
THE UNIVERSITY
*of* EDINBURGH

# Assessing urban air quality through measurements and modelling and its implications for human exposure assessment

Hao Wu


THE UNIVERSITY
*of* EDINBURGH

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
The University of Edinburgh

2017

# Lay Summary

Outdoor air pollution is estimated to cause 3.7 million premature deaths worldwide in 2012 and also contributes to a range of adverse health effects. To study the effects of air pollution on human health, researchers examine the relationship between adverse health outcomes and people's exposure to ambient air pollution. The concentrations of air pollutants to which people are exposed can be assessed using air quality measurements and modelling. Both approaches are subject to uncertainties in estimating an individual's true exposure. Accurate exposure assessment is, however, vital to the quantification of overall health effects associated with air pollution. The aim of this thesis is to investigate how different measurement and modelling techniques can improve currently applied exposure assessment methods.

Emerging low-cost portable air quality monitors have the potential to provide detailed exposure assessment at personal level. However, the data produced by these monitors are sometimes questionable and not directly comparable with those produced by more sophisticated instruments that are accepted as the reference method. In this thesis three types of low-cost air quality monitors that measure three different air pollutants – the gases nitrogen dioxide ($NO_2$) and ozone ($O_3$), and the 'respirable' fraction of particles suspended in the air (called $PM_{2.5}$) – are evaluated against their respective reference instruments. Data produced from all three types of monitors investigated here required substantial post-processing in order to compare with the reference measurements.

Another advantage of a portable air quality monitor is its capability to study the local-scale spatial variability of air pollution that is not achievable by the fixed-site monitoring stations that constitute the present approach to determining levels of air pollution. Mobile measurements utilising three ways of characterising the amount of particulate matter (PM) in the air – the mass concentration, the number concentration and the amount of 'soot'-like particles – were conducted in different urban environments in Edinburgh. It was found that local traffic emissions strongly influenced the number and 'soot'-like particles concentrations, but the mass

concentration was dominated by long-range transported sources. The varying contributions from different sources resulted in varying spatial variabilities of the three types PM concentrations in the urban environment.

Two modelling methods that are used to simulate pollutant concentrations at relatively high spatial resolution within urban areas are dispersion models and land-use regression (LUR) models. The former type of model aims to simulate the processes that link the emissions of pollutants from individual sources and their atmospheric dispersion and transport. The uncertainties in these models mainly derive from their need for very detailed emissions and meteorological input data. The effect of using different input data on the output of a dispersion model applied to Edinburgh was investigated and validated against measurements. The dispersion model simulated well the general spatial and temporal trends of air pollution but failed to capture variation caused by very local effects such as queueing traffic and road junctions.

LUR models, on the other hand, are empirical models that rely on existing measurements to derive the statistical relationship between pollutant concentrations at a given location and local predictor variables related to the emission and dispersion of air pollutants (for example, local population density and distance to nearest major road). The quality of LUR models depends on the choice of measurement data and the selection of predictor variables. The effect of the choice of measurement data on a series of LUR models' capabilities at estimating $NO_2$ concentrations at residential addresses in Edinburgh was investigated with the help of a dispersion model. LUR models developed from $NO_2$ concentrations at highly populated areas and roadside locations better estimated the residential $NO_2$ concentrations.

In the final part of the thesis, the findings and limitations from the measurement and modelling studies undertaken in this work are discussed in the context of current air pollution exposure assessment studies. The challenges for the development of a more comprehensive exposure assessment framework are discussed.

# Abstract

Outdoor air pollution is a major contributor to adverse health effects of citizens, in particular those living in urban environments. Air quality monitoring networks are set up to measure air quality in different environments in compliance with national and European legislation. Generally, only a few fixed monitoring sites are located within a city and thus cannot represent air pollutant concentrations in urban areas accurately enough to allow for a detailed human exposure assessment. Other approaches to derive detailed urban air pollutant concentration estimates exist, such as dispersion models and land-use regression (LUR) models. Low-cost portable air quality monitors are also emerging, which have the potential to add value to existing monitoring networks by providing measurements at greater spatial resolution and also to provide individual-level exposure assessment. The aim of this thesis is to demonstrate how measurements and modelling in combination allow detailed investigations of the variability of air pollutants in space and time in urban area, and in turn improve on the current exposure assessment methods.

Three types of low-cost portable monitors measuring $NO_2$, $O_3$ (Aeroqual monitors) and $PM_{2.5}$ (microPEM monitor) were evaluated against their respective reference instruments. The Aeroqual $O_3$ monitor showed very good correlation ($r^2 > 0.9$) with the respective reference instruments, but biases in the slope and intercept coefficients indicated that calibration of Aeroqual $O_3$ monitor was needed. The Aeroqual $NO_2$ monitor was subject to cross-sensitivity from $O_3$, which, as demonstrated, can be effectively corrected by making $O_3$ and $NO_2$ measurements in tandem. Correlation between the microPEM monitor and its reference instrument was poor ($r^2 < 0.1$) when $PM_{2.5}$ concentrations were low (< 10 µg $m^{-3}$), but significantly improved ($r^2 > 0.69$) during periods with elevated $PM_{2.5}$ concentrations. Relative humidity was not found to affect the raw results of $PM_{2.5}$ measurements in a consistent manner. All three types of monitors cannot be used as equivalent or indicative methods instead of reference methods in studies that require quantification of absolute pollutant concentrations. However, the generally good correlations with reference instruments reassure their application in studies of relative trends of air pollution.

Concentrations of $PM_{2.5}$, ultrafine particles (UFP) and black carbon (BC) were quantified using portable monitors through a combination of mobile and static measurements in the city of Edinburgh, UK. The spatial variability of UFP and BC was large, of similar magnitude and about 3 times higher than the spatial variability of $PM_{2.5}$. Elevated concentrations of UFP and BC were observed along streets with high traffic volumes whereas $PM_{2.5}$ showed less variation between streets and a footpath without road traffic. Both BC and UFP significantly correlated with traffic counts, while no significant correlation between $PM_{2.5}$ and traffic counts was observed. The relationships between UFP, $NO_2$ and inorganic components of $PM_{2.5}$ were further investigated through long-term measurements at roadside, urban background and rural sites. UFP moderately correlated with $NO_x$ ($NO_2$ + NO) and showed varying relationships with $NO_x$ depending on the particle size distribution. Principal component analysis and air-mass back trajectory analysis revealed that $PM_{2.5}$ concentrations were dominated by long-range transport of secondary inorganic aerosols, whereas UFP were mainly related to varying local emissions and meteorological conditions. These findings imply the need for different policies for managing human exposure to these different particle components: control of much BC and UFP appears to be manageable at local scale by restricting traffic emissions; however, abatement of $PM_{2.5}$ requires a more strategic approach, in cooperation with other regions and countries on emissions control to curb long-range transport of $PM_{2.5}$ precursors.

A dispersion model (ADMS-Urban) was used to simulate high resolution $NO_2$ and $O_3$ concentrations in Edinburgh. The effects of different emission and meteorological input datasets on the resulting modelled $NO_2$ concentrations were investigated. The modelled $NO_2$ and $O_3$ concentrations using the optimal model setup were validated against reference instrument and diffusion tube measurements. Temporal variability of $NO_2$ was predicted well at locations that were not heavily influenced by local effects, such as road junctions and bus stops. Temporal variability of $O_3$ was predicted better than for $NO_2$. Long-term spatial variability of $NO_2$ was found to correlate well with diffusion tube measurements, while modelled spatial variability of $O_3$ in ADMS-Urban compared poorly with diffusion tube measurements.

However, it was found that the $O_3$ diffusion tube measurements may be subject to some unidentified biases affecting their accuracy.

Land-use regression (LUR) models are widely used to estimate exposure to air pollution in urban areas. An appropriately sized and designed monitoring network is an important component for the development of a robust LUR model. Concentrations of $NO_2$ were simulated by ADMS-Urban at 'virtual' monitoring sites in 54 different network designs of varying numbers and types of site, using a 25 $km^2$ area including much of the Edinburgh city area. Separate LUR models were developed for each network. These LUR models were then used to estimate ambient $NO_2$ concentrations at all residential addresses, which were evaluated against the ADMS-Urban modelled concentration at these addresses. The improvement in predictive capability of the LUR models was insignificant above ~30 monitoring sites, although more sites tended to yield more precise LUR models. Monitoring networks containing sites located within highly populated areas better estimated $NO_2$ concentrations across all residential locations. LUR models constructed from networks containing more roadside sites better characterised the high end of residential $NO_2$ concentrations but had increased errors when considering the whole range of concentrations. No particular composition of monitoring network resulted in good estimation simultaneously across all residential $NO_2$ concentration and of the highest $NO_2$ levels implying a lack of spatial contrast in LUR-modelled pollution surface compared with the dispersion model.

Finally, the results from the measurement and modelling studies presented in thesis are synthesised in the context of current exposure assessment studies. Low-cost air-quality monitors currently do not possess and are unlikely in the near future to provide the robustness and accuracy to replace the existing routine monitoring network. Development of the low-cost air-quality should be aiming at upgrading them as the indicative method as defined in the data quality objective in the EU directive. The monitoring sites used to build LUR models should capture well the population distribution in the study area as opposed to capturing the greatest pollution contrast. The traditional methods of evaluating LUR models are also ineffective in characterising the models' capability at estimating pollutant

concentration at residential address. Given that the dispersion models are also subject to the availability and uncertainties in the input data, future air quality model development should endeavour to incorporate both dispersion and land-use regression models, where the uncertainty in the input data can be reduced by using LUR models built on actual measurements, and the limitation in the statistical modelling can be replaced by adopting the deterministic approach used in the dispersion model.

# Acknowledgements

First of all, I am tremendously grateful to my supervisors, Mathew Heal and Stefan Reis, for giving me the opportunity to study for the PhD at the University of Edinburgh. The scholarships they have applied for me made it possible for me to study overseas. I truly appreciate their support and timely feedbacks, which have made this journey exceptionally smooth. I have learned a lot from them not only in the field of air quality but also about the attributes it takes to become a scientist.

I would also like to thank Chun Lin from our research group who has given me incredible technical supports throughout my work and useful life tips for living in the UK as an overseas student, which both immensely helped me to progress especially at the early confusing stage of my PhD study.

The regular teleconference with Strathclyde colleagues is another source of inspiration for my PhD. Work reports circulated by Jonathan Gillespie and Nicola Masey at the University of Strathclyde set great examples for me on how to conduct experiments and write a scientific report. The MACAQUE group meeting with colleagues from the School of GeoSciences was a constant motivation for me to produce interesting results to share and discuss with fellow atmospheric scientists. My presentation and public speaking skills really have sharpened thanks to these meetings. I also acknowledge the support from the staff at the Centre for Ecology & Hydrology on allowing me to attend many technical workshops. I would also like to thank Archie Forrest at the City of Edinburgh Council for giving me access to the automatic monitoring sites.

I also thank my partner Xiao Tang for her moral support and company in this once-in-a-life-time overseas study experience.

Last but not least, without the financial and moral support from my parents from the beginning, I could not be where I am now. I owe my greatest debt to them.

# Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Parts of this work have been published in

Wu, H., Reis, S., Lin, C., Beverland, I.J., Heal, M.R., 2015. Identifying drivers for the intra-urban spatial variability of airborne particulate matter components and their interrelationships. Atmos. Environ. 112, 306–316. doi:10.1016/j.atmosenv.2015.04.059

Wu, H., Reis, S., Lin, C., Heal, M.R., 2017. Effect of monitoring network design on land use regression models for estimating residential NO2 concentration. Atmos. Environ. 149, 24–33. doi:10.1016/j.atmosenv.2016.11.014.

Some subjects of this thesis are also related to other publications I have contributed to,

Kenagy, H.S., Lin, C., Wu, H., Heal, M.R., 2015. Greater nitrogen dioxide concentrations at child versus adult breathing heights close to urban main road kerbside. Air Qual. Atmosphere Health, doi: 10.1007/s11869-015-0370-3.

Steinle, S., Reis, S., Sabel, C.E., Semple, S., Twigg, M.M., Braban, C.F., Leeson, S.R., Heal, M.R., Harrison, D., Lin, C., Wu, H., 2015. Personal exposure monitoring of $PM_{2.5}$ in indoor and outdoor microenvironments. Sci. Total Environ. 508, 383–394.

Hao Wu

April 2017

# Contents

# List of Figures

# List of Tables

# Chapter 1 Introduction

Air pollution can cause harm to human health, agricultural production, and the natural or built environment, and also contribute to climate change. Some common air pollutants, which are regulated by national and international legislation, include carbon monoxide (CO), nitrogen dioxide ($NO_2$), ozone ($O_3$), sulphur dioxide ($SO_2$) and particulate matter (PM). The World Health Organisation (WHO) has recommended limits for ambient concentrations of a number of pollutants for the protection of human health (WHO, 2006). Based on these guidelines, and consideration of other evidence, limit values are set in the EU Directives (EC, 2008), to which member states are legally bound. Table 1.1 summarises the limit values for different pollutants set in the Directive 2008/50/EC.

*Table 1.1 Limit values for the ambient concentration of different air pollutants set in the Directive 2008/50/EC. Note that the concentrations for $O_3$ and $PM_{2.5}$ are target values, which means a level to be attained where possible, but that is not legally binding.*

| Pollutant | Measured as | Limit or target concentration | Maximum number of exceedances allowed |
|---|---|---|---|
| **CO** | Maximum daily running 8-hour mean | 10 mg/m$^3$ | |
| **$NO_2$** | Annual mean | 40 µg/m$^3$ | |
| | Hourly mean | 200 µg/m$^3$ | Not to be exceeded more than 18 hours a year |
| **$O_3$** | Maximum daily running 8-hour mean | 120 µg/m$^3$ | Not to be exceeded more than 25 days a year |
| **$SO_2$** | Daily mean | 125 µg/m$^3$ | Not to be exceeded more than 24 days a year |
| | Hourly mean | 350 µg/m$^3$ | Not to be exceeded more than 3 hours a year |
| **$PM_{2.5}$** | Annual mean | 25 µg/m$^3$ | |
| **$PM_{10}$** | Annual mean | 40 µg/m$^3$ | |
| | Daily mean | 50 µg/m$^3$ | Not to be exceeded more than 35 days a year |

Member states transpose the detail of the EU Directives into their national laws. In addition, member states can set more stringent Air Quality Directives (AQD) than that specified in the EU Directives. To comply with the AQD, governments need to demonstrate that the ambient concentrations of the air pollutants are below the limit values by using measurement or modelling. Currently in the UK there are 156 air quality monitoring stations in operation for compliance reporting against the AQD, and these constitute the Automatic Urban and Rural Network (AURN). In addition to these national government-operated monitoring stations, many local authorities may also set up additional monitoring sites within towns and cities for local air quality management. These often form regional monitoring networks [e.g. the Scottish Air Quality Network (SAQN) or the London Air Quality Network (LAQN)]. The monitoring stations provide real-time (hourly) measurements of a range of regulated air pollutants, but the majority of monitoring stations only measure a subset, sometimes only one, of the regulated air pollutants. The instruments are subject to rigorous and consistent quality assurance methodologies. However due to the cost and labour-intensive maintenance of the instruments, generally only a few of these monitoring stations are distributed across a city therefore providing limited information on the spatial variation of air pollution.

Many epidemiological studies, which focus on the relationship between air pollution and human health, utilise the data from these stations to estimate population exposure. Due to the spatial scarcity of the measurement data, researchers either have to make the assumption that the pollutant of interest is homogenously distributed in the study area or have to use some form of modelling to estimate pollutant concentrations at un-sampled locations. This may introduce one of the major uncertainties in epidemiological studies – exposure misclassification (Özkaynak et al., 2013; Sheppard et al., 2012).

To eliminate exposure misclassification, it is important to understand how different air pollutants vary in space and time in the urban environment. In this introduction, the measurement techniques and variability of different air pollutants in the urban environment are discussed first, followed by a brief overview of the existing evidence on the adverse health impacts of different air pollutants. Exposure

assessment methods used in these health studies and their limitations are then reviewed. Current efforts in improving exposure assessment facilitated by the development of low-cost sensors, computation and communication technologies are discussed later. Finally, the aims of this thesis are outlined.

## 1.1 Measurement techniques and spatiotemporal variability of urban air pollution

The Directive 2008/50/EC specifies the reference measurement methods that member states shall apply to measure the regulated air pollutants. Also specified in Directive 2008/50/EC are the data quality objectives that the reference methods are required to meet. The data quality objectives include the uncertainties of measurement in the region of the limit/target values for each pollutant, time coverage of the measurements in relation to the reference period of the limit/target value, and the effective measurement time. Other measurement methods may be used if equivalent results can be demonstrated. The principles and methodologies to be used for demonstration of equivalent method are published by the European Commission (EC, 2008). This section describes the reference or reference-equivalent methods used in the routine monitoring networks (e.g. AURN, SAQN and LAQN) for measurement of $NO_2$, $O_3$ and $PM_{2.5}$, followed by an overview of the major sources and distribution characteristics of $NO_2$, $O_3$, $PM_{2.5}$, ultrafine particles (UFP) and black carbon (BC) in the urban environment.

### 1.1.1 Nitrogen dioxide

During any combustion process nitrogen and oxygen in the air combine at high temperature to produce nitric oxide (NO). NO transforms to $NO_2$ on reaction with atmospheric oxidants such as $O_3$. Together, NO and $NO_2$ are referred to as $NO_x$. The reference measurement method for $NO_2$ is the chemiluminescence analyser. In this method light at visible and near infrared wavelengths resulting from the fluorescence of electronically excited $NO_2$ formed in the $NO + O_3$ reaction is measured (Dunlea et al., 2007). The intensity of the light is proportional to the $NO_2$ concentration formed through this reaction, and therefore also to the initial NO concertation in the air sample if excess $O_3$ is present. Excess $O_3$ is usually generated within a separate air

stream into the chemiluminescence analyser via corona discharge. The $NO_2$ concentration in the air sample is determined by diverting the incoming air through an $NO_2$ to NO converter (usually a heated molybdenum oxide catalyst) every few minutes so that the chemiluminescence intensity is then proportional to $NO_x$. The $NO_2$ concentration is obtained from the difference between $NO_x$ and NO. Calibration of the chemiluminescence analyser is required, which involves sampling of a known NO standard to determine the absolute response of the instrument.

Anthropogenic emission is the major source of NOx in the atmosphere. Sources include the combustion of fossil fuels for heating, power generation (including for industry) and transportation (Vardoulakis et al., 2010). The majority of the NOx emissions are in the form of NO. However, given enough time and enough $O_3$ to react with, most NO is converted to $NO_2$. In the urban environment, the variability of $NO_2$ concentrations is closely related to combustion processes, primarily traffic emissions. Higher $NO_2$ concentrations are expected near roads as a result of a proportion of the NOx emission from traffic exhaust being emitted directly as primary $NO_2$ and also the fast reaction between NO and ambient $O_3$. Concentrations of $NO_2$ usually decrease rapidly away from road sources due to dispersion. Hence considerable spatial variation in $NO_2$ concentration is observed in the urban environment, associated with distance from roads (Cyrys et al., 2012; Lin et al., 2016; Vardoulakis et al., 2009).

### 1.1.2 Ozone

Ozone is a secondary pollutant formed through chemical reactions between precursor gases including volatile organic compounds (VOCs), CO and $NO_x$. The reference measurement method for measuring $O_3$ is the UV absorbance spectrometer. This method relies on the UV absorbing property of $O_3$ (AQEG, 2009). An air sample passing through a tubular cell of length ($l$) causes attenuation of UV light at 254 nm ($I_1$) due to the presence of $O_3$. The intensity of the light ($I_0$) in the absence of $O_3$ is also measured by alternately passing the air sample through an $O_3$ scrubber before it enters the absorption cell. The concentration of $O_3$ ($c$) is quantified by Beer-Lambert law:

$$I_1 = I_0 e^{-\sigma l c}$$

where $\sigma$ = $O_3$ absorption coefficient at 254 nm.

While $O_3$ in the stratosphere protects humans from the adverse effects of UV radiation, it is harmful to human health, agricultural crops and vegetation at ground level. Ambient concentration of $O_3$ is determined by a series of complex chemical reactions between $NO_x$, VOCs and hydrogen oxide radicals ($HO_x$ = H + OH + $HO_2$). Figure 1.1 shows a schematic of tropospheric $O_3$ chemistry (The Royal Society, 2008). The production and destruction of $O_3$ start with photolysis of $O_3$ itself. OH radicals play an important role in $O_3$ chemistry as they react with VOCs to initiate reaction cycles that produce and remove $O_3$. In moderately polluted areas (e.g. rural areas in most industrialised countries, $NO_x$ < 1 ppb), $O_3$ is characterised by net formation. Peroxy radicals (exemplified by $CH_3O_2$ and $HO_2$ in Figure 1.1) oxidise NO to $NO_2$, which, through the latter's photolysis, produces $O_3$, forming part of the free-radical-propagated (FRP) $O_3$-forming cycle (blue box in Figure 1.1). Increased $NO_x$ in these areas (while still below around 1 ppb) allows a greater number of FRP $O_3$-forming cycles to occur, generally outcompeting the radical termination reactions that result in $O_3$ destruction (green boxes in Figure 1.1). Hence $O_3$ concentrations increase. The $O_3$ formation as a result of VOC oxidation takes a long time (hours, days or even longer), which typically happens at a regional scale usually many kilometres downwind of its precursor emissions (AQEG, 2009). Further increases in $NO_x$ concentration ($NO_x$ > 1 ppb) leads to another radical termination reaction (OH + $NO_2$, orange box in Figure 1.1) becoming more dominant and this can interrupt the FRP $O_3$-formation cycle. This results in lower $O_3$ concentrations in more polluted urban environments. Within the urban environment the destruction of $O_3$ via reaction with NO also strongly influences the $O_3$ concentration locally (i.e. near road source). This is another pathway for $O_3$ removal other than the radical termination that forms nitric acid ($HNO_3$) (orange box in Figure 1.1). Occasionally, significant $O_3$ formation may also occur in urban areas during summer when the solar intensity is high due to $NO_2$ photolysis although the concentrations of VOCs are low (Jenkin, 2008). At the regional scale (100 – 1000 km), dry deposition of $O_3$ to the terrestrial surface also plays an important role on determining $O_3$ concentrations (The Royal Society, 2008).

$O_3$ dry deposition depends on the nature of the surface (e.g. vegetation/non-vegetation) and meteorological conditions that determine the magnitude of vertical flux.

Overall, as a secondary pollutant, with a lag in production following precursor emissions, $O_3$ is generally considered to be more spatially homogeneous than other primary pollutants. However, limited studies have shown considerable intra-urban spatial variation of $O_3$ due to local destruction of ozone near sources of high NO (Lin et al., 2016; Malmqvist et al., 2014; McConnell et al., 2006).

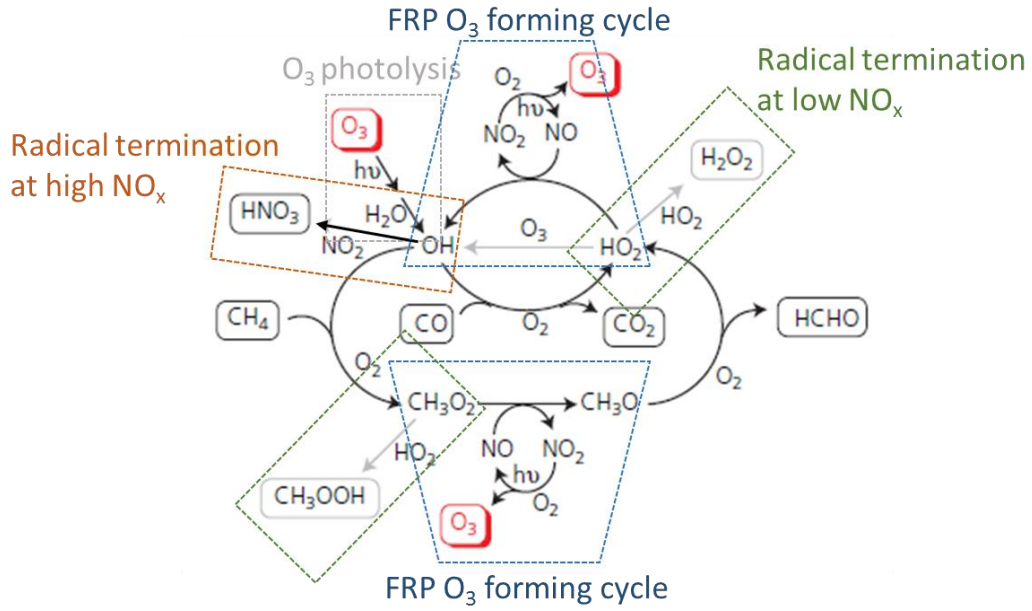



*Figure 1.1 Schematic representation of $O_3$ production and loss processes during the free-radical mediated atmospheric oxidation of Methane ($CH_4$) and CO. $CH_4$ is used here as a surrogate of VOCs. (Source: The Royal Society (2008))*

## 1.1.3 Particulate matter

$PM_{2.5}$ is defined as the mass concentration of particles with aerodynamic diameter less than 2.5 μm. The aerodynamic diameter differs from the physical diameter of particles and is related to the sampling technique of the particulate matter. The trajectory of particles in an air steam is determined by inertia (keeping particles on a straight line), pressure gradient (keeping particles along the air stream) and viscosity

(slowing down the movement). Smaller (lighter) particles tend to follow the air steam, while larger (heavier) particles impact on the object that the air stream encounters. Based on this principle cascade impactor or cyclone are designed so that particles smaller than the specified transmission curve stay with the air stream whilst larger particles impact on to a surface due to inertia (Heal et al., 2012). The reference method for quantifying the mass of $PM_{2.5}$ involves drawing air sample through an impactor or cyclone inlet, to separate the relevant particle sizes, and then through a filter that is weighed before and after the sampling under specified temperature and relative humidity conditions (Heal et al., 2012). The sampling period is typically 24 h. Since this method is labour intensive and only provides relatively long time-averaged data, UK monitoring stations use the reference-equivalent method implemented in the TEOM-FDMS (Tapered Element Oscillating Microbalance-Filter Dynamics Measurement System) to measure hourly $PM_{2.5}$ concentrations. The TEOM method indirectly measures the mass of PM by measuring the change in oscillation frequency of the filter resulting from the accumulation of sampled PM (Heal et al., 2012). As ambient PM typically contains a range of chemical components with different volatility, the mass of PM changes according to the partition of volatile or semi-volatile components in the gas and condensed phase under different environmental conditions. To control the variability in measurement caused by the volatility of PM components, the FDMS improves on the TEOM method by taking extra measures that better estimate the volatile component of PM and the volatilisation and condensation processes taking place on the filter during the weighing (Heal et al., 2012).

Ambient particular matter consists of solid and/or liquid particles that vary in size, shape and chemical compositions. The diverse physical and chemical properties of ambient PM mean that PM can be classified and characterised in different ways. Considering the different processes of particle formation, ambient particles can be classified into four modes: nucleation (1 – 30 nm), Aitken (20 – 100 nm), accumulation (30 – 1000 nm) and coarse (> 1 μm) (Heal et al., 2012). For particle size characterisation, however, the first two modes are usually considered together, giving the following size classifications: ultrafine particles (UFPs) (particles with diameter less than 100 nm), $PM_{2.5}$ (particles with aerodynamic diameter less than 2.5

μm) and $PM_{10}$ (particles with aerodynamic diameter less than 10 μm). Particles in different size fractions exhibit different characteristics in terms of particle number, surface area and mass (Figure 1.2). Most importantly Figure 1.2 shows that the majority of particle number concentration is contributed by the UFPs, whereas the majority of particle mass comes from the accumulation and coarse mode particles. For this reason, UFPs are usually quantified by the number concentration, whilst $PM_{2.5}$ and $PM_{10}$ are quantified by the mass concentration.

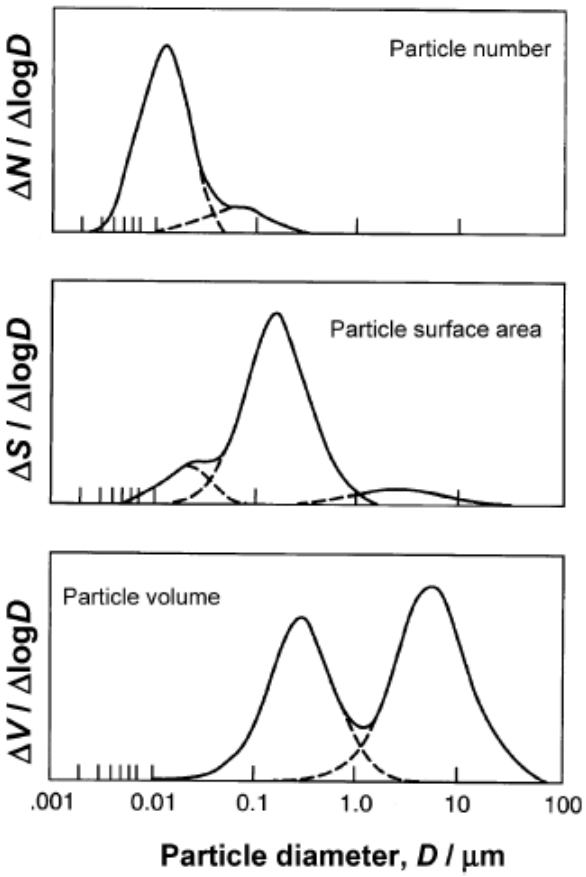



*Figure 1.2 A typical ambient particle distribution as a function of particle size expressed by particle number, surface area, and volume. The latter is equivalent to a mass distribution when variation in particle density is small. Vertical scaling is individual to each panel. (Source: Heal et al. (2012))*

### 1.1.3.1 Fine particles ($PM_{2.5}$)

Ambient $PM_{2.5}$ generally consists of a vast array of constituents including ions ($SO_4^{2-}$, $NO_3^-$, $NH_4^+$, $Na^+$ and $Cl^-$), elemental carbon, organic carbon, mineral material and water (Heal et al., 2012). Each constituent generally comes from a different source, thus $PM_{2.5}$ compositions will vary at different locations because of varying contributions from different sources. For example, in maritime areas $Na^+$ and $Cl^-$ might predominate due to primary emissions of sea-salt particles, whereas in farming areas agricultural emissions contribute most to the $NH_4^+$ component. In the urban environment, especially near roadside, elemental carbon and organic carbon are the components of most interest since they directly come from exhaust emission. These carbonaceous particles are initially produced in the nucleation mode by homogeneous, heterogeneous or reactive condensation in the exhaust emissions (Heal et al., 2012). The nucleation mode particles have a short lifetime as they coagulate or adsorb water and organic vapour to grow into accumulation mode particles (Kittelson et al., 2004) which remain stable in the atmosphere for several days (Hinds, 1999). $NO_x$ emission from transport or other combustion sources can be oxidised to $HNO_3$ over time, which reacts with $NH_3$ from agriculture contributing to the $NO_3^-$ component. Stringent exhaust emission standards and upgraded technology have decreased the contribution of tailpipe emissions to the total PM concentration, therefore increasing the relative contribution of non-exhaust traffic emissions (Kumar et al., 2013; Pant and Harrison, 2013). Non-exhaust emissions generally comprise tyre wear, brake wear, road dust and road surface abrasion which contribute to particles mostly larger than 1 μm (Pant and Harrison, 2013).

A number of factors affecting $PM_{2.5}$ variability in the urban environment were summarised by Pinto et al. (Pinto et al., 2004). These include: 1) local sources of primary PM; 2) topography; 3) transient emissions events; 4) meteorological phenomena; and 5) differences in the behaviour of semi-volatile components. Several studies have investigated the extent to which different factors affect $PM_{2.5}$ variability. A positive relationship was observed between $PM_{2.5}$ and traffic volume in London and Madrid, and elevated $PM_{2.5}$ concentrations during the cold season due to higher atmospheric stability and reduced mixing height (Kassomenos et al., 2014). Road

traffic can substantially elevate $PM_{2.5}$ concentration at immediate kerbside, but not a few metres from the kerb (AQEG, 2012). In the UK high $PM_{2.5}$ concentrations are frequently associated with air transported from continental Europe accompanied with high secondary inorganic components (AQEG, 2012). On the whole, $PM_{2.5}$ variability is not only governed by local emissions but also by the synoptic-scale meteorological phenomena which affects both the dispersion and deposition characteristics of the PM and the extent of long-range transport from regional sources.

### 1.1.3.2 Ultrafine particles

There is no reference method or standard metric to quantify UFP. As mentioned above, particle number concentration (UFPNC) is perhaps the most straightforward metric to measure UFP. One common method to count particle number is by measuring the light scattered by the particles as they pass through a laser beam. Scattered light is directly proportional to the number concentration of particles for a given particle size (Ramachandran et al., 2003). For UFPs in particular, this involves enlarging the particles by condensation of vapour (usually butanol or water) prior to the detection as particles smaller than ~300 nm are insensitive to optical scattering (Heal et al., 2012). Instrumentation using this technique is referred to as condensation particle counter (CPC), and is most widely used in studies quantifying UFP (Peters et al., 2014; Reche et al., 2011; Wallace and Ott, 2011).

The main source of UFP in the urban environment is road traffic which contributes about 90% of the total number concentration (Heal et al., 2012). Contributions from other anthropogenic sources, such as ship exhausts, cooking, biomass burning and aircraft emission are considered to be moderate compared with road traffic emission (Kumar et al., 2013). The UFP size fraction consists of nucleation and Aitken modes particles. Vehicle exhaust contributes to the nucleation mode particles via condensation and nucleation of vapour in the exhaust gases consisting of sulphates, nitrates and organic compounds (Heal et al., 2012). Formation of UFP from this route has diminished over the last decade due to reduction in the sulphur content of diesel fuel (Jones et al., 2012). Aitken mode particles are mainly made of a soot/ash core and formed through growth and coagulation of nucleation mode particles, and

are also from primary vehicle exhausts (Heal et al., 2012). Rapid transformation processes, including nucleation, coagulation, evaporation, condensation and deposition, happen within seconds after the emission of exhaust and continuously change the number concentration as the particles disperse (Kumar et al., 2011). Given the short lifetime of UFP, sharp concentration gradients are found with distance to the road traffic (HEI, 2010). Street canyons, which are common in the densely populated urban environments, introduce complex flow and dispersion characteristics that further increase the spatial variability. It has been reported that UFPNC in street canyons can be 27 times higher than the average background level not directly affected by near-source anthropogenic emissions; an even higher ratio was found in road tunnels (64 times higher than the background level) (Morawska et al., 2009). High-intensity secondary particle formation resulting in bursts in number concentration was also observed when traffic emissions were relatively low in the midday (Reche et al., 2011). These kind of nucleation episodes often result from intensive photochemical reactions accompanied with high $O_3$ levels.

#### 1.1.3.3 Black carbon

Although not regulated in UK or EU legislation, the darkness of ambient particles, has been measured in the UK for many decades, originally in the form of the black smoke metric and latterly as the black carbon (BC) metric. BC is considered to be an important component of $PM_{2.5}$ for being a marker of combustion sources and its impact on human health (WHO, 2012). BC refers to the light-absorbing carbonaceous particles that have strong absorption across a wide spectrum of visible wavelengths (Bond and Bergstrom, 2006). Based on this property, BC is quantified by the absorption of light by PM collected on a filter, which is then converted to the mass concentration. Although in theory the Beer-Lambert law predicts a linear relationship between light absorbance and BC concentration, in reality this relationship can be non-linear especially when excessive BC are collected on the filter, which is known as the loading effect (Virkkula et al., 2007). The mass extinction coefficient used to derive BC concentration has been shown to vary with the composition mix of the particles and the type of filters, which both add

uncertainties into calculating BC concentration through this method (Davy et al., 2017).

In the urban environment BC mainly originates from incomplete combustion of diesel fuels and wood burning from domestic heating. Contrary to UFP, BC is relatively inert and not likely degraded under atmospheric conditions so may transport over long distances (Venkatachari et al., 2006). A study in southern Scotland showed BC lifetime varied between 4 to 12 days depending on the precipitation which acts as a major removal process (Cape et al., 2012). High temporal variability was shown in the urban environment between days (Peters et al., 2014; Viidanoja et al., 2002) and within a day (Fruin et al., 2008; Peters et al., 2014), which was attributed to the wind speed, mixing height, background concentration fluctuation and diesel traffic density. Large spatial variability was also reported from measurements at 27 sites in New York (Maciejczyk et al., 2004) and mobile cycling measurements in Belgium (Peters et al., 2014). Significant difference was shown in BC concentration between streets (Peters et al., 2014), and the authors concluded that traffic density, street topology and distance between receptor and traffic sources had dominant effects on the spatial variation.

## 1.2 Health studies on air pollution

Evidence for adverse health effects of air pollution from both epidemiological and toxicological studies have been well documented in the past few decades. Epidemiological studies examine both short-term and long-term effects of exposure to air pollutants. The usual study design for investigating short-term effects of air pollution is time-series analysis, which regresses hourly or daily population exposure against disease data typically for a whole city (Moolgavkar et al., 2013). The long-term effects of air pollution are determined from cohort or spatial ecological studies, which regress annual exposure to air pollution against variation in diseases between urban areas or within a single urban area at smaller spatial units such as the addresses in the cohort (Schikowski et al., 2014) or the electoral wards (Haining et al., 2010). This section gives an overview of existing evidence on the short-term and long-term

effects of air pollution and different exposure assessment methods used in epidemiological studies.

### 1.2.1 Evidence on health effects of air pollution

A wide range of adverse health effects are associated with exposure to ambient air pollution. WHO periodically reviews the evidence on health aspects of air pollution to update its air quality guidelines. The most recent review on health aspects of air pollution (REVIHAAP) was published in 2013 (WHO, 2013). The following summary is mainly based on the content of this report unless specified otherwise.

A number of time-series studies have found associations between hourly and daily $NO_2$ concentrations and mortality. Studies have also shown associations of long-term $NO_2$ exposures with respiratory and cardiovascular mortality. Associations between adverse effects and $NO_2$ concentrations below the current EU limit values (both hourly and annual average) have been discovered in many studies, which suggests that lower guideline values for $NO_2$ are required. However, the question remains as to whether $NO_2$ is responsible for the adverse health outcomes on its own. In many studies the association between $NO_2$ and short-term health effects remained robust after adjustments for $PM_{2.5}$, $PM_{10}$ and sometimes black smoke. It is particularly difficult to separate the independent effect of $NO_2$ from other traffic-related air pollutants because the correlations among these pollutants are usually high. Proximity to roads (i.e. traffic-related air pollutants) were associated with adverse effects on health after adjusting for socioeconomic status and noise. Available evidence does not discern any individual or combinations of traffic-related air pollutants to be related to different health outcomes. However, based on the consistent epidemiological evidence on short-term and long-term effects of $NO_2$ and some mechanistic insights from toxicological studies, especially for respiratory effects, it is reasonable to infer that ambient exposure to $NO_2$ has some direct effects on human health.

Many epidemiological studies have assessed the effects of short-term exposure to $O_3$. Association was found between daily $O_3$ levels (maximum 1-hour or maximum rolling 8-hour mean) with all-cause, cardiovascular and respiratory mortality.

Adverse effects of short-term exposure to $O_3$ on both respiratory and cardiovascular hospital admissions have also been reported after adjustment for the effects of $PM_{10}$. There are fewer epidemiological studies on the long-term effects of $O_3$ on human health, but existing studies showed association between long-term exposure to $O_3$ and respiratory and cardiorespiratory mortality (Atkinson et al., 2016). Due to insufficient studies on the long-term health effect of $O_3$, a threshold for the long-term guideline for $O_3$ cannot be identified.

The adverse effects of PM are especially well documented. Accumulating evidence has suggested effects of both short-term and long-term exposure to $PM_{2.5}$ on mortality and morbidity. No evidence was found for a safe level of $PM_{2.5}$ below which no adverse health effects occur. There was also little evidence that one particular property of PM was responsible for the adverse health outcomes. Based on the results from health studies on PM comprising varying fractions and types, it is likely that both chemical composition and physical properties of particle are responsible for the observed health effects. For example, evidence has shown association between the BC component of PM with cardiovascular health effects and premature mortality for both short-term (daily) and long-term (annual) exposures, even after adjustment of $PM_{2.5}$ mass. Evidence on the association between short-term exposures to UFP and cardiorespiratory health has been emerging in recent years. Due to their small size toxicological studies suggested that UFPs act through mechanisms not shared with larger particles (e.g. $PM_{2.5}$ and $PM_{10}$) that dominate PM mass. Up to 2013 there have been no studies on the long-term effect of UFP, mainly due to the limited UFP monitoring in time and space that are essential for long-term exposure assessment (HEI, 2013).

### 1.2.2 Exposure assessment methods

Health studies investigating the short-term and long-term effects of air pollution use different exposure assessment methods. Studies of short-term effects of air pollution typically assign the pollutant concentration at the background fixed monitoring site as the population exposure. This will result in larger exposure estimate errors for pollutants with high spatial variability (e.g. $NO_2$ and UFP) than for pollutants with less spatial variability ($PM_{2.5}$) and $PM_{10}$). However, in time-series studies temporal

exposure contrast is the variable regressed against health data. It may be plausible to use fixed site monitoring data as the measure of exposure if the monitoring data can represent the temporal variability of the pollution over the study area even if large spatial contrast exists. Nevertheless, it is important in time-series studies to assess the suitability of the reliance on a fixed monitoring site to represent the temporal variability of the studied pollutant.

Studies of long-term effects of air pollution usually require pollutant concentration at much finer scale, which cannot be provided by the routine monitoring network. Therefore, different types of modelling methods are used to estimate pollutant concentration at the desired spatial scale. Three common types of models used in exposure assessment of ambient air pollution are reviewed here.

Interpolation models estimate the pollutant concentration at the desired locations from a set of monitoring sites distributed in the study area based on geostatistical techniques. The most common technique of this kind is the Kriging method (Jerrett et al., 2005). The Kriging method exploits spatial dependence in the measured pollutant concentration to estimate a continuous pollution field. Beyond the random error in the data, the spatial dependence is considered to be determined by two effects. The first effect is a measure of the broad trend in the pollutant concentration over the entire study area. The second effect is a measure of local variations as a function of distance between measurement points (Jerrett et al., 2005). A major advantage of the Kriging method compared with other geostatistical technique is the ability to quantify the degree of uncertainty in the predicted values. However, a major drawback of Kriging is the requirement of a reasonably dense monitoring network in order to obtain desired errors in estimates.

Land-use regression (LUR) models compute the relationship between pollutant concentration measured at a certain site and surrounding land use and traffic characteristics using regression analysis. The relationship is then applied to un-sampled locations to extrapolate the pollutant concentration. LUR models have been increasingly used in recent epidemiological studies of long-term exposure to ambient air pollution because they are relatively inexpensive to implement and can provide reliable estimates of pollutant concentration, especially for traffic-related air

pollutants (Hoek et al., 2008). However, as with other empirical models, the major limitation of LUR models is the transferability of models to other locations and time periods. Models developed in different areas usually differ in the terms in their regression equations, which is partly due to different contributions from different sources but is also related to the selection of monitoring sites and predictor variables to build the model (Hoek et al., 2008). The measurement data used to build LUR models typically come from passive samplers for $NO_x$ or active samplers for PM, both of which only give long-term average (weekly or monthly) pollutant concentration. Therefore, the LUR models developed from a set of measurements may only represent well the period when the measurements were taken. Hence temporal stability of the air pollution determines the capability of using LUR models to extrapolate pollutant concentrations in time. As an empirical model, the quality of a LUR model heavily depends on size and range of the measurement data (Hoek et al., 2008). While the effect of number of monitoring sites has been investigated before in a few studies (Wang et al., 2012), the effect of the distribution of monitoring sites on LUR models has not been explored. Chapter 5 of the thesis addresses this question and recommends ideal monitoring strategies for building LUR models and model validation process.

Dispersion models used in epidemiological studies for exposure assessment generally rely on Gaussian plume equations (Jerrett et al., 2005). The plume emitted from a constant stream of source can be mathematically proven to follow Gaussian distribution in lateral and vertical directions when the wind speed, wind direction and the propensity of atmosphere to disperse pollutant are considered to be constant (De Visscher, 2014). Although in reality none of these assumptions is fulfilled, the hourly average concentration is very close to the Gaussian profile in many cases (De Visscher, 2014). As a deterministic model, dispersion models estimate the pollutant concentration based on fundamental processes governing the emission and dispersion of air pollutants. Therefore, extensive data on emissions, meteorology and topography are required for dispersion models to carry out the calculation. Typically, there are three main aspects of data feeding into dispersion models: background concentration, meteorological data and emission data. Background data are usually obtained from routine monitoring sites at rural or urban background locations, or

larger scale model results. Meteorological data should include information on wind speed, wind direction, ambient temperature, solar radiation and cloud cover in order to estimate the atmospheric stability. Traffic emissions are usually estimated from traffic counts and emission factors. For point emissions (i.e. industrial sources), a number of parameters are required such as the height and diameter of the stack, temperature and exit velocity of the plume. Emissions data are usually reported on an annual basis. Therefore, to model pollutant concentration at high temporal resolution requires applying appropriate time-varying emissions factors to the emissions data. Dispersion models have distinct advantages over the previous two methods in providing high spatiotemporal pollution estimates and transferability. However, the performance of dispersion models depends on the availability and quality of the large amount of input data. These aspects often limit the application of dispersion models in wider areas. For this reason, a dispersion model (ADMS-Urban) is evaluated using UK-wide available emission and meteorological data to investigate the dispersion-modelling capability if a nationwide exposure assessment is required. Being a commercial dispersion model, different modules of the ADMS-Urban have been tested and validated in various case studies (CERC, 2016). However, the performance of ADMS-Urban in a general UK city context using national level input data has not been reported in literature. This is investigated in Chapter 4 of the thesis.

## 1.3 The changing paradigm of exposure assessment

The previous section has summarised different methods used in health studies to estimate human exposure. In fact, it is more accurate to describe them as methods to estimate pollution concentrations rather than human exposure. It is important to differentiate between the concepts of concentration and exposure, because exposure has a further meaning indicating whether interactions between environment and human exist. Once the intersection of air pollution and human is established, it may be also required to know the amount of pollutant that is actually inhaled and retained in the body to better understand the exposure-response relationship of air pollution. However, due to various technical and practical challenges current epidemiological studies are only capable of assigning the pollutant concentration as the exposure metric. Recent exposure assessment studies are starting to take account of the effect

of the time-activity pattern on population exposure (Dhondt et al., 2012; Smith et al., 2016). Although people spend only a small proportion of time (~ 5%) commuting, it was found that a much larger proportion (11 – 30%, depending on the pollutant) of the daily pollution exposure occurred in transit(de Nazelle et al., 2013; Dons et al., 2012, 2011). These studies highlight the importance of considering population mobility in the exposure assessment. It was suggested by the Health Effects Institute (HEI, 2010) that a hybrid exposure model, combining space-time-activity data, air quality model and personal measurement, would come closest to a "best" estimate of human exposure.

Traditional fixed-site measurements may be adequate to validate models predicting pollution concentrations at fixed points. Validating the hybrid models requires the air quality measurements to incorporate the population mobility as well. Personal exposure measurement is the ultimate link between ambient air pollution and human health effects and the ideal approach to validate the hybrid models. Due to the rapid development of lower-cost and portable air pollution monitors, personal exposure studies have been increasing in recent years and have shown varying levels of exposure to ambient air pollution in different micro-environments (Bekö et al., 2015; Wallace and Ott, 2011). Despite the promising future of personal exposure measurements, there remain a number of practical and technical challenges facing this emerging area of science, namely the development of robust low-cost sensors that produce high quality data, a rigorous evaluation method against reference instrument for the personal monitors, integration of data of different quality from various sources (e.g. reference instruments and low-cost sensors), and the scalability to a population level (Snyder et al., 2013; Steinle et al., 2013). It should be noted that the measurement principles of most personal monitors differ from that of their associated reference instrument. Therefore, assessment of their compatibility with reference instruments must be conducted before any application.

