counts alone explain only a small share of the hour-to-hour variation in NO_2 at Connolly. Departures have the clearest link, but the explained variance remains modest compared with the total variability in NO_2.

3.3.3.4 Idling time vs NO_2 at Connolly

Hourly cumulative carriage-adjusted idling times were calculated for Regions 1–3 at Connolly. Across the 90-day period, total idling amounts were approximately 157,500 minutes in Region 1, 83,700 minutes in Region 2 and 29,200 minutes in Region 3. Region 1, covering the head of the enclosed platforms near the NO_2 monitor, experiences the longest idling durations.

Using hour-of-day means, regressions between NO_2 and mean idling in Region 1 give $R^2 = 0.56$ and a slope of 0.40 μg/m³ per idling minute per hour (Table 3.12). On average, an additional 30 minutes of carriage-adjusted idling in Region 1 during a given hour is associated with an increase of ≈12 μg/m³ in mean NO_2.

For Regions 2 and 3, relationships are weaker. Regression between NO_2 and mean idling is $R^2 \approx 0.02$ and 0.08 with slopes of 0.10 and −0.50 μg/m³ per idling minute per hour, respectively. These low R^2 values indicate that idling further from the enclosed monitoring location contributes less to the Platform 4 NO_2 measurements and is more strongly modulated by ventilation and background conditions.

Boxplots of hourly NO_2 grouped by idling-time deciles (Figure 3.28) confirm this pattern. In Region 1, median NO_2 increases across idling deciles, with higher idling associated with higher typical NO_2 and more frequent extreme values. In Regions 2 and 3, trends are weak and non-monotonic.

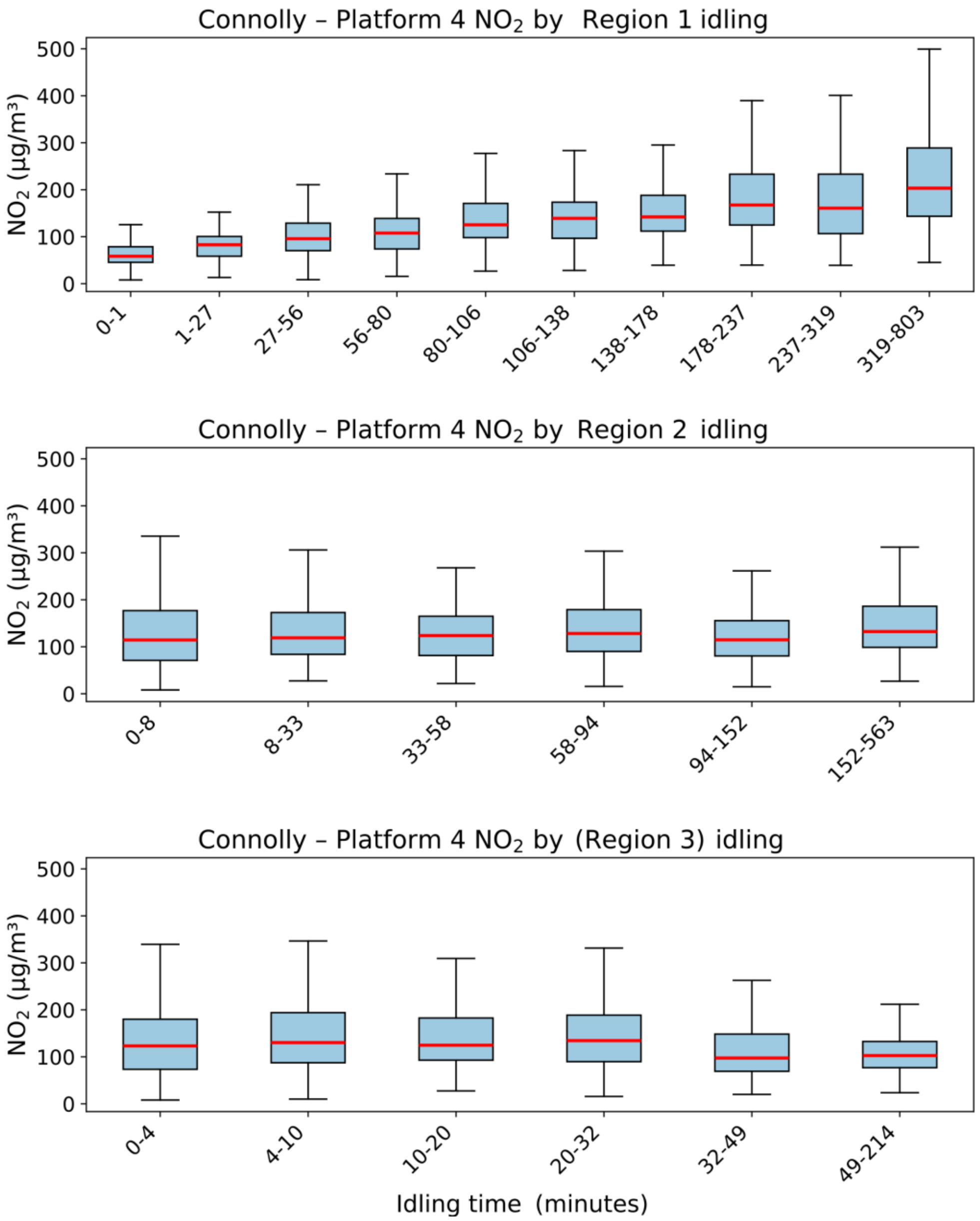


Figure 3.28 Boxplots of NO_2 at Connolly Platform 4 by idling-time in Regions 1–3

3.3.3.5 Comparison between Heuston and Connolly

At Heuston, which operates only diesel services, simple operational indicators such as total train frequency and cumulative idling time explain a large share of the variability in platform $PM_{2.5}$ and PM_{10}. Using hour-of-day averages, regressions between platform PM and total trains or idling give

R^2 values above 0.65 in the main regions. This means that the number of trains and the duration of idling are good proxies for changes in particulate pollution at this terminal diesel station.

At Connolly, NO_2 levels on the enclosed diesel platforms are much higher in absolute terms, but they are not strongly controlled by simple train counts. Using the same hour-of-day averaging approach, the regression between mean NO_2 and total trains per hour yields $R^2 \approx 0.00$, indicating no detectable linear relationship. Regressions using arrivals and pass-through trains as predictors likewise produce very low R^2 values. Departures have a moderate positive association with NO_2 ($R^2 \approx 0.16$), but still leave most of the variation unexplained. In contrast, carriage-adjusted idling in Region 1 shows a stronger link to NO_2 ($R^2 \approx 0.56$), while idling in Regions 2 and 3 remains weakly related.

These results suggest that, in a diesel-only terminal station like Heuston, total train movements and idling times are good operational indicators for managing $PM_{2.5}$ and PM_{10} at the enclosed platform. In a mixed diesel–electric station with enclosed and open platforms like Connolly, NO_2 is driven by a combination of idling at enclosed platforms, background concentrations and meteorology, rather than by the simple number of arrivals, departures or pass-through trains. Control strategies that focus only on reducing train counts are therefore unlikely to be sufficient at Connolly; measures that limit idling and improve ventilation around the enclosed diesel platforms are likely to be more effective.

3.3.4 Random Forest results for PM at Heuston

Kruskal–Wallis screening for the Heuston $PM_{2.5}$ and PM_{10} datasets identified background concentrations (µg/m³), meteorological variables (air temperature (℃), atmospheric pressure (kPa), wind speed(m/s) and wind direction), total train count, adjusted regional idling times (mins) and several train-type-specific arrival and departure counts as significant predictors (Table 3.13 and Table 3.14). Variables with very weak or redundant effects (e.g. some rarely occurring set-length categories and relative humidity for the $PM_{2.5}$ model) were excluded to simplify the models. Exclusion in this context reflects statistical redundancy or limited independent information rather than a judgement that these variables are unimportant in physical terms; their effects are largely captured by correlated predictors that remain in the model.

Table 3.13 Kruskal–Wallis test results for hourly $PM_{2.5}$ at Heuston platform (Variables marked in

red are excluded)

Variable	Chi-square	p-value	df	Significance
Background_PM2.5_Category	230.41	0.0000	9	***
Pres_Category	118.85	0.0000	9	***
Idtime_Region 1_Adjusted_Category (Min)	59.92	0.0000	9	***
Wind Direction_Category	50.24	0.0000	9	***
Tem_Category	49.57	0.0000	9	***
Wind Speed_Category	48.15	0.0000	9	***
Idtime_Region 3_Adjusted_Category	38.51	0.0000	8	***
5 DMU_Departure_Category	38.02	0.0000	3	***
3 DMU_Departure_Category	34.66	0.0000	2	***
3 DMU_Arrival_Category	32.36	0.0000	4	***
Idtime_Region 2_Adjusted_Category	25.24	0.0027	9	**
5 DMU_Arrival_Category	20.62	0.0001	3	***
4 DMU_Arrival_Category	19.42	0.0006	4	***
6 DMU_Arrival_Category	17.27	0.0006	3	***
8 MKIV_Arrival_Category	9.84	0.0073	2	**
4 DMU_Departure_Category	8.09	0.0175	2	*
6 DMU_Departure_Category	4.18	0.1236	2	+
Humidity	13.11	0.1576	9	+
7 DMU_Departure_Category	0.36	0.8362	2	
7 DMU_Arrival_Category	0.22	0.8963	2	

Table 3.14 Kruskal–Wallis test results for hourly PM_{10} at Heuston platform (Variables marked in red are excluded)

Variable	Chi-square	p-value	df	Significance
Background_PM10_Category	160.53	0.0000	9	***
Wind Speed_Category	151.35	0.0000	9	***
Temp_Category	135.95	0.0000	9	***

Variable	Chi-square	p-value	df	Significance
humidity_Category	116.04	0.0000	9	***
Wind Direction_Category	105.86	0.0000	9	***
Idtime_Region 1_Category	89.59	0.0000	9	***
Pressure_Category	86.71	0.0000	9	***
5 ICR_Departure_Category	68.27	0.0000	3	***
3 ICR_Departure_Category	60.36	0.0000	2	***
3 ICR_Arrival_Category	60.61	0.0000	3	***
Idtime_Region 3_Category	43.06	0.0000	7	***
Idtime_Region 2_Category	41.42	0.0000	9	***
5 ICR_Arrival_Category	40.22	0.0000	3	***
8 MKIV_Departure_Category	64.24	0.0000	1	***
8 MKIV_Arrival_Category	25.31	0.0000	2	***
4 ICR_Arrival_Category	27.59	0.0000	3	***
6 ICR_Arrival_Category	21.38	0.0000	2	***
4 ICR_Departure_Category	16.66	0.0002	2	**
6 ICR_Departure_Category	14.71	0.0006	2	**
7 ICR_Departure_Category	5.28	0.0714	2	+
7 ICR_Arrival_Category	0.07	0.7916	1	

The $PM_{2.5}$ model shows good performance, with an R^2 of 0.71 and an RMSE of 2.86 μg/m³ on the test data. The ranking of variable importance (Figure 3.29) confirms that urban background $PM_{2.5}$ is the dominant predictor, explaining about 62% of the variance in the model. Idling time in Region 1 is the second most important variable (9% of total importance), followed by pressure, wind direction, temperature and wind speed. Idling times in Regions 2 and 3 also contribute, but to a smaller extent. Individual train movements (arrivals and departures for different set lengths) have much lower importance scores, each contributing less than 1 % of the total variable importance.

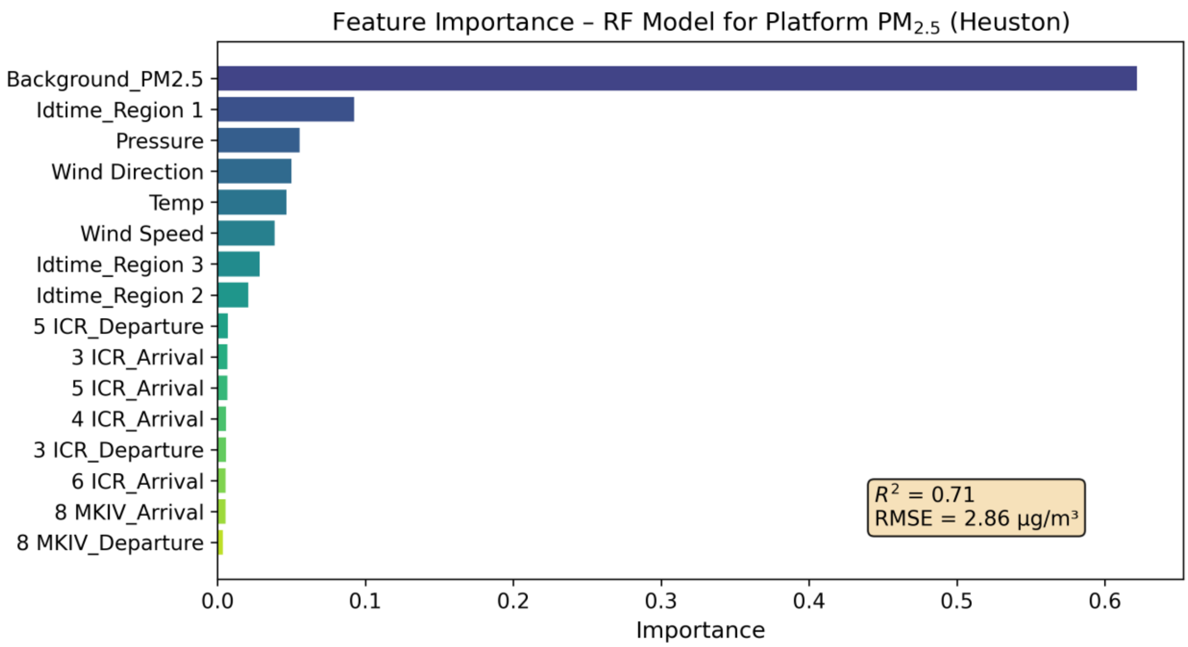
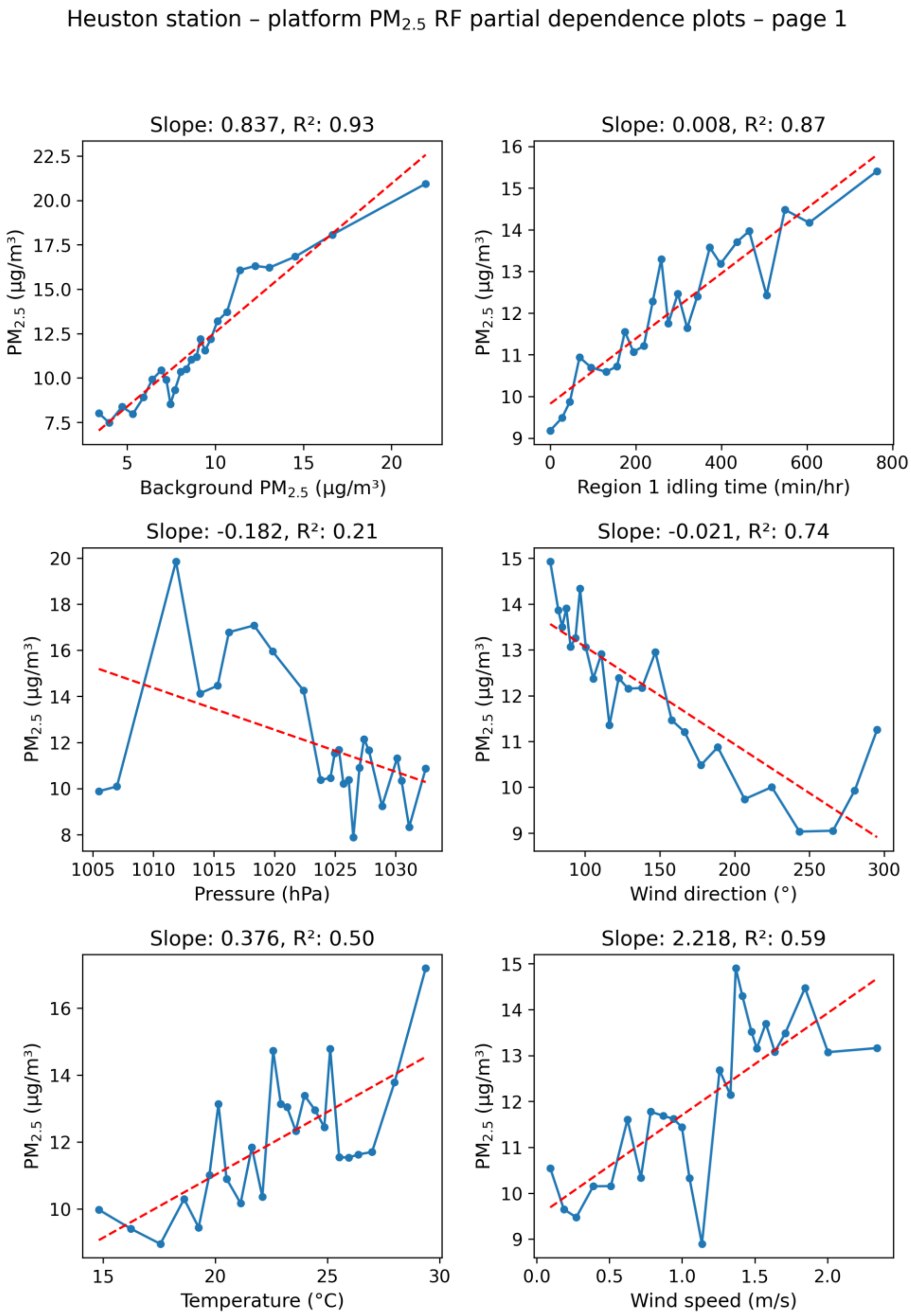
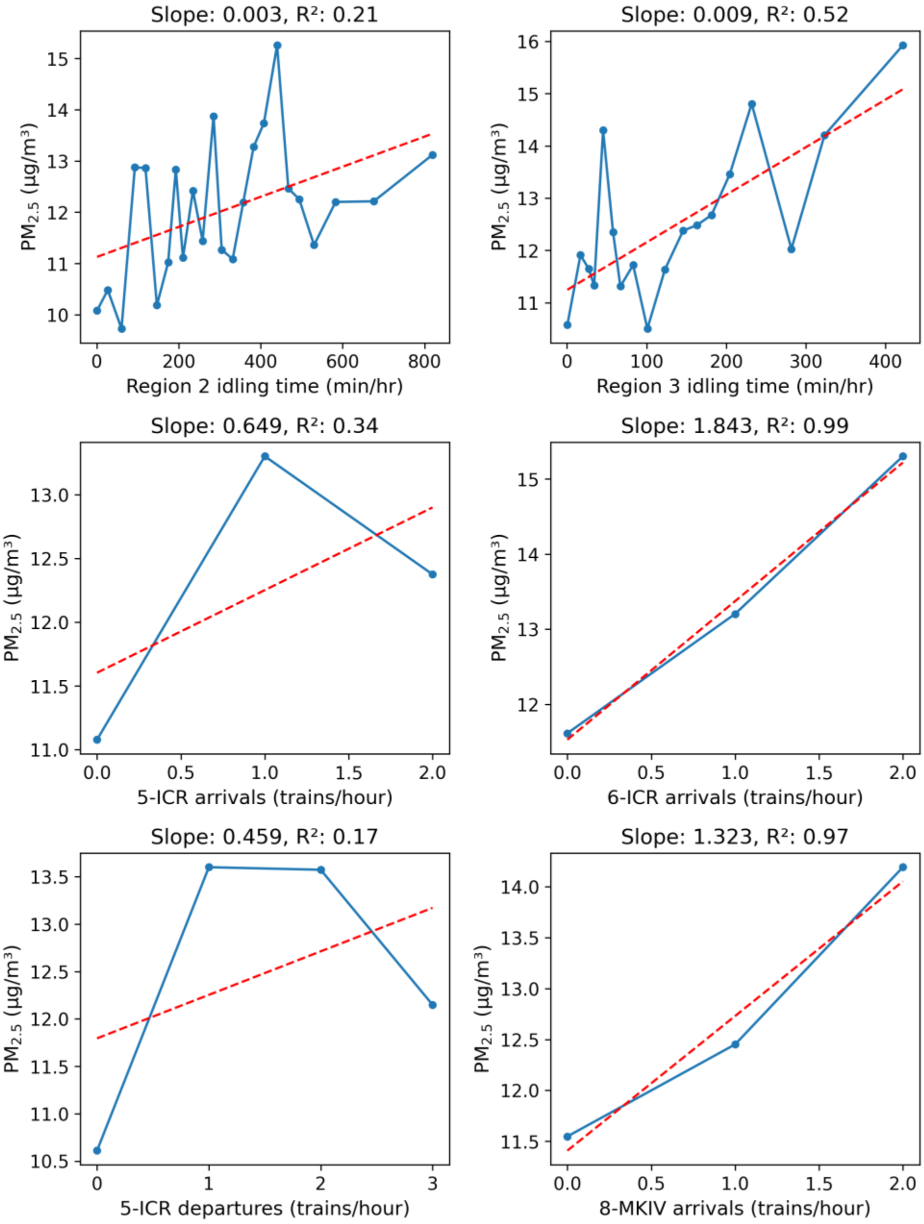


Figure 3.29 Variable importance for the Random Forest $PM_{2.5}$ model at Heuston platform

The partial dependence plots for $PM_{2.5}$ (Figure 3.30) show that platform concentrations increase almost linearly with background $PM_{2.5}$, with a slope of about 0.84 µg/m³ per 1 µg/m³ increase in the background, and a high fit to the partial trend ($R^2 \approx 0.93$). This confirms that regional air quality strongly controls the baseline pollution level inside the station. Idling time in Region 1 also has a clear positive association with platform $PM_{2.5}$ (slope ≈ 0.008 µg/m³ per minute, $R^2 \approx 0.87$). Although the slope is small, the cumulative idling time in Region 1 can reach several hundred carriage-minutes per hour, so the overall effect on platform concentrations is meaningful. Idling times in Regions 2 and 3 show weaker but still positive trends, consistent with their lower importance scores. Meteorological effects are mixed: higher temperatures and higher wind speeds tend to increase $PM_{2.5}$, while higher pressure and wind directions associated with better dispersion show negative slopes. The partial dependence curves for train counts indicate that increases in the hourly number of arriving ICR and MKIV sets are associated with modest increases in $PM_{2.5}$, particularly for ICR arrivals and 8-car MKIV movements, which show near-linear and steep positive trends.


Heuston station – platform PM2.5 RF partial dependence plots – page 1
Slope: 0.837, R²: 0.93
Background PM2.5 (µg/m³)
PM2.5 (µg/m³)
Slope: 0.008, R²: 0.87
Region 1 idling time (min/hr)
PM2.5 (µg/m³)
Slope: -0.182, R²: 0.21
Pressure (hPa)
PM2.5 (µg/m³)
Slope: -0.021, R²: 0.74
Wind direction (°)
PM2.5 (µg/m³)
Slope: 0.376, R²: 0.50
Temperature (°C)
PM2.5 (µg/m³)
Slope: 2.218, R²: 0.59
Wind speed (m/s)
PM2.5 (µg/m³)

Heuston station – platform $PM_{2.5}$ RF partial dependence plots – page 2



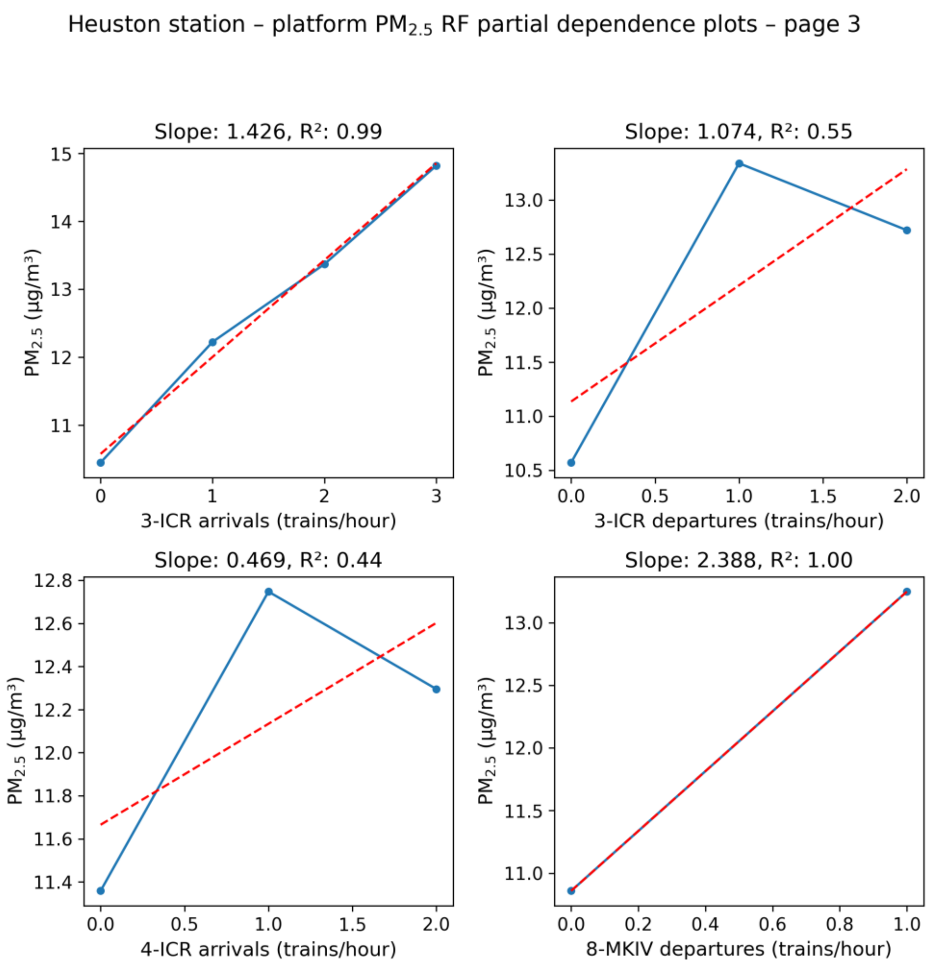


Figure 3.30 Partial dependence of hourly $PM_{2.5}$ at Heuston platform on key predictors in the Random Forest model

The PM_{10} model performs slightly less well than the $PM_{2.5}$ model but still explains a substantial part of the variance, with an R^2 of 0.54 and an RMSE of 5.66 µg/m³. The variable importance plot (Figure 3.31) again identifies background PM_{10} as the main driver, with an importance of about 0.39. Temperature and relative humidity are the next most influential predictors, followed by wind speed, pressure and idling time in Region 1. Wind direction and idling times in Regions 2 and 3 have smaller but non-negligible contributions. Train arrivals and departures of different ICR and MKIV sets show low importance, with individual importance values below 0.01, which indicates that their impact on hourly PM_{10} is more modest than that of background conditions and idling.

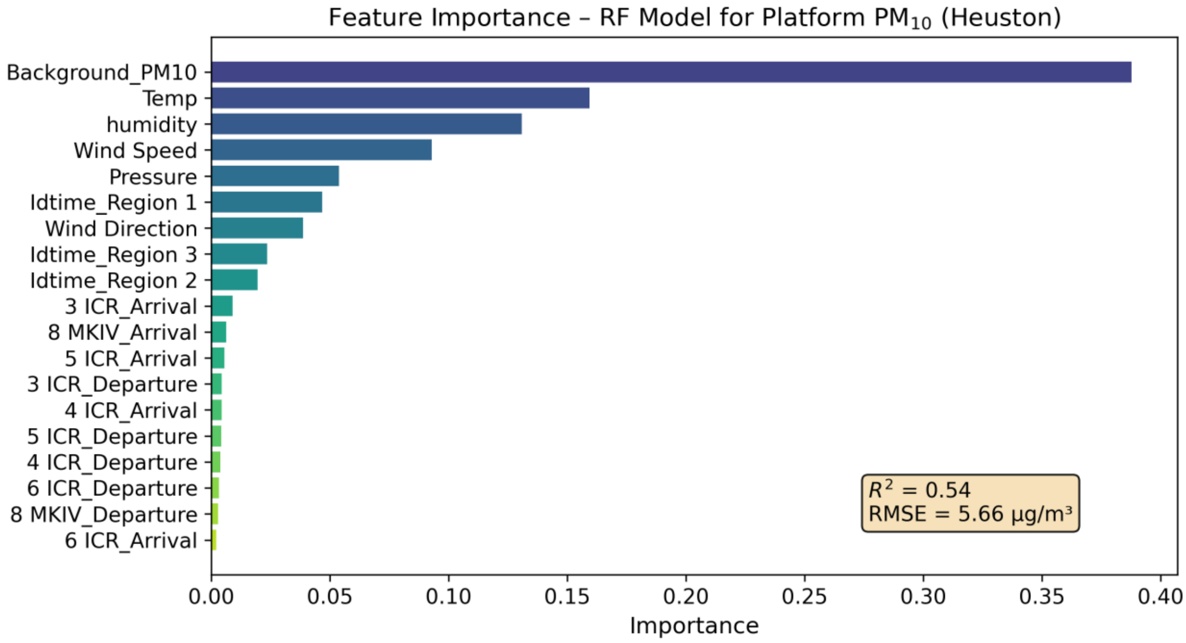
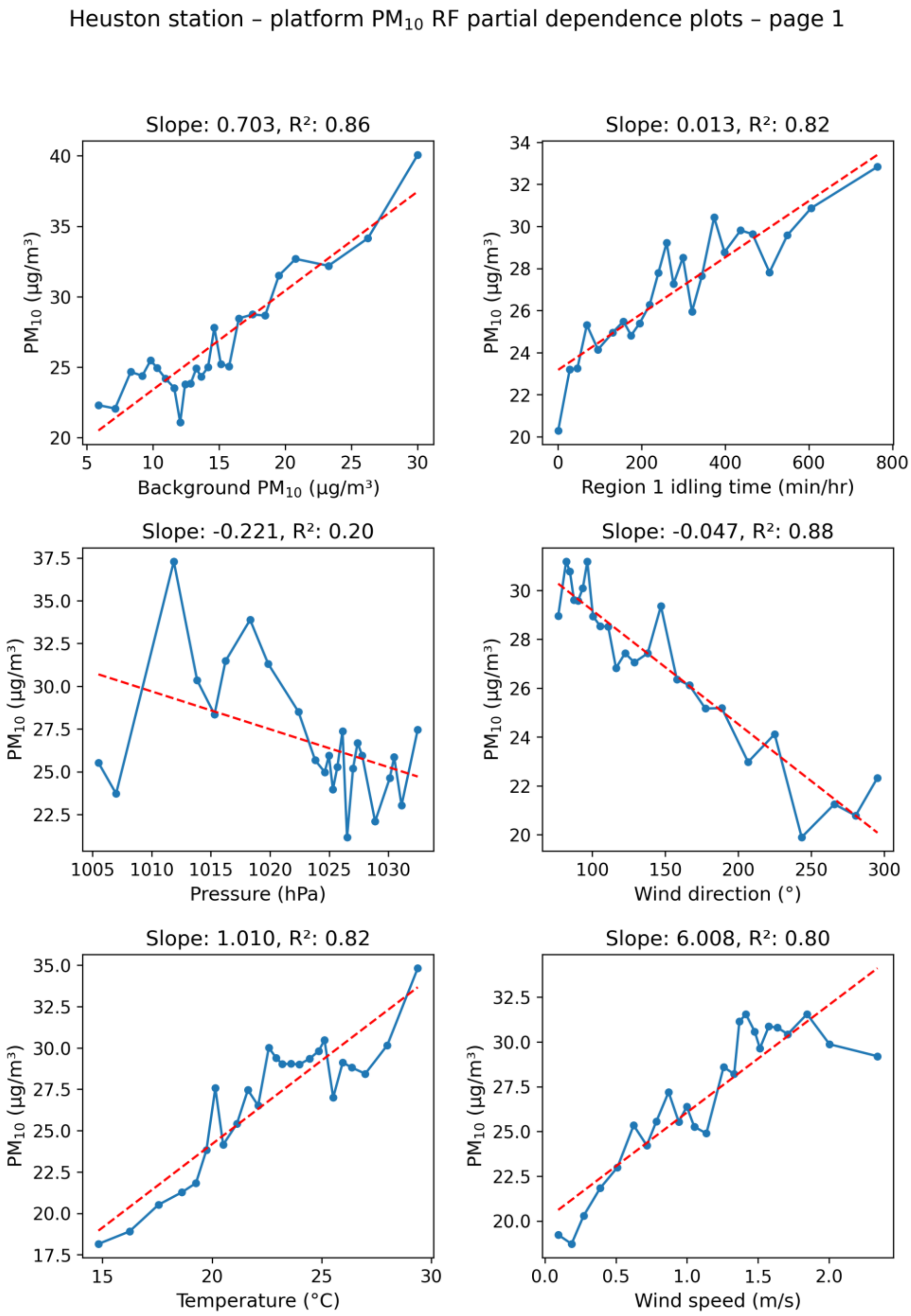
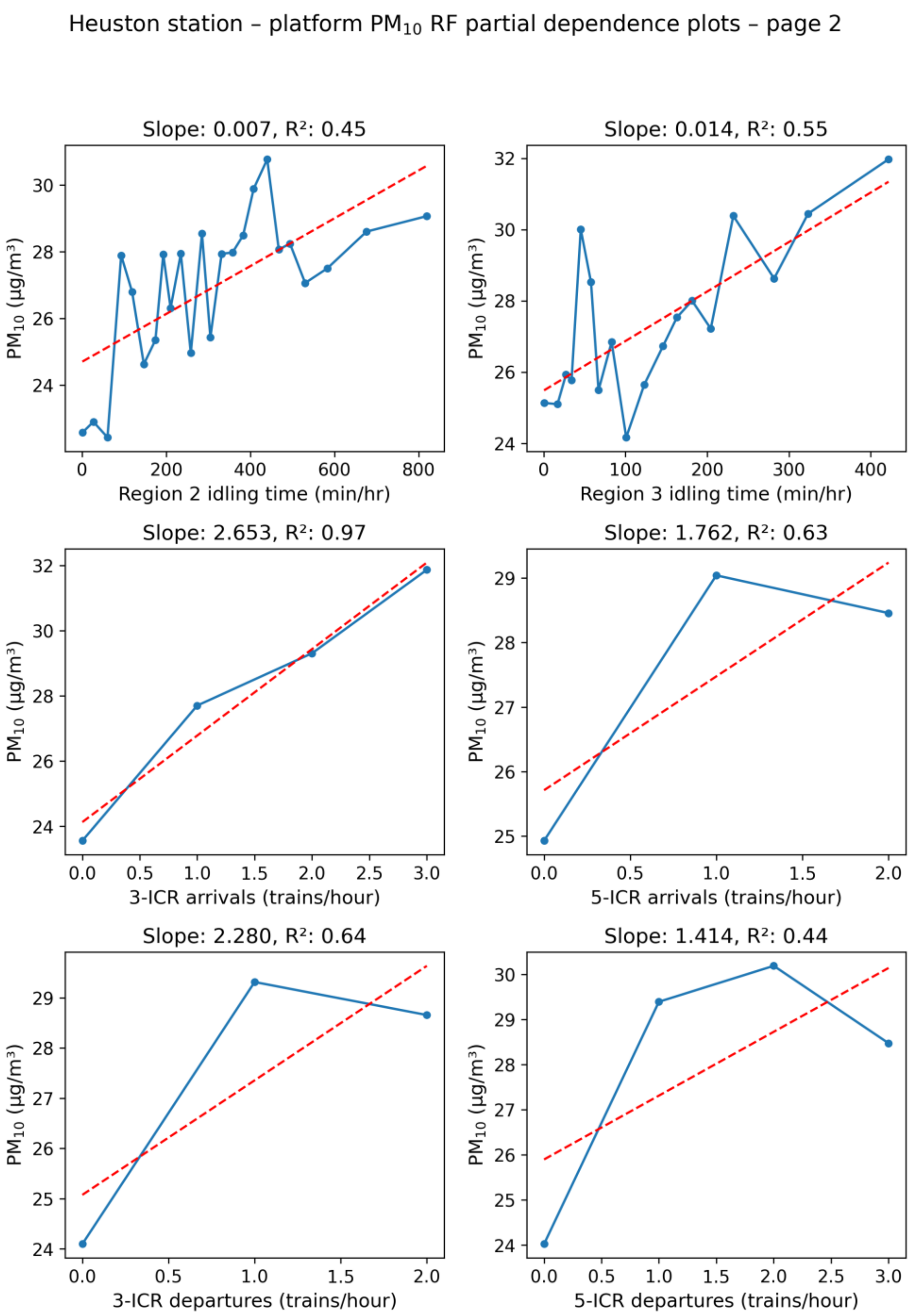


Figure 3.31 Variable importance for the Random Forest PM_{10} model at Heuston platform

The partial dependence plots for PM_{10} (Figure 3.32) show patterns that are broadly consistent with the $PM_{2.5}$ model. Platform PM_{10} rises with background PM_{10} (slope ≈ 0.70 µg/m³ per 1 µg/m³ increase, $R^2 \approx 0.86$) and with idling time in Region 1 (slope ≈ 0.013 µg/m³ per minute, $R^2 \approx 0.82$). Higher temperatures and wind speeds are associated with higher PM_{10}, while higher pressure and certain wind directions are linked to lower concentrations, reflecting the influence of mixing and advection. Idling times in Regions 2 and 3 also show positive slopes, though with more variability. Partial dependence on train counts suggests that additional ICR arrivals and departures can raise PM_{10} by a few µg/m³, especially for 3-, 5- and 6-car sets, but these effects remain secondary compared with background and idling.


Heuston station – platform PM_{10} RF partial dependence plots – page 1
Slope: 0.703, R^2: 0.86
PM_{10} (µg/m³)
Background PM_{10} (µg/m³)
Slope: 0.013, R^2: 0.82
PM_{10} (µg/m³)
Region 1 idling time (min/hr)
Slope: -0.221, R^2: 0.20
PM_{10} (µg/m³)
Pressure (hPa)
Slope: -0.047, R^2: 0.88
PM_{10} (µg/m³)
Wind direction (°)
Slope: 1.010, R^2: 0.82
PM_{10} (µg/m³)
Temperature (°C)
Slope: 6.008, R^2: 0.80
PM_{10} (µg/m³)
Wind speed (m/s)


Heuston station – platform PM10 RF partial dependence plots – page 2
Slope: 0.007, R²: 0.45
Region 2 idling time (min/hr)
Slope: 0.014, R²: 0.55
Region 3 idling time (min/hr)
Slope: 2.653, R²: 0.97
3-ICR arrivals (trains/hour)
Slope: 1.762, R²: 0.63
5-ICR arrivals (trains/hour)
Slope: 2.280, R²: 0.64
3-ICR departures (trains/hour)
Slope: 1.414, R²: 0.44
5-ICR departures (trains/hour)
PM10 (µg/m³)

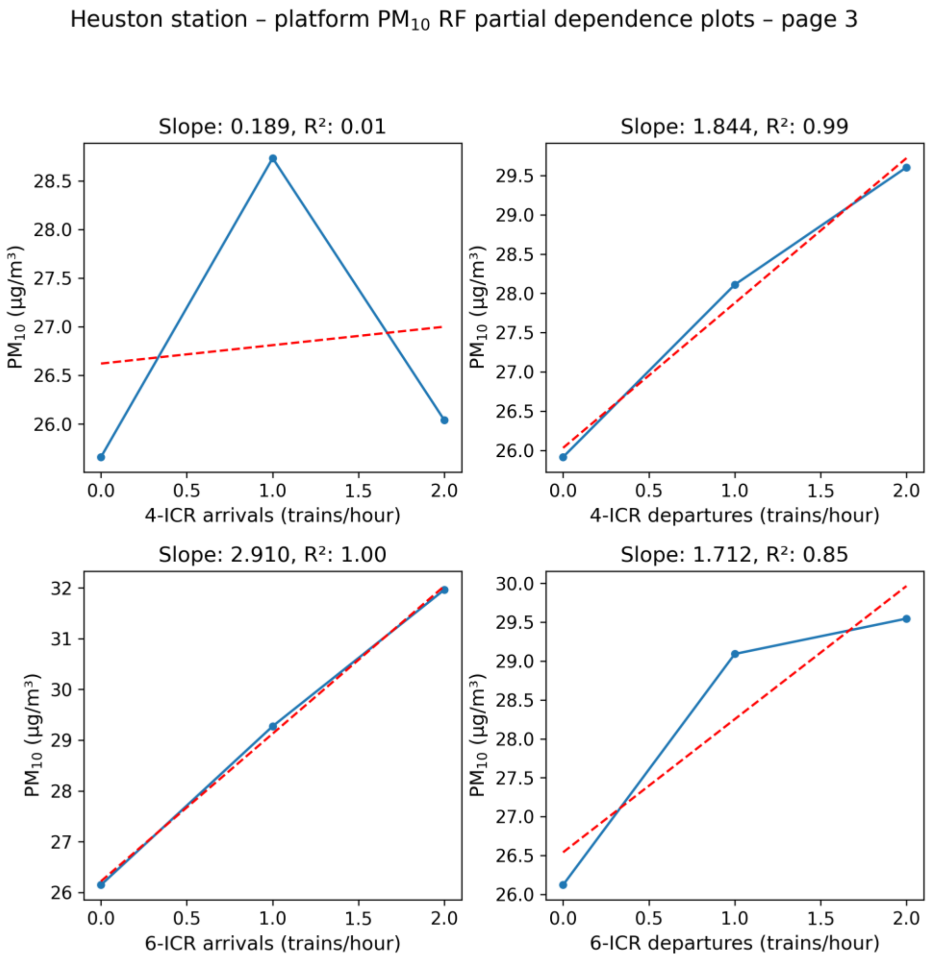


Figure 3.32 Variable importance for the Random Forest PM_{10} model at Heuston platform

Overall, the $PM_{2.5}$ and PM_{10} random forest models at Heuston show that regional background concentrations set the baseline air quality inside the station, while local operational factors, particularly idling in Region 1, increase concentrations on the platform. Meteorology also plays an important role, whereas the number of individual train movements has a more minor but still measurable influence.

3.3.5 Random Forest results for NO_2 at Connolly

For Connolly, Kruskal–Wallis tests were applied to hourly NO_2 at Platform 4 and the candidate predictors. Background NO_2, temperature, wind speed, wind direction, relative humidity, total trains, total arrivals, total departures, total passes, and idling times in Regions 1 and 3 all showed statistically significant differences in NO_2 between categories (Table 3.12). Idling time in region 2 is excluded from the model.

Table 3.12 Kruskal–Wallis test results for hourly NO_2 at Connolly platform

Variable	Chi-square	p-value	df	Significance
idling_region1_time_Category	619.30	0.0000	9	***
Total_Train_Category	264.14	0.0000	17	***
Total_Pass_Category	235.22	0.0000	13	***
temp_Category	122.22	0.0000	9	***
Total_Departure_Category	107.28	0.0000	7	***
Total_Arrival_Category	71.84	0.0000	8	***
windspeed_Category	65.39	0.0000	8	***
humidity_Category	58.30	0.0000	9	***
winddriction_Category	53.72	0.0000	9	***
idling_region3_time_Category	40.17	0.0000	5	***
no2_background_Category	28.77	0.0079	9	**
idling_region2_time_Category	14.60	0.0122	5	*

The Random Forest model for hourly NO_2 achieved an R^2 of 0.51 and an RMSE of 23.98 μg/m³ on the test data. Capturing around half of the variance in hourly NO_2 is reasonable given the strong short-term fluctuations driven by individual train events and ventilation conditions.

The variable-importance ranking (Figure 3.33) shows that idling time in Region 1 is by far the most influential predictor, with an importance score of about 0.39, roughly three times that of any other variable. This is consistent with the regression in Section 3.3.3.4 and Kruskal–Wallis results and with the conclusion that idling of trains in the enclosed terminus zone has a strong impact on NO_2 concentrations.

Background NO_2 is the second most important variable (importance ≈ 0.13), followed by temperature (≈ 0.11), wind speed (≈ 0.09), wind direction (≈ 0.09) and relative humidity (≈ 0.07). Among the train-activity metrics, total departures have the highest importance (≈ 0.05), while idling time in Region 3 and total numbers of trains, passes and arrivals have smaller contributions.

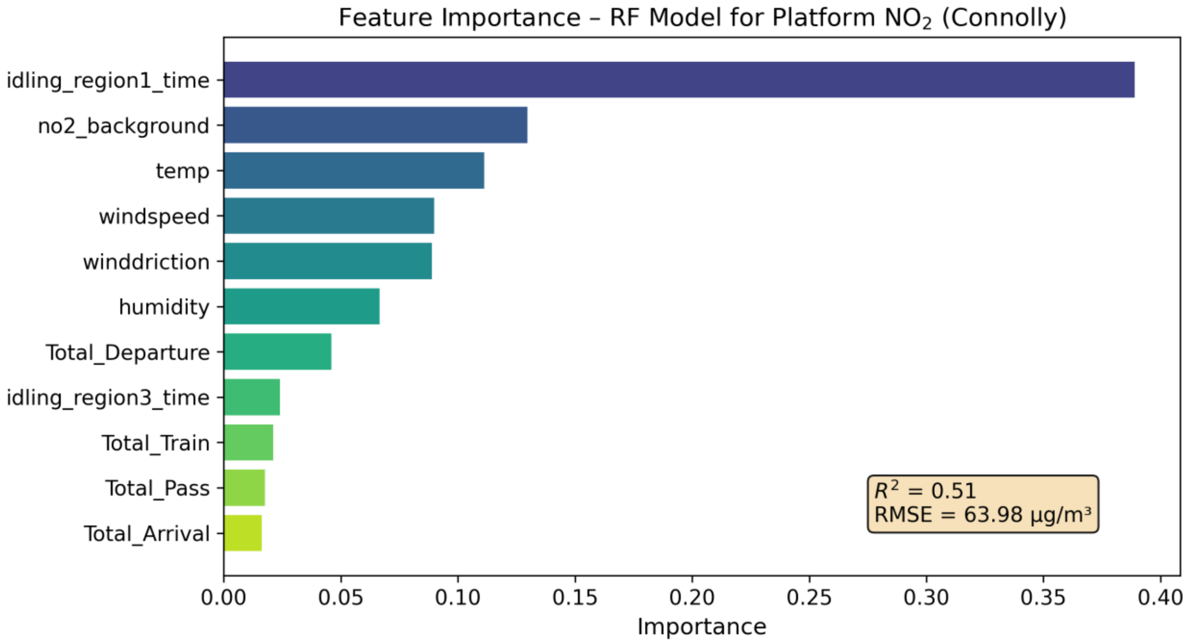
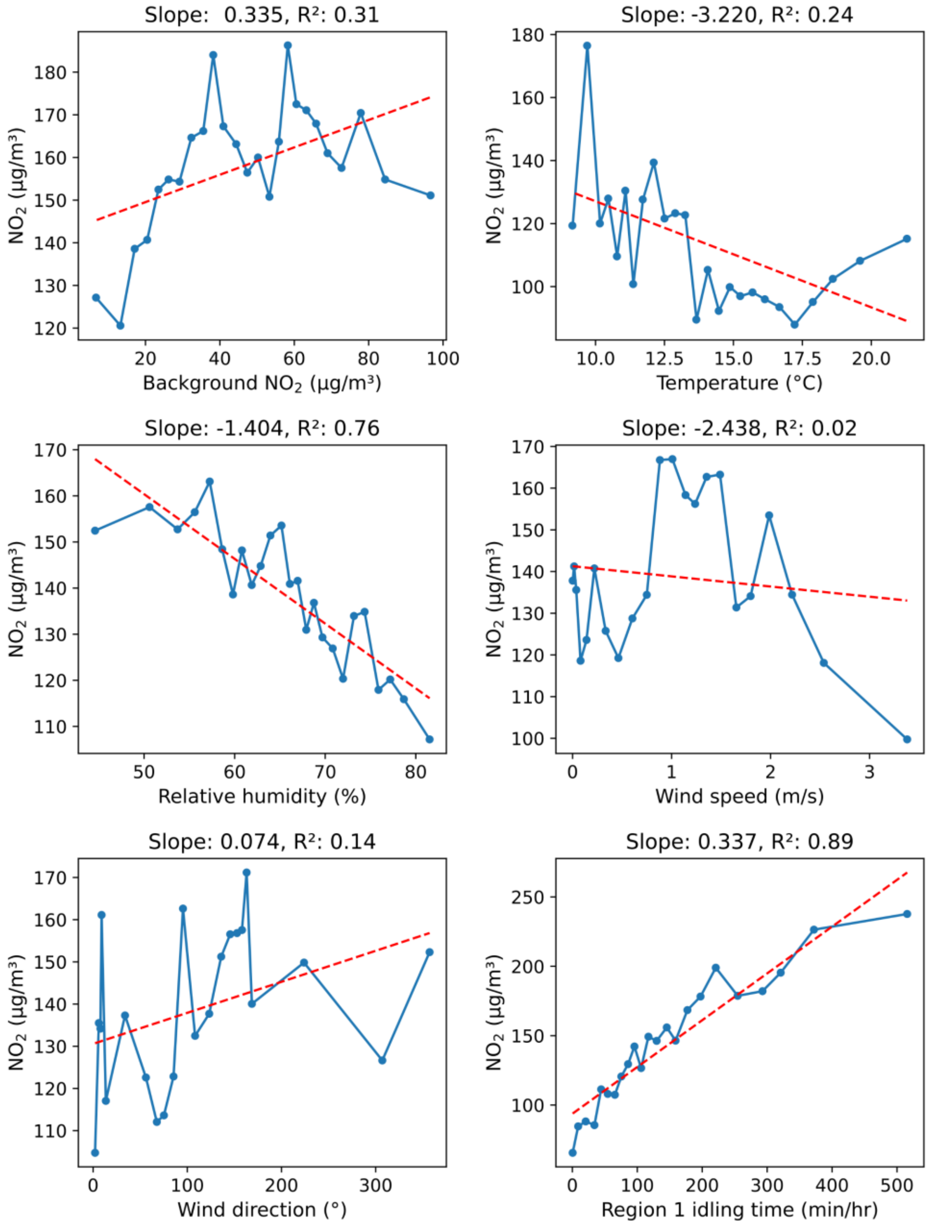


Figure 3.33 Variable importance for the Random Forest NO_2 model at Connolly (Platform 4)

Partial dependence plots (Figure 3.34) reveal several patterns. Relative humidity shows a strong negative association with NO_2 (slope ≈ −1.4 μg/m³ per 1% increase in humidity, R^2 ≈ 0.76), suggesting that humid conditions favour dilution or removal of NO_2 or co-occur with meteorological conditions that promote dispersion. The background concentration of NO_2 has a lower correlation with the station than PM (R^2 ≈ 0.24).

Connolly station – Platform 4 NO_2 RF partial dependence plots – page 1



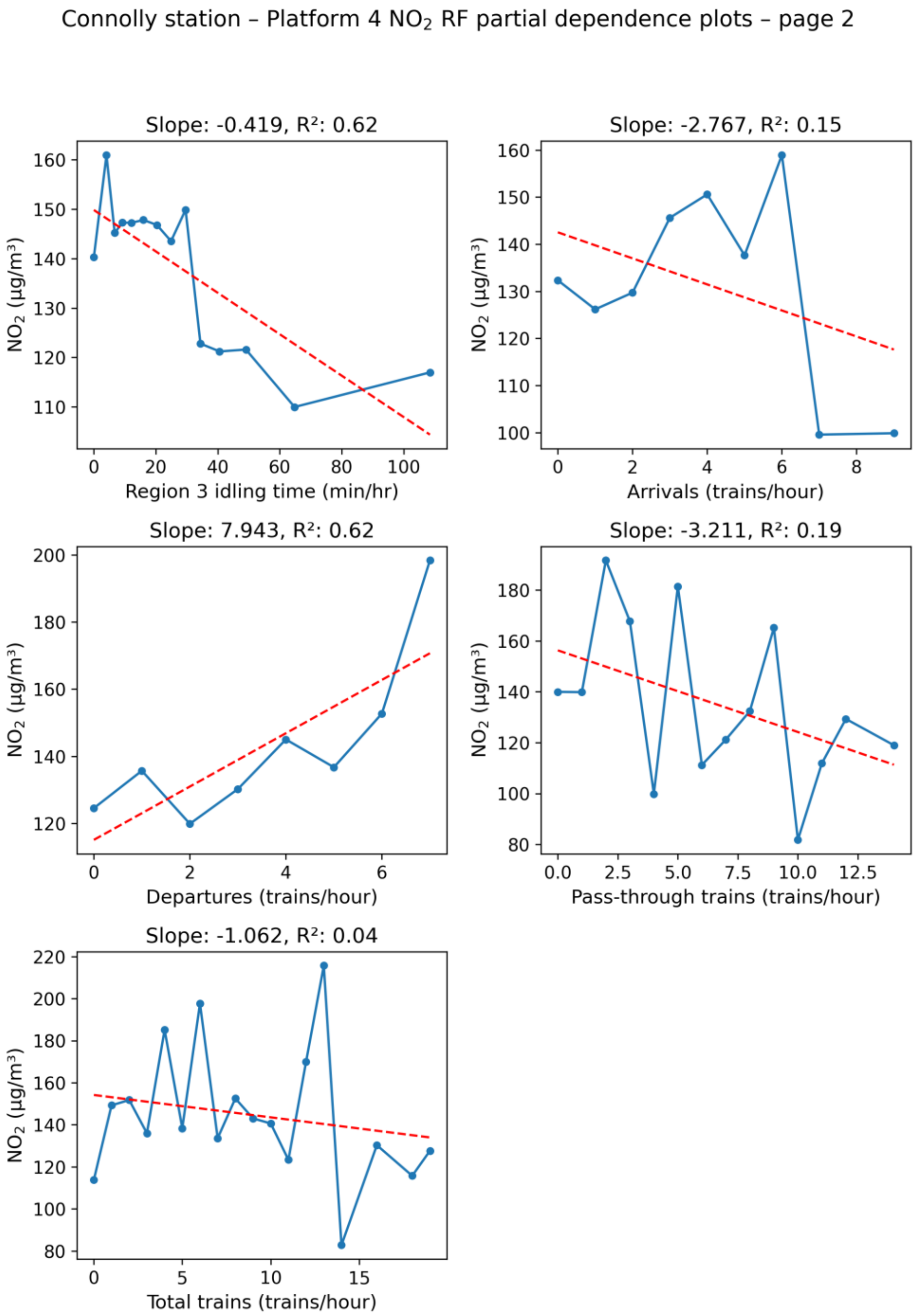


Figure 3.34 Partial dependence of hourly NO_2 at Connolly on key predictors in the Random Forest model

The strongest operational effect is the influence of idling in Region 1. The partial dependence curve is almost linear, with a slope of about 0.34 µg/m³ per minute and an R^2 of about 0.89. This means that an increase of 100 minutes of cumulative idling per hour in Region 1 is associated with an increase in NO_2 of around 34 µg/m³ at Platform 4. Total departures are positively associated with NO_2 (slope ≈ 7.9 µg/m³ per additional departing train, R^2 ≈ 0.62), consistent with the strong emissions produced during acceleration out of the station. In contrast, total arrivals and total passes show small negative slopes with low R^2, again indicating that, once background, idling and departures are accounted for, arrivals and passes carry little independent explanatory power and that the apparent negative slopes mainly reflect timing and correlation patterns rather than true mitigating effects.

3.4 Discussion

This chapter provides a detailed characterisation of air quality in two major Dublin railway stations and quantifies how concentrations respond to train operations and external conditions. The results demonstrate that enclosed or semi-enclosed platforms with intensive diesel operation are important exposure hotspots and that cumulative idling time in these spaces is a key operational driver, superimposed on a strong influence from regional background air quality.

Across both stations, background PM and NO_2 emerge as the dominant predictors in the Random Forest models, confirming that station air quality is anchored by the wider urban environment. Nevertheless, idling time in the most enclosed regions (Region 1 at both stations) consistently ranks as the most important local operational variable. At Heuston, partial dependence plots indicate that platform $PM_{2.5}$ and PM_{10} rise almost linearly with idling time under the roof, even after controlling for background and meteorology. At Connolly, an analogous pattern is observed for NO_2 in the enclosed terminus hall, with increments of several tens of µg/m³ associated with plausible changes in cumulative idling per hour.

Meteorological variables play a secondary but non-negligible role. Higher wind speeds and certain wind directions tend to reduce concentrations by enhancing ventilation and dispersion, while temperature, pressure and humidity modulate mixing and removal pathways. These influences partly explain the residual variance that remains unexplained even in the Random Forest models and highlight the importance of considering both operations and weather when interpreting station measurements or designing interventions.

The analyses also clarify the role of different types of train movements. Simple counts of arrivals and pass-through trains show weak or inconsistent relationships with platform concentrations, especially for NO2 at Connolly. This does not imply that these movements are harmless, but rather that, relative to idling and departures, their incremental contribution to hourly average concentrations is small and often masked by background and meteorological variability. In some regressions, arrivals and pass-through counts exhibit weak negative slopes; given the very low R^2

values, these should be interpreted as indicating a lack of clear positive association, rather than a protective effect. They likely reflect correlations with other variables (such as time-of-day or wind conditions) and the fact that arrival and coasting phases generally emit less than prolonged idling or acceleration.

Comparing Heuston and Connolly highlights the combined influence of enclosure and traction mix. Heuston is a fully diesel terminal with a semi-enclosed canopy and vents, while Connolly combines an enclosed diesel terminus with open through platforms served primarily by electric trains. At both stations, elevated $PM_{2.5}$ and PM_{10} and high $PM_{2.5}/PM_{10}$ ratios are found in the diesel-enclosed areas, whereas more open locations (Heuston platforms 6–8, Connolly Platform 6/7 and the concourse) show lower levels and lower $PM_{2.5}/PM_{10}$ ratios, closer to background. NO_2 at Connolly Platform 4 is particularly extreme, with all days exceeding the WHO 24-hour guideline and a substantial fraction of hours exceeding the 1-hour guideline. These patterns are consistent with previous work highlighting high exposures in enclosed or semi-enclosed stations with diesel traction and limited ventilation, but the present analysis adds a more explicit decomposition of background and operational contributions and quantifies the role of idling in different station regions.

There are several limitations to note. First, the monitoring campaigns provide detailed coverage for specific months rather than a full year. The levels and patterns reported here may therefore not fully capture seasonal extremes such as winter inversions or summer photochemical episodes. For example, PM and NO_2 levels could be higher during cold, stagnant conditions or lower during periods of enhanced ventilation and lower background. Second, while low-cost sensors offer high spatial and temporal resolution for PM, their NO_2 performance is more limited; this chapter mitigates that limitation by combining them with diffusion tubes and an MCERTS-grade NO_2 monitor, but some uncertainty remains, particularly for short-term NO_2 peaks. Third, the Random Forest models are empirical; they describe observed relationships but do not explicitly represent physical processes such as plume rise or complex flow patterns within station halls. The models therefore complement, but do not replace, the physics-based dispersion modelling developed later in the thesis.

3.5 Conclusion

This chapter combined detailed in-station monitoring with statistical analysis to characterise air pollution in two major Dublin rail stations and to identify the main drivers of pollutant variability. The main conclusions are:

1. Stations as exposure hotspots. Enclosed and semi-enclosed platforms at Heuston and Connolly experience substantially higher $PM_{2.5}$, PM_{10} and NO_2 levels than adjacent open platforms and concourse areas. At Heuston, 24-hour average $PM_{2.5}$ frequently exceeded the WHO guideline on both the platform and roof. At Connolly, $PM_{2.5}$ and PM_{10} were elevated on enclosed Platforms 2–4, and hourly NO_2 on Platform 4 often reached very high levels, with frequent exceedances of the WHO 1-hour guideline. These findings confirm that rail stations can act as important short-term exposure hotspots for passengers and staff.

2. Background vs local contributions. Random Forest models demonstrate that urban background concentrations are the dominant predictor of in-station PM and NO_2, establishing the baseline upon which station-specific contributions are superimposed. Local operational factors, particularly cumulative idling time in the most enclosed regions, add significant increments. Meteorology further shapes these relationships by controlling dispersion and mixing.

3. Role of idling and train movements. Cumulative idling time under the roof (Region 1) is consistently the most important local operational driver of platform concentrations at both stations. Departures from enclosed areas also contribute positively, especially for NO_2 at Connolly. In contrast, counts of arrivals and pass-through trains have weaker and sometimes negligible associations with concentrations, once background, idling and meteorology are taken into account. This indicates that reducing idling and managing departures in enclosed spaces are likely to be more effective mitigation strategies than simply reducing the total number of train movements.

4. Implications for station operation and design. The results suggest that measures such as minimising dwell time with engines running, relocating or shortening idling in the most enclosed zones, improving ventilation paths, and prioritising cleaner rolling stock on enclosed platforms could all contribute to lower pollutant levels in stations. Because background pollution remains a major driver, broader regional air-quality improvements will also directly benefit station environments.

These findings provide a quantitative basis for the operational analysis of idling behaviour in Chapter 4 and for the emissions and dispersion modelling developed in Chapters 5 and 6. Together, the chapters support a more integrated understanding of how station design, operations and wider urban air quality combine to shape exposure in rail environments and where interventions can yield the greatest benefit.

4. Operational drivers of diesel train idling in Dublin rail stations: timetable constraints, driver behaviour, and mitigation

4.1 Introduction

In Chapter 3, statistical models of station air quality identified train idling time as the most important operational factor associated with elevated $PM_{2.5}$ and NO_2 at the station, after accounting for background and meteorology. However, those models did not explain why idling varied, or which parts of a typical station cycle were responsible for the observed differences in exposure. Building on that evidence, this chapter examined the timing, duration and frequency of idling events for specific diesel multiple unit (DMU) classes. It then investigated how these idling patterns related to timetable structure (e.g., scheduled dwell and turnaround) and driver behaviour (e.g., power/notching choices during approach and layover). This operational analysis provided the link between the statistical relationships reported in Chapter 3 and the emissions modelling developed in the subsequent sections.

4.1.1 Regulatory definitions and operating rules

Several regulations and company rules seek to limit “long-duration idling” of diesel trains. The U.S. EPA instead highlights long periods of locomotive idling at yards, stations and sidings as an important source of fuel use and emissions, and promotes idling-reduction technologies and operating policies for railroads.(US EPA, 2016a). Locomotives fitted with automatic engine stop–start systems are required to shut down after no more than 30 minutes of continuous idling, except where idling is needed to prevent engine damage, maintain brake air pressure or recharge batteries (Transportation Environmental Resource Centre, 2020). Freight operators set idle time limits, often around 30 minutes. For example, CSX rules state that locomotives may not idle for more than 30 minutes in Massachusetts above 7 °C, except for specific operational reasons such as crew changes or repairs (CSX Transportation, 2010).

In the UK, the Rail Delivery Group Guidance Note: Reducing Diesel Emissions in Stations and Depots asks operators of diesel traction to prepare a Diesel Engine Idling Reduction Plan (Rail Delivery Group, 2021). These plans should identify sensitive locations (such as enclosed stations, depots and residential-area sidings), set maximum idling times for those locations, and combine operational, behavioural and technical measures to reduce unnecessary idling (Rail Delivery Group, 2021).

At the EU level, no harmonised maximum idling time rule for rail traction was identified; instead, EU policy primarily addresses diesel rail emissions via engine emission limit values and type-approval requirements for non-road mobile machinery engines (Stage V), which cover locomotive and railcar engine categories ((EU) No 167/2013, and Amending and Repealing Directive 97/68/EC, 2016).

For passenger rail operations, idling reduction cannot be considered only as an emissions issue. Engine shutdown may reduce fuel use and local emissions, but unsuccessful or delayed restart could create high operational risks, including delayed departures, platform occupation, missed paths and wider timetable disruption. Drivers may therefore be reluctant to shut down engines during short or uncertain layovers, especially where restart reliability, brake air pressure, battery condition, auxiliary power demand or passenger comfort are operational concerns.

In practice, actual idling times reflect written rules, timetable constraints (TC), rolling stock characteristics, local habits, and concerns about comfort and reliability.

4.1.2 Emission characteristics of idling and transient operation

Experimental studies show that emissions during and immediately after idling can be substantial. Weaver (2006) measured start-up and idling emissions for two locomotives and found that, for modern engines, shutting down and restarting the engine produces lower total emissions than remaining at low idle, provided the shutdown period is longer than about eight minutes; longer shutdowns give correspondingly larger emission savings.

Subsequent work has examined how transient operating modes influence energy use and emissions. Lebedevas et al. (2015) describe how short, frequent idling episodes and transitions between idle and load affect diesel engine fuel consumption and exhaust characteristics. Rakopoulos and Giakoumis (2006, 2009) review the unsteady operation of turbocharged diesel engines and link post-idle transients to poorer driveability and increases in particulate and gaseous emissions. High-frequency measurements of diesel particle number emissions further demonstrate that these transients can generate sharp, short-lived emission peaks that would be underestimated by steady-state test cycles (Kamarianakis et al., 2011).

Overall, these studies indicate that idling and the transitions into and out of idle can make an important contribution to both average and peak emissions, and that controlling idle time can help improve local air quality.

4.1.3 Technical and behavioural solutions to reduce idling

A wide range of technical and organisational measures has been proposed to reduce idling in diesel engines. For rail, Brecher et al. (2014) discuss locomotive energy-efficiency measures including automatic engine stop–start systems, auxiliary power units (APUs) and shore-power connections. Similar technologies have also been evaluated in road-vehicle studies, including APUs, direct-fired heaters, battery-based cab-comfort systems, electrified parking and automatic or soft start-stop systems (Rahman et al., 2013; Shancita et al., 2014). Automatic or soft start-stop systems shut down the engine during suitable idle periods and restart it when power is required, reducing fuel use and exhaust emissions. However, their use depends on reliable restart performance and sufficient auxiliary power to maintain onboard systems (Hu et al., 2022; Karrouchi et al., 2023). Similar principles could be applied to diesel rail vehicles, but stronger safeguards would be needed because restart delay or failure could affect brake air pressure, battery charge, passenger comfort, platform occupation and timetable reliability. Overall, direct-fired heaters, APUs and electrified parking or shore-power connections are generally identified as the most effective idle-reduction measures for reducing fuel use and exhaust emissions.

Behavioural strategies also play a role. The US Environmental Protection Agency (US EPA, 2016b) emphasises training staff to avoid unnecessary idling and setting clear time limits when the main engine is not required. EPA has also provided financial support for idle reduction technologies; in 2016, it funded up to 40% of the cost of eligible, verified locomotive idle reduction technologies (United States Environmental Protection Agency, 2016). At the EU level, there is no equivalent, harmonised programme targeting rail idling behaviour specifically.

Related work in behavioural science shows that simple prompts and feedback can reduce idling in road vehicles. Abrams et al. (2021), for example, used psychological interventions to reduce unnecessary car idling. Comparable behavioural approaches have not yet been widely applied or assessed in railway operations, where drivers work under tighter safety and TC.

4.1.4 Evidence on locomotive idling patterns

Most research on railway idling concerns freight and shunting locomotives. Several studies report average idle times during shunting, or the share of total engine operating time spent at idle based on annual averages (Kortas & Kropiwnicki, 2013; Skoglund, 2011; Souten, 2008).

A series of measurements in Polish industrial sidings between 2009 and 2013 provides detailed power time series for diesel shunting locomotives (Kałuża, 2016). These data show that diesel engines can spend 55–90% of daily operating time at idle, with average idle fuel consumption around 160 L per day. This suggests a large operational and economic impact of idling in shunting operations(Kałuża, 2016).

Further analysis of the same data shows that idle time distributions for shunting locomotives are heavy tailed (Kałuża & Kucharski, 2018): long idling events occur much more often than simple models would predict. The study, however, focuses mainly on descriptive statistics and does not explore specific operational changes. More broadly, there is very limited work on individual locomotive idle periods and their sampling distributions, especially for passenger trains in stations rather than freight locomotives in yards.

4.1.5 Research gaps and motivation for this chapter

Current research on railway idling is still limited. Existing work shows that locomotive engines spend a large fraction of their time idling, that emissions during and after idling are harmful, and that idle time distributions can be heavy tailed. Technical and behavioural measures are available, but there is little quantitative evidence on how specific operational practices, such as scheduling and driver habits, relate to observed idling patterns. Behavioural interventions that have reduced idling in road transport have not yet been systematically adapted to rail.

For urban rail systems and passenger trains in particular, few studies link detailed operational data with air quality concerns. In the Dublin case, modelling has identified idling as the key traffic driver of station air quality. However, there is no systematic analysis of how long trains idle in the station area, how this varies by train model and by timetable gap, or how closely current operating rules are followed.

This chapter addresses these gaps using detailed train movement records from Dublin's main stations. First, it identifies individual idling events for specific diesel train models and describes their duration and frequency within the station area. Second, it uses Bayesian models to examine how driver behaviour and TCs are associated with these idling durations, with particular attention to the scheduled gap between consecutive services. By combining operational data, statistical modelling, and existing evidence on emissions and behaviour, the chapter aims to support practical idling reduction strategies in the Irish rail context and in similar urban rail systems.

4.2 Data and definition

For this study, Irish Rail provided a dataset of Class 22000 DMU operation records covering four months (25 May to 26 June 2023 at Heuston;1 March to 31 May 2024 at Connolly). The dataset includes train speed, engine rpm, and GPS information. Using the GPS data, idling events occurring within Heuston Station were identified as periods when train speed was zero while engine rpm remained greater than 0, and the absolute duration of each continuous idling period was calculated.

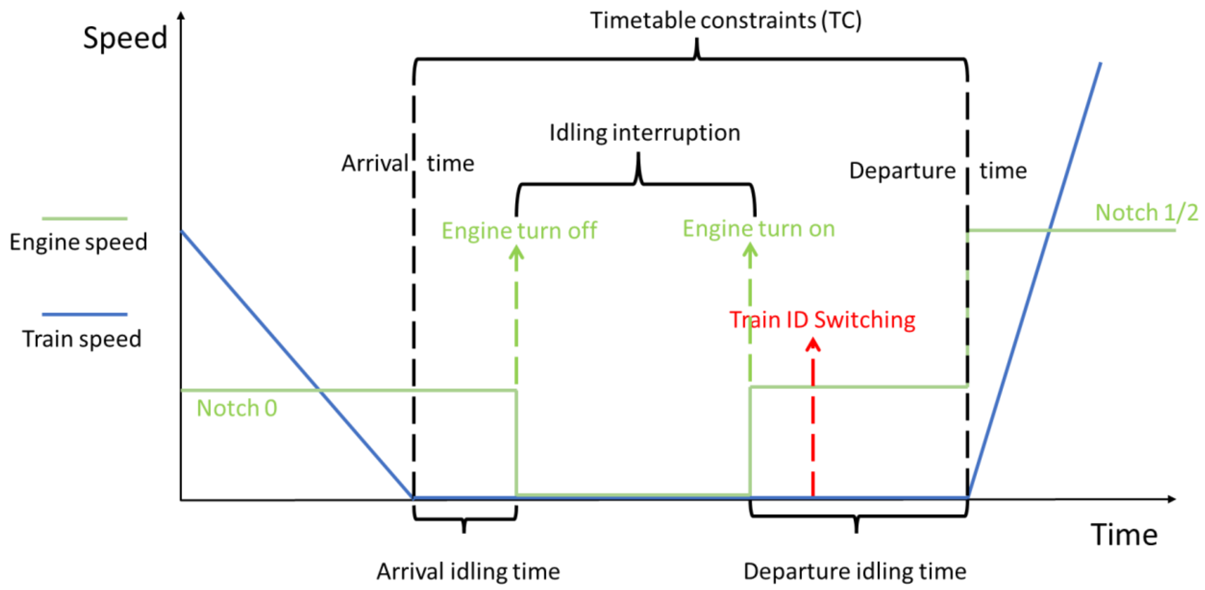


Figure 4.1 Illustration of the idling definition

In this research, idle time refers to the duration when the train's speed is zero while the engine is still running (engine rpm > 0). After a train arrives at the station, the duration it remains stopped at the platform with the engine running is defined as arrival idle time. The duration when the engine is started in advance before departure is defined as departure idle time. It should be noted that not all DMUs shut down and restart their engines at the station; in such cases, arrival and departure idling times are not separated. Figure 4.1 shows the engine speed and train speed over time for each phase.

The TC between each DMU's arrival and departure are restricted by the train timetable set by the Irish Rail operations department. The first and last services of each DMU at this station are not under scheduling pressure, so in this thesis, they are considered as driver behaviour information without TC influence.

Based on the speed profile before and after each event, idle periods were classified as arrival idling (preceded by deceleration), departure idling (followed by acceleration), or continuous idling without engine shutdown and restart. Each idling period was then matched to its corresponding time and DMU number to calculate the total idle time for each DMU's entry and exit at the station.

Notably, if a DMU remains running throughout its station dwell, its total idle time is equal to the TC. If the engine is shut down and restarted, the total idle time equals the sum of arrival and departure idling, which is shorter than the TC. As the monthly timetable is consistent, the dataset is considered representative of average operations.

To provide a clearer operational example, Figure 4.2 and Figure 4.3 show representative time-series windows for one train movement at Heuston and one at Connolly. Each panel includes wheel-derived speed, GPS-derived speed, engine RPM, power notch and the local line gradient. These variables illustrate how arrival, dwell/idling and departure phases were identified from the OTMR data. Periods with zero speed and non-zero engine RPM were classified as idling, while changes in speed and notch indicate approach and departure behaviour. The gradient information is included because it affects traction demand during station entry and exit.

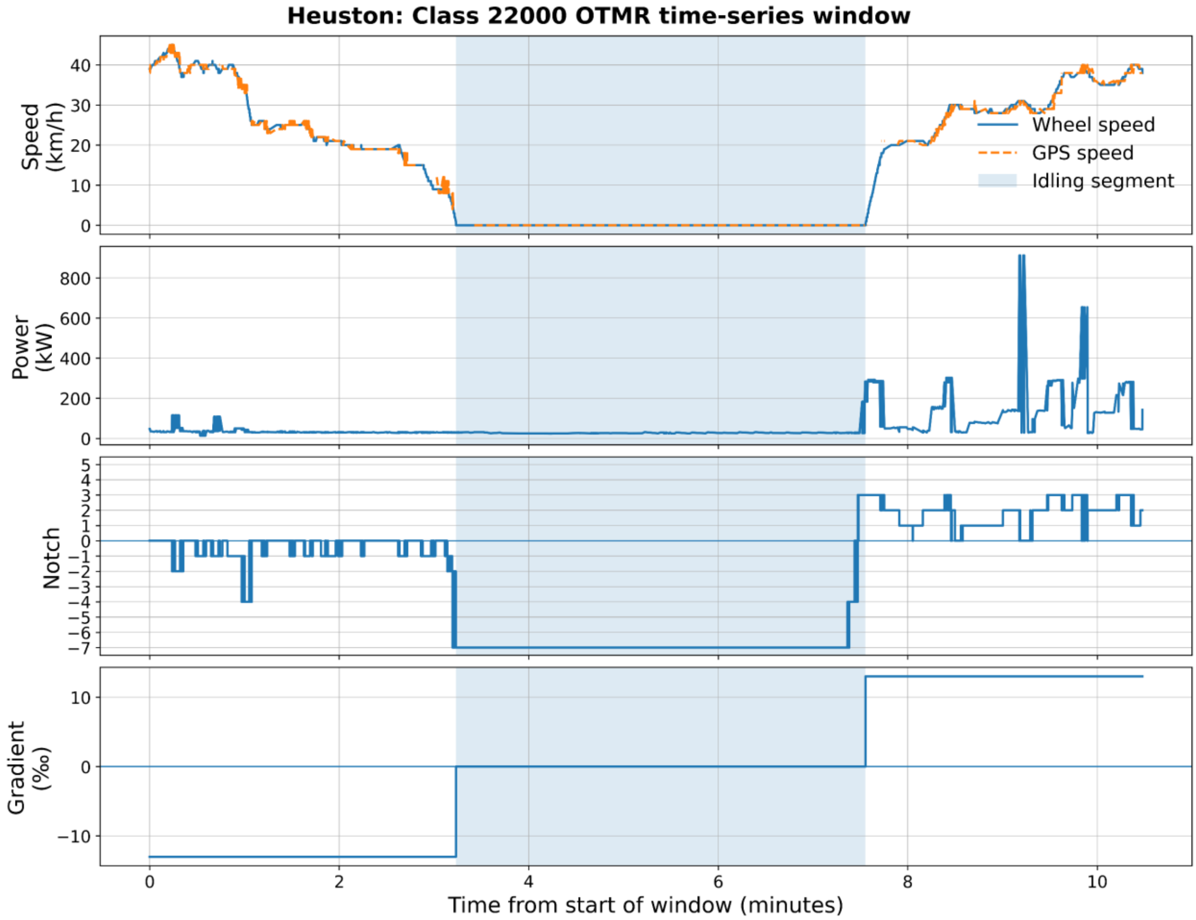


Figure 4.2 Heuston Class 22000 OTMR time-series window showing wheel speed, GPS speed, engine power, notch position (Positive notch values indicate power notches, while negative

notch values indicate braking) and track gradient.

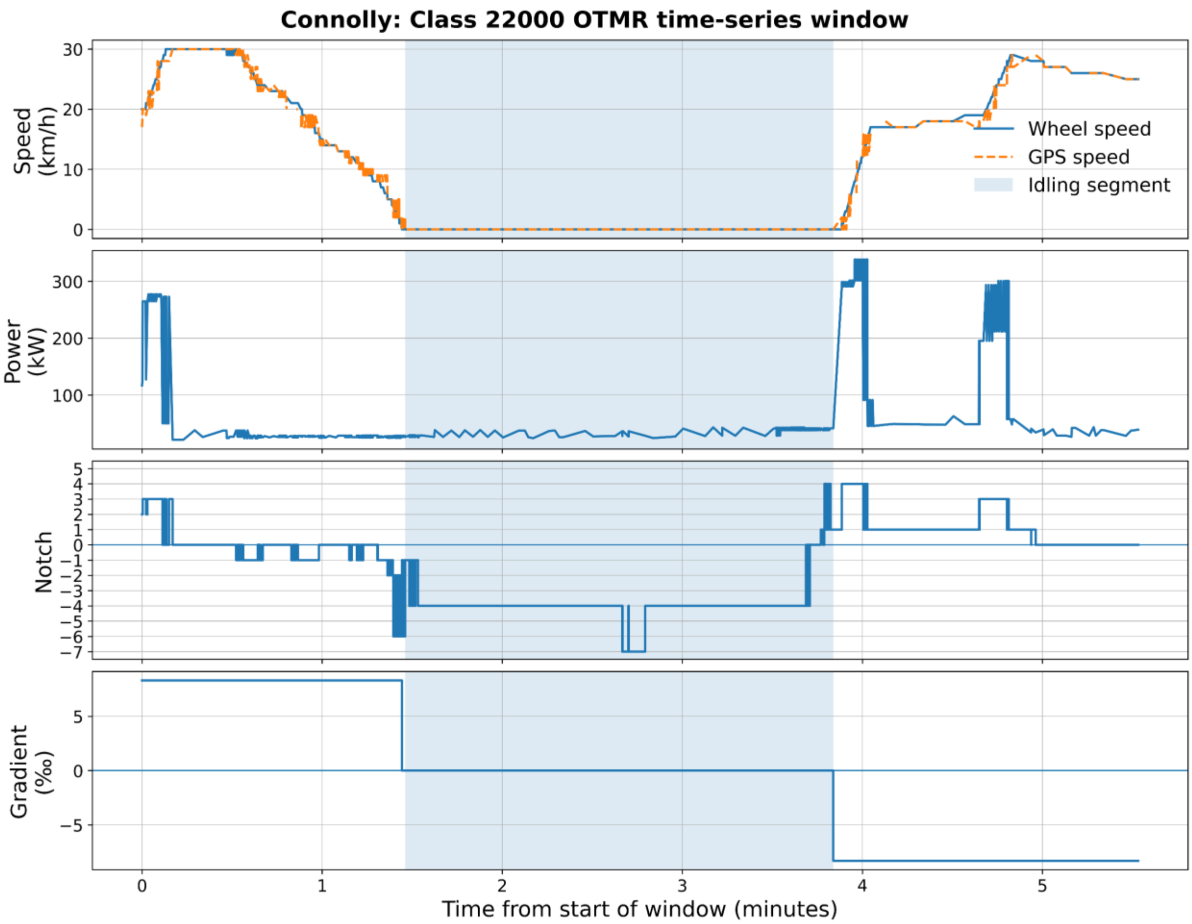


Figure 4.3 Connolly Class 22000 OTMR time-series window showing wheel speed, GPS speed, engine power, notch position (Positive notch values indicate power notches, while negative notch

4.3 Distribution of idling time

For the first part of the analysis, each continuous idling period identified in the OTMR logs for Class

22000 DMUs was treated as a separate observation, regardless of whether the same set idled again later in the day. The OTMR dataset provided second-by-second records of speed, power notch and other operating variables for routine services at Connolly and Heuston. Within these time series, an idling period was defined as any uninterrupted interval during which the train was stationary while the engine remained running. Each Idling_ID therefore represents one such continuous idling episode.

For each idling period, the Train ID and train type were taken from the movement that accounted for the longest share of the idle period. Train types were grouped into two broad categories using the Irish Rail alpha-code system. Codes A, B, C, D, J and P were treated as passenger-related services: A refers to scheduled mainline passenger trains, B to special mainline passenger trains, C to empty diesel suburban services, D to diesel suburban services to outer destinations, J to empty mainline passenger trains, and P to diesel suburban services to inner destinations. These were grouped as passenger services because they represent either passenger-carrying trains or empty passenger stock movements. All other codes were treated as non-passenger or auxiliary movements, such as shunting, radio-test trains, freight, departmental trains, or unusual movements. The full code classification is provided in Appendix 2.

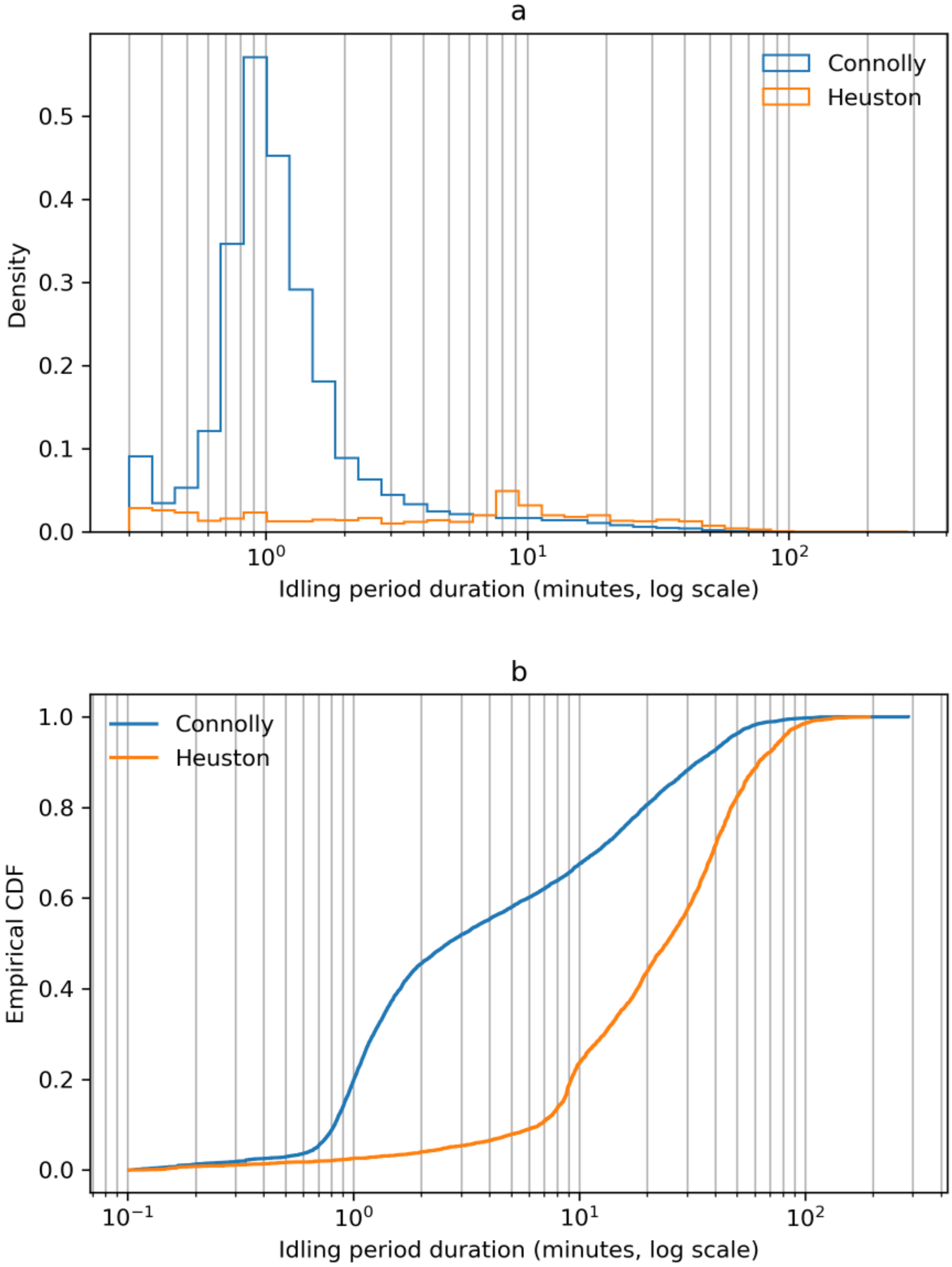


Figure 4.4 Distributions of idling duration for class 22000 DMUs at Dublin Connolly and Heuston

(a) Overlaid histogram of individual idling periods by station, plotted on a logarithmic time axis to show both short and long events. (b) Empirical cumulative distribution functions (ECDFs) of idling duration by station, also on a log scale.

Figure 4.4 summarises the overall distribution of idling duration for class 22000 DMUs at Dublin Connolly and Heuston. Panel (a) shows an overlaid histogram of idling time by station, plotted on a logarithmic x-axis to accommodate the strong right-skewness and wide range of durations. Panel (b) shows the corresponding empirical cumulative distribution functions (ECDFs), again on a log-scaled time axis. Together, these plots highlight that most idling periods are short, but a non-negligible fraction extends to tens or even hundreds of minutes. Across both stations, there are 12431 idling periods in total, of which 8188 occur at Connolly and 4243 at Heuston. Descriptive statistics for each station are reported in Table 4.1.

Table 4.1 Idling time statistics by station (all 22000-series DMUs)

Station	N	Mean	Median	P10	P25	P75	P90	Max	%>15 mins	%>30 mins	%>60 mins
Connolly	8188	11.0	2.6	0.8	1.1	15.3	33.4	286.5	25.5	11.8	1.8
Heuston	4243	30.2	24.3	6.7	10.6	42.9	62.3	192.6	65.5	42.7	11.4

The ECDF in Figure 4.4b makes the differences between the two stations particularly clear. At Connolly, idling periods are concentrated at very short durations: the median idle period is 2.6 minutes, the mean is 11.0 minutes, and the 90th percentile is 33.4 minutes. Only about 25.5% of periods exceed 15 minutes, 11.8% exceed 30 minutes, and 1.8% exceed 60 minutes. At Heuston, idling periods are substantially longer on average. The median idle period is 24.3 minutes, the mean is 30.2 minutes, and the 90th percentile is 62.3 minutes. Approximately 65.5% of periods exceed 15 minutes, 42.7% exceed 30 minutes, and 11.4% exceed one hour. These heavy-tailed distributions, with many short events and a small number of very long ones, are consistent with previous work on rail idling, but here they are documented for passenger DMUs in large urban stations rather than shunting locomotives in industrial sidings (Kałuża, 2016; Kałuża & Kucharski, 2018).

The contrast between Connolly and Heuston reflects their different operational roles. Connolly functions as a mixed pass-and-terminus station. The log-histogram in Figure 4.4a shows a strong peak at durations around one to three minutes, corresponding to brief passenger stops or short dwell times before departure. Longer idle periods are present but relatively rare. Heuston, in contrast, operates as a total terminal for the services studied. Trains typically arrive, unload and then wait for their next working at or near the same platform, resulting in systematically longer layovers. This is visible in Figure 4.4a as a higher density of periods in the 10–60 minute range and in Figure 4.4b as a noticeably slower rise of the ECDF at shorter durations. The longest single idle event at Connolly reaches 286.5 minutes and is associated with a non-passenger movement, whereas the maximum at Heuston is 192.6 minutes. Although such extreme events are rare, they contribute a substantial share of total idle hours and thus of fuel use and emissions.

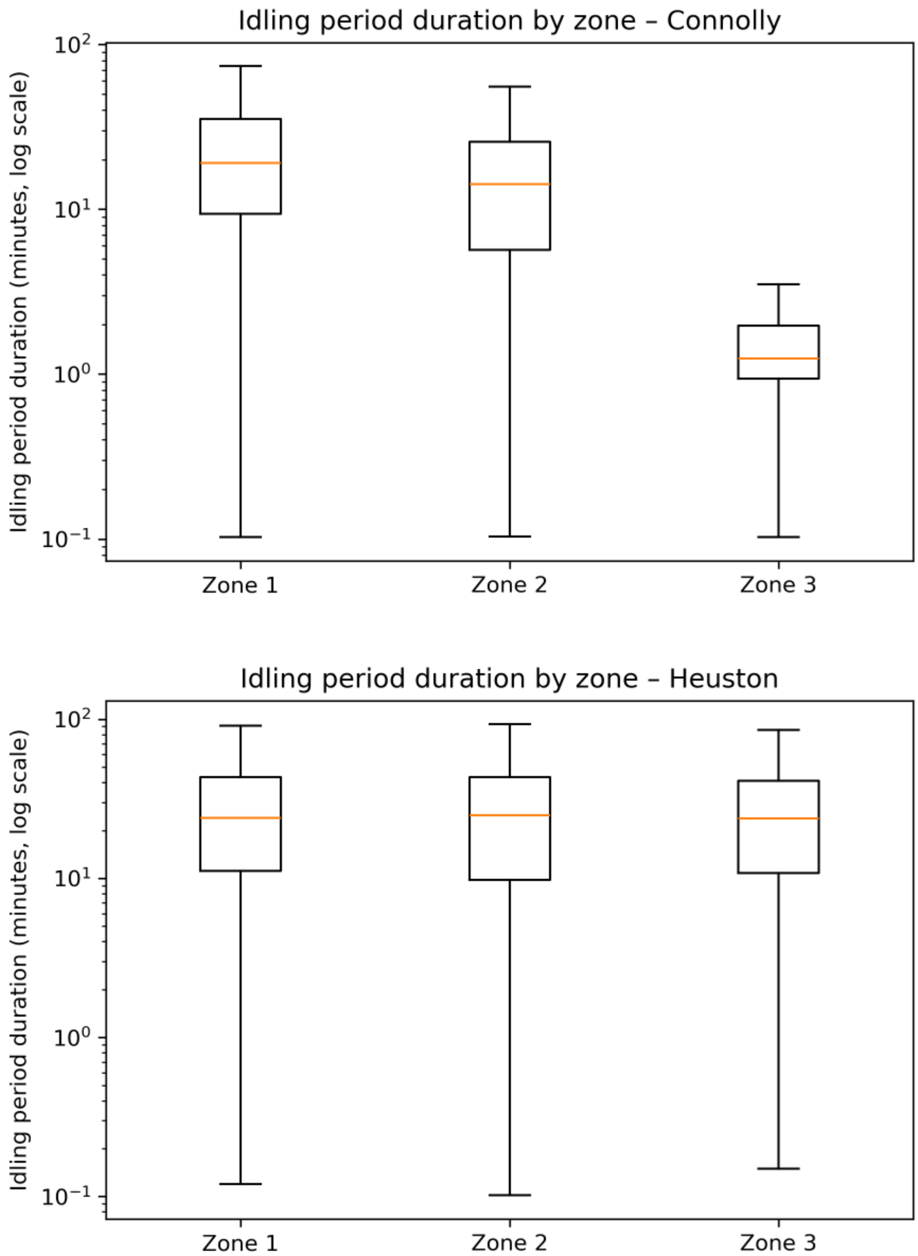


Figure 4.5 Idling duration by platform zone for 22000-series DMUs

To examine how idling behaviour varies within each station, the platform area is divided into three zones based on the most common GPS region occupied during each idling period. At Connolly, Zone 1 corresponds to the roofed portion of platforms 1–4, Zone 2 to the uncovered extension of these platforms, and Zone 3 to the covered through-platform area on platforms 5–7. At Heuston, Zone 1 is the covered part of platforms 2–5, Zone 2 the uncovered extension of these platforms, and Zone 3 the open platforms 6–8. Figure 4.5 shows boxplots of idling duration by zone for each station on a logarithmic y-axis, and the corresponding summary statistics are presented in Table 4.2

Table 4.2 Idling time statistics by station and zone

Station	Zone	N	Mean	Median	P10	P25	P75	P90	Max	%>15 mins	%>30 mins	%>60 mins
Connolly	1	2062	24.8	19.1	4.5	9.4	35.4	50.8	286.5	60.5	31.8	5.8
	2	1616	18.1	14.2	1.3	5.7	25.6	42.0	120.8	48.0	19.2	1.7
	3	4510	2.2	1.2	0.8	0.9	2.0	4.3	92.0	1.4	0.1	0.0
Heuston	1	1655	30.6	23.9	7.0	11.2	43.2	63.0	154.6	66.2	42.6	12.0
	2	1883	30.5	24.9	5.6	9.8	43.1	62.9	192.6	64.8	43.7	11.4
	3	705	28.6	23.7	7.3	10.8	41.1	59.3	140.1	65.8	40.3	9.8

Within Connolly, idle durations differ strongly between zones. In Zone 3, where most trains are pass services, idling periods are very short: the median duration is 1.2 minutes, and the mean is 2.2 minutes, with fewer than 1.4% of periods exceeding 15 minutes and almost none exceeding 30 minutes. This pattern is typical of brief passenger stops with little slack for extended idling. Zone 2, representing the uncovered extension of platforms 1–4, shows intermediate behaviour, with a median of 114.2 minutes and a mean of 18.1 minutes; around 48% of periods exceed 15 minutes, and 19.2% exceed 30 minutes. The longest idle periods at Connolly are concentrated in Zone 1, the covered part of platforms 1–4, which functions more like a terminal area. Here, the median idle duration is 19.1 minutes, and the mean is 24.8 minutes, with 60.5% of periods longer than 15 minutes, 31.8% longer than 30 minutes, and about 5.8% longer than one hour. Long-duration idling at Connolly is therefore not evenly distributed but is concentrated in the terminal platforms closest to the main concourse.

Heuston exhibits a different picture. As shown in Figure 4.5 and Table 4.2, all three zones display broadly similar idling behaviour, reflecting the fact that Heuston operates as a terminus across almost all platforms. Median idle durations range from 23.7 to 24.9 minutes across Zones 1–3, with mean values between 28.6 and 30.6 minutes. In each zone, roughly 65% of periods exceed 15 minutes, around 40-44% exceed 30 minutes, and 10–12% exceed 60 minutes. Zone 2, the uncovered extension of platforms 2–5, has the longest mean duration (30.5 minutes) and the highest share of very long periods, but the differences between zones within Heuston remain much smaller than at Connolly. For exposure and mitigation, this means that idling at Heuston cannot be addressed simply by relocating services to a "cleaner" part of the station; long layovers are embedded throughout the platform layout.

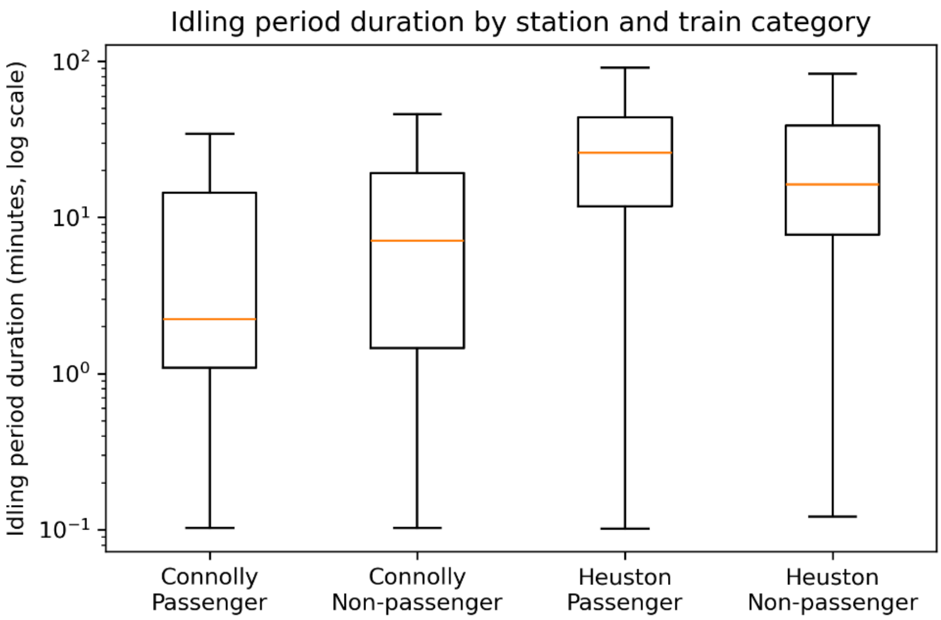


Figure 4.6 Idling duration by station and train category

The analysis also distinguishes passenger and non-passenger uses of the class 22000 DMUs. Figure 4.6 plots the distributions of idling duration by train category for each station, using boxplots on a logarithmic y-axis, and Table 4.3 summarises the corresponding statistics. At Connolly, passenger services account for most idling periods. Passenger and non-passenger trains have broadly similar distributions in the short-to-moderate range, but the very longest idle events are more often associated with non-passenger movements such as shunting or radio test trains (operational testing of onboard or railway communication systems). Passenger idle periods at Connolly have a median duration of 2.2 minutes and a mean of 10.4 minutes; about 24.1% exceed 15 minutes, and 11.2% exceed 30 minutes. Non-passenger idle periods have a slightly higher mean (14.5 minutes) and a higher share of very long events, with 3.3% exceeding 60 minutes. This indicates that, while most day-to-day exposure at Connolly is driven by regular passenger dwell, extreme idling events that dominate fuel use and emissions often arise from auxiliary operations.

Table 4.3 Idling time statistics by station and train category

Station	Train category	N	Mean	Median	P10	P25	P75	P90	Max	%>15 mins	%>30 mins	%>60 mins
Connolly	Non-passenger	1306	14.5	7.1	0.8	1.5	19.2	39.8	286.5	32.4	15.2	3.3
	Passenger	6882	10.4	2.2	0.8	1.1	14.4	32.1	117.5	24.1	11.2	1.5
Heuston	Non-passenger	710	25.6	16.3	2.2	7.7	38.6	59.4	166.0	52.0	34.1	9.7
	Passenger	3533	31.2	25.9	7.7	11.8	43.6	63.2	192.6	68.3	44.4	11.7

At Heuston, long-duration idling is primarily a passenger issue. Passenger services have a median idle duration of 25.9 minutes and a mean of 31.2 minutes, with around 68.3% of periods longer than 15 minutes and 44.4% longer than 30 minutes. Non-passenger movements are shorter on average (median 16.3 minutes, mean 25.6 minutes) and less likely to exceed 30 minutes, with about 52.0% of periods longer than 15 minutes, 34.1% longer than 30 minutes, and 9.7% exceeding one hour. In other words, at the terminal station the core scheduled passenger timetable, rather than occasional non-passenger moves, is still responsible for most of the idle time. This has direct implications for policy design: interventions that target only shunting or departmental activities would miss the majority of potential savings at Heuston, whereas at Connolly both passenger operations and auxiliary moves contribute materially to long idle periods.

Overall, the distributional analysis using log-scaled histograms and ECDFs confirms three key points that inform the later modelling and the concluding discussion of the thesis. First, idle time distributions at both stations are heavy tailed, with a small number of very long events that account for a large share of total idle hours. Second, station function matters: Connolly's mixed role leads to a combination of very short through-service stops and fewer, concentrated long layovers in specific terminal platforms, whereas at Heuston long-duration idling is common across the platform area. Third, the relative contribution of passenger and non-passenger services differs by station, with non-passenger movements dominating the extreme tail at Connolly, but regular passenger layovers driving long idling at Heuston. The next sections focus on passenger services and link these idling patterns more directly to TCs and driver behaviour, using regression models to quantify how changes in TCs translate into changes in arrival and departure idling.

4.4 Factors influencing passenger services' idling time

The previous chapter showed that platform air quality is dominated by emissions from diesel trains idling within the station area, and that a small number of long idle periods account for a large share of total engine-running time. Reducing unnecessary idling is therefore one of the most direct levers for improving air quality at Connolly and Heuston. Section 4.4 focuses on the operational mechanisms behind these idle periods, asking how the working timetable and drivers' shutdown decisions combine to produce the observed distributions of arrival, departure and total idling. The aim is to move from descriptive statistics to a behavioural understanding of when and why trains idle, and to quantify how much of the timetable slack is typically converted into engine-off time versus additional idling.

Idling within the station varies widely across locations, and a small number of long periods account for a large share of total idling time. This section focuses on idling linked to passenger services that terminate or originate at the station and change Train ID. These switching movements are operationally important because they correspond to planned turnrounds in the working timetable and thus provide a natural setting to examine how timetable design and driver decisions jointly shape idling.

For each such movement at Connolly and Heuston, the processed datasets link timetable information (arrival and departure Train IDs, scheduled gap between arrival and departure) with measured idling within the station area. The key variables are the TC, defined as the scheduled layover between arrival and departure; an indicator of whether the engine is shut down and restarted during this period (idling interruption); and the associated arrival, departure, and total idling times. At both stations, the sample includes only terminating or originating passenger services that change Train ID, with idling of through services excluded.

This subsection examined how these factors interacted. First, the influence of TC on the probability of an idling interruption was assessed under the assumption that when no interruption occurred, the idling duration was effectively equal to TC. The analysis then considered how TC and interruption status jointly determined the length and allocation of idling between arrival and departure, and whether these relationships differed between Connolly and Heuston. This linked the descriptive distributions in Section 4.3 to the subsequent modelling of strategies to reduce unnecessary passenger idling in terminal stations.

4.4.1 Timetable constraint and the probability of idling interruption

The first step in analysing the idling time of passenger services' trains is to examine how the scheduled layover influences whether drivers shut down the engine during the period between arrival and departure. For switching movements at Connolly and Heuston, each switch was treated as a single observation, defined by the matched arrival and departure Train IDs for the same unit on the same day. The TC is taken from the working timetable as the scheduled gap between arrival and departure. When no idling interruption occurs, the engine is assumed to remain running for essentially the full layover, and the idling duration is therefore taken as equivalent to TC. Under this assumption, TC is the primary determinant of whether a shutdown is initiated: short gaps rarely justify switching off, whereas longer gaps provide both greater opportunity and stronger incentive to shut down.

To quantify this relationship, the probability of an idling interruption was modelled as a function of TC for each station. The binary response variable was coded as 1 when the period was classified as interrupted (engine shut down and restarted) and 0 otherwise. A logistic regression of the following form was fitted:

Equation 4.1:

$$Pr(interruption = 1 \mid TC) = logit^{-1}(\alpha + \beta \cdot TC)$$

Where $logit^{-}1(x) = e^{x}/(1 + e^{x})$

The function was fitted separately for Connolly and Heuston using the switching-train datasets. In addition, for descriptive purposes, the data are grouped into 10-minute TC bins, and the empirical interruption probability in each bin is plotted together with binomial 95% confidence intervals. The logistic curves are drawn on the same axes to provide a smooth representation of how interruption probability rises with TC (Figure 4.7).

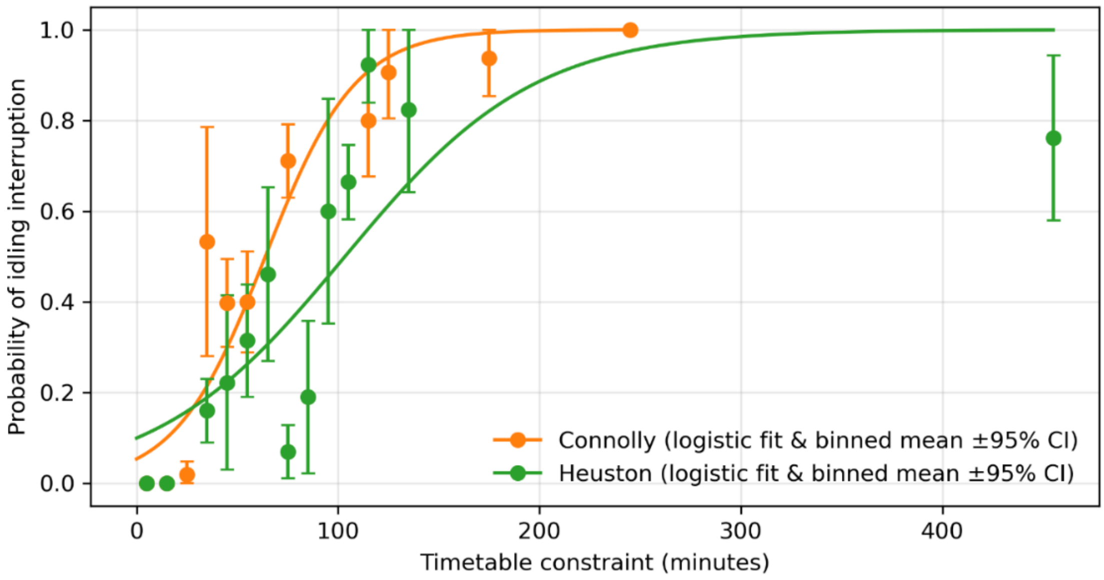


Figure 4.7 Idling interruption vs timetable constraint (logistic fit)

The fitted logistic models show that longer TCs are associated with a greater probability that an idling period will be interrupted (the train engine is shut down and restarted while waiting for the next service). At Connolly, the estimated coefficient for TC is about 0.045 per minute (standard error ≈ 0.0038), corresponding to a multiplicative increase in the odds of interruption of roughly 1.57 for every additional 10 minutes of TC. At Heuston, the TC coefficient is smaller, about 0.021 per minute (standard error ≈ 0.0026), implying an odds ratio of approximately 1.24 per 10 minutes. Thus, while TC has a positive effect at both stations, the sensitivity is stronger at Connolly.

The observed data points in Figure 4.7 show a similar pattern. For both stations, idling interruptions are virtually absent when TC is below about 20 minutes: almost all short layovers are handled with the engine kept running throughout. As TC increases into the 30–60 minute range, the probability of interruption begins to rise. At Connolly, the logistic fit suggests that the probability of an interruption grows from around 0.12 at TC = 20 minutes to about 0.46 at TC = 60 minutes, and exceeds 0.9 for very long layovers around 120 minutes. At Heuston, the increase is more gradual: the fitted probability is around 0.15 at TC = 20 minutes, 0.28 at 60 minutes and about 0.59 at 120 minutes. For very long constraints, both stations show high interruption probabilities, but Connolly tends to reach these levels at shorter TC.

These results support the behavioural assumption used in the rest of the analysis. When TC is short, passenger trains almost always idle continuously, so idling time is effectively equal to the scheduled layover. As TC becomes longer, the probability of an interruption increases, especially at Connolly, and the simple equivalence between TC and idling time starts to break down. In the next subsections, TC and the interruption indicator are used together to explain variation in arrival, departure and total idling time, with the logistic results here providing the motivation for treating TC both as a direct determinant of idling duration and as a key driver of shutdown decisions.

4.4.2 Total interrupted idling and timetable constraint

Section 4.1.1 shows that the timetable TC strongly influences whether an idling interruption occurs. When no interruption is observed, idling duration is effectively equal to TC. This subsection focuses on the complementary case of interrupted idling periods, in which the engine is shut down and restarted between arrival and departure. For these periods, total idling time is, by definition, less than the timetable gap, and the key question is how the remaining idling varies with TC. Understanding this relationship is important, because it determines how much idling can realistically be removed by promoting shutdowns during longer layovers and provides the basis for the Bayesian model in Section 4.4.3.

The analysis is restricted to periods classified as having an interruption in the Connolly and Heuston switching-train datasets. For each period, TC is taken from the working timetable as in Section 4.4.1, and total idling time is the sum of the arrival and departure idling attributed to the corresponding arrival and departure Train IDs. Figure 4.8 plots total idling time against TC for interrupted periods only. Individual observations are shown as pale scatter points. To stabilise the mean pattern, TC is grouped into 20-minute bins and the mean total idling in each bin is plotted together with 95% confidence intervals based on within-bin variability. To avoid imposing a specific parametric form (such as a linear or polynomial relationship), a locally weighted regression (LOWESS) smoother is fitted separately for Connolly and Heuston. LOWESS is a simple non-parametric method that does not assume linearity and is well suited to exploratory plots of this kind, providing a flexible trend line that follows the scatter and the binned means without being overly influenced by individual outliers (Cleveland, 1979; Cleveland & Devlin, 1988).

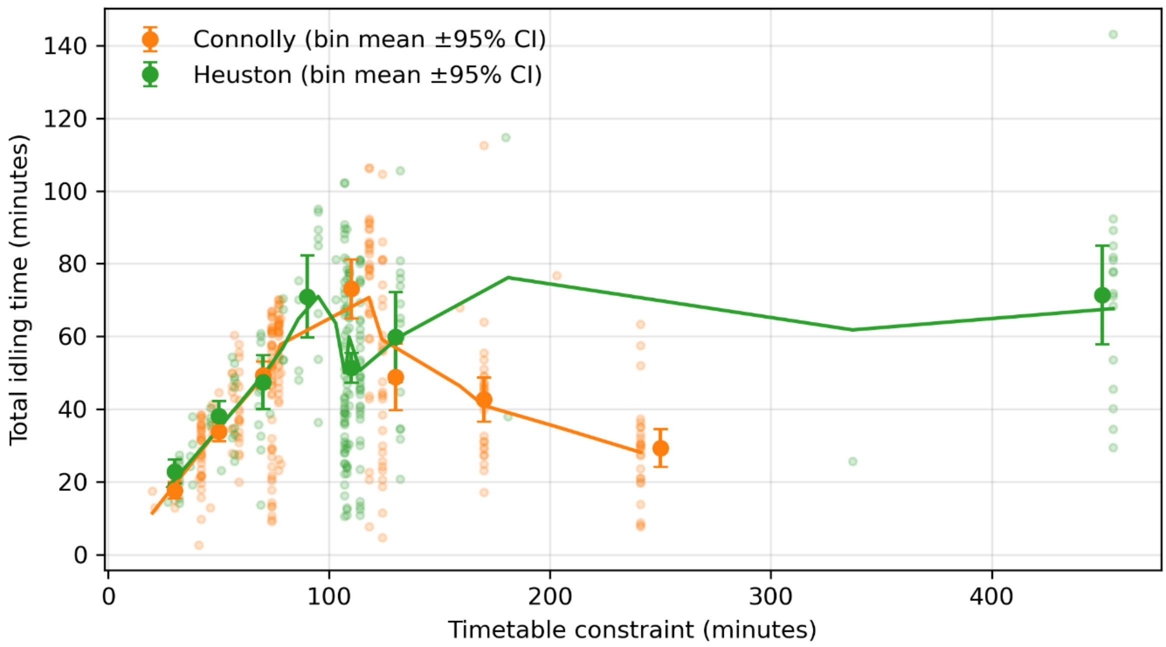


Figure 4.8 Total interrupted idling vs timetable constraint (LOWESS smooth)

The smoothed curves show a clear but non-linear relationship between TC and total interrupted idling, and they differ noticeably between the two stations. At Connolly, the mean total idling in interrupted periods increases from about 20 minutes for TC around 30 minutes (bin centre 30) to roughly 34 minutes at 50 minutes and about 46 minutes at 70 minutes. The highest bin mean is observed for TC around 110 minutes, where the average total idling reaches about 71 minutes. Beyond this point, the curve turns downwards: for TC around 130 minutes and 170 minutes, the mean total idling falls to about 45 minutes and 42 minutes, respectively, and for the longest observed layovers (TC ≈ 250 minutes) the mean is only about 28 minutes, with a relatively narrow confidence interval. This pattern suggests that once a shutdown occurs, drivers at Connolly do not allow total idling to grow in proportion to TC. Instead, interrupted periods tend to cluster around an upper band of roughly one hour of idling, and very long TCs are largely absorbed as non-idling time (engine off).