Research on the performance of low-cost air quality monitors has only become popular recently. There is no clear definition on what can be classified as an 'low-cost' air quality monitor. A great variety of commercial air quality monitors are emerging on the market. While they are considerably cheaper than the reference

instrument used in the routine monitoring networks, their prices range from a few hundred dollars to a few thousand dollars. The AQ-SPEC program is one of the first attempts aiming to provide a test centre for the 'low-cost' air quality monitors and to establish performance standards by which the air monitors are evaluated (http://www.aqmd.gov/aq-spec/home). So far a wide range of PM and gas air quality monitors have been evaluated in the AQ-SPEC program. In general, monitors that measure gaseous species, especially $NO_2$, performed worse than PM monitors. It is therefore of more interest to further investigate the reason causing the poor performance.

The REVIHAAP report also identified as a critical gap in current health studies which is the assessment of air pollution as multiple independent pollutants rather than as a mixture (WHO, 2013). Understanding the relationship between different pollutants and characterisation of the pollution mix would improve the assessment of independent and/or synergistic effects of ambient air pollutants (Dominici et al., 2010). Emerging low-cost portable monitors may be a solution to this problem as it may become practical to measure several pollutants simultaneously at wider spatial coverage. This is demonstrated in Chapter 3.

## 1.4 Aims and structure of this thesis

The aims of this thesis are in line with the future direction of exposure assessment identified in the previous section, namely investigating the capability and limitation of using personal monitors to improve human exposure assessment, understanding the relationship between different pollutant metrics and demonstrating how the current exposure assessment method can be improved by integration of high-resolution air quality modelling. The structure of this thesis is as follows.

Chapter 2 evaluates the performance of three types of low-cost portable monitors that measure $NO_2$, $O_3$ and $PM_{2.5}$, respectively. Special considerations regarding using these monitors in short-term monitoring campaigns or long-term personal monitoring are discussed. Chapter 3 demonstrates how short-term measurements with portable monitors can be used in conjunction with the data from fixed monitoring station to study the spatiotemporal variability and interrelationship between BC, UFP and

$PM_{2.5}$ in an urban area. Detailed relationships between UFP, $NO_2$ and inorganic components of $PM_{2.5}$ are investigated through longer-term measurements at roadside, urban background and rural sites. Chapter 4 describes the application of an atmospheric dispersion model (ADMS-Urban) in simulating high resolution pollutant concentrations in Edinburgh and its performance is evaluated with respect to the choices of different available input data. Chapter 5 presents a novel method to build and validate LUR models with the help of the dispersion model. Finally, Chapter 6 summaries the main results from previous chapters and their implications for exposure assessment. Challenges facing the evolution of exposure assessment are discussed in the context of findings from previous chapters.

# References

AQEG, 2012. Fine Particulate Matter (PM2.5) in the United Kingdom. Defra, UK. URL http://uk-air.defra.gov.uk/library/reports?report_id=727

AQEG, 2009. Ozone in the United Kingdom. Defra, UK. URL https://uk-air.defra.gov.uk/assets/documents/reports/aqeg/aqeg-ozone-report.pdf

Atkinson, R.W., Butland, B.K., Dimitroulopoulou, C., Heal, M.R., Stedman, J.R., Carslaw, N., Jarvis, D., Heaviside, C., Vardoulakis, S., Walton, H., Anderson, H.R., 2016. Long-term exposure to ambient ozone and mortality: a quantitative systematic review and meta-analysis of evidence from cohort studies. BMJ Open 6, e009493. doi:10.1136/bmjopen-2015-009493

Bekö, G., Kjeldsen, B.U., Olsen, Y., Schipperijn, J., Wierzbicka, A., Karottki, D.G., Toftum, J., Loft, S., Clausen, G., 2015. Contribution of various microenvironments to the daily personal exposure to ultrafine particles: Personal monitoring coupled with GPS tracking. Atmos. Environ. 110, 122–129. doi:10.1016/j.atmosenv.2015.03.053

Bond, T.C., Bergstrom, R.W., 2006. Light Absorption by Carbonaceous Particles: An Investigative Review. Aerosol Sci. Technol. 40, 27–67. doi:10.1080/02786820500421521

Cape, J.N., Coyle, M., Dumitrean, P., 2012. The atmospheric lifetime of black carbon. Atmos. Environ. 59, 256–263. doi:10.1016/j.atmosenv.2012.05.030

CERC, 2016. ADMS Model validation [WWW Document]. URL http://cerc.co.uk/environmental-software/model-validation.html (accessed 17-November-2016)

Cyrys, J., Eeftens, M., Heinrich, J., Ampe, C., Armengaud, A., Beelen, R., Bellander, T., Beregszaszi, T., Birk, M., Cesaroni, G., Cirach, M., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Dėdelė, A., Dimakopoulou, K., Eriksen, K., Galassi, C., Grąulevičienė, R., Grivas, G., Gruzieva, O., Gustafsson, A.H., Hoffmann, B., Iakovides, M., Ineichen, A., Krämer, U., Lanki, T., Lozano, P., Madsen, C., Meliefste, K., Modig, L., Mölter, A., Mosler, G., Nieuwenhuijsen, M., Nonnemacher, M., Oldenwening, M., Peters, A., Pontet, S., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Sugiri, D., Stephanou, E.G., Taimisto, P., Tsai, M.-Y., Vaskövi, É., Villani, S., Wang, M., Brunekreef, B., Hoek, G., 2012. Variation of NO2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study. Atmos. Environ. 62, 374–390. doi:10.1016/j.atmosenv.2012.07.080

Davis, J.C., 2011. Statistics and data analysis in geology, 3rd ed. John Wiley & Sons.

Davy, P.M., Tremper, A.H., Nicolosi, E.M.G., Quincey, P., Fuller, G.W., 2017. Estimating particulate black carbon concentrations using two offline light absorption methods applied to four types of filter media. Atmos. Environ. 152, 24–33. doi:10.1016/j.atmosenv.2016.12.010

de Nazelle, A., Seto, E., Donaire-Gonzalez, D., Mendez, M., Matamala, J., Nieuwenhuijsen, M.J., Jerrett, M., 2013. Improving estimates of air pollution exposure through ubiquitous sensing technologies. Environ. Pollut. 176, 92–99. doi:10.1016/j.envpol.2012.12.032

De Visscher, A., 2014. Air dispersion modeling: foundations and applications. John Wiley & Sons, Inc, Hoboken, New Jersey.

Dhondt, S., Beckx, C., Degraeuwe, B., Lefebvre, W., Kochan, B., Bellemans, T., Int Panis, L., Macharis, C., Putman, K., 2012. Health impact assessment of air pollution using a dynamic exposure profile: Implications for exposure and health impact estimates. Environ. Impact Assess. Rev. 36, 42–51. doi:10.1016/j.eiar.2012.03.004

Dominici, F., Peng, R.D., Barr, C.D., Bell, M.L., 2010. Protecting human health from air pollution: shifting from a single-pollutant to a multipollutant approach. Epidemiol. Camb. Mass 21, 187–194. doi:10.1097/EDE.0b013e3181cc86e8

Dons, E., Int Panis, L., Van Poppel, M., Theunis, J., Wets, G., 2012. Personal exposure to Black Carbon in transport microenvironments. Atmos. Environ. 55, 392–398. doi:10.1016/j.atmosenv.2012.03.020

Dons, E., Int Panis, L., Van Poppel, M., Theunis, J., Willems, H., Torfs, R., Wets, G., 2011. Impact of time–activity patterns on personal exposure to black carbon. Atmos. Environ. 45, 3594–3602. doi:10.1016/j.atmosenv.2011.03.064

Dunlea, E.J., Herndon, S.C., Nelson, D.D., Volkamer, R.M., San Martini, F., Sheehy, P.M., Zahniser, M.S., Shorter, J.H., Wormhoudt, J.C., Lamb, B.K., others, 2007. Evaluation of nitrogen dioxide chemiluminescence monitors in a polluted urban environment. Atmospheric Chem. Phys. 7, 2691–2704.

EC, 2010. Guide to the demonstration of equivalence of ambient air monitoring methods. European Commission. URL http://ec.europa.eu/environment/air/quality/legislation/pdf/equivalence.pdf

EC, 2008. DIRECTIVE 2008/50/EC OF The European Parliament And Of The Council of 21 May 2008 On Ambient Air Quality And Cleaner Air For Europe. European Commission. URL http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008L0050&qid=1471948031872&from=en

Fruin, S., Westerdahl, D., Sax, T., Sioutas, C., Fine, P.M., 2008. Measurements and predictors of on-road ultrafine particle concentrations and associated pollutants in Los Angeles. Atmos. Environ. 42, 207–219. doi:10.1016/j.atmosenv.2007.09.057

Haining, R., Li, G., Maheswaran, R., Blangiardo, M., Law, J., Best, N., Richardson, S., 2010. Inference from ecological models: Estimating the relative risk of stroke from air pollution exposure using small area data. Spat. Spatio-Temporal Epidemiol., GEOMED ConferenceGEOMED Conference 1, 123–131. doi:10.1016/j.sste.2010.03.006

Heal, M.R., Kumar, P., Harrison, R.M., 2012. Particles, air quality, policy and health. Chem. Soc. Rev. 41, 6606. doi:10.1039/c2cs35076a

HEI, 2013. Understanding the Health Effects of Ambient Ultrafine Particles. Health Effects Institute, Boston, Massachusetts. URL https://www.healtheffects.org/system/files/Perspectives3.pdf

HEI, 2010. Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects. Health Effects Institute, Boston, Massachusetts. URL https://www.healtheffects.org/publication/traffic-related-air-pollution-critical-review-literature-emissions-exposure-and-health

Hinds, W.C., 1999. Aerosol technology: properties, behavior, and measurement of airborne particles. Wiley.

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. Atmos. Environ. 42, 7561–7578. doi:10.1016/j.atmosenv.2008.05.057

Jenkin, M.E., 2008. Trends in ozone concentration distributions in the UK since 1990: Local, regional and global influences. Atmos. Environ. 42, 5434–5445. doi:10.1016/j.atmosenv.2008.02.036

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., Giovis, C., 2005. A review and evaluation of intraurban air pollution exposure models. J. Expo. Sci. Environ. Epidemiol. 15, 185–204. doi:10.1038/sj.jea.7500388

Jones, A.M., Harrison, R.M., Barratt, B., Fuller, G., 2012. A large reduction in airborne particle number concentrations at the time of the introduction of "sulphur free" diesel and the London Low Emission Zone. Atmos. Environ. 50, 129–138. doi:10.1016/j.atmosenv.2011.12.050

Kassomenos, P.A., Vardoulakis, S., Chaloulakou, A., Paschalidou, A.K., Grivas, G., Borge, R., Lumbreras, J., 2014. Study of PM10 and PM2.5 levels in three European cities: Analysis of intra and inter urban variations. Atmos. Environ. 87, 153–163. doi:10.1016/j.atmosenv.2014.01.004

Kittelson, D.B., Watts, W.F., Johnson, J.P., 2004. Nanoparticle emissions on Minnesota highways. Atmos. Environ. 38, 9–19. doi:10.1016/j.atmosenv.2003.09.037

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. J. Aerosol Sci. 42, 580–603. doi:10.1016/j.jaerosci.2011.06.001

Kumar, P., Pirjola, L., Ketzel, M., Harrison, R.M., 2013. Nanoparticle emissions from 11 non-vehicle exhaust sources – A review. Atmos. Environ. 67, 252–277. doi:10.1016/j.atmosenv.2012.11.011

Lin, C., Feng, X., Heal, M.R., 2016. Temporal persistence of intra-urban spatial contrasts in ambient NO2, O3 and Ox in Edinburgh, UK. Atmospheric Pollut. Res. 7, 734–741. doi:10.1016/j.apr.2016.03.008

Maciejczyk, P.B., Offenberg, J.H., Clemente, J., Blaustein, M., Thurston, G.D., Chi Chen, L., 2004. Ambient pollutant concentrations measured by a mobile laboratory in South Bronx, NY. Atmos. Environ., Particulate Matter: Atmospheric Sciences, Exposure and the Fourth Colloquium on PM and Human Health - Papers from the AAAR PM Meeting 38, 5283–5294. doi:10.1016/j.atmosenv.2004.02.062

Malmqvist, E., Olsson, D., Hagenbjörk-Gustafsson, A., Forsberg, B., Mattisson, K., Stroh, E., Strömgren, M., Swietlicki, E., Rylander, L., Hoek, G., Tinnerberg, H., Modig, L., 2014. Assessing ozone exposure for epidemiological studies in Malmö and Umeå, Sweden. Atmos. Environ. 94, 241–248. doi:10.1016/j.atmosenv.2014.05.038

McConnell, R., Berhane, K., Yao, L., Lurmann, F.W., Avol, E., Peters, J.M., 2006. Predicting residential ozone deficits from nearby traffic. Sci. Total Environ. 363, 166–174. doi:10.1016/j.scitotenv.2005.06.028

Moolgavkar, S.H., Roger O. McClellan, Anup Dewanji, Jay Turim, E. Georg Luebeck, Melanie Edwards, 2013. Time-Series Analyses of Air Pollution and

Mortality in the United States: A Subsampling Approach. Environ. Health Perspect. 121, 73–78. doi:10.1289/ehp.1104507

Morawska, L., Wang, H., Ristovski, Z., Jayaratne, E.R., Johnson, G., Cheung, H.C., Ling, X., He, C., 2009. JEM Spotlight: Environmental monitoring of airborne nanoparticles. J. Environ. Monit. 11, 1758–1773. doi:10.1039/B912589M

Özkaynak, H., Baxter, L.K., Dionisio, K.L., Burke, J., 2013. Air pollution exposure prediction approaches used in air pollution epidemiology studies. J. Expo. Sci. Environ. Epidemiol. 23, 566–572. doi:10.1038/jes.2013.15

Pant, P., Harrison, R.M., 2013. Estimation of the contribution of road traffic emissions to particulate matter concentrations from field measurements: A review. Atmos. Environ. 77, 78–97. doi:10.1016/j.atmosenv.2013.04.028

Peters, J., Van den Bossche, J., Reggente, M., Van Poppel, M., De Baets, B., Theunis, J., 2014. Cyclist exposure to UFP and BC on urban routes in Antwerp, Belgium. Atmos. Environ. 92, 31–43. doi:10.1016/j.atmosenv.2014.03.039

Pinto, J.P., Lefohn, A.S., Shadwick, D.S., 2004. Spatial variability of PM2.5 in urban areas in the United States. J. Air Waste Manag. Assoc. 1995 54, 440–449.

Ramachandran, G., Adgate, J.L., Pratt, G.C., Sexton, K., 2003. Characterizing Indoor and Outdoor 15 Minute Average PM 2.5 Concentrations in Urban Neighborhoods. Aerosol Sci. Technol. 37, 33–45. doi:10.1080/02786820300889

Reche, C., Querol, X., Alastuey, A., Viana, M., Pey, J., Moreno, T., Rodríguez, S., González, Y., Fernández-Camacho, R., de la Campa, A.M.S., de la Rosa, J., Dall'Osto, M., Prévôt, A.S.H., Hueglin, C., Harrison, R.M., Quincey, P., 2011. New considerations for PM, Black Carbon and particle number concentration for air quality monitoring across different European cities. Atmospheric Chem. Phys. 11, 6207–6227. doi:10.5194/acp-11-6207-2011

Schikowski, T., Adam, M., Marcon, A., Cai, Y., Vierkötter, A., Carsin, A.E., Jacquemin, B., Kanani, Z.A., Beelen, R., Birk, M., Bridevaux, P.-O., Brunekeef, B., Burney, P., Cirach, M., Cyrys, J., Hoogh, K. de, Marco, R. de, Nazelle, A. de, Declercq, C., Forsberg, B., Hardy, R., Heinrich, J., Hoek, G., Jarvis, D., Keidel, D., Kuh, D., Kuhlbusch, T., Migliore, E., Mosler, G., Nieuwenhuijsen, M.J., Phuleria, H., Rochat, T., Schindler, C., Villani, S., Tsai, M.-Y., Zemp, E., Hansell, A., Kauffmann, F., Sunyer, J., Probst-Hensch, N., Krämer, U., Künzli, N., 2014. Association of ambient air pollution with the prevalence and incidence of COPD. Eur. Respir. J. 44, 614–626. doi:10.1183/09031936.00132213

Sheppard, L., Burnett, R.T., Szpiro, A.A., Kim, S.-Y., Jerrett, M., Pope, C.A., Brunekreef, B., 2012. Confounding and exposure measurement error in air pollution epidemiology. Air Qual. Atmosphere Health 5, 203–216. doi:10.1007/s11869-011-0140-9

Smith, J.D., Mitsakou, C., Kitwiroon, N., Barratt, B.M., Walton, H.A., Taylor, J.G., Anderson, H.R., Kelly, F.J., Beevers, S.D., 2016. London Hybrid Exposure Model: Improving Human Exposure Estimates to NO2 and PM2.5 in an Urban Setting. Environ. Sci. Technol. 50, 11760–11768. doi:10.1021/acs.est.6b01817

Snyder, E.G., Watkins, T.H., Solomon, P.A., Thoma, E.D., Williams, R.W., Hagler, G.S.W., Shelow, D., Hindin, D.A., Kilaru, V.J., Preuss, P.W., 2013. The

Changing Paradigm of Air Pollution Monitoring. Environ. Sci. Technol. 47, 11369–11377. doi:10.1021/es4022602

Steinle, S., Reis, S., Sabel, C.E., 2013. Quantifying human exposure to air pollution -- Moving from static monitoring to spatio-temporally resolved personal exposure assessment. Sci. Total Environ. 184. doi:10.1016/j.scitotenv.2012.10.098

The Royal Society, 2008. Ground-level ozone in the 21st century: future trends, impacts and policy implications. The Royal Society, London.

Vardoulakis, S., Lumbreras, J., Solazzo, E., 2009. Comparative evaluation of nitrogen oxides and ozone passive diffusion tubes for exposure studies. Atmos. Environ. 43, 2509–2517. doi:10.1016/j.atmosenv.2009.02.048

Vardoulakis, S., Phoon, X., Ochieng, C., 2010. Health Effects of Air Pollutants, in: Lazaridis, M., Colbeck, I. (Eds.), Human Exposure to Pollutants via Dermal Absorption and Inhalation, Environmental Pollution. Springer Netherlands, pp. 143–184.

Venkatachari, P., Zhou, L., Hopke, P.K., Felton, D., Rattigan, O.V., Schwab, J.J., Demerjian, K.L., 2006. Spatial and temporal variability of black carbon in New York City. J. Geophys. Res. Atmospheres 111, D10S05. doi:10.1029/2005JD006314

Viidanoja, J., Sillanpää, M., Laakia, J., Kerminen, V.-M., Hillamo, R., Aarnio, P., Koskentalo, T., 2002. Organic and black carbon in PM2.5 and PM10: 1 year of data from an urban site in Helsinki, Finland. Atmos. Environ. 36, 3183–3193. doi:10.1016/S1352-2310(02)00205-4

Virkkula, A., Mäkelä, T., Hillamo, R., Yli-Tuomi, T., Hirsikko, A., Hämeri, K., Koponen, I.K., 2007. A simple procedure for correcting loading effects of aethalometer data. J. Air Waste Manag. Assoc. 1214. doi:10.3155/1047-3289.57.10.1214

Wallace, L., Ott, W., 2011. Personal exposure to ultrafine particles. J. Expo. Sci. Environ. Epidemiol. 21, 20–30. doi:10.1038/jes.2009.59

Warton, D.I., Wright, I.J., Falster, D.S., Westoby, M., 2006. Bivariate line-fitting methods for allometry. Biol. Rev. 81, 259–291. doi:10.1017/S1464793106007007

Wang, M., Beelen, R., Eeftens, M., Meliefste, K., Hoek, G., Brunekreef, B., 2012. Systematic Evaluation of Land Use Regression Models for NO2. Environ. Sci. Technol. 46, 4481–4489. doi:10.1021/es204183v

WHO, 2013. Review of evidence on health aspects of air pollution – REVIHAAP project: final technical report. WHO, Copenhagen, Denmark. URL http://www.euro.who.int/en/health-topics/environment-and-health/air-quality/publications/2013/review-of-evidence-on-health-aspects-of-air-pollution-revihaap-project-final-technical-report

WHO, 2012. Health effects of black carbon. WHO, Copenhagen, Denmark. URL http://www.euro.who.int/__data/assets/pdf_file/0004/162535/e96541.pdf

WHO, 2006. Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide. WHO, Copenhagen, Denmark. URL http://www.euro.who.int/__data/assets/pdf_file/0005/78638/E90038.pdf

# Chapter 2 Evaluation of low-cost portable air quality monitors

## 2.1 Introduction

Air pollutant concentrations are monitored by national governments, agencies or local authorities to assess compliance with air quality legislation. The common monitoring strategy adopted by governments has been to make measurements with reference or reference-equivalent method at a small number of locations notionally representative of urban or rural environments. These monitoring data typically have high temporal resolution and are widely used in epidemiological studies to investigate short-term health impacts of air pollution (Samoli et al., 2016). Studies of long-term effects of air pollution require more spatially resolved data which usually come from a monitoring network purposely designed by the researchers. In these studies, gaseous pollutants such as $NO_x$ and VOCs are generally measured with passive samplers, whereas PM is measured through active sampling onto a filter followed by weighing. These measurement techniques are relatively low-cost but can only give long-term averaged (weekly or monthly) pollutant concentrations. Advancing the current exposure assessment method requires integration of individual human mobility, which demands the measurements to have both great spatial and temporal resolution that the two monitoring strategies discussed above do not possess.

Research utilising portable 'low-cost' air quality monitors that measure real-time pollutant concentrations have been increasing in recent years, driven by the incentive to achieve more detailed source attribution (Heimann et al., 2015) and better estimation of personal exposure (Steinle et al., 2015). The extensive measurements in space and time achievable by these air quality monitors are attractive to researchers in many aspects including better human exposure assessment and insight into high-resolution model performance. It is recognised from the outset that these sensors are unlikely to replace current monitoring networks or to provide equivalent measurements in terms of the robustness, precision and accuracy. However the

capability of these sensors should still be scrutinised and classifications should be made to inform potential users whether the sensors produce indicative data or only qualitative data for educational use (Lewis and Edwards, 2016).

A protocol for the calibration and evaluation of low-cost gas sensors has been published by the European Commission (Spinelle et al., 2013). This document detailed the evaluation procedures for determination of whether a monitor can be accepted as an indicative method as defined in the Directive 2008/50/EC (2008). An indicative method is defined as that the relative expanded uncertainty of a method is within the acceptable uncertainty set out in the Directive, which are 30% for $O_3$, 25% for $NO_2$ and 50% for $PM_{2.5}$ at their respective limit or target value (EC, 2008). The evaluation procedures involve the development of gas sensors in laboratory testing, to select important sensor parameters to convert to air pollutant concentration, and field testing which compares sensor's output with the reference method. The commercial monitors evaluated in this chapter can be classified as 'black box' monitors, which means that they directly output a pollutant concentration whilst the internal algorithms converting the sensor's raw response to a pollutant concentration are unknown to the user. It is therefore only possible to evaluate the monitor's performance in field testing, which includes calculation of the regression between the monitor and reference measurements and the comparison of uncertainty of the monitor with the data quality objective (DQO) for the indicative method defined in the Directive 2008/50/EC. In reality there are many other aspects of the monitor's performance that need to be considered in addition to complying with the DQO such as the extent of agreement between duplicate monitors, the stability of the monitor's response over time, limits of detection and sensitivity to potential confounding variables such as temperature or other pollutants.

This chapter evaluates three types of relatively low-cost portable air quality monitors that measure $NO_2$, $O_3$ and $PM_{2.5}$. These monitors are Aeroqual series 500 $O_3$ and $NO_2$ portable monitors (~$500) and RTI microPEM $PM_{2.5}$ monitor ($2,000). The Aeroqual $O_3$ and microPEM monitors have previously been tested in the AQ-PEC program and had $R^2$ of 0.65 – 0.90 and 0.85, respectively, with their respective reference instruments (http://www.aqmd.gov/aq-spec/evaluations/summary). This

result compared favourably with other 'low-cost' monitors, and therefore set the benchmark for other air quality monitors. The evaluation of these monitors in this chapter follows the protocol published by Spinelle et al. (2013), but also extend to the field usability of the monitors. Capabilities and limitations on their application in exposure assessment studies are discussed.

## 2.2 Evaluation of Aeroqual Series 500 portable monitors

### 2.2.1 Introduction

Aeroqual series 500 monitors use electrochemical and metal oxide semiconductor (MOS) sensors to measure $NO_2$ and $O_3$, respectively (Aeroqual, 2016). An electrochemical sensor converts current resulting from a redox reaction of the selected gas on the working electrode to the gas concentration (Mead et al., 2013). A MOS sensor relies on detecting the conductivity changes of heated semiconducting oxide induced by the presence of reactive gases (Williams et al., 2009). Cross interference with $O_3$ has been shown to be a significant issue for the $NO_2$ electrochemical sensor (Mead et al., 2013). Cross-sensitivity to other gases including chlorine ($Cl_2$) and hydrogen sulphide ($H_2S$) are also significant for $O_3$ MOS sensor (Williams et al., 2009). However due to the extremely low concentrations of these gases in the ambient atmosphere, the MOS ozone sensor is essentially free from interference under ambient conditions. Electrochemical and MOS sensors are in general sensitive to change in relative humidity and temperature (Snyder et al., 2013). It is therefore important to maintain constancy of these environmental variables inside the monitors during operation.

### 2.2.2 Method

To compare the performance of the Aeroqual monitors against their respective reference instruments, Aeroqual $NO_2$ and $O_3$ monitors were deployed at the Edinburgh St Leonards monitoring station during the summer of 2015. The St Leonards monitoring station is part of the UK national network (AURN) that provides hourly averaged measurements of $NO_2$, $O_3$ and $PM_{2.5}$. The deployment schedule is summarised in Table 2.1. The logging interval was 1 min during most

deployments, which allowed continuous monitoring of ~6 days' worth of data before memory capacity was reached. During the last deployment, a 5 min logging interval was used to expand the continuous monitoring to 13 days. The analysis was conducted on hourly averaged data with a data capture requirement per hour of 75%. A portable weather sensor Kestrel 4500 (http://www.r-p-r.co.uk/kestrel/k4500.php) was co-located with the Aeroqual monitors to measure temperature and relative humidity. The measurement frequency of the Kestrel 4500 was 5 min. Hourly averaged data were calculated for the analysis of the effect of temperature and relative humidity on the Aeroqual monitors' performance.

*Table 2.1 Summary of Aeroqual S500 monitor deployment periods in 2015 at the Edinburgh St Leonards AURN site.*

| Deployment period | Start date | End date | No. of days | Logging interval |
|---|---|---|---|---|
| **P1** | 06-09 | 06-15 | 6 | 1 min |
| **P2** | 07-21 | 07-28 | 7 | 1 min |
| **P3** | 07-28 | 08-03 | 6 | 1 min |
| **P4** | 08-05 | 08-11 | 6 | 1 min |
| **P5** | 08-21 | 08-27 | 6 | 1 min |
| **P6** | 08-27 | 09-09 | 13 | 5 min |

### 2.2.3 Performance of the Aeroqual Series 500 $O_3$ monitor

Figure 2.1 shows the scatter plot and the major axis (MA) regression results comparing the Aeroqual $O_3$ monitor and reference $O_3$ measurement for each deployment period. Details on the MA regression method are explained in Appendix I. Correlation between Aeroqual $O_3$ monitor and reference instrument was very high during each deployment period ($r^2 > 0.90$). The gradients calculated from the regression analysis for each period were similar, but the intercept showed a generally decreasing trend. This suggests that the baseline of the Aeroqual $O_3$ monitor may have drifted between deployments. Further investigation did not find this drift of baseline to be related to ambient temperature ($T$) or relative humidity ($RH$). As it is not technically feasible for the user to correct the Aeroqual $O_3$ baseline, the best practice for using the Aeroqual to measure $O_3$ in field measurements would be to

time the calibration of the Aeroqual monitor as close to the actual measurement period as possible to reduce the impact of potential sensor drift on the observed concentrations. In addition, different calibration equations may be required for the Aeroqual $O_3$ measurements if field measurements span an extensive period of time in order to account for potential sensor drift.

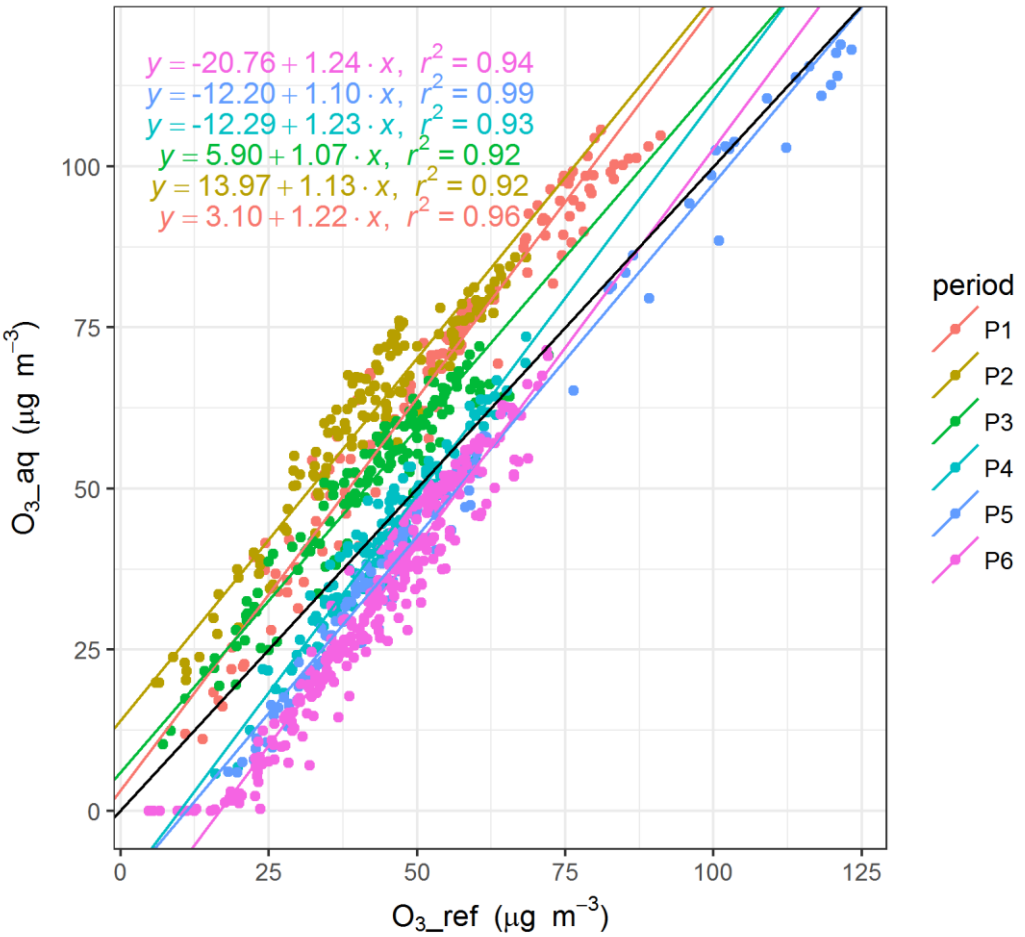



*Figure 2.1 Scatter plot of $O_3$ concentrations measured by Aeroqual monitor and reference instrument for different deployment periods. MA regression equations for each deployment period are shown in the top-left of the graph. 95% confidence interval of the regression coefficients are summarised in Appendix II Table 1. The black line shows the 1:1 line.*

To investigate the sensor's performance after calibration with the reference analyser, P2 and P5 were selected as calibration periods, from which the resulting calibration equations were applied to the data in P3 and P6, respectively. P2/P3 and P5/P6 are two pairs of consecutive deployment periods that represent the situation where the sensor calibration is undertaken close in time to the actual measurements. Figure 2.2

shows the scatter plots of calibrated Aeroqual $O_3$ data against the reference instrument for the two periods. The statistics assessing the agreement are shown in Table 2.2.

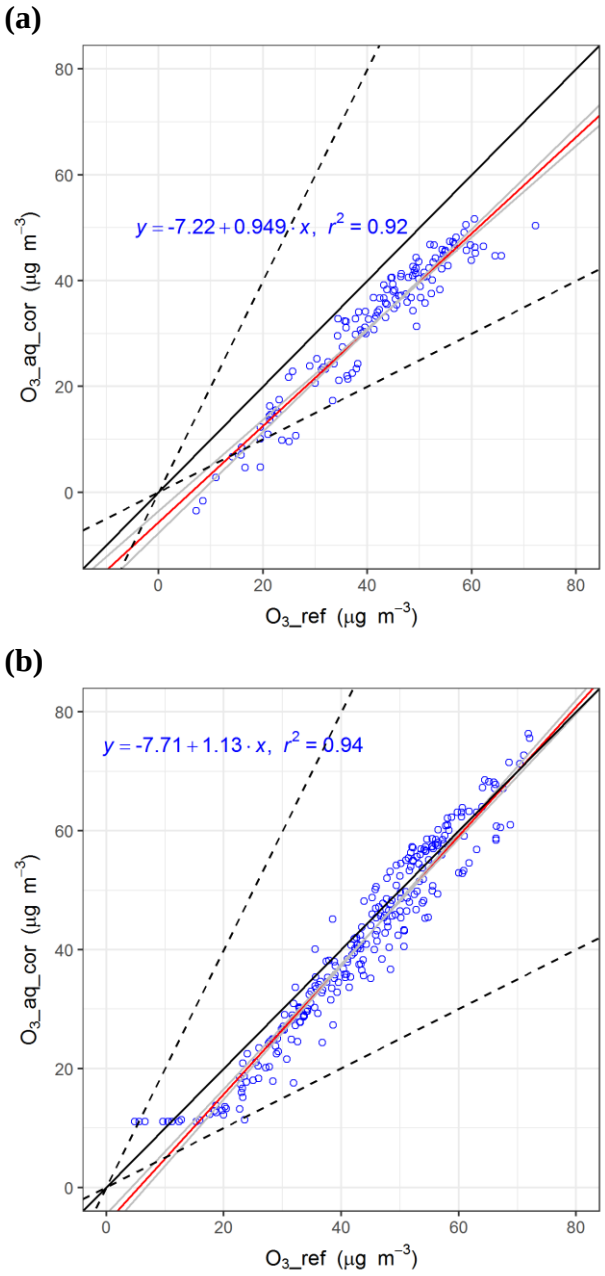



*Figure 2.2 Scatter plots of hourly-averaged calibrated Aeroqual $O_3$ vs reference $O_3$. (a) Data for P3 calibrated using data from P2. (b) Data for P6 calibrated using data from P5. The red line shows linear relationship between the two variables calculated from MA regression analysis with the grey lines showing the 95% confidence interval. Solid black line shows the 1:1 line, and dashed lines show the 1:2 and 2:1 lines.*

*Table 2.2 Summaries of pairwise statistics between hourly-averaged calibrated Aeroqual $O_3$ and reference $O_3$. RMSE is root mean squared error. FAC2 is the percentage of the ratios of the paired data that are within 0.5 to 2. The definitions of all the evaluation statistics are given in Appendix I.*

| **Calibration period** | **Validation period** | ***N*** | **FAC2** | **Mean Bias ($\mu g\ m^{-3}$)** | **RMSE ($\mu g\ m^{-3}$)** | ***$r^2$*** |
|---|---|---|---|---|---|---|
| **P2** | P3 | 131 | 92% | –9.4 | 10.1 | 0.92 |
| **P5** | P6 | 259 | 99% | –2.3 | 4.7 | 0.94 |

Figure 2.2 shows that the calibration equation derived from P2 was not suitable for the data in P3 as indicated by the scatter points being systematically under the 1:1 line. This observation suggests that even between two consecutive deployment periods the Aeroqual $O_3$ monitor may still be affected by a significant amount of baseline drift. This drift appeared to be consistent in direction, but by an inconsistent rate (Figure 2.1). As a result, comparison between Aeroqual $O_3$ monitor and reference instrument before and after field measurement is necessary to examine if significant drift in baseline occurs.

### 2.2.4 Performance of the Aeroqual series 500 $NO_2$ monitor

The Aeroqual $NO_2$ monitor has previously been identified to be subject to $O_3$ cross sensitivity (Lin et al., 2015). $NO_2$ measured directly by the Aeroqual monitor correlated poorly with the reference instrument, however the difference between $NO_2$ measured by Aeroqual and $NO_2$ measured by reference instrument ($\Delta NO_2$) was found to correlate well with $O_3$ concentration (Lin et al., 2015). This allows the calibration of Aeroqual $NO_2$ measurements if simultaneous $O_3$ measurements are also present. Lin et al. (2015) regressed $\Delta NO_2$ against $O_3$ measurements, which is equivalent to regressing $NO_2$ measured by Aeroqual [$NO_2$(Aq)] against $NO_2$ measured by reference analyser [$NO_2$(ref)] and $O_3$ measured by Aeroqual [$O_3$(aq)] with the coefficient of $NO_2$(ref) fixed to 1 (Fit1 in Table 2.3). The Lin et al. equation assumes that one unit change in the reference $NO_2$ will result in one unit change in the Aeroqual $NO_2$, keeping $O_3$ constant. However, it is likely that the response in the Aeroqual signal solely due to $NO_2$ may not be 1:1 to the response of reference $NO_2$. Hence a multiple linear relationship of Aeroqual $NO_2$ against reference $NO_2$ and

Aeroqual $O_3$ (shown as Fit2 in Table 2.3) is also plausible and has not been investigated before.

*Table 2.3 The equation used to calibration Aeroqual $NO_2$ measurements in Lin et al. (2015) and the two variations of regression equations used in this study to calibrate Aeroqual $NO_2$ measurements. Fit1 has the same underlying algebraic form as the Lin et al. (2015) approach.*

| | |
|---|---|
| **Lin et al. (2015)** | $\Delta NO_2\ [NO_2(Aq) - NO_2(ref)] = a \cdot O_3(Aq) + b$ |
| **Fit1** | $NO_2(Aq) = NO_2(ref) + a \cdot O_3(Aq) + b$ |
| **Fit2** | $NO_2(Aq) = a_1 \cdot NO_2(ref) + a_2 \cdot O_3(Aq) + b$ |

Figure 2.3 shows the variation in $NO_2(Aq)$ ($r^2$) explained by the two regression models for each deployment period. For most of the deployment periods, $r^2$ values calculated by the two regression models were similar. However ANOVA analysis suggested that the sum of squared residuals (SSR) for Fit2 was significantly smaller ($P < 0.05$) than the SSR for Fit1 for all deployment periods except P2, suggesting that Fit2 may be the better model to calibrate $NO_2(Aq)$.

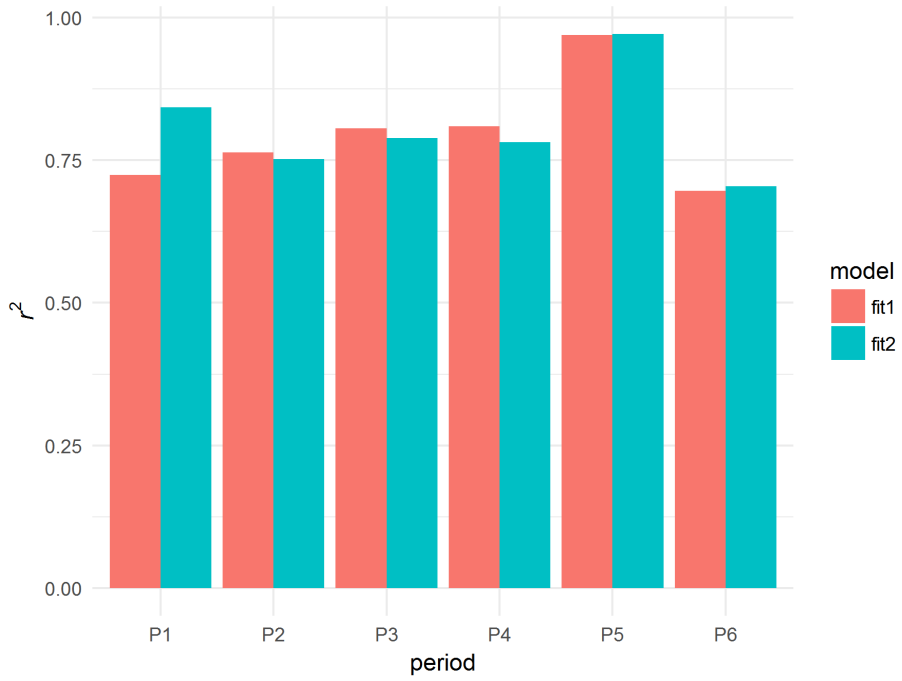



*Figure 2.3 $r^2$ values of the two regression models in Table 2.3 for different deployment periods.*

Regression coefficients for $NO_2$(ref) and $O_3$(Aq) and the intercepts for the two regression models are summarised in Figure 2.4. Apart from the second deployment period (P2), regression coefficients were significantly different between the two models. The coefficient for $NO_2$(ref) was relatively consistent among P2-P5 at around 0.8 for Fit2. The coefficients for $O_3$(Aq) varied to a larger extent between deployment periods for both models, which may be due to the sensor drift observed in section 2.2.3 Performance of the Aeroqual Series 500 $O_3$ monitor.

It is noted that in Fit2 P1 and P6 were associated with relatively low coefficients for $NO_2$(ref) and $O_3$(Aq) and relatively large intercepts, suggesting that Aeroqual $NO_2$ measurements during these periods were largely dominated by the monitor's baseline concentration. To investigative whether this was caused by building the model using the uncorrected (for baseline drift) $O_3$(Aq) measurements, $O_3$(ref) measurements were used instead of $O_3$(Aq) to calibrate $NO_2$(Aq). Although in practice it is unlikely to use reference $O_3$ measurements to calibrate Aeroqual $NO_2$ monitor at locations other than the reference monitoring station, for testing purpose, using $O_3$(ref) can largely eliminate uncertainties in the $O_3$ measurements, therefore allowing better understanding of the factors influencing the Aeroqual $NO_2$ measurements. Regression coefficients for the two regression models (Fit1 and Fit2) using $O_3$(ref) instead of $O_3$(Aq) are summarised in Figure 2.5. Significant variation in the coefficients for $NO_2$(ref) and $O_3$(ref) and intercept between deployment periods was retained for both models, suggesting that the Aeroqual $NO_2$ monitor in fact responded to $NO_2$ and $O_3$ concentrations to varying extents in different deployment periods.

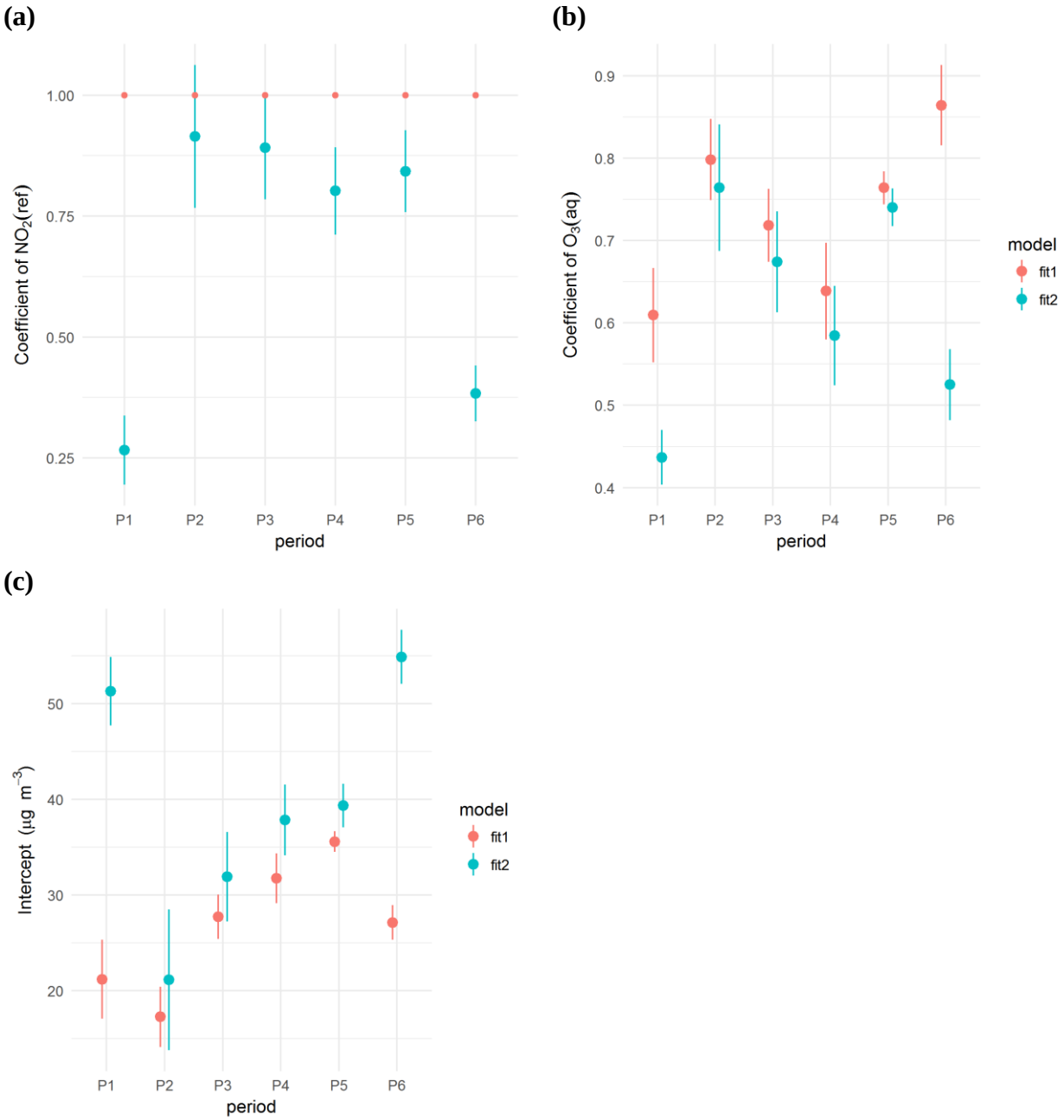



*Figure 2.4 Regression coefficients for (a) $NO_2$(ref), (b) $O_3$(aq) and (c) intercept calculated from the two regression models Fit1 and Fit2 given in Table 2.3. The error bar indicates the 95% confidence interval associated with the coefficient.*

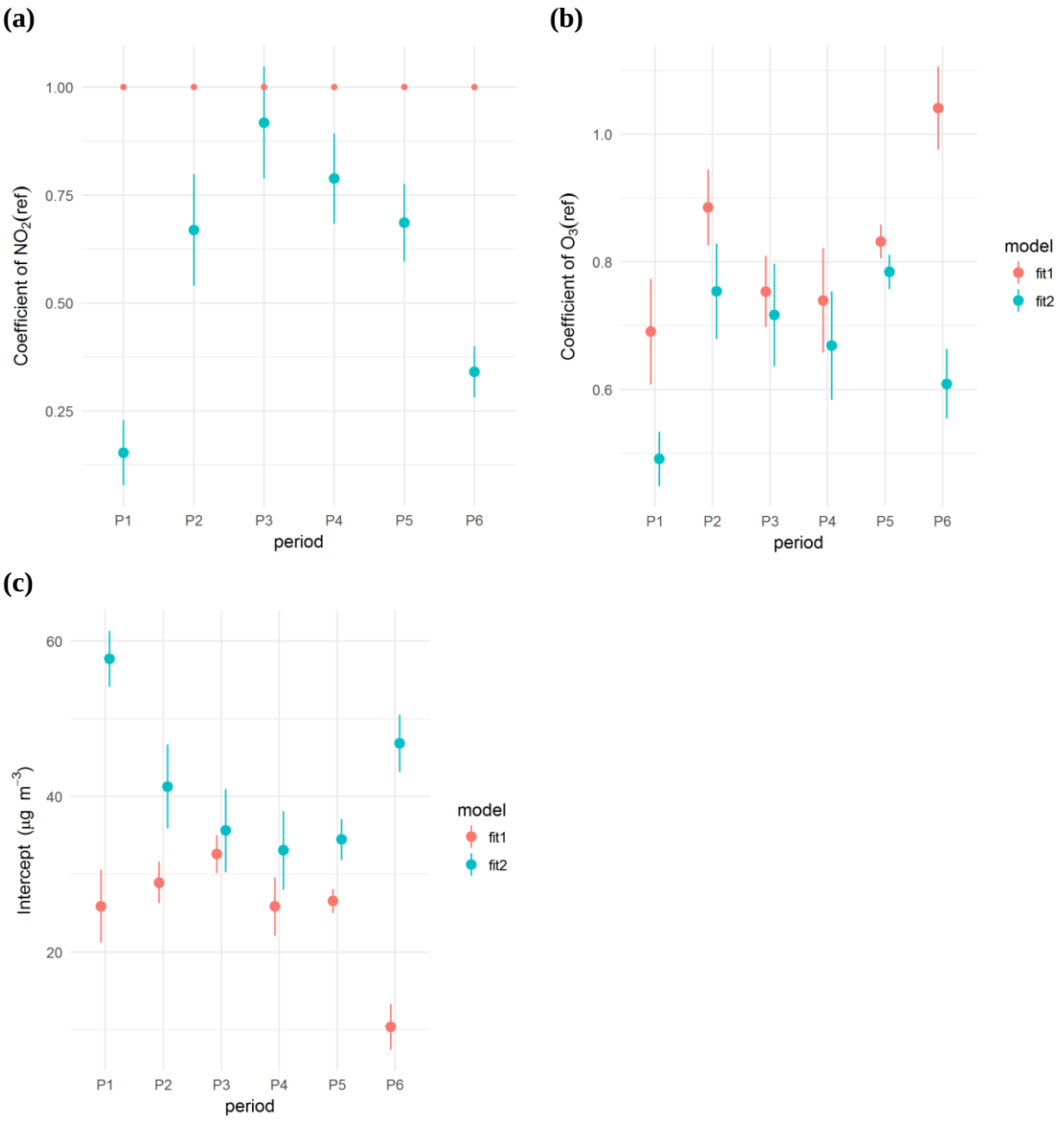



*Figure 2.5 Regression coefficients for (a) $NO_2$(ref), (b) $O_3$(ref) and (c) intercept calculated from the two regression models Fit1 and Fit2 given in Table 2.3. The error bar indicates the 95% confidence interval associated with the coefficient.*

It is evident in this study and the previous study (Lin et al., 2015) that the Aeroqual $NO_2$ monitors suffer from significant cross-interference from $O_3$. In practice the calibration of Aeroqual $NO_2$ measurements require simultaneous $O_3$ measurements (mostly likely from another Aeroqual $O_3$ monitor). The reliability of applying a calibration equation derived from one period to another depends not only on how

stable the response to $NO_2$ and $O_3$ is for the Aeroqual $NO_2$ monitor but also on how stable the Aeroqual $O_3$ monitor is. Although the response to $NO_2$ and $O_3$ from the Aeroqual $NO_2$ monitor was relatively consistent for most deployment periods (regression coefficients for Fit2 during periods P2-P5 in Figure 2.5), significant variations in the regression coefficients for $NO_2$ and $O_3$ were observed during other periods without apparent reasons. This implies that calibration of Aeroqual $NO_2$ monitors benefits from extended deployment periods in order to ascertain the stable calibration equation and that frequent calibrations before, after and even during a field measurement campaign are necessary.