In practical terms, this means that for long TCs at Connolly, most additional timetable slack beyond about one hour is already being used as engine-off time rather than as extra idling. The analysis reveals a "peak pollution window" for medium-duration layovers, with TCs of roughly 1 to 2 hours, where ambiguity about the remaining wait leads to disproportionately high engine-on time compared with longer, clearly defined breaks. Very long TCs appear to be largely self-regulating, because the extended gap removes much of the hesitation to shut down and encourages engines to be switched off. Mitigation efforts should therefore focus on these intermediate layovers, by reducing the one-hour idle band and providing clearer guidance on when engines are expected to be shut down during medium-length waits.

Heuston exhibits a different profile. For short to medium layovers, mean total idling rises from about 28 minutes at TC ≈ 30 minutes to around 47 minutes at 50 minutes and 51 minutes at 70 minutes. The bin centred on 90 minutes shows the highest mean of roughly 75 minutes, indicating that interrupted periods with TC around one and a half hours still contain long blocks of idling. For TC ≈ 110 minutes and 130 minutes, mean total idling drops back to around 55–60 minutes, but remains substantial. In the longest bin, centred on 450 minutes, the mean total idling is again high at about 70 minutes, although the confidence interval is relatively wide due to limited sample size.

Overall, Heuston shows a tendency for interrupted periods with large TC to retain long idling components, even when the engine has been shut down at some point in the layover. This profile suggests that shutdowns at Heuston often coexist with substantial residual idling blocks during long layovers. Compared with Connolly, a larger share of timetable slack is still spent with engines running, even when an interruption occurs. From a mitigation perspective, this indicates that Heuston offers greater potential to cut total idling without major timetable changes, by tightening shut-down procedures and encouraging drivers to minimise engine-on time within long interrupted periods.

Two features stand out in Figure 4.8. First, even in interrupted periods, total idling remains substantial: for most of the range where data are dense, the smoothed curves lie between 30 and 70 minutes, and at Heuston, they exceed 100 minutes for some of the longest layovers. Second, the relationship between TC and total interrupted idling is clearly non-linear and rather noisy, with much of the variability unexplained by TC alone. This reinforces the idea that, once shutdowns occur, local operating practices and driver decisions play a large role in determining how much of the layover is actually spent idling.

These observations motivate the Bayesian model developed in Section 4.4.3. Rather than assuming a simple linear effect of TC, the model treats total idling time in interrupted periods as a stochastic outcome with a flexible mean function of TC (for example, represented by spline terms) and station-specific effects, while also allowing for substantial residual variation around this mean. In this way, the model can capture both the overall increasing–flattening pattern seen in the LOWESS curves and the wide spread of idling times observed for any given TC.

4.4.3 Bayesian regression model

The preceding subsections have examined the effects of timetable constraints on idling behaviour in two complementary ways. First, descriptive statistics and non-parametric smoothing showed that total idling time during interrupted periods responds to TC in a non-linear way, with clear changes in behaviour at longer layovers. Second, logistic regression models quantified how TC influences the probability that an idling spell is interrupted by an engine shutdown. Together, these results suggest that both the occurrence of shutdowns and the amount of idling conditional on shutdown are shaped by TC, but in ways that are heterogeneous across drivers and stations.

This subsection developed a Bayesian framework to bring these components together and quantify how TCs shape idling behaviour at Connolly and Heuston. Earlier descriptive and analyses showed that idling times are strictly positive, heavily right-skewed, and often heavy-tailed, and that their response to TC is non-linear with clear constraint changes. Classical linear models are poorly suited to such data and provide only point estimates, whereas here need (i) flexible likelihoods for skewed idling times, (ii) a principled way to incorporate independent prior information from first departures and last arrivals, and (iii) full uncertainty quantification when combining TC–idling relationships with TC-dependent interruption probabilities.

In this context, Bayesian models are particularly attractive because the posterior distribution naturally expresses uncertainty arising from heterogeneous and partially random driver decisions; similar interpretations of posterior variation as behavioural randomness are widely used in transport applications such as Bayesian travel-time and dwell-time modelling (Isukapati et al., 2021; Jintanakul et al., 2009; Tebaldi & West, 1998). Thus, the Bayesian framework adopted here allows us both to represent the stochasticity of driver behaviour and to propagate this uncertainty through to predicted idling under alternative timetable scenarios.

4.4.3.1 Station-specific thresholds in the timetable–idling relationship

Sections 4.4.1 and 4.4.2 showed that (i) the probability of an idling interruption increases with TC and (ii) total idling time initially rises with TC but appears to saturate beyond a certain range (Figure 4.8). To formalise this pattern, the TC–idling relationship was treated as a two-constraint process, and station-specific breakpoints were estimated using total idling time as the response.

Segmented (broken-stick) regression was used, following Muggeo's framework for regression models with unknown breakpoints (Muggeo, 2003). For each station, total idling time T_i was modelled as a piecewise log-linear function of TC:

Equation 4.2:

$$\mathrm{Log}\, T_i = \rho_0 + \beta_1 TC_i + \beta_2 (TC_i - c)_+ + \varepsilon_i$$

Where $(TC_i - c)_+ = \max(\mathrm{TC_i} - \mathrm{c}, 0)$ and c is the unknown breakpoint. This specification allows the slope of $\mathrm{Log}\ T_i$ with respect to TC to change at TC = c, while remaining continuous at the breakpoint. Rather than estimating c iteratively, a transparent information-theoretic strategy was used. For each station, the model was fitted over a grid of candidate breakpoint values spanning the observed TC range (excluding the extreme tails so that both sides contained at least 10 observations) and selected the value of c that minimised the AIC. An approximate confidence band for the breakpoint was defined as the set of candidate values with ΔAIC ≤ 2, corresponding to a standard likelihood-support region.

Table 4.4 Idling time statistics by station and train category

Station	Breakpoint min	AIC_band lower	AIC_band upper	Slope before	Slope after
Connolly	78	77	78	0.021	-0.0036
Heuston	57	46	79	0.030	0.00097

Table 4.4 summarises the segmented regression results using five statistics for each station. "Breakpoint min" is the value of TC (in minutes) at which the AIC is minimised. "AIC_band lower" and "AIC_band upper" give the lower and upper limits of the support region with ΔAIC ≤ 2, that is, the range of breakpoint values that are statistically indistinguishable from the optimum on an AIC basis. "Slope before" and "Slope after" are the estimated slopes of log total idling with respect to TC below and above the breakpoint, respectively; positive slopes indicate that idling increases with TC in those conditions, whereas near-zero or negative slopes indicate saturation or a slight decline in idling as TC increases further.

At Connolly, the best breakpoint occurred at 78 min of TC, with an AIC-based support band of 77–78 min. Here, "best breakpoint" denotes the TC value at which a piecewise linear model with different slopes below and above the breakpoint provides the best balance of fit and parsimony, as measured by the Akaike Information Criterion. In practical terms, it marks the point where the marginal effect of extra TC on total idling changes regime. Below this point, the fitted slope of log total idling with respect to TC was clearly positive (≈0.021 per minute), whereas above 78 min the slope became slightly negative (−0.004 per minute).

At Heuston, the estimated breakpoint was 57 min, with a broader support band of 46–79 min. Here the pre-breakpoint slope was again strongly positive (0.030 per minute), but beyond 57 min the slope flattened to approximately zero (0.001 per minute). The piecewise fits superimposed on the scatterplots (Figure 4.9): total idling increases rapidly with TC up to around one hour at Heuston and roughly 80 minutes at Connolly, after which further timetable slack produces little additional increase and may even allow some reduction in platform idling.

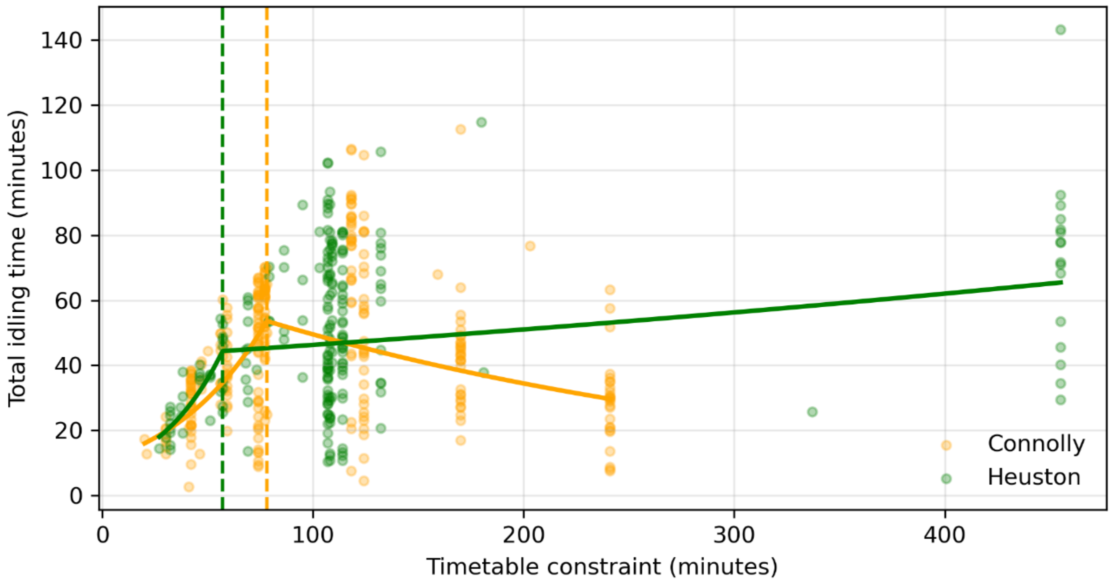


Figure 4.9 Piecewise log-linear fits of total idling vs timetable constraint

To test whether arrival and departure idling have distinct thresholds, the segmented analysis was repeated for arrival and departure idling at each station. Overall, the estimated breakpoints were either poorly constrained (wide support bands) or overlapped the total-idling threshold.

At Connolly, the arrival breakpoint (≈57 min; support band 56–78 min) was not clearly separable from the total-idling threshold (77–78 min). The departure breakpoint (≈159 min) was unstable and appeared to be driven by a small number of extreme layovers. At Heuston, the departure breakpoint (≈46 min; 38–57 min) fell within the total-idling support band, while the arrival breakpoint was too imprecise (support band 56–223 min) to justify a distinct threshold.

Using a simple representation with one TC threshold per station, defined from the total-idling models: approximately 60 min at Heuston and 80 min at Connolly. In the remainder of this section, TC below these thresholds is treated as a tight-timetable constraint, in which additional slack has a clear marginal effect on idling, while TC above the threshold is treated as a relaxed-timetable constraint, where further slack has little additional influence on total idling behaviour.

4.4.3.2 Prior information and selection of likelihoods

The breakpoint analysis in Section 4.4.3.1 showed that the marginal effect of TC on total idling changes around station-specific thresholds. In the Bayesian regression models that follow, making use of independent prior information on idling in situations where trains are not under timetable pressure. This section describes these prior datasets and examines which parametric families provide an adequate statistical description of them, forming the basis for the likelihood choices in Section 4.4.3.3.

Description of prior datasets

For each terminus, two prior datasets were assembled from operating records:

Arrival prior: duration of the idle state after final arrival for terminating trains, measured from the point at which the train has arrived and is stably in idle (no further movement, but engine running) until operational activities associated with the last arrival are complete. These cases are not constrained by a subsequent departure under time pressure.

Departure prior: duration of the idle state before first departures originating at the station, measured from the onset of stable idling prior to departure until the train leaves the platform.

At Connolly, the arrival prior comprises 239 observations (mean 16.2 min, standard deviation 10.9 min, skewness 1.55), while the departure prior includes 215 observations (mean 18.5 min, standard deviation 17.3 min, skewness 1.77). At Heuston, the arrival prior contains 151 observations (mean 27.4 min, standard deviation 24.7 min, skewness 1.88), and the departure prior likewise has 151 observations (mean 29.2 min, standard deviation 11.6 min, skewness 0.30).

Fitting candidate distributions to the prior data

To identify suitable likelihood families, three standard two-parameter distributions were fitted to each prior dataset:

- Log-normal (parameters μ, σ on the log scale);
- Gamma (shape k, scale θ);
- Weibull (shape c, scale λ).

The log-normal, Gamma and Weibull distributions were chosen because they are standard models for positive event-time data. Reliability guidance lists these distributions as basic lifetime models for non-repairable populations (NIST/SEMATECH, 2013; Taketomi et al., 2022). They are defined on the positive real line and are designed for non-negative, right-skewed survival times (Taketomi et al., 2022). This shape matches the long right tails observed in the idle-time data. In public-transport research, log-normal, Gamma and Weibull distributions are widely tested as candidate models for skewed quantities such as bus running times, dwell times and route travel times (Büchel & Corman, 2020; Luo et al., 2022; Shen et al., 2019). Using the same three families here therefore aligns the train idle-time modelling with established practice in both reliability and transport applications.

Parameters were estimated by maximum likelihood. For each fit, the Akaike Information Criterion (AIC) and the Kolmogorov–Smirnov (KS) goodness-of-fit statistic were computed (Table 4.5). Lower AIC values indicate a better trade-off between fit and parsimony, while the KS statistic and its p-value provide a simple check for obvious discrepancies between the fitted distribution and the empirical cumulative distribution. Because parameters are estimated from the same data, KS p-values are interpreted as approximate diagnostics rather than strict hypothesis tests.

Table 4.5 Fits of candidate distributions to prior idling data
(AIC = Akaike Information Criterion; KS = Kolmogorov–Smirnov statistic. μ, σ are log-normal parameters on the log scale; k, θ are Gamma shape and scale; c, λ are Weibull shape and scale.)

Station	Type	Distribution	n	AIC	KS stat	KS p	Main parameters (approx.)
Connolly	Arrival	Log-normal	239	1751.7	0.054	0.468	$\mu = 2.53$, $\sigma = 0.84$
		Gamma	239	1808.7	0.106	0.008	$k = 2.10$, $\theta = 7.73$
		Weibull	239	1750.4	0.044	0.721	$c = 1.55$, $\lambda = 18.04$
Heuston	Arrival	Log-normal	151	1275.6	0.123	0.019	$\mu = 2.99$, $\sigma = 0.82$
		Gamma	151	1283.6	0.133	0.008	$k = 1.70$, $\theta = 16.07$
		Weibull	151	1291.0	0.111	0.046	$c = 1.27$, $\lambda = 29.75$
Connolly	Departure	Log-normal	215	1687.3	0.117	0.005	$\mu = 2.31$, $\sigma = 1.18$
		Gamma	215	1680.1	0.052	0.596	$k = 0.96$, $\theta = 19.24$
		Weibull	215	1686.0	0.108	0.013	$c = 0.94$, $\lambda = 17.87$
Heuston	Departure	Log-normal	151	1220.1	0.106	0.061	$\mu = 3.27$, $\sigma = 0.51$
		Gamma	151	1185.0	0.069	0.454	$k = 5.08$, $\theta = 5.74$
		Weibull	151	1171.3	0.041	0.950	$c = 2.69$, $\lambda = 32.73$

The results show that all three families can reproduce the main body of the prior distributions. For idling after the last arrival at Connolly, Log-normal and Weibull distributions achieve lower AIC and higher KS p-values than the Gamma. At Heuston, the log-normal has the lowest AIC for arrival idle, with the Gamma and Weibull models providing slightly worse. For idling before first departures, the Gamma fits Connolly best according to both AIC and KS, while at Heuston the Weibull distribution gives the lowest AIC and the highest KS p-value, with the Gamma model again close behind.

Choice of likelihoods for the regression models

Station-specific prior fits sometimes favour different distributions (for example, Weibull for some Connolly priors and log-normal for some Heuston priors). However, using different likelihood families at each terminus would make the later regression models harder to compare and interpret. The main role of the prior information here is to characterise typical magnitudes and variability of arrival idling after arrival and departure idling before first departures, and to inform baseline parameters in the Bayesian regression models, rather than to capture every detail of the far tails.

To retain parsimony and comparability across stations, a single likelihood family is adopted for each type of idling. For arrival idling, a log-normal distribution is used at both Connolly and Heuston. In all four prior datasets, the logarithm of arrival idle times is approximately symmetric, and the log-normal fits provide an adequate description of the centre and right tail of the distributions. Even where Gamma or Weibull models achieve slightly lower AIC (as at Connolly), the differences are modest and largely confined to extreme values. The log-normal form also has a straightforward interpretation in terms of multiple small multiplicative factors affecting post-arrival idle time.

For departure idling, a Gamma distribution with a log link is used at both termini. The Gamma model consistently delivers good AIC and KS diagnostics for the departure priors and matches the pattern of positively skewed waiting times with variance increasing in the mean. At Heuston, Weibull fits are marginally better by AIC, but the Gamma model remains competitive and is more convenient for the regression structure adopted in the next subsection.

In the Bayesian TC–idling models that follow, log-normal likelihoods are therefore used for arrival idling and Gamma likelihoods for departure idling, with their parameters informed by the priors described above.

4.4.3.3 Bayesian regression of tight timetable constraints and idling

Building on the station-specific thresholds identified in Section 4.4.3.1 and the likelihood choices in Section 4.4.3.2, this section specifies the Bayesian regression models used to quantify how TC influences arrival and departure idling within the tight-timetable constraint. Recall that this constraint is defined as $TC \leq 80$ min at Connolly and $TC \leq 60$ min at Heuston, where total idling increases most strongly with TC.

Arrival idling and departure idling were modelled separately for each station, using log-normal and Gamma likelihoods, respectively. The aim is to estimate, for each terminus, how expected idling changes with TC when timetable pressure is present but has not yet reached the relaxed constraint (threshold) identified in Section 4.4.3.1, and to assess the sensitivity of those estimates to the use of prior information.

Model structure

For each observation i at station $s \in$ {Connolly, Heuston} in the tight-timetable constraint, let:

- A_i is the arrival idle time after arrival (minutes),
- D_i is the departure idle time before departure (minutes),
- TC_i is the TC (minutes).

To improve numerical stability and to give the regression coefficients a simple interpretation, a centred and rescaled TC covariate is used.

Equation 4.3:

$$\widetilde{TC_i} = \frac{TC_i - \overline{TC}_s}{10}$$

where $T\bar{C}_s$ is the mean TC in the tight-timetable constraint at station s. A one-unit increase in $\widetilde{TC_i}$ corresponds to a 10-minute increase in TC at that station.

The arrival model assumes a log-normal likelihood:

Equation 4.4:

$$\log A_i \sim N\left(\mu_i^{(A)}, \sigma_S^2\right), \quad \mu_i^{(A)} = \alpha_s^{(A)} + \beta_s^{(A)} \widetilde{TC_i}$$

with station-specific residual standard deviation σ_S . Here $\alpha_s^{(A)}$ is the typical log arrival idling time at mean TC, and $\beta_s^{(A)}$ measures the proportional change in arrival idling for a 10-minute increase in TC. On the original scale, a 10-minute increase in TC multiplies the expected arrival idling by approximately exp ($\beta_s^{(A)}$)

The departure model assumes a Gamma likelihood with a log link:

Equation 4.5:

$$D_i \sim Gamma\left(mean = \mu_i^{(D)}, \phi_s\right), \quad log\mu_i^{(D)} = \alpha_s^{(D)} + \beta_s^{(D)}\widetilde{TC}_\iota$$

Where ϕ_s is a station-specific shape parameter controlling the dispersion. In this formulation $\alpha_s^{(D)}$ is the log mean departure idling time at mean TC, and $\beta_s^{(D)}$ gives the proportional change in mean departure idling for a 10-minute increase in TC (exp($\beta_s^{(D)}$) is the multiplicative effect on the mean.)

Both models are fitted separately per station, with four sets of regression coefficients: $\left\{\alpha_{Con}^{(A)}, \beta_{Con}^{(A)}\right\}$,$\left\{\alpha_{Heu}^{(A)}, \beta_{Heu}^{(A)}\right\}$, and pairs for departures.

Prior distributions

The prior datasets described in Section 4.4.3.2 provide estimates of typical idling durations in situations without timetable pressure. Let $m_S^{(A)}$ and $m_S^{(D)}$ denote the mean arrival idle after the last arrival and the mean departure idle before the first departure in the prior information for the station. These values are used to anchor the intercepts in the regression models.

On the log scale, the intercept priors are

Equation 4.6:

$$a_s^{(A)} \sim N\left(\log m_S^{(A)}, \tau^2\right),\ a_s^{(D)} \sim N\left(\log m_S^{(D)}, \tau^2\right),$$

with τ=0.4. This variance corresponds to an approximate ±50 % range around the prior mean on the original scale, providing a weakly informative anchor that can still be substantially updated by the TC-dependent data.

The regression slopes are given weakly informative priors centred at zero,

Equation 4.7:

$$\beta_s^{(A)} \sim N(0, 0.1^2),\ \beta_s^{(D)} \sim N(0, 0.1^2)$$

which allows for a wide range of plausible effects (both positive and negative) while ruling out unrealistically steep changes in idling for modest variations in TC.

For the dispersion parameters, half-Normal priors were used

Equation 4.8:

$$\sigma_s \sim Half - Normal(0,1),\ \ \phi_s \sim Half - Normal(0,5),$$

which gently regularise the residual variance and Gamma shape without imposing strong prior beliefs.

To assess how strongly the empirical priors influence the results, a parallel set of sensitivity models was estimated in which the station-specific intercepts for arrivals and departures, $a_s^{(A)}$ and $\alpha_s^{(D)}$, were given diffuse priors,

Equation 4.9:

$$a_s^{(A)}, \alpha_s^{(D)} \sim N(0, 10^2)$$

while the priors on the slope and dispersion parameters were left unchanged. The intercepts control the typical level of idling when the TC is close to zero, so replacing the empirical priors with a very wide Normal prior largely removes prior information about baseline idling. Posterior slopes and fitted idling–TC curves (predicted idling duration as a function of TC were then compared between the empirical-prior and diffuse-prior models to assess whether the shape or level of the relationship changed. If the curves and slopes are similar under both specifications, the main conclusions about how idling responds to timetable pressure can be regarded as robust to the choice of priors on the intercepts. This “diffuse-prior” specification therefore acts as a prior-sensitivity check, following standard Bayesian practice of refitting models under alternative priors to assess the influence of prior assumptions on posterior inferences (Gelman et al., 2013; Lunn et al., 2012; McElreath, 2020; Vehtari et al., 2024).

Estimation and diagnostics

All models are estimated in Python using the No-U-Turn Sampler (NUTS) implemented in PyMC.

NUTS is an adaptive Hamiltonian Monte Carlo (HMC) algorithm that automatically tunes both the step size and the number of leapfrog steps, improving sampling efficiency and reducing the need for manual tuning (Hoffman & Gelman, 2014; Salvatier et al., 2016). For each station and direction, four parallel chains were run, with a warm-up period followed by sufficient post-warm-up iterations to obtain effective sample sizes well above 400 for all regression parameters. Convergence is assessed using the Gelman–Rubin statistic ($\hat{R}$), which remained close to 1 for all parameters.

Model fit was assessed using posterior predictive checks. For each fitted model, replicated idling times were drawn from the posterior predictive distribution and their 0.1, 0.5 and 0.9 quantiles were compared with the corresponding empirical quantiles of the observed data (Table 4.6). For arrival idling, the log-normal models reproduced the positive, right-skewed distributions reasonably well: at Connolly, the posterior predictive quantiles were generally within a few minutes of the observed quantiles across the lower, middle and upper parts of the distribution, while at Heuston the overall scale and skewness were captured but discrepancies were larger in the lower and upper tails. For departure idling, the Gamma models reflected the positively skewed, mean-dependent spread of the data but tended to under-predict typical (median) idling and over-predict the longest idling times at both stations. In all cases, informative- and diffuse-prior versions produced very similar posterior slopes and posterior predictive quantiles, indicating that the empirical priors are broadly consistent with the TC-dependent data and do not materially drive the fitted idling–TC relationships.

Table 4.6 Posterior predictive checks for arrival and departure idling models. Observed and posterior predictive 0.1, 0.5 and 0.9 quantiles of idling time (minutes) for each station, direction and prior specification.

Station	Direction	Prior	q	Observed idling (min)	Posterior predictive (min)
Connolly	Arrival	Informative	0.1	5.1	5.1
Connolly	Arrival	Informative	0.5	12.8	13.9
Connolly	Arrival	Informative	0.9	38.8	36.4
Connolly	Arrival	Diffuse	0.1	5.1	5.1
Connolly	Arrival	Diffuse	0.5	12.8	13.9
Connolly	Arrival	Diffuse	0.9	38.8	36.2
Connolly	Departure	Informative	0.1	5.1	4.8
Connolly	Departure	Informative	0.5	21.9	17.9
Connolly	Departure	Informative	0.9	42.5	46.7
Connolly	Departure	Diffuse	0.1	5.1	4.8
Connolly	Departure	Diffuse	0.5	21.9	18.0
Connolly	Departure	Diffuse	0.9	42.5	46.7
Heuston	Arrival	Informative	0.1	8.6	5.9
Heuston	Arrival	Informative	0.5	10.8	13.2
Heuston	Arrival	Informative	0.9	35.9	30.0
Heuston	Arrival	Diffuse	0.1	8.6	5.7
Heuston	Arrival	Diffuse	0.5	10.8	12.8
Heuston	Arrival	Diffuse	0.9	35.9	28.9
Heuston	Departure	Informative	0.1	10.0	9.5
Heuston	Departure	Informative	0.5	21.5	19.0
Heuston	Departure	Informative	0.9	28.8	35.2
Heuston	Departure	Diffuse	0.1	10.0	9.4
Heuston	Departure	Diffuse	0.5	21.5	18.9
Heuston	Departure	Diffuse	0.9	28.8	34.9

Posterior TC effects in the tight-timetable constraint

Table 4.7 summarises the posterior estimates of the TC effects $\beta_s^{(A)}$ and $\beta_s^{(D)}$ for arrival idling and departure idling at each station. These coefficients are defined on the centred and rescaled covariate $\widetilde{TC}_i$, so that a one-unit increase in TC_i corresponds to an additional 10 minutes of timetable constraint. The slopes β therefore represent the change in log-expected idling time per 10-minute increase in TC, while the multiplicative effects exp(β) reported in the table give the proportional change in expected idling for each extra 10 minutes of TC. For each station and idling type, results are shown for both the informative-prior and diffuse-prior specifications. The estimates from these two specifications are almost identical: posterior means differ only in the third decimal place, and the 95% credible intervals largely overlap. This similarity indicates that the data on TC and idling are strong enough to determine the slopes and that the empirical priors are not in conflict with them. In practice, the empirical priors mainly stabilise the intercepts by anchoring baseline idling at plausible levels and slightly reducing their uncertainty, while leaving the estimated TC effects β and their credible intervals essentially unchanged.

Table 4.7 TC effects on arrival and departure idling in the tight-timetable constraint

Station	Phase	Prior type	β(per 10 min)	95% CrI for β	Multiplicative effect exp (β)	95% CrI for exp (β)
Connolly	Arrival idling	Informative	0.232	0.171 – 0.288	1.26	1.19 – 1.33
		Weak	0.231	0.168 – 0.290	1.26	1.18 – 1.34
	Departure idling	Informative	0.106	0.034 – 0.176	1.11	1.04 – 1.19
		Weak	0.107	0.038 – 0.180	1.11	1.04 – 1.20
Heuston	Arrival idling	Informative	0.101	−0.036 – 0.230	1.11	0.97 – 1.26
		Weak	0.103	−0.027 – 0.234	1.11	0.97 – 1.26
	Departure idling	Informative	0.221	0.100 – 0.332	1.25	1.11 – 1.39
		Weak	0.224	0.106 – 0.339	1.25	1.11 – 1.40

For Connolly, the posterior TC effect on arrival idling is clearly positive and well identified. Under the informative model, the slope is $\beta_{Con}^{(A)}$ = 0.232 (95% CrI: 0.171–0.288), with an almost identical estimate under the weak-prior model (0.231; 95% CrI: 0.168–0.290). On the original scale, this corresponds to a multiplicative factor of exp (0.232) ≈ 1.26 in expected arrival idle for each additional 10 minutes of TC in the tight-timetable constraint. A 30-minute increase in TC approximately doubles arrival idle (exp (3 × 0.232) ≈ 2.0). Thus, at Connolly, when TC is below the threshold identified in Section 4.3.1, additional slack is systematically absorbed into longer idle periods after arrival.

For departure idling at Connolly, the TC effect is smaller but still credibly positive. The informative model provides $\beta_{Con}^{(D)}$ = 0.106 (95% CrI: 0.034–0.176), with the weak-prior model again matching very closely (0.107; 95% CrI: 0.038–0.180). This implies that each extra 10 minutes of TC increases mean departure idling by a factor of about exp (0.106) ≈ 1.11, or roughly 11%. Over a 30-minute increase in TC, expected departure idling rises by around 37% (exp (3 × 0.106) ≈ 1.37). Compared with the arrival model, this indicates that Connolly tends to allocate a larger share of additional TC to the post-arrival idle phase, with a more moderate but still systematic extension of pre-departure idling.

At Heuston, the posterior pattern is different. For arrival idling, the posterior mean slope is again positive but substantially more uncertain: $\beta_{Heu}^{(A)}$ = 0.101 (95% CrI: −0.036–0.230) under the informative model and 0.103 (95% CrI: −0.027–0.234) under the weak-prior model. On average, this corresponds to an 11% increase in arrival idle per additional 10 minutes of TC (exp(0.10) ≈ 1.11), but the credible intervals span both modest positive and near-zero effects. In other words, within the tight-timetable constraint at Heuston, arrival idle after arrival may increase with TC, but the data are consistent with a relatively weak or even negligible systematic effect, suggesting that other operational factors and driver behaviour introduce considerable randomness into post-arrival idling.

By contrast, for departure idling at Heuston, the TC effect is both large and precisely estimated. The informative model gives $\beta_{Heu}^{(D)}$= 0.221 (95% CrI: 0.100–0.332); the weak-prior model again agrees closely with a mean of 0.224 (95% CrI: 0.106–0.339). This implies that, in the tight-timetable constraint, each additional 10 minutes of TC increases the expected departure idling time by about 25% (exp (0.22) ≈ 1.25), and a 30-minute increase in TC almost doubles departure idling (exp (3 × 0.22) ≈ 1.94). The credible intervals exclude zero by a wide margin, indicating a strong association between timetable slack and pre-departure idling at Heuston.

Across all four station–direction combinations, the estimates obtained with and without informative priors are nearly indistinguishable: the posterior means differ only in the third decimal place, and the 95% credible intervals almost entirely overlap. This confirms that the empirical prior information from idle states after the last arrival and before the first departure is compatible with the TC-stratified idling data and serves mainly to stabilise the intercepts, rather than altering the estimated TC effects. Connolly allocates a substantial share of additional TC to the idle state after arrival, with a smaller but still systematic extension of departure idling, whereas Heuston exhibits a weak and uncertain TC effect on arrival idle but a strong and almost linear increase in departure idling as TC grows.

These relationships are illustrated in Figure 4.10, which shows how expected arrival and departure idling evolve with TC at each terminus in the tight-timetable constraint, together with their associated credible intervals and empirical prior information.

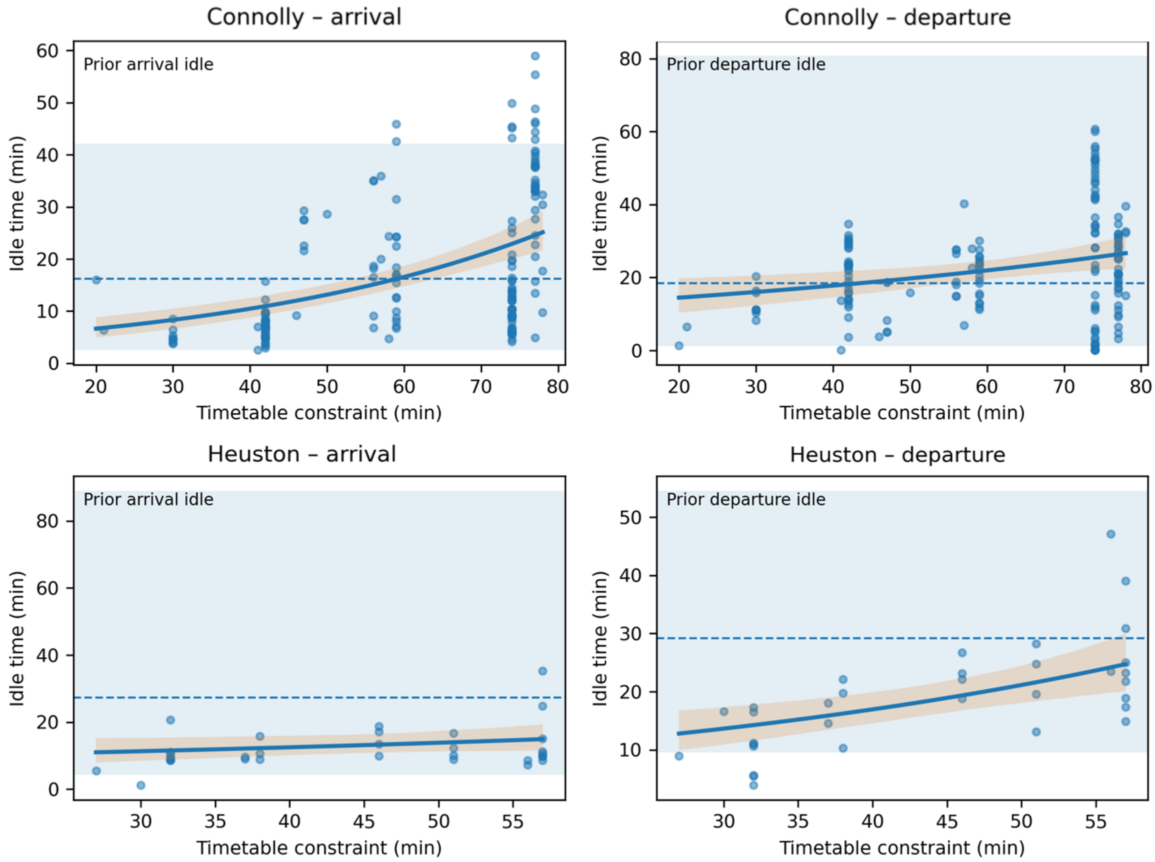


Figure 4.10 Expected arrival and departure idling evolve with tight TC

Each point represents an observed idling. Solid lines show the posterior mean expected idling from the Bayesian regression models, and the shaded bands show the corresponding 95% credible intervals. The light blue horizontal bands represent the empirical prior distributions for idle times without timetable pressure. Mean indicated by the dashed line, 2.5–97.5% range by the shaded area.

In the next section, these phase-specific TC effects are combined with the TC-dependent interruption probabilities from Section 4.4.1 to quantify how timetable design shapes overall idle durations at each station.

4.4.3.4 Combining interruption probability and idling duration

The previous sections treated two components of passenger idling behaviour separately:

1. The probability that an idle period is interrupted (engine shut down and later restarted) as a function of TC (Section 4.4.1)
2. The length of arrival and departure idling in the tight-timetable, conditional on interruption, is modelled with Bayesian regressions (Section 4.4.3.3)

In practice, the station experiences a mixture of interrupted and uninterrupted idle periods. To assess how timetable design influences total engine-running time, these two components were combined into a single measure of expected total idling as TC varies.

For a given station s and TC, two conditions were distinguished:

No interruption: the engine runs throughout the whole connection; by definition the total idling time is equal to the timetable gap

Equation 4.10:

$$T_{no-int}(TC) = TC$$

With interruption: the train idles after arrival, shuts down, and idles again before departure. The total idling time is then

Equation 4.11:

$$T_{int}(TC) = A_s(TC) + D_s(TC)$$

where $A_s(TC)$ and $D_s(TC)$ are the arrival and departure idling durations predicted by the Bayesian models in Section A.3.3 for that station.

Let $P_s(TC)$ denote the probability of an interruption at station *s* for TC, obtained from the logistic models in Section 4.3.1. The expected total idling at that station and TC is then

Equation 4.12:

$$E[T_s(TC)] = (1 - p_s(TC)) \cdot TC + p_s(TC) \cdot (E[A_s(TC)] + E[D_s(TC)])$$

In words, this is a weighted average of "always idling for the full TC" and "idling only during the arrival and departure phases", with weights given by the probability of an interrupted idle period.

Within the Bayesian framework, this quantity was obtained via Monte Carlo simulation. For each posterior draw, the fitted logistic interruption model and the arrival and departure regressions were evaluated on a grid of TC values to compute the corresponding $T_s(TC)$ for that draw. The posterior mean and 95% credible interval for $E[T_s(TC)]$ were then obtained by summarising $T_s(TC)$ across draws. This procedure propagates uncertainty from both the interruption probabilities and the idling-duration models.

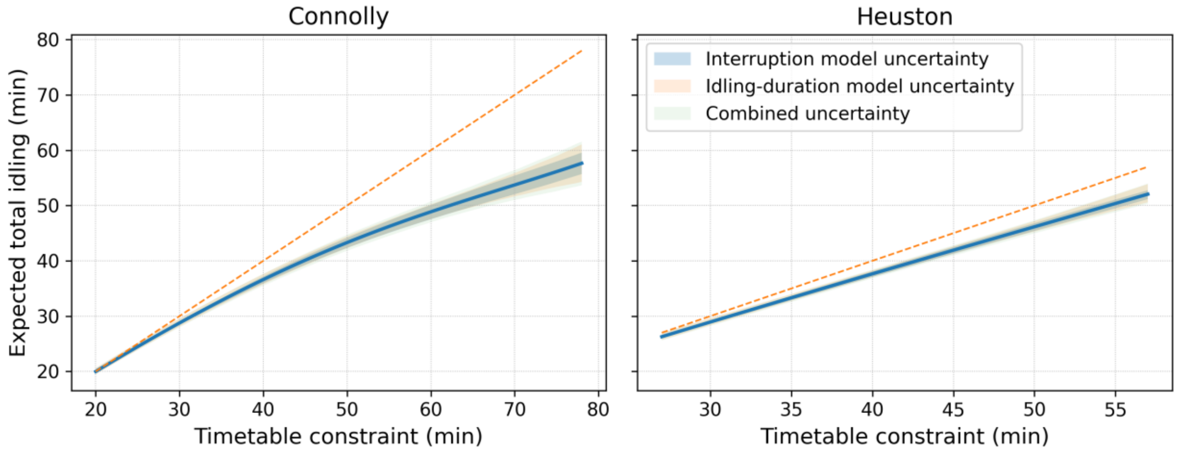


Figure 4.11 Expected total idling time vs timetable constraint, with decomposed uncertainty

Figure 4.11 shows how this total expected idling varies across the tight-timetable constraint at Connolly and Heuston. In each panel, the solid curve represents the posterior mean expected total idling as a function of TC. The dashed line is a 1:1 reference showing the hypothetical case where the engine idles continuously for the full TC (that is, no shut-down and restart). The shaded bands decompose the uncertainty in $E[T_s(TC)]$into three parts. The darkest band shows the 95% interval arising from uncertainty in the interruption (logistic) model alone, with the idling-duration models fixed at their posterior mean parameters. The medium band shows the 95% interval arising from the Bayesian idling-duration models alone, with the interruption model fixed at its maximum-likelihood estimate. The light outer band shows the combined 95% interval when both sources of uncertainty are propagated together.

It is important to note that the intervals in Figure 4.11 describe uncertainty in the expected mean idling time for a given TC, not the full spread of individual idle durations. In constructing these bands, uncertainty in the interruption model (logistic parameters) and in the Bayesian arrival–departure regressions (regression coefficients) are propagated, but the residual variability of the lognormal and gamma likelihoods for individual trains is not simulated. As a result, the confidence intervals around the mean curves are comparatively narrow, even though the underlying observations show substantial scatter; the wide variation between trains at the same TC is captured by the residual terms in the models rather than by these mean-curve intervals.

Taken together, the analyses in Section 4.4.1 show how timetable design and driver shut-down decisions translate into total engine-running time during passenger layovers. The logistic models quantify how the probability of switching the engine off rises with TC; the Bayesian regressions then show how, when shutdowns occur, arrival and departure idle periods lengthen with additional slack in different ways at Connolly and Heuston. Combining these components provides simple curves of expected total idling versus TC, which reveal how much of each extra 10–30 minutes of timetable slack is typically converted into idling rather than genuine engine-off time. This provides a behavioural explanation for the heavy-tailed idle distributions in Section 4.3 and, more importantly, gives a policy-relevant link from timetable decisions to idle minutes that can be used in the mitigation scenarios in Section 4.5.

4.5 Strategies for reducing passenger-train idling

The combining models developed in Section 4.4 describe how TC affects both the probability that an idle period is interrupted and the duration of arrival and departure idling at each terminus. They also form the basis for quantifying how much idling can be reduced by changing TC and by influencing driver decisions through training and reminders.

4.5.1 Optimising timetables to reduce idling

To provide a simple, policy-relevant summary, the effect of reducing timetable slack is evaluated near the station-specific thresholds identified in Section 4.4.3.1. For each station, the modelled expected total idling is compared between the breakpoint TC and a connection that is 10 minutes shorter. All other aspects of the models are held constant, so the difference reflects only the impact of a shorter timetable gap on interruption probability and on arrival and departure idling.

At Connolly, reducing TC from 80 minutes to 70 minutes leads to a mean reduction in expected total idling of about 5 minutes per affected connection. The posterior mean reduction is approximately 5.0 minutes, with a 95% credible interval from about 3.0 to 7.2 minutes. This means that cutting 10 minutes of slack in this part of the timetable does not translate into a one-to-one reduction in idling. Roughly half of the removed slack becomes genuine engine-off time, while the remainder is absorbed through changes in interruption behaviour and shifts in how idling is divided between the arrival and departure phases.

At Heuston, reducing TC from 60 minutes to 50 minutes has a stronger effect. The models predict a mean reduction in expected total idling of around 8.4 minutes per connection, with a 95% credible interval from roughly 7.1 to 9.8 minutes. Here the reduction in idling is close to proportional to the reduction in TC. This reflects the combination of a high probability of shutdown for long connections and a strong TC effect on departure idling: when TC is shortened, both the opportunity to idle and the likelihood of spending the whole gap with the engine running are reduced.

4.5.2 Training and reminders for driver behaviour

The regression results distinguish two TC conditions. When TC is below the station-specific threshold, tightening or relaxing the layover has a clear effect on idling: changes in TC shift the interruption probability and the lengths of arrival and departure idling in a predictable way. When TC is above the threshold, this relationship becomes weak, and variation in total idling is dominated by how individual drivers use the available slack rather than by the timetable itself. This pattern is consistent with broader evidence from road-transport and fleet operations, where driver behaviour is a key determinant of engine idling alongside schedule design and technical constraints (Lalot et al., 2025; Shancita et al., 2014b). In the rail context, guidance on diesel emissions in stations and depots similarly stresses that technical measures such as automatic shut-down systems and auxiliary power units need to be supported by clear operating procedures and staff engagement (Rail Delivery Group, 2021; U.S. Environmental Protection Agency, 2025).

Behavioural interventions in road-transport fleets show that targeted information, feedback and training can reduce unnecessary idling, with reported reductions of around 5–10% or more when programmes are actively maintained (Abrams et al., 2021; Lalot et al., 2025; Sigurjonsdottir et al., 2022). Although these studies focus mainly on car, van and truck drivers, they highlight mechanisms that are also relevant to rail operations: simple “switch off when waiting” messages, personalised feedback on idling performance and occasional prompts or social-norm cues (Abrams et al., 2021; Lalot et al., 2025). Anti-idling campaigns in construction and other fleet settings likewise stress awareness-raising, on-site reminders and routine reporting of fuel use as practical tools to support behaviour change (Centre for Low Emission Construction, 2019; U.S. Environmental Protection Agency, 2025).

In this study context, the high-TC regime suggests that further timetable optimisation alone is unlikely to deliver reliable reductions in idling once layovers are already long. Practical mitigation therefore needs to focus on how drivers use the available slack. A first step is to define clear, simple rules for long layovers, for example specifying that engines should be shut down when TC exceeds a station-specific threshold and restarted only a few minutes before departure, with explicit operational exceptions (such as very cold weather or known coupling moves). These rules can be reinforced through periodic training and briefings that explain why idling matters for air quality around the station, supported by simple visual reminders at platforms or in cabs. OTMR-derived idling metrics can then be used to provide regular feedback at depot or driver level, for example by summarising average idling times in different TC bands and highlighting good practice. Together, these measures offer a feasible way to shift behaviour in the high-TC towards the lower end of the modelled idle-time distributions, without major changes to rolling stock or infrastructure.

Driver training and reminders should be supported by clear operational safeguards. Guidance should specify when engine shutdown is appropriate, taking account of layover length, restart reliability, brake air pressure, battery charge, auxiliary loads and passenger comfort. In the longer term, soft stop-start systems and platform electrical connections could reduce the reliance on driver judgement alone, while maintaining timetable reliability.

In parallel with these behavioural measures, there is also a scope to reduce idling at Connolly and Heuston through targeted vehicle modifications, drawing on the technical options outlined in Section 4.1.3. For the existing diesel fleet, the most realistic short- to medium-term measures are retrofits that allow the main engines to be shut down while maintaining cab comfort and onboard services, such as automatic engine shut-down/start-up systems, direct-fired heaters and small auxiliary power units. Where platform layout and electrical capacity permit, providing shore power at selected layover locations would further reduce the need for engines to idle solely to supply “hotel” loads. In practice, any combination of these technologies would need to be evaluated for cost, reliability and compatibility with Irish Rail’s current rolling stock, but the model results suggest that, once long layovers are unavoidable, such technical measures are likely to be an important complement to behavioural interventions in cutting extended idling at the two Dublin stations.

4.6 Conclusion

This chapter analysed diesel train idling of Class 22000 DMUs at Dublin Connolly and Heuston using detailed movement and timetable data. It defined idling and TC, linked them to the air quality results in Chapter 3, and described the distributions of idle periods by station, platform zone and train category. The results showed heavy-tailed idle times, shorter and more concentrated idling at Connolly’s through platforms, longer layovers at Heuston and Connolly terminal platforms, and clear differences between passenger and non-passenger movements.

The chapter then focused on passenger services that change the Train ID. It modelled the probability of an idling interruption as a function of TC and showed that this probability increases with TC at both stations. For interrupted periods, it described how total, arrival and departure idling vary with TC and identified approximate TC thresholds of about 80 minutes at Connolly and 60 minutes at Heuston, within which extra TC is associated with longer total idling.

Using Bayesian regression models, the chapter linked TC to arrival and departure idling in the tight-timetable range. It then combined these results with the interruption models to calculate expected total idling as TC changes and to estimate the effect of reducing TC near the station-specific thresholds. Finally, it outlined timetable changes and driver training or reminders as practical ways to reduce unnecessary idling at two stations.

Taken together, these analyses show that timetable design and driver behaviour play distinct roles in controlling idling. Within the tight-timetable range, TC is an effective lever: modest reductions in slack can provide predictable reductions in expected total idling, especially for services close to the identified thresholds. Beyond these thresholds, variation in idling is largely driven by how drivers use the available layover time, so behavioural and technical measures become more important than further timetable compression. These findings are carried forward into Chapter Five, which uses detailed OTMR records, notch–power maps and the station zoning from Chapter Three to convert train movements and layovers into high-resolution emission inventories for NOx and $PM_{2.5}$ at Connolly and Heuston. The idling patterns and TC regimes identified here provide the temporal structure and operational context for interpreting those emission fields and for assessing how changes in fleet mix, operating practice and fuel choice translate into local air-quality impacts.

5. High-Resolution Emission Calculation for Diesel Train in stations

5.1 Introduction

Previous chapters established a link between air pollution peaks and train activities (Chapter 3) and measured the duration of idling events (Chapter 4). However, analysing time alone is not enough to determine environmental impact. The relationship between engine activity and emissions is non-linear; emission levels differ significantly between idling, coasting, and acceleration modes (Ma et al., 2020; J. Tu et al., 2013; Zhai et al., 2020). For the Irish rail system, only emissions estimates based on fuel consumption are available(Irish Rail, 2023). This estimate cannot accurately allocate emissions to the station area.

This chapter shifts the focus from duration to mass quantification. A physics-based exhaust calculation method is developed to convert operational data into mass emissions of NOx and PM2.5 in the station area. These outputs serve as the source term for the air dispersion model in Chapter 6.

5.1.1 Why a physics-based approach is needed

Regional inventories typically use distance-based emission factors (g/km) (da Fonseca-Soares et al., 2024; Fridell et al., 2010b; Fuller, 2014; Perdikopoulos et al., 2025). This approach assumes pollution scales linearly with distance, an assumption that fails in station environments where trains are frequently stationary but still emitting. A distance-based model incorrectly assigns zero emissions to a stopped train (Rail Safety and Standards Board, 2020). This estimate cannot accurately allocate emissions to the station area.

In reality, a stationary train operates under load to power auxiliary systems, generating continuous emissions (Biess et al., 2009; Mallemoggala et al., 2024). Furthermore, transient events, such as accelerating a heavy train from a standstill, generate sharp pollution pulses that can exceed steady-state cruising emissions (Kim et al., 2020; Yuan & Frey, 2021). Capturing these dynamics requires a formulation linked to engine power and time rather than distance alone (Rail Safety and Standards Board, 2020). An energy-based approach (g/kWh) is therefore more appropriate for station-scale analysis.

5.1.2 Data reality across the Irish fleet

The fleets operating at Heuston and Connolly differ in age, engine type, and onboard recording systems. Modern Class 22000 multiple units provide rich OTMR telemetry, including engine speed, torque, speed, GPS, and fuel flow, enabling high-resolution energy and emission estimates. Older fleets provide sparser data: Class 29000 units record speed and notch signals but no engine speed; Class 201 locomotives at Heuston have no accessible onboard telemetry in scope for this study.

These differences mean that a single calculation method cannot be applied uniformly. Instead, a flexible framework is used, applying high-precision power-based calculations where OTMR supports them and robust proxy methods where only partial data or timetables exist. In the absence of indigenous emission tests for the Irish fleet, technically similar UK fleets are used as proxies, matched by engine type, transmission, and build era.

The remainder of this chapter is organised as follows. Section 5.2 describes the fleet data, the energy-based computational framework, the tiered calculation methods, and key sources of bias. Section 5.3 presents station totals, spatial patterns, diurnal profiles, cycle-level characteristics, and the HVO fuel scenario results. Section 5.4 discusses the implications of these findings in the context of national inventories, local hotspots, mitigation options, and limitations. Section 5.5 summarises the main conclusions and links to the dispersion modelling in Chapter 6.

5.2 Methods: Data Processing and Calculation Framework

5.2.1 Fleet information and data availability

The analysis covers the diesel fleets that work the platforms at Heuston and Connolly during the monitoring periods. OTMR content differs by class and determines which calculation tier can be applied.

At Heuston, Class 22000 provides a complete telemetry set for 25 May–26 June 2023: gear/notch (HD3), engine RPM and torque, GPS position, two speed channels (plus HVAC speed), charge-air pressure and temperature, fuel flow, and engine/auxiliary indicators. These signals allow notch-resolved, power-based estimates within platform polygons and the 300 m approach buffer. Class 201 + Mk4 has no onboard telemetry in scope; only arrival and departure times are available, so movements are represented from timetables using synthetic drive cycles calibrated from Connolly evidence.

At Connolly, Class 22000 provides the same complete telemetry set as Heuston for 1 March–31 May 2024. Class 201 records engine-run status, per-notch "Idle/1–8" flags, GPS position and speed; RPM is not recorded. Class 29000 records three binary "notch wires" (8–10), speed, brake-handle position, GPS position, and engine warning lines; RPM is not recorded. The notch-wire pattern is decoded to an integer notch before applying power-based factors. DD/09 are unpowered De Dietrich coaches (as with Mk4) and contribute no traction emissions; where present they are treated as coach stock within locomotive-hauled consists.

Because no dedicated emission tests exist for the Irish fleets, proxy emission factors were selected from UK fleets based on engine technical specifications, transmission types, and build years. Selection prioritised engine architecture (which governs combustion chemistry and g/kWh emission rates), followed by transmission type and era.

Class 22000: Uses MTU 6H 1800 engines with hydraulic transmission. The UK Class 172, fitted with the identical MTU 6H 1800 engine (mechanical transmission), was selected as the proxy to match combustion behaviour.

Class 201: Shares the EMD 12-710 engine and diesel-electric transmission architecture with the UK Class 66. The UIC1/UIC2 variant closest in build date to the Class 201 was used as a proxy.

Class 29000: Uses a MAN D2876 engine with no direct rail test data. The UK Class 170, a contemporary diesel-hydraulic DMU with a similar heavy-duty six-cylinder engine, was selected as the closest equivalent.

A summary of fleet, OTMR availability, and proxy assignments is presented in Table 5.1.

Table 5.1 Fleet and data overview (Heuston and Connolly)

Irish Fleet	OTMR	Engine	Transmission	Date built	Emission factor
Class 22000 (ICR DMU)	Heuston (25 May–26 Jun 2023): gear/notch (HD3), RPM, torque, GPS, speed (incl. HVAC), charge-air P/T, fuel flow, engine/aux indicators. Connolly (1 Mar–31 May 2024): same complete set as Heuston.	MTU 6H 1800 R83 (one per car)	Voith-T-312 R hydraulic	2007–2011; 2019–2022	Proxy: UK Class 172 notch-based g/kWh
Class 201 + Mk4/DD09	Heuston: timetables only (no OTMR). Connolly (1 Mar–31 May 2024): engine-run, per-notch flags (Idle/1–8), GPS, speed (no	EMD 12-710G3B (JT42HCW)	Diesel-electric (DC traction)	1994–1995	Proxy: UK Class 66 g/kWh UIC2 (primary);

Irish Fleet	OTMR	Engine	Transmission	Date built	Emission factor
	RPM).				bench ppm/PN for sensitivity only
Class 29000 (DMU)	Connolly (1 Mar–31 May 2024): notch wires 8–10, speed, brake-handle, GPS, engine warning lines (no RPM).	MAN D2876 LUH01 (per power car)	Diesel-hydraulic (Voith)	2002–2005	Proxy: UK Class 170 notch-based g/kWh

The emission-factor proxy assignments in Table 5.1 were based on the RSSB CLEAR methodology and associated rail traction emission-factor data (Rail Safety and Standards Board, 2020). The detailed notch- and load-based emission-factor tables used in the calculations were provided by RSSB under a confidential data-sharing agreement.

5.2.2 Computational framework

This study employs an energy-based approach (g/kWh) aligned with the Rail Safety and Standards Board (RSSB) CLEAR methodology for estimating rail traction emissions. CLEAR (Carbon Low Emissions Assessment and Reporting) expresses emissions as grams of pollutant per kilowatt-hour of engine output and then multiplies these factors by the time-integrated engine power over each movement or dwell (Rail Safety and Standards Board, 2020). In practice, this means that emission estimates are based on the instantaneous power demand reconstructed from notch or throttle settings and speed, integrated over the duration of the event, rather than on simple distance- or time-based averages.

The total mass emission E for a specific pollutant p over a defined train movement is calculated by integrating the power output over time:

Equation 5.1:

$$\boldsymbol{E_p = \sum_{t=0}^{T} \big(P_{set}(t) + P_{aux}(t)\big) \Delta t EF_{p,state}(t)}$$

Where:

- E_p: Total mass of pollutant p emitted (g)
- P_{set}: Whole DMU set power out

- P_{aux}: Auxiliary "Hotel" load (kW) required for onboard systems such as HVAC, lighting, and door controls. This load is critical during station dwell times

- Δt: Duration of the timestep (hours), derived from the frequency of the OTMR or timetable records

- $EF_{p,state}$ The specific emission factor (g/kWh) for pollutant p corresponding to the engine's operating mode. In Tier 1, $EF_{p,state}$ is indexed by load fraction $\theta = Peng/Pmax$; in Tier 2 or 3, it is indexed by $notch\ n \in \{0, \dots, 8\}$

Different calculation methods are assigned to each fleet based on the availability of specific variables:

Tier 1 – Direct power-based calculation. Complete OTMR (RPM, torque or equivalent load signal). Applied to Class 22000 at Heuston and Connolly.

Tier 2 – Notch-based reconstruction. No RPM; power recovered from a notch to the power map using load-fraction curves and decoded notch. Applied to Class 29000 and Class 201 (Connolly).

Tier 3 – Synthetic drive cycle. No OTMR; timetable-based station cycle derived from Connolly behaviour. Applied to Class 201 (Heuston).

Table 5.2 lists the assigned tier for each fleet and station.

Table 5.2 Tier assignment for each class

Station	Fleet class	Key signals available	Assigned tier
Heuston	Class 22000	RPM, torque/load, notch/controls, speed, GPS	Tier 1
	Class 201	Timetable only	Tier 3
Connolly	Class 22000	RPM, torque/load, notch/controls, speed, GPS	Tier 1
	Class 29000	Binary notch wires, speed, GPS (no RPM)	Tier 2
	Class 201	One-hot notch flags, speed, GPS (no RPM)	Tier 2

5.2.3 Tier-based emission calculation

5.2.3.1 Tier 1 Direct power-based calculation (data-rich fleets)

This tier applies where station area OTMR provides high-frequency engine speed, engine torque, power-controller wires (TDA/TDB/TDC), speed, and GPS (e.g., Class 22000 at both stations). The method converts measured engine load into instantaneous power, maps that load to engine-basis emission factors, and integrates over time to obtain mass emissions. Because the logging stream typically comes from a single vehicle within a multi-car set, the results are upscaled to the full train set level.

Records are time-ordered and converted to step durations Δt. To avoid inflating totals due to station-only logging gaps, long intervals are removed with a gap cap ($\Delta t \leq 5$ s). Notch wires are decoded (N, 1–5) for diagnostics; the calculation itself uses measured power. GPS is retained for mapping; rows without valid coordinates contribute to the time series but not to maps. Timestamp parsing handles naive time-zones and standardises to Dublin.

Instantaneous engine power is computed from torque and engine speed,

Equation 5.2:

$$P(t) = \frac{Torque(t) \times PRM(t)}{9549} \qquad [kw]$$

constrained to non-negative values. Load is expressed as a fraction of the rated engine power, $\theta(t) = \min\{1, \max[0, P(t)/P_{max}]\}$, where P_{max} is taken from the engine's certified rating.

For each pollutant $x \in \{NO_x, PM_{2.5}\}$, the engine-basis curve $EF_x(\theta)$ in g/ kWh is used, interpolated over $\theta \in [0,1]$. This captures high-idle and transient operation without assuming a fixed notch-to-load ratio.

Instantaneous rates follow directly from power and the factor:

$$\dot{m}_p(t) = EF_p[\theta(t)]\frac{P(t)}{3600} \qquad [g/s]$$

Mass over each step is $m_{p,i} = \dot{m}_p \Delta t_i$. Summing across steps provides totals for any interval (event, hour, or day).

When only one vehicle stream is logged for a unit, step masses are multiplied by the car count (one engine per car) to obtain set-level emissions. If multiple vehicles for the same unit are present at the same timestamp, their observed powers are summed, and no further scaling is applied. This prevents both under-counting and double counting.

5.2.3.2 **Tier 2 Notch-based reconstruction with proxy emission factors (Class 29000; Class 201 at Connolly)**

Tier 2 is used when the OTMR records traction state but not engine RPM or torque. The recorded traction commands are first converted to a single integer notch per time step. For Class 29000 and Class 201 units, three “notch wires” are decoded using a predefined truth table.

Power is then reconstructed from a validated hydraulic-DMU notch to power map at an engine basis. For a contemporary UK diesel-hydraulic proxy (e.g., Class 170), typical magnitudes are idle ≈ 26 kW per engine and top notch ≈ 315 kW per engine, with NO_x falling from ≈ 15.4 g/kWh at idle to ≈ 2.7 g/kWh at high notch, and PM in the ≈ 0.07–0.24 g/kWh range. Set power is the per-engine value multiplied by engines in the consist; for Class 29000, only the two driving cars are motored, so $N_{engines}$=2. Step durations use timestamp differences with the same short cap. When speed is ≤ 1 km/h, each time step is treated as dwell and the traction notch is set to 0 unless an explicit auxiliary indicator is present; this prevents overestimating power when a low notch is briefly commanded at a stand.

It should be noted that the notch-based NO_x values are expressed as specific emission factors in g/kWh, rather than absolute emission rates in g/s or g/h. The idle notch may show a relatively high g/kWh value because the engine produces very low useful power while still emitting NO_x. This does not mean that idling emits more NO_x per unit time than the highest power notch.

Mass rates follow directly:

Equation 5.3:

$$\dot{m}_p(t) = EF_p[notch(t)]\frac{P(t)}{3600} \qquad [g/s]$$

Emission is calculated separately for NOx and PM, and subsequently aggregated to quantify the total mass (in grams) per time interval, per hour, and per day within the defined station boundaries.

For NOx and PM separately, the total emitted mass was calculated as $Ep = \sum \dot{m}p(t)\Delta t$ and summarised per interval, per hour, and per day within the station boundary.

5.2.3.3 Tier 3 Timetable-based synthesis with learned drive cycles (Class 201 at Heuston)

Tier 3 is used when no OTMR data are available for a station, and only the working timetable is known. In this case, a realistic traction profile for each train service is generated by learning a template from the Class 201 OTMR data at Connolly (Tier 2) and then applying this template to Heuston.

First, the weekly pattern (Monday–Friday / Saturday / Sunday) is expanded to the modelling period, and arrivals are paired with the next departure on the same day. Each arrival–departure pair defines one station event with a known dwell. At Connolly, OTMR records are then used to learn typical approach–dwell–departure cycles. Dwells are detected inside a small station bounding box using a low-speed threshold and a minimum stationary duration. Around each dwell, three segments are extracted: (i) an approach segment before the stop, (ii) a short high-idle period immediately before brake release while stationary, and (iii) a departure segment after the stop. For each segment, the notch time series is collapsed to a simple ladder of consecutive notches with associated durations. A template cycle is then constructed by taking the median of these ladders across events (mode notch at each step with median duration), and by using median window lengths for the approach and departure segments and the median duration of the final contiguous high-idle period before departure. This procedure provides a compact “approach–dwell-idle–pre-depart–depart” cycle for each station service that reflects observed passenger operation rather than assumed settings. The main steps of this workflow are summarised in Figure 5.1 as a flow chart.

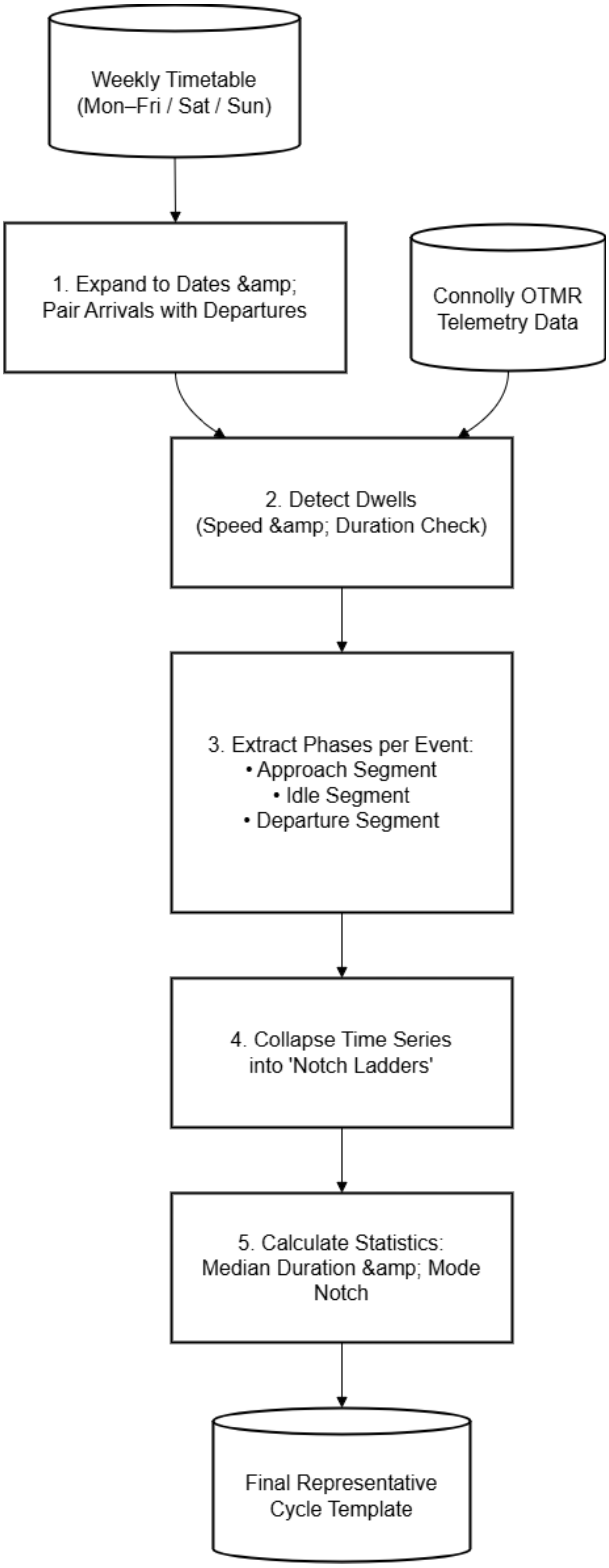


Figure 5.1 Workflow for reconstructing approach–dwell–departure cycles from timetable and OTMR data

For Heuston, which is a terminal station, each timetable service is mapped onto this traction–idling cycle. If the scheduled dwell time is shorter than the learned pre-departure duration, the pre-departure segment is truncated and the remaining dwell is treated as true idle; if only an arrival or only a departure is present, only the relevant parts of the cycle are applied. Each time step has a fixed notch and duration, so set power is obtained from the Class 66 proxy notch-to-power table (engine basis, auxiliaries included) and multiplied by the number of engines (one for Class 201). Instantaneous and cumulative emissions are then calculated using the Tier 2 formula.

5.2.4 Data outputs and quality assurance

Each tier pipeline produces four types of outputs:

(i) per-step records of emissions, power, duration, and coordinates;

(ii) hourly totals and daily average hourly profiles;

(iii) 10-minute average emission rates;

(iv) station-area emission maps obtained by summing step-level masses on a spatial grid.

Quality checks include verifying that peak set power does not exceed the rated multi-car maximum, confirming that the Δt cap is active, and checking that higher decoded notches correspond to higher measured power. Map coverage is compared with the proportion of records containing valid GPS coordinates, and obvious outlier emission entries are removed from the final output database.

5.2.5 Zoning, temporal aggregation, and preprocessing

Emissions are reported by location using the zoning introduced in Chapter 3. Connolly has five zones: Region 1–3 inside the station, Region_Depot at the maintenance and sidings yard south-east of platform 1, and Region_Line for adjacent railway approaches. Heuston has four zones: Regions 1–3 for platforms, plus Region_Line. At Heuston, regional results include only Class 22000; Class 201 lacks GPS and is counted in station totals but not split by zone.

Region_Depot and Region_Line close two gaps in coverage. The depot hosts trains that idle during maintenance and turn-back, while the line zone captures approach and departure emissions immediately outside the platforms where acceleration, queuing, and short dwells occur. Without these, emissions would “leak” beyond Regions 1–3, and the inventory would understate the near-station emissions.

The time windows reflect OTMR availability. Connolly uses 1 March–31 May 2024 for Classes 22000, 29000, and 201. Heuston uses 25 May–26 June 2023 for Class 22000 OTMR; Class 201 is timetable-based only and appears in station totals.

GPS is logged at the leading cab. For distributed-power DMUs, each point is shifted to the train centre before assigning zones (half the set length for Class 22000; the same treatment for Class 29000 with engines in the two driving cars). This prevents a bias toward buffer stops and sharpens the split between interior and line areas.

Idle status follows the recorded data. The is_idle flag is used where present (Classes 22000 and 29000). At Connolly, Class 201 uses its notch stream to classify idle. At Heuston, Class 201 has no GPS recording, so it is excluded from zoned tables; its idle share is handled only in station-level totals via the learned cycle.

All inputs are harmonised to one-minute sums in grams per minute and then aggregated to hourly totals and weekday or weekend daily average hourly profiles. Hours and days with no records are treated as zero so averages cover the full station clock. These zoned aggregates feed both the temporal analyses in this chapter and the dispersion modelling in Chapter 6.

5.2.6 Bias and sources of deviation

This emission calculation method can differ from reality in three ways: (i) proxy emission factors may not match the Irish fleets; (ii) Tier 2 power rebuilt from notches may not match real power; (iii) Tier 3 station cycles may not match what actually operations.

5.2.6.1 Proxy emission factors vs. actual fleets

Proxy g/kWh values may not reflect the real engines. Reasons include different engine tuning, age, or maintenance; different after-treatment; different fuels or weather; and different test cycles. Auxiliary loads (HVAC, doors, compressors) also vary by place and season, so emission factors from a proxy may not fit local practice.

In this study, the OTMR telemetry windows (Heuston: late May–June 2023; Connolly: March–May 2024) exclude the coldest winter months. As a result, the calculated emissions may underestimate annual-average values, because they do not capture the additional idling needed for winter heating (HVAC) or any extra fuel use associated with cold starts. The monitoring campaigns were conducted in sequential phases, leading to non-overlapping observation periods at Heuston and Connolly. Longer or aligned OTMR extracts could not be obtained, despite formal requests to the railway operator, because older records were subject to archiving limits and data-handling constraints. However, the underlying fleet composition, timetables and engine-management protocols remained stable over the study period, so the emission profiles derived from these spring/early-summer snapshots are considered representative of standard operating conditions, while recognising that they reflect relatively favourable (non-winter) conditions rather than a worst-case winter scenario.

5.2.6.2 Tier 2 power maps vs. actual power

Rebuilding power from notches can misrepresent real loads if the proxy map uses a different engine rating or transmission. Low-speed behaviour, derating rules, and controller logic (eco-notch, creep, anti-slip) all affect load at a given notch. Notch wiring can differ across fleets; decoding errors shift the assigned notch. A fixed hotel-load baseline may be too high or too low for real dwell conditions. Gradients, curves, and operational manoeuvres (e.g. coupling) also change load at the same command.

5.2.6.3 Tier 3 reconstructed cycles vs. actual cycles

Track gradient is another source of uncertainty in the reconstructed emission cycles because it affects traction power during arrival and departure. At Heuston, the western approach falls from west to east towards Heuston Station at approximately 1.3%; therefore, arrivals into Heuston are downhill, while departures from Heuston are uphill over this section (Irish Rail, 2020). At Connolly, the track gradient is 8.3‰. For the movements considered in this study, the approach into Connolly is uphill, while the exit from Connolly is downhill.

Where direct OTMR power or notch data were available, these gradient effects were partly reflected in the recorded engine load. However, for timetable-based reconstructed cycles, gradient was not modelled separately. This may lead to some uncertainty in the estimated emissions, with possible underestimation during uphill acceleration and overestimation during downhill movements. This limitation is most relevant for reconstructed station entry and exit cycles, rather than for idling periods where gradient has little direct effect on engine load.

5.3 Results: Station-area emissions

5.3.1 Station totals and spatial patterns

Hotspot maps for Connolly and Heuston (Figure 5.2, Figure 5.3) show average daily NOx and $PM_{2.5}$ across the station-related area (Regions 1–3, Region_Line, and Region_Depot where applicable). Emissions from the step-level data are summed onto a 1 m × 1 m grid and averaged over the study windows. Colours represent mean grams per day per grid cell. $PM_{2.5}$ hotspot maps are provided in Appendix 3; their spatial patterns closely follow those for NOx.

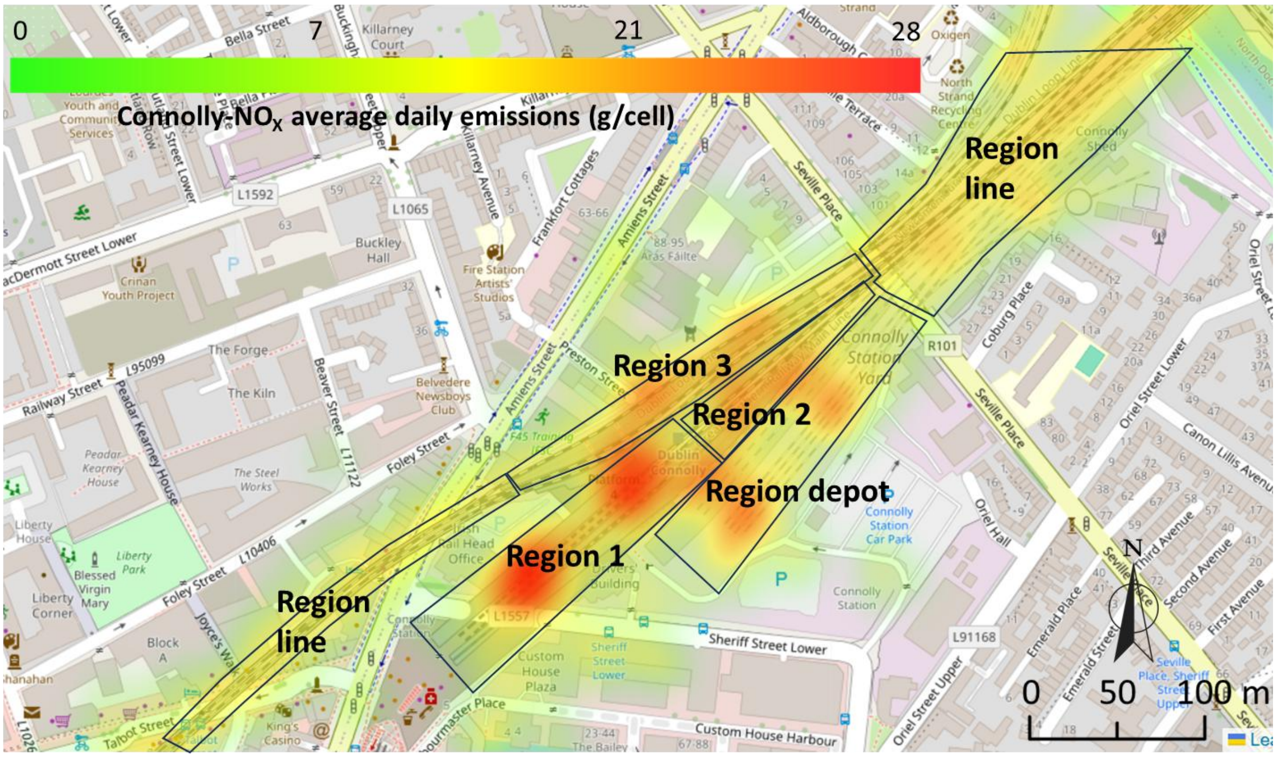


Figure 5.2 NOx heatmap at Connolly

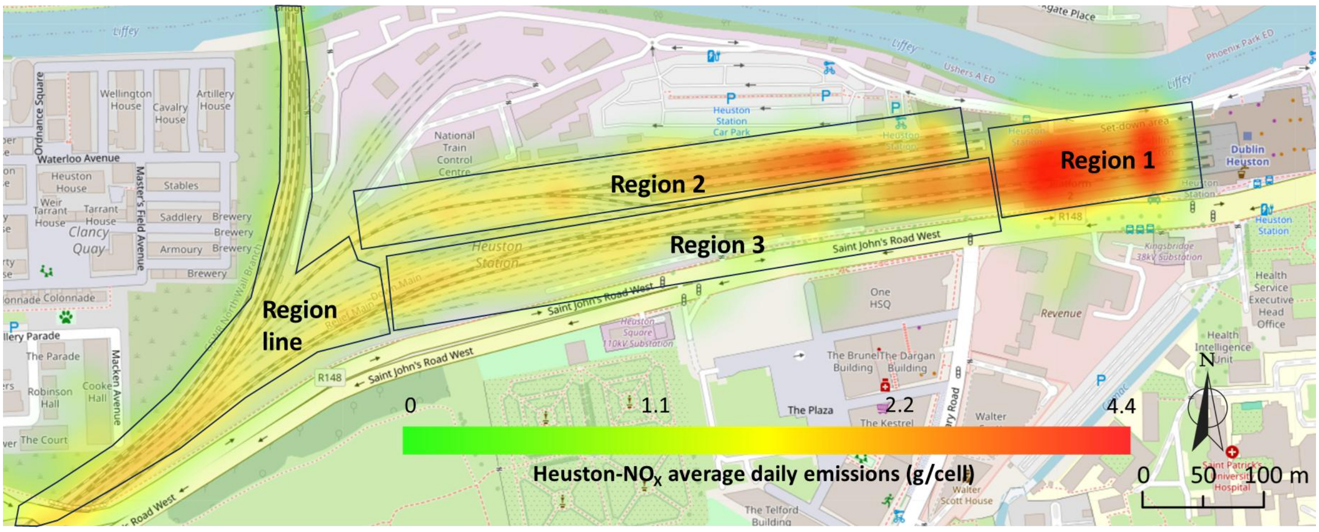


Figure 5.3 NOx heatmap at Heuston

Over 1 March–31 May 2024, Connolly's station-related area emits 7.04 t NOx and 125.8 kg $PM_{2.5}$ (Table 5.3). Over 25 May–26 June 2023, Heuston emits 0.54 t NOx and 20.2 kg $PM_{2.5}$ from Class 22000 in the station-related area, with a further 0.42 t NOx and 13.1 kg $PM_{2.5}$ from Class 201 in the whole-station totals (Table 5.3). In terms of daily averages, this corresponds to 75.7 kg NOx day^{-1} at

Connolly and 16.3 kg NOx day^{-1} at Heuston over the station-related area.

Table 5.3 Station totals and idle shares by class

Station	Class	Hours with data	Idle time (%)	NOx (t)	NOx share of station (%)	$PM_{2.5}$ (kg)	$PM_{2.5}$ share of station (%)	Time share of station (%)
Connolly	201	874.2	84.4	1.485	21.1	46.9	37.3	9.6
Connolly	22000	2906.4	60.0	0.555	7.9	21.0	16.7	31.9
Connolly	29000	5334.7	80.2	5.001	71.0	57.9	46.0	58.5
Heuston	201	253.8	27.4	0.417	43.7	13.1	39.3	8.7
Heuston	22000	2669.8	59.8	0.538	56.3	20.2	60.7	91.3

For comparisons with national and external models, only the station areas (Regions 1–3) are considered. Within these zones, summed emissions from all classes amount to 4.02 t NOx and 71.8 kg $PM_{2.5}$ at Connolly, and 0.33 t NOx and 13.1 kg $PM_{2.5}$ from Class 22000 at Heuston (Table 5.4 and Table 5.5). Treating the Connolly and Heuston monitoring windows as representative of a full year implies station-area NOx totals of 15.8 t/yr and 3.63 t/yr, respectively. Together they give 19.4 t NOx/yr, which is about 1% of the 1.94 kt NOx/yr reported nationally for railways (N. Rahman et al., 2025). This is a lower bound, because the station-area contribution from Class 201 at Heuston is not zoned.

Heuston is also represented in the national air-quality modelling. The MapEire inventory assigns 1.5 t NOx/yr, 0.04 t $PM_{2.5}$/yr and 0.04 t PM_{10}/yr to the grid square covering the uncovered section of the station, represented as a 3 m-deep volume source (EPA, 2023). Using only the station-area emissions from Class 22000 at Heuston and scaling the 33-day window to a full year gives 3.63 t NOx/yr (Table 5.2), already more than double the MapEire value. Adding the locomotive contribution from Class 201 would increase this further.

The regional split (Table 5.4) shows where within each station-related area emissions are concentrated. At Connolly, Region_Line carries 2.13 t NOx (30.3%) and 39.9 kg $PM_{2.5}$ (31.7%). Region 2, the central platforms, carries 1.62 t NOx (23.0%) and 38.9 kg $PM_{2.5}$ (30.9%). Region 1 and Region 3 together account for 2.40 t NOx (34.1%) and 33.0 kg $PM_{2.5}$ (26.2%). Area 3, which serves as the pass and electric train platform, has much lower emissions than the other areas.

Region_Depot adds 0.89 t NOx (12.6%) and 14.1 kg $PM_{2.5}$ (11.2%). In the Connolly emission heatmap (Figure 5.2), this appears as a strong band at the platform throat, with secondary hotspots at the busiest platform faces in Regions 1–2 and around the stabling and turn-back area in Region_Depot.

Table 5.4 Regional split at Connolly (all classes)

Region	NOx (t)	NOx share of station (%)	$PM_{2.5}$ (kg)	$PM_{2.5}$ share of station (%)
R1	1.550	22.0	21.4	17.0
R2	1.616	23.0	38.9	30.9
R3	0.852	12.1	11.5	9.1
R_DEPOT	0.889	12.6	14.1	11.2
R_LINE	2.133	30.3	39.9	31.7
Total	7.041	100.0	125.8	100.0

At Heuston, the regional split (Table 5.5) is based on Class 22000 only, because Class 201 has no GPS and is included only in the station totals. For Class 22000, Region_Line carries 0.209 t NOx (38.9%) and 7.17 kg $PM_{2.5}$ (35.4%). Region 1 carries 0.131 t NOx (24.3%) and 5.22 kg $PM_{2.5}$ (25.8%). Region 2 carries 0.139 t NOx (25.8%) and 5.51 kg $PM_{2.5}$ (27.2%). Region 3 carries 0.059 t NOx (11.0%) and 2.33 kg $PM_{2.5}$ (11.5%). The Heuston emission heatmaps (Figure 5.3) describe the spatial pattern of DMU emissions, with the largest hotspots centred on the under-roof platforms (Region 1).

Table 5.5 Regional split at Heuston (Class 22000 only).

Region	NOx (t)	NOx share of 22000 (%)	$PM_{2.5}$ (kg)	$PM_{2.5}$ share of 22000 (%)
R1	0.131	24.3	5.2	25.8
R2	0.139	25.8	5.5	27.2
R3	0.059	11.0	2.3	11.5
R_LINE	0.209	38.9	7.2	35.4
Total	0.538	100.0	20.2	100.0

The class breakdown in Table 5.3 explains much of the difference between stations. At Connolly, Class 29000 provides about 59% of the observed hours but 71% of NOx and 46% of $PM_{2.5}$. Class 201 appears for only about 10% of hours but still contributes 21% of NOx and 37% of $PM_{2.5}$. Class 22000 covers about 32% of hours but only 8% of NOx and 17% of $PM_{2.5}$, consistent with a cleaner DMU type. At Heuston, Class 22000 dominates the timetable, with around 91% of hours and about 56% of NOx and 61% of $PM_{2.5}$, while Class 201 contributes 9% of hours but 44% of NOx and 39% of $PM_{2.5}$.

In terms of time on site and idle fractions (Table 5.3), Class 201 and Class 29000 at Connolly spend about 84% and 80% of their time idling, respectively, while Class 22000 idles for about 60% of its time. At Heuston, Class 22000 idles for about 60% of its time, but Class 201 idles for only about 27%. Combined with the higher per-minute emission rates of Class 201, this means that a short 201 presence can contribute a large share of both NOx and $PM_{2.5}$, especially at Connolly.

5.3.2 Diurnal emission profiles

Diurnal profiles are built from the one-minute zoned data by summing to hourly totals in Regions 1–3 and then averaging by hour over all weekdays and all weekends (Table 5.6, Figure 5.4). For Connolly station-area NOx on weekdays, the mean between 00:00 and 04:00 is 1.82 kg/h. The profile peaks at 13:00 with 2.89 kg/h and stays above 2 kg/h through most of the daytime. At weekends, the 00:00–04:00 mean is 1.15 kg/h, with a peak again at 13:00, slightly higher at 2.99 kg/h.

Table 5.6 Summary of diurnal NOx profiles (station area, Regions 1–3 only)

Station	Day type	Mean 00:00–04:00 NOx (kg/h)	Peak hour	Peak NOx (kg/h)
Connolly	Weekday	1.82	13:00	2.89
Connolly	Weekend	1.15	13:00	2.99
Heuston	Weekday	0.12	15:00	0.92
Heuston	Weekend	0.08	16:00	0.74

For Heuston station-area NOx, weekday nights are much lower, with 0.12 kg/h between 00:00 and 04:00. The weekday peak occurs at 8:00 and 15:00 with about 0.9 kg/h. At weekends, the 00:00–04:00 mean is 0.08 kg/h, and the peak shifts to 16:00 with 0.74kg/h. The $PM_{2.5}$ profiles follow the same timing but at lower magnitudes (Figure 5.4). These profiles indicate that a small number of daytime hours contribute disproportionately to station-area emissions.

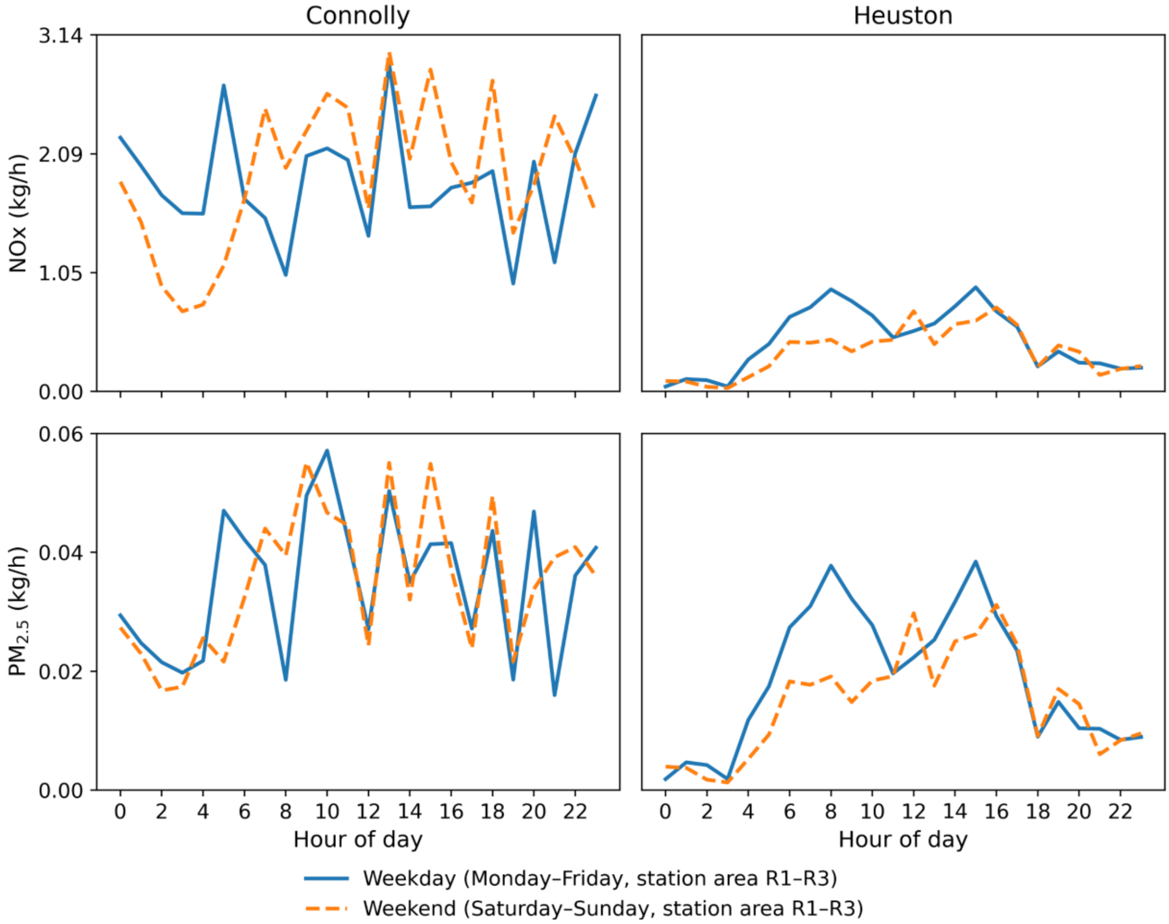


Figure 5.4 Diurnal profiles of NOx and PM2.5 at Connolly and Heuston

5.3.3 Cyclic emission characteristics during station entry, dwell, and exit

Section 5.3.2 showed large differences in per-minute station emissions across train classes. This subsection compares representative approach–dwell–departure cycles at Connolly and Heuston to identify where these class differences arise within the station movement sequence. The Class 22000 examples also serve as Tier 1 calculation (Section 5.2.3) examples for the two stations, showing how measured OTMR power, phase duration and load-dependent emission factors determine the phase-level emissions.

For each station, the reference cycle was taken from a Class 22000 OTMR record. The cycle was cut from the station approach, through the zero-speed dwell period, to the point where the train had left the platform area. The change from moving to stopped was used to mark arrival, and the next change from stopped to moving was used to mark departure. For comparison, Class 201 and Class 29000 emissions were rebuilt on the same time base. The Class 22000 notch sequence was retained, but class-specific notch–power and emission-factor tables were applied. The results were scaled to whole-train level: Class 201 as one locomotive, Class 29000 as a two-engine DMU, and Class 22000 as a four-car set.

At Connolly, the representative cycle consisted of 79.7 s of approach, 93.9 s of dwell and 71.7 s of departure. Table 5.7 shows that the Class 22000 emitted 25.8 g NOx and 0.93 g $PM_{2.5}$ over the complete cycle. The Class 29000 emitted 101.3 g NOx and 1.27 g $PM_{2.5}$, around four times the NOx of the 22000. The Class 201 emitted 216.0 g NOx and 4.39 g $PM_{2.5}$, around eight times the NOx of the 22000.

Table 5.7 Phase-level NOx emissions for the representative Connolly cycle

Class	Total NOx (g)	Total $PM_{2.5}$ (g)	Approach NOx (g, %)	Dwell NOx (g, %)	Departure NOx (g, %)
22000	25.8	0.93	10.1 (39%)	8.1 (31%)	7.7 (30%)
29000	101.3	1.27	33.1 (33%)	36.8 (36%)	31.4 (31%)
201	216.0	4.39	67.4 (31%)	50.9 (24%)	97.6 (45%)

For the Class 22000 and 29000 at Connolly, departure emissions are slightly lower than approach emissions. This is counterintuitive because departure normally requires acceleration from standstill and higher traction power. However, the emission-segment records show that the result is caused by the combined effect of gradient, phase duration and load-dependent NOx emission factors. At Connolly, the approach into the station is uphill, while the exit is downhill. The uphill approach can require some engine load before final deceleration, whereas the downhill exit reduces the traction demand during departure. In addition, the approach segment contains a longer period of low-load, neutral, braking and engine-on operation. These low-load records are assigned higher NOx emission factors, so they can accumulate more NOx even when engine power is lower.

This is shown clearly for the Class 22000 in Table 5.8. Departure used more energy than approach, 5.84 kWh compared with 3.56 kWh. However, the energy-weighted NOx emission factor during departure was much lower, 1.31 g/kWh compared with 2.84 g/kWh during approach. Therefore, the lower departure total does not mean that departure was less power-demanding in physical terms. It means that the higher-power departure operated at a lower NOx factor per unit of work, while the approach accumulated emissions during low-load operation.

Table 5.8 Energy and NOx-factor explanation for the Connolly Class 22000 cycle

Phase	Duration (s)	Mean power (kW)	Energy (kWh)	NOx (g)	Energy-weighted NOx factor (g/kWh)
Approach	79.7	134.5	3.56	10.11	2.84
Dwell	93.9	107.1	2.68	8.06	3.00
Departure	71.7	528.5	5.84	7.67	1.31

The Class 201 behaves differently at Connolly. Its departure emissions are higher than approach emissions, despite the downhill exit, because the locomotive requires much greater traction power during acceleration and has substantially higher absolute emission rates. This creates the steep post-departure increase shown in Figure 5.5.

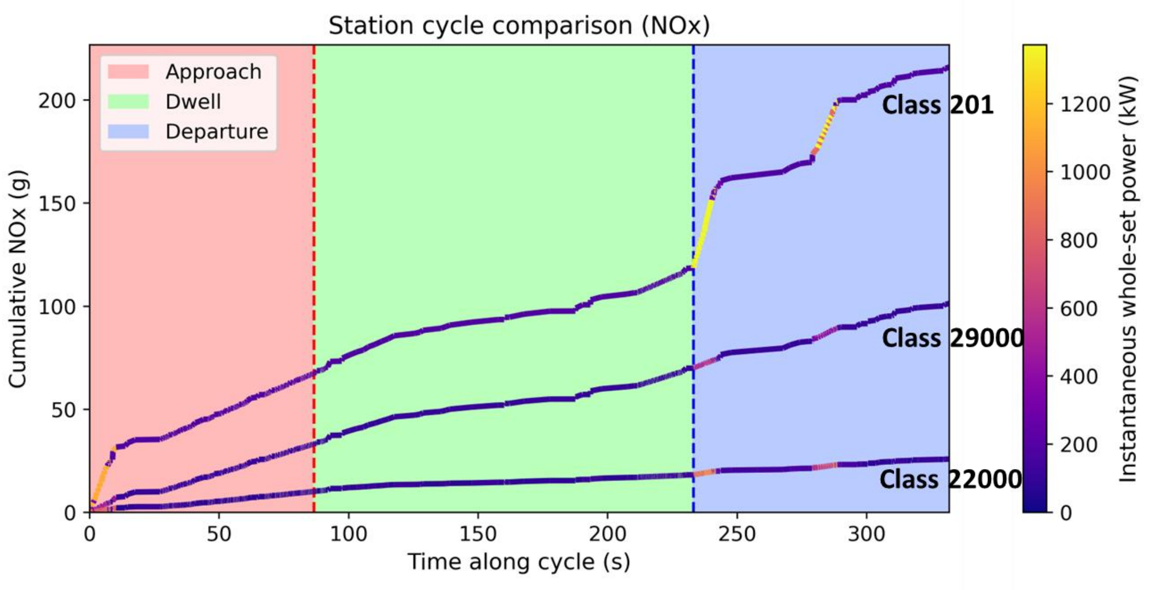


Figure 5.5 Cumulative NOx over the typical approach–dwell–departure cycle for Classes 201, 29000, and 22000 at Connolly

The same analysis was repeated for Heuston. The representative Heuston cycle consisted of 199.55 s of approach, 95.53 s of dwell and 158.46 s of departure. Unlike Connolly, the Heuston approach is downhill and the exit is uphill. This produces a more intuitive phase pattern, with departure contributing more strongly to total emissions.

Table 5.9 shows that the Class 22000 emitted 54.4 g NOx and 1.81 g $PM_{2.5}$ over the complete Heuston cycle. The Class 29000 emitted 197.5 g NOx and 2.48 g $PM_{2.5}$, while the Class 201 emitted 470.2 g NOx and 10.92 g $PM_{2.5}$. For the Class 22000 and 29000, approach and departure NOx contributions are broadly similar, although departure is slightly higher. For the Class 201, departure dominates strongly, contributing 63.8% of NOx and 73.7% of $PM_{2.5}$.

Table 5.9 Phase-level emissions for the representative Heuston cycle

Class	Phase	Duration (s)	NOx (g)	NOx share (%)	$PM_{2.5}$ (g)	$PM_{2.5}$ share (%)
22000	Approach	199.55	19.98	36.8	0.809	44.8
22000	Dwell	95.53	12.53	23.1	0.491	27.2
22000	Departure	158.46	21.85	40.2	0.507	28.0
29000	Approach	199.55	78.14	39.6	0.904	36.5
29000	Dwell	95.53	38.90	19.7	0.452	18.3
29000	Departure	158.46	80.48	40.7	1.120	45.2
201	Approach	199.55	110.36	23.5	1.854	17.0
201	Dwell	95.53	59.98	12.8	1.019	9.3
201	Departure	158.46	299.90	63.8	8.043	73.7

The detailed Class 22000 Heuston emission-segment results explain why the departure increase in NOx is smaller than the increase in energy use. As shown in Table 5.10, departure used 21.30 kWh, compared with only 5.59 kWh during approach. However, departure produced only slightly more NOx than approach, 22.14 g compared with 19.98 g. This is because the energy-weighted NOx factor was much lower during departure, 1.04 g/kWh, compared with 3.57 g/kWh during approach. Departure mean power was 4.80 times higher than approach mean power, and departure energy use was 3.81 times higher. However, departure NOx was only 1.11 times higher. This is because the approach NOx factor was 3.44 times higher than the departure NOx factor.

Table 5.10 Phase-level energy and emission-factor results for the Heuston Class 22000 cycle

Phase	Duration used in calculation (s)	Mean power (kW)	Energy (kWh)	Energy share (%)	NOx (g)	NOx share (%)	$PM_{2.5}$ (g)	$PM_{2.5}$ share (%)	Energy-weighted NOx EF (g/kWh)
Approach	199.55	100.90	5.59	19.15	19.98	36.75	0.81	44.77	3.57
Dwell	95.53	87.09	2.31	7.91	12.24	22.52	0.49	26.91	5.30
Departure	158.46	483.90	21.30	72.94	22.14	40.73	0.51	28.32	1.04

Overall, the discharge at the departure can be lower than at the approach when the departure is downhill and the train requires less sustained traction power after leaving the station. This is the case for the Class 22000 and 29000 at Connolly. However, gradient alone is not sufficient to explain the results. The emission-factor structure is also important. Low-load approach operation is assigned higher NOx factors per kWh, while higher-power departure operation can have lower NOx factors per kWh. Therefore, departure emissions can be lower than approach emissions even where departure has higher instantaneous power.

Figure 5.5 and Figure 5.6 show these patterns as cumulative NOx curves. Figure 5.5 shows the Connolly cycle, where the 22000 and 29000 curves rise more evenly across approach, dwell and departure, while the 201 curve steepens after departure. Figure 5.6 shows the Heuston cycle, where the departure segment becomes more important, especially for the 201. The comparison confirms that the largest class differences arise from the combination of engine type, whole-set power, load-dependent emission factors and station-specific gradient. It also shows that departure emissions can be lower than approach emissions at Connolly because trains enter uphill but leave downhill, while the opposite pattern is observed at Heuston.

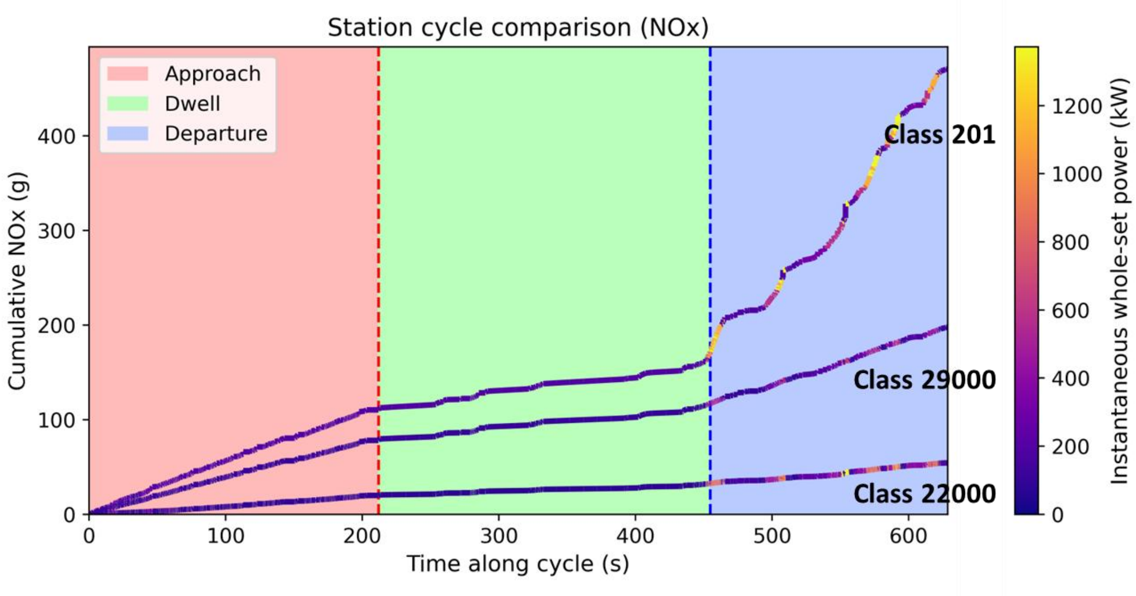


Figure 5.6 Cumulative NOx over the typical approach–dwell–departure cycle for Classes 201, 29000, and 22000 at Heuston

A further experiment examines how idling during a long dwell changes emissions for a single 22000 set. Here, the timetable-control (TC) window is defined as a 60-minute dwell only; the approach and departure profiles are taken from the reference cycle and added before and after the TC window, but do not change between scenarios. Within the dwell, the measured 22000 idle pattern is repeated in a loop so that power and emission factors during idling vary realistically. Two strategies are compared. In the "long idle" case, the set idles for the entire 60-minute dwell. In the "clean driving" case, the set idles for the first 10 minutes after arrival, is shut down for the middle 35 minutes, and is restarted for the final 15 minutes before following the same departure profile as in the reference cycle.

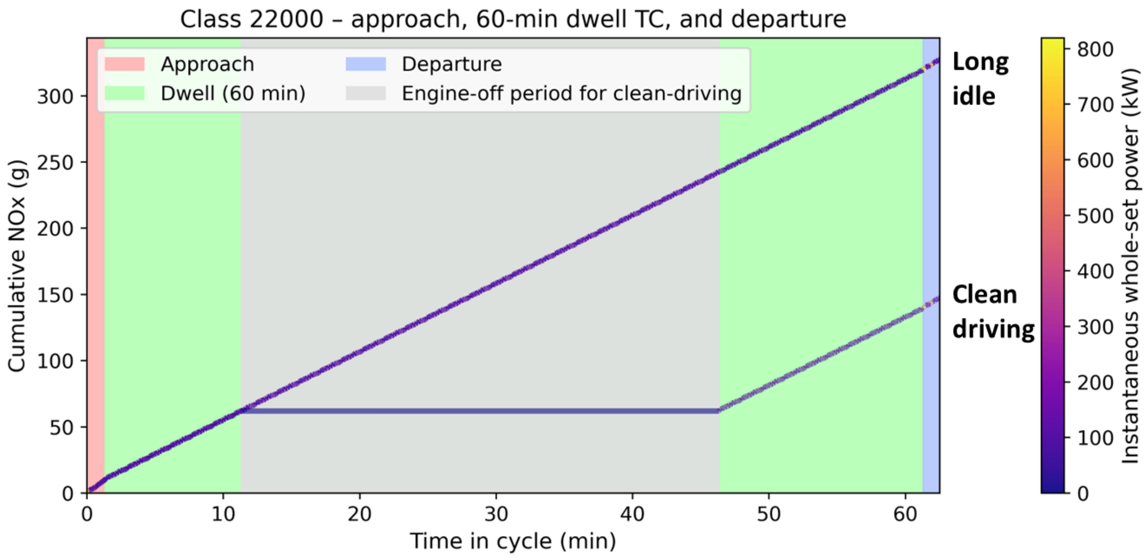


Figure 5.7 Class 22000 60-minute TC window, cumulative NOx for long-idle and clean-driving scenarios

Figure 5.7 shows cumulative NOx over the full cycle for these two strategies. The time axis starts at the beginning of the approach, continues through the 60-minute dwell, and ends after the departure. Phase shading marks the approach, dwell, and departure segments, and a lighter band within the dwell highlights the engine-off period in the clean-driving scenario. The colour along each line again shows instantaneous whole-set power, while line opacity distinguishes long idle and clean driving. In the long-idle case, the cumulative curve rises smoothly in the approach, then almost linearly during the hour-long dwell, before steepening slightly again during departure. In the clean-driving case, the curve tracks the long-idle line through the approach and the first 10 minutes of dwell, then becomes almost flat while the engines are shut down, and finally steepens again when the engines restart for the last 15 minutes and through the departure.

Over the 60-minute dwell window, the long-idle strategy emits 314.2g NOx, whereas the clean-driving strategy emits 147.3g. The difference of 166.9g is a 53% reduction in dwell-period NOx for the same timetable and the same approach and departure profile, achieved solely by shutting down the set for the middle part of the dwell. The extra emissions from approach and departure are identical between the two strategies, so the relative saving is controlled entirely by the idling practice within the TC window.

5.3.4 HVO Fuel emission reduction

Hydrotreated vegetable oil (HVO) is a paraffinic diesel fuel produced by hydrotreating vegetable oils and waste fats. It has very low sulphur and aromatic content and a high cetane number, and can usually be used as a drop-in replacement for conventional diesel in compression-ignition engines with little or no hardware modification (Aatola et al., 2008; Dobrzyńska et al., 2020). Engine and vehicle studies for road and non-road applications show that replacing fossil diesel with HVO typically reduces particulate mass and number, carbon monoxide and unburned hydrocarbons, and tailpipe CO_2, while effects on NO_x are smaller and more engine-dependent (Aatola et al., 2008; Bohl, 2018; Dobrzyńska et al., 2020; Pirjola et al., 2014; Ropkins et al., 2007).

Evidence for rail traction is more limited but points in the same direction. Dynamometer and in-use measurements on passenger locomotives running on biodiesel blends report substantial reductions in particulate emissions and more modest changes in NO_x relative to petroleum diesel, underlining the importance of engine-specific testing and duty-cycle weighting for rail applications (Graver et al., 2016; Graver & Frey, 2013). These results support the use of controlled engine tests to derive scaling factors that can then be applied to inventory emission factors.

The HVO emission-reduction factors used in this section were based on an internal Inchicore Depot emissions trial shared by Irish Rail (Green Biofuels Ltd, 2023). The trial tested Irish Rail Locomotive Fleet Number 218 using diesel on 25 September 2023 and Gd+ HVO on 26 September 2023. Emissions were measured during a stationary load-bank test, with the locomotive operated through throttle settings from idle to notch 7 and back to idle. At each notch, the engine was stabilised for 10 minutes, followed by a 10-minute measurement period. Notch 8 was used only for load comparison because of load-bank restrictions, with the same maximum load of 1898 hp recorded for both fuels.

Gaseous emissions were measured using an MGA Prime Q portable emissions measurement system, which recorded NO, NO_2, SO_2, CO_2, CO, N_2O, CH_4 and hydrocarbons. Particle number concentrations were measured using a TSI HC NPET 3795 condensation particle counter, which measures particles down to 23 nm. The reported values were based on stable readings at each notch and corrected notch-weighting factors. The trial reported reductions with Gd+ HVO relative to diesel of 57.4% for ultrafine particle number, 13.4% for NO_x, 34.4% for CO, 10.7% for CO_2 and 15.0% for CH_4 (Green

Biofuels Ltd, 2023). Table 5.11 summarises the percentage reductions relative to diesel and the scaling factors applied to the Class 201 notch-based emission factors. Because these percentage changes are independent of the trial's original units (ppm, % v/v, mg/m, particle number concentration), they can be applied directly as multiplicative scalars to the g/kWh values used in the inventory.

Table 5.11 Example results from the Inchicore 218 trial: weighted-cycle reduction with Gd⁺ HVO relative to diesel, and scaling factors used for the Class 201 factors in this study.

Pollutant	Trial unit (c)	Weighted-cycle reduction with HVO (%)	Scaling factor applied in inventory*
NOx	ppm	13.4	0.866
CO_2	% v/v	10.7	0.893
CO	ppm	34.4	0.656
CH_4	ppm	15.0	0.850
PN (ultrafine)	Particles/cm	57.4	0.426 (HVO-optimistic $PM_{2.5}$ case)

***Scaling factor = 1 − reduction; g/kWh values for Class 201 are multiplied by these factors in the HVO scenario.**

To estimate the effect on the Connolly inventory, the NO_x scaling factor of 0.866 and the "HVO-optimistic" $PM_{2.5}$ factor of 0.426 are applied to the Class 201 notch table and the Connolly totals are recomputed. Table 5.12 shows the resulting Class 201 NO_x and $PM_{2.5}$ totals for Connolly, both for station areas (R1–R3) and for station plus station-related areas (R1–R3 + line + depot). For Connolly station areas, the baseline Class 201 NO_x total over the March–May 2024 window is 0.93 t; in the HVO scenario, this falls to 0.80 t, a reduction of 0.12 t (13.4 %). When the line and depot regions are included, the Class 201 NO_x total falls from 1.48 t to 1.29 t, a reduction of 0.20 t.

These changes are modest at the national scale but more important locally. Recent national emission inventories attribute roughly 2 kt NO_x per year to the Irish rail sector (Environmental Protection Agency, 2024). Even if the Connolly reductions were sustained for a full year, the avoided NO_x would correspond to only about 0.01 % of the national rail total. However, the reductions are concentrated in a dense urban corridor where trains dwell under canopies and close to platforms and adjacent streets. In this context, a 13 % reduction in locomotive NO_x at Connolly is more relevant as a local air-quality and exposure benefit than as a contribution to national totals.

For $PM_{2.5}$ mass, there are no direct HVO measurements for Class 201, only the 57.4 % reduction in ultrafine particle number measured in the Inchicore trial. Treating this as an “optimistic” proxy for $PM_{2.5}$, the Connolly Class 201 $PM_{2.5}$ in station areas falls from 29.3 kg to 12.5 kg (a reduction of 16.8 kg), and the combined station–line–depot total falls from 46.9 kg to 20.0 kg (a reduction of 26.9 kg). Because particle number and mass are not perfectly correlated, these values are best interpreted as an upper bound on what HVO might deliver for $PM_{2.5}$.

Table 5.12 Connolly Class 201 totals and implied HVO reductions for the inventory period (March–May 2024).

Scope (Connolly, Class 201)	Pollutant	Baseline (diesel)	HVO scenario	Absolute reduction	Relative reduction
Station areas R1–R3	NOx (t)	0.93	0.80	0.12 t	13.4 %
Station areas R1–R3 (HVO-optimistic)	$PM_{2.5}$ (kg)	29.3	12.5	16.8 kg	57.4 %
Station + line + depot (R1–R3+DEPOT+LINE)	NOx (t)	1.48	1.29	0.20 t	13.4 %
Station + line + depot (HVO-optimistic)	$PM_{2.5}$ (kg)	46.9	20.0	26.9 kg	57.4 %

5.4 Discussion

5.4.1 Station emissions in the context of national inventories

The station-area totals derived here represent only a small share of national rail NOx emissions when scaled to a full year (around 1 % of the 1.94 kt Nox/yr reported nationally). This is expected because the national inventory includes emissions along the entire network, including long inter-urban sections where trains cruise at higher speeds. The key contribution of this chapter is not to change the national budget, but to resolve where and when emissions occur within and around major stations.