### 2.2.5 Effect of temperature and *RH* on Aeroqual Series 500 $NO_2$ monitor

Temperature and relative humidity are known in principle to affect an electrochemical sensor's sensitivity and baseline (Mead et al., 2013). It was recently found that *T* and *RH* had a more pronounced effect on NO sensors than on $NO_2$ electrochemical sensors (Popoola et al., 2016). In this study the effect of *T* and *RH* on $NO_2$(Aq) in addition to $NO_2$(ref) and $O_3$(Aq) was investigated by adding *T* and *RH* to the linear regression (Table 2.4). One caveat of adding *RH* and *T* to the linear regression is that sometimes *RH* and *T* may correlate well with each other and with $O_3$, which makes it difficult to distinguish the independent effect of each variable. This multicollinearity issue in each model was examined with variance inflation factor (VIF). VIF describes the inflation of the confidence interval for a variable's regression coefficient relative to a model with uncorrelated variables. Generally a variable with VIF > 4 is considered to have a multicollinearity issue (Kabacoff, 2011). In this study *RH* was frequently found to have a multicollinearity issue with *T* (P1, 2 and 4). *T* was not found to have multicollinearity issue with $O_3$(Aq) or $NO_2$(ref) for all periods.

The two regression models also considering *T* and *RH* (Fit3 and Fit4) are compared with the model excluding *T* and *RH* (Fit2), and the effectiveness of the different models was evaluated using the Akaike Information Criterion (AIC) (Figure 2.6). The AIC index evaluates a model's statistical fit and the number of variables included to achieve this fit. Smaller AIC value of a model indicates an adequate fit

with fewer predictor variables and therefore is preferred (Kabacoff, 2011). AIC values suggest that Fit3 and Fit4 improved on the Fit2 only in P1 and P4 and by a very small amount (Figure 2.6). Considering the multicollinearity issue for *RH*, Fit3 might be the better model to correct $NO_2$(Aq). However the effect of *T* on $NO_2$(Aq) was inconsistent and weak if contribution of *T* is assumed to be linear to the response of $NO_2$(aq) as described in Fit3, consistent with the findings by Lin et al. (2015).

*Table 2.4 Linear regression models used to calibrate Aeroqual $NO_2$ monitor measurements including T or T and RH.*

| | |
|---|---|
| **Fit3** | $\mathrm{NO_2(Aq)} = a_1 \cdot \mathrm{NO_2(ref)} + a_2 \cdot \mathrm{O_3(Aq)} + a_3 \cdot T + b$ |
| **Fit4** | $\mathrm{NO_2(Aq)} = a_1 \cdot \mathrm{NO_2(ref)} + a_2 \cdot \mathrm{O_3(Aq)} + a_3 \cdot T + a_4 \cdot RH + b$ |

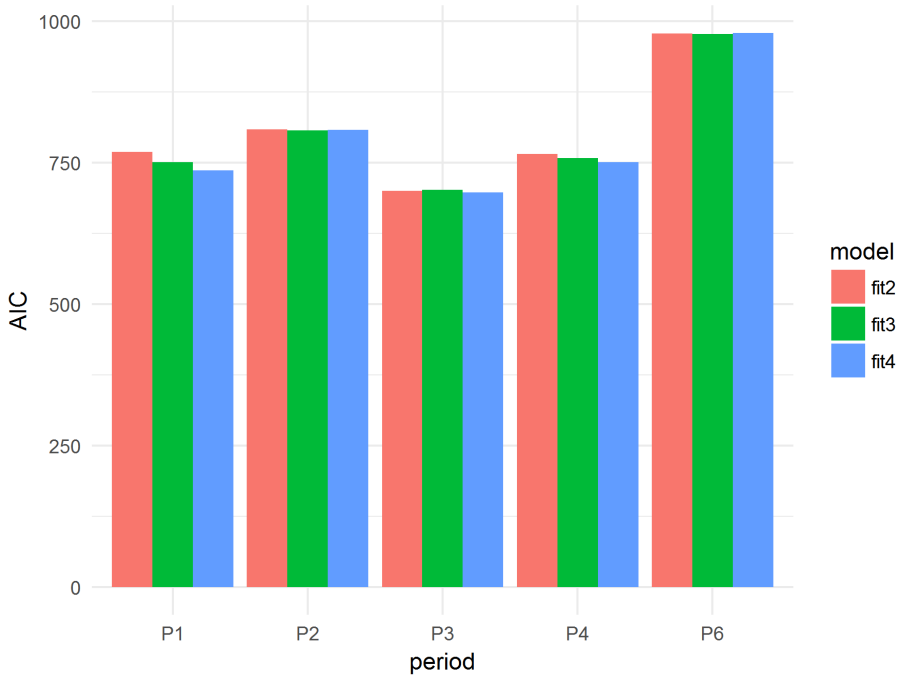



*Figure 2.6 AIC values of three regression models for calibrating $NO_2$(Aq) for different deployment periods. Data in P5 were not included in the analysis due to lack of T and RH measurements.*

## 2.3 Evaluation of RTI microPEM personal $PM_{2.5}$ monitors

### 2.3.1 Introduction

The microPEM (micro personal exposure monitor) monitor designed by RTI international (RTI, 2016) measures particles using the light scattering technique. Air samples are drawn by a miniaturised pump at a flowrate of 0.5 L per minute through dual impactor stages that select particles with aerodynamic diameter less than 2.5 μm before reaching the testing chamber. Light scattered by the particles passing through the laser beam shinning across the testing chamber is detected by a photodiode. The light intensity is converted into a voltage which is in turn proportional to the mass concentration of $PM_{2.5}$. The intensity of light scattered depends on particle size, the index of refraction and light-absorbing characteristics of the particles (TSI, 2012). Without knowing the properties of the particle samples, the conversion between the voltage and $PM_{2.5}$ mass concentration relies on a constant conversion factor obtained from calibration with known mass concentration of test aerosol that has a wide size distribution and is intended to be representative of a wide variety of ambient aerosols. It is assumed that the wide range of particle sizes averages the effect of particle size dependence on the measured signal (TSI, 2012). However any given ambient particle mixture may have properties that differ from the test aerosols on which the conversion factor is based. This contributes intrinsic (and unquantifiable) uncertainty to the monitors measuring PM mass concentration based on light scattering technique. In addition many studies have also found that relative humidity can affect the PM concentrations measured by the optical method compared to the gravimetric method (Fischer and Koshland, 2007; Soneja et al., 2014).

### 2.3.2 Method

Duplicate microPEM monitors (S/N: 586N and 618N) were deployed at St Leonards monitoring station in different seasons of 2014 and 2015 (Table 2.5). Data were recorded at 10-second intervals. In the subsequent analysis the 10-second data were first averaged to one minute and then to hourly data based on the one-minute data with a 75% data capture threshold per hour. The default $PM_{2.5}$ mass concentrations

output by the microPEM monitor has already been corrected for relative humidity using the method developed by (Chakrabarti et al., 2004). To investigate the effectiveness of this correction method, uncorrected $PM_{2.5}$ data and *RH* data recorded internally by the microPEM monitor were also extracted and analysed. In the 2015 deployment, a weather monitor (Kestrel 4500) was deployed on-site with the microPEM monitors to record ambient temperature and *RH*. Correction of the microPEM $PM_{2.5}$ data using *RH* recorded by the Kestrel was explored to investigate whether the precision and accuracy of the $PM_{2.5}$ data could be improved.

It is recommended in the operation manual that the nephelometer in the microPEM monitor should be zeroed frequently. This was done by adjusting the zero offset value for each microPEM monitor with a HEPA filter attached at the beginning of each deployment. Although the reading of the nephelometer stayed at zero for a few minutes when the HEPA filter was attached, it was discovered that in some instances the nephelometer zero reading drifted by ~3 μg $m^{-3}$ after the deployment, suggesting that such fluctuation in baseline might also occur during the measurements. As a result the zero offset values for each microPEM monitor were not consistent between the deployment periods (Table 2.5). Considering the zero offset values from the past several deployments, it appears that offset values of 47 and 56 for the 586N and 618N units were mostly used, suggesting that these are the stable settings for each microPEM monitor and should be used throughout.

*Table 2.5 Summary of the deployment periods and baseline settings of the duplicate microPEM monitors (586N and 618N) at Edinburgh St Leonards AURN station.*

| Deployment period | | Start/End date | Zero offset (586N, 618N) |
|---|---|---|---|
| **Summer 2014** | **P1.1** | 2014-07-30 ~ 2014-08-01 | (47, 56) |
| | **P1.2** | 2014-08-08 ~ 2014-08-14 | (47, 56) |
| | **P1.3** | 2014-08-25 ~ 2017-09-09 | (46, 56) |
| **Spring 2015** | **P2.1** | 2015-03-11 ~ 2015-03-16 | (47, 56) |
| | **P2.2** | 2015-03-23 ~ 2015-03-26 | (47, 57) |
| **Summer 2015** | **P3.1** | 2015-07-15 ~ 2015-07-21 | (47, 56) |
| | **P3.2** | 2015-07-21 ~ 2015-08-03 | (47, 56) |
| | **P3.3** | 2015-08-03 ~ 2015-08-27 | (47, 56) |

### 2.3.3 Precision of the microPEM monitors

The precisions of the three outputs (*RH*-corrected $PM_{2.5}$, *RH*-uncorrected $PM_{2.5}$ and *RH*) between duplicate microPEM monitors were assessed by MA regression analysis. The results for the 2014 summer deployment are shown in Figure 2.7. The correlation coefficients for all three parameters were high ($r^2 > 0.99$) during each measurement period. However the gradient from the regression analysis for both *RH*-corrected and *RH*-uncorrected $PM_{2.5}$ was significantly different from unity. The gradient of the *RH*-corrected $PM_{2.5}$ for all three periods was similar, but the intercept for the period P1.3 was roughly 3 μg $m^{-3}$ lower than the intercept for periods P1.1 and P1.2, coinciding with the amount of drift in baseline that was observed in zeroing. The cause of the difference in intercept was likely use of an incorrect zero offset value for microPEM 586N during P1.3.

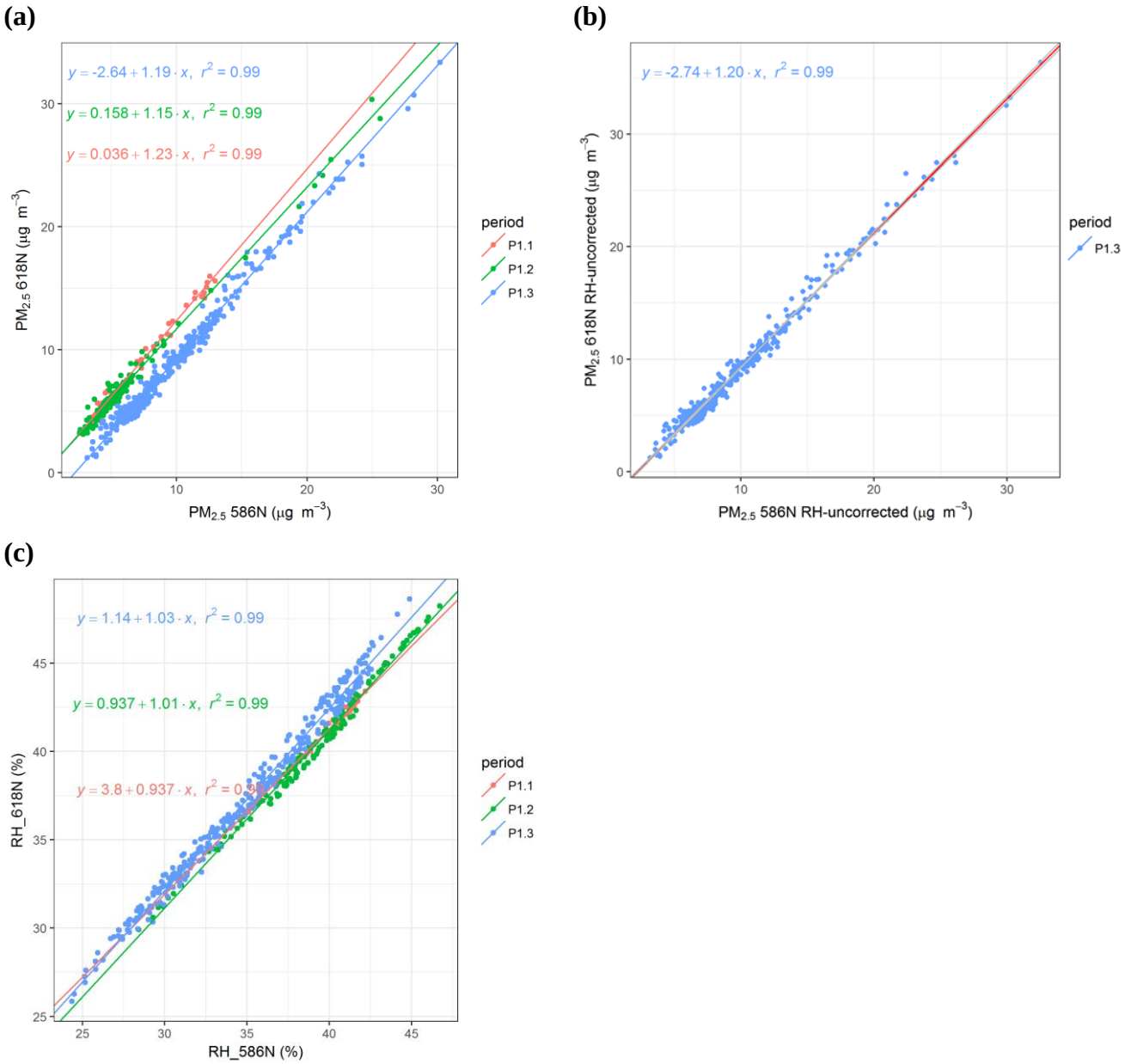



*Figure 2.7 Scatter plots for the summer 2014 measurements of the hourly average (a) RH-corrected $PM_{2.5}$, (b) RH-uncorrected $PM_{2.5}$ and (c) RH recorded by the duplicate microPEM monitors. Due to the loss of raw measurement files for periods P1.1 and P1.2, it is not possible to extract the RH-uncorrected $PM_{2.5}$ data for these two periods. 95% confidence interval of the regression coefficients are summarised in Appendix II Table 2.*

Similar to the deployment in 2014, correlations between duplicate microPEM monitors were high for *RH*-corrected $PM_{2.5}$, *RH*-uncorrected $PM_{2.5}$ and *RH* measurements in spring 2015 (Figure 2.8). However the relationships between duplicate *RH* and uncorrected $PM_{2.5}$ measurements were noticeably different for the two separate periods. Interestingly the relationship between duplicated *RH*-corrected $PM_{2.5}$ measurement was similar between the two periods. Applying the *RH* correction method on the uncorrected PM data using the *RH* recorded by the Kestrel weather monitor should eliminate any errors due to inconsistent response of the microPEM internal *RH* sensors (Figure 2.8d). The *RH* recorded by the Kestrel 4500 (min: 44%; mean: 82%; max: 100%) had a wider range and was generally higher than the *RH* recorded by the microPEM monitors (min: 27%; mean: 51%; max: 68%). The resulting corrected $PM_{2.5}$ concentrations based on *RH* data from the Kestrel weather monitor were therefore generally lower (Figure 2.8d).

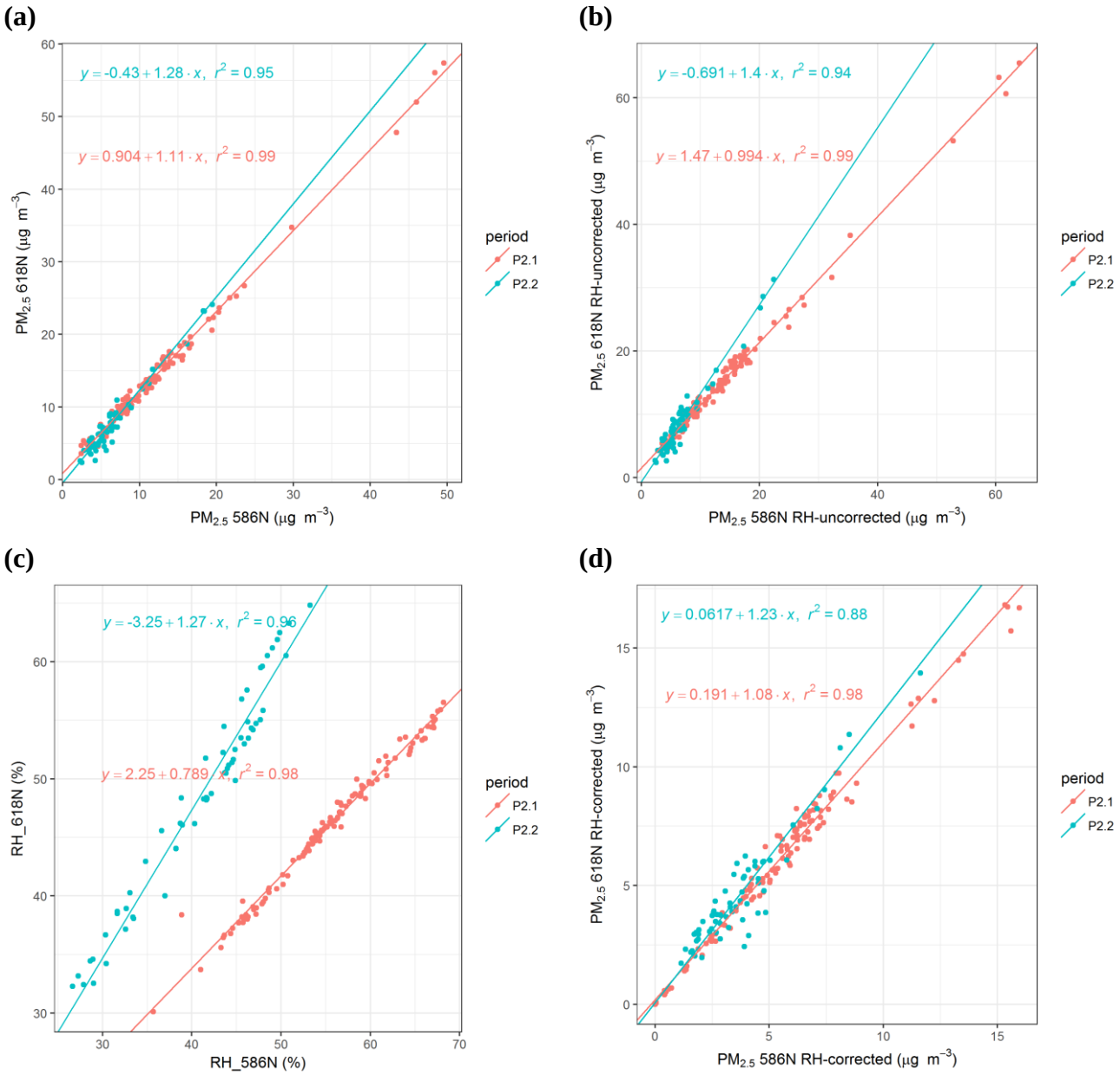



*Figure 2.8 Scatter plots for the spring 2015 measurements of the (a) RH-corrected $PM_{2.5}$, (b) RH-uncorrected $PM_{2.5}$, (c) RH and (d) RH-corrected $PM_{2.5}$ based on RH data from the Kestrel weather monitor. 95% confidence interval of the regression coefficients are summarised in Appendix II Table 2.*

In the summer 2015 deployment, *RH* recorded by the duplicate microPEM monitors again showed varying relationships between different periods (Figure 2.9c). The variation of relationship was also evident even within the same continuous period (P3.2 and P3.3). This resulted in reduced $r^2$ values of *RH*-corrected $PM_{2.5}$ compared to uncorrected $PM_{2.5}$ for P3.2 and P3.3. Despite using the *RH* data, correction of $PM_{2.5}$ concentrations using *RH* measurements from the Kestrel weather monitor did

not consistently improve the precision of the microPEM monitors (Figure 2.9d), suggesting that the correction method used as default in the microPEM monitors may not be appropriate in this setting. Both 2015 deployments showed that although different responses between duplicate microPEM measurements of *RH* and uncorrected $PM_{2.5}$ were evident, the corrected $PM_{2.5}$ measured by microPEM monitors was more consistent between deployment periods.

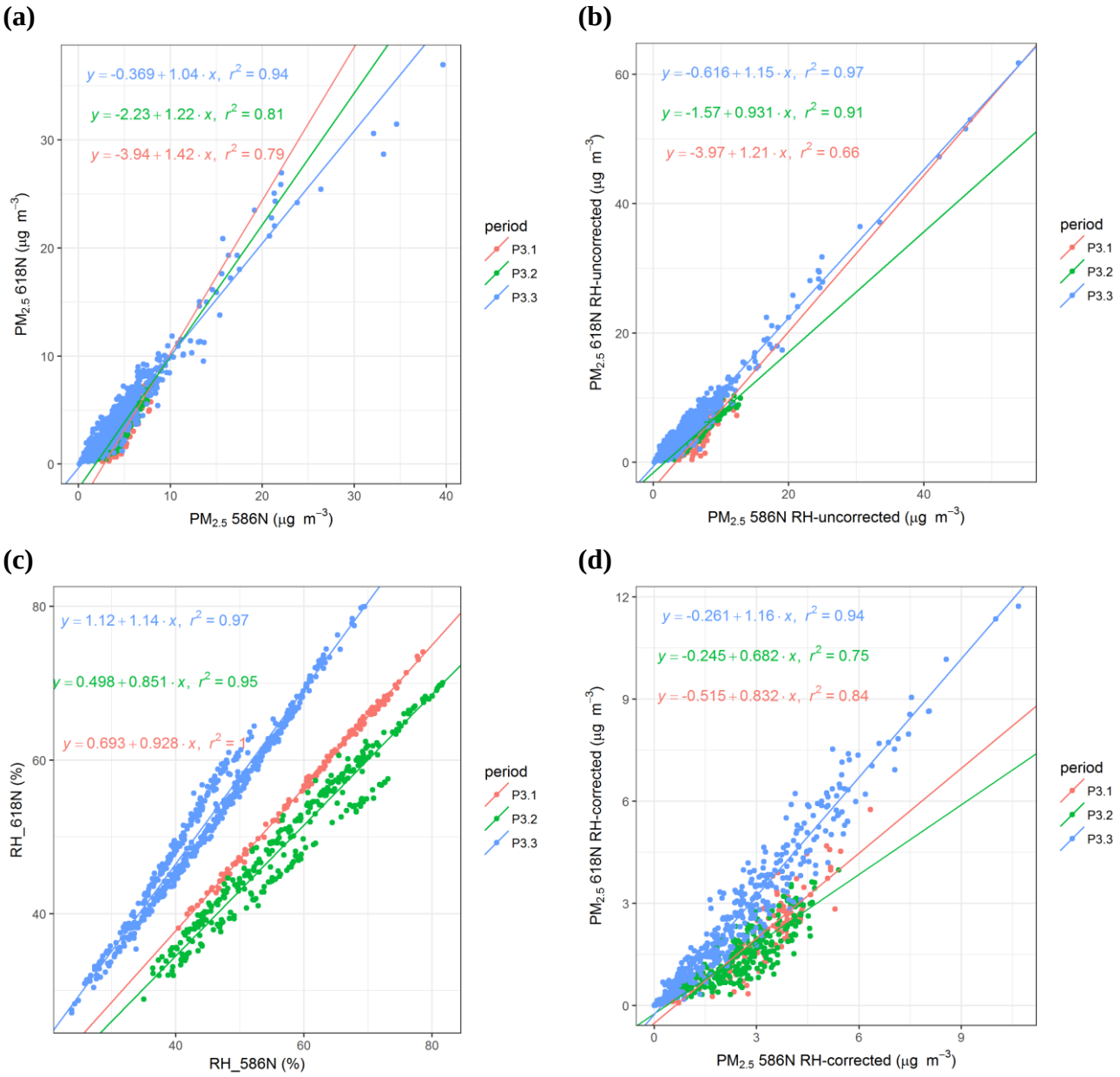



*Figure 2.9 Scatter plots for the summer 2015 measurements of the (a) RH-corrected $PM_{2.5}$, (b) RH-uncorrected $PM_{2.5}$, (c) RH and (d) RH-corrected $PM_{2.5}$ based on RH data from the Kestrel weather monitor. 95% confidence interval of the regression coefficients are summarised in Appendix II Table 2.*

### 2.3.4 Accuracy of the microPEM monitors

Since the comparisons between duplicate microPEM monitors have shown good correlations in both *RH*-corrected and uncorrected $PM_{2.5}$ values, the evaluation of the accuracy focuses only on the comparison between 586N and TEOM-FDMS. The total $PM_{2.5}$ mass concentration measured by TEOM-FDMS is used for comparison, which is the sum of volatile and non-volatile components of $PM_{2.5}$ that are measured independently during two separate cycles (AQEG, 2012). Figure 2.10 shows the scatter plots of *RH*-corrected and uncorrected $PM_{2.5}$ against TEOM-FDMS measurements for each deployment period. Correlations of TEOM-FDMS measurements with *RH*-corrected and uncorrected microPEM measurements were similar. Good correlations between microPEM and TEOM-FDMS measurements ($R^2 > 0.6$) were observed in most periods with the exception of P1.2, P3.1 and P3.2 where the correlation was poor ($r^2 < 0.1$). Comparing the $PM_{2.5}$ observed in different periods showed that during these periods average $PM_{2.5}$ was only around 5 μg $m^{-3}$ and the range of variation was small (Figure 2.11). At such low concentration the noise in the microPEM values (e.g. fluctuation of the baseline) would have more impact on the measurement data therefore affecting the accuracy. In addition the TEOM-FDMS measurements of such low $PM_{2.5}$ concentrations may also be subject to greater uncertainties (AQEG, 2012).

Table 2.6 shows the statistics of the comparison of *RH*-corrected and uncorrected MicroPEM $PM_{2.5}$ measurements with TEOM-FDMS measurements. In the periods when there was significant correlation between microPEM and TEOM-FDMS measurements, the mean bias (MB) and normalised mean bias (NMB) for the *RH*-corrected $PM_{2.5}$ measurements were not consistently lower than the uncorrected measurements. However, in some instances (P2.1, P3.1 and P3.2) the uncorrected data showed almost twice the error of the *RH*-corrected data. The *RH*-corrected measurements also showed consistent improvement in the root mean squared error (RMSE). Without calibration the *RH*-corrected microPEM measurements were on average between −20% – 47% of the TEOM-FDMS for the periods when $PM_{2.5}$ concentrations were low (P1.2, P3.1 and P3.2) and between −11% – 43% for the rest of the periods.

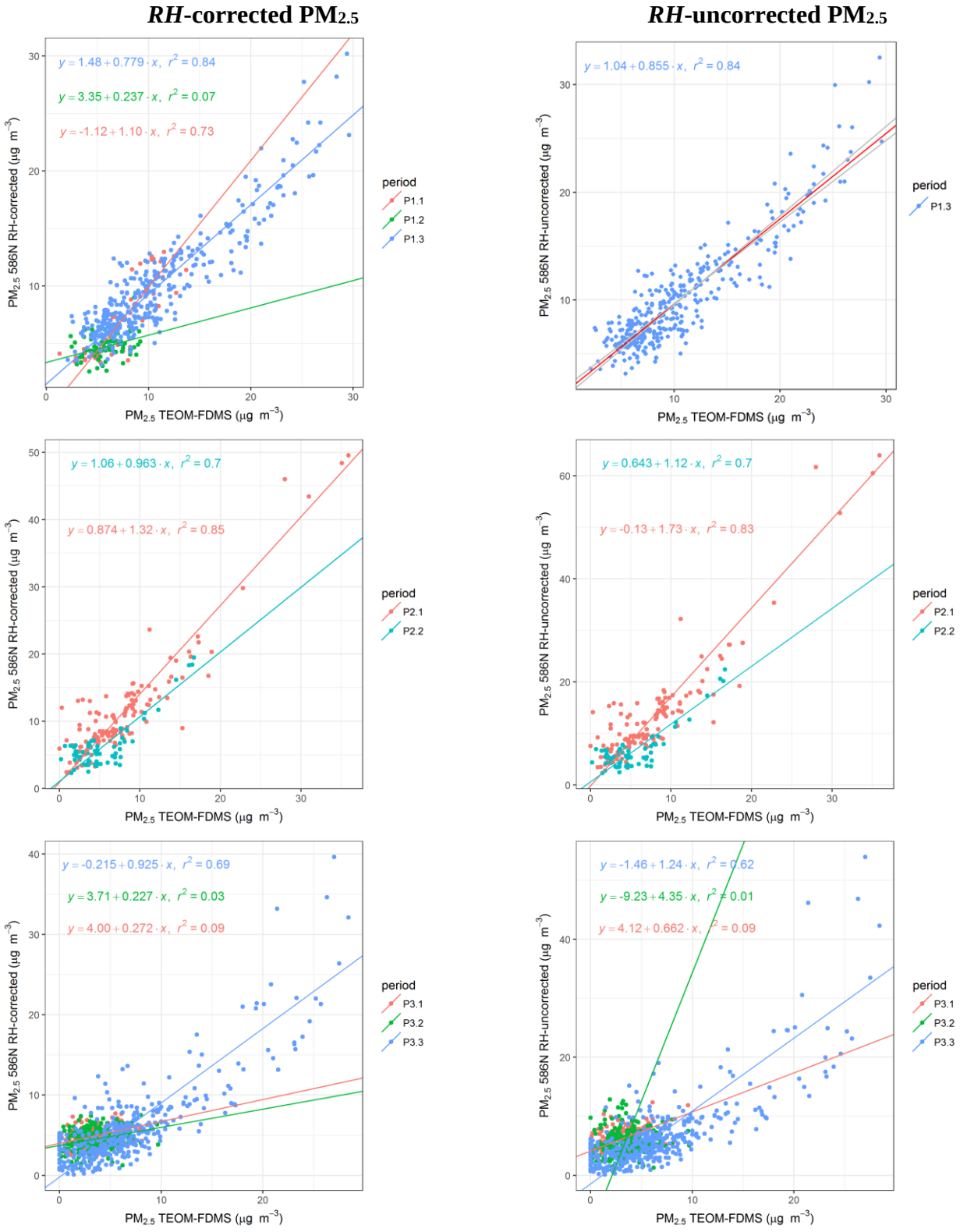



*Figure 2.10 Scatter plots for each measurement period of RH-corrected and uncorrected $PM_{2.5}$ measured by microPEM 586N against $PM_{2.5}$ measured by TEOM-FDMS. 95% confidence interval of the regression coefficients are summarised in Appendix II Table 3.*

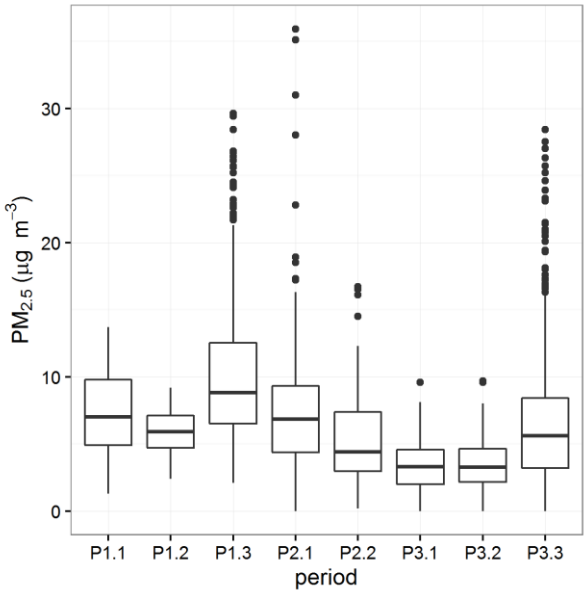



*Figure 2.11 Boxplot of the hourly $PM_{2.5}$ concentrations measured by TEOM-FDMS during the different deployment periods. The whiskers extend to the values that are within 1.5 times the IQR on each side of the median.*

*Table 2.6 Statistics for the RH corrected (uncorrected) microPEM measurements against TEOM-FDMS measurements. The grey shading highlights the periods when there was no correlation between microPEM and TEOM-FDMS measurements. The definitions of all the evaluation statistics are included in Appendix I.*

| Period | FAC2 | MB (µg $m^{-3}$) | NMB | RMSE (µg $m^{-3}$) | $r^2$ |
|---|---|---|---|---|---|
| **P1.1** | 0.96 (NA) | -0.37 (NA) | -0.05 (NA) | 1.76 (NA) | 0.73 (NA) |
| **P1.2** | 0.93 (NA) | -1.15 (NA) | -0.2 (NA) | 2.08 (NA) | 0.07 (NA) |
| **P1.3** | 0.99 (0.99) | -0.85 (-0.5) | -0.08 (-0.05) | 2.54 (2.35) | 0.84 (0.84) |
| **P2.1** | 0.81 (0.73) | 3.45 (5.75) | 0.43 (0.71) | 4.81 (7.78) | 0.85 (0.83) |
| **P2.2** | 0.82 (0.82) | 0.85 (1.30) | 0.16 (0.24) | 2.29 (2.62) | 0.70 (0.70) |
| **P3.1** | 0.68 (0.51) | 1.57 (2.99) | 0.47 (0.90) | 2.48 (3.68) | 0.09 (0.09) |
| **P3.2** | 0.69 (0.58) | 1.05 (2.33) | 0.30 (0.68) | 2.27 (3.55) | 0.03 (0.01) |
| **P3.3** | 0.74 (0.73) | -0.72 (0.12) | -0.11 (0.02) | 3.21 (4.00) | 0.69 (0.62) |

### 2.3.5 Investigation of the *RH* dependence of the microPEM monitor

The default $PM_{2.5}$ output from the microPEM monitor is corrected for the potential issue that high *RH* may cause the nephelometer to overestimate the particle concentration. The precision between duplicate microPEM monitors for *RH*-corrected $PM_{2.5}$ therefore also depends on the precision of the *RH* sensors between the two units. In several deployment periods in 2015 there was noticeable difference in the response of the *RH* sensors in the two units. However, the *RH*-corrected $PM_{2.5}$ concentrations appeared not to be significantly affected by the inconsistency of *RH* measurements, possibly because the *RH* recorded by microPEM monitors was generally lower and did not significantly influence the corrected values. Although in theory the uncorrected data should be more reliable in terms of precision, the analysis did not show a significant improvement in the precision of *RH*-uncorrected data over the *RH*-corrected data. The algorithm applied by the microPEM monitor to correct potential overestimation at high *RH* was investigated by plotting the ratio of uncorrected $PM_{2.5}$ to reference $PM_{2.5}$ against *RH* (Figure 2.12). No clear pattern was found in the bias of uncorrected $PM_{2.5}$ as a function of *RH* recorded either by the Kestrel weather monitor or the microPEM monitor. The *RH*-correction method adopted by microPEM monitors did not account well for the overestimation of the uncorrected $PM_{2.5}$. This explains the findings that the *RH*-correction did not significantly improve the correlation between *RH*-corrected and TEOM-FDMS measurements compared to the uncorrected data (Table 2.6). However, there was evidence that *RH*-corrected data reduced the error (RMSE) comparing to the TEOM-FDMS measurements and also effectively reduced the occasional large overestimations if no correction was made. The *RH*-corrected microPEM measurement also requires calibration due to the occasional large NMB observed (e.g. P2.1). However, the varying regression coefficients between the deployments mean that erroneously applying the calibration equation derived from one period to another may introduce greater error than the un-calibrated data (Figure 2.13). Given that in many occasions the slope of regression analysis between microPEM and TEOM-FDMS measurements was not significantly different from unity and the intercept was only around 1 $\mu g\ m^{-3}$, it is plausible to use the un-calibrated microPEM

measurement if the uncertainties (MB, NMB and RMSE) compared to reference instrument are also clearly stated.

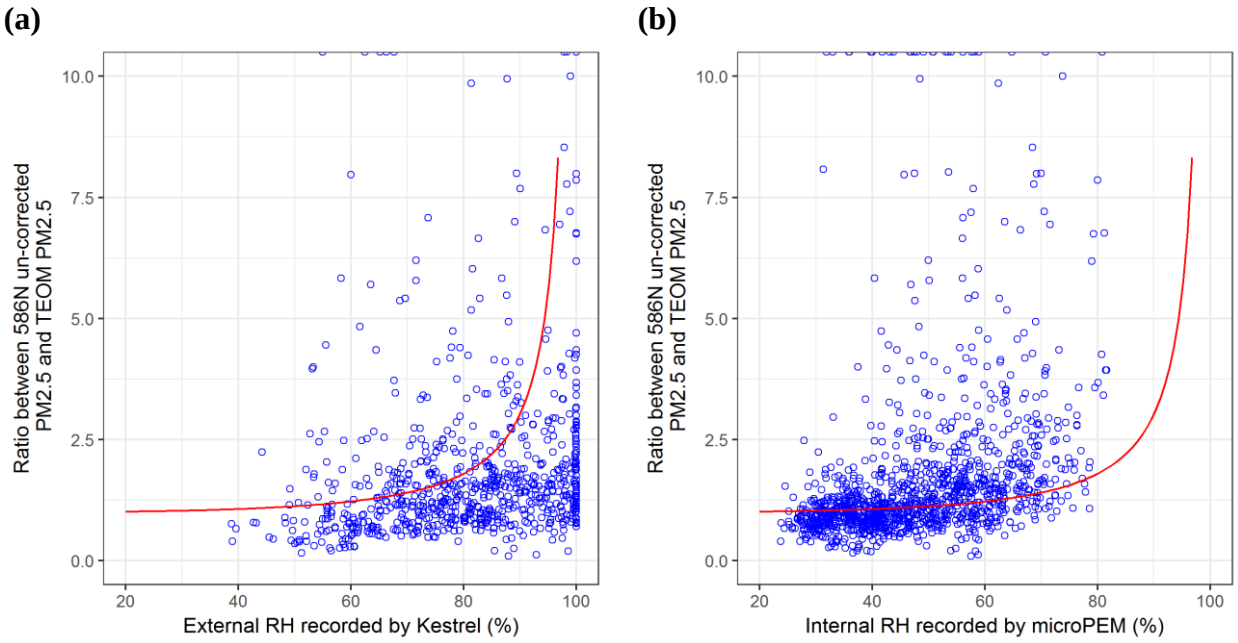



*Figure 2.12 Ratio of uncorrected $PM_{2.5}$ measured by microPEM 586N to reference $PM_{2.5}$ measurement as a function of RH recorded by (a) the Kestrel weather monitor and (b) the microPEM monitor for all the deployment periods. The red curve shows the RH-correction function used by the microPEM monitor.*

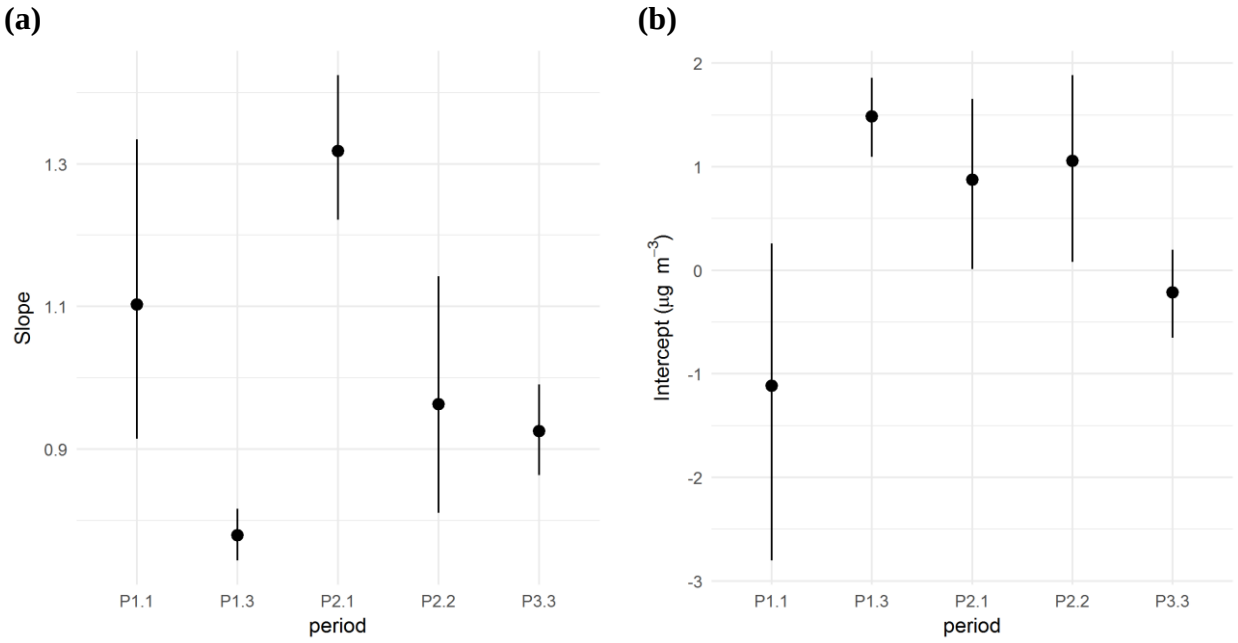



*Figure 2.13 Slope and intercept coefficients (with 95% confidence interval indicated by error bar) of the MA regression between RH-corrected microPEM 586N measurements and TEOM-FDMS measurements for the periods when there was significant correlation between the two.*

## 2.4 Discussion

The performances of Aeroqual $O_3$ and $NO_2$ and microPEM $PM_{2.5}$ monitors were evaluated against respective reference instruments through a series of co-location deployments spanning a few months. The Aeroqual $O_3$ monitor tested here showed very good correlation ($r^2 > 0.90$) with the reference instrument. This compares favourably than the results published in AQ-SPEC ($r^2 \sim 0.85$) and is generally higher than other portable $O_3$ monitors (http://www.aqmd.gov/aq-spec/evaluations/summary). Yet there were biases in the slope and intercept coefficients that needed correction. The baseline of the Aeroqual $O_3$ changed between different deployment periods, suggesting that the correction equation derived in the calibration period may not be suitable for a separate measurement period. Therefore, calibration between Aeroqual $O_3$ monitor and reference instrument before and after the measurement period is essential to determine if baseline drift occurs. Consistent with previous findings, the Aeroqual $NO_2$ monitor is subject to

cross-sensitivity from $O_3$. Temperature and relative humidity were not found to have significant impacts on the performance of Aeroqual $NO_2$ and $O_3$ monitors. Aeroqual $NO_2$ measurement can be effectively calibrated using reference $NO_2$ and Aeroqual $O_3$ measurements. However, in practice the calibrated $NO_2$ data may be associated with greater uncertainties contributed by the instability of both $O_3$ and $NO_2$ monitors. The Aeroqual $O_3$ monitor can provide very comparable measurement with reference analyser if frequent calibration and quality control processes are carried out. Aeroqual $NO_2$ monitor measurements have to be accompanied by contemporaneous $O_3$ measurements to allow calibration of raw Aeroqual $NO_2$ data.

The analysis of the inter-comparison between duplicate microPEM monitors uncovered a few practical and technical considerations for its field application. First, an optimal zero offset setting for each microPEM monitor should be determined before the measurements. Although the baseline might fluctuate between deployments, it is advantageous to keep the same setting in order to ensure comparability of measurements in different periods. The response of the internal *RH* sensor in duplicate microPEM monitors was inconsistent between deployments. However, since the absolute values of *RH* were low (< 80%) in both microPEM monitors, the resulting *RH*-corrected data reported by duplicate microPEM monitors were not significantly affected and were relatively consistent between different periods. Correlation between microPEM monitor and reference instrument was good ($r^2$ = 0.69 – 0.85) compared to other monitors tested in AQ-SPEC in most testing periods, but was poor ($r^2$ < 0.1) when the concentration was low (< 10 μg $m^{-3}$) and the range of concentration was small. This could be due to increased uncertainties in both microPEM and TEOM-FDMS measurements. Calibration of microPEM monitor against reference analyser should cover the range of concentrations that might be encountered in the actual measurements whenever possible. The *RH*-corrected $PM_{2.5}$ measurements slightly improved on the un-corrected data in terms of RMSE and NMB. Although occasionally large overestimation (43%) was observed in microPEM versus TEOM-FDMS measurements, the measurement error (NMB) was small (–11% ~ 16%) for the rest of comparisons (excluding the periods when $PM_{2.5}$ concentrations were less than 10 μg $m^{-3}$). Caution should be taken when using the microPEM monitor to measure $PM_{2.5}$ at very low concentrations due to the poor

correlation with TEOM-FDMS and greater measurement error. Calibration of microPEM monitor against reference analyser should cover the range of concentrations that might be encountered in the actual measurements whenever possible. The *RH*-corrected $PM_{2.5}$ measurements slightly improved on the un-corrected data in terms of RMSE and NMB. Although occasionally large overestimation (43%) was observed in microPEM versus TEOM-FDMS measurements, the measurement error (NMB) was small (−11% ~ 16%) for the rest of comparisons (excluding the periods when $PM_{2.5}$ concentrations were less than 10 μg $m^{-3}$). Caution should be taken when using the microPEM monitor to measure $PM_{2.5}$ at very low concentrations due to the poor correlation with TEOM-FDMS and greater measurement error

It is acknowledged that the small number of each type of monitors tested here has limitation on the generalisability of the results. However, many findings in this chapter, such as the instability of the monitor's regression coefficient and the cross-interference to $O_3$ concentration for Aeroqual $NO_2$ monitor, are consistent with other studies (Masey et al., 2017; Lin et al., 2016) conducted in different locations and seasons, suggesting that these are common issues for the Aeroqual monitors. The choice of a single St. Leonard background monitoring site as the location of evaluation may fail to test the monitors' performance at relatively high pollution level. Exposure studies intending to use these monitors to quantify human exposure in commuting therefore requires further evaluation of these monitors at a roadside monitoring station. The evaluation of Aeroqual monitors in this chapter was only conducted in the summer period. Seasonality has a bigger impact on the $O_3$ concentration, where higher $O_3$ concentrations are expected in the summer than in winter. A wide range of $O_3$ concentration was observed during this evaluation, and therefore the effect of seasonality should have minimal effect on the monitor's performance in different concentration ranges. However, the low ambient temperature during winter may also affect the response of the Aeroqual $O_3$ monitors, which is more difficult to quantify given the instability of monitor's response with time.

To determine whether an air quality monitor can be accepted as an indicative method requires the relative expanded uncertainty of the method (both from laboratory and field testing) calculated in the way specified by EC (2010) to be within the acceptable uncertainty set in the DQOs, which are 30% for $O_3$, 25% for $NO_2$ and 50% for $PM_{2.5}$ at their respective limit or target value. The protocol drafted by Spinelle et al. (2013) for evaluation of gas sensors also detailed a series of steps leading up to the compliance testing with the DQO, one of which is to demonstrate that the slope and intercept coefficients from the orthogonal regression analysis between the monitor and reference method are not significantly different from 1 and 0, respectively. In many cases the comparison of Aeroqual $O_3$ and microPEM monitors with their respective reference methods showed significant biases in slope and intercept coefficients. Although appropriate calibration of the data from the monitors may resolve this issue, the inconsistent calibration equations derived from different periods complicates the choice of the correct calibration equation to use for a certain measurement period. For this reason, none of the monitors evaluated here can be accepted as an indicative method. However, their general high correlation with the reference methods and their portability show usefulness in complementing the existing air quality monitoring framework.

## 2.5 Conclusions

In general, the three types of battery-operated portable monitors investigated here are relatively easy to use with minimal maintenance required. However, there are many caveats limiting the immediate usability of the data recorded. The time and resources required for finding a stable parameter in the monitors' settings, post-processing of the monitors' raw data and the frequent calibration with reference instruments are not trivial undertakings and also offset some of their "low-cost" advantage against more expensive reference instruments. It should be recognised that any calibration exercise carried out in the field by comparison to reference measurements is valid only under some specific conditions such as a particular pollution mix, a particular range of environment variables and a particular setting for the monitor. It is impossible to evaluate and calibrate a portable monitor with respect to the reference measurement whilst on the move. Therefore, it should not be expected that the portable monitors

can provide data to the same level of quality as provided by the reference instrument. The portable monitors investigated here cannot be used as an equivalent or even indicative method to a reference measurement in circumstances where absolute quantification is needed, hence they cannot replace the existing monitoring network. Nevertheless, good correlation with reference instrument can reassure the confidence in relative trends within a set of mobile measurements made by the portable monitors, which is demonstrated to be useful in Chapter 3 to study the spatial variability of air pollutants in urban environment using short-term mobile measurements.

# References

Aeroqual, 2016. Aeroqual-Portable-Fixed-Gas-Sensor-Specifications-V6.pdf. URL http://www.aeroqual.com/wp-content/uploads/Aeroqual-Portable-Fixed-Gas-Sensor-Specifications-V6.pdf (accessed 28-March-2017).

AQEG, 2012. Fine Particulate Matter (PM2.5) in the United Kingdom. Defra, UK. URL http://uk-air.defra.gov.uk/library/reports?report_id=727

Chakrabarti, B., Fine, P.M., Delfino, R., Sioutas, C., 2004. Performance evaluation of the active-flow personal DataRAM PM2.5 mass monitor (Thermo Anderson pDR-1200) designed for continuous personal exposure measurements. Atmos. Environ. 38, 3329–3340. doi:10.1016/j.atmosenv.2004.03.007

EC, 2010. Guide to the demonstration of equivalence of ambient air monitoring methods. European Commission. URL http://ec.europa.eu/environment/air/quality/legislation/pdf/equivalence.pdf

EC, 2008. DIRECTIVE 2008/50/EC OF The European Parliament And Of The Council of 21 May 2008 On Ambient Air Quality And Cleaner Air For Europe. European Commission. URL http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008L0050&qid=1471948031872&from=en

Fischer, S.L., Koshland, C.P., 2007. Field performance of a nephelometer in rural kitchens: effects of high humidity excursions and correlations to gravimetric analyses. J. Expo. Sci. Environ. Epidemiol. 17, 141–150. doi:10.1038/sj.jes.7500486

Heimann, I., Bright, V.B., McLeod, M.W., Mead, M.I., Popoola, O.A.M., Stewart, G.B., Jones, R.L., 2015. Source attribution of air pollution by spatial scale separation using high spatial density networks of low cost air quality sensors. Atmos. Environ. 113, 10–19. doi:10.1016/j.atmosenv.2015.04.057

Kabacoff, R., 2011. R in action: data analysis and graphics with R. Manning, Shelter Island, NY.

Lewis, A., Edwards, P., 2016. Validate personal air-pollution sensors. Nature 535, 29–31. doi:10.1038/535029a

Lin, C., Gillespie, J., Schuder, M.D., Duberstein, W., Beverland, I.J., Heal, M.R., 2015. Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide. Atmos. Environ. 100, 111–116. doi:10.1016/j.atmosenv.2014.11.002

Masey, N., Gillespie, J., Ezani, E., Lin, C., Wu, H., Ferguson, N., Hamilton, S., Heal, M.R., Beverland, I.J., 2017. Temporal changes in field calibration relationships for Aeroqual S500 O3 and NO2 sensors. (under review)

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. Atmos. Environ. 70, 186–203. doi:10.1016/j.atmosenv.2012.11.060

Popoola, O.A.M., Stewart, G.B., Mead, M.I., Jones, R.L., 2016. Development of a baseline-temperature correction methodology for electrochemical sensors and its implications for long-term stability. Atmos. Environ. 147, 330–343. doi:10.1016/j.atmosenv.2016.10.024

RTI, 2016. MicroPEM$^{TM}$ Sensor for Measuring Exposure to Air Pollution [WWW Document]. RTI. URL https://www.rti.org/impact/micropem-sensor-measuring-exposure-air-pollution (accessed 28-March-2017).

Samoli, E., Atkinson, R.W., Analitis, A., Fuller, G.W., Green, D.C., Mudway, I., Anderson, H.R., Kelly, F.J., 2016. Associations of short-term exposure to traffic-related air pollution with cardiovascular and respiratory hospital admissions in London, UK. Occup. Environ. Med. 73, 300–307. doi:10.1136/oemed-2015-103136

Snyder, E.G., Watkins, T.H., Solomon, P.A., Thoma, E.D., Williams, R.W., Hagler, G.S.W., Shelow, D., Hindin, D.A., Kilaru, V.J., Preuss, P.W., 2013. The Changing Paradigm of Air Pollution Monitoring. Environ. Sci. Technol. 47, 11369–11377. doi:10.1021/es4022602

Soneja, S., Chen, C., Tielsch, J.M., Katz, J., Zeger, S.L., Checkley, W., Curriero, F.C., Breysse, P.N., 2014. Humidity and Gravimetric Equivalency Adjustments for Nephelometer-Based Particulate Matter Measurements of Emissions from Solid Biomass Fuel Use in Cookstoves. Int. J. Environ. Res. Public. Health 11, 6400–6416. doi:10.3390/ijerph110606400

Spinelle, L., Aleixandre, M., Gerboles, M., 2013. Protocol of evaluation and calibration of low-cost gas sensors for the monitoring of air pollution. (No. EUR 26112 EN). European Commission, Luxembourg. URL http://dx.publications.europa.eu/10.2788/9916 (accessed 2-February-2017).

Steinle, S., Reis, S., Sabel, C.E., Semple, S., Twigg, M.M., Braban, C.F., Leeson, S.R., Heal, M.R., Harrison, D., Lin, C., Wu, H., 2015. Personal exposure monitoring of PM2.5 in indoor and outdoor microenvironments. Sci. Total Environ. 508, 383–394. doi:10.1016/j.scitotenv.2014.12.003

TSI, 2012. Dusttrak aerosol monitor theory of operation. URL http://www.tsi.com/uploadedFiles/_Site_Root/Products/Literature/Application_Notes/ITI-036.pdf

Williams, D.E., Salmond, J., Yung, Y.F., Akaji, J., Wright, B., Wilson, J., Henshaw, G.S., Wells, D.B., Ding, G., Wagner, J., Laing, G., 2009. Development of low-cost ozone and nitrogen dioxide measurement instruments suitable for use in an air quality monitoring network, in: 2009 IEEE Sensors. Presented at the 2009 IEEE Sensors, pp. 1099–1104. doi:10.1109/ICSENS.2009.5398568

# Chapter 3 Identifying drivers for the intra-urban variability of air pollutants and their interrelationships through mobile and fixed-site measurements

*Section 3.2 of this chapter 'Study of the variability and interrelationships of different PM metrics through mobile short-term measurements' is based on a research paper published in 'Atmospheric Environment' (Wu, H., Reis, S., Lin, C., Beverland, I.J., Heal, M.R., 2015. Identifying drivers for the intra-urban spatial variability of airborne particulate matter components and their interrelationships. Atmos. Environ. 112, 306–316. doi:10.1016/j.atmosenv.2015.04.059). I undertook the experiment, data analysis and drafting of the manuscript. Dr Mathew Heal, Dr Stefan Reis and Dr Iain Beverland gave advice on the methodology, presentation of results and manuscript editing. Dr Chun Lin helped with the field measurements and methodology of the experiment.*

## 3.1 Introduction

Evidence continues to accumulate of the adverse health impacts of $PM_{2.5}$, the mass concentration of airborne particulate matter (PM) with an aerodynamic diameter of less than 2.5 μm (WHO, 2013). However, other components of ambient PM, including ultrafine particles (UFP, particles of diameter <100 nm) and black carbon (BC) concentrations, are emerging as important in terms of their association with health effects (Heal et al., 2012). UFPs are usually quantified by particle number concentration (UFPNC), since smaller particles dominate the number concentration, while contributing only a minor fraction to particle mass concentrations. Lung deposited surface area (LDSA) is another metric to quantify UFP, given the high surface area to mass ratio for UFPs and its toxicological significance in term of its direct physical interaction with cells (Oberdörster et al., 2005).

High variability of air pollutant concentrations are shown in urban environments due to the dynamics of emissions and urban topographies (Van den Bossche et al., 2015). Street canyons, which are ubiquitous in many urban environments, introduce complex dispersion characteristics that increase the spatial variability of BC and UFP

(Peters et al., 2014; Rakowska et al., 2014). A relevant issue is the extent to which UFP and BC concentrations vary within populated areas, since a shortcoming in many epidemiological studies is assumption of homogenous exposure within the study area. This assumption might be plausible for pollutants with less spatial variability but could result in significant bias in exposure-response relationships for highly spatially variable pollutants (Hoek et al., 2002). In this context, variables related to the contribution of emissions from road sources (e.g. traffic intensity, or distance to the road) are commonly identified as significant predictors for a range of traffic-related air pollutants in many studies applying land-use regression models (Hoek et al., 2008), the validity of which might be influenced by the underlying causes of the variability of different pollutants. Thus one of the aims of this work was to evaluate the extent to which potential factors affect the spatiotemporal variability of ambient BC, UFP and $PM_{2.5}$ in an urban area. These factors include local traffic, street topography and synoptic meteorology which, although recognised in the literature, have rarely been compared in terms of their influences on different metrics.

Given the increasing evidence for the harmfulness of UFP (WHO, 2013) with its ability to penetrate deep into the airways (Knibbs et al., 2011), investigation on the relationship between UFP and other regulated pollutants (e.g. $PM_{2.5}$ and $NO_2$) can provide insight on the extent to which current policy can effectively protect human health. Understanding the relationship between UFPs and other pollutants is also critical for the assessment of their independent effects on human health. Another of the aims of this work was to investigate the inter-relationships between the different metrics of PM and $NO_2$.

Traditional fixed-site monitoring stations are rarely sufficient in number to enable detailed study of intra-urban variability of air pollutants. One way to monitor pollutant concentrations at high spatial resolution is by use of portable monitoring instruments, which also has good prospects for wider application in the assessment of human exposure to air pollution in the future (Steinle et al., 2013).

In this chapter, analyses of two separate sets of measurement data are presented. In the first set of analyses, pairs of portable instruments were used to measure $PNC_{0.5\text{-}2.5}$

(used here as a measure of $PM_{2.5}$), UFPNC and BC concentrations within the city of Edinburgh (Scotland) in two series of measurement campaigns in winter and in spring. Analyses of data from a combination of mobile and stationary measurements were used to evaluate possible causes of the spatial and temporal variations in the concentrations of the different PM metrics. In the second set of analyses, continuous stationary measurements of UFPNC, LDSA and mean particle size were taken at background and roadside sites. Their relationships with $NO_2$, $PM_{2.5}$ and inorganic components of $PM_{2.5}$ were investigated to understand different drivers for their variation.

## 3.2 Study of the variability and interrelationships of different PM metrics through mobile short-term measurements

### 3.2.1 Methods

#### 3.2.1.1 Study design

BC, UFPNC and $PNC_{0.5\text{-}2.5}$ concentrations were measured across the south of the city of Edinburgh, UK (55.9° N, 3.2° W, population ~480,000) in two separate campaigns using duplicate units of the following instruments: microAeth AE51 (AE51), TSI 3007 Condensation Particle Counter (CPC) and Dylos Corp. DC1700 (Dylos).

In the winter campaign, between December 2013 and January 2014, the measurements were conducted three times on Mondays and once on Sunday primarily near roads by walking between and pausing at designated sites (Figure 3.1a). The sites were selected to cover potential hotspots, urban background and different street topographies (open or built-up) over an area of about 6 $km^2$. In a typical walk, the measurements started at around 10:00 a.m. and proceeded through the designated sites to the final location roughly in the order from south to north and from east to west (Figure 3.1a). At each site a 5-min static measurement was conducted, during which the number and type of vehicles (car, van, heavy goods vehicle and bus) passing the observer were recorded. Throughout each walk, measurements for each pollutant were taken in parallel (with duplicate instruments)

on both sides of the road. To evaluate the duplicate precision, inter-comparison between the pairs of instruments was conducted in a separate trial on Mon 3rd Feb 2014 from 10:00 a.m. to 12:00 p.m. by walking through the same route but with the duplicate instruments carried by one person. The weather conditions on this day were similar to other measurement days (Table 3.1). Traffic characteristics are assumed to be similar to other Monday measurements.

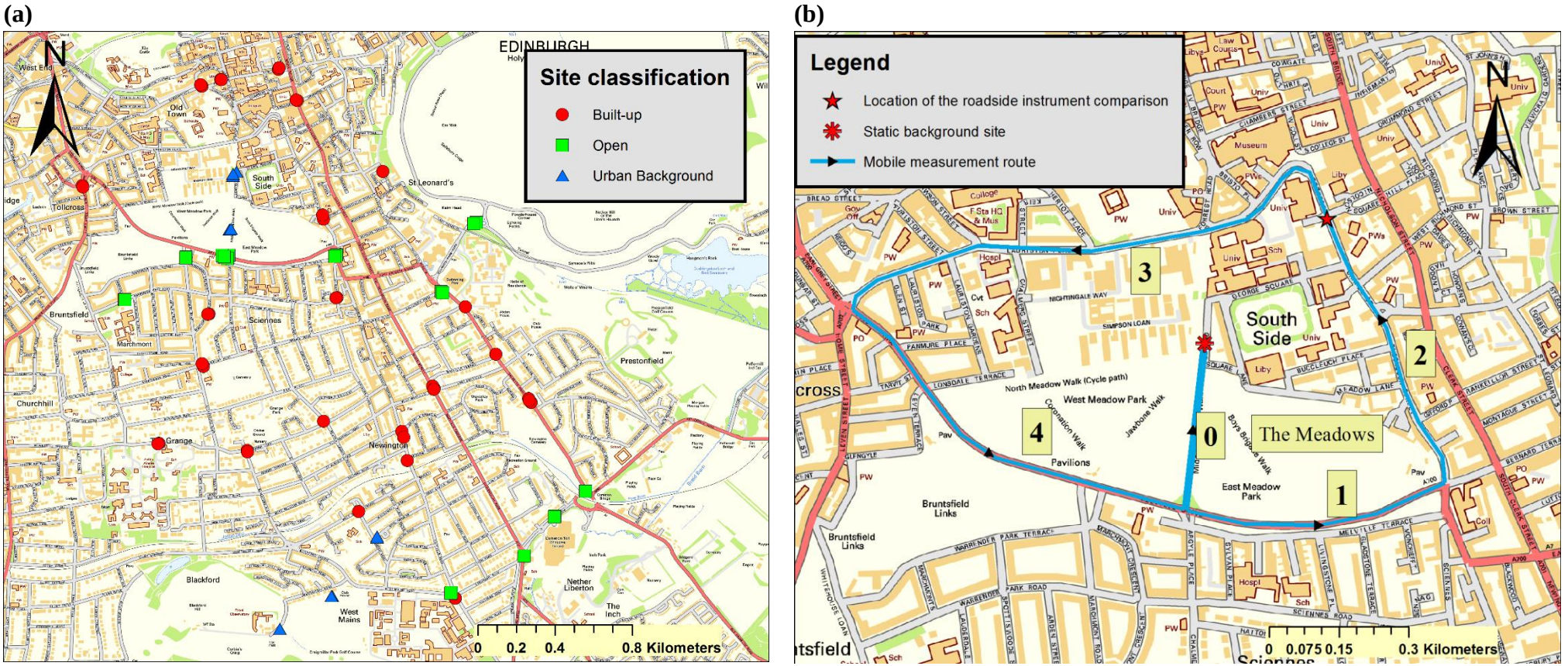



*Figure 3.1 (a) Location and classification of the static measurement sites. Streets with buildings on both sides are classified as built-up. Streets with buildings on only one side or no buildings on either side are classified as open. Background sites are at least 130 m away from the nearest major road. (b) Mobile measurement route and location of the contemporaneous static background measurements. Segments of the mobile route are labelled from 0 to 4. Base map from Edina Digimap®.*

*Table 3.1 Summary of meteorological data measured on the rooftop of a seven storey building located ~3 km to the south of the Meadows (55.92°N, 3.17°W) during each measurement period and during the instrument inter-comparison period in the winter campaign.*

| Date | rainfall (mm) | wind speed (m/s) | wind direction (°) | surface temperature (°C) | Relative humidity (%) | solar flux (kW/m$^2$) |
|---|---|---|---|---|---|---|
| **Winter Campaign** | | | | | | |
| Mon 2nd Dec 13 | 0 | 3.6 | 215 | 7.3 | 82 | 0.04 |
| Mon 9th Dec 13 | 0 | 4.9 | 223 | 10.4 | 79 | 0.04 |
| Sun 19th Jan 14 | 1.2 | 1.7 | 180 | 6.3 | 81 | 0.07 |
| Mon 27th Jan 14 | 0.6 | 1.6 | 147 | 5.3 | 80 | 0.06 |
| **Winter instrument inter-comparison** | | | | | | |
| Mon 3rd Feb 2014 | 0 | 3.1 | 132 | 5.9 | 72 | 0.22 |
| **Spring Campaign** | | | | | | |
| Sun 6th Apr 14 | 0.4 | 5.4 | 163 | 13.0 | 83 | 0.31 |
| Thu 10th Apr 14 | 0 | 4.9 | 252 | 9.9 | 60 | 0.45 |
| Fri 18th Apr 14 | 0 | 0.5 | 47 | 8.8 | 46 | 0.57 |
| Wed 23rd Apr 14 | 0 | 1.4 | 52 | 10.9 | 86 | 0.30 |
| Wed 30th Apr 14 | 0 | 1.1 | 53 | 7.0 | 100 | 0.03 |
| Wed 7th May 14 | 0.4 | 6.1 | 234 | 9.6 | 72 | 0.36 |

In the spring campaign, between April and May 2014, measurements were taken around a park area (~ 1 km$^2$), referred to locally as 'the Meadows', focusing on understanding the contributions from traffic-related sources and local background sources to BC, UFPNC and $PNC_{0.5\text{-}2.5}$ concentrations on typical urban streets (Figure 3.1b). The road on the south edge of the Meadows had an annual average daily flow of 13,272 vehicles in 2013 (DfT, 2015). The pollution level associated with traffic was monitored by walking along a route surrounding the Meadows. To measure temporal variation in the background concentrations, a duplicate set of instruments was located at a static site inside the route circuit during the collection of the mobile measurements. This background location had perpendicular distance between 160 – 480 m to the three sides of the triangle route (Figure 3.1b). The measurements were conducted on one Sunday and five weekdays. On each day the mobile measurements started together with the static measurements at the background site and proceeded in the directions indicated in Figure 3.1b. Two trips were carried out in the morning (~ 9–10 am) and early afternoon (~ 1–2 pm) during each day, except for adverse weather conditions on one of the weekday afternoons. Other incomplete sets of

measurements during each day were due to instrument failure. The route of the mobile measurements was divided into five segments with different street topographies and traffic densities, as labelled on Figure 3.1b. The total traffic passing the observer in the direction of the route was counted for each segment. Traffic flow in the opposite direction of the route is assumed to be similar. The duplicate instruments were compared against each other during the last four measurement days by co-location for at least 20 min either at the static background site or near a busy roadside (Figure 3.1b). The inclusion of both background and roadside sites was to cover a range of concentrations for the evaluation of duplicate instrument precision.

### 3.2.1.2 Instrumentation

The AE51 determines BC concentration from absorption of 880 nm laser light by particles continuously collected on a glass-fibre filter. The CPC measures UFP of particles between 0.01 and 1 μm in diameter by using laser light scattering after condensing particles with super-saturated isopropanol vapour. Although UFP is usually defined as particles smaller than 100 nm, since the number concentration is dominated by ultrafine particles the measurement from the CPC can be considered to represent UFP number concentration. The Dylos measures the particle number concentration using a laser light scattering technique in two size ranges, >0.5 and >2.5 μm. Only particles between 0.5 and 2.5 μm in diameter were included in this study and are thus referred to as $PNC_{0.5\text{-}2.5}$. The use of the terminology $PM_{2.5}$ elsewhere in this paper refers to the mass of all particles <2.5 μm as defined for air quality standards.

In the winter campaign the measurement intervals for the AE51, CPC and Dylos monitors were 1 min, 1 s and 1 min, respectively. Because of the shorter duration of a trip in the spring campaign the resolution for the AE51 was increased to 30 s to ensure sufficient data points for each segment. The Optimised Noise-reduction Averaging (ONA) algorithm was used to reduce the noise in the data recorded by the AE51 (Hagler et al., 2011). The ONA algorithm conducts adaptive time-averaging of the BC data, with the incremental light attenuation (ΔATN) through the instrument's internal filter determining the time window of averaging. The ΔATN thresholds were set at 0.01 and 0.05 for winter and spring measurements, respectively, as a result of

the different proportions of clean background areas and sampling resolutions in two campaigns. Negative values recorded after the smoothing were omitted from further analyses (consisting of ~5% of the whole data set), which mostly occurred when the measured concentrations were <100 ng/m$^3$.

Major axis (MA) regression analysis was carried out to test the equivalence between duplicate instruments, assuming that the uncertainties in the duplicate instruments are similar (Warton et al., 2006). A statistical summary of instrument inter-comparison results is given in Table 3.2. Correlations between duplicate instruments were highly significant; however, the 95% confidence interval of the slopes between the pairs of AE51 and CPC instruments in the spring campaign, and the pair of Dylos instruments in both campaigns, did not encompass unity. Therefore, corrections based on the slope and intercept from the MA regression analyses were applied to the corresponding instruments to allow comparison between duplicate instruments. It is noted that the slope coefficient for duplicate Dylos instruments derived at the end of the winter campaign (3rd Feb 2014) was significantly different from that derived during a 3-day comparison in an urban background environment at the beginning of the winter campaign (3rd – 5th Dec 2013) (Table 3.2). The regression results derived on 3rd Feb 2014 was used to correct Dylos data in the winter campaign. The effect of using a single equation to correct Dylos data in the winter campaign are discussed in the results section.

*Table 3.2 Statistical summary of major axis regression analyses for the instrument inter-comparisons. ** indicates correlation at >99% significance. UFPNC data from CPC were 1 min averaged in winter campaign, while in the spring campaign UFPNC data were measured every second, consistent with other analyses in 3.2.2 Results and discussion. The correction applied to Dylos data in the winter campaign was based on the regression results on 3rd Feb 2014.*

| Winter campaign | | | | | | |
|---|---|---|---|---|---|---|
| **Instrument** | **Location** | **Date** | **Correlation coefficient (*r*)** | **Slope (95% confidence interval)** | **Intercept (95% confidence interval)** | **Number of data points** |
| **Dylos** | Urban background | 3rd Dec 2014 | 0.97** | 0.60 (0.58 – 0.61) | -0.07 (-0.13 – -0.02) | 447 |
| | | 4th Dec 2014 | 0.98** | 0.77 (0.76 – 0.77) | -0.11 (-0.14 – -0.08) | 1,440 |
| | | 5th Dec 2014 | 0.99** | 0.76 (0.75 – 0.77) | -0.04 (-0.09 – 0.01) | 633 |
| **AE51** | Mobile route in winter campaign | 3rd Feb 2014 | 0.97** | 0.97 (0.93 – 1.02) | 65 (-42 – 168) | 111 |
| **CPC** | | | 0.99** | 0.98 (0.96 – 1.00) | 78 (-240 – 391) | 107 |
| **Dylos** | | | 0.90** | 0.44 (0.40 – 0.47) | 1.6 (1.4 – 1.8) | 113 |
| **Spring campaign** | | | | | | |
| **AE51** | Background and roadside sites in spring campaign | 18th, 23rd ,30th Apr & 7th May 2014 | 0.80** | 1.1 (1.03 – 1.21) | -20 (-137 – 87) | 354 |
| **CPC** | | | 0.99** | 1.35 (1.35 – 1.35) | -317 (-389 – -246) | 8,302 |
| **Dylos** | | | 0.99** | 0.48 (0.47 – 0.49) | 1.49 (1.35 – 1.62) | 177 |

### 3.2.1.3 Additional data

The average $PM_{2.5}$ concentrations measured by a TEOM-FDMS instrument at the St. Leonards AURN station (55.945589° N, 3.182186° W) located 600 m to the east of the Meadows during each set of measurements are summarised in Figure 3.2. The TEOM-FDMS is a reference-equivalent instrument for quantifying gravimetric $PM_{2.5}$ for statutory purposes. Meteorological data for the period of measurement for each day in both campaigns and for the period when instrument inter-comparisons were carried out in the winter campaign were obtained from a weather station on the rooftop of a seven storey building (JCMB) located ~3 km to the south of the Meadows (55.92° N, 3.17° W) and are summarised in Table 3.1.

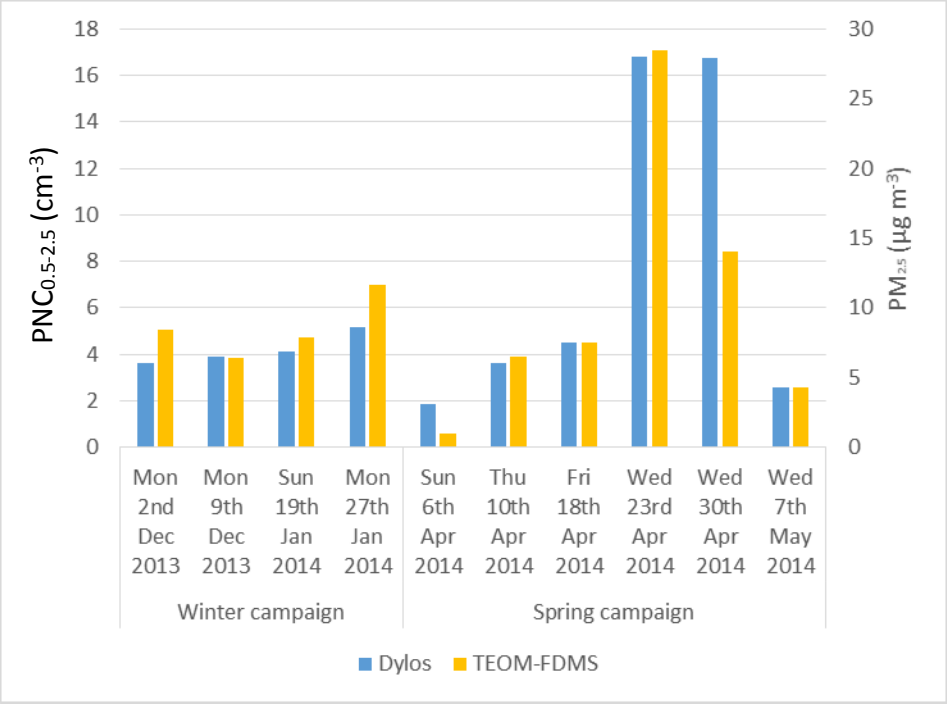



*Figure 3.2 Summary of average $PM_{2.5}$ mass concentration measured by TEOM-FDMS at the Edinburgh St. Leonards network monitor and $PNC_{0.5\text{-}2.5}$ measured by Dylos at static background site for spring period and the same Dylos for winter period.*

### 3.2.1.4 Data analyses

Reduced major axis (RMA) regression analysis was used to investigate the correlation between different pollutants, since the magnitudes of the data values and

their uncertainties are not the same for different instruments (Ayers, 2001). Difference between RMA and MA regression analysis is discussed in detail in Appendix I. Given the skewed nature of pollutant distributions, the non-parametric Mann-Whitney U test and the Kruskal-Wallis test were used to determine whether median concentrations differed significantly between two, and more than two samples, respectively. Data analyses were performed using R software (R Core Team, 2015). Air-mass back trajectory data was imported from pre-calculated trajectory data using the HYSPLIT trajectory model via the 'importTraj' function in openair (Carslaw and Ropkins, 2012), an R package for air quality data analysis.

## 3.2.2 Results and discussion

### 3.2.2.1 Spatiotemporal variability of BC, UFPNC and $PNC_{0.5-2.5}$

Distributions of BC, UFPNC and $PNC_{0.5-2.5}$ concentrations in the winter measurements are summarised in Figure 3.3. The distributions of pollutant concentrations are highly skewed, especially for BC and UFPNC, so the interquartile range (IQR) is used to illustrate the variation in distribution. Statistical summaries of the median and IQR for measurements on both sides of the road are listed in Table 3.3a. BC and UFPNC concentrations in the spring campaign and $PNC_{0.5-2.5}$ concentrations in both campaigns were corrected according to the MA regression analyses results in Table 3.2 to allow comparison between duplicate instruments. The ratio between IQR and median is used here as a metric of the spatiotemporal variability of each pollutant during each measurement trip. Table 3.3a shows that BC had the highest variability in the winter campaign (average IQR/median ratio of 1.34), which is about twice as high as for the metric with lowest variability, $PNC_{0.5-2.5}$ (average IQR/median ratio of 0.56). $PNC_{0.5-2.5}$ showed noticeable difference in distributions between two sides of the road. This is likely due to the use of a single linear equation to correct the Dylos data. Due to lack of inter-comparison between the Dylos instruments during each measurement day in winter campaign, it is difficult to quantify the genuine difference in $PNC_{0.5-2.5}$ concentrations between the two sides of road in Figure 3.3c. This also demonstrated that great caution needs to

be taken to compare absolute results from different 'low-cost' monitors. The discrepancies in the relationship between duplicate CPC and AE51 units on different days were considered to be less significant than for the Dylos units, since the correlation coefficients between CPC and AE51 units ($r > 0.97$) in the mobile winter inter-comparison were much higher and the gradients (0.98 and 0.97, respectively) were much closer to unity than the Dylos equivalents ($r = 0.90$, gradient = 0.44) (Table 3.2).

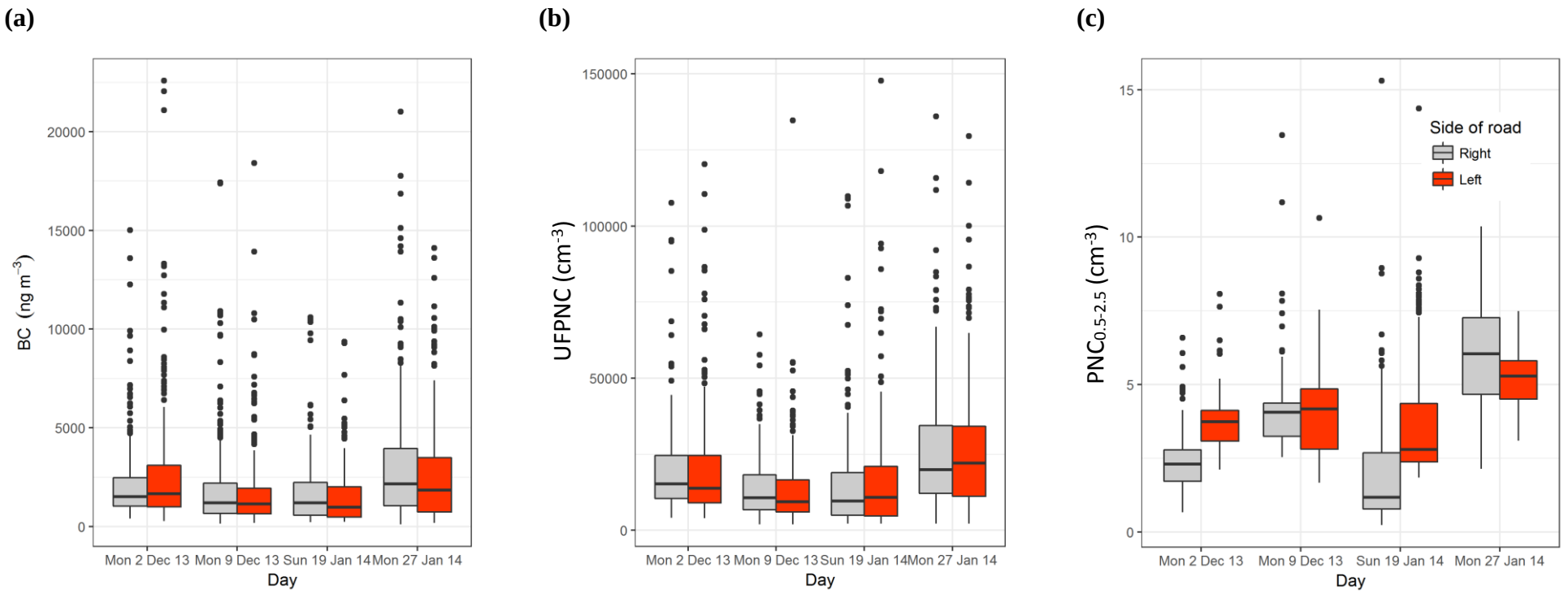



*Figure 3.3 Distributions of (a) BC, (b) UFPNC and (c) $PNC_{0.5-2.5}$ concentrations measured on both sides of the road during each week in the winter campaign. The bold horizontal line denotes the median, and the box demarcates the interquartile range. The whiskers extend to the values within 1.5 times the IQR on each side of the median. The $PNC_{0.5-2.5}$ concentration measured by one of the Dylos instruments was corrected based on the statistics from MA regression analysis of instrument co-deployment during the winter campaign (Table 3.2). Side of road is defined with respect to the walking direction in the mobile measurements. The UFPNC concentrations in (b) are 1 min averages of the raw 1 s data.*

*Table 3.3 The median, IQR and ratio between IQR and median on each day for BC, UFPNC and $PNC_{0.5\text{-}2.5}$ for (a) winter and (b) summer campaign. Left and right side of the road is defined with respect to the walking direction in the mobile measurements. Adjustment derived from MA regression analyses of instrument co-deployments (Table 3.2) was applied to one set of $PNC_{0.5\text{-}2.5}$ data in both campaigns, and one set of BC and UFP data in the spring campaign.*

**(a) Winter campaign**

| **BC [Right (Left)]** | | | |
|---|---|---|---|
| **Date** | **Median (ng $m^{-3}$)** | **IQR (ng $m^{-3}$)** | **IQR/Median** |
| Mon 2nd Dec 13 | 1,481 (1,668) | 1,540 (2,112) | 1.04 (1.27) |
| Mon 9th Dec 13 | 1,210 (1,105) | 1,521 (1,336) | 1.26 (1.21) |
| Sun 19th Jan 14 | 1,195 (951) | 1,664 (1,571) | 1.39 (1.65) |
| Mon 27th Jan 14 | 2,155 (1,842) | 2,900 (2,834) | 1.35 (1.54) |
| **UFPNC [Right (Left)]** | | | |
| **Date** | **Median ($cm^{-3}$)** | **IQR ($cm^{-3}$)** | **IQR/Median** |
| Mon 2nd Dec 13 | 15,251 (13,971) | 14,132 (15,870) | 0.93 (1.14) |
| Mon 9th Dec 13 | 10,664 (9,481) | 11,412 (10,759) | 1.07 (1.13) |
| Sun 19th Jan 14 | 9,649 (10,916) | 14,030 (16,715) | 1.45 (1.53) |
| Mon 27th Jan 14 | 19,890 (22,502) | 22,349 (23,490) | 1.12 (1.04) |
| **$PNC_{0.5\text{-}2.5}$ [Right (Left)]]** | | | |
| **Date** | **Median ($cm^{-3}$)** | **IQR ($cm^{-3}$)** | **IQR/Median** |
| Mon 2nd Dec 13 | 2.3 (3.74) | 1.07 (1.04) | 0.46 (0.28) |
| Mon 9th Dec 13 | 4.06 (4.18) | 1.13 (2.05) | 0.28 (0.49) |
| Sun 19th Jan 14 | 1.18 (2.8) | 1.91 (1.98) | 1.62 (0.71) |
| Mon 27th Jan 14 | 6.05 (5.29) | 2.6 (1.29) | 0.43 (0.24) |

**(b) Spring campaign**

| **BC [Mobile (Static)]** | | | |
|---|---|---|---|
| **Date** | **Median (ng $m^{-3}$)** | **IQR (ng $m^{-3}$)** | **IQR/Median** |
| Sun 6th Apr 14 | 251 (23) | 282 (3) | 1.12 (0.14) |
| Thu 10th Apr 14 | 955 (276) | 1,027 (470) | 1.08 (1.7) |
| Fri 18th Apr 14 | 1,035 (540) | 6,54 (391) | 0.63 (0.72) |
| Wed 23rd Apr 14 | 2,124 (1,779) | 2,208 (594) | 1.04 (0.33) |
| Wed 30th Apr 14 | 1,672 (1,154) | 1,850 (288) | 1.11 (0.25) |
| Wed 7th May 14 | 926 (217) | 1,208 (635) | 1.3 (2.92) |
| **UFPNC [Mobile (Static)]** | | | |
| **Date** | **Median ($cm^{-3}$)** | **IQR ($cm^{-3}$)** | **IQR/Median** |
| Sun 6th Apr 14 | 1,757 (624) | 1,891 (509) | 1.08 (0.82) |
| Thu 10th Apr 14 | 10,276 (6,081) | 8,824 (6,560) | 0.86 (1.08) |
| Fri 18th Apr 14 | 38,009 (30,128) | 24,034 (22,487) | 0.63 (0.75) |
| Wed 23rd Apr 14 | 11,370 (1,6115) | 8,674 (15,642) | 0.76 (0.97) |
| Wed 30th Apr 14 | 9,633 (4,547) | 17,816 (2,172) | 1.85 (0.48) |
| Wed 7th May 14 | 6,431 (3,286) | 8,353 (8,66) | 1.3 (0.26) |
| **$PNC_{0.5\text{-}2.5}$ [Mobile (Static)]** | | | |
| **Date** | **Median ($cm^{-3}$)** | **IQR ($cm^{-3}$)** | **IQR/Median** |
| Sun 6th Apr 14 | 0.61 (1.87) | 0.42 (0.21) | 0.69 (0.11) |
| Thu 10th Apr 14 | 3.7 (3.46) | 0.51 (0.22) | 0.14 (0.06) |
| Fri 18th Apr 14 | 4.64 (4.41) | 0.47 (0.43) | 0.1 (0.1) |
| Wed 23rd Apr 14 | 19.46 (18.23) | 3.67 (5.99) | 0.19 (0.33) |
| Wed 30th Apr 14 | 15.31 (16.42) | 3.49 (3.05) | 0.23 (0.19) |
| Wed 7th May 14 | 2 (2.49) | 0.75 (0.44) | 0.38 (0.18) |

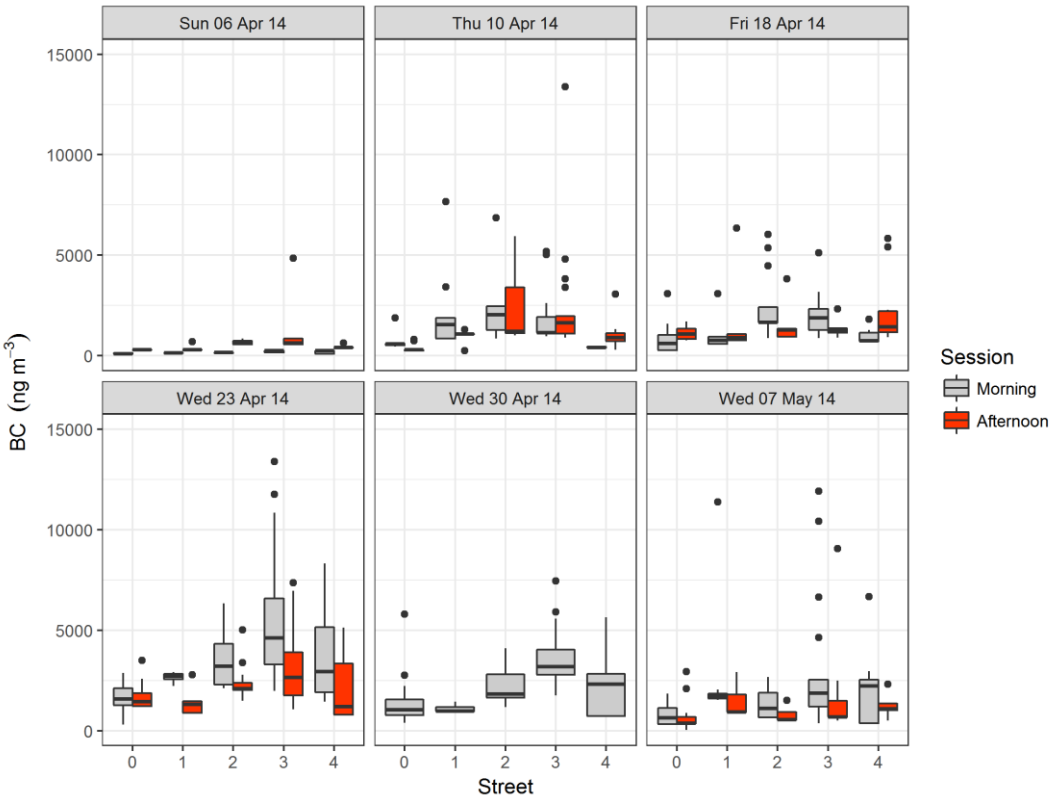
Distributions of BC, UFPNC and $PNC_{0.5\text{-}2.5}$ concentrations in the spring measurements are summarised in Figure 3.4. Jittered points are plotted for UFPNC to reveal the overlapping of large numbers of data points. In order to discern the details of the boxes in Figure 3.4a and b, BC and UFPNC concentrations greater than 15,000 $ng/m^3$ and 100,000 $cm^{-3}$, respectively, are not included. However, extremely high concentrations were observed, extending to 50,000 $ng/m^3$ for BC and 400,000 $cm^{-3}$ for UFPNC. Elevated BC concentrations were observed on streets with traffic (route segments 1 to 4) compared with the footpath through an area of urban park (route segment 0). The median concentrations for each route segment also varied between different times of a day. Only on Thu 10th Apr 2014 were BC concentrations not significantly different between morning and noon sessions (Mann-Whitney U test). BC concentration on Sun 6th Apr 2014 (median = 251 $ng/m^3$) was significantly lower than on weekdays (median = 1281 $ng/m^3$) (one-tailed Mann-Whitney U test, $P < 0.01$), which is similar to findings in the winter measurements.

The large range of outliers in the UFPNC distributions (Figure 3.4b) implied that UFPNC was greatly influenced by the emissions of nearby traffic. Despite the fact that the outlier measurements on each street were usually a factor 3 higher than their median, the difference in the medians between streets was relatively small, suggesting that elevated levels due to occurrence of exhaust plumes only transiently affected the UFPNC. Using ratios between median concentrations on streets and on the footpath it was calculated that, on average, a pedestrian was exposed to 1.8 times higher concentration on streets than on the footpath for UFPNC, and 2.3 times higher for BC. The markedly high UFPNC level on Fri 18th Apr 2014 at noon was associated with the lowest wind speed, highest solar flux (Table 3.1) and relatively high $O_3$ concentration of ~70 $\mu g/m^3$ (recorded at the St. Leonards AURN monitoring station) compared to other days. All these conditions are likely to have promoted secondary particle formation from photochemical reactions (Reche et al., 2011).