The comparison with MapEire at Heuston illustrates this point. The detailed station-scale inventory provides NOx totals that are more than twice the value assigned to the relevant MapEire grid cell even before locomotive contributions are added. This difference reflects the use of OTMR-derived power profiles, explicit treatment of idling and auxiliary loads, and finer spatial resolution, rather than a change in underlying fuel use.

5.4.2 Drivers of local hotspots

The regional splits show that emissions are concentrated along platform-adjacent track and in depot and approach zones. At Connolly, Region_Line and Region_Depot together account for nearly half of station-related NOx and $PM_{2.5}$, with additional hotspots at the busiest platform faces in Regions 1–2. At Heuston, the under-roof platforms (Region 1) and the approach/departure tracks in Region_Line dominate.

The class breakdown highlights the disproportionate role of locomotive-hauled services. Even though Class 201 services account for a small fraction of observed hours, they contribute a large share of NOx and $PM_{2.5}$, particularly at Connolly. The cycle analysis confirms that for a fixed timetable pattern, a 201-hauled train emits several times more NOx and $PM_{2.5}$ over a typical approach–dwell–departure cycle than either DMU type. These high per-train emissions are then concentrated at locations where trains dwell under canopies and close to passenger platforms and adjacent streets.

The diurnal profiles reinforce the importance of specific hours. At Connolly, daytime NOx remains well above night-time levels for most of the operating day, with peaks around midday, while Heuston shows lower night-time baselines and peaks aligned with morning and afternoon service peaks. This implies that a small set of zones and hours dominates near-station diesel exposure.

5.4.3 Implications for mitigation

The cycle experiment demonstrates that idling practice during long dwell times can materially alter emissions for the same timetable. For a 60-minute dwell, shutting down a Class 22000 set for the middle 35 minutes reduces dwell-period NOx by about half compared with continuous idling, without changing approach or departure profiles. This complements the timetable and idling analyses in Chapter 4: reducing unnecessary idling during long layovers offers a practical way to cut emissions, particularly in high-exposure zones.

The HVO scenario shows that fuel switching can further reduce emissions, especially for locomotive-hauled services. Applying measured HVO reduction factors to the Class 201 notch table provides double-digit percentage reductions in NOx and CO_2 and potentially much larger reductions in particulate emissions. Although the absolute NOx reduction is small at national scale, it is concentrated at a small number of urban stations where trains dwell for extended periods, so the local air-quality benefits could be significant.

Taken together, these results suggest a mixed strategy for station-scale mitigation:

(i) managing dwell behaviour and idling in long-layover services;

(ii) prioritising cleaner rolling stock (e.g. DMUs or electrified services) on the most exposed platforms; and

(iii) considering low-carbon, low-emission fuels such as HVO for locomotive-hauled fleets serving dense urban stations.

In addition to HVO and idling reduction, other technologies could help reduce station-area emissions. Short-term options include engine soft stop-start systems and supplying auxiliary loads from local electrical connections or shore power at platforms, reducing the need for diesel engines to remain running during dwell periods. Medium- to long-term options include parallel hybrid powertrains and hydrogen fuel-cell trains for non-electrified routes, although these would require greater infrastructure and fleet investment.

5.4.4 Limitations and future work

The analysis is constrained by OTMR availability, reliance on proxy emission factors, and the use of reconstructed cycles for some fleets. Spring and early-summer telemetry snapshots do not capture winter heating loads and cold-start behaviour, so annual averages may be underestimated. Proxy engines may differ from the Irish fleets in tuning, after-treatment, and auxiliary loads, and Tier 2 and Tier 3 methods inevitably smooth over operational variability.

Future work could extend OTMR coverage to a full year, incorporate direct emission measurements for the Irish fleets (particularly Class 29000 and Class 201), and refine the Tier 3 templates using additional stations and service types. Integrating measured station-area concentrations with these emission fields would also allow more direct evaluation of model performance and exposure.

5.5 Conclusion

Chapter 5 developed a high-resolution, energy-based emission inventory for Connolly and Heuston stations, combining OTMR data, proxy emission factors, and the zoning scheme from Chapter 3. The methods translated detailed train operations into time-resolved emission fields for NOx and $PM_{2.5}$ at the scale of individual platforms, depots, and approach tracks.

The results show that, while station-area emissions represent only a small share of national rail NOx when scaled to annual totals, they are highly concentrated in specific zones and hours. Connolly exhibits larger station-area emissions than Heuston over the study windows, driven by frequent DMU services and significant locomotive activity, while Heuston shows a stronger contribution from

Class 22000. Across both stations, Class 201 locomotives stand out for their high per-train emissions and their disproportionate contribution to NOx and $PM_{2.5}$ in local hotspots.

Cycle-level analysis reveals that a typical Class 201 station event emits several times more NOx and $PM_{2.5}$ than an equivalent Class 22000 or 29000 event, and that a substantial share of locomotive emissions is concentrated in short departure phases. The 60-minute dwell experiment demonstrates that shutting down trains for part of a long layover can halve dwell-period NOx without altering the timetable. The HVO fuel-switch scenario, based on Inchicore engine tests, indicates that further reductions in NOx and especially particulate emissions are achievable for locomotive-hauled services.

Overall, this chapter provides the emission fields, temporal profiles, and fleet-specific behaviours needed for the dispersion modelling in Chapter 6. The next chapter uses these zoned, time-resolved inventories as inputs to quantify how the calculated emissions translate into pollutant concentrations in nearby streets and platforms, and to evaluate how operational changes or fuel switching might improve air quality around Dublin's main rail stations.

6. Air quality dispersion modelling around the Dublin rail station area

6.1 Introduction

This chapter brings together the monitoring and modelling work developed in the previous chapters to quantify how diesel train activity influences outdoor air quality in the neighbourhood around Dublin's Connolly Station. Chapters 3 and 4 showed that enclosed platforms can experience pronounced short-term pollution peaks and that idling in roofed zones is a key operational driver. Chapter 5 then converted train activity data into time- and space-resolved emission inventories for NO_x and $PM_{2.5}$ in and around the station. While these inventories provide the required source terms, they do not on their own indicate how far pollutant increments extend into surrounding streets, or how large the station signal is relative to local road traffic and regional background. The aim of this chapter is therefore to translate the calculated emissions into concentrations and to characterise the near-field footprint of diesel rail operations around Connolly.

Dispersion modelling provides the link between emissions and exposure by combining source strength with meteorology and urban form (Holmes & Morawska, 2006; Kumar et al., 2011). In dense city-centre settings such as Connolly, this link is less direct because buildings can constrain ventilation, emissions are released close to pedestrians and building facades, and rail and road sources overlap in space and time (Chan et al., 2001; Kumar et al., 2011). Street-canyon studies show that concentrations can vary sharply over short distances and are sensitive to both background winds and traffic conditions, which reinforces the need to represent canyon processes when interpreting near-source gradients (Boddy et al., 2005; Vardoulakis et al., 2003). A dispersion model supports two core tasks in this thesis. First, it allows the station emission inventory derived in Chapter 5 to be evaluated against independent monitoring data, providing a check on whether the magnitude, timing and spatial gradients implied by the inventory are plausible. Second, it enables scenario testing, where rail and road emissions can be adjusted separately to quantify the station contribution.

Other approaches could be used to link station emissions with outdoor concentrations around transport hubs. Data-driven statistical and machine-learning models, including land-use regression and tree-based algorithms, are now widely used to predict traffic-related air pollution from monitoring data, land use, traffic and meteorology at street and neighbourhood scales (Hoek et al., 2008; Shen et al., 2022; Shi et al., 2020). Recent work highlights their ability to capture fine-scale spatial contrasts in urban $PM_{2.5}$ and NO_2, including near traffic hotspots and key intersections.

However, they depend on relatively dense monitoring, provide limited physical control over how individual source groups are represented in space, and are less well suited to counterfactual scenarios in which specific source categories (such as diesel rail versus road traffic) are adjusted independently while meteorology and background remain unchanged.

More detailed representations of flow and dispersion outdoors are possible with computational fluid dynamics (CFD), which has been applied to traffic emissions around busy intersections, real-world street canyons and roadside hot spots (Ahmad et al., 2025; Gallagher, 2016; Qin et al., 2024; Rodríguez-Camarena & Gonzalez-Olivardia, 2023). CFD can resolve complex building-scale flow structures and local source–receptor relationships with high spatial resolution, making it useful for designing and testing mitigation at specific junctions or corridors. For the Connolly case, however, full-year, multi-scenario simulations over a 2 × 2 km inner-city domain would be computationally demanding and would require detailed, time-varying boundary and input data that go beyond what is available, especially given the need to couple rail and road sources consistently.

At larger scales, regional chemistry–transport models such as WRF-Chem and WRF-CMAQ are increasingly used to assess the air-quality impacts of road transport and land-use change over cities and regions, simulating secondary pollutant formation and long-range transport on kilometre-scale grids (Torres-Vazquez et al., 2022; Yarragunta et al., 2025). These models are appropriate for regional source attribution and policy assessment, and can be coupled to local models, but they are not optimised for resolving the very steep, near-field gradients between a single rail terminal and adjacent streets within a small inner-city domain. They also require full regional emission inventories and chemical mechanisms, whereas in this study, urban background concentrations are already provided by measurements, and the focus is on primary NO_2 and $PM_{2.5}$ increments from local transport.

Taken together, these alternatives show that a wide range of tools exists for assessing outdoor air quality around transport hubs, but their data demands, computational cost and scale focus (either microscale CFD or regional chemistry–transport models) make them less suited to the present objective: kilometre-scale mapping of primary pollutants around one station, with explicit, source-resolved rail and road scenarios. This motivates the use of a local-scale Gaussian dispersion model with urban-canopy and street-canyon treatments.

In this chapter, that role is filled by ADMS-Urban. ADMS-Urban is a widely used operational dispersion model designed for urban applications, representing dispersion from explicit point, line, area and volume sources over domains ranging from individual streets to whole cities (Dėdelė & Miškinytė, 2015; Righi et al., 2009; Zhong et al., 2023). It incorporates boundary-layer physics, building-downwash and a street-canyon module and has been applied extensively to assess traffic-related air quality and to support local air-quality management. Recent studies show that, when driven by appropriate emissions and background fields, ADMS-Urban reproduces temporal variability and spatial patterns of NO_2 and particulate pollution at roadside and urban-background sites with useful skill (Dėdelė & Miškinytė, 2015; Zhong et al., 2023). For the Connolly case, ADMS-Urban offers a pragmatic compromise between physical realism, data requirements and computational cost. It operates at exactly the spatial scale of interest (a few kilometres around a complex source), can represent the bespoke rail emissions inventory from Chapter 5 and the link-based road-traffic inventory in a single consistent framework, and includes the building and canyon parameterisations needed to capture the constrained flow along the station–road corridor.

The remainder of this chapter applies ADMS-Urban to simulate NO_2 and $PM_{2.5}$ concentrations over a 2 × 2 km domain around Connolly Station. The model is driven by station emission sources derived from the Chapter 5 inventory, combined with a link-based road-traffic inventory and an hourly urban background series. The set-up incorporates 3D building data, urban-canopy parameters and explicit street-canyon representation for key roads. Model performance is evaluated using continuous EPA monitoring stations and a local NO_2 diffusion-tube deployment. The results section then (i) checks the representativeness of the chosen meteorological data for the station area, (ii) evaluates NO_2 and $PM_{2.5}$ model performance, (iii) separates rail and road contributions through scenario runs, and (iv) presents annual-average concentration maps to place the point-based results into a spatial context.

6.2 Study area and domain definition

Several considerations motivated the choice of Connolly as the case study for detailed dispersion modelling. First, Connolly is located in a more central and heavily trafficked part of the city than Heuston, with a denser mixture of commercial and residential land uses, making it a critical location for potential exposure. Second, four fixed-site air quality monitoring stations operated by the EPA lie within approximately 1 km of Connolly. This dense monitoring coverage provides multiple independent time series for model evaluation and for assessing the performance of different background and emission configurations. Third, the Connolly rail emission inventory developed in Chapter 5 is based on a longer and more complete observation period and has fewer assumptions than the corresponding Heuston inventory, giving greater confidence in the spatial and temporal patterns of emissions. Finally, practical constraints on project time and modelling resources meant that a single detailed dispersion case study was feasible within this thesis, and Connolly offered the most informative combination of central location, monitoring coverage and emissions data for that purpose.

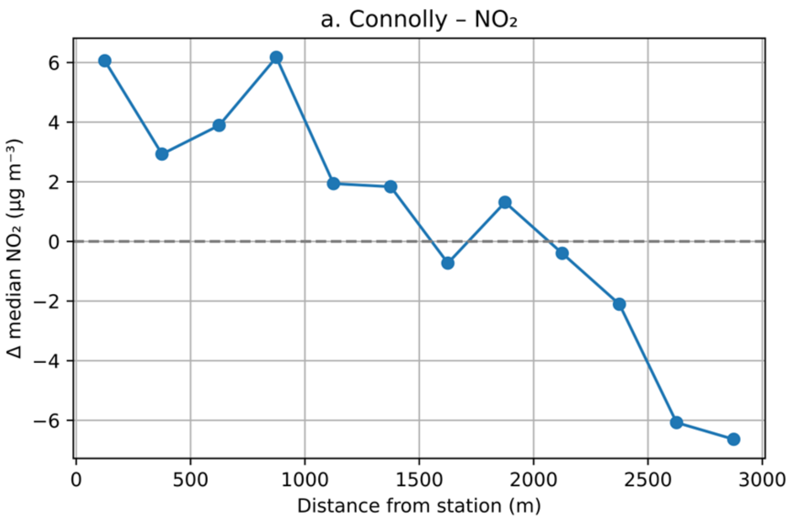
The selection of the spatial extent for the modelling is based on the Google Project Air View dataset for Dublin (Google & Dublin City Council, 2023). This dataset provides street-by-street measurements collected by an instrumented electric vehicle between May 2021 and August 2022, aggregated to approximately 50 m road segments. For the present analysis, the road-segment shapefile (AirView_DublinCity_RoadData_ugm3.shp) is used, which contains median NO_2 and $PM_{2.5}$ concentrations for each segment together with its geometry. These data are used solely to identify the neighbourhood over which elevated concentrations associated with the station–road corridor are observed, not to evaluate absolute modelled concentrations.

The Dublin Air View campaign and its quality assurance procedures are described in detail by Chen et al. (2024), who show that research-grade instruments on an all-electric Street View vehicle, combined with systematic calibration and drift checks, produce stable street-level NO_2 and $PM_{2.5}$ patterns. Consistent findings from earlier Google mobile monitoring work (Apte et al., 2017; Whitehill et al., 2020), together with the Dublin-specific analysis of Tafidis et al. (2024), indicate that the Air View measurements provide reliable, high-resolution maps of urban air pollution suitable for this study.

The Air View campaign period does not coincide exactly with the monitoring and train-activity periods used to construct the rail emission inventories in Chapters 3–5. As a result, absolute concentration levels and background conditions may differ between the two datasets. However, the spatial pattern of long-term traffic-related pollution around the stations is controlled primarily by the road network and urban form, which change slowly over time, so the Air View dataset is expected to provide a robust indication of where the highest traffic-related concentrations occur, even if year-to-year mean levels differ. In addition, the present analysis uses the Air View data in a relative, median-based way—examining differences from the Dublin-wide median in distance bands rather than raw concentrations—which reduces sensitivity to any systematic bias or temporal mismatch. These choices mean that the conclusions drawn here depend mainly on the robustness of the spatial patterns, not on the exact absolute levels during the Air View campaign.

Station centres are defined by the mid-point of the main platform groups, at 53.352354° N, 6.247512° W for Connolly and 53.3465° N, 6.2927° W for Heuston.

After re-projecting the Air View data to the Irish Transverse Mercator (ITM) system– the national grid coordinate system for Ireland, which represents locations in metres on a planar grid –the city-wide median of all Dublin road segments are calculated for each pollutant. For every segment, the minimum distance to the relevant station point is then computed, and segments are grouped into 250 m distance bands from 0 to 3 km. Within each band, the median of the segment concentrations is taken and expressed as a difference from the Dublin median. Medians are used instead of means at both stages because the distribution of street-level concentrations is skewed and includes hotspot segments with very high values. Means would be strongly influenced by these outliers and would overstate typical conditions within each band, whereas medians provide a more robust measure of the central tendency for streets at a given distance.



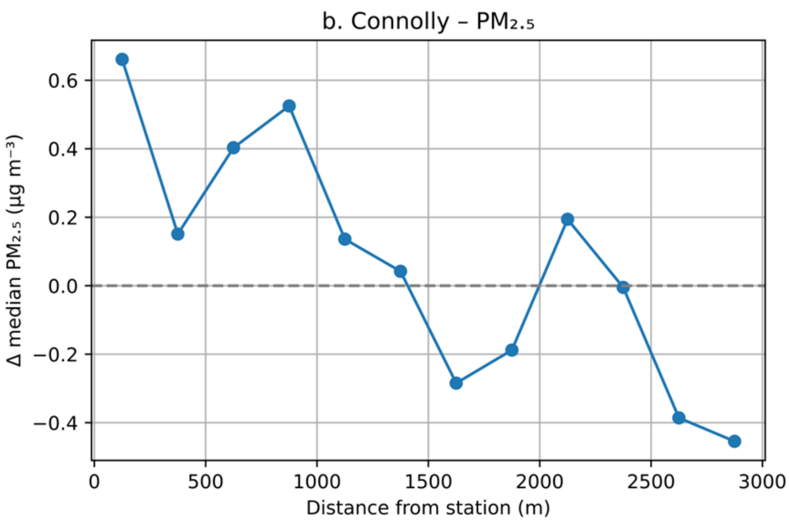


Figure 6.1 Median NO_2 /PM2.5 vs distance from station (difference from Dublin median)

The resulting distance–decay profiles for NO_2 and $PM_{2.5}$ are shown in Figure 6.1 a and b. For Connolly, median street-level NO_2 within the first kilometre is about 2–6 µg/m³ above the Dublin median, with $PM_{2.5}$ about 0.1–0.65 µg/m³ above the median. Both pollutants decrease with distance and approach the city-wide median between roughly 1.5 and 2 km. These radial patterns indicate that the main zone where streets show higher NO_2 and $PM_{2.5}$ than typical Dublin streets lie within the immediate 1–1.5 km neighbourhood of the station.

The Air View measurements represent total ambient concentrations on each road segment and do not distinguish between contributions from trains, buses, cars or other local sources. High NO_2 and $PM_{2.5}$ values close to the stations reflect a combination of effects. Without a dedicated dispersion or source-apportionment analysis, it is not possible to assign the observed increments uniquely to rail. In this chapter, the Air View data are used only to identify the spatial scale over which streets near the stations show higher concentrations than typical Dublin streets.

Based on these profiles, the modelling domain for Connolly is defined as a 2 × 2 km square centred on the station point and aligned with the ITM axes (Figure 6.2). The square extends 1 km north–south and east–west from the station centre and encompasses the zone where the Air View road segments show clear positive differences relative to the Dublin median.

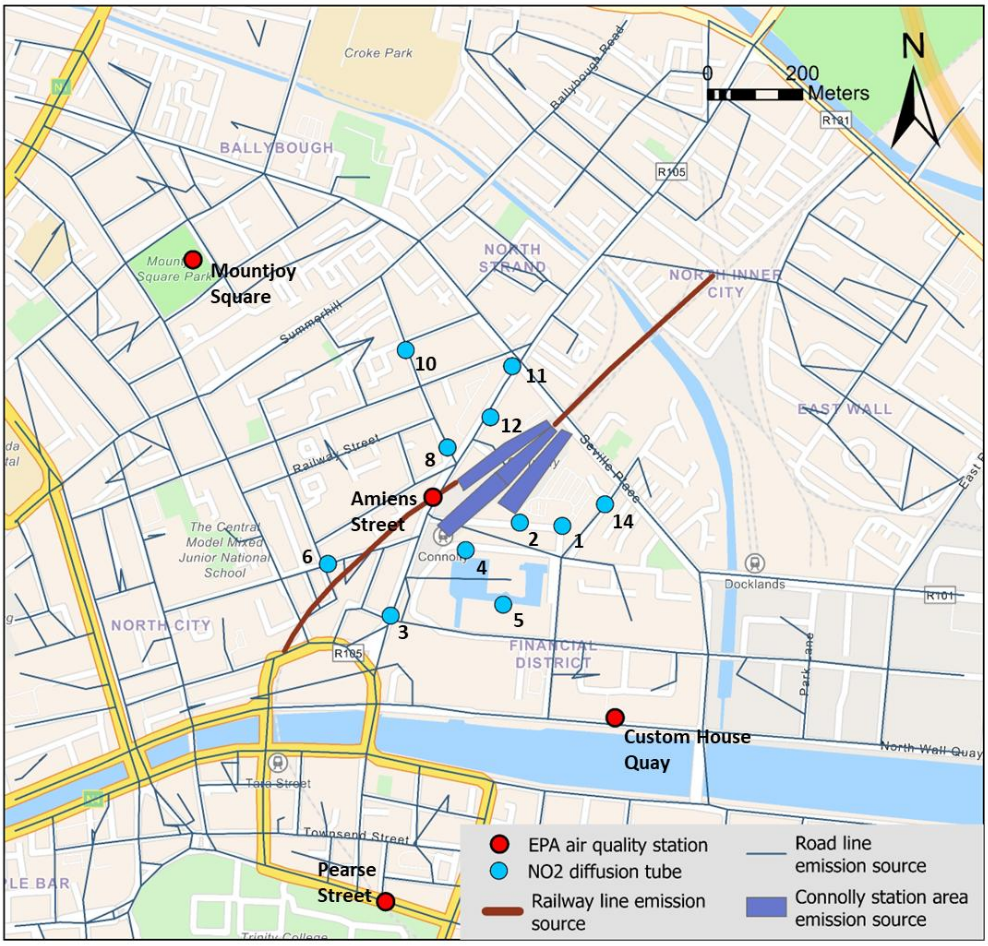


Figure 6.2 Dispersion modelling domain for Connolly station and schematic of emission sources and monitoring points

A larger 3 × 3 km domain was considered but rejected. The Air View analysis indicates that NO_2 and $PM_{2.5}$ differences relative to the city-wide median become small beyond 1–1.5 km, so extending the domain to 3 × 3 km would mainly add streets where concentrations are already close to typical city values. In addition, receptor counts and run times in ADMS-Urban scale approximately with the domain area for a given grid spacing. For the fine spatial resolution required to resolve street-scale gradients, moving from 2 × 2 km to 3 × 3 km would more than double the number of grid receptors and substantially increase computational cost, without materially improving the representation of the near-station impact. The 2 × 2 km domains are used to clip the road network, buildings and station-related emission sources and to define the ADMS-Urban source and receptor grids described.

6.3 ADMS-Urban model inputs and validation

This section describes the datasets and model configuration used to implement the Connolly study domain in ADMS-Urban. The modelling setup required spatial inputs to define the domain, road network, buildings, urban canopy, street-canyon parameters and station source regions, together with emission time series for rail and road sources, meteorological data, receptor locations and background concentration time series.

Figure 6.3 summarises the ADMS-Urban dispersion modelling workflow used in this chapter. Local rail emissions from Chapter 5 and road-traffic emissions were modelled as explicit sources within the Connolly domain. Background concentrations were included separately to represent wider urban and regional pollution, and to support the final concentration calculation and $NOx–NO_2–O_3$ chemistry. They were not treated as local emission sources in the source-apportionment scenarios.

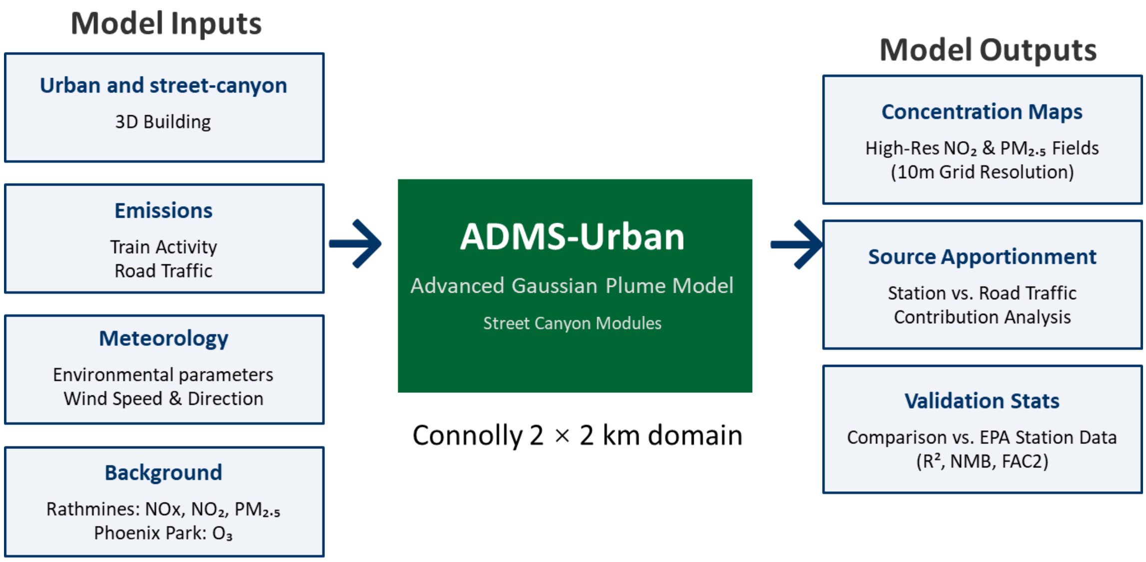


Figure 6.3 Workflow of ADMS-Urban dispersion modelling for Connolly Station

The scenario runs were used to separate the contribution of station emissions from nearby road-traffic emissions. The main outputs included receptor concentrations, comparison with monitoring data, source-contribution estimates and annual-mean contour maps.

6.3.1 3D Building

A key requirement is building geometry with height information. These 3D building data are used in two ways. First, they support street-canyon parameter calculations for roads, including canyon width, building height on each side of the road and porosity. Second, they support urban-canopy parameter calculations on a regular grid, including average building height, the fraction of the ground covered by buildings and directional frontal-area indices. The building polygons and heights are assembled into a single 3D buildings layer for Connolly, which is then used to derive the canyon and canopy inputs required by ADMS-Urban.

Building shapes are taken from OpenStreetMap building data for Dublin (OpenStreetMap, 2015). The OSM planet extract (18 May 2021) is filtered to keep polygons tagged as buildings within the 2 × 2 km Connolly domain. These polygons provide the building boundaries and plan area used in later processing and ensure that all buildings in the modelled area are represented consistently in a single layer.

Building heights are assigned using two sources, with a preference for the Dublin Building Database Points dataset (Zenodo record 7757659). This point dataset covers the city centre between the canals and provides building height information compiled for building-energy and urban morphology applications. Within the Connolly domain, these height points are matched to the nearest OSM building polygon, and the recorded height is written to a HEIGHT_M field. Heights from this dataset are treated as the first-choice values because they are internally consistent across the city-centre area and were produced specifically for building-scale applications.

Some buildings near Connolly have no matching height point. For these polygons, heights are derived from airborne LiDAR products made available through the Open Topographic LiDAR Data service and published by Geological Survey Ireland (GSI) and partners. The LiDAR point clouds are provided in gridded form as a Digital Terrain Model (DTM), representing the ground surface, and a Digital Surface Model (DSM), representing the top surface including buildings, vegetation and other structures. The raster tiles are supplied as GeoTIFFs at resolutions between about 0.13 m and 2 m, depending on the survey, and elevations are referenced to the Malin Head Vertical Datum. For this case study, all DSM and DTM tiles intersecting the Connolly domain are mosaicked into continuous rasters (dublin_DSM.tif and dublin_DTM.tif). A height-above-ground surface is then calculated as:

Equation 6.1:

$$CHM = DSM - DTM$$

For each building polygon without a valid HEIGHT_M from the Dublin Building Database Points, a robust LiDAR-based height is computed from the CHM values inside the polygon using zonal statistics. The 95th percentile of CHM is used rather than the maximum to represent a typical roof height while reducing sensitivity to isolated spurious returns. Implausible values (for example, unusually tall heights that do not match the surrounding building context) are flagged and, where appropriate, re-estimated using the LiDAR statistics.

The resulting buildings_heights layer contains OpenStreetMap building polygons with heights assigned from the Dublin Building Database Points where available, and LiDAR-derived heights otherwise. This single 3D buildings dataset provides the building inputs used in the Connolly ADMS-Urban simulations.

6.3.2 Urban-canopy and street-canyon parameters

The 3D building dataset described in Section 6.3.1 is used to compute the urban-form inputs required by ADMS-Urban's Urban Canopy Flow and Street Canyon modules. All parameters are derived in ArcGIS Pro using Python ArcPy workflows, so that the same rules are applied consistently across the Connolly domain and across the simplified road network used for dispersion modelling.

6.3.2.1 Urban-canopy grid and variables

Urban-canopy parameters are calculated on a regular grid covering the 2 × 2 km Connolly domain centred on (716657 m E, 734987 m N) in ITM. The grid resolution is 40 m × 40 m, giving approximately 50 columns by 50 rows depending on the exact domain boundary. For each grid cell, the variables required by the ADMS-Urban UCDATAVERSION1 format are computed and stored using the cell centroid coordinates as identifiers.

The centroid coordinates X and Y are recorded in ITM. Mean building height H (m) is calculated as the average of HEIGHT_M for building polygons that intersect the cell. A characteristic canyon width G (m) is also assigned for each cell to represent typical near-road spacing. This is estimated from the local pattern of roads and building setbacks by using road centrelines from the processed road network (Section 6.3.4.2) and measuring representative distances between road centrelines and nearby building edges; G is stored as the average canyon width across the road segments intersecting the cell.

The plan-area fraction of buildings, λP, is computed as the ratio of total building area within the cell to the cell area:

Equation 6.2:

$$\lambda_P = \frac{\Sigma A_P}{A_T}$$

where ΣA_P is the sum of the plan area of building polygons intersecting the cell and A_T is the total cell area (40 m × 40 m). Directional frontal-area indices, λ_F, are computed for a set of wind-direction sectors (for example, 0–45°, 45–90°, …, 315–360°). For each sector, building façades are represented by building height (HEIGHT_M) multiplied by the projected plan-length perpendicular to the sector direction. The sector-specific frontal areas are summed across buildings in the cell to obtain $A_F(\phi_1, \phi_2)$, and the index is calculated as:

Equation 6.3:

$$\lambda_F(\phi_1, \phi_2) = \frac{A_F(\phi_1, \phi_2)}{A_T}$$

where A_T is the cell area.

The resulting table is exported to the ADMS-Urban urban canopy input format. The file includes a UCDATAVERSION1 header and a GRID definition (including $Xmin, Ymin, Xmax$, $Ymax$, and the grid spacing), followed by a VARIABLES list containing X, Y, G, H, λ_P and the $\lambda_F$sectors, and then a DATA section with one comma-separated record per grid cell.

6.3.2.2 Street-canyon parameters along roads

Street-canyon parameters are computed for each road link in the Connolly network (Section 6.3.4.2) to characterise near-road geometry on both sides of each link in a form compatible with the ADMS-Urban Advanced Street Canyon model. For each road link, a buffer zone (100 m) is generated around the road centreline to define the building search area. Building polygons intersecting this zone are then assigned to the left or right side of the road based on the digitised road direction, using line orientation and side-of-line geometric tests.

For each roadside, summary height statistics are calculated from HEIGHT_M, including average, minimum and maximum building height (m). Canyon porosity is also calculated as the fraction of the road-front length that is not blocked by buildings, capturing gaps caused by junctions, courtyards and open spaces. Canyon width is defined to keep the road geometrically "inside" the canyon representation. In practice, it is taken as the minimum of (i) the effective carriageway width derived from lane counts and lane-width assumptions and (ii) a representative distance between opposing building edges, so that canyon widths do not exceed what the surrounding building pattern can support.

These street-canyon metrics are written to the ADMS-Urban Advanced Canyon input format and linked to the corresponding road sources using a unique road identifier. For road links with limited building presence on one or both sides (for example, along open plazas or park edges), the derived parameters move toward open-road conditions (large width, low height and high porosity), which reduces canyon effects in the dispersion calculations.

6.3.3 Meteorological data

Meteorological input to ADMS-Urban is taken from hourly observations at the Met Éireann synoptic station at Dublin Airport (Met Éireann, 2023). The dataset provides wind speed and wind direction at 10 m above ground, screen-level air temperature and total cloud cover for the full modelling period. These hourly values are reformatted into an ADMS-Urban meteorological file containing, for each time step, the year, Julian day and hour, wind speed (U), wind direction (PHI), temperature and cloud cover (CL). The series is treated as a continuous hourly record and is assumed to represent the synoptic-scale meteorological conditions over the Dublin area during the simulations. Local modifications to near-surface wind speed and turbulence associated with buildings and street canyons are handled within ADMS-Urban through the urban canopy and advanced street canyon options described in Sections 6.3.1–6.3.2.

Dublin Airport is used as the primary meteorological station because it provides a long, continuous and quality-controlled record with standard exposure and instrumentation. Conditions at Connolly Station can differ from those at the airport due to the influence of surrounding buildings, tracks and nearby streets. To assess how representative the airport record is for the station environment, a weather station (HOBO U30, Onset Computer Corporation, Massachusetts, USA) was installed at the open end of platform 2 at Connolly from 1 to 31 March 2024. This station recorded wind speed and direction at 5-minute intervals. The measurements were aggregated to hourly values and compared with the corresponding airport series. This comparison is used to test the consistency of wind speed and direction between the two sites; the Connolly measurements are used only for this evaluation and are not ingested directly into ADMS-Urban. The results of this comparison are reported in Section 6.4.1.

Using Dublin Airport as the sole meteorological input has some limitations. Dublin Airport is located approximately 10 km north of Connolly Station, in a more open environment, whereas the station is embedded in a dense urban environment. As a result, local wind speeds at Connolly are typically lower and directions can be more affected by buildings and street orientation than at the airport, especially under light-wind or highly stable conditions. In addition, the Connolly HOBO record covers only one month (March 2024), so the representativeness assessment is limited to that period and cannot exclude seasonal differences in airport–city contrasts. In the ADMS-Urban set-up, the airport observations are therefore interpreted as characterising the larger-scale synoptic flow, while local effects of surface roughness and buildings are handled through the model's urban and street-canyon parameterisations.

6.3.4 Emission inventories

This study considers emissions from two main source categories within the Connolly case-study domain: (i) area and line sources representing Connolly Station activities, and (ii) road line sources representing motor-traffic emissions in the 2 × 2 km Connolly domain. All emission rates are specified for NOx and $PM_{2.5}$. Temporal variation is introduced through separate time-varying factors (6.3.4.3).

6.3.4.1 Connolly station emissions

Connolly Station emissions are represented by four area sources and two track-aligned line sources (Figure 6.2).

Regions R1–R3 cover the station platforms and extended areas of platforms not covered by roofs, and R_DEPOT represents the adjacent depot area. R_LINE follow the main approach tracks and is treated as a width-line source in ADMS-Urban.

Constant emission strengths for these sources are derived from the activity-based estimates in Section 5.3.2. For each region, total emissions over the March–May 2024 period (g) are divided by the total operating time in seconds to obtain a mean emission rate in g/s. This mean rate is then divided by the plan area (m^2) for area sources or by the line length (m) for the track-aligned source to give the required ADMS-Urban units of g/m^2/s and g/m/s, respectively. The resulting values are summarised in Table 6.1.

Table 6.1 Connolly station emission sources and constant emission rates used in ADMS-Urban

Station	Region	Source type	Size (m^2 or m)	Emission-rate unit	NOx emission rate	$PM_{2.5}$ emission rate
Connolly	R1	Area	11 210 m^2	$g/m^2/s$	1.78×10^{-5}	2.46×10^{-7}
Connolly	R2	Area	4 753 m^2	$g/m^2/s$	4.37×10^{-5}	1.05×10^{-6}
Connolly	R3	Area	5 485 m^2	$g/m^2/s$	2.00×10^{-5}	2.69×10^{-7}
Connolly	R_DEPOT	Area	7 761 m^2	$g/m^2/s$	1.47×10^{-5}	2.33×10^{-7}
Connolly	R_LINE	Line	1 000 m	g/m/s	2.74×10^{-4}	5.13×10^{-6}

These constant rates represent typical NOx and $PM_{2.5}$ emissions averaged over the March–May 2024 period. The train service timetable is effectively stable across the year, so this three-month window is taken as representative of typical annual operating conditions. Sub-daily and day-of-week variability is introduced through the emission factors described in Section 6.3.4.3.

6.3.4.2 Road-traffic emissions

Road-traffic emissions in the Connolly domain were taken from the EPA/CERC regional-to-local road inventory developed for national air-quality modelling in Ireland (Cambridge Environmental Research Consultants & Environmental Protection Agency, 2023). These emissions were not recalculated in this thesis. Instead, the existing road-link emission rates from the EPA/CERC inventory were used directly as the road-emission input to ADMS-Urban. Local diurnal and weekly variation was then represented using 2024 Sydney Coordinated Adaptive Traffic System (SCATS) traffic profiles, as described in Section 6.3.4.3.

The EPA/CERC road-emission inventory was based on two main spatial and traffic datasets: the National Transport Authority (NTA) Regional Modelling System (National Transport Authority, 2024) and the Ordnance Survey Ireland Prime2 road network. The NTA Regional Modelling System is a strategic transport planning model that provides traffic flows by time period, congestion information and average speed data for road links. For the EPA/CERC inventory, these modelled traffic outputs were converted into Annual Average Weekday Traffic (AAWT) flows, split by vehicle type, and then mapped from the original straight-line NTA link geometries onto the more accurate Prime2 road geometry.

The Prime2 dataset was used to define real-world road source locations and road characteristics. Roads were separated into major and minor road datasets. In Dublin, roads with AAWT greater than 2,500 were treated as explicitly modelled major roads, while remaining Prime2 roads were included in the minor-road emissions grid. This distinction was made in the original EPA/CERC processing and was retained in this thesis when selecting road links within the Connolly 2 × 2 km ADMS-Urban domain.

Fleet composition was also taken from the EPA/CERC inventory rather than measured directly in this thesis using local automatic number plate recognition data. In the EPA/CERC method, different fleet profiles were applied to each Irish region. These profiles were based on monitoring data from the Five Cities Demand Management Study (Department of Transport, 2021), which provided vehicle-class splits and year-of-manufacture information. The year of manufacture was used to assign Euro emission classes, while supplementary vehicle-split information, including more detailed heavy goods vehicle categories, was taken from the UK Emission Factor Toolkit fleet data for Northern Ireland (DEFRA, 2023). This approach provides a consistent regional fleet estimate, but it may not fully represent the actual vehicles operating near Connolly Station. For example, newer vehicles may be driven more frequently and over longer distances, while older vehicles may be driven less often before retirement or scrappage. The city-centre fleet may also differ from the wider regional fleet because of taxis, buses, delivery vehicles and commuting patterns. Road-traffic emission rates were calculated in the EPA/CERC inventory using CERC's Emissions Inventory Toolkit (EMIT) (DEFRA, 2023).

For the present study, all road links from the EPA/CERC inventory intersecting the Connolly 2 × 2 km domain were selected and exported in ADMS-Urban line-source format. The original NOx and $PM_{2.5}$ emission rates from the inventory were preserved and used directly in ADMS-Urban, while hourly temporal factors derived from nearby SCATS traffic counters were applied to represent local traffic variation during the modelled period.

6.3.4.3 Time-varying emission profiles

The constant emission rates in the railway and road traffic represent long-term mean conditions.

ADMS-Urban applies time-varying factors from emission-factor files (.fac) to reproduce typical diurnal and weekly patterns. Separate profiles are derived for road traffic and for station-related rail activity, but in both cases the factors are non-dimensional and normalised so that the mean over a full 7 × 24-hour week is equal to 1. The same profile for each source type is applied to both NOx and $PM_{2.5}$.

Road-traffic factors from SCATS

Time-varying road traffic factors are derived from the SCATS (Sydney Coordinated Adaptive Traffic System) traffic management data operated by Dublin City Council, which provides detector-based traffic volumes at signalised junctions across the city (Dublin City Council, 2024). The SCATS Traffic Volumes dataset contains time-stamped records for each detector, including site and detector identifiers, the number of vehicles counted in each interval (Sum_Volume) and basic metadata on junction location. For this study, all SCATS sites within approximately 1 km of Connolly are selected and used as a representative sample of local traffic patterns around the station. Saturday and Sunday are separated because traffic intensity is higher on Saturdays than on Sundays in the SCATS data. Treating them as distinct day types allows the temporal factors to reflect this difference and therefore represent the variation in road-traffic emissions more accurately in the model.

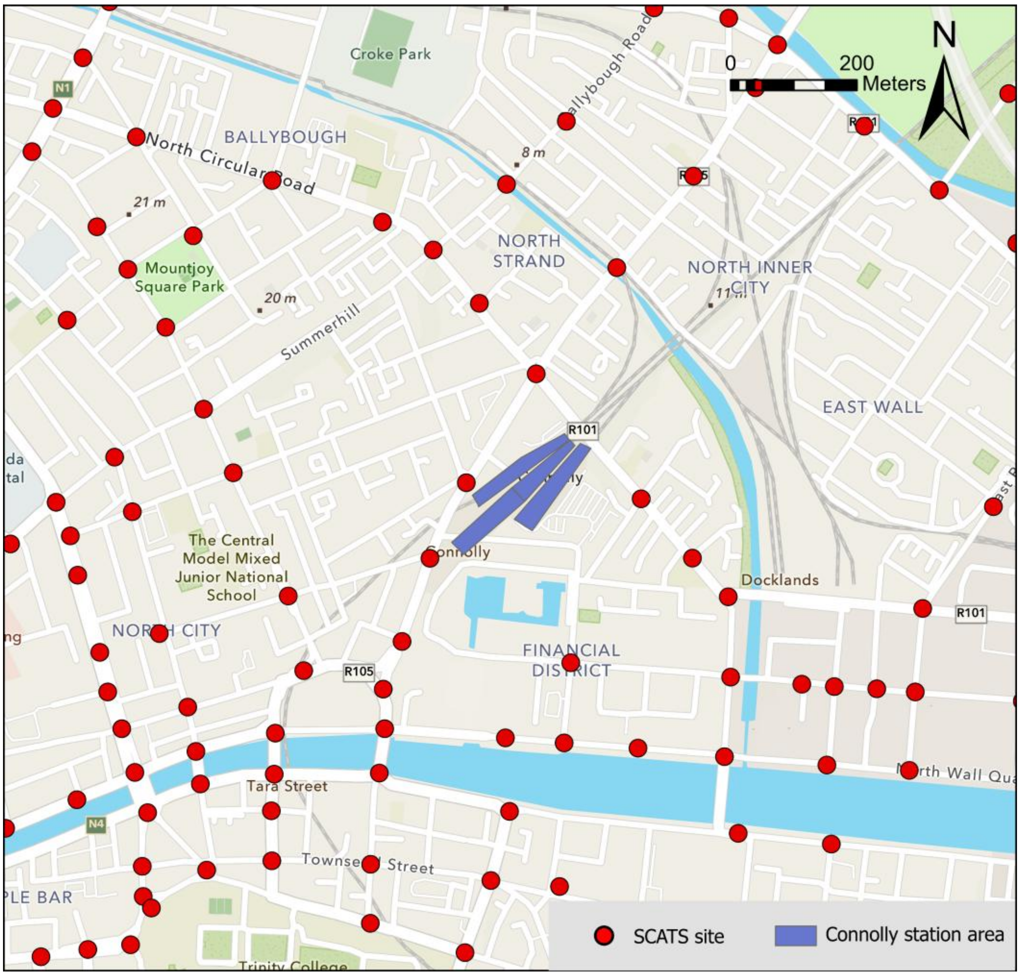


Figure 6.4 SCATS traffic detectors within 1 km of Connolly Station

For road traffic, hourly temporal factors are derived from the 2024 SCATS records. For each selected site, detector counts (Sum_Volume) are first aggregated to hourly totals and, where multiple detectors operate at the same junction, summed to give an hourly volume per site. The data are then grouped by day of week and hour of day to obtain mean volumes for three day types: Monday–Friday (“weekday”), Saturday and Sunday. The mean volumes for each of the 7 × 24 hour–day combinations are converted to coefficients by dividing by the overall weekly mean volume, ensuring that the average coefficient over the week is 1. These coefficients are then used to scale the road-traffic emissions in the ADMS-Urban runs, so that the temporal profile reflects typical local traffic behaviour rather than a single station. Figure 6.4 maps the locations of the SCATS sites used in this analysis, showing the density of detectors in the streets surrounding Connolly and their proximity to the main station–road corridor.

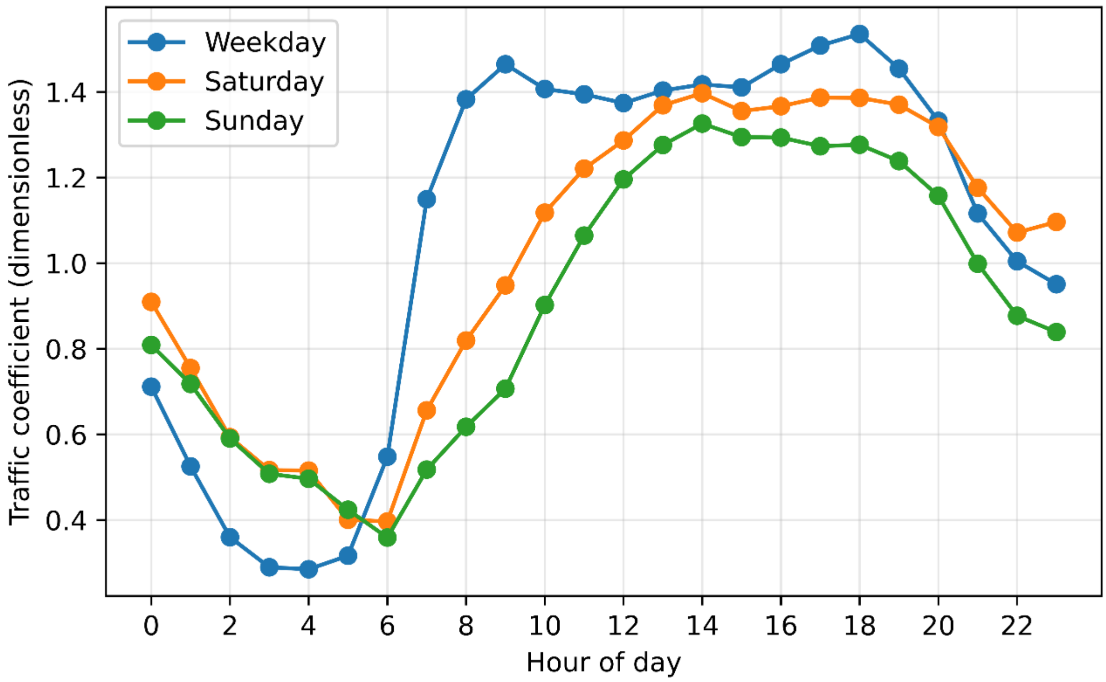


Figure 6.5 Road traffic emission coefficients in the Connolly domain

The resulting day-type profiles are shown in Figure 6.5. Weekdays show strong morning and evening peaks, with coefficients around 1.4–1.5 during the 08:00–09:00 and 17:00–18:00 periods and much lower values overnight. Saturday traffic has a flatter profile with a late-morning maximum and consistently lower nighttime values than weekdays. Sunday volumes are lowest overall, particularly in the early morning and late evening. In the dispersion model, these coefficients multiply the constant NOx and $PM_{2.5}$ emission rates for all road links in the Connolly domain.

Railway factors from train activity data

For station-related rail emissions, hourly NOx totals from the region hourly emission dataset are first summed over all Connolly regions (R1–R3 and R_DEPOT) to obtain a single station-level time series. This series is then grouped by day of week (0–6) and hour of day (0–23) to compute mean NOx emissions for each of the 7 × 24 combinations. The 7 × 24 matrix is normalised by its overall mean so that the weekly average coefficient is 1; the same matrix is used for $PM_{2.5}$. For plotting and interpretation, the 7 days are collapsed into three day types (weekday, Saturday and Sunday) by averaging the corresponding coefficients.

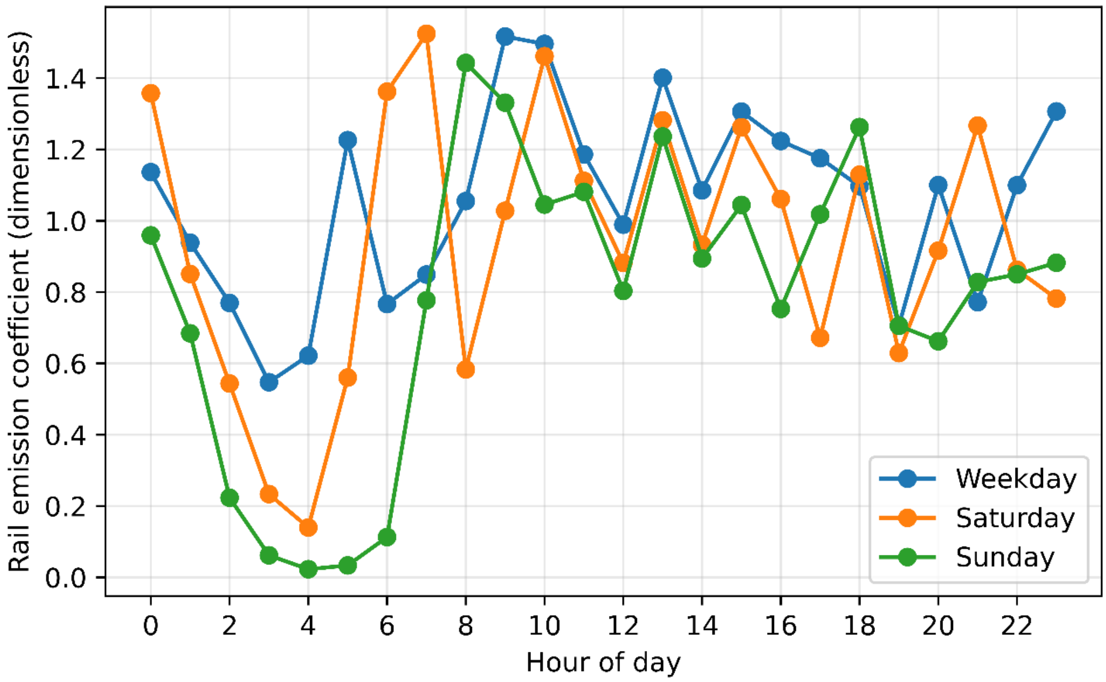


Figure 6.6 Connolly station emission coefficients

The resulting day-type profiles are shown in Figure 6.6. Weekdays exhibit pronounced peaks in the morning and early evening, with coefficients exceeding 1.4 around 08:00–10:00 and again near 18:00, and reduced emissions during the late night and early morning. Saturday shows similar peak hours but slightly lower coefficients during the inter-peak period. Sunday emissions are lowest overall, particularly between midnight and 06:00, reflecting reduced train services. These coefficients are applied in ADMS-Urban to the constant emission rates in Table 6.1 for all Connolly station sources, ensuring that the model represents both the spatial distribution of emissions and their typical temporal variation across the week.

6.3.5 Background

Background concentrations for the Connolly modelling were taken from the EPA monitoring network. An ADMS model verification and sensitivity study shows that using observations from urban background stations as the model background generally gives smaller bias and higher correlation than relying only on regional chemical-transport model fields (Porwisiak et al., 2024).

Hourly NOx, NO_2 and $PM_{2.5}$ for March–May 2024 were taken from the Rathmines urban background station, and hourly O_3 for the same period was taken from the Phoenix Park suburban background station. These time series were used as the background input to ADMS-Urban for all model hours.

Rathmines was chosen because it is classified by the EPA as an urban background site for the Dublin city area. The station is set back from major roads (approximately 65 m from Rathmines Road), is not directly influenced by heavy local traffic, and has been treated in earlier Dublin dispersion studies as a suitable site for representing city-scale NOx, NO_2 and $PM_{2.5}$ (Environmental Protection Agency, 2019, 2020; Jin et al., 2024). Rathmines lies in the southern inner suburbs, about 3.5 km south-southwest of Connolly Station, and is therefore expected to capture the wider urban background rather than the immediate influence of the station–road corridor. Its location relative to Connolly is shown in Figure 3.1.

Phoenix Park was selected for O_3 to represent regional photochemical conditions over the Greater Dublin Area. The station is located within a large urban park to the west of the city centre, around 7 km west-north-west of Connolly, and is less affected by local NO titration than inner-city sites. As a result, it better reflects the upwind O_3 field required by the ADMS-Urban NOx–NO_2 chemistry scheme. The location of Phoenix Park relative to Connolly is also shown in Figure 3.1.

6.3.6 Other ADMS-Urban model settings

Additional ADMS-Urban options were used to improve the model：

First, the time-varying emissions history option was enabled for road and rail sources. This allows ADMS-Urban to take account of the travel time between sources and receptors when applying hourly emission factors, so that near-field concentrations respond to the current-hour emission factor, while locations further downwind are influenced by the factors in preceding hours. This is particularly relevant for rush-hour peaks in NOx emissions from traffic and train operations.

Second, the daylight saving time adjustment module was activated in the additional input file. Hourly emissions factors in the diurnal profiles were defined in local solar time, while observed data and modelling period use Irish clock time. The daylight saving time option shifts the application of hourly factors during the summer period so that emission peaks remain correctly aligned with local activity patterns (e.g. morning and evening commuting) despite the clock change.

For chemistry and dispersion, the standard NOx–NO_2–O_3 scheme was used with default reaction parameters, and calm or extremely low-wind hours were processed using the model's built-in treatment rather than being removed. No special options for wet or dry deposition were activated, as the focus is on near-field NO_2 and $PM_{2.5}$ rather than long-range removal.

6.3.7 Model validation

Model performance was evaluated using two independent observation datasets within the Connolly 2 × 2 km domain: a short-term network of NO_2 diffusion tubes installed specifically for this study and hourly measurements from four EPA automatic monitoring stations (Figure 6.2, Table 6.2). In all cases, modelled concentrations were formed by adding the Connolly ADMS-Urban increment to the background time series described in Section 6.3.5.

Model performance was assessed using mean bias (MB), normalised mean bias (NMB), normalised mean error (NME), RMSE, the Pearson correlation coefficient and the fraction of predictions within a factor of two (FAC2). MB (µg/m^3) is the average signed difference between modelled and observed concentrations, so values close to 0 indicate little overall over- or under-prediction. NMB expresses this bias relative to the mean observation, with an ideal value of 0, positive values indicating systematic over-prediction and negative values indicating under-prediction. NME is the mean absolute error normalised by the mean observation and reflects the typical magnitude of the error, so smaller values are preferred. FAC2 is the proportion of model values lying within a factor of two of the observations, and values approaching 1 indicate that most predictions fall within an acceptable range of the measured NO_2 concentrations.

6.3.7.1 NO_2 diffusion tubes

Fifteen Palmes-type passive diffusive tubes (PDTs) were deployed around Connolly Station to characterise the spatial pattern of NO_2 in the near field of the road and rail sources. Tubes were mounted at approximately 2–3 m above ground level on street furniture or building façades, away from direct obstructions. Due to access and handling issues, eleven tubes were successfully retrieved and provided with valid measurements at the locations shown in Figure 6.2.