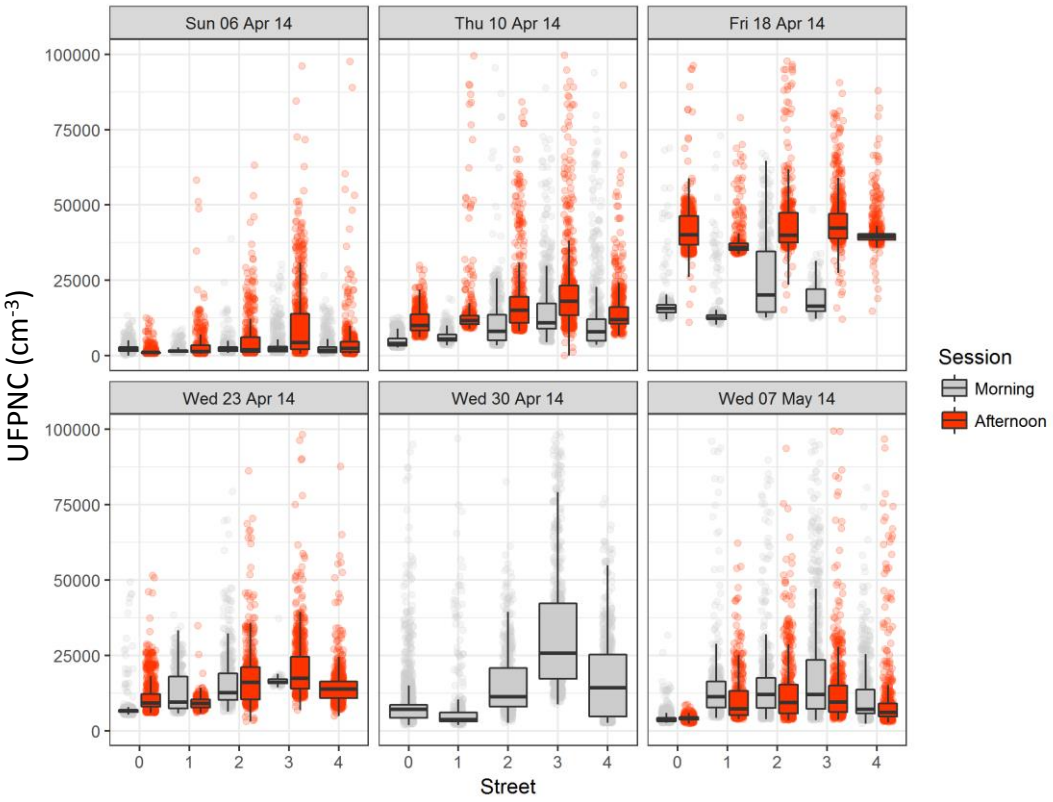
The distribution of $PNC_{0.5\text{-}2.5}$ concentrations (Figure 3.4c) exhibited a different pattern compared with BC and UFP concentrations. The range of concentration on each street was smaller, and the relative variation in medians between streets within the same day was lower. Only on Sunday and Wednesday 1 were $PNC_{0.5\text{-}2.5}$

concentrations highly significantly different between footpath and the streets (Mann-Whitney U test, $P < 0.01$), indicating that local traffic is frequently not a dominant contributor to $PNC_{0.5\text{-}2.5}$ concentrations in the locations where measurements were made. Significant variation in $PNC_{0.5\text{-}2.5}$ concentrations was noticed between all spring working day samples (Kruskal-Wallis test, $\chi^2 = 297$, df = 3, $P < 0.01$), which implies that $PNC_{0.5\text{-}2.5}$ is more influenced by regional sources and meteorological conditions on a particular day rather than local traffic (primary) emissions. This is consistent with observations in other cities; e.g. relatively low contribution (13%) to $PM_{2.5}$ from local traffic was also found in Paris (Skyllakou et al., 2014). Despite the Dylos instrument not measuring particles smaller than 0.5 µm, $PNC_{0.5\text{-}2.5}$ variation between days was in very good relative agreement with that of the $PM_{2.5}$ concentration measured by TEOM-FDMS at the St Leonard's national network urban background site (Figure 3.2), i.e. highest $PM_{2.5}$ levels on Mon 27$^{th}$ Jan 2014 in winter, and on Wed 23$^{rd}$ Apr 2014 and Wed 30$^{th}$ Apr 2014 in spring, and lower on Sundays in both campaigns.

**(a)**


Sun 06 Apr 14
Thu 10 Apr 14
Fri 18 Apr 14
Wed 23 Apr 14
Wed 30 Apr 14
Wed 07 May 14
BC (ng m$^{-3}$)
15000
10000
5000
0
0 1 2 3 4
Street
Session
Morning
Afternoon

**(b)**


Sun 06 Apr 14
Thu 10 Apr 14
Fri 18 Apr 14
Wed 23 Apr 14
Wed 30 Apr 14
Wed 07 May 14
UFPNC (cm$^{-3}$)
100000
75000
50000
25000
0
0 1 2 3 4
Street
Session
Morning
Afternoon

**(c)**

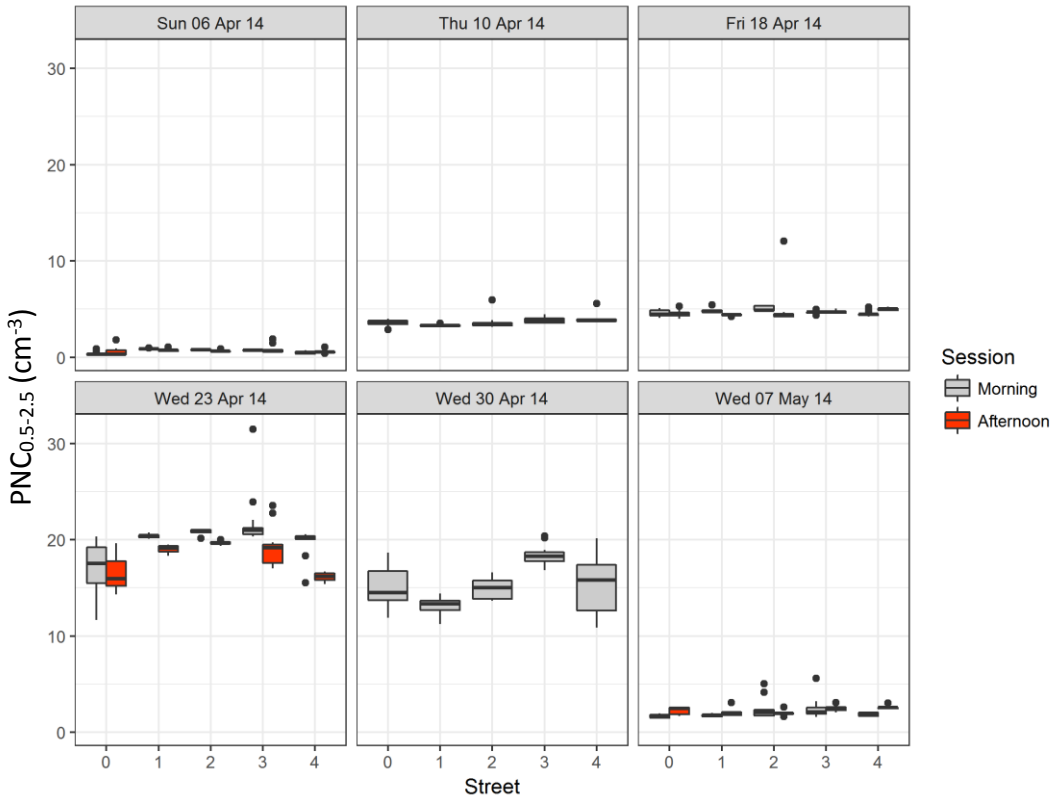



*Figure 3.4 Box plots for (a) BC, (b) UFPNC and (c) $PNC_{0.5\text{-}2.5}$ concentrations grouped by different days and sessions in the spring campaign. Data visualisations as defined in Figure 3.3. Jitter points are plotted in Figure 3.4b to reveal the extent of data in the outliers.*

Assuming that UFPNC represents contribution from local traffic sources, the ratio between median UFPNC and $PNC_{0.5\text{-}2.5}$ approximates the relative contribution of local sources to $PM_{2.5}$ concentration during that day. However, caution should be taken regarding this interpretation since any photochemical new particle formation (e.g. Fri 18$^{th}$ Apr 2014 in the spring campaign) may lead to overestimation of the relative level of local contribution compared to days when conditions are unfavourable to such formation occurring. In the winter measurements, the highest UFPNC/$PNC_{0.5\text{-}2.5}$ ratio (3,999), was associated with relatively low $PM_{2.5}$ concentration (8 µg/m$^3$) on Sunday (Figure 3.5). In the spring measurements, the highest ratio was on Fri 18$^{th}$ Apr 2014 (3,715), which was about ten times the ratio on Wed 23$^{rd}$ Apr 2014 (349), despite the fact the $PM_{2.5}$ concentration on Fri 18$^{th}$ Apr 2014 (~9 µg/m$^3$) was only a third of that on Wed 23$^{rd}$ Apr 2014 (~27 µg/m$^3$). It is

noted that in both campaigns highest local contributions did not coincide with the highest $PM_{2.5}$ concentrations. This observation again indicates a strong regional component in $PM_{2.5}$ observed in urban areas, as also noted by AQEG (2012).

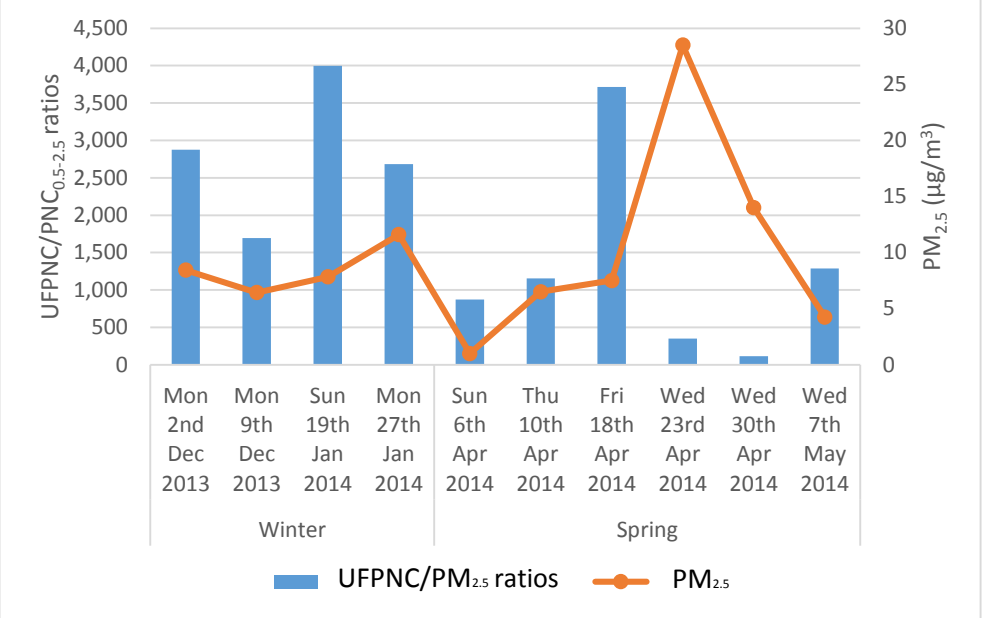



*Figure 3.5 Ratios between median UFPNC and $PNC_{0.5\text{-}2.5}$ concentrations and average $PM_{2.5}$ concentration measured by TEOM-FDMS on each day. The ratios in the spring campaign were calculated from the static measurements. The same instruments were used to calculate the ratios in the winter campaign for the purpose of comparison between the ratios in the two campaigns.*

Four-day air-mass back trajectories arriving at Edinburgh at 09:00 a.m. and 12:00 p.m. on each measurement day are plotted in Figure 3.6. The trajectories associated with the highest $PM_{2.5}$ concentrations in the winter and spring campaigns originated from North America and northern continental Europe, respectively. The highest $PM_{2.5}$ concentrations in winter coincided with a relatively large contribution from local sources (i.e. large $UFPNC/PNC_{0.5\text{-}2.5}$ ratio of 2,682). In contrast, the south-easterly trajectory originating from northern continental Europe (Figure 3.6) was associated with high $PM_{2.5}$ concentrations and low local contributions. It was noted that the $PM_{2.5}$ concentrations on the days associated with relatively high local emissions were close to the Scottish annual air quality objective threshold concentration for $PM_{2.5}$ of 12 µg/m$^3$ to be achieved by 2020 (AQEG, 2012), and that

the two days with $PM_{2.5}$ concentrations breaching the limit (Weds 23rd and 30th Apr 2014 in spring, Figure 3.2) had the largest regional contribution. In the UK context, transboundary import of inorganic aerosol components resulting in high particle concentrations was reported in a modelling study (Vieno et al., 2014).

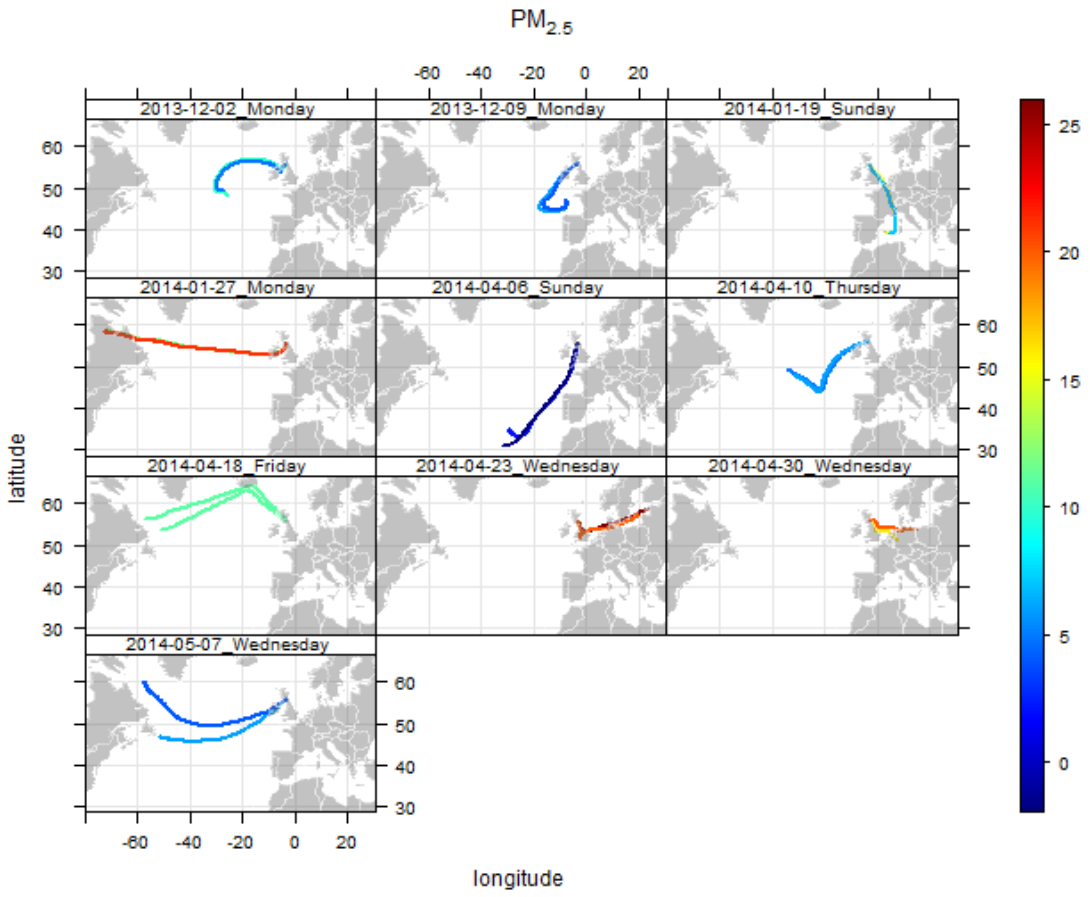



*Figure 3.6 Four-day air-mass back trajectories arriving in Edinburgh at 0900 and 1200 (GMT for winter and BST for spring) for each measurement day, coloured by the $PM_{2.5}$ mass concentration ($\mu g/m^3$) measured by the TEOM-FDMS instrument at Edinburgh St. Leonards at 0900 and 1200 (GMT for winter and BST for spring).*

The evidence presented here highlights that the effective management of ambient $PM_{2.5}$ concentrations in the UK may require international-scale cooperation in the control of emission sources contributing precursors of secondary particle formation and long-range transport. The good agreement between Dylos and TEOM-FDMS measurements, together with the above analysis on the potential sources of high

$PM_{2.5}$ episodes during each campaign, confirm that the Dylos instrument is capable of measuring elevated $PM_{2.5}$ arising from both regional and local influences (Steinle et al., 2015). However, this analysis does not reveal the contribution from local traffic emission at street level. This is in line with the findings of Price et al. (Price et al., 2014) who showed that number concentration of particles larger than 262 nm was more closely related to meteorological conditions whilst UFPs was more closely associated with traffic variables. Nevertheless this work presents a novel approach to apportioning $PM_{2.5}$ to local and regional sources by comparing the relationship between UFPNC and $PM_{2.5}$ with the help of back-trajectory analysis.

The IQR/median ratio was used to represent the variability of each pollutant in the spring measurements (Table 3.3b), in which the IQR/median ratio from the static measurements quantifies only the temporal variation between different times of a day, and the ratio in the mobile measurement reflects both spatial and temporal variation between different streets (with the former expected to be greater than the latter). Consistent with the observations in the winter measurements, the spatial variations of BC and UFPNC (average mobile IQR/median ratios of 1.05 and 1.08, respectively) were much larger than spatial variations of $PNC_{0.5\text{-}2.5}$ (average mobile IQR/median ratio of 0.29). The within-day variations of BC and UFPNC (range of static IQR/median ratios 0.14 – 2.92 and 0.26 – 1.08, respectively) were also generally larger than equivalent within-day variations of $PNC_{0.5\text{-}2.5}$ (range of static IQR/median ratios 0.06 – 0.33) but were of varying magnitudes. This implies that changes in local traffic counts and atmospheric conditions within a day have more pronounced effects on variations in BC and UFPNC than on variations in $PNC_{0.5\text{-}2.5}$. However $PNC_{0.5\text{-}2.5}$ varied more between days (static IQR/median ratio for the whole spring campaign of 3.27) than BC and UFPNC (equivalent ratios of 1.77 and 2.19, respectively). Collectively this evidence suggests that the variability of all three PM metrics is subject to the varying background concentrations. For BC and UFP in particular, the geographical locations, namely proximity to trafficked roads, also contribute a significant part of the spatial variability. Contrary to another study (Sullivan and Pryor, 2014), which found that the spatial variability of $PM_{2.5}$ (defined as the relative standard deviation for the mobile measurements on different routes) was 2 – 3 times greater than the sub-daily temporal variability (defined as the RSD

for the measurements when stationary), results from this study suggest that the spatial variability of $PNC_{0.5-2.5}$ was of similar magnitude to the sub-daily temporal variability. Possible explanations of this discrepancy include the larger geographical area in the former study, and the potential for bias from a few extremely high concentrations when using RSD rather than IQR as an indicator for variability.

#### 3.2.2.2 Pollutant concentration in relation to street topography and traffic counts

To understand the relationship between traffic counts and the pollutant concentration, reduced major axis (RMA) regression analyses were conducted between the mean of the concentrations on both sides of the road at each spot measurement in the winter campaign and traffic counts. Figure 3.7 shows the scatter plots of BC and UFPNC against traffic counts grouped by the classification of each site. Correlations of 5-min average BC and UFP concentrations with traffic counts were moderate and highly significant ($r$ = 0.56 and 0.39, respectively, $P < 0.01$, $n$ = 72). $PNC_{0.5-2.5}$ was not significantly correlated with traffic counts ($r$ = 0.17, $n$ = 72).

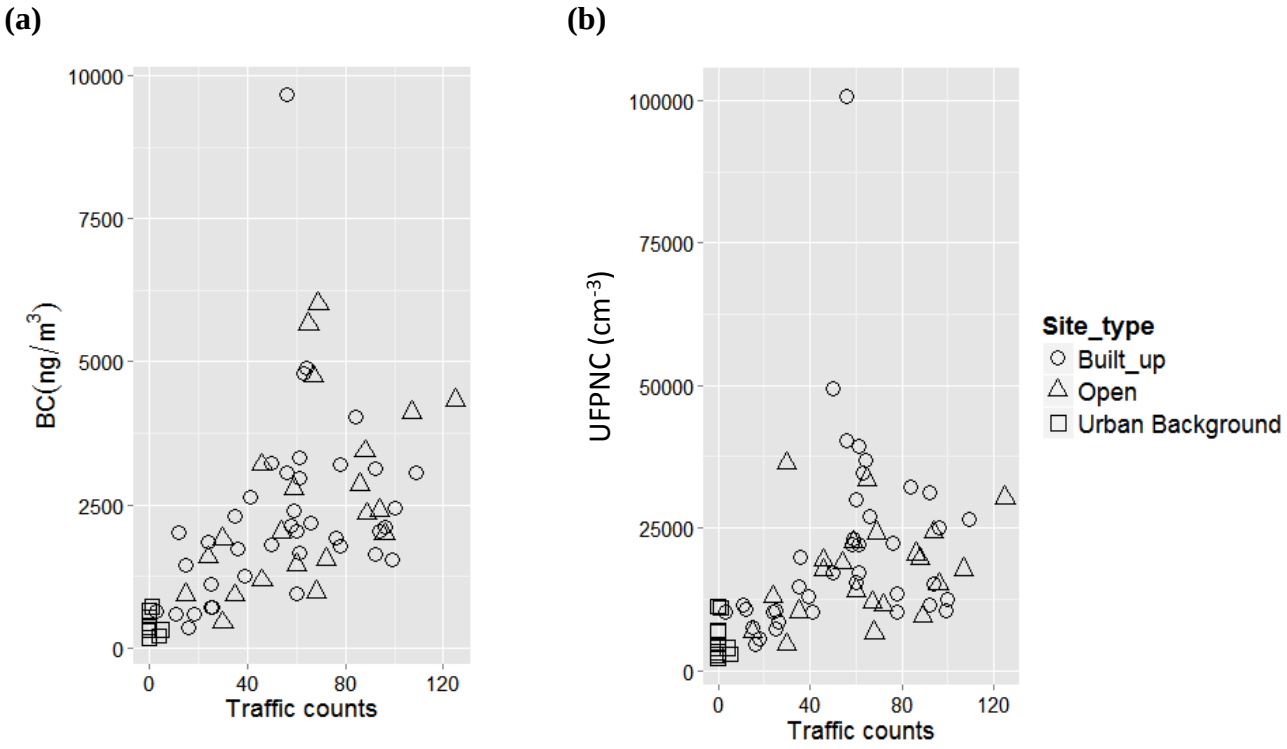



*Figure 3.7 Scatter plots of 5-min averaged (a) BC and (b) UFPNC vs. traffic counts at each measurement site in the winter campaign.*

Ratios between mobile and static measurements in the spring campaign were calculated for each timestamp to represent the elevation in the pollutant concentration due to traffic. The distribution of the ratios for BC and UFPNC were grouped by streets and plotted in Figure 3.8. Only ratios in the range 0 to 30 are shown in Figure 3.8 to avoid extreme values skewing the plots. Median mobile/static ratios are higher on Sun 6$^{th}$ Apr, Thu 10$^{th}$ Apr and Wed 7$^{th}$ May 2014 than on other days, an observation more pronounced for BC than for UFPNC. These days coincide with the days with higher wind speed (Table 3.1), which explains the greater contrast between roadside and background concentrations as the high wind speed facilitates the mixing of clean air at the background site while the immediate roadside concentration still stays at relatively high level. A summary of the median values of the mobile/static measurement ratio for each street and the traffic density on each street are tabulated in Table 3.4. The data show a common pattern in the mobile/static ratios for BC and UFPNC: streets 2 and 3 had the highest median

mobile/static ratios. Considering that the traffic density in streets 2 and 3 were not the highest among all the streets, the difference in street topography is the most likely explanation of the greater mobile/static ratios. Streets 2 and 3 are mostly characterised by street canyons with aspect ratios in the range of 0.6 – 1.3, whereas streets 1 and 4 are beside the Meadows urban park with an open terrain. As a result, dispersion in streets 2 and 3 is likely to be reduced compared to dispersion in streets 1 and 4. Similar results were reported in a study in Hong Kong characterised by high-rise buildings, where notably elevated BC and UFPNC concentrations were observed in deep street canyons compared to an open road although the traffic flow were significantly lower in street canyons (Rakowska et al., 2014).

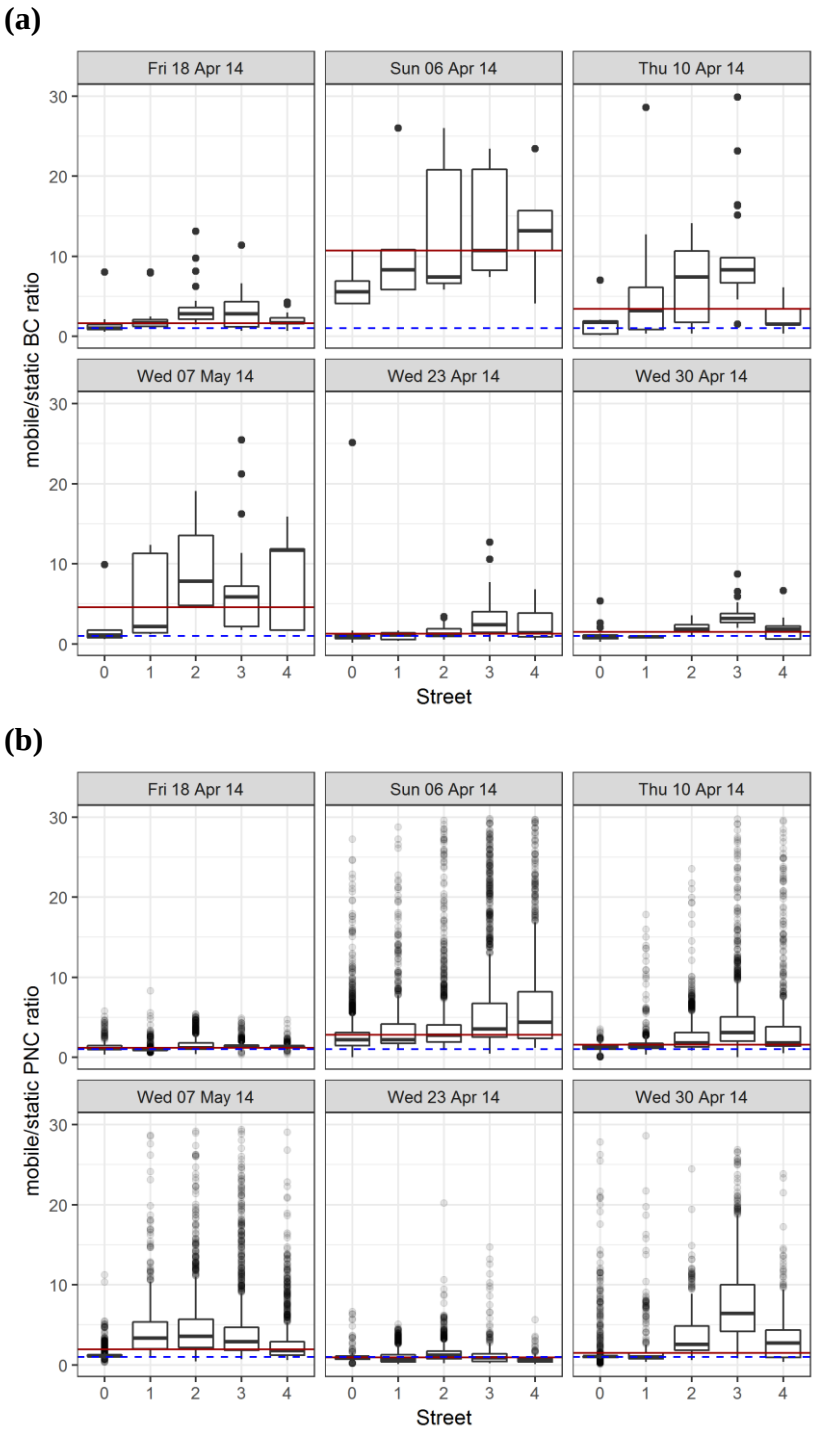



*Figure 3.8 Distributions of mobile/static measurement ratios in different streets for (a) BC, and (b) UFPNC, in the winter campaign. Solid red lines denote the median mobile/static measurement ratio for each day. The dashed blue line denotes a ratio of one to highlight the elevated concentrations on streets. Measurements at the static background were corrected based on the MA regression analyses results in Table 3.2.*

*Table 3.4 Median ratios of mobile/static measurements in each street and the traffic density for each street.*

| Street | Median of the mobile/static ratios for all the days (range of medians for each day) | | Mean traffic density (range of traffic density for each session) (traffic $km^{-1}$) |
|---|---|---|---|
| | BC | UFPNC | |
| 0 | 1.1 (0.9 – 5.6) | 1.1 (1.0 – 2.2) | 0 |
| 1 | 1.4 (0.9 – 8.3) | 1.4 (0.7 – 3.4) | 27 (16 – 38) |
| 2 | 3.0 (1.2 – 7.8) | 2.0 (1.2 – 3.5) | 33 (24 – 42) |
| 3 | 4.2 (2.4 – 15.8) | 2.8 (0.9 – 7.0) | 27 (18 – 35) |
| 4 | 1.8 (1.4 – 13.2) | 1.8 (0.6 – 4.7) | 39 (28 – 48) |

The correlation of averaged BC and UFPNC concentration with traffic counts for each route segment was calculated in the spring campaign. BC was again significantly correlated with traffic counts ($r = 0.45$, $P < 0.01$, $n = 50$), but UFPNC was not ($r = 0.25$, $n = 48$, $P = 0.09$). The lower correlation in spring compared to results of the winter campaign could be attributed to the fact that traffic counting on one side of the street in spring campaign may not have represented total traffic composition adequately. The low correlation between traffic count and UFPNC may have been caused by the secondary particle formation mentioned in Section 3.1. This interpretation was supported by the increased correlation when the data from Friday was excluded ($r = 0.66$, $P < 0.01$, $n = 39$). Despite the moderate to high correlation coefficients of BC and UFPNC with traffic counts found in this study, mixed conclusions have been drawn in the literature for the relationships between traffic and BC or UFP as a result of characteristics of specific measurement sites, the consistency of traffic flow and the formation/transformation of particles governed by environmental conditions (Kumar et al., 2008; Peters et al., 2014; Price et al., 2014; Rakowska et al., 2014). Further investigation of relationships between BC, UFPNC and traffic would provide beneficial information to inform the potential use of nearby traffic flow data to predict BC and UFPNC. Street topography effects on BC and UFPNC are similarly important for further investigation.

### 3.2.2.3 Correlation between BC and UFPNC

The correlations between BC and UFPNC were evaluated using RMA regression. The results summarised in Table 3.5 were calculated from 1-min averaged concentrations with BC as the *y* variable and UFPNC as the *x* variable.

The correlations between BC and UFPNC during each week ranged between 0.49 – 0.77 and in all cases were significant ($P$ <0.01). During working and non-working days, slope coefficients were not significantly different between measurements taken on different sides of the road in the winter campaign or between mobile and static measurements in the spring campaign. The agreement of the BC/UFPNC relationship between roadsides or between busy roads and local background suggests that BC and UFPNC vary similarly as they disperse away from traffic sources. The slopes on non-working days were significantly lower than the slopes on working days, indicating a decrease in BC/UFPNC ratios on non-work days. This variation in the BC/UFPNC ratios was not only observed in the measurements taken near road but also in the static background measurements during the spring campaign. The reason for this variation in BC/UFPNC relationship is likely due to the decrease in heavy goods vehicles (HGV) on the road during the weekend as HGV had the lowest share on weekends (3.2%) compared with weekdays (~6.8%). This observation suggests that HGV contribute relatively more to BC than to UFPNC. Therefore a policy targeting reduction of UFPNC may not be effective if it only focuses on restricting HGV. On the other hand policies aimed at reduction of BC should focus on controlling HGV emissions, which has direct implications for instance for the design and implementation of Low Emission Zones in urban areas. Considering that recent epidemiological studies suggest that BC was more strongly associated with adverse health effects than was the $PM_{2.5}$ mass (Grahame et al., 2014; WHO, 2013), traffic control strategies targeting HGVs may effectively contribute to alleviation of population health burdens arising from urban air pollution.

*Table 3.5 RMA regression analyses for 1-min averaged BC and UFPNC. The shading in the table represents the data collection period (grey: working days; white: non-working days). Left and right side of the road is defined with respect to the walking direction in the mobile measurements. ** indicates correlation at >99% significance.*

| **Winter campaign** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Left side of road** | | | | **Right side of road** | | | |
| ***r*** | **Slope (ng × $10^{-6}$)** | **Intercept ($ng/m^{-3}$)** | **Number of data points** | ***r*** | **Slope (ng × $10^{-6}$)** | **Intercept ($ng/m^{-3}$)** | **Number of data points** |
| 0.66** | 0.15 (0.14 – 0.16) | -575 (-726 – -432) | 833 | 0.65** | 0.16 (0.15 – 0.17) | -737 (-900 – -583) | 822 |
| 0.58** | 0.08 (0.07 – 0.09) | 149 (-1 – 283) | 223 | 0.58** | 0.11 (0.09 – 0.12) | 137 (-43 – 298) | 225 |
| **Spring campaign** | | | | | | | |
| **Mobile** | | | | **Static** | | | |
| ***r*** | **Slope (ng × $10^{-6}$)** | **Intercept ($ng/m^{-3}$)** | **Number of data points** | ***r*** | **Slope (ng × $10^{-6}$)** | **Intercept ($ng/m^{-3}$)** | **Number of data points** |
| 0.49** | 0.13 (0.12 – 0.14) | -102 (-289 – 69) | 319 | 0.50** | 0.11 (0.10 – 0.13) | 148 (61 – 226) | 270 |
| 0.62** | 0.05 (0.04 – 0.06) | -48 (-158 – 50) | 166 | 0.77** | 0.04 (0.03 – 0.04) | -15 (-46 – 13) | 181 |

## 3.3 Study of the correlation and drivers of UFP and other pollutants through fixed-site long-term measurements

### 3.3.1 Methods

#### 3.3.1.1 Instrumentation

Continuous monitoring of UFPNC, LDSA, mean particle size and $PM_{2.5}$ were conducted at St Leonards and Gorgie automatic monitoring sites in 2015. The St Leonards monitoring site is classified as an urban background site within the AURN network and provides hourly measurements of $NO_x$, $O_3$ and $PM_{2.5}$. Gorgie monitoring station is a roadside site within the SAQN and provides hourly measurements of $NO_x$. The instruments used to monitor UFPNC, LDSA, mean particle size and $PM_{2.5}$ are described below.

UFPNC, LDSA and mean particle size were measured by the miniature diffusion size classifier (DiSCmini, matter Aerosol AG, Switzerland). The device operates by polarising particles using a diffusion charger (10 nA, 3 – 5 kV) followed by detection of currents due to the charged particles at two stages: a diffusion stage ($i_d$) and a filtration stage ($i_f$) (Bau et al., 2012). The mean particle size is expected to be proportional to $i_f/i_d$. UFPNC and LDSA are related to the sum of two currents ($i_d$ + $i_f$). The range of particle size detectable by DiSCmini is approximately 10 – 700 nm and the range of UFPNC quantifiable is $10^3$ – $10^6$ $cm^{-3}$ (Fierz et al., 2011). In this analysis, 1-minute average data were exported from the DiSCmini. Measurements associated with errors (mostly flow rate error) were filtered out. Additionally, spurious measurements (i.e. UFPNC < $10^3$ or > $10^6$ particles $cm^{-3}$; mean size < 10 or ≥ 300 nm) were also discarded. The resulting 1-min-average data were averaged by hour using a 75% data capture threshold.

$PM_{2.5}$ was measured by reference analyser (TEOM-FDMS) at St Leonards and by microPEM monitor at both St Leonards and Gorgie. The microPEM monitor recorded data every 10 seconds. Due to the instability of the baseline some measurements were negative. These values were set to 0 which comprised 0.07% and 1.2% of total data recorded by the microPEM 586N and 618N units at St Leonards

and 0.3% by the microPEM 618N unit at Gorgie. Raw 10 s data were averaged to 1 min and then to hourly data using a 75% data capture threshold.

Additional measurements of 8 inorganic ions in the $PM_{2.5}$ fraction at the Auchencorth Moss monitoring site (18 km South of Edinburgh) were included in this analysis. The 8 inorganic ions were measured by a MARGA instrument (Measurement of Aerosols and Reactive Gases Analyser, Metrohm Applikon B.V., NL) utilising ion chromatography. The MARGA instrument is operated and maintained by the Centre for Ecology & Hydrology (CEH) (Twigg et al., 2015). These inorganic ions include the secondary inorganic aerosol (SIA) components: nitrate ($NO_3^-$), ammonium ($NH_4^+$) and sulphate ($SO_4^{2-}$), which are formed through reactions of gaseous pollutants $NO_x$, $NH_3$ and $SO_2$ (Vieno et al., 2014). The other inorganic ions are chloride ($Cl^-$), sodium ($Na^+$), magnesium ($Mg^{2+}$) (components of sea salt), calcium ($Ca^{2+}$) (part of crustal source), and potassium ($K^+$) (an indicator of biomass combustion) (Viana et al., 2008).

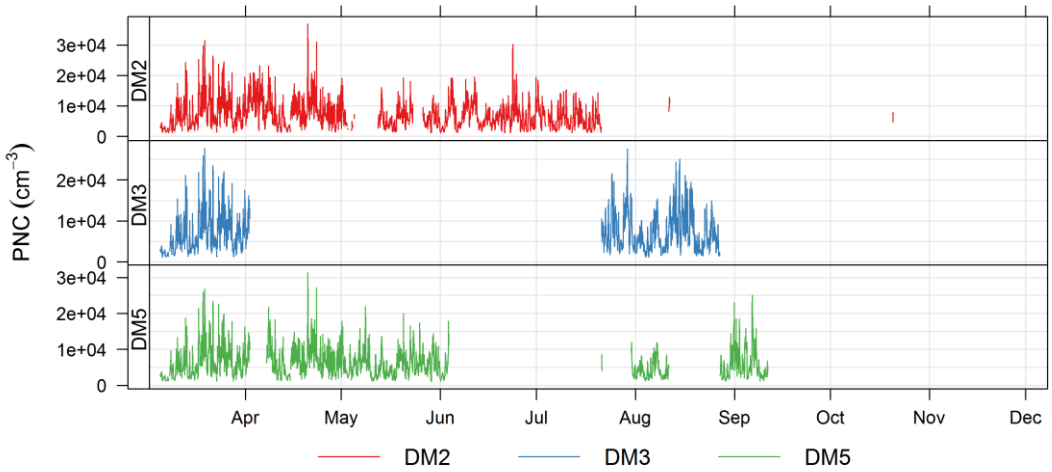
Hourly meteorological data, including wind speed/direction, temperature (*T*) and relative humidity (*RH*), are available at Gogarbank, a rural area west of the Edinburgh conurbation (more details in section ). Hourly solar radiation data were downloaded from the JCMB weather station already described in section 3.2.1.3 Additional data.

### 3.3.1.2 QA/QC procedures

Summaries of the data from the deployment periods of duplicate DiSCminis (DM2, DM3 and DM5) at St Leonards and Gorgie are shown in Figure 3.9. Monitoring at St Leonards was undertaken from March to September 2015 by different DiSCmini instruments. Measurements at the Gorgie monitoring site were intermittent due to instrument malfunction but covered a total period of around 2 months. Since DM5 had the longest co-location with both DM2 and DM3, it is considered as the reference instrument, against which the other two DiSCminis were calibrated. UFPNC and LDSA measured by duplicate DiSCminis appeared to follow a linear relationship. Mean particle size appeared to be described well by fitting a quadratic relationship (Figure 3.10). Hence calibration equations for UFPNC and LDSA were

calculated using MA regression analysis and mean particle size was calculated using $2^{nd}$-order polynomial regression. The co-location periods between DM3 and DM5 were separated by a few months (Figure 3.9a). There was a clear deviation in the calibration equations between the two periods (Figure 3.10). Therefore DM3 measurements before July were corrected using the Mar/Apr calibration equation, and measurements after July were corrected using the Aug calibration equation. Co-location between DM2 and DM5 was mostly continuous without long-term separation in-between (Figure 3.9a). The relationship between DM2 and DM5 did not show significant change during the co-location period (Figure 3.11). Therefore the equations derived in Figure 3.11 were applied to all the DM2 measurements.

**(a)**




**(b)**

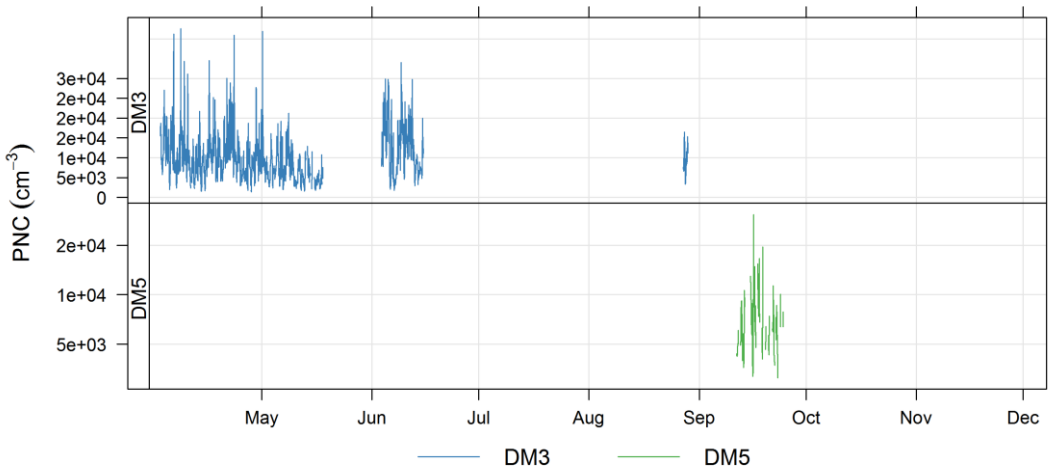



*Figure 3.9 Summary of data from the deployment periods for different DiSCmini units at (a)St Leonards and (b) Gorgie.*

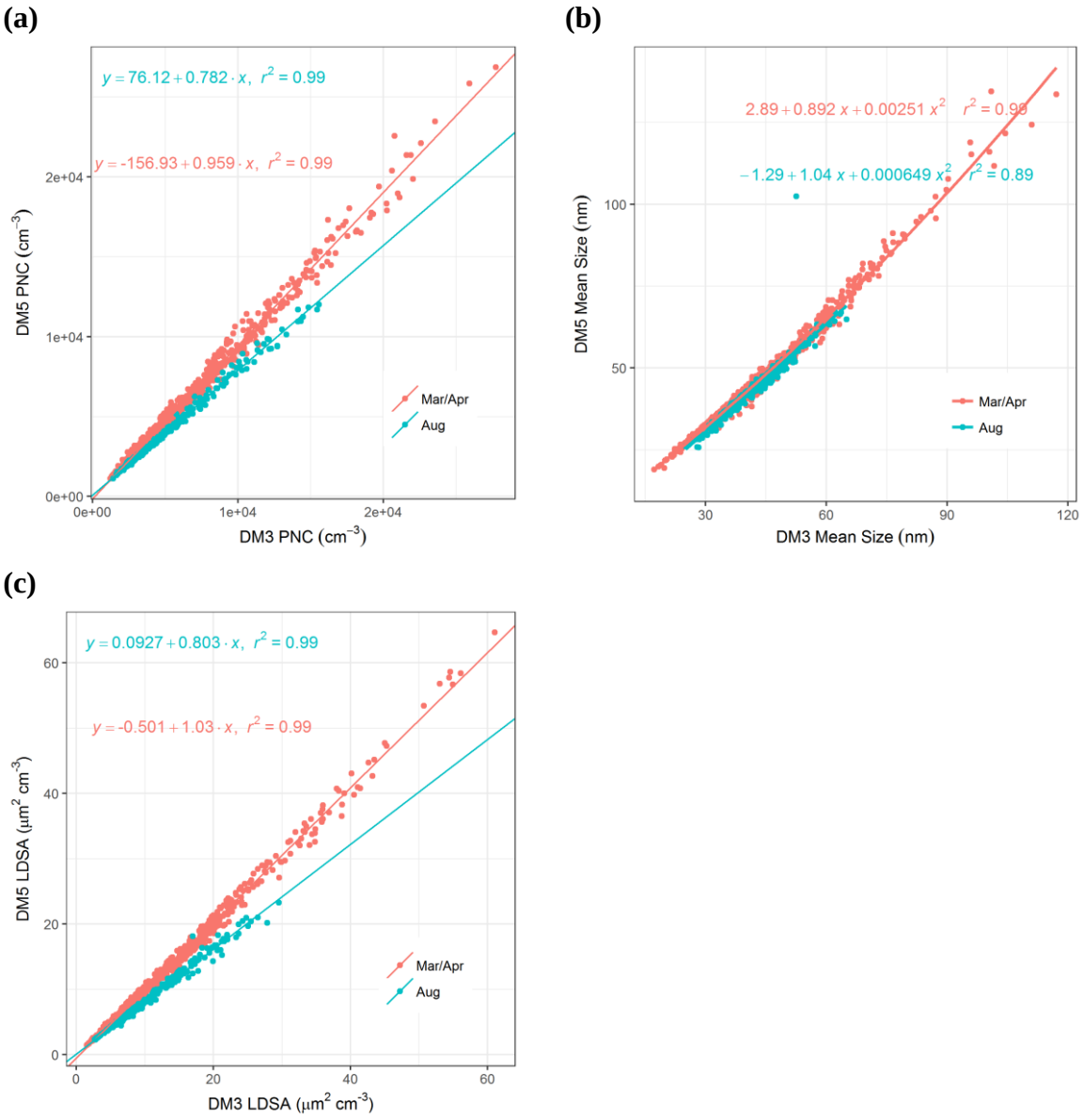



*Figure 3.10 Inter-comparison and calibration equations for DiSCmini units DM5 and DM3.*

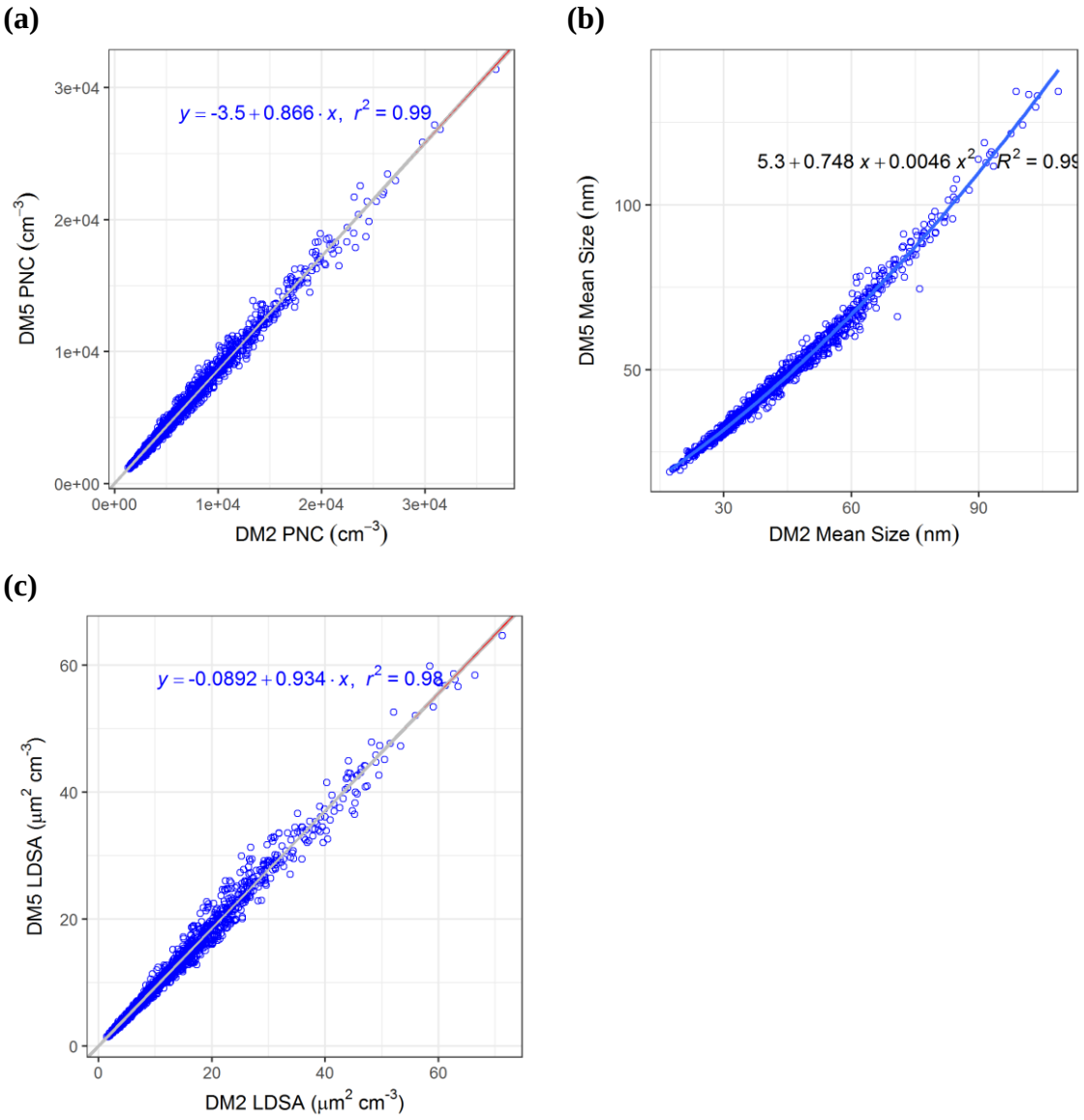
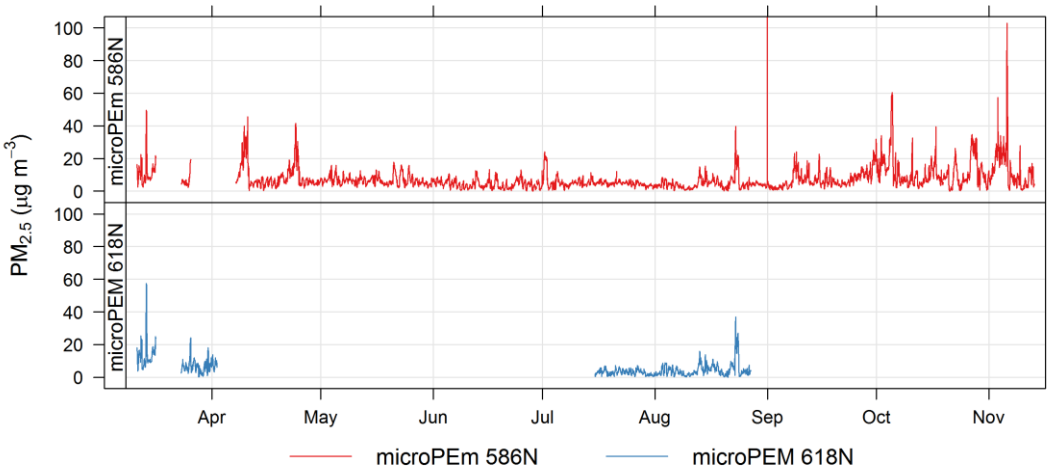



*Figure 3.11 Inter-comparison and calibration equations for DiSCmini units DM5 and DM2.*

Measurements from the deployment periods for microPEM units at both sites are summarised in Figure 3.12. Precision between duplicate microPEM units was discussed in detail in Chapter 2. MA regression analysis was conducted on all the co-location data in 2015 (Figure 3.13). The derived equation was used to correct microPEM 618N.

**(a)**




**(b)**

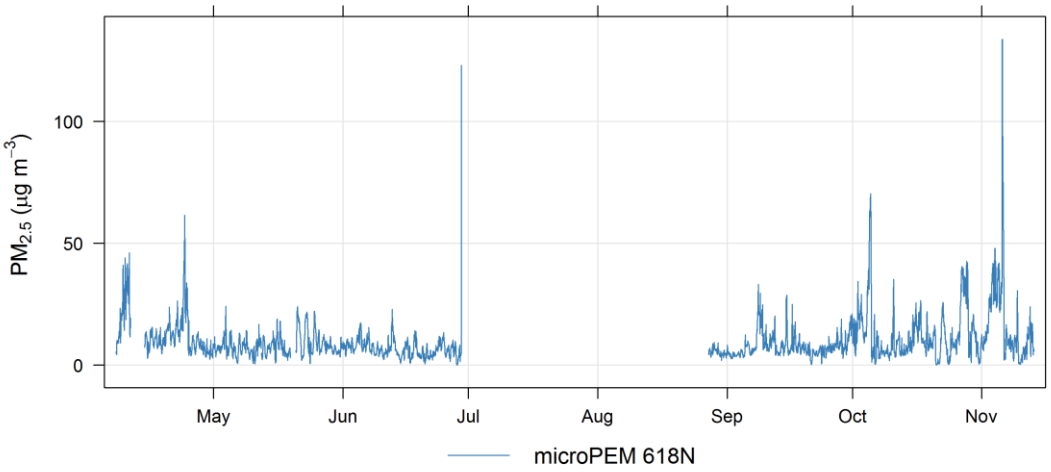



*Figure 3.12 Summary of the deployment periods for duplicate microPEM units at (a) St Leonards and (b) Gorgie.*

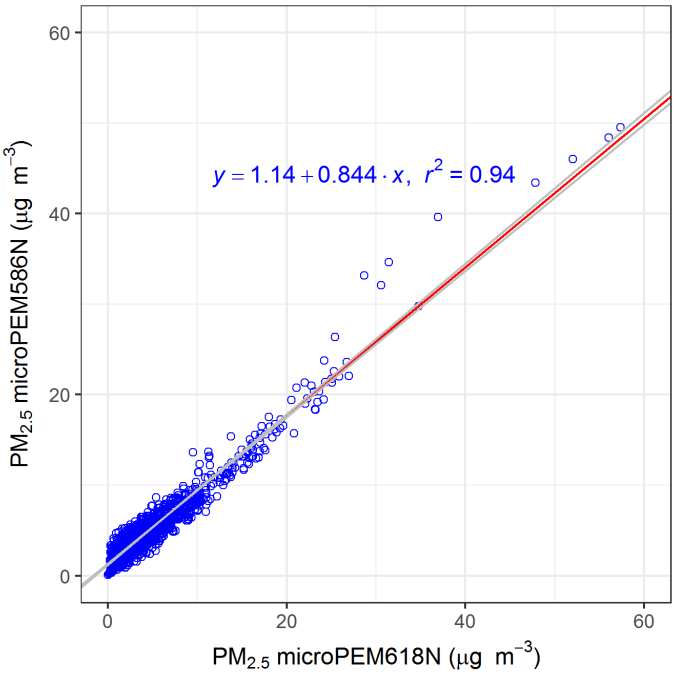



*Figure 3.13 Inter-comparison and calibration equations for duplicate microPEM units.*

## 3.3.2 Results and discussion

### 3.3.2.1 Correlations of UFPs and other pollutants

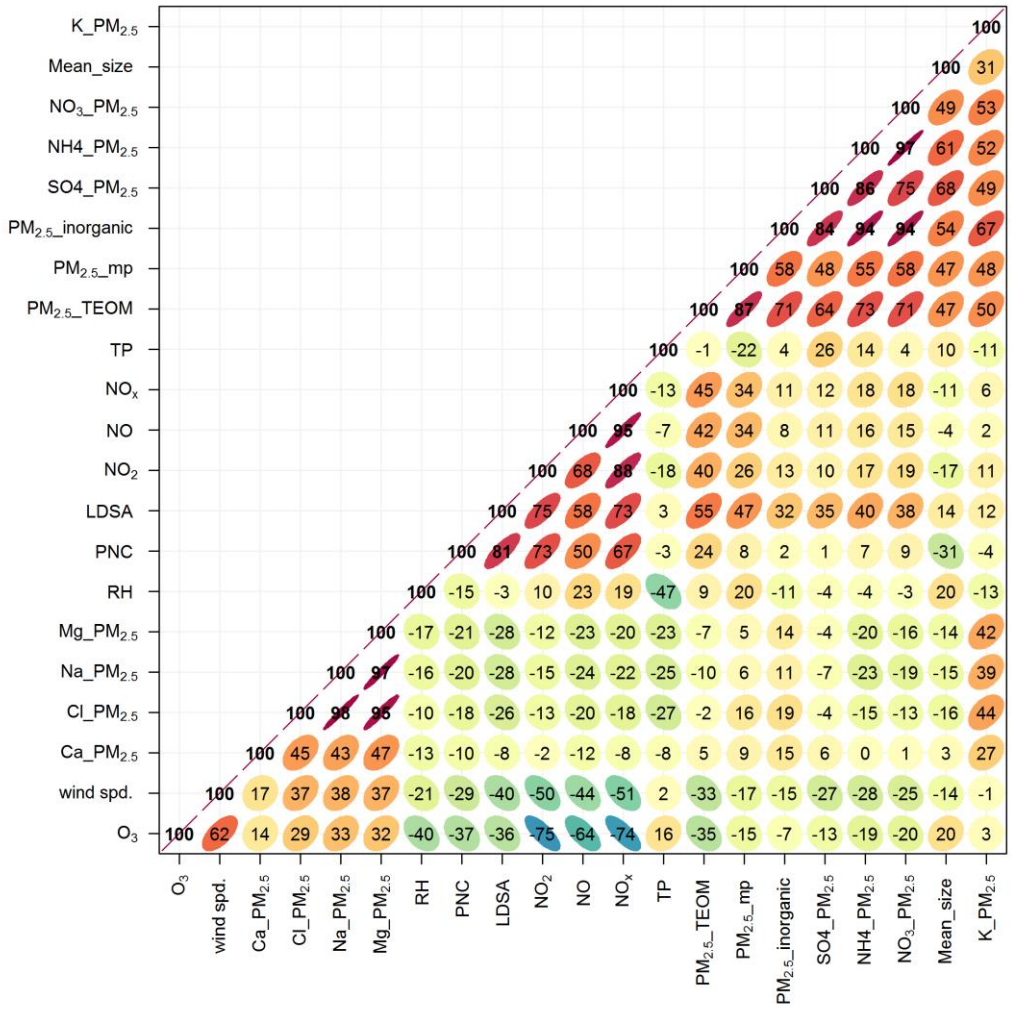
Pairwise Pearson correlation coefficients among all the pollutants measured at St Leonards and Gorgie, PM component measurements at Auchencoth and other meteorological variables (wind speed, *T* and *RH*) are summarised in Figure 3.14. The correlation matrix revealed three groups of highly correlated pollutants or meteorological variables at both sites (i.e. three separate rectangular triangles with correlation coefficient values visualised by dark orange and red colours).

One group corresponded to the pollutants related to traffic emissions including UFPNC, LDSA and $NO_x$. In this group UFPNC had the highest correlation with LDSA ($r > 0.8$), which may be partly due to the measuring principles of these two metrics by the DiSCmini instrument both relying on the sum of $i_d$ and $i_f$ (see 3.3.1.1 Instrumentation). However, it is also generally assumed that UFPNC is a good proxy for particle surface area, since smaller particles have much larger total surface area than larger particles for a give mass and smaller particles dominate the UFPNC. Correlation between UFPNC and $NO_x$ was generally high ($r > 0.67$) at both locations. $NO_2$ correlated more strongly with UFPNC at St Leonards than at Gorgie, and vice versa for NO. This observation confirms that UFPNC is strongly linked to traffic emissions since a greater proportion of $NO_x$ is in the form of NO at roadside.

Another group included mean particle size, $PM_{2.5}$, $NH_4^+$, $SO_4^{2-}$, $NO_3^-$ and $K^+$. Mean particle size showed moderate correlation with $PM_{2.5}$ concentration ($r \sim 0.50$) at both sites and with secondary inorganic $PM_{2.5}$ components measured at Auchencoth ($NH_4^+$, $SO_4^{2-}$ and $NO_3^-$).

The last group included the sea-salt components of $PM_{2.5}$ ($Na^+$, $Cl^-$ and $Mg^{2+}$), which weakly correlated with pollutants from the other groups. It is evident from the correlation matrix that UFPNC and LDSA are dominated by local traffic emissions whereas mean particle size is related to SIA originated from regional sources.

**(a)**


K_PM2.5
Mean_size
NO3_PM2.5
NH4_PM2.5
SO4_PM2.5
PM2.5_inorganic
PM2.5_mp
PM2.5_TEOM
TP
NOx
NO
NO2
LDSA
PNC
RH
Mg_PM2.5
Na_PM2.5
Cl_PM2.5
Ca_PM2.5
wind spd.
O3
O3 wind spd. Ca_PM2.5 Cl_PM2.5 Na_PM2.5 Mg_PM2.5 RH PNC LDSA NO2 NO NOx TP PM2.5_TEOM PM2.5_mp PM2.5_inorganic SO4_PM2.5 NH4_PM2.5 NO3_PM2.5 Mean_size K_PM2.5

**(b)**

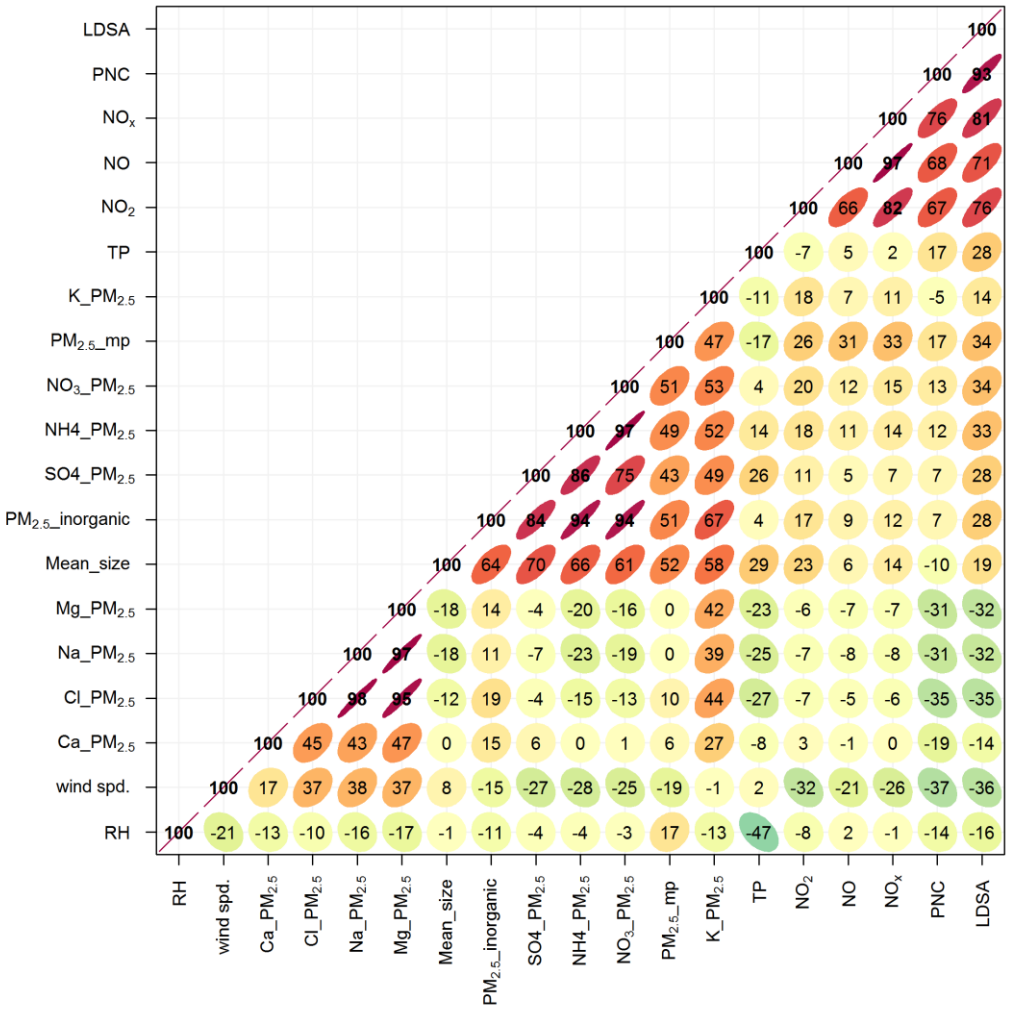



*Figure 3.14 Correlation coefficients for hourly average concentrations of various pollutants at (a) St Leonards and (b) Gorgie.*

### 3.3.2.2 Air-mass back trajectory analysis

To understand the geographic origins of different pollutants, 4-day air-mass back trajectories arriving at hourly intervals at Edinburgh were calculated for 2015. Six trajectory clusters were calculated based on the similarity of the angles from trajectory points to the trajectory origin using the *trajCluster* function in openair (Figure 3.15) (Carslaw, 2016). The determination of the number of trajectory clusters is a subjective opinion, which in reality is a balance between too few clusters that cannot uncover the distinct origins of air masses and too many clusters that unnecessarily differentiate very similar air masses. In this analysis, six trajectory clusters were found to be sufficient to separate distinct air masses but not overly complicated. Three of the trajectory clusters spent most of the time over marine locations (C2, C4 and C5), which are referred as 'Marine' clusters. Of the other three clusters, one had trajectories originating from the Atlantic and passing through Ireland before arriving at Edinburgh (C3), one had trajectories originating from northern Europe (C6), and the third had trajectories originating from western Europe and passing across the UK before arriving at Edinburgh (C1). These three clusters are referred as 'Marine-IRE', 'N/Europe' and 'UK-W/Europe' in the later discussion.

Boxplots of UFPNC, $PM_{2.5}$ and mean particle size corresponding to the six back-trajectory clusters at both St Leonards and Gorgie are summarised in Figure 3.16. $PM_{2.5}$ and mean particle size showed notable difference between different trajectory clusters. The 'UK-W/Europe' cluster was associated with significantly higher $PM_{2.5}$ concentration and mean particle size than the other clusters. Previous studies also showed that high $PM_{2.5}$ concentrations were associated with air masses from continental Europe carrying a large proportion of SIA (Harrison et al., 2012; Vieno et al., 2014). UFPNC did not show any correlation with respect to trajectory clusters but significantly higher concentrations were observed at the roadside than at the background site. $PM_{2.5}$ concentrations at the roadside were also consistently higher than at the background site but to a lesser extent than the UFPNC. Difference in mean particle size between roadside and background sites was inconsistent. In many instances mean particle size at roadside were higher than at the background site, which contradicts the expectation that roadside particles are dominated by freshly

emitted smaller nucleation particles. However, non-exhaust emissions from brake, tyre wear and road abrasion generate larger accumulation particles, which may also explain the elevation of mean particle size at Gorgie (Thorpe and Harrison, 2008). Overall, the air-mass back trajectories revealed that regional sources had greater influence than local sources on $PM_{2.5}$ concentration and mean particle size, whereas UFPNC was dominated by local traffic emissions.

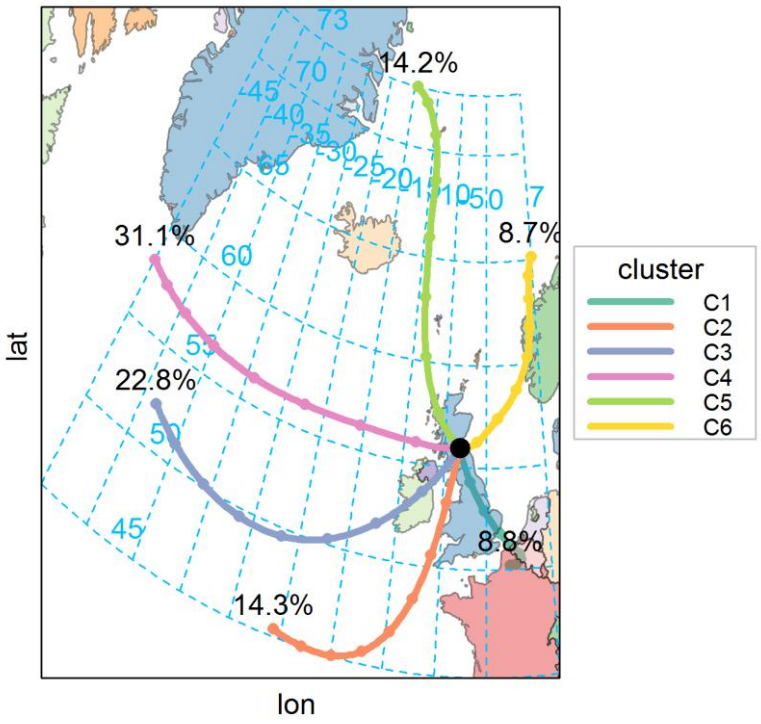
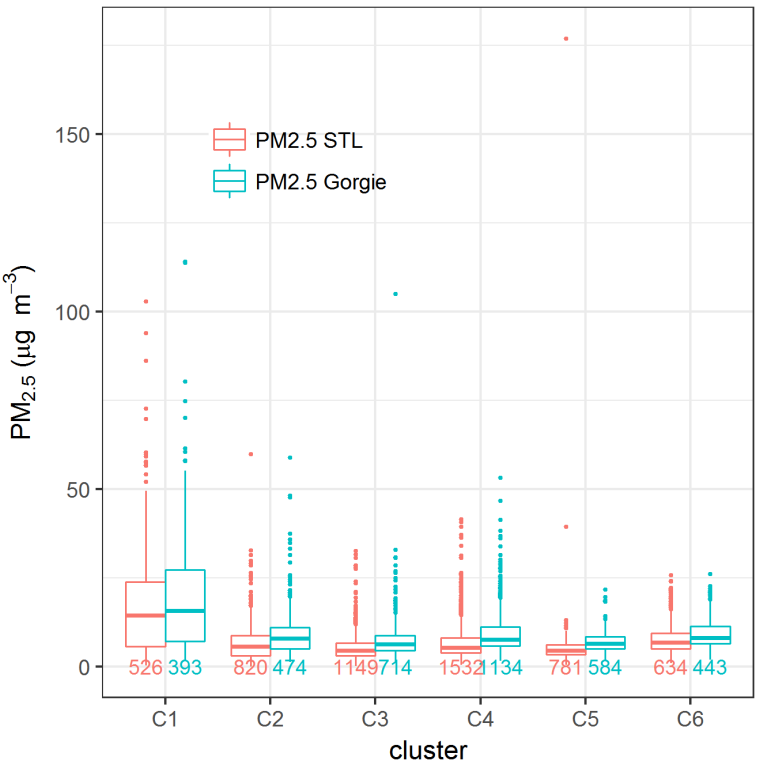
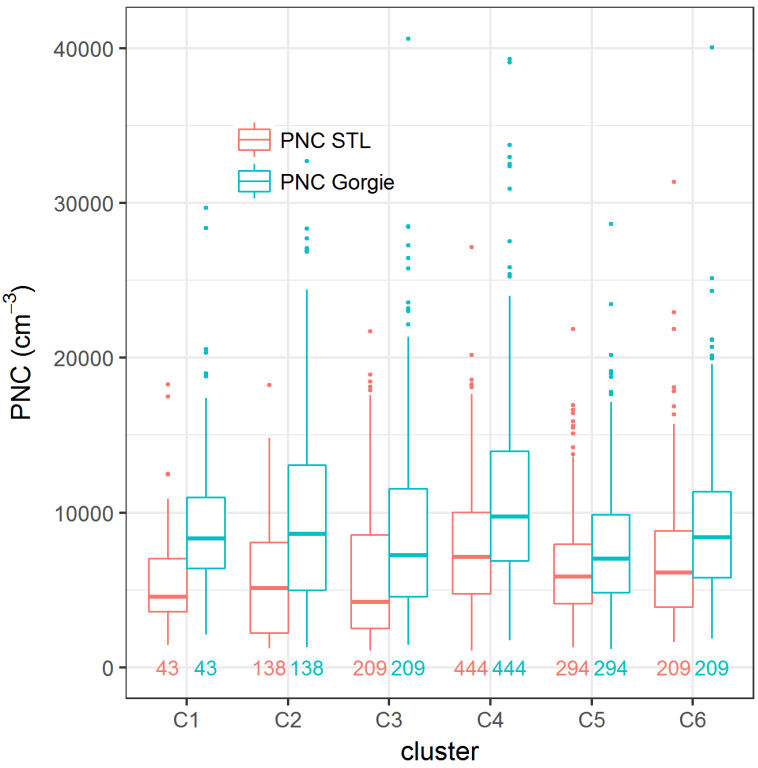



*Figure 3.15 Four-day air-mass back trajectories arriving at Edinburgh in 2015 clustered according to the similarity of the angles from trajectory points to the trajectory origin. The coloured lines show the average back trajectory for the trajectories contributing to that cluster and the percentage value is the contribution of all back trajectories to that cluster.*

**(a)**


PM2.5 STL
PM2.5 Gorgie
PM2.5 (μg m-3)
150
100
50
0
526 393
820 474
1149 714
1532 1134
781 584
634 443
C1
C2
C3
C4
C5
C6
cluster

**(b)**


PNC STL
PNC Gorgie
PNC (cm-3)
40000
30000
20000
10000
0
43 43
138 138
209 209
444 444
294 294
209 209
C1
C2
C3
C4
C5
C6
cluster

**(c)**

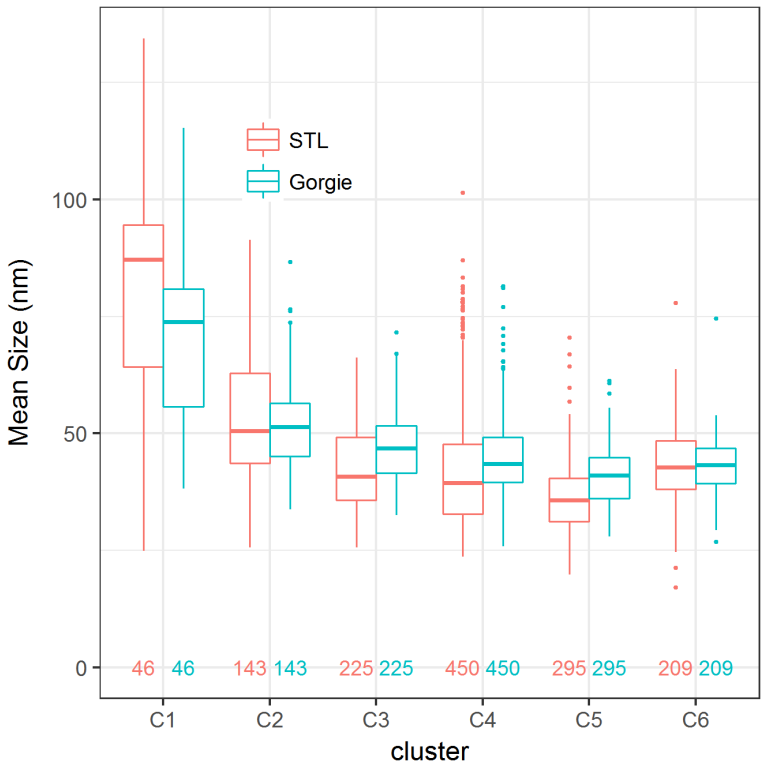


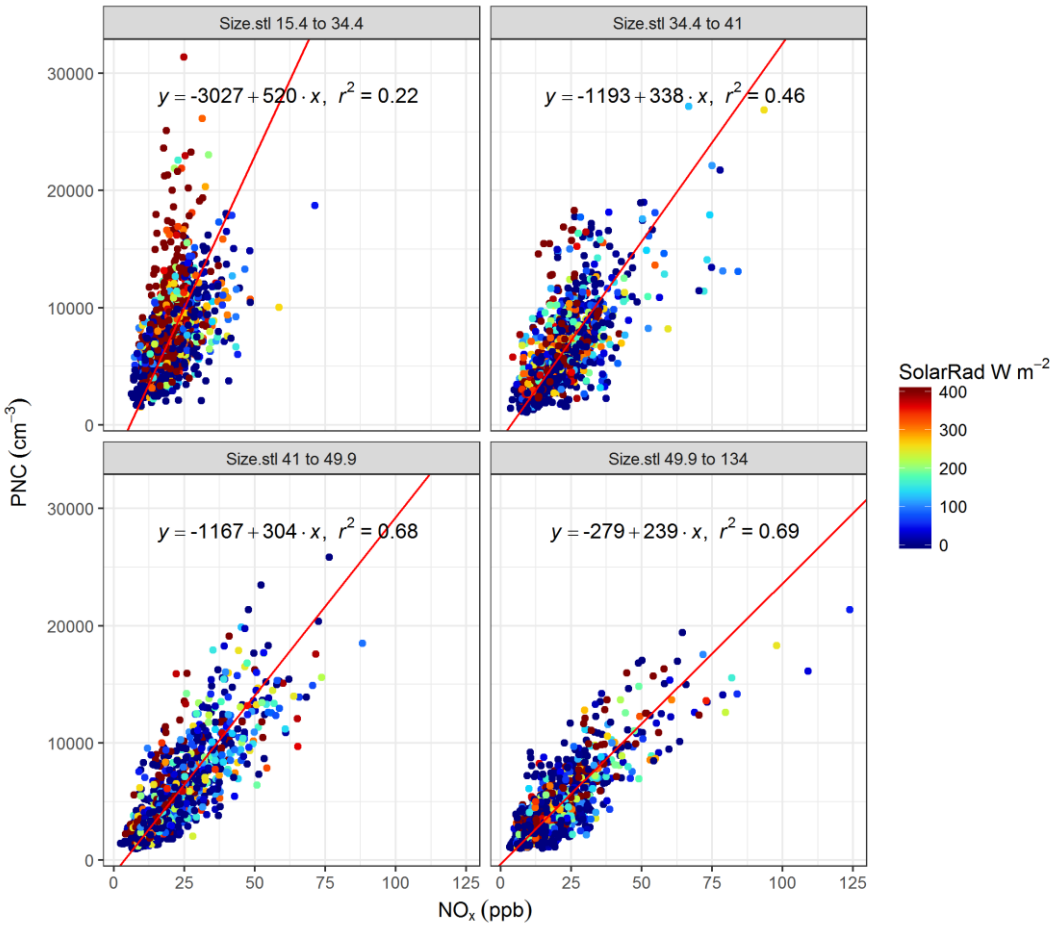

*Figure 3.16 Boxplots of (a) $PM_{2.5}$ measured by microPEM units and (b) UFPNC and (c) mean particle size for different back trajectory clusters. The number below each box shows the number of data points (hourly concentrations) contributing to each box.*

### 3.3.2.3 Relationship between UFPNC and $NO_x$ in relation to particle size

Given the strong association between UFPNC and traffic emissions, the relationship between UFPNC and $NO_x$ was further investigated to determine if $NO_x$ can be considered as a proxy to the UFPNC in epidemiological studies. Once emitted from exhaust, unlike $NO_x$, UFPs undergo a series of complex transformations that alter both the number and the size of the particles (Kumar et al., 2010). Typically, fresh nucleation particles (< 50 nm) almost completely evolve into larger accumulation mode particles (> 50 nm) and drastically decrease in number concentration after dispersing a few hundreds of metres from the sources (HEI, 2013). Therefore, UFPNC in different size ranges may have different relationships with $NO_x$ despite both being governed by similar dispersion processes. Scatter plots and RMA regression analysis between UFPNC and $NO_x$ in different mean particle size ranges at both sites are shown in Figure 3.17. Correlation between UFPNC and $NO_x$ decreased with decreasing mean particle size at St Leonards but no trend was found at Gorgie. Regression coefficients were significantly different between different mean particle size ranges. Slopes were generally lower for large particle sizes than for the small particle sizes. Different atmospheric conditions and emission activities lead to varying particle number-size distribution in air arriving at the monitoring sites. Deriving a quantitative relationship between $NO_x$ and UFPNC was shown in Figure 3.17 to be not robust if a broad size range of particles (10 – 700 nm as measured by DiSCmini) was considered.

At St Leonards anomalously high UFPNC coincided with small mean particle size and high solar radiation (top left panel of Figure 3.17a). These conditions suggested a potential new particle formation (NPF) event, as observed in many previous studies (Charron et al., 2007; Dall'Osto et al., 2013; Ma and Birmili, 2015). NPF events, as a result of gas-to-particle conversion due to build-up of gaseous precursors, generally occur at a large regional scale (Dall'Osto et al., 2013). However UFPNC at Gorgie was not found to be anomalously high during the same hours as the potential NPF event occurred at St Leonards.

**(a)**


Size.stl 15.4 to 34.4
Size.stl 34.4 to 41
Size.stl 41 to 49.9
Size.stl 49.9 to 134
y = -3027 + 520 · x, r² = 0.22
y = -1193 + 338 · x, r² = 0.46
y = -1167 + 304 · x, r² = 0.68
y = -279 + 239 · x, r² = 0.69
PNC (cm⁻³)
NOx (ppb)
SolarRad W m⁻²
400
300
200
100
0
30000
20000
10000
0
0
25
50
75
100
125

**(b)**

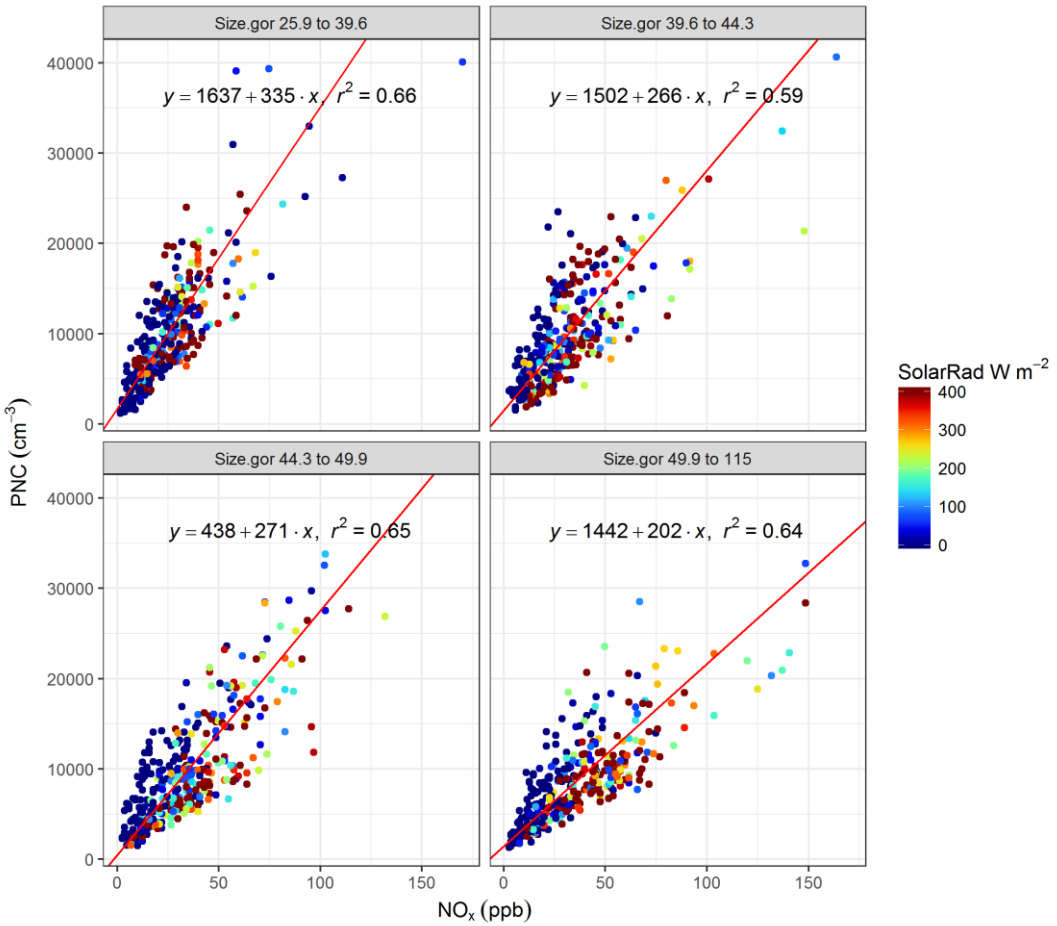



*Figure 3.17 (a) Scatterplot of hourly UFPNC against $NO_x$ concentration at (a) St Leonards and (b) Gorgie grouped by quartile ranges of mean particle size and coloured by solar radiation.*

### 3.3.2.4 Principal component analysis

Principal component analysis (PCA) was performed on the $PM_{2.5}$ component measurements at Auchencoth, gaseous pollutants and DiSCmini measurements at St Leonards and Gorgie and other meteorological variables to understand how different PM metrics are related to long and short-range sources. PCA is a data-reduction technique that reduces a dataset with a large number of potential correlated variables to a small number of uncorrelated variables (principal components) while retaining as much information from the original dataset as possible (Kabacoff, 2011). The number of principal components (PC) to calculate is usually determined by the eigenvalues associated with each PC. Typically, components with an eigenvalue greater than one are considered as principal components. The output of PCA gives a matrix of loadings of each variable for each PC. The loadings indicate the correlation (anti-correlation) of the each variable with a particular PC.

Four principal components with eigenvalues greater than one were calculated for both the St Leonards and Gorgie datasets. To aid the interpretation of the components Varimax rotation was applied to the loading matrix to ensure each component has a few variables with high loadings and many variables with small loadings. The loadings of each variable for each rotated principal components (RC) are summarised in Figure 3.18. The first RC in both datasets was highly correlated with SIA, $PM_{2.5}$ at St Leonards and mean particle size at the respective sites. RC1 can be interpreted as contribution from long-range transport sources. Despite being located close to traffic sources, mean particle size at Gorgie was still highly correlated with RC1, suggesting that long-range transport PM carrying large particles in the accumulation mode significantly influences the mean particle size across the city more than the impact from the local traffic. The second RC at both sites explained the variation in the sea-salt components of PM, which weakly correlated with other pollutants but moderately correlated with wind speed. The third RC explained the variation in the traffic related air pollutants measured locally at each site. UFPNC and LDSA strongly correlated with RC3 together with $NO_2$ and $NO_x$. $PM_{2.5}$ measured at St Leonards also had notable correlation ($r \sim 0.4$) with RC3 comparing with other PM components or mean particle size, indicating a small

contribution to $PM_{2.5}$ from local traffic emissions. The last RC in both datasets mainly explained the variance in various meteorological variables.

Consistent with the results in the previous back-trajectory analysis, the PCA results highlight the distinct sources governing the variability of different PM metrics. UFPNC and LDSA are solely related to local traffic emissions, whereas mean particle size is governed by $PM_{2.5}$ episodes associated with large contribution from long-range transported SIA.

**(a)**


RC1
RC2
RC3
RC4
Loading
1.0
0.5
0.0
-0.5
NO2
O3
NOx
PM2.5 (STL)
PNC
Mean size
LDSA
Temp
RH
Solar Rad
wind speed
NH4
Na
K
Ca
Mg
Cl
NO3
SO4

**(b)**

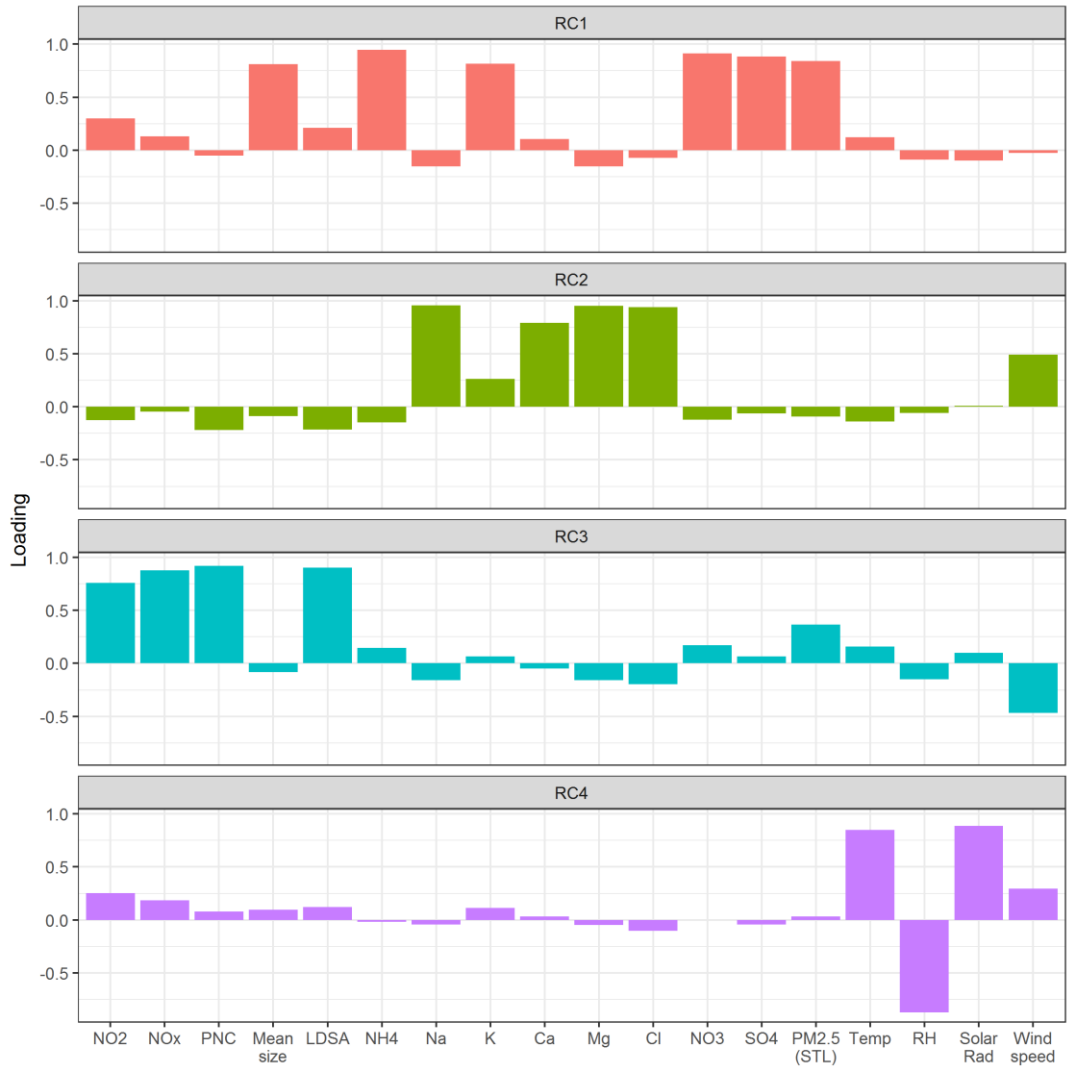
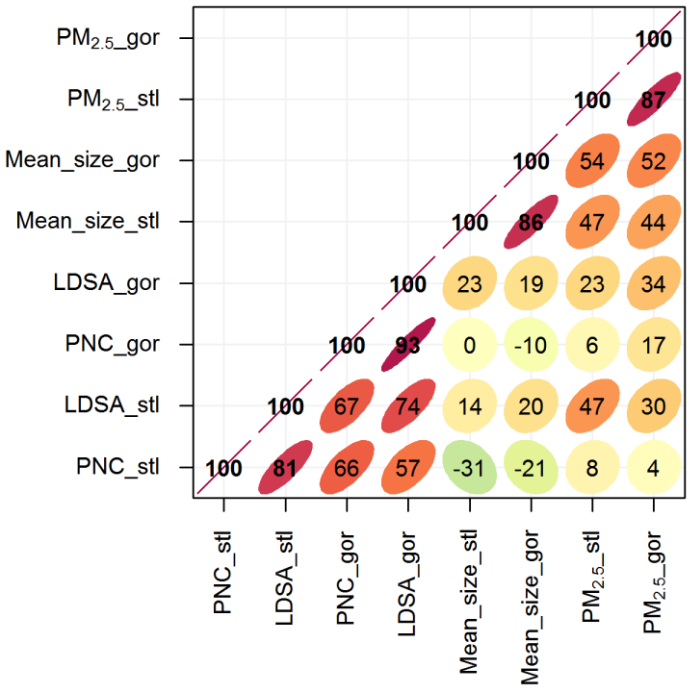



*Figure 3.18 Loadings of each pollutant and environment variables from the PCA analysis. The $NO_x$ and UFP measurements were made at (a) St Leonards and (b) Gorgie. The $PM_{2.5}$ concentrations used at both locations were measurement at St Leonards by TEOM-FMDS.*

### 3.3.2.5 Correlation of UFPs between different sites

Despite the large spatial variation in UFPNC in an urban environment, temporal variation of UFPNC at different locations in a city might be influenced by similar diurnal emission patterns and meteorological conditions. Consequently, high temporal correlation in UFPNC between monitoring sites may be observed, which reassures the use of a fixed monitoring site for estimation of population exposure to examine short-term effect of UFPNC. Pearson correlation coefficients for hourly and daily average $PM_{2.5}$, UFPNC, LDSA and mean particle size between St Leonards and Gorgie are shown in Figure 3.19. High correlations were found in hourly $PM_{2.5}$ and mean particle size between St Leonards and Gorgie ($r > 0.86$). Correlations in hourly UFPNC and LDSA between the two sites were moderate ($r = 0.66$ and 0.74, respectively). Correlations of daily average concentration between the sites improved for all pollutants, reaching $r = 0.76$ for UFPNC. The daily average is the usual averaging period used in many epidemiological studies (Atkinson et al., 2010; Breitner et al., 2009). Some previous studies demonstrated high temporal correlations of daily UFPNC (median correlation range 0.67 – 0.76) among central site and home addresses in four European cities (Puustinen et al., 2007). High correlations ($r > 0.8$) were also reported for hourly UFPNC concentration among four traffic sites in Germany (Cyrys et al., 2008). However, a range of modest correlations (median hourly correlation range 0.3 – 0.56) were found among 14 various types monitoring sites in Los Angeles (Moore et al., 2009). Overall it is demonstrated that high correlation of UFPNC can be found among certain monitoring sites within a city. However, the variability in the results highlights that these correlations are site-specific. Therefore, direct evaluation at each study location is needed by examining UFPNC measurements at multiple locations if one fixed monitoring site is used to estimate population exposure to UFPNC.

**(a)**




**(b)**

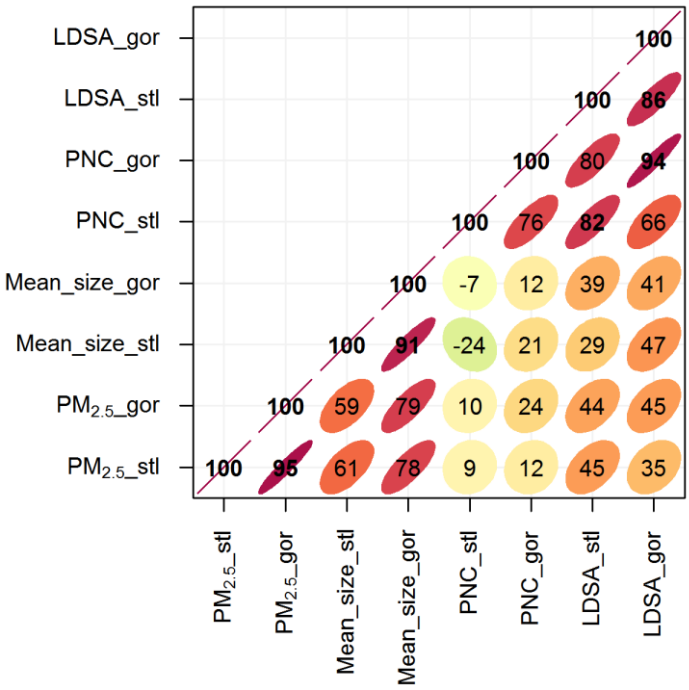



*Figure 3.19 Correlation plots between different PM metrics measured at St Leonards and Gorgie at the temporal resolutions of (a) hourly average and (b) daily average.*

## 3.4 Conclusions

In section 3.2, mobile measurements confirmed the literature that BC and UFPNC exhibited much higher (~ 3 times greater) spatial variability than $PM_{2.5}$ in the urban environment. Both BC and UFPNC were highly correlated with traffic counts, but not for the $PNC_{0.5\text{-}2.5}$. This suggests that concentrations of BC, UFP and $NO_x$ can be modelled effectively using local sources and meteorological variables. On the other hand, $PM_{2.5}$ concentrations cannot be modelled adequately using only local sources. Statistically significant difference was found between the slopes of BC vs. UFPNC regression analyses between working days and non-working days. During non-working days the reduction of HGV flows resulted in a decrease in the BC/UFPNC ratio, suggesting that HGV may contribute more to BC than to UFP concentrations. Therefore, control of HGVs may be effective in reducing the negative health effects associated with BC.