The deployment period was 4–28 May 2024, coinciding with a dedicated ADMS-Urban simulation for the same dates. Laboratory analysis provided a single period-average NO_2 concentration for each tube. For model–measurement comparison, a receptor was placed at each tube location at 2 m height. Hourly modelled NO_2 at these receptors (increment plus background) was averaged over the deployment window to obtain a predicted period-mean value.

Performance against the diffusion tube network was assessed using the spatial correlation between observed and modelled period-mean NO_2, the mean and normalised mean bias across the eleven sites, and the proportion of tubes for which the modelled concentration lay within a factor of two of the observation. These metrics indicate whether the combined emissions and urban geometry produce a realistic spatial gradient from the station and main roads into the surrounding street network.

6.3.7.2 EPA automatic monitoring sites

Within the Connolly study area, four EPA stations provide hourly data (Table 6.2): three urban traffic sites (Amiens Street, Custom House Quay, Pearse Street) and one urban background site (Mountjoy Square). Amiens Street and Pearse Street report both NO_2/NOx and $PM_{2.5}$, Custom House Quay reports $PM_{2.5}$, and Mountjoy Square reports $PM_{2.5}$ only.

Table 6.2 EPA Air quality monitoring stations within the Connolly modelling domain

Location	Type	Pollutant	Distance to Connolly（m）
Amiens Street, Dublin 1	Urban Traffic	PM2.5,NO2	20
Custom House Quay, Dublin 1	Urban Traffic	PM2.5	540
Mountjoy Square, Dublin 1	Urban Background	PM2.5	782
Pearse Street, Dublin 2	Urban Traffic	PM2.5,NO2	920

For each site, a model receptor was placed at the instrument location at 2 m height. Hourly measured data for March–May 2024 were downloaded, screened for obvious errors and negative values, and hours flagged as invalid by the EPA were removed. For a given pollutant, a station was excluded if data capture over the evaluation period was below 75 %. For NO_2, modelled concentrations at Amiens Street and Pearse Street were taken directly from the ADMS-Urban NOx–NO_2 scheme, using the same background NOx/O_3 pairing as elsewhere in the domain. For $PM_{2.5}$, modelled road and rail increments were added to the Rathmines background time series, and hourly totals were compared with EPA measurements at all four sites with $PM_{2.5}$ data.

Model performance was quantified separately for NO_2 and $PM_{2.5}$ at each station using mean bias, normalised mean bias, root-mean-square error and the Pearson correlation coefficient between hourly time series, together with the fraction of hours for which modelled and observed concentrations lay within a factor of two. Daily mean time series and hourly scatter plots were also inspected to identify systematic diurnal discrepancies, such as underestimation of morning peaks or overestimation at night, and to check consistency across the three traffic sites and the single urban-background site.

6.3.7.3 Summary of validation role

In combination, the diffusion tubes provide a dense spatial snapshot of NO_2 around Connolly for one month, while the EPA stations provide continuous temporal information at a smaller number of locations for both NO_2 and $PM_{2.5}$. Together they allow the Connolly ADMS-Urban set-up to be checked for realistic spatial gradients in NO_2 around the station and adjacent streets, realistic hourly and daily variability at nearby traffic and background monitoring sites, and consistency of model skill across pollutants and site types. These validation results are presented and discussed in Section 6.4.2-6.4.3 and form the basis for interpreting the subsequent scenario analysis.

6.3.8 Output setting

Model outputs from ADMS-Urban were configured to provide (i) time series at discrete receptors coincident with monitoring locations and (ii) annual-mean contour fields over the Connolly 2 × 2 km domain. Outputs were generated for NO_2 (as NO_2) and $PM_{2.5}$ using the standard ADMS-Urban source-group options, with separate source-group reporting retained at discrete receptors and an all-sources field used for mapping.

6.3.8.1 Receptors at monitoring locations

15 receptors were placed at all monitoring sites used for model evaluation: the 11 Palmes-type NO_2 diffusion tube locations around Connolly and the four EPA automatic stations within the domain (Amiens Street, Custom House Quay, Mountjoy Square, Pearse Street). Receptor coordinates were taken from the GIS database and EPA metadata so that each receptor coincides with the corresponding tube or analyser location. All receptors were specified at a height of 2.5 m above ground level, consistent with the mounting height of the diffusion tubes and instruments, to enable direct comparison at façade and kerbside locations.

For each of these receptors, ADMS-Urban was configured to output hourly concentrations for the full simulation period (March–May 2024), together with period-average values consistent with the subsequent evaluation (e.g. the diffusion-tube exposure period and seasonal or multi-month averages). Outputs were written for three source-group combinations: all sources (combined road and rail/station emissions), the road-traffic group, and the rail/station group. This allows direct comparison with observations using the all-sources output, and separate attribution of road and rail contributions when interpreting the modelled increments.

6.3.8.2 Annual-mean contour outputs

To visualise spatial patterns of modelled concentrations across the Connolly study area, a regular Cartesian grid of receptors was defined covering the 2 × 2 km domain used for the emission inventory and building datasets. The grid was aligned with the Connolly-centred square described in Section 6.2, with a uniform spacing of the order of 5 metres. All grid receptors were specified at a height of 2 m so that the contour plots represent typical breathing-zone concentrations at street level. For this grid, ADMS-Urban was configured to output annual-mean concentrations of NO_2 and $PM_{2.5}$ for all sources combined.

6.4 Results

This section presents the performance of the ADMS-Urban Connolly model and the resulting concentration fields for NO_2 and $PM_{2.5}$. Following the modelling objectives, the analyses focus on (i) how well the model reproduces measured spatial and temporal patterns at street scale, and (ii) what the validated model implies for the relative roles of road traffic and rail station emissions for local air quality around Connolly.

6.4.1 Representativeness of Dublin Airport meteorology for Connolly Station

The suitability of Dublin Airport as the sole meteorological input for the Connolly simulations was assessed using the overlapping period when airport and platform measurements are both available (1–31 March 2024). Hourly wind speed and direction from the airport were matched with hourly aggregates derived from the Connolly platform station. Figure 6.7 shows wind roses for the two locations, summarising the frequency of wind direction and speed classes over this period.

At Dublin Airport, the wind rose indicates a broad distribution of directions, with a clear emphasis on easterly to south-easterly sectors and secondary contributions from southerly, south-westerly and westerly flow. A wide range of wind speeds occurs, including frequent hours in the 6–9 m/s class and a noticeable fraction above 9 m/s, reflecting the relatively open and exposed site conditions. The Connolly wind rose displays a more focused pattern: winds are strongly concentrated between east and south, with very few hours from south-west through west. Wind speeds are generally lower, with most observations in the 0–3 and 3–6 m/s classes and only occasional hours above 9 m/s. This contrast is consistent with the influence of surrounding buildings and track alignments, which channel and slow the flow in the station area.

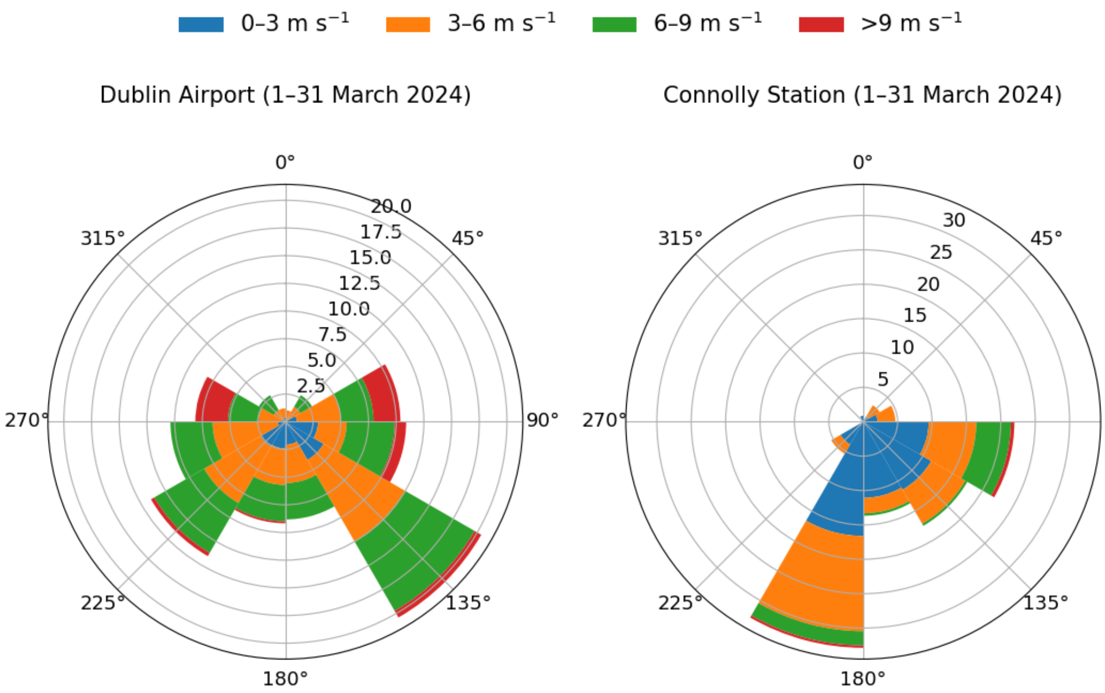


Figure 6.7 Wind roses for Dublin Airport and Connolly Station

To quantify the directional agreement, hourly winds at both sites were classified into eight 45° sectors (N, NE, E, SE, S, SW, W, NW). At Connolly, the prevailing direction in each hour was defined as the modal sector from the underlying minute-resolution data, while at the airport, the reported predominant direction was used directly. Over the 578 common hours, the Connolly sector matches the airport sector exactly in 39 % of cases and lies within one neighbouring sector (±45°) in about 70 % of cases. In 82 % of hours, the difference is within two sectors (±90°). When sector centres are converted to angles, the median absolute difference between the two series is 45°. In Figure 6.7, the concentric circles are labelled with the percentage of all hours in the record (e.g. 5, 10, 15%), providing the frequency scale for the wind roses.

Wind speed also shows a consistent temporal pattern. Mean hourly wind speed at Dublin Airport during the comparison period is higher than at Connolly by about 2–2.5 m/s, but the two series are well correlated (Pearson r = 0.74), indicating that windy and calm periods occur at the same times even though magnitudes differ because of local roughness effects.

Overall, the wind-rose patterns and hourly statistics indicate that Dublin Airport captures the large-scale wind that also affects Connolly Station. The main synoptic sectors are shared between the two sites, while the station environment modifies the local speed and direction through the building. These local effects are represented in ADMS-Urban through the urban-canopy and advanced street-canyon options rather than by altering the input meteorological record. On this basis, the Dublin Airport observations are used as the single meteorological driver for the Connolly dispersion simulations, and the Connolly platform data serve to confirm the representativeness of this choice.

6.4.2 NO_2 model evaluation

Model performance for NO_2 was evaluated using all sources output at two roadside EPA stations (Amiens Street and Pearse Street; Table 6.2) and at the 11 diffusion-tube locations around Connolly. For the EPA stations, both hourly and daily mean statistics were calculated for the period March–May 2024. For the diffusion tubes, period-mean concentrations over 4–28 May 2024 were compared with modelled means for the same dates.

6.4.2.1 Comparison with EPA monitoring stations

Table 6.3 summarises hourly statistics for Amiens Street and Pearse Street. For March–May 2024, the observed hourly mean NO_2 is 41.1 μg/m³ at Amiens and 35.4 μg/m³ at Pearse, while the corresponding model means are 39.1 and 33.6 μg/m³. Mean bias is small at both sites (−2.0 and −1.8 μg/m³, i.e. around 5 % low), and the fraction of hourly values within a FAC2 is 0.72 at Amiens and 0.76 at Pearse. Normalised mean error is around 40 % at Amiens and 36 % at Pearse, with hourly correlation coefficients of 0.58 and 0.62.

Table 6.3 Hourly NO_2 model-evaluation statistics at Amiens Street and Pearse Street (March–May 2024)

Site	N	Mean obs (μg/m³)	Mean mod (μg/m³)	MB (μg/m³)	NMB	NME	RMSE (μg/m³)	R	FAC2
Amiens street	2181	41.11	39.14	−1.97	−0.05	0.40	21.08	0.58	0.72
Pearse street	2179	35.40	33.61	−1.79	−0.05	0.36	16.30	0.62	0.76

The hexagon frequency scatter plots in Figure 6.8 Hourly frequency scatter plots of modelled vs observed NO_2 at Amiens Street and Pearse Street (March–May 2024), with 1:1 and factor-of-two lines show the distribution of hourly points. At both stations, the highest densities are clustered close to the 1:1 line in the range 20–60 µg/m³, indicating that the model reproduces the most frequent concentration range reasonably well. The factor-of-two lines enclose most of the hexagons. At Amiens Street, there is a visible tendency for the model to underestimate the highest observed hours (>70–80 µg/m³), which contributes to the negative mean bias and higher RMSE (21 µg/m³) at this site. At Pearse Street, the scatter is slightly tighter, with fewer very high observed hours and a lower RMSE (16 µg/m³).

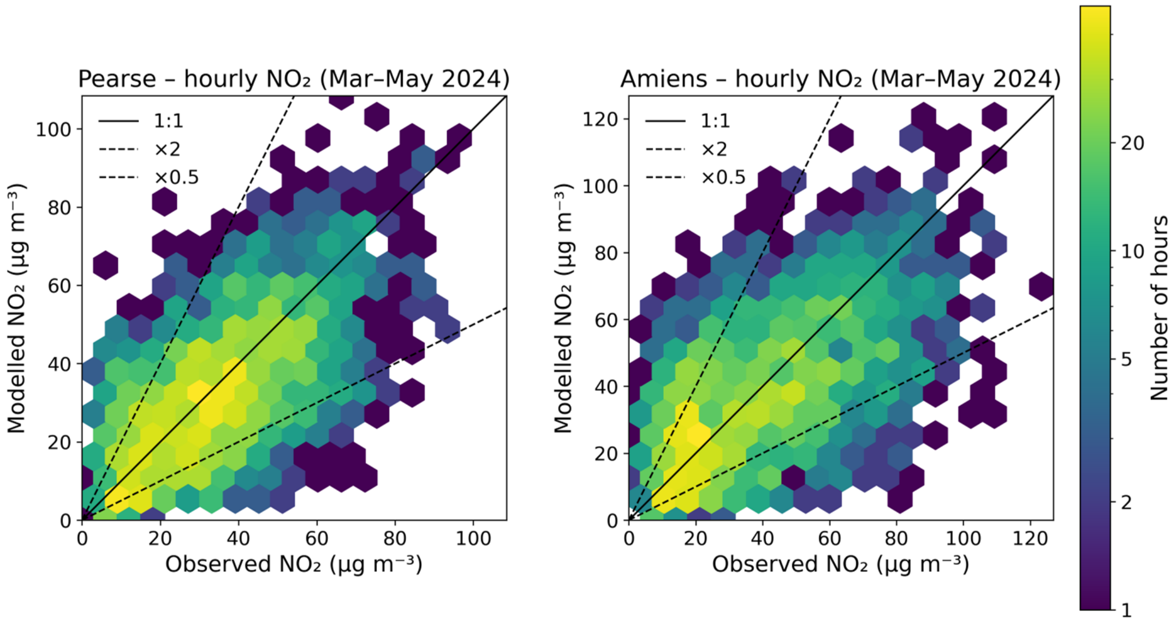


Figure 6.8 Hourly frequency scatter plots of modelled vs observed NO_2 at Amiens Street and Pearse Street (March–May 2024), with 1:1 and factor-of-two lines

Mean diurnal profiles (Figure 6.9) illustrate how the model reproduces the daily cycle. At both sites the observations show a clear morning peak between about 07:00 and 09:00, a mid-day plateau and a second, broader increase in the late afternoon and early evening. The model captures this overall structure, including lower concentrations during the night and early morning.

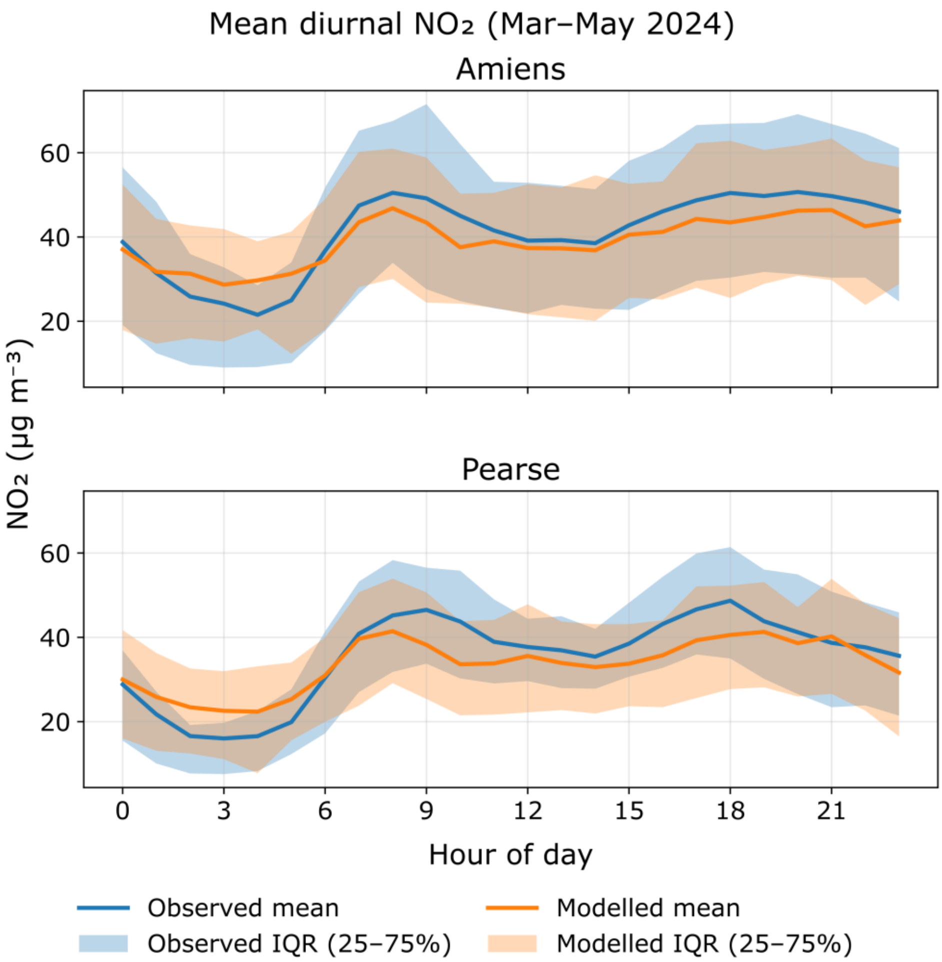


Figure 6.9 Mean diurnal NO_2 profiles at Amiens Street and Pearse Street, showing mean and inter-quartile range for March–May 2024.

At Amiens Street, the morning peak is slightly smoothed and under-predicted by 3–5 µg/m³, and the evening shoulder is also a little low. At Pearse Street, the timing of the morning and evening peaks is well reproduced, but the model tends to underestimate the mid-day and evening maxima by a few µg/m³. The shaded bands in Figure 6.9 show the inter-quartile range (25th–75th percentiles) across the March–May period; observed and modelled envelopes overlap substantially, indicating that day-to-day variability in the diurnal cycle is captured in magnitude, though with some smoothing of extremes.

Table 6.4 Daily-mean NO_2 model-evaluation statistics at Amiens Street and Pearse Street

(March–May 2024)

Site	N	Mean obs (μg/m³)	Mean mod (μg/m³)	MB (μg/m³)	NMB	NME	RMSE (μg/m³)	R	FAC2
Amiens	92	41.33	39.23	−2.11	−0.05	0.20	10.24	0.77	0.97
Pearse	92	35.49	33.74	−1.75	−0.05	0.15	6.60	0.76	0.99

Daily mean statistics indicate closer agreement (Table 6.4). The daily mean NO_2 concentrations are 41.3 (observed) versus 39.2 μg/m³ (modelled) at Amiens and 35.5 versus 33.7 μg/m³ at Pearse. Daily biases remain small (about −2 μg/m³ at both sites, 5 % low), while normalised mean errors reduce to 20 % at Amiens and 15 % at Pearse. Correlation coefficients for daily means are 0.77 and 0.76, and 97–99 % of daily values lie within a factor of two of the measurements. The daily scatter plots in Figure 6.8 show that most points sit close to the 1:1 line, with only modest spread at higher concentrations.

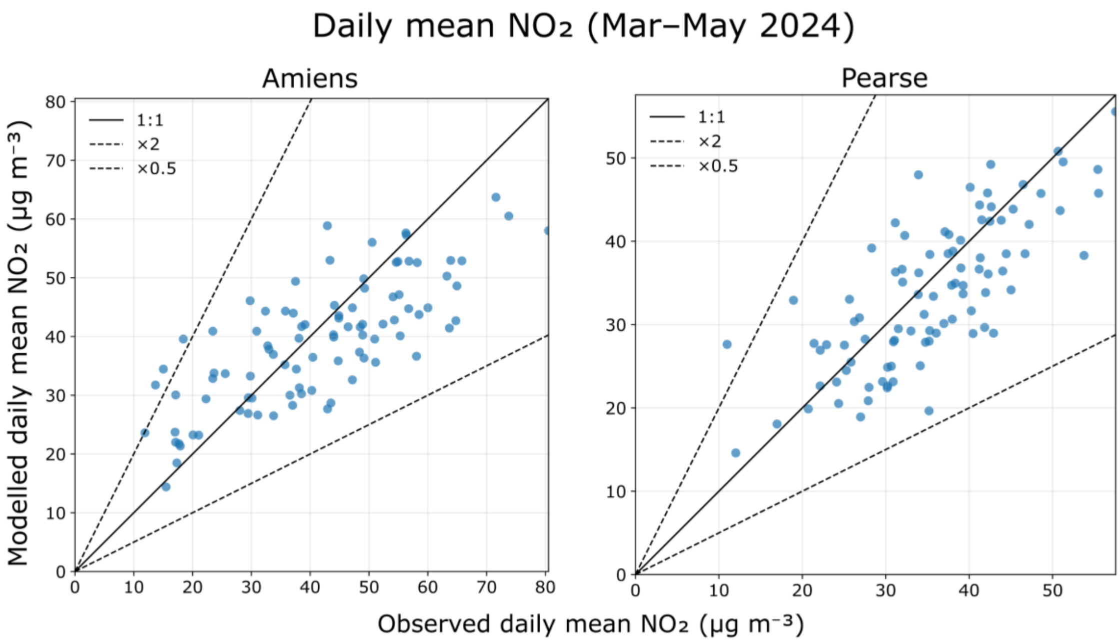


Figure 6.10 Daily-mean NO_2 scatter plots at Amiens Street and Pearse Street (March–May 2024)

The daily time-series plots (Figure 6.11) show that the model tracks the observed day-to-day variability quite well at both sites. Peaks associated with periods of stagnant or easterly conditions in late March and early May are represented, as are lower-NO_2 periods in mid-April and towards the end of May. At Amiens Street some of the sharpest observed peaks are under-predicted, especially during episodes with very high daily means (>60–70 µg/m³), consistent with the hourly scatter. At Pearse Street the model tends to slightly smooth both peaks and troughs, but captures the timing of most episodes.

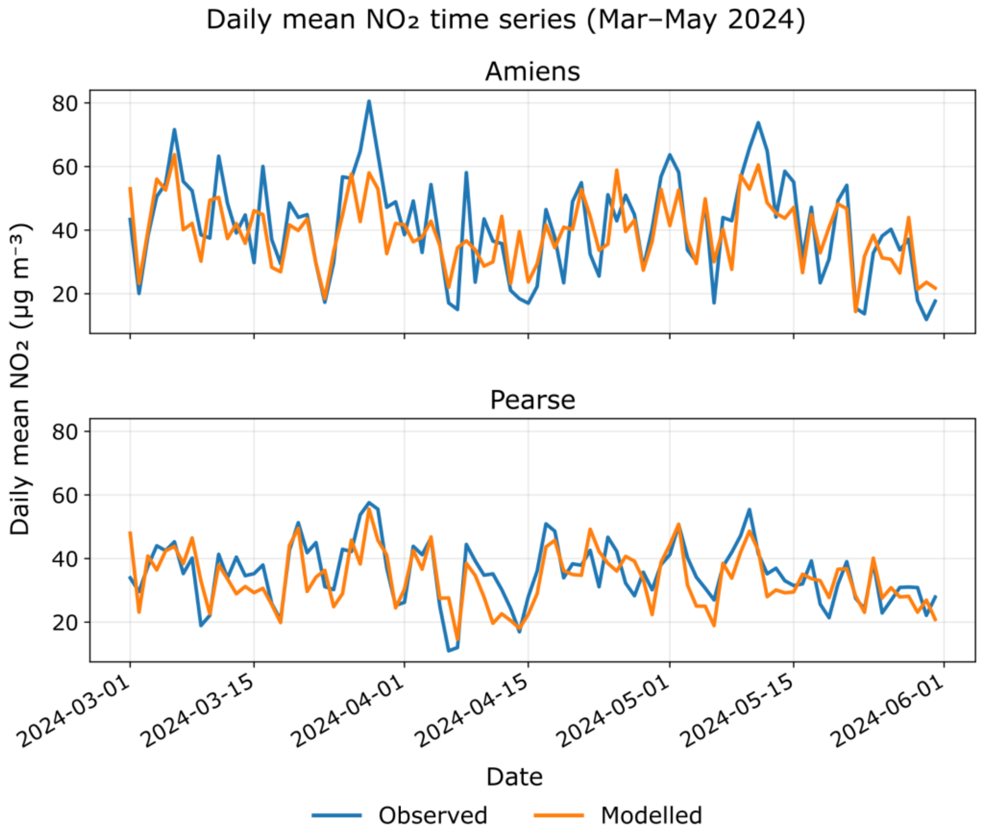


Figure 6.11 Daily-mean NO_2 time series at Amiens Street and Pearse Street

6.4.2.2 Comparison with NO_2 diffusion tubes

The period-mean comparison with the 11 NO_2 diffusion tubes around Connolly is summarised in Table 6.5and Figure 6.12. Over the 4–28 May 2024 deployment, the observed tube means range from about 18 to 32 µg/m³, with a domain-average of 25.9 µg/m³. The corresponding modelled means range from roughly 16 to 31 µg/m³, with an average of 22.8 µg/m³. The mean bias across all tubes is −3.1 µg/m³ (about 12 % low), the normalised mean error is 17 %, and the RMSE is 5.3 µg/m³. The correlation between observed and modelled tube means is 0.62, and all 11 tubes fall within a factor of two of the observations (FAC2 = 1.0).

Table 6.5 Period-mean NO_2 statistics for the Connolly diffusion tubes (4–28 May 2024)

Dataset	N	Mean obs (µg/m³)	Mean mod (µg/m³)	MB (µg/m³)	NMB	NME	RMSE (µg/m³)	R	FAC2
Diffusion tubes	11	25.93	22.80	−3.14	−0.12	0.17	5.28	0.62	1.00

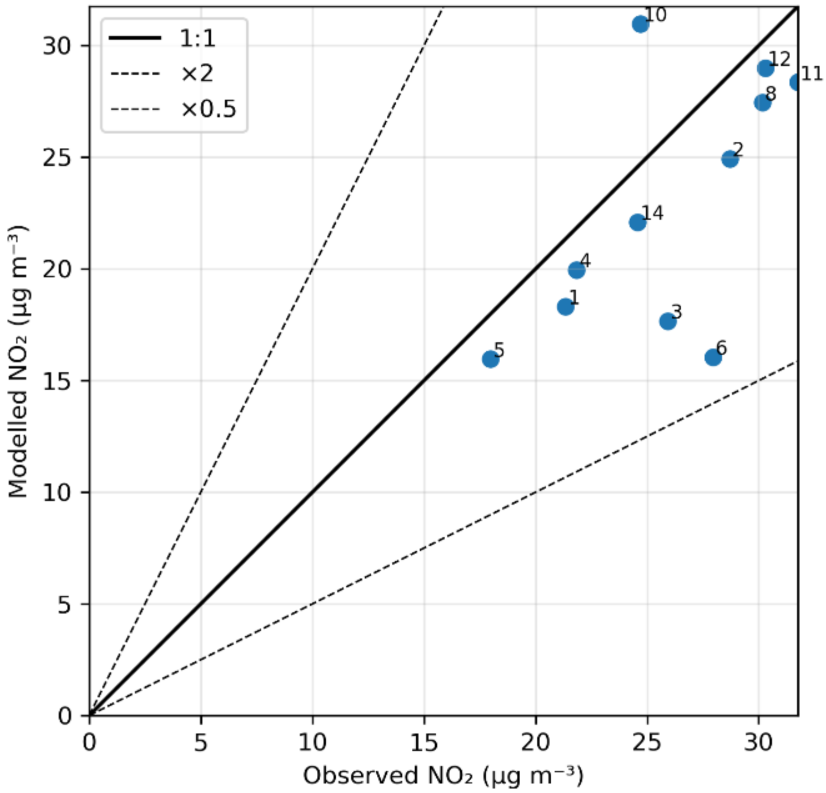


Figure 6.12 Period-mean NO_2 scatter plot for the 11 Connolly diffusion tubes (4–28 May 2024)

The scatter plot in Figure 6.10 shows that the model reproduces the ranking of sites fairly well: locations with the highest observed NO_2 (e.g. tubes closest to the main rail alignment and heavily trafficked streets) also have the highest modelled means, while lower values are predicted at more sheltered or residential locations. The systematic negative bias is most apparent for the highest-concentration tubes, where modelled values fall a few µg/m³ below the 1:1 line. This pattern is consistent with partial underestimation of near-source concentrations, which may arise from residual uncertainties in station emission strengths, canyon parameters or local traffic flows.

6.4.3 $PM_{2.5}$ model evaluation

Modelled $PM_{2.5}$ was evaluated against hourly data from four EPA stations within or close to the

Connolly domain: Amiens Street and Pearse Street (urban traffic), Custom House Quay (urban traffic, quayside) and Mountjoy Square (urban background). Statistics were computed for March–May 2024 on both hourly and daily-mean data, using the same metrics as for NO_2 (Section 6.3.7).

6.4.3.1 Hourly $PM_{2.5}$

Table 6.6 summarises the hourly statistics for all four sites. Across all stations, the model shows a positive bias, with normalised mean bias between +0.15 and +0.43 and normalised mean error between 0.43 and 0.65. The overestimation is largest at Custom House Quay and Amiens Street, and smallest at Pearse Street. Hourly correlations are moderate to good (R ≈ 0.56–0.76), with Pearse Street and Amiens Street showing the strongest agreement. Between 53 % and 72 % of hourly values fall within a factor of FAC2.

Table 6.6 Hourly $PM_{2.5}$ model evaluation statistics at four EPA sites (Mar–May 2024)

Site	N	mean_obs (µg/m3)	mean_mod (µg/m3)	MB (µg/m3)	NMB	NME	RMSE (µg/m3)	R	FAC2
Amiens	2183	9.44	12.12	+2.68	0.28	0.65	9.44	0.74	0.53
Pearse	2183	8.63	9.90	+1.27	0.15	0.43	4.59	0.76	0.72
Custom	741	6.46	9.25	+2.79	0.43	0.56	4.39	0.56	0.70
Mount	2205	8.35	10.07	+1.72	0.21	0.58	6.16	0.65	0.58

The hexagonal frequency scatter plots in Figure 6.13 show that the highest-density cells at all sites lie close to the 1:1 line, with the majority of hours within the factor-of-two envelope. At Pearse Street and Mountjoy Square, the cloud is comparatively narrow, indicating consistent relative performance across the full concentration range. At Amiens Street and Custom House Quay, a small number of hours with elevated modelled concentrations appear above the 1:1 line, consistent with the higher positive bias reported in Table 6.6.

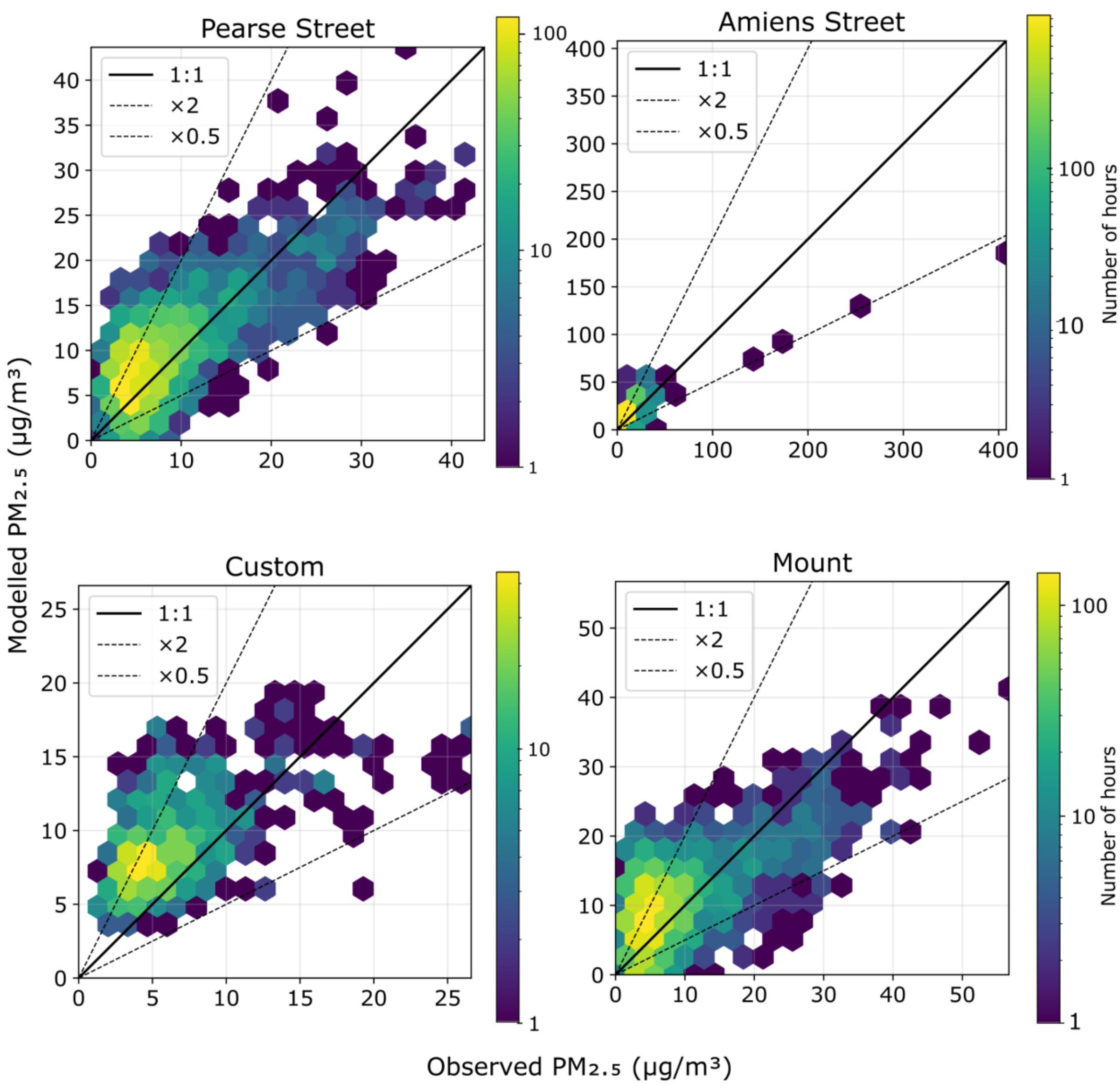


Figure 6.13 Frequency scatter plots (hexbin) of hourly $PM_{2.5}$ at Amiens Street, Pearse Street, Custom House Quay and Mountjoy Square

The diurnal profiles in Figure 6.14 indicate that the model reproduces the relatively flat daily pattern expected for $PM_{2.5}$, with only modest peaks during the morning and evening traffic periods. At Amiens Street and Pearse Street, observed and modelled curves are well aligned in phase, with the model capturing both the timing and broad shape of the diurnal cycle. The main discrepancy is a systematic offset: modelled concentrations tend to sit a few µg/m3 above the observations, particularly in the late afternoon and evening. At Custom House Quay and Mountjoy Square, diurnal variation is weaker in both model and observations, but the model again shows a consistent positive offset through most of the day.

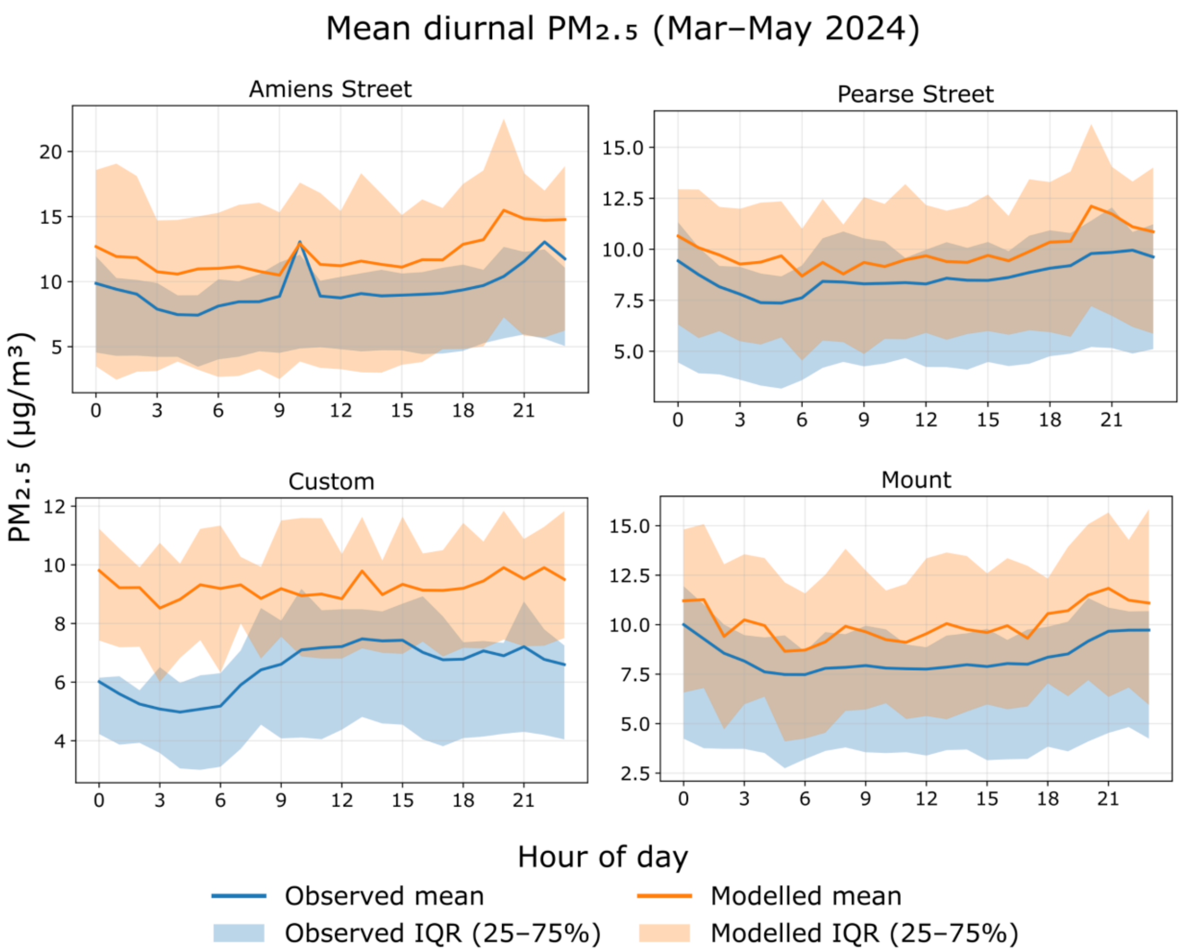


Figure 6.14 Mean diurnal $PM_{2.5}$ at the four EPA stations for March–May 2024

6.4.3.2 Daily mean $PM_{2.5}$

Daily-mean statistics for March–May are given in Table 6.7.

Table 6.7 Daily mean $PM_{2.5}$ model evaluation statistics at four EPA sites (Mar–May 2024)

Site	N	mean_obs (µg/m3)	mean_mod (µg/m3)	MB (µg/m3)	NMB	NME	RMSE (µg/m3)	R	FAC2
Amiens	92	9.47	12.14	+2.67	0.28	0.50	6.21	0.83	0.69
Pearse	92	8.66	9.91	+1.25	0.15	0.33	3.48	0.82	0.91
Custom	31	6.46	9.25	+2.79	0.43	0.46	3.56	0.62	0.84
Mount	92	8.35	10.06	+1.72	0.21	0.41	4.32	0.78	0.78

Averaging to daily means reduces random variability and improves agreement. Correlations increase to R ≈ 0.82 at both traffic sites and 0.62–0.78 at Custom House Quay and Mountjoy Square. FAC2 rises to 0.78–0.91, indicating that almost all daily means are within a factor of two of the observations, and most are substantially closer. However, the positive bias persists, with daily normalised mean bias values similar to the hourly case.

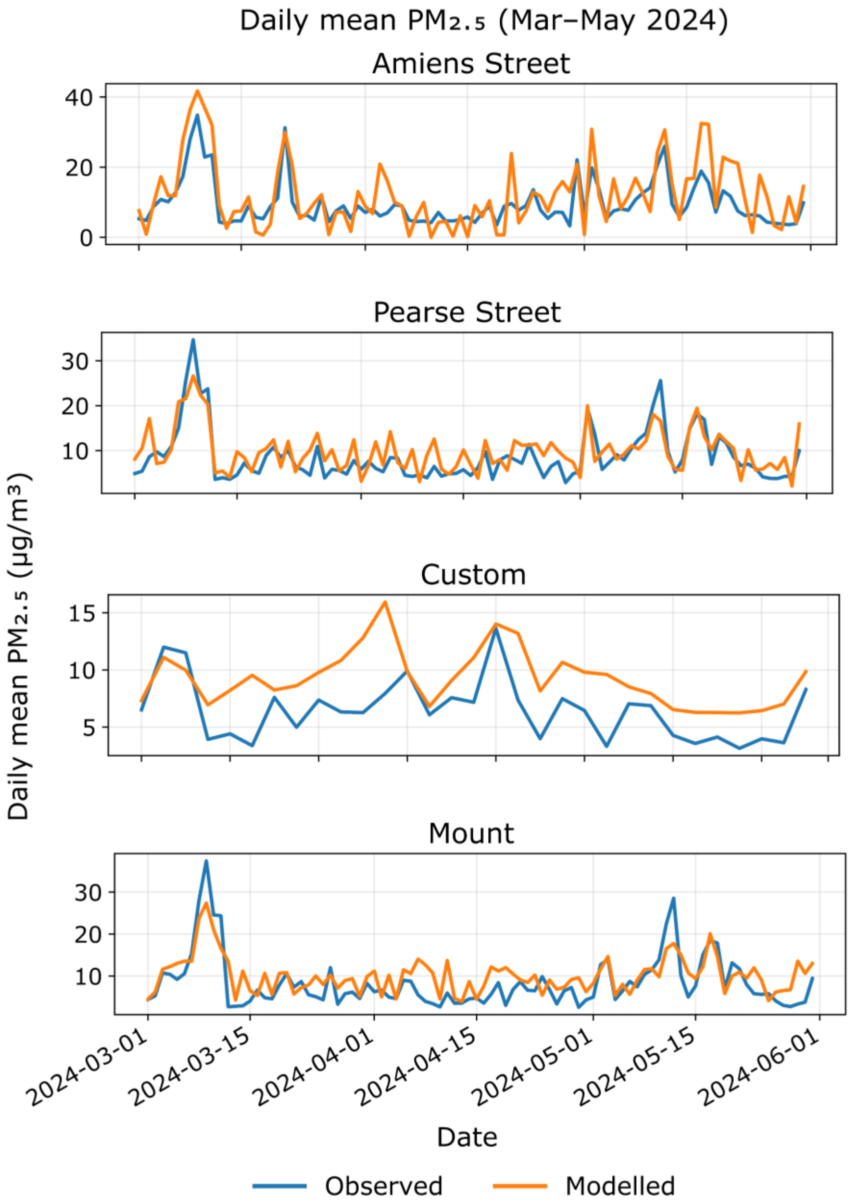


Figure 6.15 Daily mean $PM_{2.5}$ time series at the four EPA stations (observed and modelled) for

March–May 2024.

The time-series plots in Figure 6.15 show that the model captures the timing of the main $PM_{2.5}$ episodes at all four sites, including elevated concentrations in early March and mid-May. At Amiens Street and Pearse Street, modelled and observed curves track each other closely from day to day, with the model generally reproducing both sharp peaks and lower background periods, but tending to overshoot during some of the higher episodes. At Custom House Quay and Mountjoy Square, the episodes are less pronounced, and the amplitude of the observed variation is smaller; the model tends to overestimate during low-concentration days but still tracks the main synoptic-scale fluctuations.

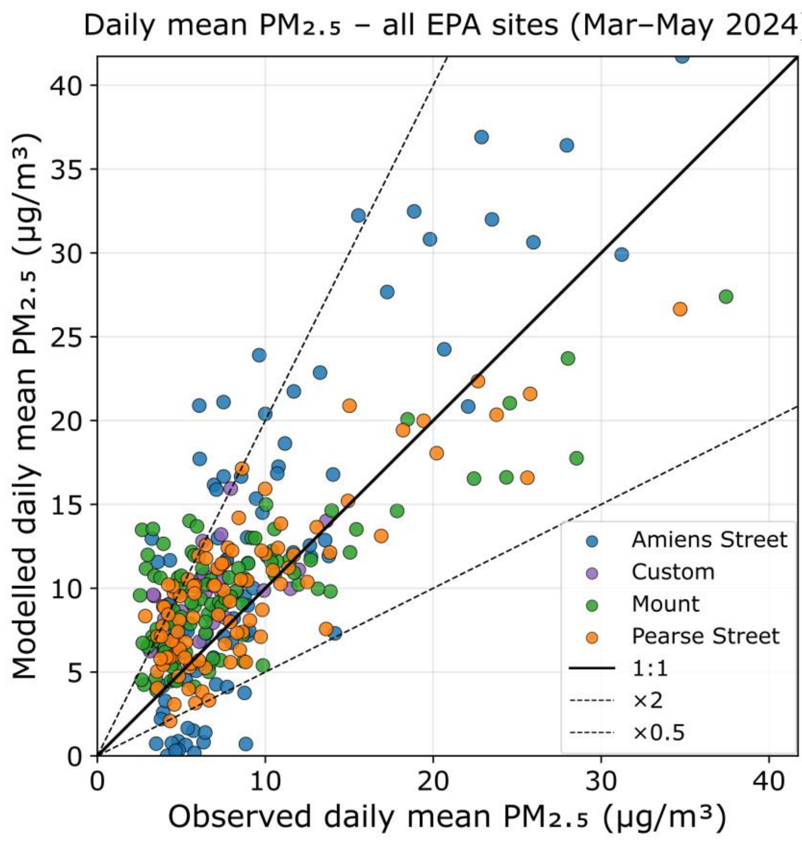


Figure 6.16 Scatter plot of daily mean modelled versus observed $PM_{2.5}$ for all four EPA stations combine

The combined daily-mean scatter plot for all sites (Figure 6.16) confirms these patterns. Points cluster around the 1:1 line for the bulk of the observed range, with a tendency for Amiens Street and Custom House Quay to lie above the line at moderate concentrations (10–25 µg/m3), reflecting the positive bias at these locations. Pearse Street shows the tightest cloud around 1:1, consistent with the lowest daily NME and highest FAC2 in Table 6.7. Mountjoy Square points sit between the traffic sites, indicating an intermediate level of bias and scatter.

Overall, the $PM_{2.5}$ evaluation indicates that the Connolly ADMS-Urban set-up reproduces the main temporal features of $PM_{2.5}$ at both traffic and urban-background sites during March–May 2024. Hourly and daily correlations are generally in the range 0.6–0.8, and most values fall within a factor of two of the observations. The principal limitation is an overestimation of $PM_{2.5}$ by roughly 15–45 % across the four stations, particularly at Custom House Quay and Amiens Street. This bias is most evident during low to moderate concentration periods and in the afternoon/evening part of the diurnal cycle.

6.4.4 Scenario analysis: role of station versus road traffic emissions

Three emission scenarios were used to separate the influence of Connolly station from surrounding road traffic within the 2 × 2 km domain: a base-case with both source groups active, a road-only run with rail emissions set to zero, and a station-only run with all local roads removed. For each receptor, the contribution of the station sources was estimated as the difference between the base-case and road-only concentrations, and the road contribution as the difference between the base-case and station-only runs. Results are reported as both absolute increments (µg/m3) and percentages of the base-case concentration.

6.4.4.1 Station contribution at EPA monitoring sites

Table 6.8 summarises the March–May mean NO_2 and $PM_{2.5}$ concentrations at the four EPA sites, together with the decomposed road and station contributions. The spatial pattern matches the site locations in Figure 6.2 and the prevailing easterly–southerly winds (Section 6.4.1). These winds most often carry emissions toward the north and northwest of the station (downwind). In contrast, the southern part of the domain is usually upwind.

Table 6.8 Mean modelled NO_2 and $PM_{2.5}$ and source contributions at EPA sites (Mar–May 2024)

Site	Type	NO_2 base (µg/m^3)	NO_2 from station (µg/m^3)	NO_2 station share (%)	NO_2 from roads (µg/m^3)	NO_2 road share (%)	$PM_{2.5}$ base (µg/m^3)	$PM_{2.5}$ from station (µg/m^3)	$PM_{2.5}$ station share (%)	$PM_{2.5}$ from roads (µg/m^3)	$PM_{2.5}$ road share (%)
Amiens St	Urban traffic	39.2	11.7	30	13.8	35	12.1	2.1	18	2.7	22
Pearse St	Urban traffic	33.6	0.4	1	11.9	36	9.9	0.1	1	1.5	16
Custom House	Urban traffic	–	–	–	–	–	8.6	0.3	3	0.4	4
Mountjoy Sq	Urban background	–	–	–	–	–	10.1	0.8	7	2.0	20

At Amiens Street, which lies immediately northwest of Connolly on the main downwind axis for the dominant easterly and southerly flows, the station contributes on average 11.7 µg/m^3 to NO_2 (around 30 % of the base-case concentration), comparable to the 13.8 µg/m^3 (35 %) attributed to surrounding roads. For $PM_{2.5}$, the station increment is smaller in absolute terms, about 2.1 µg/m^3 (18 %), but still similar in magnitude to the road increment of 2.7 µg/m^3 (22 %). This confirms that Amiens Street is strongly influenced by both emission groups, consistent with its position at the junction of the rail corridor and a heavily trafficked canyon street.

Further from the station, the $PM_{2.5}$ increments at Mountjoy Square and Custom House Quay are more modest. Mountjoy Square, located several hundred metres downwind of Connolly in the prevailing wind sectors, receives about 0.8 µg/m^3 (7 %) of its mean $PM_{2.5}$ from the station and 2.0 µg/m^3 (20 %) from local roads. At Custom House Quay, which is offset to the southeast and more often cross-wind from the station, the station contributes only about 0.3 µg/m^3 (3 %) to $PM_{2.5}$, and road traffic around 0.4 µg/m^3 (4 %).

Pearse Street lies to the south-southwest of Connolly (Figure 6.2), generally upwind of the station for the dominant easterly and southerly flows. The scenario results therefore show only a very small influence from the station at this site: about 0.4 µg/m^3 NO_2 and 0.1 µg/m^3 $PM_{2.5}$, both corresponding to around 1 % of the base-case concentration. In contrast, local road traffic at Pearse

Street accounts for approximately 11.9 µg/m^3 NO_2 (36 %) and 1.5 µg/m^3 $PM_{2.5}$ (16 %). This strong asymmetry between Amiens Street and Pearse Street is consistent with their relative upwind and downwind positions and supports the interpretation of the station as a significant but spatially confined source whose influence decays rapidly upwind of Connolly.

Overall, Table 6.8 indicates that within the traffic-dominated canyon environment at Amiens Street the station and roads make comparable contributions to both NO_2 and $PM_{2.5}$, whereas at more remote or upwind sites the signal from the station is much smaller and road traffic and background contributions dominate.

6.4.4.2 Spatial gradients around Connolly from diffusion tubes

The Palmes tube network provides a dense set of receptors in the near field of Connolly, allowing the spatial footprint of the station plume to be examined in more detail. Table 6.9 summarises the March–May mean NO_2 concentrations and the inferred station increments for each tube. Tubes are ordered approximately from the immediate station frontage to locations further along the rail corridor and onto quieter residential streets (Figure 6.2).

Table 6.9 Mean NO_2 and $PM_{2.5}$ and station contribution at diffusion tube locations (Mar–May 2024)

Tube ID	NO_2 base (µg/m³)	NO_2 from station (µg/m³)	NO_2 Station share (%)	$PM_{2.5}$ base (µg/m³)	$PM_{2.5}$ from station (µg/m³)	$PM_{2.5}$ Station share (%)
Tube 1	17.7	1.8	10	8.7	0.3	3
Tube 2	24.1	4.7	20	9.1	0.4	5
Tube 3	17.0	0.8	5	8.7	0.3	3
Tube 4	19.4	1.4	7	8.8	0.3	3
Tube 5	15.2	0.4	3	8.6	0.3	3
Tube 6	15.2	0.5	3	8.6	0.3	3
Tube 8	26.6	4.0	15	9.3	0.4	4
Tube 10	30.1	7.8	26	9.4	0.6	6
Tube 11	28.8	7.3	25	9.3	0.5	6
Tube 12	29.5	7.6	26	9.3	0.5	6
Tube 14	21.2	3.4	16	8.9	0.4	4

Station-related NO_2 increments are largest along the rail corridor to the north and northeast of Connolly. Tubes 10–12, which are located several hundred metres along this corridor in the main downwind direction for easterly and southerly winds, show station increments of 7–8 µg/m^3 and station shares of around 25 % of local NO_2. Intermediate increments of 3–5 µg/m^3 (15–20 %) occur at tubes 2, 8 and 14, situated close to the station entrances and at the junction with Amiens Street. On quieter streets to the west and south (tubes 1, 3, 4, 5 and 6), station contributions fall below 2 µg/m^3, representing less than about 3–10 % of local NO_2.

This spatial pattern mirrors the wind climatology: downwind receptors aligned with the rail line and the Amiens Street canyon receive a substantial additional NO_2 burden from station operations, whereas sites on the upwind side of the station or shielded by intervening buildings experience only small increments. $PM_{2.5}$ increments at the tube locations are everywhere smaller in absolute terms, typically below 0.6 µg/m^3 and contributing less than about 7 % of local $PM_{2.5}$, which is consistent with the more diffuse nature of particulate emissions and the relatively larger regional background for $PM_{2.5}$.

Taken together, the scenario results show that Connolly station is an important local source of NO_2 in the immediate downwind corridor, with contributions comparable to those from nearby roads at Amiens Street and along the rail line, but that its influence diminishes rapidly with distance and is almost negligible at upwind locations such as Pearse Street. For $PM_{2.5}$, the station signal is weaker but still measurable at the closest traffic and background sites.

6.4.5 Annual-average concentration mapping

Annual-average concentration fields for NO_2 and $PM_{2.5}$ over the Connolly 2 × 2 km domain are shown in Figure 6.17 and Figure 6.18. These maps are based on gridded ADMS-Urban outputs for all local sources, combined with the Rathmines background time series (Section 6.3.5) and then averaged over 2024. Both pollutants show a clear localisation of impacts around the station–road corridor, with much weaker gradients elsewhere in the domain.

6.4.5.1 Annual-average NO_2 concentration field around Connolly

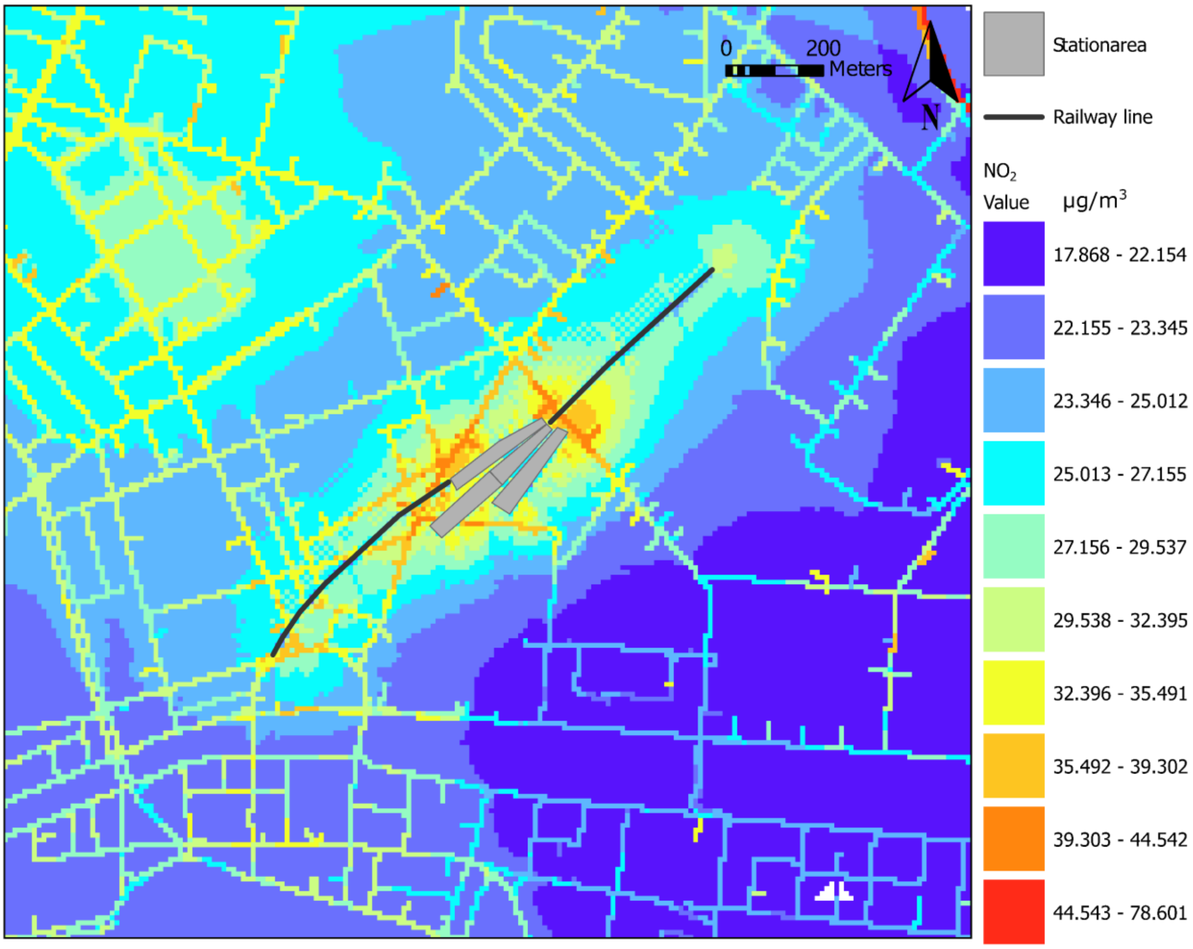


Figure 6.17 Annual mean NO_2 concentration (µg/m3) across the Connolly 2 × 2 km domain for 2024

The NO_2 map (Figure 6.17) shows a pronounced ridge of elevated annual mean concentrations extending from the station throat towards the southwest along the main approach roads. The highest grid-cell values (orange to red classes, about 45–80 µg/m^3) occur in two main areas: immediately around the station platforms and depot, and along the heavily trafficked road around Connolly.

A secondary band of elevated NO_2 (yellow to light orange, about 33–45 µg/m^3) follows the rail alignment to the north and northeast. This is consistent with the downwind corridor identified in the scenario analysis (Section 6.4.4): concentrations are highest close to the station throat and gradually decrease along the line, but a distinct plume remains visible several hundred metres away, especially where the track runs parallel to narrow, canyon-like streets.

Away from the station and the principal road–rail axis, NO_2 falls back rapidly. Most residential streets to the west and northwest of Connolly lie in the cyan to light-green range (about 23–30 $\mu g/m^3$), only slightly above the background field, while much of the Docklands waterfront to the southeast shows near-background values in the darker blue classes (about 18–23 $\mu g/m^3$). The sharp contrast between orange/yellow grid cells along the main corridor and surrounding blue/green cells highlights the strong street-scale gradients in NO_2.

This spatial pattern agrees well with the point-based evaluation. Amiens Street, which sits within the high-concentration ridge, lies in the 35–45 $\mu g/m^3$ band and receives sizeable contributions from both road and station sources. Pearse Street, located further south and largely outside the ridge, lies within the lower blue–cyan field and is dominated by road traffic, with only a very small station increment. Diffusion tube locations 10–12, positioned along the downwind rail corridor, fall within the yellow–orange bands where the scenario analysis suggested station shares of roughly 20–30 %. Overall, the NO_2 map confirms that Connolly station and the adjacent roads create a narrow, elongated hotspot aligned with the rail corridor and the main traffic routes, with annual means approaching urban-background levels within roughly 200–300 m of these features.

6.4.5.2 Annual-average $PM_{2.5}$ concentration field around Connolly

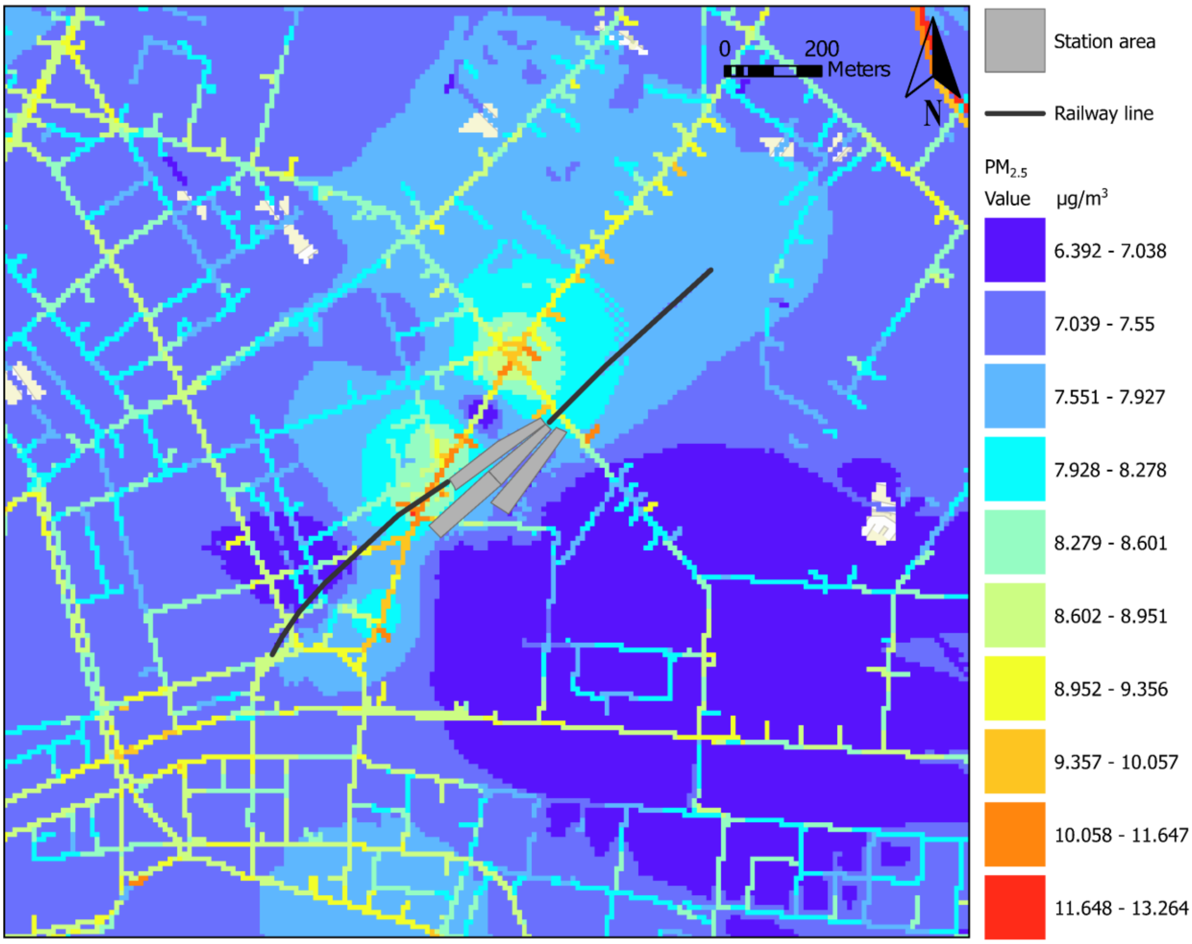


Figure 6.18 Annual mean $PM_{2.5}$ concentration (µg/m3) across the Connolly 2 × 2 km domain for 2024

The annual-average $PM_{2.5}$ field (Figure 6.18) is markedly smoother than NO_2 and shows smaller absolute contrasts across the domain. Most grid cells fall between about 7 and 9 µg/m^3 (blue to cyan classes), reflecting the dominance of the regional and city-wide background for fine particles.

Even so, the model resolves a coherent pattern of modest local enhancement along the main road–station corridor. A narrow band of elevated $PM_{2.5}$ (yellow to orange classes, about 9–11.5 µg/m^3) runs from Custom House Quay through Amiens Street towards the station frontage. Slightly higher values appear immediately around the platforms and depot and along short sections of the rail line, but the rail-related enhancement is much weaker than for NO_2, in line with the smaller station contribution to $PM_{2.5}$ diagnosed in Section 6.4.4.

Beyond this corridor, $PM_{2.5}$ quickly returns to near-background levels. Residential areas to the north and northwest and much of the Docklands to the east and southeast remain in the lowest three colour classes (about 6.4–8.3 µg/m^3). The typical gradient between the busiest links and surrounding streets is of order 1–2 µg/m3, consistent with the source apportionment at the EPA sites: road and station increment together account for roughly 20–30 % of total $PM_{2.5}$ at Amiens Street but only a few percent at more distant or upwind locations such as Pearse Street and Mountjoy Square.

These spatial patterns mirror the $PM_{2.5}$ evaluation in Section 6.4.4. Amiens Street and Custom House Quay are embedded in the moderate yellow–orange band and experience the largest local increments, matching their positive model bias and larger modelled source contributions. Pearse Street and Mountjoy Square lie closer to the cyan–green field, where $PM_{2.5}$ is only slightly above background and the Connolly source groups contribute relatively little.

Taken together, the annual-average maps reinforce the main conclusions from the point-based evaluation and the scenario analysis:

For NO_2, Connolly station and the adjacent road network generate a distinct, street-scale hotspot that is tightly confined to the rail–road corridor. Annual means in this band are substantially above the domain background and are shaped by the combination of station plumes and canyon-enhanced road emissions. Concentrations drop rapidly with distance, so much of the 2 × 2 km domain experiences NO_2 levels only modestly above the Rathmines urban background.

For $PM_{2.5}$, annual means are dominated by regional and city-wide contributions. The Connolly sources add a relatively small increment of about 1–3 µg/m^3 near the busiest links and very little elsewhere, leading to weak spatial gradients away from the immediate station–road corridor.

6.5 Conclusion

This chapter used ADMS-Urban to translate the station-area emission inventories developed in Chapter 5 into near-field concentrations around Connolly Station, and to evaluate the scale of station-related increments relative to local road traffic and background. The preliminary analysis of meteorology indicated that Dublin Airport provides a workable representation of the wind direction and speed classes relevant to dispersion around Connolly, supporting its use as the primary meteorological input for the model.

For NO_2, the model reproduced the observed day-to-day variability and achieved good agreement with measured daily means at the two EPA fixed sites within the modelling domain. During the May 2024 evaluation period, daily mean NO_2 was modelled as 39.2 µg/m³ at Amiens Street compared with an observed 41.3 µg/m³, and 33.7 µg/m³ at Pearse Street compared with an observed 35.5 µg/m³. Correlations were strong at both sites ($r \approx 0.77$ and $r \approx 0.76$, respectively), with normalised mean errors of the order of 15–20%.

The short diffusion-tube deployment provided an additional spatial check on the near-station gradient. Across the 11 tube sites, the mean bias was modest (−3.1 µg/m³) and all locations fell within a factor of two of the measured periods means (FAC2 = 1.0). The main remaining discrepancy was that the model tended to underestimate the highest tube means in the immediate station corridor, which is consistent with the sensitivity of near-field NO_2 to short-term operational peaks and street-canyon ventilation.

For $PM_{2.5}$, the model captured broad temporal changes but showed a systematic tendency to overestimate concentrations at the monitoring sites. Across the two EPA stations, mean positive bias was of the order of 15–45%, despite correlations remaining moderate to strong (approximately 0.6–0.8). This performance is consistent with the stronger role of regional background and the additional uncertainty introduced by simplified treatment of primary $PM_{2.5}$ emissions and secondary formation. The results indicate that the model is suitable for comparing spatial patterns and scenario differences, but that absolute $PM_{2.5}$ values should be interpreted cautiously and with attention to background specification.

The scenario analysis demonstrated that diesel rail operations at Connolly can be a material local contributor to NO_2 in the immediate station corridor, but that the contribution decays rapidly with distance and is highly dependent on receptor location and prevailing winds. At Amiens Street, the modelled station increment was 11.7 µg/m³ (about 30% of total NO_2), comparable to the road-traffic increment of 13.8 µg/m³ (about 35%). For $PM_{2.5}$ at the same site, the station increment was 2.1 µg/m³ (about 18%) versus 2.7 µg/m³ (about 22%) from roads. In contrast, at Pearse Street the station increment was close to negligible (around 1% of total for both pollutants). Annual-average concentration maps reinforced this conclusion by showing the strongest gradients aligned with the rail corridor and adjacent arterial roads, with more modest spatial variation away from these corridors. Overall, the chapter provides a quantitative basis for prioritising mitigation measures and frames how operational or fuel-based changes to rail activity might translate into outdoor air quality benefits in nearby streets.

7. Conclusion

7.1 Summary of Research

This thesis set out to examine air quality in and around railway station environments, using Connolly and Heuston as Irish case studies, and to develop a modelling framework that links station operations to both indoor exposure and outdoor neighbourhood impacts. The overall motivation is that Ireland remains highly dependent on diesel traction, and its main Dublin terminals combine semi-enclosed architecture with high service frequency and heavy surrounding road traffic.

Chapter 2 established the context through a systematic literature review of modelling approaches used for THs. While dispersion, emission and statistical models are widely applied to airports, ports and road corridors, the review found that modelling studies focused on rail stations are comparatively limited and that existing emissions databases generally lack the spatial and temporal resolution needed to assess station-scale impacts. This gap provided the rationale for combining monitoring, operation-based emissions calculation and dispersion modelling within a single integrated workflow.

Chapter 3 analysed air quality monitoring data inside the stations and used statistical models to separate regional and local drivers. Across pollutants, urban background concentrations were the dominant predictor of platform conditions, but diesel operations added a clear increment. Random Forest models indicated that cumulative idling time in the enclosed platform region is the most influential operational factor, with near-linear increases in platform PM and NO_2 as idling accumulates. These results show that station air quality is not simply a reflection of city-wide conditions: operational choices within the station footprint can measurably change exposure for passengers and staff.

Chapter 4 focused on the operational behaviour underlying these station increments by analysing individual diesel multiple units idling events and their relationship to TCs. The analysis showed that idling durations are heavy-tailed and that long idling events occur more often than would be expected under simple dwell-time assumptions. Bayesian modelling linked idling behaviour to the scheduled gap between arrival and departure, and demonstrated that reducing unnecessary engine-on time during long dwell periods is a realistic mitigation lever that does not require timetable changes.

Chapters 5 and 6 translated these behavioural and operational insights into quantified emissions and neighbourhood concentrations. Chapter 5 developed a high-resolution emissions calculation tool based on OTMR, timetables and mode-specific emission factors, producing zoned and temporal emissions fields that highlighted strong spatial concentration at platforms and depot areas, and clear morning and evening peaks linked to service patterns. The results also showed pronounced fleet effects, with Class 201 locomotive-hauled services emitting substantially more NO_x and $PM_{2.5}$ per movement than the diesel multiple unit fleets, and with large reductions possible through fuel switching to HVO for the highest-emitting classes. Chapter 6 then coupled the rail emissions with road traffic and meteorology in ADMS-Urban to map outdoor concentrations around Connolly. The evaluation supported the ability of the framework to reproduce observed NO_2 patterns and to attribute a material station contribution immediately downwind of the rail corridor (for example, around 30% of total NO_2 at Amiens Street during the evaluation period), while showing negligible station influence at receptors outside the main corridor. Together, the chapters provide an evidence base for targeted mitigation that can deliver local air quality benefits alongside longer-term decarbonisation through electrification.

7.2 Fulfilment of Research Objectives

This thesis addressed the six research objectives set out in Chapter 1. Table 7.1 summarises how each objective was fulfilled and identifies the main chapters where the corresponding work was presented.

Table 7.1 Fulfilment of research objectives

Research objective	Thesis outcome
Objective 1: Undertake a systematic literature review of modelling approaches to assess air quality at transport hubs and identify literature gaps.	Chapter 2 reviewed statistical, emission and dispersion modelling approaches used for airports, ports, bus terminals, underground systems and rail stations. It identified key gaps in station-scale rail emission inventories, idling representation, and integrated rail–road dispersion modelling.
Objective 2: Collect air quality and traffic data at Heuston and Connolly stations.	Chapter 3 presented monitoring campaigns at both stations, including $PM_{2.5}$, PM_{10} and NO_2 measurements, meteorological data, EPA background data, train frequency data and detailed train operation records.

Objective 3: Apply statistical modelling methods to evaluate the effect of diesel trains and idling time on in-station air quality.	Chapter 3 used descriptive analysis, statistical screening and Random Forest models to show that background concentration set the main baseline, while diesel idling in enclosed platform areas was the most important local operational driver of PM and NO_2.
Objective 4: Develop a diesel train emissions calculation tool using train movement data and engine-state emission factors.	Chapter 5 developed a high-resolution emission calculation framework using train movement records, fleet information, engine state, notch-specific emission factors and station zoning to estimate NO_x and $PM_{2.5}$ emissions from diesel trains.
Objective 5: Quantify the spatial distribution of transport-related air pollution around stations and separate station and road contributions using ADMS-Urban.	Chapter 6 coupled the rail emission inventory with road traffic emissions, meteorology and background concentrations in ADMS-Urban. Scenario runs were used to distinguish rail-station and road-traffic contributions around Connolly Station.
Objective 6: Investigate technological or policy measures to reduce air pollution from rail stations.	Chapters 4–6 identified practical mitigation options, including idling reduction, improved shutdown guidance, cleaner fuels such as HVO, fleet transition, and better management of diesel operations in enclosed station areas.

Overall, the thesis fulfilled its objectives by linking monitoring, operational analysis, emission calculation and dispersion modelling into a single framework for assessing air quality in and around diesel railway stations. This integrated approach provides evidence for both local operational interventions and wider rail decarbonisation and air-quality policy.

7.3 Contribution to Knowledge

This thesis makes three main methodological contributions and two applied contributions to the evidence base on rail-station air quality.

First, it connects strands that previous work has treated separately. Existing station studies at London Paddington, Edinburgh Waverley, King's Cross and Birmingham New Street have shown that diesel train activity and ventilation strongly influence pollutant levels in enclosed or semi-enclosed stations, using monitoring combined with statistical or simple physical models (Chong et al., 2015;

Font et al., 2020b; Hickman et al., 2018). These studies focus on the in-station environment and do not develop a full, station-centred emissions–dispersion chain for the surrounding streets. This thesis extends that literature by linking (i) in-station monitoring, (ii) statistical attribution of operational drivers, (iii) detailed idling-event analysis from OTMR, (iv) physics-based station-area emissions calculation, and (v) neighbourhood-scale dispersion modelling in ADMS-Urban into a single, practical workflow that can be used to assess local mitigation at a mainline terminal.

Second, the work advances the way diesel passenger operations in stations are analysed by treating idling as a measurable, event-level behaviour constrained by the timetable. In previous station studies, train influence is usually represented by aggregated indicators such as train counts, service type or broad time-of-day factors (Chong et al., 2015; Font et al., 2020b; Hickman et al., 2018). Separately, work on shunting locomotives has examined idle-time distributions and their impact on fuel use and emissions in yards (Kałuża, 2016; Kałuża & Kucharski, 2018). In this thesis, OTMR logs are used to reconstruct explicit idle-time and low-speed episodes for each service and to match them to timetable gaps. Bayesian models then separate dwell that is structurally required by the schedule from discretionary engine-on time. Within the statistical models, train speed within the station (capturing idling and low-speed movement) is introduced as a predictor and emerges as the most influential operational driver of platform concentrations, ahead of simple train counts. This provides a quantitative basis for identifying when engines can be shut down and restarted during long layovers without changing service frequency.

Third, the thesis develops a high-resolution emissions calculation tool tailored to the Irish intercity and commuter fleets, and to the needs of station-scale dispersion modelling. Earlier rail studies have focused on deriving emission factors using PEMS or load tests for representative locomotives, or on route-scale inventories using generic duty cycles (Frey et al., 2022; Fuller, 2014; Kim et al., 2020). These are valuable for characterising engines and average duty cycles, but do not produce time-resolved emissions fields for specific station areas. The tool developed here combines OTMR-based speed and engine-state histories with fleet-specific emission factors to generate gridded emissions at 1-s resolution along detailed track geometries in the Connolly domain. It is designed to feed directly into ADMS-Urban and to support scenario testing, such as idling reduction or HVO adoption, under Irish operating conditions.

On the applied side, the thesis adds new evidence on what drives pollution in semi-enclosed platforms and concourses in a busy European terminal. It shows how timetable structure, accumulated idle time, fleet mix and wind patterns interact to shape NO_2 and $PM_{2.5}$ levels at

Heuston and Connolly. By combining calibrated monitoring with explicit operational attribution, the work helps to explain when and why semi-enclosed areas move from background-dominated to operation-dominated conditions, and which combinations of services and dwell patterns create the highest in-station concentrations.

Finally, the Connolly dispersion modelling provides a station-centred view of how diesel rail operations affect NO_2 and $PM_{2.5}$ at street level in a dense city-centre setting. The results show that, at the most exposed receptors along the station–road corridor, station-related increments can be similar in magnitude to nearby road-traffic increments, but that these impacts are highly localised and strongly dependent on wind direction and distance from the platforms. This supports a targeted mitigation approach that focuses on specific rail–road interface streets and the most exposed parts of the station, rather than treating station emissions as a diffuse contribution spread evenly across the wider city centre.

7.4 Discussion

The results across Chapters 3–6 point to a consistent conceptual picture: rail stations in dense urban cores sit on top of a regional and city-wide background, but semi-enclosed design and diesel operating practices can create a distinct local increment in both indoor and outdoor environments. This has two implications for policy. First, improving city-wide air quality will always benefit stations, because background concentrations dominate the baseline. Second, station operators still have control over a meaningful share of exposure through operational practices, particularly idling management in roofed areas and the allocation of cleaner rolling stock to the most enclosed platforms.

A key finding is that idling is not only an emissions issue but also a management and scheduling issue. The heavy-tailed distribution of idling events means that a relatively small number of long dwell periods can account for a disproportionate share of total engine-on time and therefore of emissions. This makes idling reduction attractive as a targeted intervention: focusing on the longest dwell events and on the fleets that contribute most per minute can deliver large gains without requiring wholesale operational change. At the same time, the station environment introduces constraints that are less prominent in roadside idling contexts, such as requirements for brake air pressure, system readiness and passenger comfort. Mitigation therefore needs to be framed as an operational protocol supported by evidence, rather than as a simple idle-time limit.

The emissions results reinforce the importance of fleet composition in Irish rail decarbonisation strategies. Although total station-area emissions are modest in national terms, they are concentrated in space and time and can create localised air quality impacts at the platforms and in the immediate station corridor. The contrast between diesel multiple units and locomotive-hauled services shows that technology choice matters at the station scale, as higher-emitting classes can dominate hotspot locations even when they represent a minority of movements. The HVO scenario illustrates that near-term fuel switching can provide an immediate emissions reduction pathway for the highest emitters, and can therefore act as a bridge measure while electrification and fleet renewal programmes progress.

The dispersion modelling provides a more nuanced understanding of how station emissions translate into neighbourhood impacts. The Connolly case study shows that the station signal is strongest along specific wind-aligned corridors and at receptors that are both close to the rail line and exposed to limited dispersion, such as canyon-like roads adjacent to the station. In contrast, receptors only a short distance away but outside the primary corridor can show minimal station contribution relative to road traffic. This spatial heterogeneity suggests that mitigation and monitoring should focus on the most affected corridors (e.g., along the rail line and immediately downwind) and should explicitly consider street geometry and prevailing wind patterns.

7.5 Limitations

As with any modelling study, the findings in this thesis are constrained by the amount and quality of data that could be collected. Within the time and resources available for a PhD project it was not possible to obtain complete, year-round and network-wide datasets for train operations, emissions, air quality, meteorology and traffic. Several parts of the analysis therefore rely on short monitoring campaigns, partial operational records and proxy emission factors. These constraints introduce uncertainty and mean that the results should be interpreted within the limitations set out below.

Temporal coverage and seasonality

The work is based on a set of focused campaigns rather than continuous monitoring over several years. The Heuston platform campaign was limited to one late spring / early summer period, and the Connolly campaign covered March–May. The diffusion-tube deployment around Connolly provides good spatial coverage but only for a single month. No winter campaign was carried out.

This means the study does not capture how colder temperatures, different demand patterns and more stable boundary layers in winter affect both emissions and dispersion. In practice, winter is likely to have higher auxiliary heating loads, more cold starts and more frequent poor dispersion conditions. The emissions and concentrations reported here are therefore likely to be on the low side compared to a true annual average, and should be treated as conservative.

Emissions data and emission factors

The station emission inventories are built from OTMR, timetables and published emission factors. OTMR data are not complete for all services, days and fleets, and some movements have to be inferred from timetables and other records. The mapping from notch to engine power is based on manufacturer information and assumes typical auxiliary loads and transmission performance; actual power output will vary with loading, gradient, maintenance condition and driving style.

The emission factors themselves come from measurements on similar fleets in other countries rather than dedicated testing of the Irish fleet. This is unavoidable but introduces uncertainty in absolute NO_x and $PM_{2.5}$ emissions and in the differences between classes. Operating states that are important in stations, such as long low-notch idling or repeated stop–start movements, are also not always well represented in the underlying measurements. The HVO scenarios rely on a small number of studies and should be seen as indicative rather than definitive.

Monitoring design and pollutant scope

The monitoring strategy uses a mix of fixed reference instruments, low-cost sensors and diffusion tubes. Even with co-location and calibration, the low-cost nodes have higher noise and drift than reference instruments and can only measure a limited set of pollutants. Diffusion tubes provide only period-mean NO_2 and do not capture short-term peaks. The analysis is based on concentrations at fixed sites and heights, so it does not describe personal exposure along whole journeys or in all parts of the stations and trains.

The pollutant scope is also limited. The focus is on NO_2 and $PM_{2.5}$ mass concentrations. Other metrics that are relevant for diesel exhaust, such as black carbon, ultrafine particle number or specific toxic components, are not included. Health effects are discussed qualitatively but no formal health impact assessment is carried out. As a result, the thesis captures important aspects of air quality around the stations but not the full range of pollutants or exposure situations.

Dispersion model, meteorology and background

The dispersion modelling uses a single local-scale Gaussian model (ADMS-Urban) with parameterised chemistry and turbulence. This type of model is designed for neighbourhood-scale work but cannot fully resolve the detailed air flow inside and immediately around semi-enclosed station sheds, canopies and footbridges. Building and canyon effects are represented using simplified modules rather than full three-dimensional flow calculations, so near-source concentrations at very specific points are more uncertain than area-wide patterns.

Meteorological input comes from Dublin Airport, which is several kilometres away in a more open setting. A short comparison with measurements at Connolly suggests that the airport series captures the main wind directions but tends to overestimate local wind speeds and does not fully reflect sheltering effects from surrounding buildings. Background concentrations are taken from Rathmines (urban background) and Phoenix Park (regional ozone), which are also some distance from the case-study area and sit in different immediate surroundings. These choices are standard but mean that background and meteorology represent wider urban and regional conditions rather than exact conditions at each receptor.

Model evaluation shows a positive bias for $PM_{2.5}$. This likely reflects a combination of background specification, uncertainties in primary emissions and the simplified treatment of secondary formation. The $PM_{2.5}$ results are therefore more reliable for comparing scenarios and locations than for estimating exact absolute levels.

Another limitation of the dispersion modelling is that validation relied mainly on fixed-point monitoring locations. These point measurements are useful for checking model performance at specific receptors, but they may not fully capture short-lived or spatially narrow plumes from train movements, especially under changing wind conditions. Future work should therefore combine fixed monitoring with mobile or transect-based plume sampling around the station. Plume-regression approaches could then be used to relate measured concentration increments to wind direction, distance from the tracks, and rail activity, providing a stronger independent check on the estimated rail contribution.

Spatial scope and generalisability

The detailed dispersion work is carried out for a single 2 × 2 km domain around Connolly, and the operational and emissions analysis focuses on two Dublin termini. These stations have particular designs, degrees of enclosure, surrounding land uses and traffic patterns. Other Irish stations, especially smaller rural stops or sub-surface urban stations, have different geometries and ventilation, and may behave quite differently.

The thesis does not present a full network-wide emissions or air-quality assessment. The conclusions are most directly relevant to semi-enclosed mainline terminals in dense urban settings under the current Irish diesel fleet and timetable. Changes in rolling stock, electrification, operating patterns or background air quality will all change the balance between mechanisms and may alter the size and shape of station-related increments.

7.6 Further Research

The limitations described above point to several clear directions for further work. These would strengthen the evidence base and extend the usefulness of the framework developed in this thesis.

Seasonal and network-wide monitoring

A first priority is to extend monitoring to cover different seasons and a broader set of stations. Winter campaigns at Connolly and Heuston would allow the size of the "winter penalty" in both emissions and concentrations to be quantified, and would test whether the spatial patterns seen in the spring campaigns hold under different conditions.

Shorter campaigns at a sample of other stations with contrasting designs and locations (for example, an open rural station, a sub-surface urban station and a small commuter stop) would help to set the Dublin results in a wider context and identify which findings are general and which are station-specific.

Emission measurements for the Irish fleet

The emission inventories would benefit from direct measurements on the Irish fleet. Portable emissions measurement systems on representative Class 22000, 201 and 29000 trains, combined with information on engine load and operating mode, would allow local emission factors to be derived for the operating states that matter most for stations: approaches, idling, departures and low-speed yard movements. Where possible, these measurements could be complemented by controlled tests in depots.

Such work would reduce reliance on proxy factors from other countries, improve estimates of NO_x and $PM_{2.5}$ for individual fleets and duty cycles, and give a firmer basis for assessing fuels such as HVO under Irish conditions.

Direct Rail-Emission Estimation Using Plume Sampling and Plume Regression

Future work should also consider direct rail-emission estimation using plume sampling combined with plume regression. Farren et al. (2024) developed a plume-regression approach in which fast-response roadside instruments measure exhaust plumes from passing vehicles, and regression analysis is then used to estimate emission factors and source contributions. Although this method was developed for road vehicles, the same principle could be adapted to railway stations by placing fast-response instruments near platforms, station exits, or trackside locations to capture train exhaust plumes. Measurements of NO_x, CO_2, particle number and black carbon could be combined with train identity, fuel type, operating mode, notch setting and wind conditions to estimate rail-specific emission factors. This would provide an independent check on the emission factors used in this thesis and reduce uncertainty in future rail dispersion modelling.

Behaviour, operations and mitigation

The analysis in this thesis suggests that timetable structure does not fully explain observed idling behaviour. Future work could therefore look more directly at how driving practice, local instructions and informal habits influence idling and engine shut-down. This could involve interviews and surveys with drivers and operations staff, as well as a more detailed analysis of OTMR records before and after changes to policies or training.

On the mitigation side, pilot trials of practical measures—such as clearer rules on shutdown during long dwell times, feedback on idling performance, or simple reminders linked to timetable gaps—could be carried out at one or two stations. Changes in idling time, emissions and platform concentrations before and after such interventions would provide direct evidence of what works in practice.

Improved modelling and links to health

The ADMS-Urban framework used here is suitable for neighbourhood-scale assessment, but there is scope to refine the modelling and to link it more directly to health outcomes. For selected locations and scenarios, more detailed flow modelling (for example, using CFD around parts of the station shed and adjacent streets) could be used to check and improve the representation of very local effects.

On the impact side, combining modelled concentrations with simple health-impact calculations for staff and regular passengers would translate the concentration changes into estimated changes in health risk. This would help station operators and policymakers weigh station-focused measures against other air-quality and decarbonisation actions.

Network-level application

Finally, the methods developed here are designed to be transferable. A natural next step is to scale up the operations–emissions–dispersion framework to a larger part of the network, using OTMR and timetable data from more lines and fleets. High-resolution dispersion modelling could then be applied to a small set of priority locations, while simpler tools are used elsewhere.

Such a network-level application would show where station-related emissions and air-quality impacts are most significant, and would give a clearer picture of how operational changes, fleet renewal, electrification and fuel switching could be combined to reduce both climate and air-quality impacts from rail in Ireland.

7.7 Overall Conclusion

This thesis provides a coherent, data-driven assessment of how diesel train operations influence air quality within stations and in the surrounding urban environment. Using Dublin's Connolly and Heuston stations as case studies, it shows that station pollution is shaped by a combination of regional background conditions and local operational behaviour, with idling in semi-enclosed areas emerging as a consistent driver of elevated exposure.

By developing a high-resolution emissions calculation tool and coupling it to a validated dispersion model, the research moves beyond descriptive monitoring and provides a quantitative basis for attributing neighbourhood impacts. The results show that station contributions to outdoor NO_2 can be locally important at receptors immediately adjacent to the rail corridor, but that these impacts are spatially limited and strongly conditioned by wind and street geometry. This points towards targeted mitigation: prioritising the most enclosed platforms, the longest dwell events, and the highest-emitting fleets.

In the Irish context, where electrification remains limited, the findings support a pragmatic mitigation pathway that can operate in parallel with longer-term decarbonisation. Operational idling protocols, fleet management and near-term fuel measures such as HVO adoption offer actionable levers to reduce exposure for passengers, workers and nearby residents. More broadly, the framework developed in this thesis can be applied to other Irish stations and can support evidence-based decisions on how to manage diesel traction impacts while the rail network transitions towards lower-emission technologies.

LIST OF TABLES

LIST OF FIGURES

BIBLIOGRAPHY

Aatola, H., Larmi, M., Sarjovaara, T., & Mikkonen, S. (2008). Hydrotreated vegetable oil (HVO) as a renewable diesel fuel: Trade-off between NOx, particulate emission, and fuel consumption of a heavy duty engine. *SAE International Journal of Engines*, *1*(1), 1251–1262. https://doi.org/10.4271/2008-01-2500

Abbasi, S., Olander, L., Larsson, C., Olofsson, U., Jansson, A., & Sellgren, U. (2012). A field test study of airborne wear particles from a running regional train. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *226*(1), 95–109. https://doi.org/10.1177/0954409711408774

Abrams, D., Lalot, F., Hopthrow, T., Templeton, A., Steeden, B., Özkeçeci, H., Imada, H., Warbis, S., Sandiford, D., Meleady, R., Fell, E., Abrams, Z., Abrams, A., Ngan, X. Q., Celina, S., Tanyeri, A., Gammon, M., Abrams, B., Fischer, L., … Peckham, S. (2021). Cleaning up our acts: Psychological interventions to reduce engine idling and improve air quality. *Journal of Environmental Psychology*, *74*, 101587. https://doi.org/10.1016/j.jenvp.2021.101587

Abrutytė, E., Žukauskaitė, A., Mickevičienė, R., Zabukas, V., & Paulauskienė, T. (2014). Evaluation of NOx emission and dispersion from marine ships in Klaipeda sea port. *Journal of Environmental Engineering and Landscape Management*. (world). https://www.tandfonline.com/doi/abs/10.3846/16486897.2014.892009

Ahmad, F., Majumder, D., Ranjit, R., & Manhart, M. (2025). Preliminary study on the spread of air-borne pollutants in urban environment: A CFD simulation approach. *Scientific Reports*, *15*, 18836. https://doi.org/10.1038/s41598-025-03197-z

Air pollution. (n.d.). Retrieved 23 December 2025, from https://www.who.int/health-topics/air-pollution

Ambient (outdoor) air pollution. (n.d.). Retrieved 23 December 2025, from https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

Apte, J. S., Messier, K. P., Gani, S., Brauer, M., Kirchstetter, T. W., Lunden, M. M., Marshall, J. D., Portier, C. J., Vermeulen, R., & Hamburg, S. P. (2017). High-Resolution Air Pollution Mapping with

Google Street View Cars: Exploiting Big Data. *Environmental Science & Technology*, *51*(12), 6999–7008. https://doi.org/10.1021/acs.est.7b00891

Arunachalam, S., Valencia, A., Woody, M. C., Snyder, M. G., Huang, J., Weil, J., Soucacos, P., Webb, S., Airport Cooperative Research Program, Transportation Research Board, & National Academies of Sciences, Engineering, and Medicine. (2017). *Dispersion Modeling Guidance for Airports Addressing Local Air Quality Health Concerns* (p. 24881). Transportation Research Board. https://doi.org/10.17226/24881

Biess, L., Stewart, T., Miller, D., & Fritz, S. (2009). *Reducing Locomotive Idle Fuel Consumption and Exhaust Emissions by Applying an Auxiliary Power Unit (APU)*. 603–608. https://doi.org/10.1115/ICES2003-0550

Bo, X., Xue, X., Xu, J., Du, X., Zhou, B., & Tang, L. (2019). Aviation's emissions and contribution to the air quality in China. *Atmospheric Environment*, *201*, 121–131. https://doi.org/10.1016/j.atmosenv.2019.01.005

Boddy, J. W. D., Smalley, R. J., Goodman, P. S., Tate, J. E., Bell, M. C., & Tomlin, A. S. (2005). The spatial variability in concentrations of a traffic-related pollutant in two street canyons in York, UK–Part II: The influence of traffic characteristics. *Atmospheric Environment*, *39*(17), 3163–3176. https://doi.org/10.1016/j.atmosenv.2005.01.044

Bohac, S., Feiler, E., & Bradbury, I. (2012). Exhaust Emissions Characterization of a Turbocharged 2-Stroke Tier 0+ Locomotive Engine: NOx, Particulate Matter and Soluble Organic Fraction Composition. *Journal of Engineering for Gas Turbines and Power*, *134*, 072804. https://doi.org/10.1115/1.4006006

Bohl, T. (2018). Particulate number and NOx trade-off comparisons between HVO and mineral diesel in HD applications. *Fuel*, *215*, 90–101. https://doi.org/10.1016/j.fuel.2017.11.023

Bolourchi, A., Atabi, F., Moattar, F., & Ehyaei, M. A. (2018). Experimental and numerical analyses of particulate matter concentrations in underground subway station. *International Journal of Environmental Science and Technology*, *15*(12), 2569–2580. https://doi.org/10.1007/s13762-017-1545-2

Boogaard, H., Patton, A. P., Atkinson, R. W., Brook, J. R., Chang, H. H., Crouse, D. L., & Forastiere, F. (2022). Long-term exposure to traffic-related air pollution and selected health outcomes: A systematic review and meta-analysis. *Environment International*, *164*, 107262. https://doi.org/10.1016/j.envint.2022.107262

Borken-Kleefeld, J., & Ntziachristos, L. (2012, June). *The potential for further controls of emissions from mobile sources in Europe* [Other]. http://gains.iiasa.ac.at/images/stories/reports/TSAP/TSAP-TRANSPORT-20120612.pdf

Brecher, A., Rhoads, T., & Simpson, C. (2014). *Best Practices and Strategies for Improving Rail Energy Efficiency*. U.S. Department of Transportation, Federal Railroad Administration. https://railroads.dot.gov/elibrary/best-practices-and-strategies-improving-rail-energy-efficiency

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brown, S. S., & Stutz, J. (2012). Nighttime radical observations and chemistry. *Chemical Society Reviews*, *41*(19), 6405. https://doi.org/10.1039/c2cs35181a

Buccolieri, R., Cesari, R., Dinoi, A., Maurizi, A., Tampieri, F., & Sabatino, S. D. (2016). Impact of ship emissions on local air quality in a Mediterranean city's harbour after the European sulphur directive. *International Journal of Environment and Pollution*, *59*(1), 30–42. https://doi.org/10.1504/IJEP.2016.078061

Büchel, B., & Corman, F. (2020). Review on Statistical Modeling of Travel Time Variability for Road-Based Public Transport. *Frontiers in Built Environment*, *6*, 70.

Butterfield, D., Martin, N. A., Coppin, G., & Fryer, D. E. (2021). Equivalence of UK nitrogen dioxide diffusion tube data to the EU reference method. *Atmospheric Environment*, *262*, 118614. https://doi.org/10.1016/j.atmosenv.2021.118614

Cambridge Environmental Research Consultants & Environmental Protection Agency. (2023). *Regional-to-local scale air quality modelling of the Republic of Ireland*. Cambridge Environmental Research Consultants Ltd / Environmental Protection Agency. https://www.cerc.co.uk/environmental-software/assets/data/doc_validation/CERC_AQM_Republic_of_Ireland.pdf

Carslaw, D. C., Williams, M. L., & Barratt, B. (2012). A short-term intervention study—Impact of airport closure due to the eruption of Eyjafjallajökull on near-field air quality. *Atmospheric Environment*, *54*, 328–336. https://doi.org/10.1016/j.atmosenv.2012.02.020

Centre for Low Emission Construction. (2019). *Anti-idling campaigns*. https://clec.uk/advice/mitigation/anti-idling-campaigns

Cesari, R., Buccolieri, R., Dinoi, A., Maurizi, A., Landi, T., & Di Sabatino, S. (2018). *Influence of Ship Emissions on Ozone Concentration in a Mediterranean Area: A Modelling Approach*. 317–321. https://doi.org/10.1007/978-3-319-57645-9_50

Cesari, R., Genga, A., Buccolieri, R., Di Sabatino, S., Siciliano, M., Siciliano, T., Dinoi, A., Maurizi, A., & Ielpo, P. (2021). Combining chemical composition data and numerical modelling for the assessment of air quality in a mediterranean port city. *Applied Sciences (Switzerland)*, *11*(5), 1–26. https://doi.org/10.3390/app11052181

Chan, A. T., So, E. S. P., & Samad, S. C. (2001). Strategic guidelines for street canyon geometry to achieve sustainable street air quality. *Atmospheric Environment*, *35*(24), 4089–4098. https://doi.org/10.1016/S1352-2310(01)00212-6

Chen, D., Wang, X., Nelson, P., Li, Y., Zhao, N., Zhao, Y., Lang, J., Zhou, Y., & Guo, X. (2017). Ship emission inventory and its impact on the PM2.5 air pollution in Qingdao Port, North China. *Atmospheric Environment*, *166*, 351–361. https://doi.org/10.1016/j.atmosenv.2017.07.021

Chen, J., Mölter, A., Gómez-Barrón, J. P., O'Connor, D., & Pilla, F. (2024). Evaluating background and local contributions and identifying traffic-related pollutant hotspots: Insights from Google Air View mobile monitoring in Dublin, Ireland. *Environmental Science and Pollution Research*, *31*(44), 56114–56129. https://doi.org/10.1007/s11356-024-34903-5

Chong, U., Swanson, J., & Boies, A. (2015a). Air quality evaluation of London Paddington train station. *Environmental Research Letters*, *10*. https://doi.org/2020031315125700

Chong, U., Swanson, J. J., & Boies, A. M. (2015b). Air quality evaluation of London Paddington train station. *Environmental Research Letters*, *10*(9), 094012. https://doi.org/10.1088/1748-9326/10/9/094012

CLEAR: Fleet wide assessment of rail emissions factors (T1187). (n.d.). Retrieved 1 December 2025, from https://www.rssb.co.uk/research-catalogue/CatalogueItem/T1187

Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, *74*(368), 829–836. https://doi.org/10.1080/01621459.1979.10481038

Cleveland, W. S., & Devlin, S. J. (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*, *83*(403), 596–610. https://doi.org/10.1080/01621459.1988.10478639

CMAS: Community Modeling and Analysis System. (n.d.). Retrieved 21 January 2025, from https://www.cmascenter.org/help/documentation.cfm?MODEL=c-port&VERSION=3.0

Cohan, A., Wu, J., & Dabdub, D. (2011). High–resolution pollutant transport in the San Pedro Bay of California. *Atmospheric Pollution Research*, *2*(3), 237–246. https://doi.org/10.5094/APR.2011.030

Collett, R. S., & Oduyemi, K. (1997). Air quality modelling: A technical review of mathematical approaches. *Meteorological Applications*, *4*(3), 235–246. https://doi.org/10.1017/S1350482797000455

da Fonseca-Soares, D., Eliziário, S. A., Galvincio, J. D., & Ramos-Ridao, A. F. (2024). Greenhouse Gas Emissions in Railways: Systematic Review of Research Progress. *Buildings*, *14*(2), 539. https://doi.org/10.3390/buildings14020539

Dėdelė, A., & Miškinytė, A. (2015). The statistical evaluation and comparison of ADMS-Urban model for the prediction of nitrogen dioxide with air quality monitoring network. *Environmental Monitoring and Assessment*, *187*, 578. https://doi.org/10.1007/s10661-015-4675-0

DEFRA, F. & R. A. and D. A. (2023). *Emissions Factors Toolkit v12 user guide*. https://laqm.defra.gov.uk/wp-content/uploads/2023/11/EFTv12.0-user-guide-v1.0.pdf

Delgado-Saborit, J. M., Lim, S., Hickman, A., Baker, C., Barratt, B., Cai, X., Font, A., Heal, M. R., Lin, C., Thornes, J. E., Woods, M., & Green, D. (2022). Factors affecting occupational black carbon exposure in enclosed railway stations. *Atmospheric Environment*, *289*, 119301. https://doi.org/10.1016/j.atmosenv.2022.119301

Department for Environment, F. & R. A. (2024). *Monitoring ambient air: Techniques and standards*. https://www.gov.uk/guidance/monitoring-ambient-air-techniques-and-standards

Department of Transport, D. of T. (2021). *Five Cities Demand Management Study*. https://assets.gov.ie/static/documents/final-recommendations-report.pdf

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, L 152 Official Journal of the European Union 1 (2008). https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32008L0050

Dobrzyńska, E., Szewczyńska, M., Pośniak, M., Szczotka, A., Puchałka, B., & Woodburn, J. (2020). Exhaust emissions from diesel engines fueled by different blends with the addition of nanomodifiers and hydrotreated vegetable oil HVO. *Environmental Pollution*, *265*, 113772–113772. https://doi.org/10.1016/j.envpol.2019.113772

Dublin City Council. (2024). *Traffic volumes from SCATS traffic management system Jan–Jun 2024 (DCC SCATS detector volume Jan–Jun 2024)* [Data set]. Smart Dublin Open Data Portal. https://data.smartdublin.ie/dataset/dcc-scats-detector-volume-jan-jun-2024

Durand, A., Mehel, A., Fokoua, G., Murzyn, F., Puech, S., & Larrarte, F. (2021). Numerical and experimental investigations on brake particle dispersion in the flow generated by a train in an underground station. *Atmospheric Pollution Research*, *12*(10), 101189. https://doi.org/10.1016/j.apr.2021.101189

EEA. (2019). [Data Visualization]. https://www.eea.europa.eu/data-and-maps/daviz/contribution-of-the-transport-sector-6#tab-chart_4

EEA. (2020). [Publication]. https://www.eea.europa.eu/publications/air-quality-in-europe-2020-report

Ekmekçioğlu, A., Kuzu, S. L., Ünlügençoğlu, K., & Çelebi, U. B. (2020). Assessment of shipping emission factors through monitoring and modelling studies. *Science of The Total Environment*, *743*, 140742. https://doi.org/10.1016/j.scitotenv.2020.140742

eLichens. (2020). *A novel recalibration technology for air quality microsensors: Technical White Paper – October 2020*. eLichens. https://static1.squarespace.com/static/5a85cfd1b1ffb6a07a21b563/t/6036196d81609d07de932794/1614158194800/eLichens_Technical+White+Paper_eLos_Oct20.pdf

eLichens. (2026). *eLichens eLos—Outdoor air quality station*. https://www.elichens.com/elos-outdoor-air-quality-station

Environmental Protection Agency. (2019). *Urban Environmental Indicators: Nitrogen Dioxide Levels in Dublin* (Nos 978-1-84095-846–1). Environmental Protection Agency. https://www.epa.ie/publications/monitoring--assessment/air/urban-environmental-indicators-nitrogen-dioxide-levels-in-dublin.php

Environmental Protection Agency. (2020). *Ireland's Environment 2020 – An Assessment* (Nos 978-1-84095-953–6). Environmental Protection Agency. https://www.epa.ie/publications/monitoring--assessment/assessment/state-of-the-environment/irelands-environment-2020---an-assessment----report.php

EPA. (2023a). *Regional-to-local scale air qualitymodelling of the Republic oflreland* (p. 25).

European Commission. (2021). https://alternative-fuels-observatory.ec.europa.eu/transport-mode/rail

European Environment Agency. (2023a). *Europe's air quality status 2023*. European Environment Agency. https://www.eea.europa.eu/publications/europes-air-quality-status-2023

European Environment Agency. (2023b). *Europe's air quality status 2023* (EEA Briefing). European Environment Agency. https://www.eea.europa.eu/en/analysis/publications/europes-air-quality-status-2023

Farren, N. J., Wilson, S., Bernard, Y., Shaw, M. D., Lee, K., Crowe, M., & Carslaw, D. C. (2024). An

Ambient Measurement Technique for Vehicle Emission Quantification and Concentration Source Apportionment. *Environmental Science & Technology*, *58*(45), 20091–20100. https://doi.org/10.1021/acs.est.4c07907

Font, A., Tremper, A. H., Lin, C., Priestman, M., Marsh, D., Woods, M., Heal, M. R., & Green, D. C. (2020a). Air quality in enclosed railway stations: Quantifying the impact of diesel trains through deployment of multi-site measurement and random forest modelling. *Environmental Pollution*, *262*, 114284. https://doi.org/10.1016/j.envpol.2020.114284

Font, A., Tremper, A. H., Lin, C., Priestman, M., Marsh, D., Woods, M., Heal, M. R., & Green, D. C. (2020b). Air quality in enclosed railway stations: Quantifying the impact of diesel trains through deployment of multi-site measurement and random forest modelling. *Environmental Pollution*, *262*, 114284. https://doi.org/10.1016/j.envpol.2020.114284

Frey, H. C., Grieshop, A. P., Khlystov, A., Bang, J. J., Rouphail, N., Guinness, J., Rodriguez, D., Fuentes, M., Saha, P., Brantley, H., Snyder, M., Tanvir, S., Ko, K., Noussi, T., Delavarrafiee, M., & Singh, S. (2022). Characterizing Determinants of Near-Road Ambient Air Quality for an Urban Intersection and a Freeway Site. *Research Report (Health Effects Institute)*, (207), 1–73.

Fridell, E., Ferm, M., & Ekberg, A. (2010a). Emissions of particulate matters from railways – Emission factors and condition monitoring. *Transportation Research Part D: Transport and Environment*, *15*(4), 240–245. https://doi.org/10.1016/j.trd.2010.02.006

Fridell, E., Ferm, M., & Ekberg, A. (2010b). Emissions of particulate matters from railways – Emission factors and condition monitoring. *Transportation Research Part D: Transport and Environment*, *15*(4), 240–245. https://doi.org/10.1016/j.trd.2010.02.006

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Fuller, G. (2014). *Air pollution emissions from diesel trains in London*. https://www.londonair.org.uk/london/reports/AirPollutionEmissionsFromDieselTrainsinLondon_August2014.pdf

Fussell, J. C., Franklin, M., Green, D. C., Gustafsson, M., Harrison, R. M., Hicks, W., Kelly, F. J., Kishta, F., Miller, M. R., Mudway, I. S., Oroumiyeh, F., & Selley, L. (2022). A review of road traffic-derived non-exhaust particles: Emissions, physicochemical characteristics, health risks, and mitigation measures. *Environmental Science & Technology*, *56*(11), 6813–6835. https://doi.org/10.1021/acs.est.2c01072

Gaeta, A., Cattani, G., Di Menno di Bucchianico, A., De Santis, A., Cesaroni, G., Badaloni, C., Ancona, C., Forastiere, F., Sozzi, R., Bolignano, A., & Sacco, F. (2016). Development of nitrogen dioxide and volatile organic compounds land use regression models to estimate air pollution exposure near an Italian airport. *Atmospheric Environment*, *131*, 254–262. https://doi.org/10.1016/j.atmosenv.2016.01.052

Gallagher, J. (2016). A modelling exercise to examine variations of NOx concentrations on adjacent footpaths in a street canyon: The importance of accounting for wind conditions and fleet composition. *Science of The Total Environment*, *550*, 1065–1074. https://doi.org/10.1016/j.scitotenv.2016.01.096

Gao, Z., & Zhou, X. (2024). A review of the CAMx, CMAQ, WRF-Chem and NAQPMS models: Application, evaluation and uncertainty factors. *Environmental Pollution*, *343*, 123183. https://doi.org/10.1016/j.envpol.2023.123183

Gasser, M., Riediker, M., Mueller, L., Perrenoud, A., Blank, F., Gehr, P., & Rothen-Rutishauser, B. (2009). Toxic effects of brake wear particles on epithelial lung cells in vitro. *Particle and Fibre Toxicology*, *6*, 30. https://doi.org/10.1186/1743-8977-6-30

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd edn). Chapman and Hall/CRC.

Genikhovich, E. L. (2004). Statistical Approach to Deterministic Modelling of Air Pollution and its Applications in Russia. In C. Borrego & S. Incecik (Eds), *Air Pollution Modeling and Its Application XVI* (pp. 3–17). Springer US. https://doi.org/10.1007/978-1-4419-8867-6_1

Gibson, M. D., Kundu, S., & Satish, M. (2013). Dispersion model evaluation of PM2.5, NOx and SO2 from point and major line sources in Nova Scotia, Canada using AERMOD Gaussian plume air dispersion model. *Atmospheric Pollution Research*, *4*(2), 157–167. https://doi.org/10.5094/APR.2013.016

Google & Dublin City Council. (2023). *Google Air View data – Dublin City* [Data set]. data.gov.ie / Smart Dublin Open Data Portal. https://data.gov.ie/dataset/google-airview-data-dublin-city

Gradko International Ltd. (2024). *Nitrogen Dioxide Diffusion Tubes*. https://www.gradko.com/environmental/nitrogen-dioxide-diffusion-tubes/

Graver, B. M., & Frey, H. C. (2013). Comparison of locomotive emissions measured during dynamometer versus rail yard engine load tests. *Transportation Research Record*, *2341*(1), 23–33. https://doi.org/10.3141/2341-03

Graver, B. M., Frey, H. C., & Hu, J. (2016). Effect of biodiesel fuels on real-world emissions of passenger rail locomotives. *Environmental Science & Technology*, *50*(21), 12030–12039. https://doi.org/10.1021/acs.est.6b03567

Green Biofuels Ltd. (2023). *Comparative study of emissions of Irish Rail Locomotive Fleet Number 218 with Diesel and Gd+ HVO (Inchicore Depot)* [Unpublished confidential report shared by Irish Rail].