In section 3.3, long term stationary measurements revealed the temporal correlation of UFP with $NO_x$ and $PM_{2.5}$ and their dominating sources. Hourly concentrations of UFPNC and LDSA were highly correlated at both background and roadside. Both also moderately correlated with $NO_x$, but to a varying degree with $NO_2$ or NO depending the site type. PCA revealed the distinct sources of UFP and $PM_{2.5}$. This study adds to the literature that mean particles size also behaves similarly to the $PM_{2.5}$ concentrations, which are mainly governed by long-range transport of SIA mostly originated from continental Europe and England, whereas both UFPNC and LDSA were related to varying local emissions and meteorological conditions. Despite the strong linkage between UFPNC and $NO_x$, quantifying the total UFPNC of a broad size range can be challenging since the variation of the particle size distribution result in varying UFPNC vs $NO_x$ regression statistics. Good daily correlations were observed for UFPNC and LDSA ($r$ = 0.76 and 0.86, respectively) measured at a background and roadside site, which reassures the use of measurement from a fixed monitoring site in the study of short-term effect of UFP. However, given the highly locally-dependent nature of UFP, it is likely these high temporal correlations across sites may not be observe throughout a city. Therefore,

examination of the temporal variation of UFP at multiple locations directly in a health study area is necessary.

# References

AQEG, 2012. Fine Particulate Matter (PM2.5) in the United Kingdom. Defra, UK. URL https://uk-air.defra.gov.uk/library/reports?report_id=727

Atkinson, R.W., Fuller, G.W., Anderson, H.R., Harrison, R.M., Armstrong, B., 2010. Urban Ambient Particle Metrics and Health: A Time-series Analysis. Epidemiology 21, 501–511. doi:10.1097/EDE.0b013e3181debc88

Ayers, G.P., 2001. Comment on regression analysis of air quality data. Atmos. Environ. 35, 2423–2425. doi:10.1016/S1352-2310(00)00527-6

Bau, S., Jacoby, J., Witschger, O., 2012. Evaluation of the diffusion size classifier (meDiSC) for the real-time measurement of particle size and number concentration of nanoaerosols in the range 20–700 nm. J. Environ. Monit. 14, 1014–1023. doi:10.1039/C2EM10619A

Breitner, S., Stölzel, M., Cyrys, J., Pitz, M., Wölke, G., Kreyling, W., Küchenhoff, H., Heinrich, J., Wichmann, H.-E., Peters, A., 2009. Short-Term Mortality Rates during a Decade of Improved Air Quality in Erfurt, Germany. Environ. Health Perspect. 117, 448–454. doi:10.1289/ehp.11711

Carslaw, D.C., 2016. The openair Manual. URL https://www.dropbox.com/s/2n7wdyursdul8dk/openairManual.pdf?dl=0 (accessed 19-June-2015).

Carslaw, D.C., Ropkins, K., 2012. openair — An R package for air quality data analysis. Environ. Model. Softw. 27–28, 52–61. doi:10.1016/j.envsoft.2011.09.008

Charron, A., Birmili, W., Harrison, R.M., 2007. Factors influencing new particle formation at the rural site, Harwell, United Kingdom. J. Geophys. Res. Atmospheres 112, D14210. doi:10.1029/2007JD008425

Cyrys, J., Pitz, M., Heinrich, J., Wichmann, H.-E., Peters, A., 2008. Spatial and temporal variation of particle number concentration in Augsburg, Germany. Sci. Total Environ. 401, 168–175. doi:10.1016/j.scitotenv.2008.03.043

Dall'Osto, M., Querol, X., Alastuey, A., O'Dowd, C., Harrison, R.M., Wenger, J., Gómez-Moreno, F.J., 2013. On the spatial distribution and evolution of ultrafine particles in Barcelona. Atmos Chem Phys 13, 741–759. doi:10.5194/acp-13-741-2013

DfT, 2015. City of Edinburgh [WWW Document]. Dep. Transp. - Traffic Counts. URL http://www.dft.gov.uk/traffic-counts/cp.php?la=City+of+Edinburgh (accessed 11-March-2015).

Fierz, M., Houle, C., Steigmeier, P., Burtscher, H., 2011. Design, Calibration, and Field Performance of a Miniature Diffusion Size Classifier. Aerosol Sci. Technol. 45, 1–10. doi:10.1080/02786826.2010.516283

Grahame, T.J., Klemm, R., Schlesinger, R.B., 2014. Public health and components of particulate matter: The changing assessment of black carbon. J. Air Waste Manag. Assoc. 64, 620–660. doi:10.1080/10962247.2014.912692

Hagler, G.S.W., Yelverton, T.L.B., Vedantham, R., Hansen, A.D.A., Turner, J.R., 2011. Post-processing Method to Reduce Noise while Preserving High Time Resolution in Aethalometer Real-time Black Carbon Data. Aerosol Air Qual. Res. 11, 539–546. doi:10.4209/aaqr.2011.05.0055

Harrison, R.M., Laxen, D., Moorcroft, S., Laxen, K., 2012. Processes affecting concentrations of fine particulate matter (PM2.5) in the UK atmosphere. Atmos. Environ. 46, 115–124. doi:10.1016/j.atmosenv.2011.10.028

Heal, M.R., Kumar, P., Harrison, R.M., 2012. Particles, air quality, policy and health. Chem. Soc. Rev. 41, 6606. doi:10.1039/c2cs35076a

HEI, 2013. Understanding the Health Effects of Ambient Ultrafine Particles. Health Effects Institute, Boston, Massachusetts. URL https://www.healtheffects.org/system/files/Perspectives3.pdf

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. Atmos. Environ. 42, 7561–7578. doi:10.1016/j.atmosenv.2008.05.057

Hoek, G., Meliefste, K., Cyrys, J., Lewné, M., Bellander, T., Brauer, M., Fischer, P., Gehring, U., Heinrich, J., van Vliet, P., Brunekreef, B., 2002. Spatial variability of fine particle concentrations in three European areas. Atmos. Environ. 36, 4077–4088. doi:10.1016/S1352-2310(02)00297-2

Kabacoff, R., 2011. R in action: data analysis and graphics with R. Manning, Shelter Island, NY.

Knibbs, L.D., Cole-Hunter, T., Morawska, L., 2011. A review of commuter exposure to ultrafine particles and its health effects. Atmos. Environ. 45, 2611–2622. doi:10.1016/j.atmosenv.2011.02.065

Kumar, P., Fennell, P., Britter, R., 2008. Measurements of particles in the 5-1000 nm range close to road level in an urban street canyon. Sci. Total Environ. 390, 437–447. doi:10.1016/j.scitotenv.2007.10.013

Kumar, P., Robins, A., Vardoulakis, S., Britter, R., 2010. A review of the characteristics of nanoparticles in the urban atmosphere and the prospects for developing regulatory controls. Atmos. Environ. 44, 5035–5052. doi:10.1016/j.atmosenv.2010.08.016

Ma, N., Birmili, W., 2015. Estimating the contribution of photochemical particle formation to ultrafine particle number averages in an urban atmosphere. Sci. Total Environ. 512–513, 154–166. doi:10.1016/j.scitotenv.2015.01.009

Moore, K., Krudysz, M., Pakbin, P., Hudda, N., Sioutas, C., 2009. Intra-Community Variability in Total Particle Number Concentrations in the San Pedro Harbor Area (Los Angeles, California). Aerosol Sci. Technol. 43, 587–603. doi:10.1080/02786820902800900

Oberdörster, G., Oberdörster, E., Oberdörster, J., 2005. Nanotoxicology: An Emerging Discipline Evolving from Studies of Ultrafine Particles. Environ. Health Perspect. 113, 823–839. doi:10.1289/ehp.7339

Peters, J., Van den Bossche, J., Reggente, M., Van Poppel, M., De Baets, B., Theunis, J., 2014. Cyclist exposure to UFP and BC on urban routes in Antwerp, Belgium. Atmos. Environ. 92, 31–43. doi:10.1016/j.atmosenv.2014.03.039

Price, H.D., Arthur, R., BéruBé, K.A., Jones, T.P., 2014. Linking particle number concentration (UFPNC), meteorology and traffic variables in a UK street canyon. Atmospheric Res. 147–148, 133–144. doi:10.1016/j.atmosres.2014.05.008

Puustinen, A., Hämeri, K., Pekkanen, J., Kulmala, M., de Hartog, J., Meliefste, K., ten Brink, H., Kos, G., Katsouyanni, K., Karakatsani, A., Kotronarou, A.,

Kavouras, I., Meddings, C., Thomas, S., Harrison, R., Ayres, J.G., van der Zee, S., Hoek, G., 2007. Spatial variation of particle number and mass over four European cities. Atmos. Environ. 41, 6622–6636. doi:10.1016/j.atmosenv.2007.04.020

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rakowska, A., Wong, K.C., Townsend, T., Chan, K.L., Westerdahl, D., Ng, S., Močnik, G., Drinovec, L., Ning, Z., 2014. Impact of traffic volume and composition on the air quality and pedestrian exposure in urban street canyon. Atmos. Environ. 98, 260–270. doi:10.1016/j.atmosenv.2014.08.073

Reche, C., Querol, X., Alastuey, A., Viana, M., Pey, J., Moreno, T., Rodríguez, S., González, Y., Fernández-Camacho, R., de la Campa, A.M.S., de la Rosa, J., Dall'Osto, M., Prévôt, A.S.H., Hueglin, C., Harrison, R.M., Quincey, P., 2011. New considerations for PM, Black Carbon and particle number concentration for air quality monitoring across different European cities. Atmospheric Chem. Phys. 11, 6207–6227. doi:10.5194/acp-11-6207-2011

Skyllakou, K., Murphy, B.N., Megaritis, A.G., Fountoukis, C., Pandis, S.N., 2014. Contributions of local and regional sources to fine PM in the megacity of Paris. Atmos Chem Phys 14, 2343–2352. doi:10.5194/acp-14-2343-2014

Steinle, S., Reis, S., Sabel, C.E., 2013. Quantifying human exposure to air pollution -- Moving from static monitoring to spatio-temporally resolved personal exposure assessment. Sci. Total Environ. 184. doi:10.1016/j.scitotenv.2012.10.098

Steinle, S., Reis, S., Sabel, C.E., Semple, S., Twigg, M.M., Braban, C.F., Leeson, S.R., Heal, M.R., Harrison, D., Lin, C., Wu, H., 2015. Personal exposure monitoring of PM2.5 in indoor and outdoor microenvironments. Sci. Total Environ. 508, 383–394. doi:10.1016/j.scitotenv.2014.12.003

Sullivan, R.C., Pryor, S.C., 2014. Quantifying spatiotemporal variability of fine particles in an urban environment using combined fixed and mobile measurements. Atmos. Environ. 89, 664–671. doi:10.1016/j.atmosenv.2014.03.007

Thorpe, A., Harrison, R.M., 2008. Sources and properties of non-exhaust particulate matter from road traffic: A review. Sci. Total Environ. 400, 270–282. doi:10.1016/j.scitotenv.2008.06.007

Twigg, M.M., Di Marco, C.F., Leeson, S., van Dijk, N., Jones, M.R., Leith, I.D., Morrison, E., Coyle, M., Proost, R., Peeters, A.N.M., Lemon, E., Frelink, T., Braban, C.F., Nemitz, E., Cape, J.N., 2015. Water soluble aerosols and gases at a UK background site – Part 1: Controls of PM2.5 and PM10 aerosol composition. Atmos Chem Phys 15, 8131–8145. doi:10.5194/acp-15-8131-2015

Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J., De Baets, B., 2015. Mobile monitoring for mapping spatial variation in urban air quality: Development and validation of a methodology based on an extensive dataset. Atmos. Environ. 105, 148–161. doi:10.1016/j.atmosenv.2015.01.017

Viana, M., Kuhlbusch, T.A.J., Querol, X., Alastuey, A., Harrison, R.M., Hopke, P.K., Winiwarter, W., Vallius, M., Szidat, S., Prévôt, A.S.H., Hueglin, C., Bloemen, H., Wåhlin, P., Vecchi, R., Miranda, A.I., Kasper-Giebl, A., Maenhaut, W., Hitzenberger, R., 2008. Source apportionment of particulate

matter in Europe: A review of methods and results. J. Aerosol Sci. 39, 827–849. doi:10.1016/j.jaerosci.2008.05.007

Vieno, M., Heal, M.R., Hallsworth, S., Famulari, D., Doherty, R.M., Dore, A.J., Tang, Y.S., Braban, C.F., Leaver, D., Sutton, M.A., Reis, S., 2014. The role of long-range transport and domestic emissions in determining atmospheric secondary inorganic particle concentrations across the UK. Atmos Chem Phys 14, 8435–8447. doi:10.5194/acp-14-8435-2014

Warton, D.I., Wright, I.J., Falster, D.S., Westoby, M., 2006. Bivariate line-fitting methods for allometry. Biol. Rev. 81, 259–291. doi:10.1017/S1464793106007007

WHO, 2013. Review of evidence on health aspects of air pollution – REVIHAAP project: final technical report. WHO, Copenhagen, Denmark. URL http://www.euro.who.int/en/health-topics/environment-and-health/air-quality/publications/2013/review-of-evidence-on-health-aspects-of-air-pollution-revihaap-project-final-technical-report

# Chapter 4 Modelling of $NO_2$ and $O_3$ concentrations with ADMS-Urban

## 4.1 Introduction

Assessing population exposure to pollutants with high spatial variability such as $NO_2$ requires knowledge of the pollutant concentration at fine scales. Air quality measurements alone cannot provide the spatial resolution needed in health studies. Thus air quality models are being increasingly used to quantify human exposure to outdoor air pollution (Cai et al., 2014; Laurent et al., 2008; Maheswaran et al., 2010). These models can be roughly categorised into statistical and deterministic dispersion models. A statistical model relies on measurement data from a number of monitoring sites, upon which an algorithm is derived to characterise the spatial pattern of the pollution observed. Among the different statistical models, land-use regression (LUR) models have gained increasing popularity for the assessment of long-term exposure to air pollution (Hoek et al., 2008). A deterministic dispersion model describes the fundamental physical-chemical processes taking placing in the atmosphere by solving mathematical equations. The comprehensive coverage of dispersion models in space and time enables the incorporation of human time-activity data potentially providing more accurate personal exposure assessment that may significantly reduce exposure misclassification (Smith et al., 2016).

In this work a Gaussian dispersion model (ADMS-Urban) is used to model $NO_2$ and $O_3$ concentrations in Edinburgh. ADMS-Urban is an air quality modelling software developed by Cambridge Environmental Research Consultants (CERC) and is widely used in Europe for regulatory applications (CERC, 2016a). Although being a proprietary model, details of the dispersion calculation algorithms can be found on the CERC website (CERC, 2016b) and have been validated in various publications (CERC, 2016c). However, the performance of dispersion models in a particular area can be affected by uncertainties from a number of sources, most importantly from input data such as emission, meteorology and topology, but also from the model settings configured by the user. In this chapter available emission and meteorological

data for Edinburgh are discussed and their influences on ADMS-Urban output are investigated. Validation of the ADMS-Urban model for Edinburgh is based on both passive diffusion tube (PDT) and reference analyser data from automatic monitoring networks. The aim of this chapter is to determine the optimal setup of ADMS-Urban for modelling $NO_2$ and $O_3$ in Edinburgh (and possible for other UK cities) and to provide a basis for subsequent application of ADMS-Urban for the purpose of improving exposure assessment methodology, as shown in Chapter 5.

## 4.2 Input data for modelling

The availability and consistency of input data (traffic flow, emission inventory and meteorology) have been a major limiting factor for the application of dispersion models among different cities in health studies (ESCAPE, 2010). Although this section only discusses the available emission and meteorological data in Edinburgh, the general data availability is similar for most UK cities.

### 4.2.1 Emissions

The Department for Transport (DfT) provides annual average daily flow (AADF) of traffic in different categories for all the major roads in UK cities (DfT, 2015). This enables the calculation of emissions on the major roads provided that the average vehicle speed is known. The National Atmospheric Emission Inventory (NAEI) provides aggregated emissions for traffic and other sources at 1 $km^2$ grid resolution for the whole UK (NAEI, 2015). For the emissions from major roads, NAEI values are also based on traffic data from DfT (Defra, 2015). Traffic flow data are only available for a few minor roads in a city. Where these data are available, the minor road emission for that particular road link is calculated in NAEI. For the majority of minor roads, traffic flows are modelled based on average regional flows and fleet mix (Defra, 2015). Regional average flows are available from DfT by vehicle types. NAEI assign different vehicle types to different minor roads – B and C roads or unclassified roads. Applying NAEI emissions to model road sources requires disaggregating the gridded emissions and attributing them to the road segments, which results in all the emissions for the same road type (major or minor) being the same inside a grid. Although this method of calculating road emissions may be less

accurate than a bottom-up method based on traffic flows and emission factors, the emissions from the minor roads can at least be modelled as road sources rather than as a 1 $km^2$ grid source. The following sensitivity studies discuss the difference in the model outputs based on the two different road emission setups and some practical considerations in using each method.

### 4.2.2 Meteorology

Surface meteorological observations are provided by NOAA Integrated Surface Database (ISD) across the globe (NOAA, 2016). The comprehensive coverage of ISD stations in the UK ensures almost every city has a meteorological station nearby. The closest station to Edinburgh is located at Gogarbank (55° 55' 58.8" N, 3° 21' 00.0" W), a rural area west of the conurbation. As the wind measured at this site approaches the built-up urban area, wind speed will reduce within the building canopy. ADMS-Urban takes account of the change of wind speed by considering the mean height of the buildings in a neighbourhood (CERC, 2016a). However, the wind direction is unchanged at the dispersion site from the input meteorological measurements. The measurement at Gogarbank may represent the general weather condition of the whole city but may be less representative for a specific site within the city. For comparison, a pocket weather sensor Kestrel 4500 was mounted on the roof of the St Leonards AURN monitoring station to measure local wind speed and direction. The St Leonards station is located in a small urban park, with the closest road 35 m to the east (Figure 4.1). There are no high-rise buildings in the vicinity of the station.

In the following analysis, modelled $NO_2$ concentrations based on Gogarbank or St Leonards weather data are compared with measured $NO_2$ at St Leonards to assess whether using local meteorological measurements can improve the model performance.


*Figure 4.1 Siting of the St Leonards station with respect to the surrounding environment.*

## 4.3 Sensitivity tests

### 4.3.1 Major road emissions – NAEI gridded emission vs DfT traffic data

Major road emissions can be calculated by using the NAEI 1 $km^2$ gridded emission (top-down approach) or from traffic data provided by DfT (bottom-up approach). In this test the emissions from major roads resulting from the application of the two methods were compared.

In the top-down approach, $NO_x$ emission for each 1 $km^2$ grid cell from the major road sector in 2012 was divided by the total length of major roads in each grid to give major road emissions in g/m/s. In the bottom-up approach, major roads with traffic data were imported to the EMIT software where the emissions were calculated based on the NAEI 2012 emission inventory. The traffic data and emission inventory should be the same in both approaches. However, there are discrepancies in the vehicle classification and average vehicle speed between the two emissions

calculation methods. In the NAEI gridded emissions, traffic data are classified into 6 categories (cars, LGVs, buses, motorcycles, rigid HGVs and articulated HGVs) as opposed to 3 categories used in EMIT (LGVs, HGVs and motorcycles). The average speed used in EMIT was taken as the speed limit (48 km/h), whereas the NAEI gridded emission assigned different vehicle speeds to different road types (Defra, 2015). The discrepancy in the resulting major road emissions is therefore likely due to the combination of these two differences.

Figure 4.2 shows the difference in $NO_x$ emissions for the major roads estimated by the top-down and bottom-up approaches. There was great spatial heterogeneity in the difference of NOx emissions especially within a grid cell where roads with varying traffic flows intersected. This is anticipated since the traffic data do not have gridded resolution. However even within grids with only one road segment, the NOx emissions estimated by the top-down approach were generally higher than the bottom-up approach using an average vehicle speed of 48 km/h (Figure 4.2a). Adjusting the average speed used in EMIT to 32 km/h increased the NOx emissions estimated by the bottom-up approach and gave better agreement with the NAEI gridded NOx emissions (Figure 4.2b).

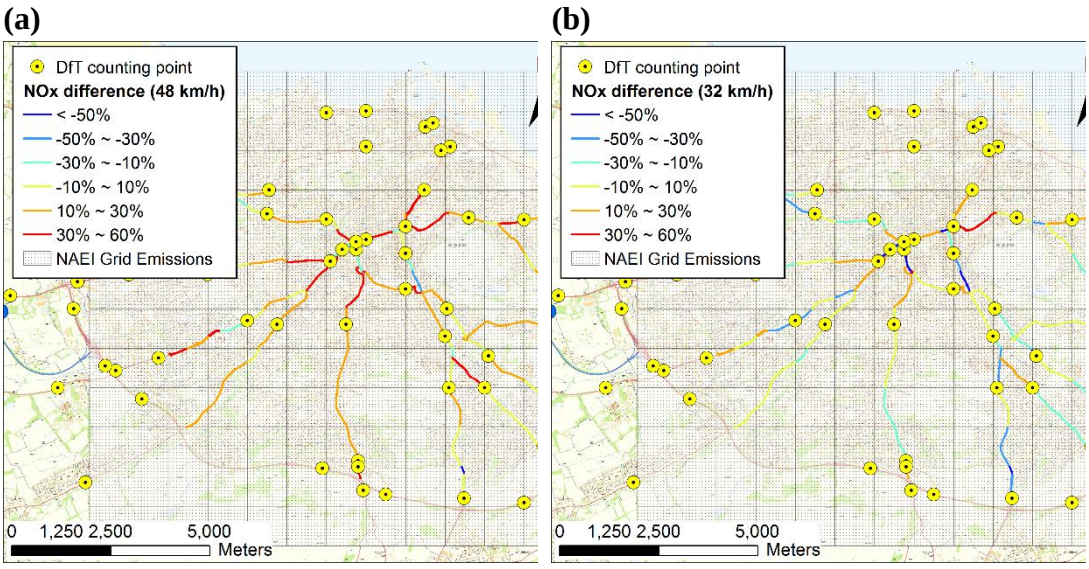



*Figure 4.2 Differences between NOx emissions estimated by the top-down and bottom-up approaches with average vehicle speed for the latter fixed at either 48 km/h (a) or 32 km/h (b). The difference was expressed as (top-down emission – bottom-up emission)/top-down emission.*

The effect of using different major road emission calculation methods on the modelled NOx concentrations was compared in Figure 4.3. Receptor points were regularly spaced at 200 m and also along the edge of the major roads and at distances from the road centreline two times the width of the road. Although the difference in the major road emissions calculated from the two methods is large for some road segments (over 30%, Figure 4.2a), the resulting modelled concentration difference reflects the emissions difference only at locations very close to the road sources (Figure 4.3). For the majority of the locations away from major roads (> 50 m), the difference in the modelled NOx was small (on average 6%, 1% and 15% for 5$^{th}$ and 95$^{th}$ percentile respectively). Despite the discrepancy in the major road emissions calculated based on these two methods, the difference in modelled pollutant concentrations will be most pronounced at short distance from the major roads.

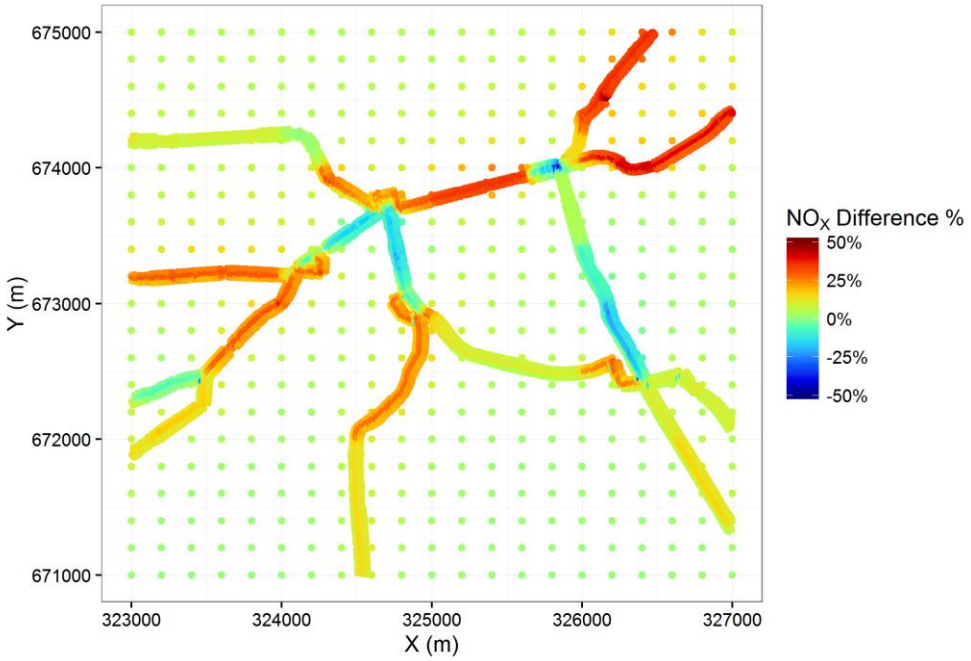



*Figure 4.3 Difference in modelled NOx concentrations between top-down and bottom-up major road emission calculation methods. The difference was expressed as (top-down – bottom-up)/top-down. In the bottom-up method an average speed of 48 km/h was applied to all traffic.*

### 4.3.2 Minor road emissions – Modelled as grid source vs road source

Traffic data on minor roads are not available in most UK cities. Modelling emissions from minor roads can either rely on the 1 $km^2$ gridded emission from NAEI or using traffic models to assign regional average traffic flow to the minor roads. Without accessing a traffic model, the former approach is investigated here. Figure 4.4 shows the ratio of NOx emission from minor roads to major roads for each 1 $km^2$ grid in Edinburgh. Within the city boundary the emissions from minor roads can be twice as high or comparable to the emissions from major roads for many 1 $km^2$ grids. This indicates the importance of including minor road emissions in the model. ADMS-Urban allows such aggregated emission to be modelled as a grid source. However, a more accurate way is to disaggregate the gridded minor road emissions to the minor roads and to model as a road source. Despite being more representative of reality there is a caveat with this approach. The minor roads defined in NAEI not only include the relatively busy B roads but also the less busy residential streets (Figure

4.5). Due to the limitation on the number of roads that can be modelled explicitly in ADMS-Urban only the B roads and the major roads can be modelled as a road source. Emission on B roads were calculated by dividing the total minor road emission in a grid by the total length of minor roads (B roads + residential streets) in that grid under the assumption that B roads have the same emission per unit length as the residential streets. As a result, the emissions on the B roads were likely to be underestimated. However, modelling minor roads as a road source add more spatial variability in the emissions as opposed to modelling them as a grid source.

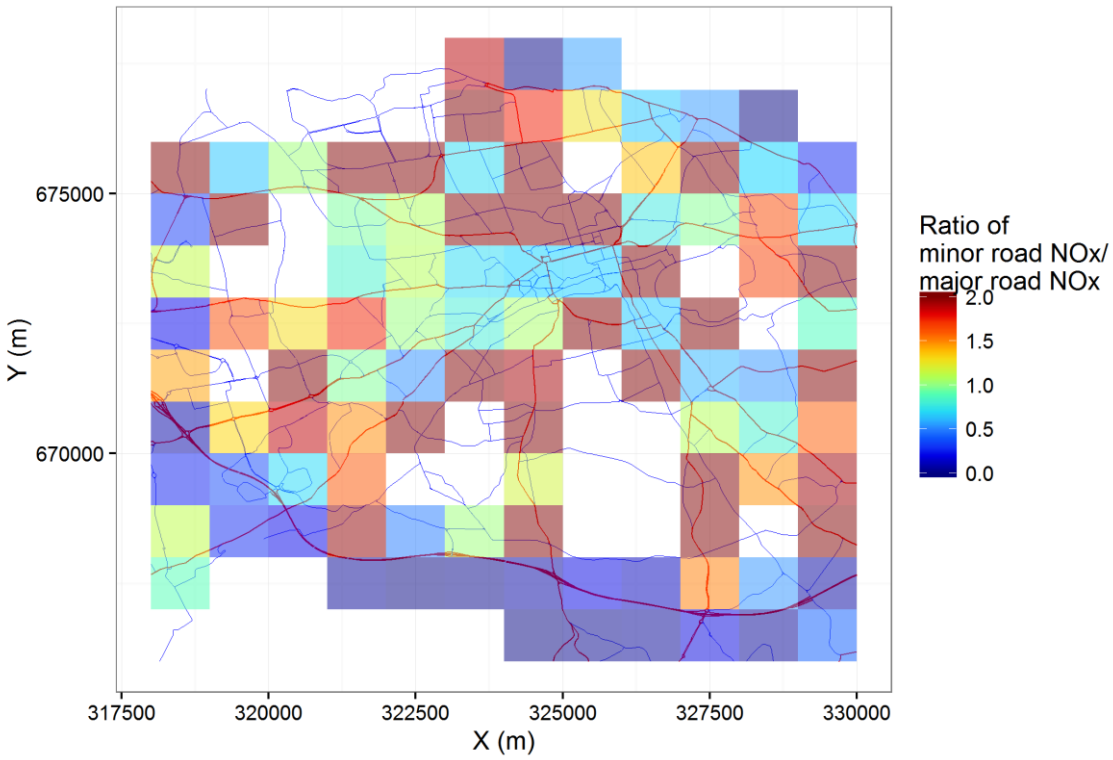



*Figure 4.4 Ratio of minor road to major road NOx emissions in 1 km$^2$ grids. Grids without a colour means that there are no major roads in that grid.*

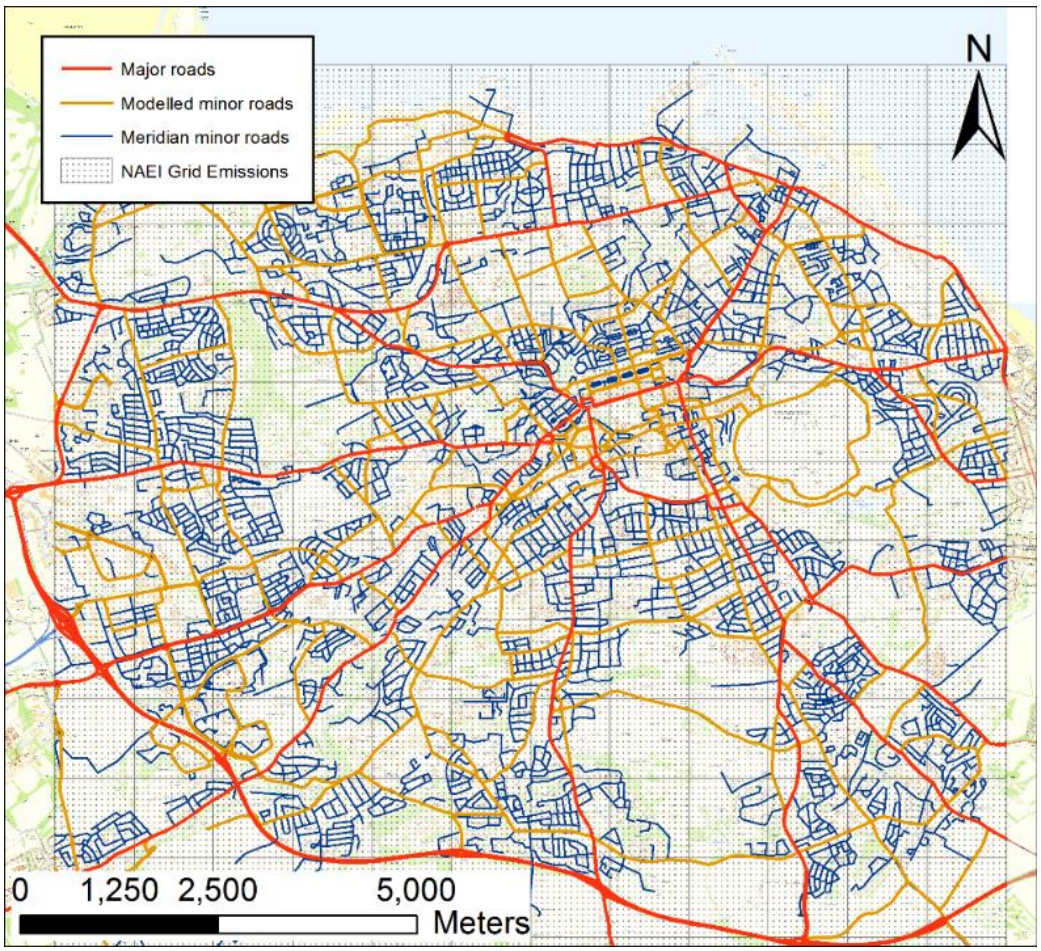



*Figure 4.5 Definition of major and minor roads in the NAEI and in this model study. The major road definition is the same between NAEI and this study. The minor roads modelled in this study only include B roads (orange), whereas in NAEI the minor roads not only include B roads but also most residential streets (blue).*

The differences in $NO_2$ concentrations resulting from the different modelling methods for minor roads are shown in Figure 4.6a. As expected, modelling minor roads explicitly resulted in higher concentrations near minor roads. However at most locations $NO_2$ concentrations were lower when minor roads were modelled as a road source because all the minor road emission dispersed as an area source in the grid source method. The difference in the modelled $NO_2$ concentration can be small between the two methods if the total minor road emission in a grid is relatively low. Large discrepancies can be expected between the two methods when the total minor road emission in a grid is high, as illustrated in the top-left grid of the output area (Figure 4.6).

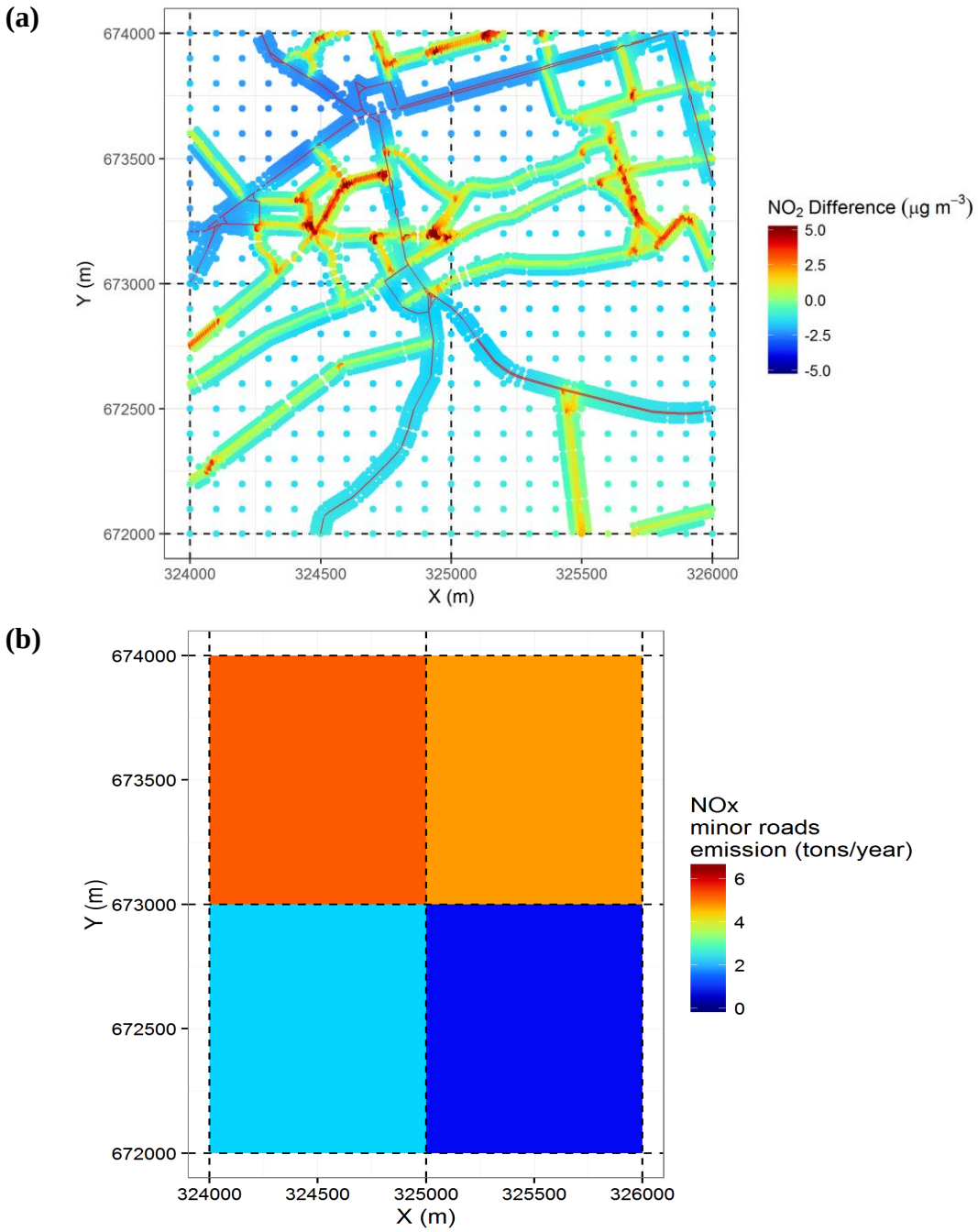



*Figure 4.6 (a) Differences in $NO_2$ concentrations between the two minor road modelling methods. The difference was expressed as $NO_2$[minor roads as road source] – $NO_2$[minor roads as grid source]. (b) The minor road NOx emissions for the four 1 km$^2$ grids corresponding to the output area in Figure 5a.*

### 4.3.3 Meteorological data – Gogarbank vs local measurements

It is standard practice to use meteorological data from a station close to, but outside, the urban area to represent the city-scale meteorology. However meteorological measurements at the dispersion site might be more suitable for modelling at a specific location. This test compares $NO_2$ concentrations at St Leonards modelled with wind speed and direction measured at St Leonards and at Gogarbank, which is situated in a rural area 10 km west of St Leonards.

In the model using Gogarbank meteorological data, urban canopy flow was enabled in order to better estimate the change of wind speed at surface relative to the measurement height at 10 m. In the model using meteorological data from St Leonards, urban canopy flow was disabled to ensure the meteorological conditions considered by the model at the receptor site were the same as the input data. Figure 4.7 compares the distributions of hourly wind speed and direction profiles from 2015-06-09 to 2015-08-21 measured with the Kestrel wind monitor at St Leonards and that simulated by ADMS-Urban based on measurements at the fixed-site automated weather station at Gogarbank.

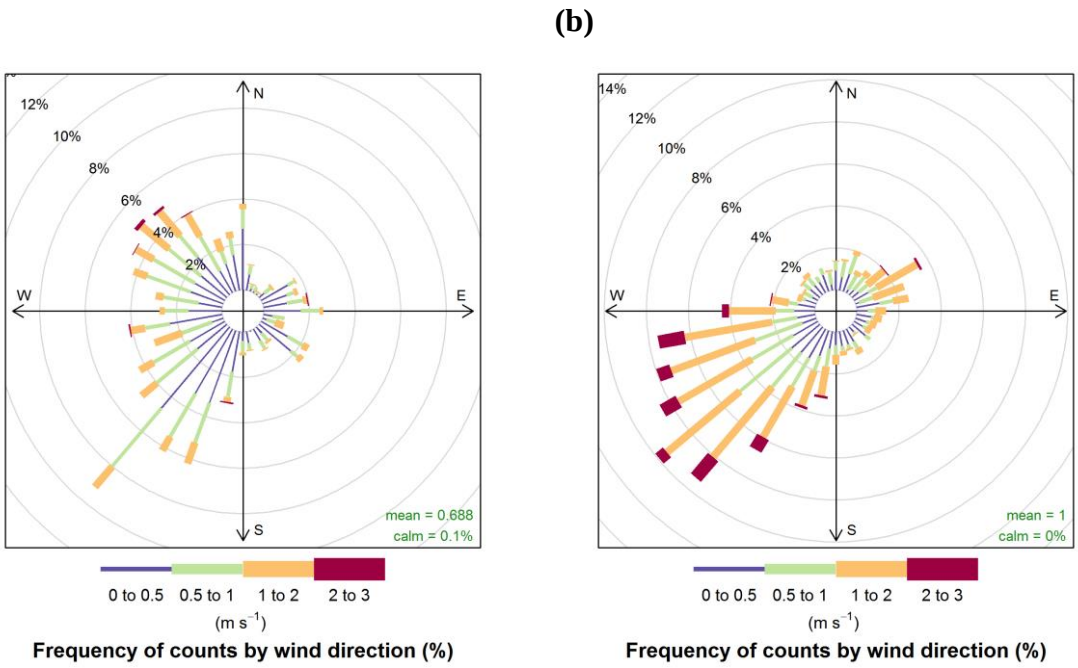



*Figure 4.7 Comparison of hourly wind profiles from 2015-06-09 to 2015-08-21 at St Leonards from direct measurement by portable Kestrel wind monitor (a) and from processing by ADMS-Urban based on measurements by automatic weather station at Gogarbank (b).*

Wind direction at Gogarbank was predominately from the southwest, whereas the local wind direction at St Leonards also had a strong component from the northwest. Despite the fact that ADMS-Urban takes into account the decreased wind speed at surface level, the wind speed estimated by the model was still on average higher than the measurements at St Leonards. A great proportion of wind speed measurements at St Leonards were below the threshold that ADMS-Urban requires for dispersion calculation. As a result wind speed below 0.75 m/s at 10 m (equivalent to 0.43 m/s at 3 m) was set to 0.75 m/s and the wind direction was set to the previous hour's, which comprised 40% of the data. The effect of these contrasting wind profiles on the modelled $NO_2$ concentrations at St Leonards are compared in Figure 4.8. As suggested in Figure 4.7, the generally lower measured wind speed at St Leonards resulted in a great number of modelled $NO_2$ concentrations higher than those modelled with Gogarbank wind data. Wind direction was not as influential as wind speed in the modelled $NO_2$ concentration at St Leonards background site. However it

is anticipated that the influence of wind direction could be larger when the receptor is closer to the road source.

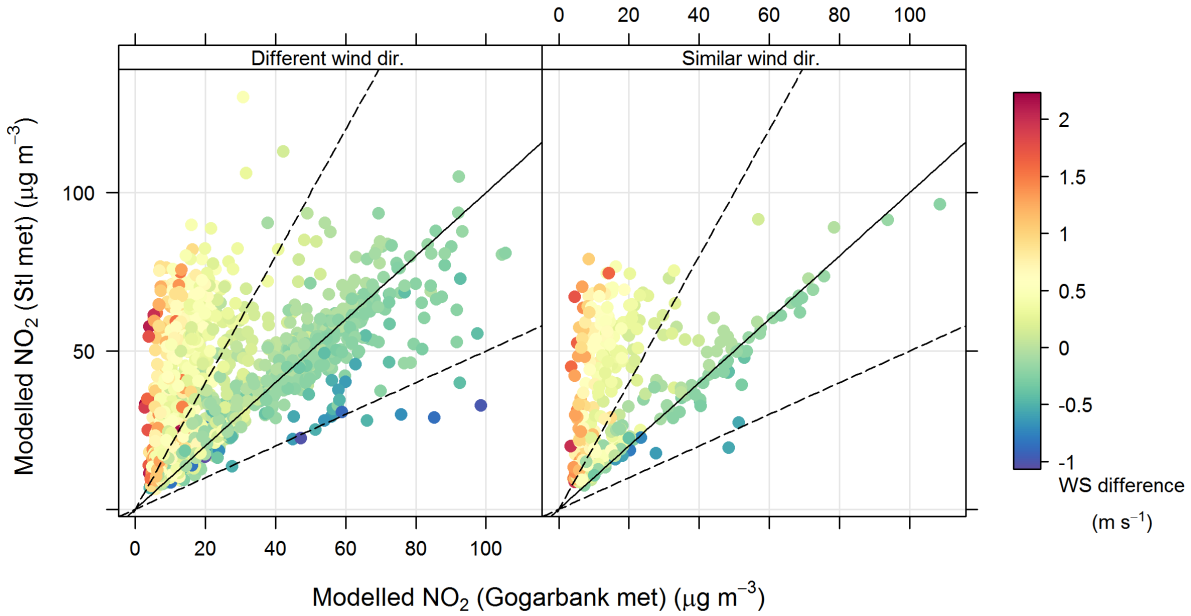



*Figure 4.8 Scatter plot of hourly $NO_2$ concentrations at St Leonards modelled with meteorological data measured at St Leonards and with meteorological data measured at Gogarbank. The two panels split the data according to whether wind direction in the two approaches was similar or different and the data points are coloured by the difference in wind speed. Similar wind direction refers to a difference in wind direction of less than $22.5^0$. The wind speed difference was calculated as the wind speed estimated from Gogarbank minus the wind speed measured at St Leonards. The solid line in each panel is the 1:1 line and the dashed lines are the 2:1 and 1:2 lines.*

To investigate whether using local meteorological data improves the accuracy of modelled concentration, modelled $NO_2$ concentrations using Gogarbank and St Leonards meteorological data are compared with measurements in Figure 4.9. The model using Gogarbank data captured well the variation in daily-average $NO_2$ due to varying meteorological conditions. However the model using local St Leonards data systematically overestimated $NO_2$ concentration as a result of the frequently measured low wind speed. The reason for the poor model performance using St Leonards meteorological data may be explained by the wind measurements not being representative of the local-scale climate. The wind sensor was mounted right on top

of the monitoring station about 3 m above ground, which was likely affected by the disturbance induced by the bulk of the station. Given that the St Leonards station was at a background location, wind measurements representing more general local-scale conditions may be more beneficial for the dispersion calculation. The results here suggest that careful selection of meteorological measurement site in the urban environment is essential for dispersion modelling. The World Meteorological Organisation has recommended measurement of wind speed/direction above the roughness sublayer (typically on tall towers) to monitor the urban local scale climate (Oke, 2006).