Groma, V. O., Ferenczi, Z., Osán, J., Török, S., & Steib, R. (2018). Verification of the EDMS model adapted to Budapest Liszt Ferenc Airport. *International Journal of Environment and Pollution*, *63*(3), 137–153. https://doi.org/10.1504/IJEP.2018.097308

Han, B., Wang, L., Deng, Z., Shi, Y., & Yu, J. (2022). Source emission and attribution of a large airport in Central China. *Science of The Total Environment*, *829*, 154519. https://doi.org/10.1016/j.scitotenv.2022.154519

Harrison, R. M. (2020). Airborne particulate matter. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *378*(2183), 20190319–20190319. https://doi.org/10.1098/rsta.2019.0319

Hart, J. E., Laden, F., Schenker, M. B., & Garshick, E. (2006). Chronic obstructive pulmonary disease mortality in diesel-exposed railroad workers. *Environmental Health Perspectives*, *114*(7), 1013–1017. https://doi.org/10.1289/ehp.8743

Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Science*, *1*(3), 297–310.

https://doi.org/10.1214/ss/1177013604

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd edn). Springer.

Heal, M., Laxen, D., & Marner, B. (2018). *A Review of Biases in the Measurement of Ambient Nitrogen Dioxide (NO2) by Palmes Passive Diffusion Tube*. Scottish Government. https://www.scottishairquality.scot/sites/default/files/orig/publications/technical-reports/pdt_bias_lit_review_v2018-11-07_final.pdf

Health Effects Institute. (2024a). *State of Global Air 2024: A special report on global exposure to air pollution and its health impacts*. Health Effects Institute. https://www.stateofglobalair.org

Health Effects Institute. (2024b). *State of Global Air 2024: Special Report*. Health Effects Institute. https://www.stateofglobalair.org/sites/default/files/documents/2024-06/soga-2024-report_0.pdf

Henry-Lheureux, T., Seers, P., Ghedhaïfi, W., & Garnier, F. (2021). Overview of Emissions at Montreal's Pierre Elliott Trudeau International Airport and Impact of Local Weather on Related Pollutant Concentrations. *Water, Air, & Soil Pollution*, *232*(5), 173. https://doi.org/10.1007/s11270-021-05087-2

Hickman, A., Baker, C., Cai, X., Delgado-Saborit, J., & Thornes, J. (2018). Evaluation of air quality at the Birmingham New Street Railway Station. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *232*(6), 1864–1878. https://doi.org/10.1177/0954409717752180

Hickman, A. L., Baker, C. J., Cai, X., Delgado-Saborit, J. M., & Thornes, J. E. (2018). Evaluation of air quality at the Birmingham New Street railway station. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *232*(6), 1864–1877. https://doi.org/10.1177/0954409717752180

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., & Briggs, D. (2008). A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, *42*(33), 7561–7578. https://doi.org/10.1016/j.atmosenv.2008.05.057

Holmes, N. S., & Morawska, L. (2006). A review of dispersion modelling and its application to the dispersion of particles: An overview of different dispersion models available. *Atmospheric Environment*, *40*(30), 5902–5928. https://doi.org/10.1016/j.atmosenv.2006.06.003

Hu, Z., Wang, Z., Lu, Z., Fu, J., & Quan, Y. (2022). Effect of idle-start-go on the economy and emissions of gasoline vehicles. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, *238*, 095440702211218. https://doi.org/10.1177/09544070221121868

IARC: DIESEL ENGINE EXHAUST CARCINOGENIC. (2012). https://www.iarc.who.int/news-events/iarc-diesel-engine-exhaust-carcinogenic

Iarnród Éireann Annual Report 2022. (2022).

Irish Rail, I. (2020). *DART+ South West Technical Optioneering Report* (p. 17). chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://assets-eu-01.kc-usercontent.com/a439a0e7-eaa3-0154-d646-f6520be9179b/8eac36c6-69f1-4a5e-be6a-04c0a70d2b49/35a3fb47-92d8-4d2d-b205-74b906db1ce3.pdf

Isakov, V., Barzyk, T. M., Smith, E. R., Arunachalam, S., Naess, B., & Venkatram, A. (2017). A web-based screening tool for near-port air quality assessments. *Environmental Modelling & Software*, *98*, 21–34. https://doi.org/10.1016/j.envsoft.2017.09.004

Isukapati, I. K., Igoe, C., Bronstein, E., Parimi, V., & Smith, S. F. (2021). Hierarchical Bayesian Framework for Bus Dwell Time Prediction. *IEEE Transactions on Intelligent Transportation Systems*, *22*, 3068–3077. https://doi.org/10.1109/TITS.2020.2979390

Izadi, T., Mehrabian, M. A., Ahmadi, G., & Abouali, O. (2021). Numerical analysis of the mirco-particles distribution inside an underground subway system due to train piston effect. *Journal of Wind Engineering and Industrial Aerodynamics*, *211*, 104533. https://doi.org/10.1016/j.jweia.2021.104533

Jin, M.-Y., Apsunde, K. A., Broderick, B., Peng, Z.-R., He, H.-D., & Gallagher, J. (2024). Evaluating the impact of evolving green and grey urban infrastructure on local particulate pollution around city square parks. *Scientific Reports*, *14*(1), 18528. https://doi.org/10.1038/s41598-024-68252-7

Jintanakul, K., Chu, L., & Jayakrishnan, R. (2009). Bayesian Mixture Model for Estimating Freeway Travel Time Distributions from Small Probe Samples from Multiple Days. *Transportation Research Record*, *2136*(1), 37–44. https://doi.org/10.3141/2136-05

Johansson, L., Karppinen, A., Kurppa, M., Kousa, A., Niemi, J. V., & Kukkonen, J. (2022). An operational urban air quality model ENFUSER, based on dispersion modelling and data assimilation. *Environmental Modelling & Software*, *156*, 105460. https://doi.org/10.1016/j.envsoft.2022.105460

Johnson, J. (2022, August 14). *An Introduction to Atmospheric Pollutant Dispersion Modelling*. https://doi.org/10.3390/ecas2022-12826

Kalhor, M., & Bajoghli, M. (2017). Comparison of AERMOD, ADMS and ISC3 for incomplete upper air meteorological data (case study: Steel plant). *Atmospheric Pollution Research*, *8*(6), 1203–1208. https://doi.org/10.1016/j.apr.2017.06.001

Kałuża, A. (2016). General statistics of diesel engines' idle time: Shunting locomotives in industrial sidings in Poland 2009…2013. *Transportation Research Part D: Transport and Environment*, *49*, 82–93. https://doi.org/10.1016/j.trd.2016.08.031

Kałuża, A., & Kucharski, R. (2018). Locomotive diesel engine idle time distributions: Case studies from Poland, 2009–2013. *Transportation Research Part D: Transport and Environment*, *62*, 524–535. https://doi.org/10.1016/j.trd.2018.02.019

Kamarianakis, Y., Gao, H. O., Holmén, B. A., & Sonntag, D. B. (2011). Robust modeling and forecasting of diesel particle number emissions rates. *Transportation Research Part D: Transport and Environment*, *16*(6), 435–443. https://doi.org/10.1016/j.trd.2011.04.005

Karrouchi, M., Rhiat, M., Nasri, I., Atmane, I., Hirech, K., Messaoudi, A., Melhaoui, M., & Kassmi, K. (2023). Practical investigation and evaluation of the Start/Stop system's impact on the engine's fuel use, noise output, and pollutant emissions. *E-Prime - Advances in Electrical Engineering, Electronics and Energy*, *6*, 100310. https://doi.org/10.1016/j.prime.2023.100310

Khreis, H., Williams, H., Abdollahpour, S. S., van den Bosch, M., Mudu, P., Tainio, M., & Hankey, S. (2024). The nexus of transportation, the built environment, air pollution and health. *Cities & Health*. https://doi.org/10.1080/23748834.2024.2376389

Kilburn, K. H. (2000). Effects of diesel exhaust on neurobehavioral and pulmonary functions. *Archives of Environmental Health*, *55*(1), 11–17. https://doi.org/10.1080/00039890009603379

Kim, M. K., Park, D., Kim, M., Heo, J., Park, S., & Chong, H. (2020). A study on characteristic emission factors of exhaust gas from diesel locomotives. *International Journal of Environmental Research and Public Health*, *17*(11), 1–12. https://doi.org/10.3390/ijerph17113788

Kim, M.-K., Park, D., Kim, M., Heo, J., Park, S., & Chong, H. (2020). A Study on Characteristic Emission Factors of Exhaust Gas from Diesel Locomotives. *International Journal of Environmental Research and Public Health*, *17*(11), 3788. https://doi.org/10.3390/ijerph17113788

Kontos, S., Liora, N., Poupkou, A., Giannaros, C., Garane, A., & Melas, D. (2017). Air-Quality Impact of Cruise and Passenger Ship Emissions in the Port of Thessaloniki. In T. Karacostas, A. Bais, & P. T. Nastos (Eds), *Perspectives on Atmospheric Sciences* (pp. 1129–1134). Springer International Publishing. https://doi.org/10.1007/978-3-319-35095-0_162

KORTAS, P., & KROPIWNICKI, J. (2013). A proposal of a description of the operating conditions of diesel-electric locomotives. *Combustion Engines*, *155*(4), 56–65. https://doi.org/10.19206/CE-116977

Kose, S., Sekban, D. M., & Ozkok, M. (2022). Determination of port-induced exhaust gas emission amounts and investigation of environmental impact by creating emission maps: Sample of Trabzon port. *International Journal of Sustainable Transportation*, *16*(3), 258–268. https://doi.org/10.1080/15568318.2020.1871130

Koulidis, A. G., Progiou, A. G., & Ziomas, I. C. (2020). Air Quality Levels in the Vicinity of Three Major Greek Airports. *Environmental Modeling & Assessment*, *25*(6), 749–760. https://doi.org/10.1007/s10666-020-09699-6

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Kumar, P., Ketzel, M., Vardoulakis, S., Pirjola, L., & Britter, R. (2011). Dynamics and dispersion modelling of nanoparticles from road traffic in the urban atmospheric environment—A review.

Journal of Aerosol Science, *42*(9), 580–603. https://doi.org/10.1016/j.jaerosci.2011.06.001

Kuzu, S. (2018). Estimation and dispersion modeling of landing and take-off (LTO) cycle emissions from Ataturk International Airport. *AIR QUALITY ATMOSPHERE AND HEALTH*, *11*(2). https://doi.org/10.1007/s11869-017-0525-5

Kuzu, S. L., Bilgili, L., & Kiliç, A. (2021). Estimation and dispersion analysis of shipping emissions in Bandirma Port, Turkey. *Environment, Development and Sustainability*, *23*(7), 10288–10308. https://doi.org/10.1007/s10668-020-01057-6

Lalot, F., Imada, H., Hopthrow, T., & Abrams, D. (2025). Driving change: A scoping review of psychological interventions to reduce engine idling and air pollution. *Journal of Environmental Psychology*, *106*, 102691–102691. https://doi.org/10.1016/j.jenvp.2025.102691

Lebedevas, S., Dailydka, S., Jastremskas, V., & Rapalis, P. (2015). The influence of locomotive diesel engine transient operating modes on energy usage. *Transportation Research Part D: Transport and Environment*, *34*, 219–229. https://doi.org/10.1016/j.trd.2014.10.011

Lee, K. K., Spath, N., Miller, M. R., Mills, N. L., & Shah, A. S. V. (2020). Short-term exposure to carbon monoxide and myocardial infarction: A systematic review and meta-analysis. *Environment International*, *143*, 105901. https://doi.org/10.1016/j.envint.2020.105901

Lim, J., Lim, C., & Jung, S. (2021). Characterizations of Size-segregated Ultrafine Particles in Diesel Exhaust. *Aerosol and Air Quality Research*, *21*(5), 200356. https://doi.org/10.4209/aaqr.200356

Liu, J., Zhou, J., Zhou, J., Li, M., Chen, E., Jiang, G., Chen, Y., Wu, J., & Yang, Q. (2020). Fine particulate matter exposure induces DNA damage by downregulating Rad51 expression in human bronchial epithelial Beas-2B cells in vitro. *Toxicology*, *444*, 152581. https://doi.org/10.1016/j.tox.2020.152581

Loxham, M., & Nieuwenhuijsen, M. J. (2019). Health effects of particulate matter air pollution in underground railway systems: A critical review of the evidence. *Particle and Fibre Toxicology*, *16*(1), 12–12. https://doi.org/10.1186/s12989-019-0296-2

Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012). *The BUGS Book: A Practical*

Introduction to Bayesian Analysis. CRC Press.

Luo, T., Liu, X., & Jin, H. (2022). Bus Queue Time Estimation Model for a Curbside Bus Stop Considering the Blocking Effect. *Scientific Reports*, *12*, 11576. https://doi.org/10.1038/s41598-022-15485-z

Lyu, P., Wang, P. (Slade), Liu, Y., & Wang, Y. (2021). Review of the studies on emission evaluation approaches for operating vehicles. *Journal of Traffic and Transportation Engineering (English Edition)*, *8*(4), 493–509. https://doi.org/10.1016/j.jtte.2021.07.004

Ma, Z., He, C., Liu, X., Li, J., Liu, M., & Wei, H. (2020). Effect of the Engine Working Condition on the Vehicle Emissions Based on Real-World Driving. In V. Jain, S. Patnaik, F. Popențiu Vlădicescu, & I. K. Sethi (Eds), *Recent Trends in Intelligent Computing, Communication and Devices* (pp. 975–981). Springer. https://doi.org/10.1007/978-981-13-9406-5_117

Makridis, M., & Lazaridis, M. (2019). Dispersion modeling of gaseous and particulate matter emissions from aircraft activity at Chania Airport, Greece. *Air Quality, Atmosphere & Health*, *12*(8), 933–943. https://doi.org/10.1007/s11869-019-00710-y

Mallemoggala, P. S., Gummadi, S. R., Vanapalli, S., Guttula, B., & Katta, J. (2024). Enhancing Fuel Efficiency and Emission Control in Diesel Locomotives through Auxiliary Power Units (APUs) in Neutral Conditions. *Journal of Climate Change*, *10*(2), 69–74. https://doi.org/10.3233/JCC240016

Mandal, A., Biswas, J., Farooqui, Z., & Roychowdhury, S. (2021). A detailed perspective of marine emissions and their environmental impact in a representative Indian port. *Atmospheric Pollution Research*, *12*(10), 101194. https://doi.org/10.1016/j.apr.2021.101194

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (2nd edn). CRC Press.

Mead, M. I., Popoola, O. A. M., Stewart, G. B., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J. J., McLeod, M. W., Hodgson, T. F., Dicks, J., Lewis, A., Cohen, J., Baron, R., Saffell, J. R., & Jones, R. L. (2013). The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmospheric Environment*, *70*, 186–203. https://doi.org/10.1016/j.atmosenv.2012.11.060

Merico, E., Dinoi, A., & Contini, D. (2019). Development of an integrated modelling-measurement system for near-real-time estimates of harbour activity impact to atmospheric pollution in coastal cities. *Transportation Research Part D: Transport and Environment*, *73*, 108–119. https://doi.org/10.1016/j.trd.2019.06.009

Merico, E., Gambaro, A., Argiriou, A., Alebic-Juretic, A., Barbaro, E., Cesari, D., Chasapidis, L., Dimopoulos, S., Dinoi, A., Donateo, A., Giannaros, C., Gregoris, E., Karagiannidis, A., Konstandopoulos, A. G., Ivošević, T., Liora, N., Melas, D., Mifka, B., Orlić, I., … Contini, D. (2017). Atmospheric impact of ship traffic in four Adriatic-Ionian port-cities: Comparison and harmonization of different approaches. *Transportation Research Part D: Transport and Environment*, *50*, 431–445. https://doi.org/10.1016/j.trd.2016.11.016

Met Éireann. (2023). *Hourly Meteorological Observations: Dublin Airport Synoptic Station (03969)*. Met Éireann, The Irish Meteorological Service. https://www.met.ie/climate/available-data

Mokalled, T., Calvé, S. L., Badaro-Saliba, N., Abboud, M., Zaarour, R., Farah, W., Adjizian-Gérard, J., Mokalled, T., Calvé, S. L., Badaro-Saliba, N., Abboud, M., Zaarour, R., Farah, W., & Adjizian-Gérard, J. (2022). Atmospheric dispersion modelling of gaseous emissions from Beirutinternational airport activities. *AIMS Environmental Science*, *9*(5), Article Environ-09-05-033. https://doi.org/10.3934/environsci.2022033

Muggeo, V. M. R. (2003). Estimating regression models with unknown break-points. *Statistics in Medicine*, *22*(19), 3055–3071. https://doi.org/10.1002/sim.1545

Munir, S., Mayfield, M., Coca, D., Mihaylova, L. S., & Osammor, O. (2020). Analysis of Air Pollution in Urban Areas with Airviro Dispersion Model—A Case Study in the City of Sheffield, United Kingdom. *Atmosphere*, *11*(3), Article 3. https://doi.org/10.3390/atmos11030285

Murena, F., Mocerino, L., Quaranta, F., & Toscano, D. (2018). Impact on air quality of cruise ship emissions in Naples, Italy. *Atmospheric Environment*, *187*, 70–83. https://doi.org/10.1016/j.atmosenv.2018.05.056

National Transport Authority. (2024). *Regional Modelling System*. https://www.nationaltransport.ie/planning-and-investment/transport-modelling/regional-modelling-system/

National Transport Authority. (2025). *National Rail Census Report 2024*. National Transport Authority. https://www.nationaltransport.ie/wp-content/uploads/2025/04/National-Rail-Census-Report-2024.pdf

NIST/SEMATECH. (2013). *E-Handbook of Statistical Methods*. NIST/SEMATECH e-Handbook of Statistical Methods. https://www.itl.nist.gov/div898/handbook/

Office of Public Affairs | Massachusetts Bay Transportation Authority to Spend Millions to Reduce Commuter Train Emissions in Clean Air Act Settlement | United States Department of Justice. (2010, August 4). https://www.justice.gov/archives/opa/pr/massachusetts-bay-transportation-authority-spend-millions-reduce-commuter-train-emissions

Orellano, P., Reynoso, J., & Quaranta, N. (2021). Short-term exposure to sulphur dioxide (SO2) and all-cause and respiratory mortality: A systematic review and meta-analysis. *Environment International*, *150*, 106434. https://doi.org/10.1016/j.envint.2021.106434

Papaconstantinou, R., Demosthenous, M., Bezantakos, S., Hadjigeorgiou, N., Costi, M., Stylianou, M., Symeou, E., Savvides, C., & Biskos, G. (2023). Field evaluation of low-cost electrochemical air quality gas sensors under extreme temperature and relative humidity conditions. *Atmospheric Measurement Techniques*, *16*(12), 3313–3329. https://doi.org/10.5194/amt-16-3313-2023

Park, G. Y. (2022). Emissions analysis of the Port Drayage Truck Replacement Program and local air quality: The case of the Port of New York and New Jersey. *Case Studies on Transport Policy*, *10*(2), 1407–1416. https://doi.org/10.1016/j.cstp.2022.05.004

Park, S., Kim, M., Kim, M., Namgung, H.-G., Kim, K.-T., Cho, K. H., & Kwon, S.-B. (2018). Predicting PM10 concentration in Seoul metropolitan subway stations using artificial neural network (ANN). *Journal of Hazardous Materials*, *341*, 75–82. https://doi.org/10.1016/j.jhazmat.2017.07.050

Pecorari, E., Mantovani, A., Franceschini, C., Bassano, D., Palmeri, L., & Rampazzo, G. (2016). Analysis of the effects of meteorology on aircraft exhaust dispersion and deposition using a Lagrangian particle model. *Science of The Total Environment*, *541*, 839–856. https://doi.org/10.1016/j.scitotenv.2015.08.147

Perdikopoulos, M., Karageorgiou, T., Ntziachristos, L., Deville Cavellin, L., Joly, F., Vigneron, J., Arfire, A., Debert, C., Sanchez, O., Gaie-Levrel, F., & Marfaing, H. (2025). Developing Emission Factors from Real-World Emissions of Euro VI Urban Diesel, Diesel-Hybrid, and Compressed Natural Gas Buses. *Atmosphere*, *16*(3), 293. https://doi.org/10.3390/atmos16030293

Pirjola, L., Pajunoja, A., Walden, J., Jalkanen, J. P., Rönkkö, T., Kousa, A., & Koskentalo, T. (2014). Mobile measurements of ship emissions in two harbour areas in Finland. *Atmospheric Measurement Techniques*, *7*(1), 149–161. https://doi.org/10.5194/amt-7-149-2014

Poplawski, K., Setton, E., McEwen, B., Hrebenyk, D., Graham, M., & Keller, P. (2011). Impact of cruise ship emissions in Victoria, BC, Canada. *Atmospheric Environment*, *45*(4), 824–833. https://doi.org/10.1016/j.atmosenv.2010.11.029

Popoola, O. A. M., Carruthers, D., Lad, C., Bright, V. B., Mead, M. I., Stettler, M. E. J., Saffell, J. R., & Jones, R. L. (2018). Use of networks of low cost air quality sensors to quantify air quality in urban settings. *Atmospheric Environment*, *194*(September), 58–70. https://doi.org/10.1016/j.atmosenv.2018.09.030

Porwisiak, P., Werner, M., Kryza, M., ApSimon, H., Woodward, H., Mehlig, D., Gawuc, L., Szymankiewicz, K., & Sawiński, T. (2024). Application of ADMS-Urban for an area with a high contribution of residential heating emissions—Model verification and sensitivity study for PM2.5. *Science of The Total Environment*, *907*, 168011. https://doi.org/10.1016/j.scitotenv.2023.168011

Priyan, S., Guo, Y., Broderick, B. M., Caulfield, B., McNabola, A., O'Mahony, M., Thornes, J., & Gallagher, J. (2025). Identifying air pollution characteristics, source apportionment methods, and air quality modeling approaches in transport hub settings: State-of-play and future directions. *Atmospheric Pollution Research*, 102709. https://doi.org/10.1016/j.apr.2025.102709

Qian, B., Yu, T., Bi, H., & Lei, B. (2020). Measurements of Energy Consumption and Environment Quality of High-Speed Railway Stations in China. *Energies*, *13*(1), Article 1. https://doi.org/10.3390/en13010168

Qin, P., Ricci, A., & Blocken, B. (2024). Modeling traffic pollutants in a street canyon by CFD: Idealized line sources versus multiple realistic car sources. *Science of the Total Environment*, *955*, 177099. https://doi.org/10.1016/j.scitotenv.2024.177099

Qiu, Z., Li, X., Hao, Y., Deng, S., & Gao, H. O. (2016). Emission inventory estimation of an intercity bus terminal. *Environmental Monitoring and Assessment*, *188*(6), 367. https://doi.org/10.1007/s10661-016-5370-8

Rahman, N., Rossiter, D., Hyde, B., Duffy, P., Monaghan, S., Streethran, N., Martini, F., & Brandonisio, A. (2025). *Ireland Informative Inventory Report 2025: Air Pollutant Emissions in Ireland 1990–2023*. Environmental Protection Agency. https://www.epa.ie/publications/monitoring--assessment/climate-change/air-emissions/IIR_Ireland-2025_web_v2.pdf

Rahman, S. M. A., Masjuki, H. H., Kalam, M. A., Abedin, M. J., Sanjid, A., & Sajjad, H. (2013). Impact of idling on fuel consumption and exhaust emissions and available idle-reduction technologies for diesel vehicles – A review. *Energy Conversion and Management*, *74*, 171–182. https://doi.org/10.1016/j.enconman.2013.05.019

Rail Delivery Group. (n.d.). *Reducing Diesel Emissions in Stations and Depots*. Retrieved https://www.raildeliverygroup.com/files/Publications/RDG-OPS-GN-004ReducingDieselEmissionsinStationsandDepots.pdf

Rail Delivery Group. (2021). *Reducing diesel emissions in stations and depots (RDG-OPS-GN-004)*. Rail Delivery Group. https://www.raildeliverygroup.com/files/Publications/RDG-OPS-GN-004ReducingDieselEmissionsinStationsandDepots.pdf

Rail, I. (n.d.). *Train Timetables by Station*. Irish Rail. Retrieved 17 June 2024, from https://www.irishrail.ie

Rail, I. (2021). *Iarnród Éireann Strategy*. Irish Rail. https://www.irishrail.ie

Rail, Irish. (2023b). *CLIMATE ACTION PLAN 2023 2030*. https://www.irishrail.ie/IrishRail/media/Content/Sustainability/climate_plan_FINAL.pdf

Rail Safety and Standards Board. (2020). *CLEAR: Fleet Wide Assessment of Rail Emissions Factors (T1187)*. Rail Safety and Standards Board. https://www.rssb.co.uk/research-catalogue/CatalogueItem/T1187

Railways, E. U. A. for. (2024, July 2). *2024 Rail Environmental Report | European Union Agency for*

Railways. https://www.era.europa.eu/content/2024-rail-environmental-report

Rakopoulos, C. D., & Giakoumis, E. G. (2006). Review of Thermodynamic Diesel Engine Simulations under Transient Operating Conditions. *SAE Technical Paper*, *2006*(01–0884). https://doi.org/10.4271/2006-01-0884

Rakopoulos, C. D., & Giakoumis, E. G. (Eds). (2009). Thermodynamic Aspects of Transient Operation. In *Diesel Engine Transient Operation: Principles of Operation and Simulation Analysis* (pp. 23–74). Springer. https://doi.org/10.1007/978-1-84882-375-4_2

Ramacher, M. O. P., Karl, M., Bieser, J., Jalkanen, J.-P., & Johansson, L. (2019). Urban population exposure to NO_x emissions from local shipping in three Baltic Sea harbour cities – a generic approach. *Atmospheric Chemistry and Physics*, *19*(14), 9153–9179. https://doi.org/10.5194/acp-19-9153-2019

Ramacher, M. O. P., Matthias, V., Aulinger, A., Quante, M., Bieser, J., & Karl, M. (2020). Contributions of traffic and shipping emissions to city-scale NOx and PM2.5 exposure in Hamburg. *Atmospheric Environment*, *237*, 117674. https://doi.org/10.1016/j.atmosenv.2020.117674

Ramacher, M. O. P., Tang, L., Moldanová, J., Matthias, V., Karl, M., Fridell, E., & Johansson, L. (2020). The impact of ship emissions on air quality and human health in the Gothenburg area – Part II: Scenarios for 2040. *Atmospheric Chemistry and Physics*, *20*(17), 10667–10686. https://doi.org/10.5194/acp-20-10667-2020

Rastogi, N., & Frey, H. C. (2021). Characterizing Fuel Use and Emission Hotspots for a Diesel-Operated Passenger Rail Service. *Environmental Science & Technology*, *55*(15), 10633–10644. https://doi.org/10.1021/acs.est.1c00273

Regulation (EU) 2016/1628 of the European Parliament and of the Council of 14 September 2016 on Requirements Relating to Gaseous and Particulate Pollutant Emission Limits and Type-Approval for Internal Combustion Engines for Non-Road Mobile Machinery, Amending Regulations (EU) No 1024/2012 and (EU) No 167/2013, and Amending and Repealing Directive 97/68/EC, Official Journal of the European Union 117 (2016). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R1628

Righi, S., Lucialli, P., & Pollini, E. (2009). Statistical and diagnostic evaluation of the ADMS-Urban model compared with an urban air quality monitoring network. *Atmospheric Environment*, *43*(25), 3850–3857. https://doi.org/10.1016/j.atmosenv.2009.05.016

Rissman, J., Arunachalam, S., Woody, M., West, J. J., BenDor, T., & Binkowski, F. S. (2013). A plume-in-grid approach to characterize air quality impacts of aircraft emissions at the Hartsfield-Jackson Atlanta International Airport. *Atmospheric Chemistry & Physics*, *13*, 9285–9302. https://doi.org/10.5194/acp-13-9285-2013

Rodríguez-Camarena, C., & Gonzalez-Olivardia, F. (2023). *Computational Fluid Dynamics Models to Estimate Pedestrian Exposure to Traffic-Related Air Pollution: A Review*. *27*, 9. https://doi.org/10.3390/ecas2023-15662

Ropkins, K., Quinn, R., Beebe, J., Li, H., Daham, B., Tate, J., Bell, M., & Andrews, G. (2007). Real-world comparison of probe vehicle emissions and fuel consumption using diesel and 5% biodiesel (B5) blend. *The Science of the Total Environment*, *376*(1–3), 267–284. https://doi.org/10.1016/j.scitotenv.2006.11.021

Sabatino, S. D., Solazzo, E., & Britter, R. (2011). The sustainable development of Heathrow Airport: Model inter-comparison study. *International Journal of Environment and Pollution*, *44*(1–4), 351–358. https://doi.org/10.1504/IJEP.2011.038436

Shancita, I., Masjuki, H. H., Kalam, M. A., Rizwanul Fattah, I. M., Rashed, M. M., & Rashedul, H. K. (2014a). A review on idling reduction strategies to improve fuel economy and reduce exhaust emissions of transport vehicles. *Energy Conversion and Management*, *88*, 794–807. https://doi.org/10.1016/j.enconman.2014.09.036

Shancita, I., Masjuki, H. H., Kalam, M. A., Rizwanul Fattah, I. M., Rashed, M. M., & Rashedul, H. K. (2014b). A review on idling reduction strategies to improve fuel economy and reduce exhaust emissions of transport vehicles. *Energy Conversion and Management*, *88*, 794–807. https://doi.org/10.1016/j.enconman.2014.09.036

Shen, Y., de Hoogh, K., Petkova, E., & Liu, Y. (2022). Europe-wide air pollution modeling from 2000 to 2019 using geographically weighted regression. *Environment International*, *168*, 107485. https://doi.org/10.1016/j.envint.2022.107485

Shen, Y., Xu, J., Wu, X., & Ni, Y. (2019). Modelling Travel Time Distribution and Its Influence over Stochastic Vehicle Scheduling. *Transport*, *34*(2), 237–249. https://doi.org/10.3846/transport.2019.8940

Shi, T., Bell, M. L., & Allen, R. W. (2020). Neighbourhood scale nitrogen dioxide land use regression modelling: Implications for exposure assessment along an urban transportation corridor. *Atmospheric Environment*, *224*, 117312. https://doi.org/10.1016/j.atmosenv.2020.117312

Sigurjonsdottir, S. S., Elnes, A. K., & Carvalho Couto, K. (2022). Turn off your engine: Reducing idling amongst professional truck drivers. *Transportation Research Interdisciplinary Perspectives*, *15*, 100654–100654. https://doi.org/10.1016/j.trip.2022.100654

Silverman, D. T., Samanic, C. M., Lubin, J. H., Blair, A. E., Stewart, P. A., Vermeulen, R., Coble, J. B., Rothman, N., Schleiff, P. L., Travis, W. D., Ziegler, R. G., Wacholder, S., & Attfield, M. D. (2012). The Diesel Exhaust in Miners Study: A Nested Case–Control Study of Lung Cancer and Diesel Exhaust. *JNCI Journal of the National Cancer Institute*, *104*(11), 855–868. https://doi.org/10.1093/jnci/djs034

Sim, S., Park, J.-H., & Bae, H. (2022). Deep collaborative learning model for port-air pollutants prediction using automatic identification system. *Transportation Research Part D: Transport and Environment*, *111*, 103431. https://doi.org/10.1016/j.trd.2022.103431

Skoglund, A. (2011). *Evaluation of Test Cycles for Freight Locomotives*. Chalmers University of Technology.

Song, S. K., & Shon, Z. H. (2012). Emissions of greenhouse gases and air pollutants from commercial aircraft at international airports in Korea. *Atmospheric Environment*, *61*, 148–158. https://doi.org/10.1016/j.atmosenv.2012.07.035

Song, S.-K., Shon, Z.-H., & Kang, Y.-H. (2015). Comparison of impacts of aircraft emissions within the boundary layer on the regional ozone in South Korea. *Atmospheric Environment*, *117*, 169–179. https://doi.org/10.1016/j.atmosenv.2015.07.010

Sonitus Systems Ltd. (2022). *DM30 Dustsens Particulate Monitor: Product Specifications*. Sonitus Systems Ltd. https://www.sonitussystems.com/docs/DM30-Dustsens.pdf

Sorte, S., Arunachalam, S., Naess, B., Seppanen, C., Rodrigues, V., Valencia, A., Borrego, C., & Monteiro, A. (2019). Assessment of source contribution to air quality in an urban area close to a harbor: Case-study in Porto, Portugal. *Science of The Total Environment*, *662*, 347–360. https://doi.org/10.1016/j.scitotenv.2019.01.185

Souten, D. R. (2008). *San Diego TAC Emissions Inventory*.

Tafidis, P., Gholamnia, M., Sajadi, P., Krishnan Vijayakrishnan, S., & Pilla, F. (2024). Evaluating the impact of urban traffic patterns on air pollution emissions in Dublin: A regression model using google project air view data and traffic data. *European Transport Research Review*, *16*(1), 47. https://doi.org/10.1186/s12544-024-00671-z

Taketomi, N., Yamamoto, K., Chesneau, C., & Emura, T. (2022). Parametric Distributions for Survival and Reliability Analyses, a Review and Historical Sketch. *Mathematics*, *10*(20), 3907. https://doi.org/10.3390/math10203907

Tan, S. T., Mohamed, N., Ng, L. C., & Aik, J. (2022). Air quality in underground metro station commuter platforms in Singapore: A cross-sectional analysis of factors influencing commuter exposure levels. *Atmospheric Environment*, *273*, 118962. https://doi.org/10.1016/j.atmosenv.2022.118962

Tang, L., Ramacher, M. O. P., Moldanová, J., Matthias, V., Karl, M., Johansson, L., Jalkanen, J.-P., Yaramenka, K., Aulinger, A., & Gustafsson, M. (2020). The impact of ship emissions on air quality and human health in the Gothenburg area – Part 1: 2012 emissions. *Atmospheric Chemistry and Physics*, *20*(12), 7509–7530. https://doi.org/10.5194/acp-20-7509-2020

Targa, J., Loader, A., & Defra Working Group on Harmonisation of Diffusion Tubes. (2008). *Diffusion Tubes for Ambient NO2 Monitoring: Practical Guidance for Laboratories and Users*. AEA Energy & Environment for the Department for Environment, Food & Rural Affairs. https://uk-air.defra.gov.uk/library/reports?report_id=499

Tebaldi, C., & West, M. (1998). Bayesian Inference on Network Traffic Using Link Count Data. *Journal of the American Statistical Association*, *93*(442), 557–573. https://doi.org/10.1080/01621459.1998.10473707

Thornes, J. E., Hickman, A., Baker, C., Cai, X., & Saborit, J. M. D. (2017). Air quality in enclosed railway stations. *Proceedings of the Institution of Civil Engineers-Transport*, *170*(2), 99–107. https://doi.org/10.1680/jtran.15.00094

Thunis, P., Miranda, A., Baldasano, J. M., Blond, N., Douros, J., Graff, A., Janssen, S., Juda-Rezler, K., Karvosenoja, N., Maffeis, G., Martilli, A., Rasoloharimahefa, M., Real, E., Viaene, P., Volta, M., & White, L. (2016). Overview of current regional and local scale air quality modelling practices: Assessment and planning tools in the EU. *Environmental Science & Policy, Multidisciplinary Research Findings in Support to the EU Air Quality Policy: Experiences from the APPRAISAL, SEFIRA and ACCENT-Plus EU FP7 Projects*, *65*, 13–21. https://doi.org/10.1016/j.envsci.2016.03.013

Torres-Vazquez, A., Pleim, J., Gilliam, R., & Pouliot, G. (2022). Performance Evaluation of the Meteorology and Air Quality Conditions from Multiscale WRF-CMAQ Simulations for the Long Island Sound Tropospheric Ozone Study (LISTOS). *Journal of Geophysical Research: Atmospheres*, *127*, e2021JD035890. https://doi.org/10.1029/2021JD035890

Transportation Environmental Resource Center. (n.d.). Retrieved 6 December 2025, from https://www.tercenter.org/engine_emissions_loco_idling.php?utm_source

Tree, D. R., & Svensson, K. I. (2007). Soot processes in compression ignition engines. *Progress in Energy and Combustion Science*, *33*(3), 272–309. https://doi.org/10.1016/j.pecs.2006.03.002

Tu, J., Wayne, W. S., & Perhinschi, M. G. (2013). Correlation Analysis of Duty Cycle Effects on Exhaust Emissions and Fuel Economy. *Journal of the Transportation Research Forum*, *52*(1). https://doi.org/10.5399/osu/jtrf.52.1.4136

Tu, M., Cha, Y., Wahlström, J., & Olofsson, U. (2019). Towards a two-part train traffic emissions factor model for airborne wear particles. *Transportation Research Part D: Transport and Environment*, *67*, 67–76. https://doi.org/10.1016/j.trd.2018.11.006

Tu, M., & Olofsson, U. (2021a). PM levels on an underground metro platform: A study of the train, passenger flow, urban background, ventilation, and night maintenance effects. *Atmospheric Environment: X*, *12*, 100134. https://doi.org/10.1016/j.aeaoa.2021.100134

Tu, M., & Olofsson, U. (2021b). PM10 in underground stations from different types of trains.

Transportation Research Part D: Transport and Environment, *95*, 102867. https://doi.org/10.1016/j.trd.2021.102867

United States Environmental Protection Agency. (2016). *Clean Diesel Funding Assistance Program FY 2016 (RFP No. EPA-OAR-OTAQ-16-02)*. U.S. Environmental Protection Agency. https://www.epa.gov/sites/default/files/2016-04/documents/rfp-epa-oar-otaq-16-02revised.pdf

U.S. Environmental Protection Agency. (2025, April 28). *Learn About Idling Reduction for Locomotives*. U.S. Environmental Protection Agency. https://www.epa.gov/verified-diesel-tech/learn-about-idling-reduction-locomotives

US EPA, O. (2016a, March 24). *Learn About Idling Reduction for Locomotives* [Other Policies and Guidance]. https://www.epa.gov/verified-diesel-tech/learn-about-idling-reduction-locomotives

US EPA, O. (2016b, March 24). *Learn About Idling Reduction for Locomotives* [Other Policies and Guidance]. https://www.epa.gov/verified-diesel-tech/learn-about-idling-reduction-locomotives

US EPA, O. (2016c, May 10). *Latest Version of MOtor Vehicle Emission Simulator (MOVES)* [Data and Tools]. https://www.epa.gov/moves/latest-version-motor-vehicle-emission-simulator-moves

Vardoulakis, S., Fisher, B. E. A., Pericleous, K., & Gonzalez-Flesca, N. (2003). Modelling air quality in street canyons: A review. *Atmospheric Environment*, *37*(2), 155–182. https://doi.org/10.1016/S1352-2310(02)00857-9

Vehtari, A., Simpson, D., Yao, Y., & Gelman, A. (2024). Prior choice recommendations: Weakly informative default priors for regression models. *Statistics and Computing*, *34*(1), 9–9. https://doi.org/10.1007/s11222-023-10381-4

Vicedo-Cabrera, A. M., Sera, F., Liu, C., Armstrong, B., Milojevic, A., Guo, Y., Tong, S., Lavigne, E., Kyselý, J., Urban, A., & et al. (2020). Short term association between ozone and mortality: Global two stage time series study in 406 locations in 20 countries. *BMJ*, *368*, m108. https://doi.org/10.1136/bmj.m108

Vichi, F., Frattoni, M., Imperiali, A., Balducci, C., Cecinato, A., Perilli, M., & Romagnoli, P. (2016). Civil aviation impacts on local air quality: A survey inside two international airports in central Italy.

Atmospheric Environment, *142*, 393–405. https://doi.org/10.1016/j.atmosenv.2016.08.005

Villeneuve, P. J., Parent, M.-É., Sahni, V., & Johnson, K. C. (2011). Occupational exposure to diesel and gasoline emissions and lung cancer in Canadian men. *Environmental Research*, *111*(5), 727–735. https://doi.org/10.1016/j.envres.2011.04.003

Voogt, M., Zandveld, P., Erbrink, H., van Dinther, D., van den Bulk, P., Kos, G., Blom, M., de Jonge, D., Helmink, H., Meydam, J., Visser, J., Middel, J., Hoek, G., van Ratingen, S., Wesseling, J., & Janssen, N. AH. (2023). Assessment of the applicability of a model for aviation-related ultrafine particle concentrations for use in epidemiological studies. *Atmospheric Environment*, *309*, 119884. https://doi.org/10.1016/j.atmosenv.2023.119884

Vutukuru, S., & Dabdub, D. (2008). Modeling the effects of ship emissions on coastal air quality: A case study of southern California. *Atmospheric Environment*, *42*(16), 3751–3764. https://doi.org/10.1016/j.atmosenv.2007.12.073

Walther, E., & Bogdan, M. (2017). A novel approach for the modelling of air quality dynamics in underground railway stations. *Transportation Research Part D: Transport and Environment*, *56*, 33–42. https://doi.org/10.1016/j.trd.2017.07.014

Wang, X., Shen, Y., Lin, Y., Pan, J., Zhang, Y., Louie, K. K. P., Li, M., & Fu, Q. (2019). Atmospheric pollution from ships and its impact on local air quality at a port site in Shanghai. *Atmospheric Chemistry and Physics*, *19*(9), 6315–6330. https://doi.org/10.5194/acp-19-6315-2019

Weaver, C. S. (2006). *Start-up and Idling Emissions from Two Locomotives*. https://www.aqmd.gov/docs/default-source/ceqa/documents/aqmd-projects/2006/rule-3501/appe.pdf?sfvrsn=4

Whitehill, A. R., Preble, C. V., Anderson, R. F., Solomon, P. A., Thoma, E. D., Beidler, L. K., & Apte, J. S. (2020). Uncertainty in collocated mobile measurements of air quality. *Environmental Advances*, *2*, 100017–100017. https://doi.org/10.1016/j.envadv.2020.100017

WHO. (2022). https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

World Health Organization. (2021a). *WHO global air quality guidelines: Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide* (No. 9789240034228). World Health Organization. https://www.who.int/publications/i/item/9789240034228

World Health Organization. (2021b). *WHO global air quality guidelines: Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide* (No. 9789240034228). World Health Organization. https://www.who.int/publications/i/item/9789240034228

World Health Organization. (2024a). *Ambient (outdoor) air quality and health*. https://www.who.int

World Health Organization. (2024b, October 24). *Ambient (outdoor) air pollution*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/ambient-%28outdoor%29-air-quality-and-health

Xu, H., Xiao, K., Pan, J., Fu, Q., Wei, X., Zhou, J., Yu, Y., Hu, X., Ren, H., Cheng, J., Peng, S., Hong, N., Ye, Y., Su, N., He, Z., & Hu, T. (2023). Evidence of aircraft activity impact on local air quality: A study in the context of uncommon airport operation. *Journal of Environmental Sciences (China)*, *125*, 603–615. https://doi.org/10.1016/j.jes.2022.02.039

Yamada, H., Misawa, K., Suzuki, D., Tanaka, K., Matsumoto, J., Fujii, M., & Tanaka, K. (2011). Detailed analysis of diesel vehicle exhaust emissions: Nitrogen oxides, hydrocarbons and particulate size distributions. *Proceedings of the Combustion Institute*, *33*(2), 2895–2902. https://doi.org/10.1016/j.proci.2010.07.001

Yang, X., Cheng, S., Lang, J., Xu, R., & Lv, Z. (2018). Characterization of aircraft emissions and air quality impacts of an international airport. *Journal of Environmental Sciences*, *72*, 198–207. https://doi.org/10.1016/j.jes.2018.01.007

Yarragunta, Y., Francis, D., Fonseca, R., & Nelli, N. (2025). Evaluation of the WRF-Chem performance for the air pollutants over the United Arab Emirates. *Atmospheric Chemistry and Physics*, *25*(3), 1685–1709. https://doi.org/10.5194/acp-25-1685-2025

Yoo, Y., Kim, J., Ga, S., Lim, J., Kim, J., & Cho, H. (2022a). Computational fluid dynamics-based optimal installation strategy of air purification system to minimize NOX exposure inside a public bus stop. *Environment International*, *169*, 107507. https://doi.org/10.1016/j.envint.2022.107507

Yoo, Y., Kim, J., Ga, S., Lim, J., Kim, J., & Cho, H. (2022b). Computational fluid dynamics-based optimal installation strategy of air purification system to minimize NOX exposure inside a public bus stop. *Environment International*, *169*, 107507. https://doi.org/10.1016/j.envint.2022.107507

Yoo, Y., Kim, J., Lee, J., & Cho, H. (2024). Air quality improvement at urban bus stops: Optimal air purification placement using CFD. *Sustainable Cities and Society*, *117*, 105937. https://doi.org/10.1016/j.scs.2024.105937

Yu, Q., Lu, L., Li, T., & Tu, R. (2022a). Quantifying the Impact of Alternative Bus Stop Platforms on Vehicle Emissions and Individual Pollution Exposure at Bus Stops. *International Journal of Environmental Research and Public Health*, *19*(11), Article 11. https://doi.org/10.3390/ijerph19116552

Yu, Q., Lu, L., Li, T., & Tu, R. (2022b). Quantifying the Impact of Alternative Bus Stop Platforms on Vehicle Emissions and Individual Pollution Exposure at Bus Stops. *International Journal of Environmental Research and Public Health*, *19*(11), 6552. https://doi.org/10.3390/ijerph19116552

Yuan, W., & Frey, H. C. (2021). Multi-scale evaluation of diesel commuter rail fuel use, emissions, and eco-driving. *Transportation Research Part D: Transport and Environment*, *99*, 102995. https://doi.org/10.1016/j.trd.2021.102995

Zhai, Z., Tu, R., Xu, J., Wang, A., & Hatzopoulou, M. (2020). Capturing the Variability in Instantaneous Vehicle Emissions Based on Field Test Data. *Atmosphere*, *11*(7), 765. https://doi.org/10.3390/atmos11070765

Zhang, J., Zhang, S., Wang, Y., Bao, S., Yang, D., Xu, H., Wu, R., Wang, R., Yan, M., Wu, Y., & Hao, J. (2021). Air quality improvement via modal shift: Assessment of rail-water-port integrated system planning in Shenzhen, China. *Science of The Total Environment*, *791*, 148158. https://doi.org/10.1016/j.scitotenv.2021.148158

Zheng, X., Orellano, P., Lin, H., Jiang, M., & Guan, W. (2021). Short-term exposure to ozone, nitrogen dioxide, and sulphur dioxide and emergency department visits and hospital admissions due to asthma: A systematic review and meta-analysis. *Environment International*, *150*, 106435. https://doi.org/10.1016/j.envint.2021.106435

Zhong, J., Harrison, R. M., Bloss, W. J., Visschedijk, A., & Denier van der Gon, H. A. C. (2023). Modelling the dispersion of particle number concentrations in the West Midlands, UK using the ADMS-Urban model. *Environment International*, *181*, 108273. https://doi.org/10.1016/j.envint.2023.108273

Zhou, Y., Zhang, Y., Ma, D., Lu, J., Luo, W., Fu, Y., Li, S., Feng, J., Huang, C., Ge, W., & Zhu, H. (2020). Port-Related Emissions, Environmental Impacts and Their Implication on Green Traffic Policy in Shanghai. *Sustainability*, *12*(10), Article 10. https://doi.org/10.3390/su12104162

Appendix 1: Photos of Irish Rail trains

Photo from Irish Rail website: https://www.irishrail.ie/en-ie/about-us/iarnrod-eireann-fleet.

22000 Class DMU

29000 Class DMU

201 Locomotive + Mark IV

Appendix 2: Train ID code classification table

Alpha code	Train type	Usage
A	Mainline Passenger	Scheduled Mainline Passenger trains as detailed in the Working Timetable.
B	Mainline Passenger Special	Special Mainline Passenger Trains as detailed in the Weekly Circular and Special Running Notices.
C	Empty Diesel Suburban	Empty Commuter Passenger Trains operating over the following Commuter lines: Pearse/M3 Parkway/Longford, Bray/Dundalk, Mallow/Cork/Cobh/Midleton and Heuston/Portlaoise.
D	Diesel Suburban (Outer Destinations)	Commuter Passenger Trains operating over the following Commuter Lines: Heuston/Portlaoise, Mallow/Cork/Cobh/Midleton, Bray/Dundalk, Pearse/Longford

Alpha code	Train type	Usage
		and Docklands/M3 Parkway. (Destinations are in the Outer Suburban area).
E	Electric (EMU)	All Electric Passenger trains. Includes Electric Special trains.
F	Empty Electric (EMU)	All Electric Empty trains.
G	Group ID’s	Reserved for Signalman’s use only.
H	Heritage Rolling Stock	All Heritage movements, diesel or steam loco hauled (Passenger and Empty). Special Instructions govern the operation of these trains.
I	Shunting	See Note 6 on Page 25.
J	Empty Mainline Passenger	Empty Mainline Passenger trains.
K	Liner Trains (Conveys Dangerous Goods)	Laden Liner Trains which convey Dangerous Goods.
L	Oil	Empty & Laden Oil Trains.
M	Mineral	Empty & Laden Mineral Trains.
N	Cement	Empty & Laden Cement Trains.
O	Stationary Vehicles	All stationary vehicles for which there is no suitable ID known or available. CTC Signalperson’s use only.
P	Diesel Suburban (Inner Destinations)	Commuter Passenger Trains operating over the following Commuter Lines: Portlaoise/Heuston, Midleton/Cobh/Cork/Mallow, Dundalk/Bray, Longford/Pearse and M3 Parkway/Docklands. (Destinations are in the Inner Suburban area).
Q	Sleeper Train	See Note 10 on Page 25.
R	Light Engine(s)	See Note 3 on Page 25.
S	Liner & Timber Trains	Empty & Laden Liner & Timber trains.
T	Special Freight Movement	Empty & Laden Special Freight Movements.

Alpha code	Train type	Usage
U	Radio Test on Locomotives and 4000 / 9000 Class DVTs	Testing of Radio Equipment.
V	Beet	Empty & Laden Beet Trains.
W	Departmental Trains	Empty & Laden Chief Civil Engineer & Chief Mechanical Engineer fully fitted (vacuum or air) Departmental Trains excluding codes Y & Z. Includes trial trains. (See Note 3 on Page 25).
X	Radio Test EMU	Testing of EMU Radio Equipment.
Y	Per Way Machines & Cars (NOT CAPABLE OF OPERATING TRACK CIRCUITS)	Permanent Way Machines & Cars NOT capable of operating track circuits.
Z	Unusual Movements	Unusual Movements

Appendix 3: PM2.5 heatmaps at Connolly and Heuston

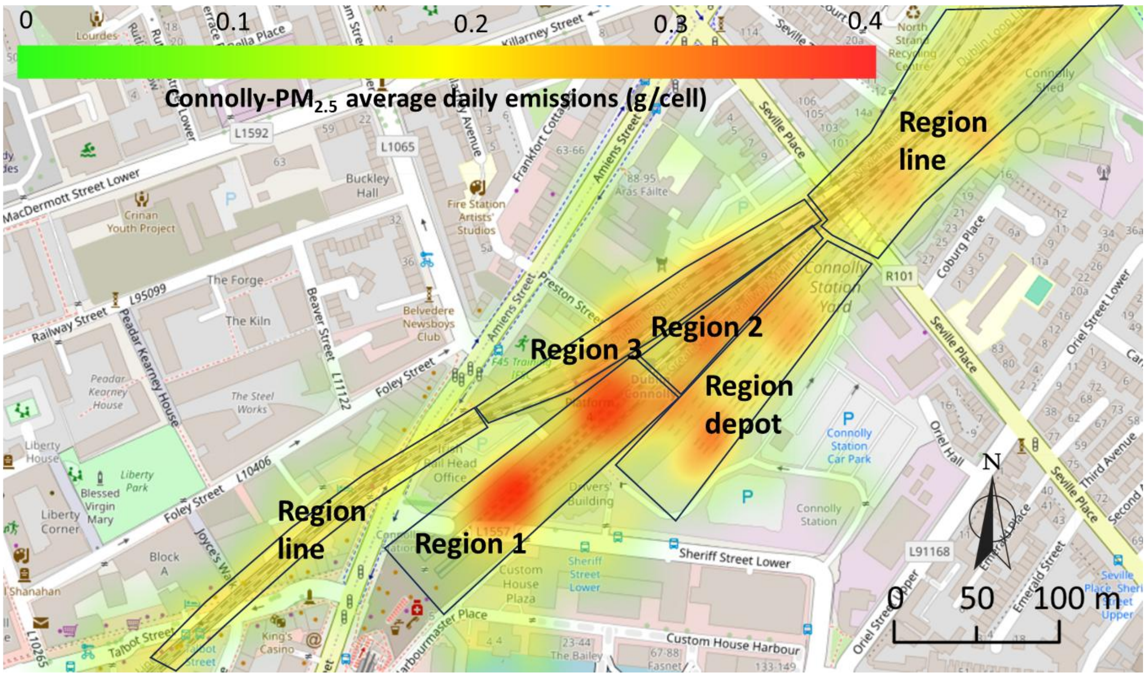


Figure A3.1 PM2.5 heatmap at Connolly

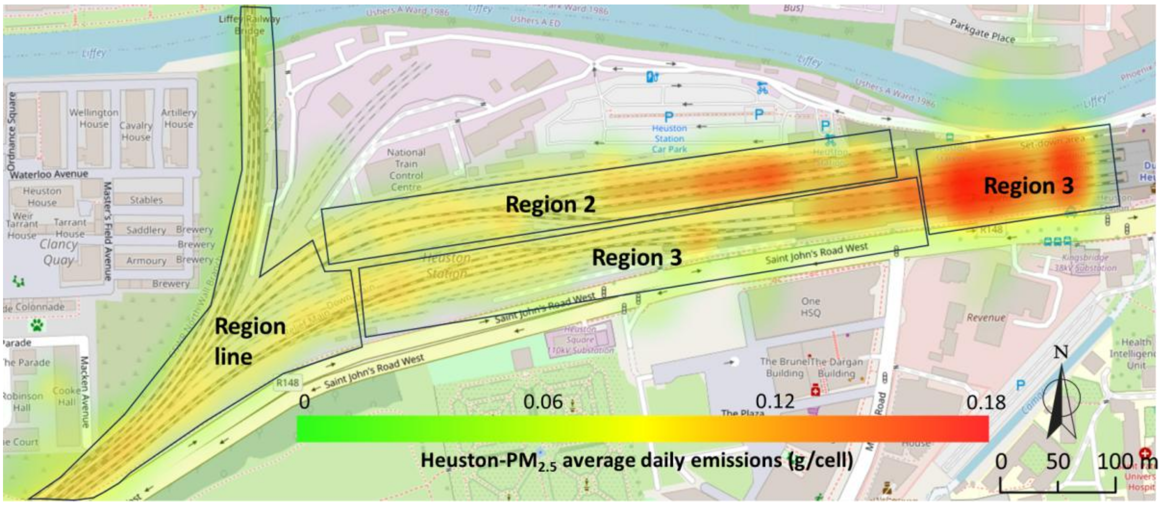


Figure A3.2 PM2.5 heatmap at Heuston