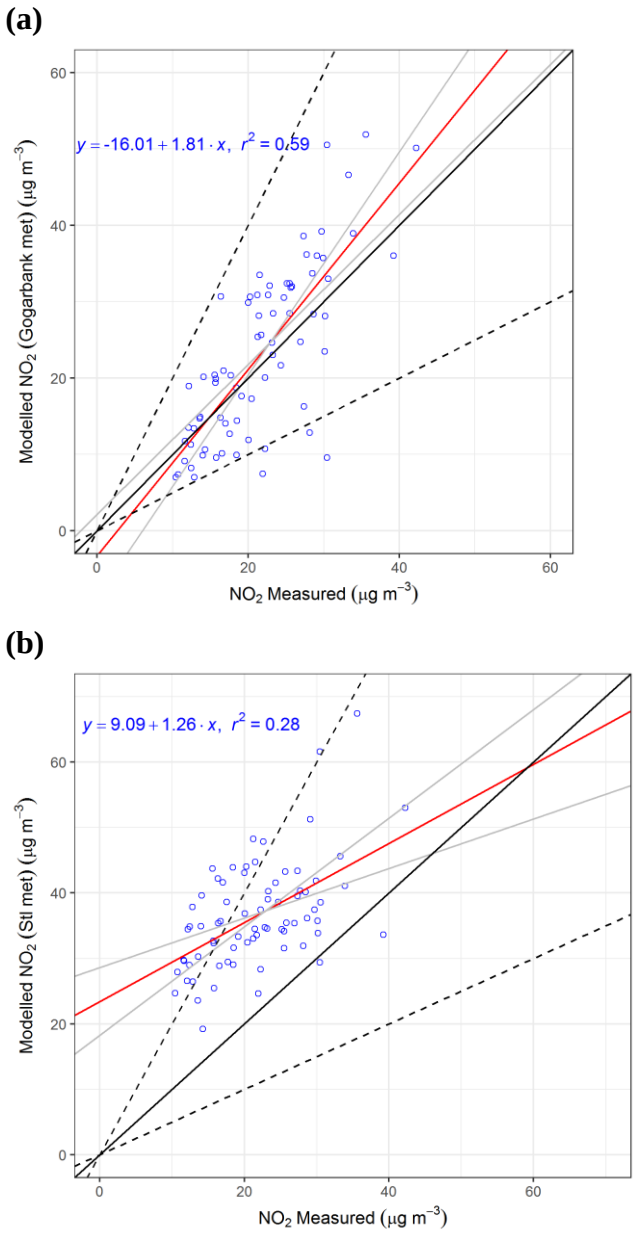



*Figure 4.9 Daily average $NO_2$ concentrations at St Leonards modelled with Gogarbank (a) and St Leonards (b) meteorological data compared with the St Leonards reference $NO_2$ measurements for the period 2015-06-09 ~ 2015-08-21. The red and grey lines show the linear fit and 95% confidence interval. The solid black line shows the 1:1 relationship and the dashed lines show 1:2 and 2:1 relationship.*

## 4.4 ADMS-Urban model validation

### 4.4.1 Model setup

Based on the results from the above sensitivity studies, the following model setups were adopted and illustrated in Figure 4.10 as the base model for the subsequent model validation:

- A and B roads were explicitly modelled as road sources;
- emissions for the road sources were calculated from the NAEI 2012 gridded emissions;
- the remainder emissions were modelled as 1 $km^2$ grid sources;
- meteorological data were taken from Gogarbank measurements;
- background concentrations were taken from measurements at Bush Estate.

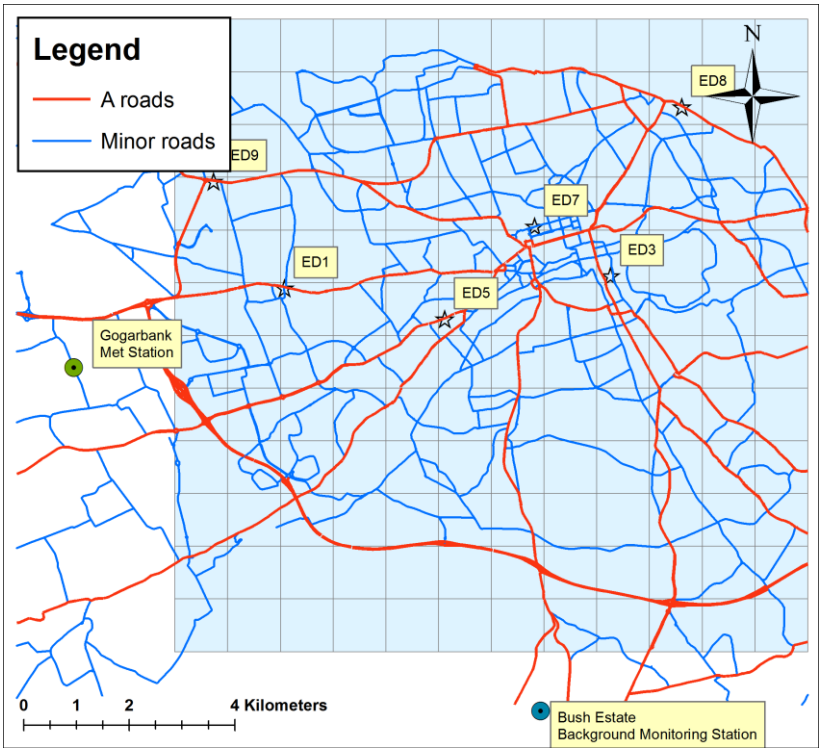



*Figure 4.10 Map of the modelling domain (12 km ×12 km) for Edinburgh showing the explicitly modelled road sources and grid sources, and the locations of the meteorological measurement station, the rural background monitoring station (Bush Estate) and the automatic monitoring stations (EDx). ED3 is the St Leonards AURN monitoring site.*

### 4.4.2 Validation against reference instruments

There are 6 automatic monitoring stations measuring NOx in Edinburgh, only one of which also measures $O_3$ (the St Leonards AURN station marked as ED3 in Figure 4.10). The measurement of real-time air pollutant concentrations from these stations allows the validation of ADMS-Urban on hourly or daily basis. Validation of modelled hourly concentration can be challenging, since hourly modelled concentrations are subject to uncertainty in the time-varying emissions and also to uncertainty in the hourly meteorological data. Validation of daily average concentration may rely less on the exact timing of emissions and reflects more of the temporal aspect of the model capability driven by the varying meteorology. Daily average concentration is also the metric used in most time-series studies quantifying the short-term effect of air pollution (Moolgavkar et al., 2013; Samoli et al., 2016). In this study, modelled daily and monthly average $NO_2$ and $O_3$ concentrations were compared with the measurement at the automatic monitoring sites for 2012. Detailed descriptions of the monitoring sites are summarised in Table 4.1.

*Table 4.1 Characteristics of the automatic monitoring sites used in model validation. The locations of these sites are marked on Figure 4.10. ED3 is the St Leonards AURN monitoring site.*

| Site | Type | Distance to nearest road (m) | Pollutant(s) measured | Comments |
|---|---|---|---|---|
| **ED1** | Kerbside | 0.5 | $NO_2$ and NOx | Inside a street canyon, at the intersection of a busy T junction. |
| **ED3** | Background | 35 | $NO_2$, NOx and $O_3$ | In a small urban park. |
| **ED5** | Roadside | 1.5 | $NO_2$ and NOx | South of a busy A road. Buildings on the north and southwest sides of the station. |
| **ED7** | Roadside | 4.5 | $NO_2$ and NOx | North of a busy B road. Buildings on the northwest and south sides of the station. |
| **ED8** | Roadside | 1.5 | $NO_2$ and NOx | South of an A road near a junction. Buildings on the south side and a wall on the north side of the station. |
| **ED9** | Roadside | 1 | $NO_2$ and NOx | South of an arterial road and next to a bus stop. Terraced houses on the south and north sides of the station. |

Commonly-used model evaluation statistics for the daily and monthly concentrations are summarised in Table 4.2 and Table 4.3. Definitions of these statistics can be found in Appendix I.

The model predicted well the daily average $NO_2$ concentration at the background site (ED3) with the smallest error (NMB: 3%) and high correlation coefficient (Table 4.2). At the roadside locations ED5, ED7 and ED8 the model generally underestimated $NO_2$, although not by much (NMB values of –9%, –2% and –15%, respectively) and had slightly smaller *r* values (0.70, 0.71 and 0.66, respectively). However, the model performed particularly poorly at ED1 and ED9 shown by the significant underestimation (NMB: ~ –50%) and large RMSE. The correlation coefficient at ED1 was poor but was high at ED9. The large negative bias suggested great underestimation in the emission, possibly caused by the unique movement of traffic at the junction for ED1 and the additional bus emission at ED9. However the good correlation at ED9 suggested that the dispersion processes were simulated well at ED9 but poorly at ED1 due to the local street canyon. $O_3$ measurement data were only available at ED3. The model simulated well urban background $O_3$ both in the daily temporal variation (high *r*) and in the absolute values (Table 4.2).

Validation results for the monthly average concentration were similar to the daily average but with slight improvement in the estimation errors since the uncertainties in the emissions were further smoothed out (Table 4.3).

*Table 4.2 Model evaluation statistics for daily average $NO_2$ and $O_3$ concentrations for 2012.*

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | **$NO_2$** | | | |
| **Site** | ***N*** | **FAC2** | **MB (μg $m^{-3}$)** | **NMB** | **RMSE (μg $m^{-3}$)** | ***r*** |
| **ED1** | 362 | 0.42 | –32.37 | –0.56 | 41.77 | 0.42 |
| **ED3** | 365 | 0.87 | 0.72 | 0.03 | 9.73 | 0.77 |
| **ED5** | 365 | 0.78 | –3.66 | –0.09 | 15.47 | 0.70 |
| **ED7** | 358 | 0.81 | –0.43 | –0.02 | 11.60 | 0.71 |
| **ED8** | 364 | 0.65 | –4.62 | –0.15 | 15.06 | 0.66 |
| **ED9** | 346 | 0.40 | –25.98 | –0.50 | 28.46 | 0.86 |
| | | | **$O_3$** | | | |
| **ED3** | 345 | 0.96 | –3.75 | –0.08 | 10.43 | 0.86 |

*Table 4.3 Model evaluation statistics for monthly average $NO_2$ and $O_3$ concentrations for 2012.*

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | **$NO_2$** | | | |
| **Site** | ***N*** | **FAC2** | **MB (μg $m^{-3}$)** | **NMB** | **RMSE (μg $m^{-3}$)** | ***r*** |
| **ED1** | 12 | 0.50 | –32.37 | –0.56 | 35.65 | 0.37 |
| **ED3** | 12 | 1.00 | 0.68 | 0.03 | 4.15 | 0.79 |
| **ED5** | 12 | 1.00 | –3.83 | –0.10 | 6.49 | 0.67 |
| **ED7** | 12 | 1.00 | –0.63 | –0.02 | 5.92 | 0.68 |
| **ED8** | 12 | 1.00 | –4.70 | –0.16 | 9.05 | 0.66 |
| **ED9** | 10 | 0.60 | –25.56 | –0.49 | 25.69 | 0.93 |
| | | | **$O_3$** | | | |
| **ED3** | 10 | 1.00 | –4.01 | –0.08 | 6.46 | 0.96 |

### 4.4.3 Validation against PDT measurements

Although the regulatory monitoring network provides real-time measurements, it is insufficient in number to capture the high spatial variability in $NO_2$ concentrations. Validation of the spatial aspect of the AMDS-Urban model output needs measurements with high spatial coverage. Passive diffusion tubes (PDTs) provide a low-cost alternative to the reference monitoring techniques in providing long-term average concentration at higher spatial resolution. Two available PDT datasets were used to validate ADMS-Urban. One was a network of 30 PDT sampling sites covering an area of ~7 $km^2$ in south Edinburgh (Lin et al., 2016). The other was more localised to a busy major road and along a perpendicular residential street (Kenagy et al., 2015). A basic summary of both datasets is shown in Table 4.4. Detail of the site locations can be found in (Lin et al., 2016) and (Kenagy et al., 2015).

*Table 4.4 Summary of the PDT datasets used for ADMS-Urban model validation.*

| **South Edinburgh** | | | |
|---|---|---|---|
| **Pollutants measured** | **Deployment period** | **Deployment interval** | **Number of sites** |
| **$NO_2$ and $O_3$** | Summer (2nd August – 13th September 2013) Winter (2nd December 2013 – 13th January 2014) | Weekly for $NO_2$ and bi-weekly for $O_3$ | 30 |
| **Craigmillar Park Road – Wilton road** | | | |
| **$NO_2$** | 12th September – 17th October 2014 | Weekly | 16 (two heights at 0.8 and 2 m at each site) |

Seasonal average concentration (i.e. mean of 6 weekly average $NO_2$ concentrations) was compared with ADMS-Urban output in the south Edinburgh dataset. Locations of the PDT in this dataset are shown in Figure 4.11. Sites containing more than one

missing weekly $NO_2$ or one bi-weekly $O_3$ were excluded from the model evaluation. Some PDTs were located next to a bus stop or at a traffic junction, where additional emissions from buses and traffic queueing are considered to be great but not modelled in the current model setup. Measurements at these sites therefore do not reflect the general predictive ability of the ADMS-Urban model and were also excluded from the evaluation. Figure 4.12 shows the relationship between modelled and measured $NO_2$ concentrations during different seasons. Overall the model underestimated $NO_2$ concentrations compared to the PDT measurements. However the spatial variation in the measured $NO_2$ was explained well by the model ($R^2$ = 73% and 77% for summer and winter, respectively) and was comparable to a previous ADMS-Urban model evaluation study (Dėdelė and Miškinytė, 2014). This indicates that although there is bias between modelled and PDT-measured $NO_2$ concentration the spatial pattern predicted by the model is consistent with the measurements. The bias could result from both the errors in the model and the errors in the PDT measurements. Large discrepancy (55% for summer and 82% for winter) between PDT measurement and reference analyser was observed during the deployment period at the ED3 co-location site. This partly explains the general underestimation in the modelled $NO_2$ compared to the PDT measurements. The positive bias in PDT measurement due to wind turbulence is well known in the literature (Heal et al., 2000). Assuming that this bias is systematic across all PDT sites, Figure 4.13 compares modelled $NO_2$ with PDT measurements corrected by the factor observed at the co-located reference analyser. The model performance significantly improved when comparing with the corrected PDT data in terms of estimation error. RMSE decreased from 12 $\mu g\ m^{-3}$ to 5 $\mu g\ m^{-3}$ and from 15 $\mu g\ m^{-3}$ to 4 $\mu g\ m^{-3}$ for summer and winter, respectively.

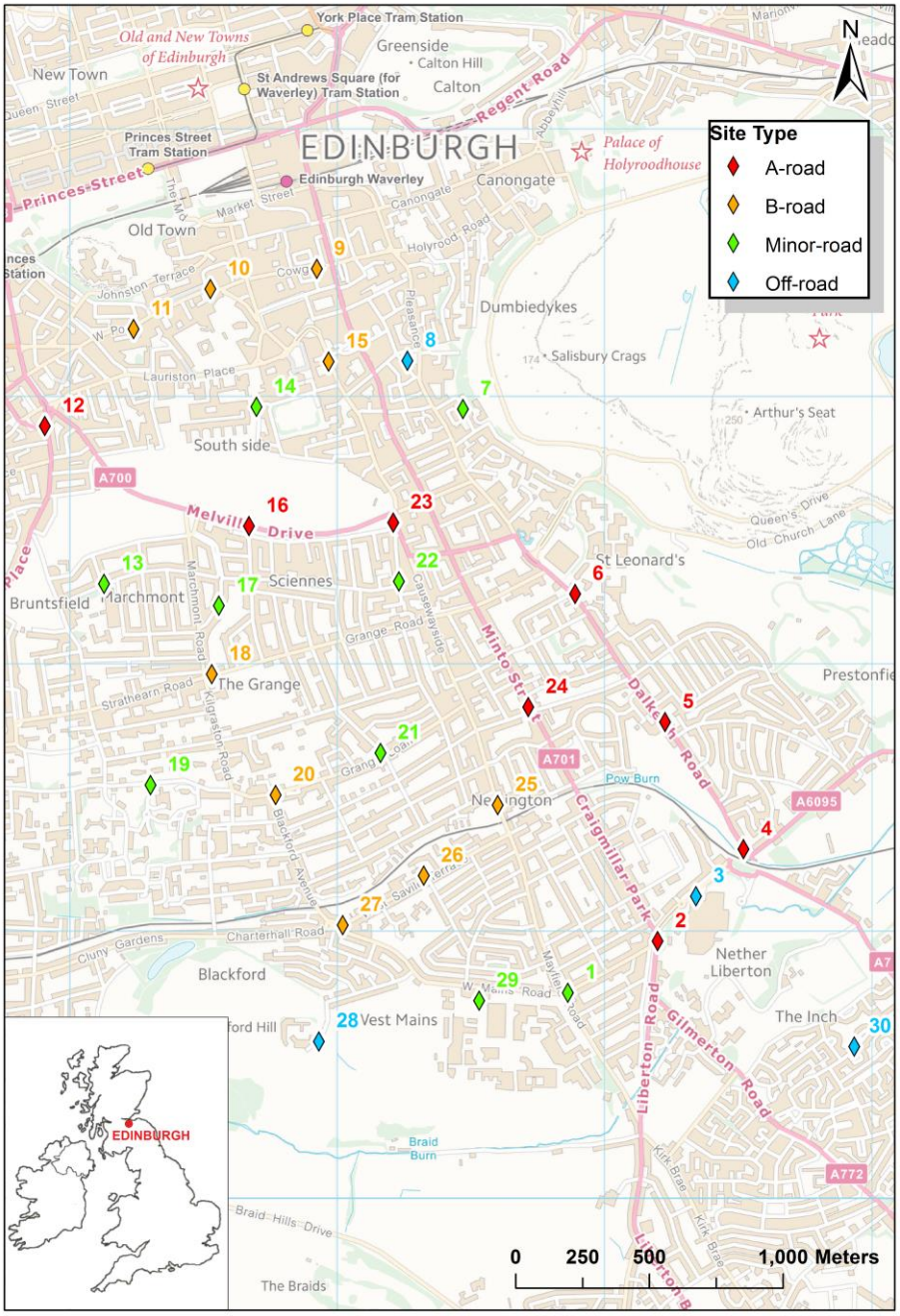



*Figure 4.11 Locations of the PDT sites in central-south Edinburgh. Site 8 is the AURN St. Leonards monitoring station. (Source: Lin et al. (2016))*

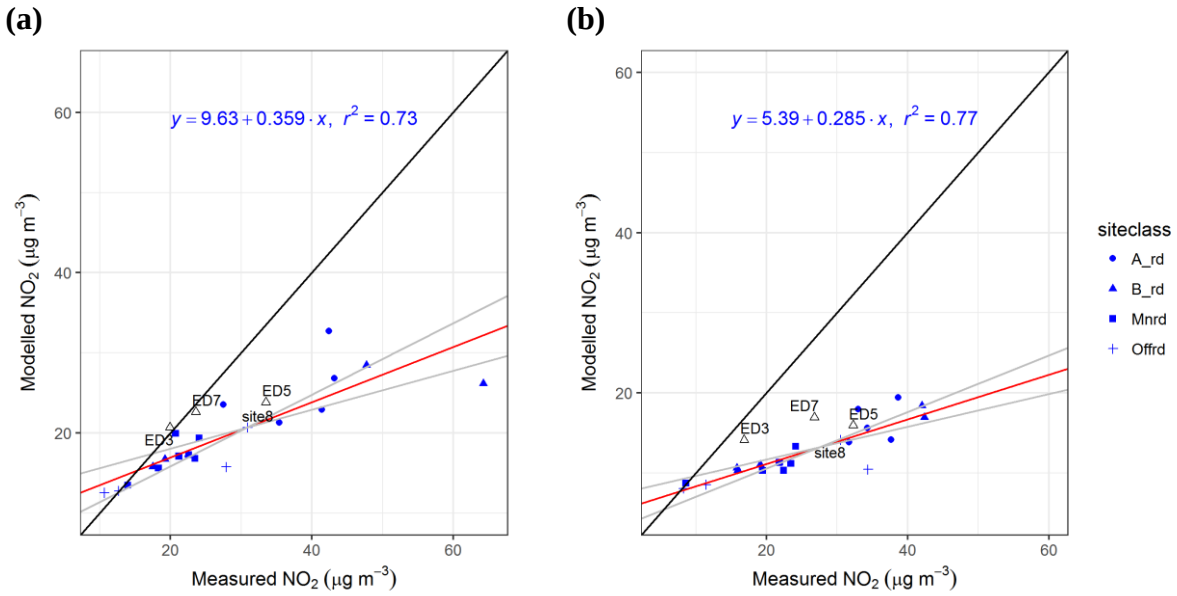



*Figure 4.12 Modelled and measured seasonal average $NO_2$ concentration for summer (a) and winter (b). Open triangle markers denote the measurements from reference analysers shown in Figure 4.10. Site 8 PDTs were co-located with reference analyser at ED3. The red and grey lines show the linear fit and 95% confidence interval. The solid black line shows the 1:1 relationship.*

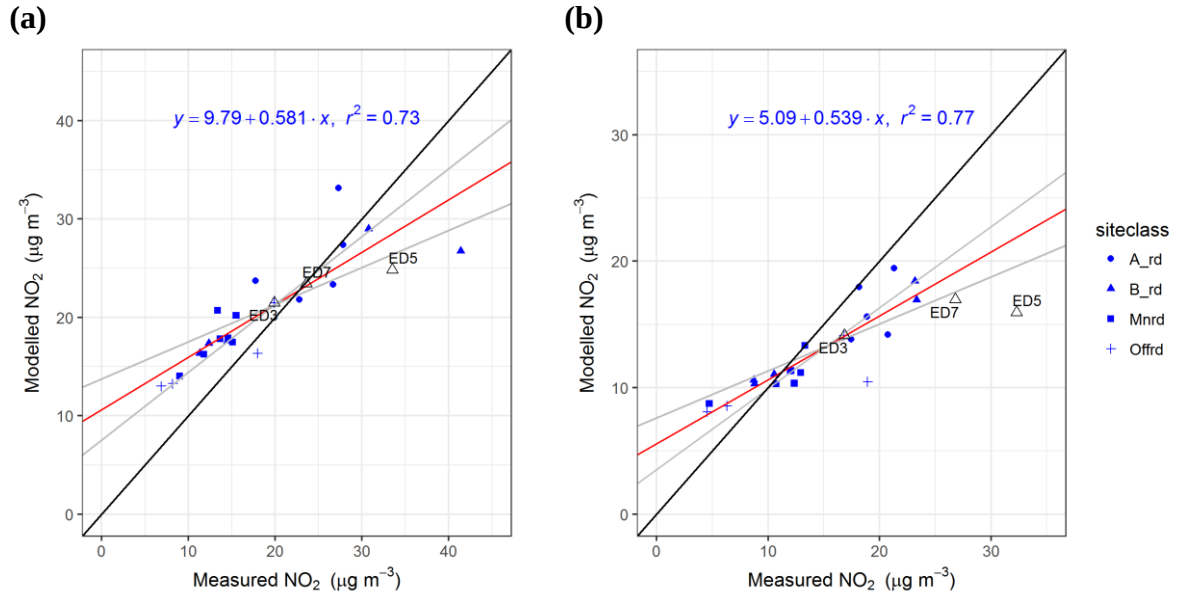



*Figure 4.13 Comparison of model and corrected PDT measurements for seasonal average $NO_2$ concentration for summer (a) and winter (b). The red and grey lines show the linear fit and 95% confidence interval. The solid black line shows the 1:1 relationship.*

In contrast to $NO_2$, correlation between measured and modelled $O_3$ concentrations was low, indicating that the spatial variability of $O_3$ was not well predicted (Figure 4.14). More importantly the range of $O_3$ concentration predicted by the model was much smaller than those measured by the PDTs. Positive bias may also have affected $O_3$ PDTs as indicated by the discrepancy between PDT and reference analyser at the co-located site 8 (Figure 4.14). Figure 4.15 shows the comparison between model and corrected $O_3$ PDT measurements. The correction eliminated some large underestimation by the model for the winter period (Figure 4.14b). However, the correlation and the spatial contrast (slope) between the model and measurement were roughly unchanged after the correction.

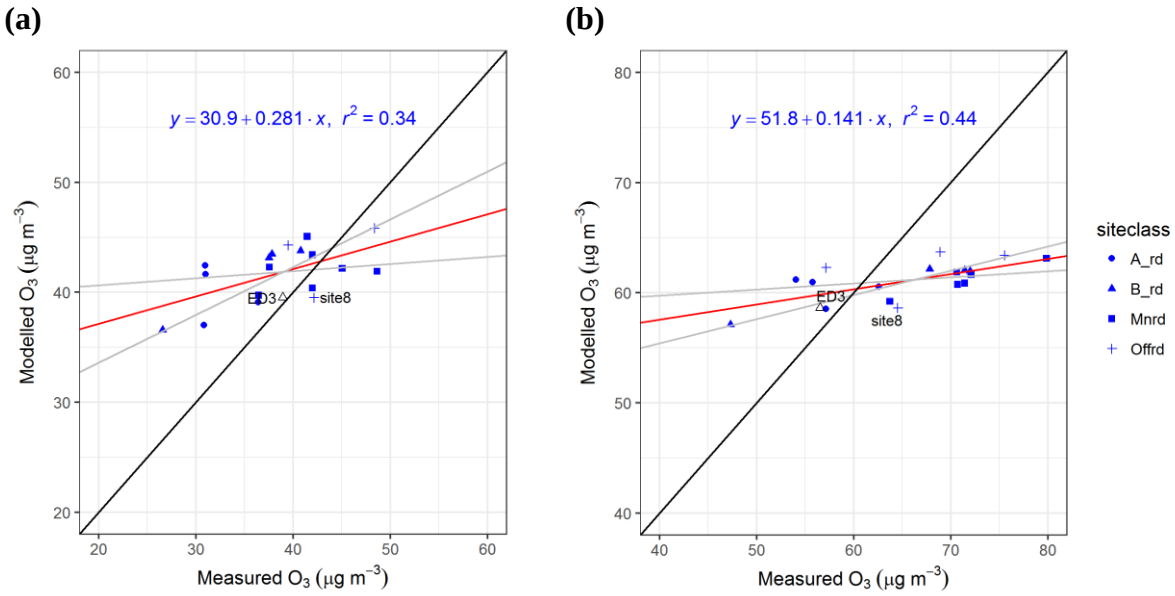



*Figure 4.14 Modelled and measured seasonal average $O_3$ concentration for summer (a) and winter (b). Site 8 PDTs were co-located with reference analyser at ED3. The red and grey lines show the linear fit and 95% confidence interval. The solid black line shows the 1:1 relationship.*

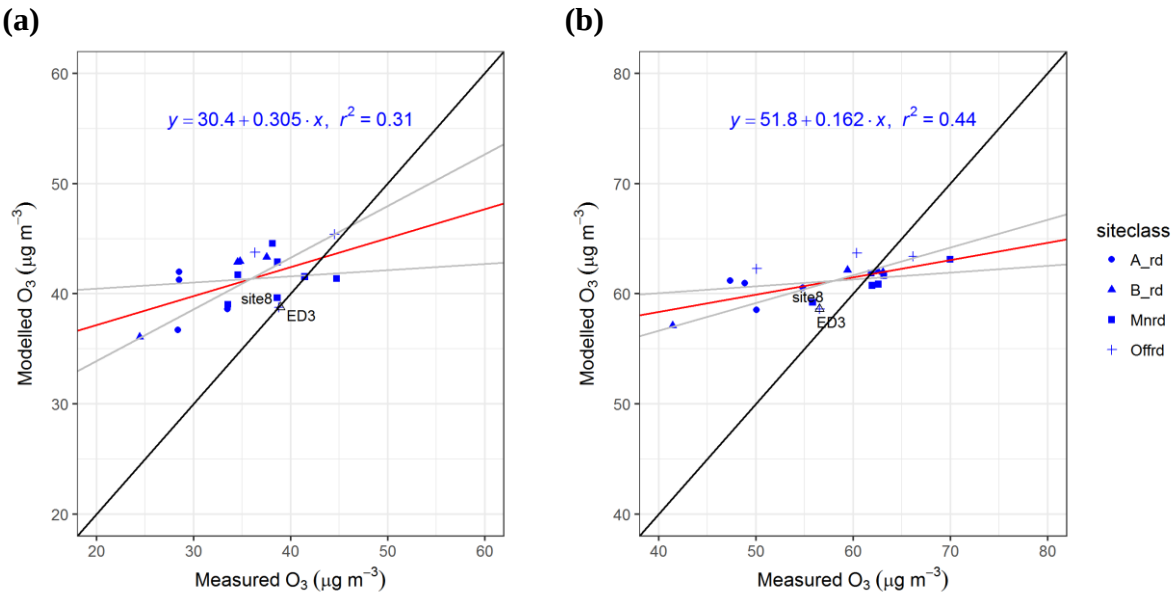



*Figure 4.15 Comparison of model and corrected PDT measurements for seasonal average $O_3$ concentration for summer (a) and winter (b). The red and grey lines show the linear fit and 95% confidence interval. The solid black line shows the 1:1 relationship.*

In the ADMS-Urban model the spatial variation of $O_3$ is mainly determined by its depletion through reaction with NO. On the other hand, spatial variation in $NO_2$ is determined not only by its generation in the NO + $O_3$ reaction but also from primary $NO_2$ emissions. In theory, at any given PDT site the increase in $NO_2$ ($\Delta NO_2$) compared to background concentration should always be equal to or larger than the decrease in $O_3$ ($\Delta O_3$) at the same site. Therefore, the difference between $\Delta NO_2$ and $\Delta O_3$ should reflect the contribution of primary $NO_2$ at that site, which is expected to be larger closer to roads and smaller away from direct emissions. Examining the $\Delta NO_2 - \Delta O_3$ quantity at a certain PDT site is equivalent to taking the difference between Ox at the PDT site and Ox at the background site, where Ox is the sum of $NO_2$ and $O_3$. This method has been previously used to estimate the primary $NO_2$ emission at traffic sites (Carslaw and Beevers, 2004). Here the $\Delta NO_2 - \Delta O_3$ values were examined for both PDT measurements and model outputs to compare the modelled and measured primary $NO_2$ emissions, thus giving insight into whether the relatively narrower range of modelled $O_3$ concentrations was caused by incorrect roadside NO concentration assumed. Measurement at Bush Estate was considered as the background concentration in the calculation of $\Delta NO_2 - \Delta O_3$. Bush Estate is roughly upwind of the prevailing wind with respect to the city of Edinburgh therefore giving a good representation of unpolluted city background.

Figure 4.16 shows the $\Delta NO_2 - \Delta O_3$ values for the model and for the corrected PDT measurements, with species concentrations converted to volume mixing ratios. As expected, modelled $\Delta NO_2 - \Delta O_3$ values were higher near A-roads reflecting significant contribution from primary $NO_2$ but gradually decreased at less busy roads. However, $\Delta NO_2 - \Delta O_3$ values calculated from PDT measurements were mostly negative even at A-roads and did not show a particular pattern between different sites. Negative values suggest the change in $O_3$ concentrations were larger than the change in $NO_2$ concentrations which contradicts the stoichiometry of the NO + $O_3$ reaction. Acknowledging that there are uncertainties in PDT measurements, $\Delta NO_2 - \Delta O_3$ values were also examined using reference measurement data. Figure 4.17 shows a selection of London Air Quality Network (LAQN) sites that provide both $NO_2$ and $O_3$ measurements and the $\Delta NO_2 - \Delta O_3$ values calculated from these sites.

Consistent with the modelled results and expectation, kerbside sites generally had higher $\Delta NO_2 - \Delta O_3$ values than urban background sites. $\Delta NO_2 - \Delta O_3$ values were consistently positive regardless of the site types. The unexpected large range of $O_3$ concentrations measured by PDTs in Edinburgh therefore cannot be explained by NO + $O_3$ chemistry. It is difficult to know whether this was caused by unknown biases in PDT measurements or by additional important $O_3$ depletion processes such as deposition onto concrete surfaces as observed by (Weissert et al., 2017).

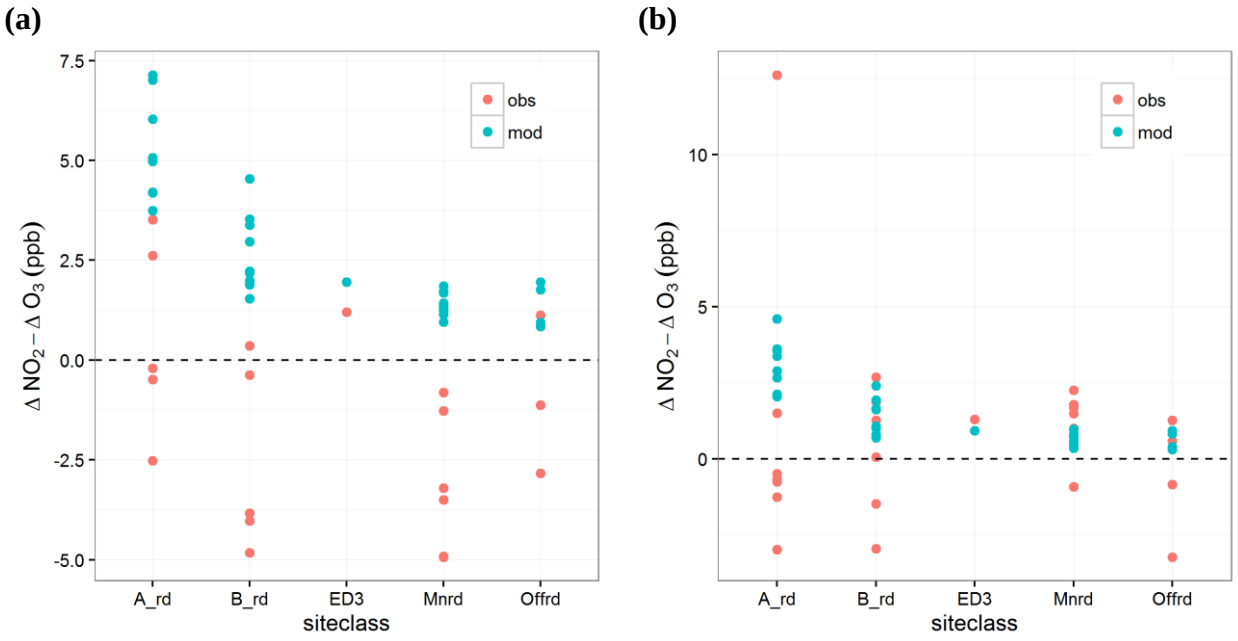
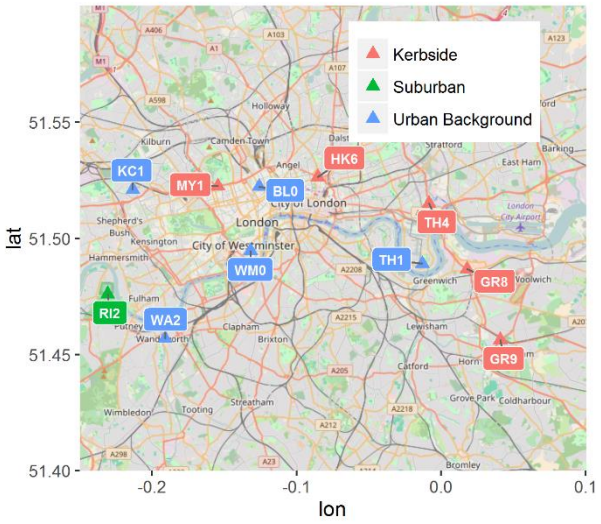



*Figure 4.16 Modelled and measured $\Delta NO_2 - \Delta O_3$ at different types of sites for summer averages (a) and winter averages (b).*

**(a)**




**(b)**

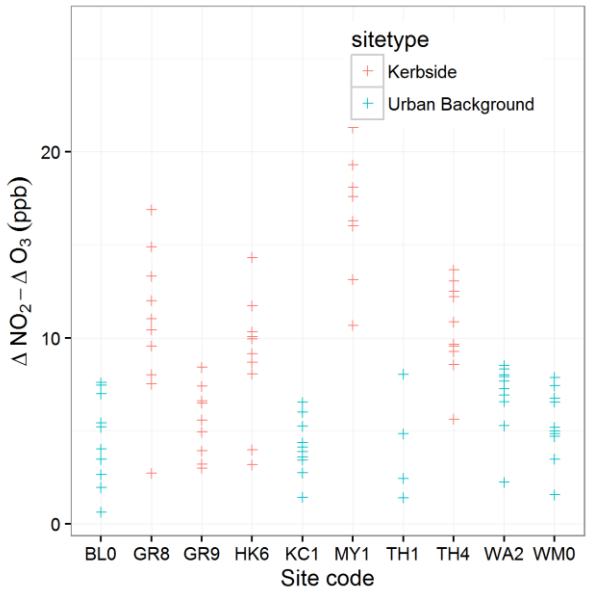


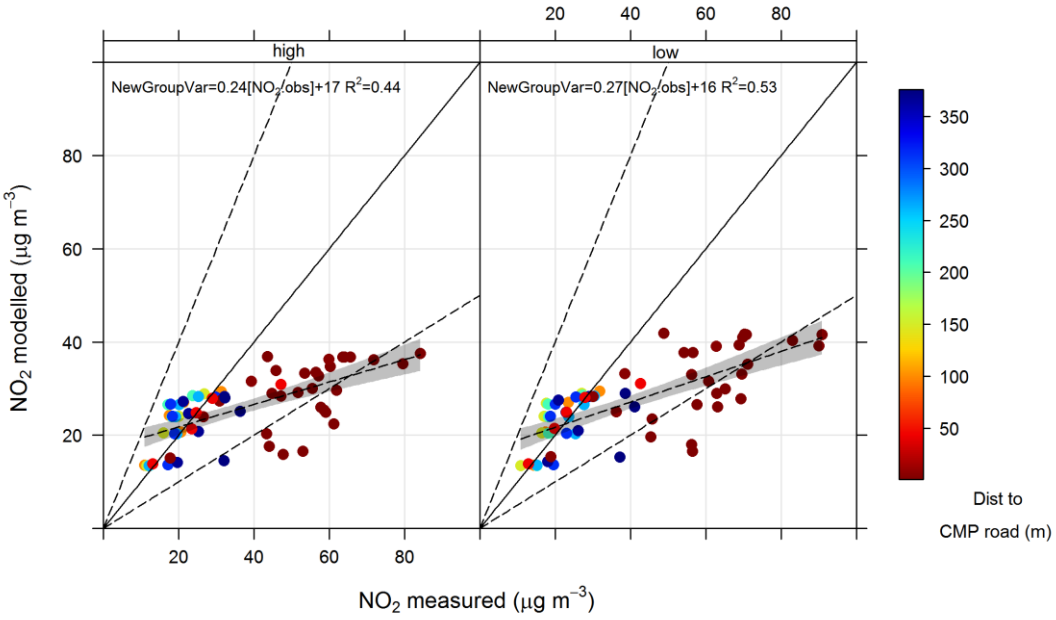

*Figure 4.17 Location of London reference monitoring stations (a) and the $\Delta NO_2 - \Delta O_3$ calculated using data from RI2 as the background site (b). $\Delta NO_2 - \Delta O_3$ at each site was calculated based on monthly average concentration in 2013.*

Validation of modelled $NO_2$ was also conducted at a very local scale where the emission was dominated by a major road source. The locations of PDT sites with respect to the major road under investigation are shown in Figure 4.18. Since Craigmillar Park road is the only major source of $NO_2$ in this small area, this study gives insights into the road emissions used in the model. Figure 4.19 shows the

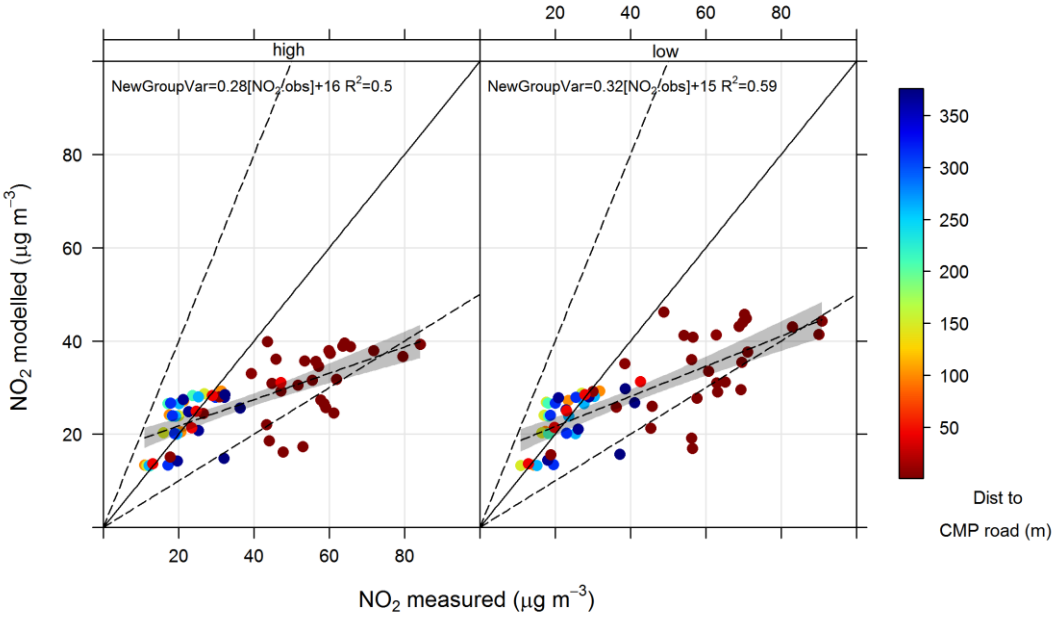
modelled and measured weekly average $NO_2$ concentrations at all the PDT sites. Figure 4.19a suggests underestimation by the model at sites closer to roads but generally good prediction at sites on the side road. This pattern can be explained by underestimation of primary $NO_2$ ($NO_2$/NOx ratio by volume) in the model. The effect of increasing primary $NO_2$ ratios on the modelled $NO_2$ concentrations was investigated. Figure 4.19b-d suggest increasing primary $NO_2$ increased roadside $NO_2$ concentrations while the concentrations on the side road remained the same. As a result model performance improved, both in terms of $R^2$ and slope coefficients, with increasing primary $NO_2$ ratios. However the roadside concentrations still tend to be underestimated even with 50% $NO_2$/NOx ratio applied in the model.


*Figure 4.18 Locations (and site labels) of the dual-height $NO_2$ measurements on Craigmillar Park (CMP) road and Wilton (W) road. At each site PDTs were deployed at 0.8 m (low) and 2 m (high) above the ground.*

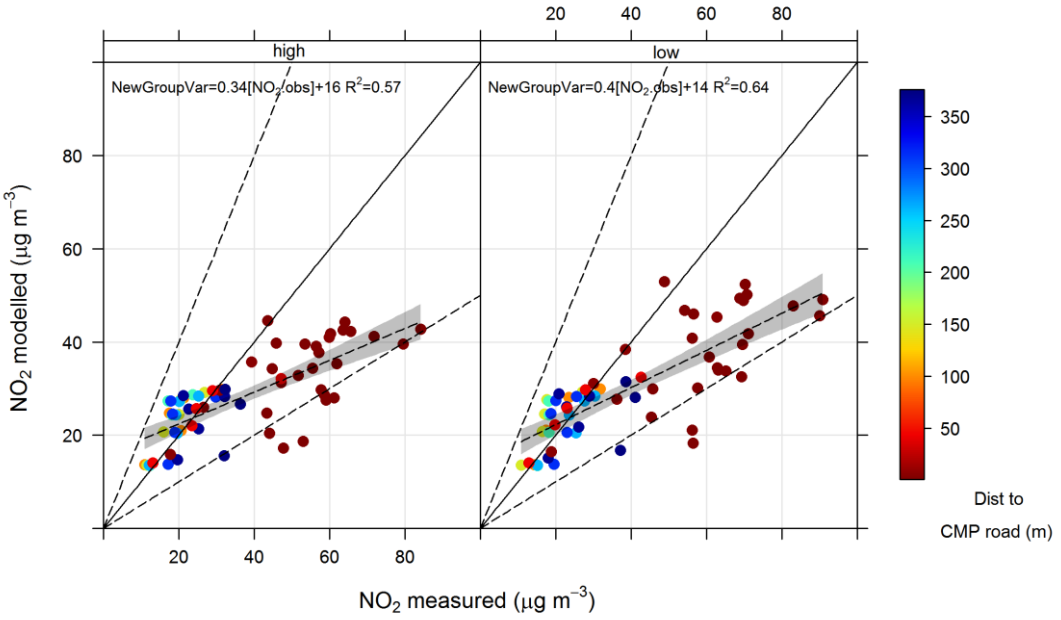
**(a)**


high
low
NewGroupVar=0.24[NO2/obs]+17 R2=0.44
NewGroupVar=0.27[NO2/obs]+16 R2=0.53
NO2 modelled (µg m−3)
NO2 measured (µg m−3)
20
40
60
80
350
300
250
200
150
100
50
Dist to
CMP road (m)

**(b)**


high
low
NewGroupVar=0.28[NO2/obs]+16 R2=0.5
NewGroupVar=0.32[NO2/obs]+15 R2=0.59
NO2 modelled (µg m−3)
NO2 measured (µg m−3)
20
40
60
80
350
300
250
200
150
100
50
Dist to
CMP road (m)

**(c)**




**(d)**

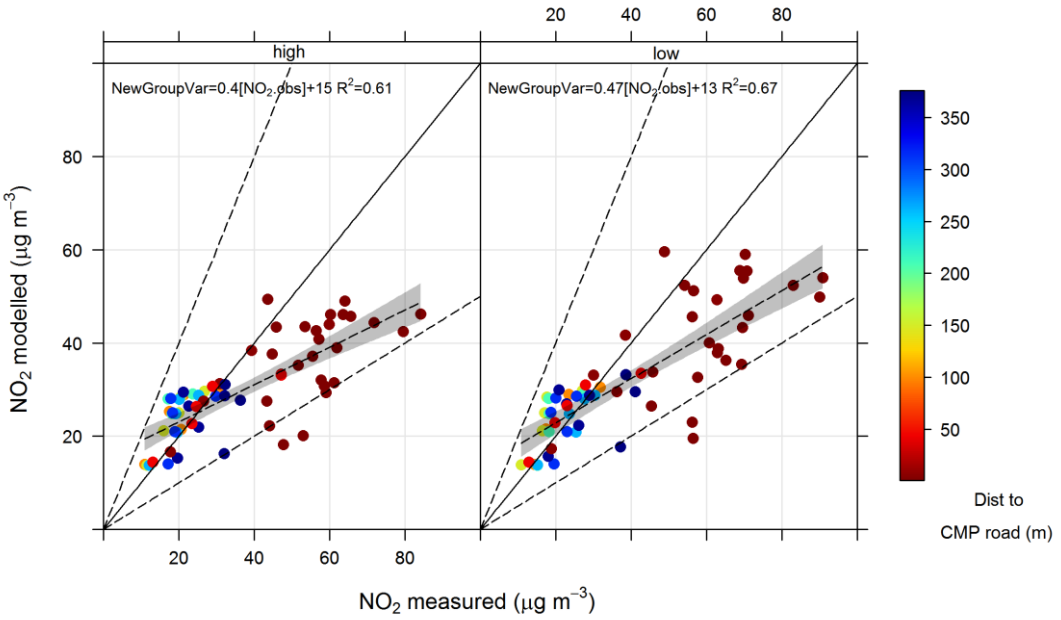



*Figure 4.19 Modelled and measured weekly average $NO_2$ concentrations. Points are colour coded by distance between PDT location and the edge of Craigmillar Park (CMP) road. The left and right plot in each row are for PDTs located at 2.2 m (high) and 0.8 m (low) above the ground, respectively. The different rows show the results of assuming different ratios of primary $NO_2$ to $NO_x$ for the modelled traffic emissions (a) 23.8%, (b) 30%, (c) 40% and (d) 50%.*

So what is the appropriate primary $NO_2$ to use in the model? The default value the AMDS-Urban model uses for the primary $NO_2$ from road emission is 23.8% suggested by the 2013 NAEI estimate for 'All UK traffic' (based on 2014 traffic fleet projection) (NAEI, 2013). Since then there has been an update in COPERT resulting in an increase in the primary $NO_2$ to 26.2% for 'All UK traffic' in the 2015 NAEI estimate. COPERT is a software NAEI uses to calculate the emission factors that form the basis of the UK emission inventory (Defra, 2015). As shown in Figure 4.20 the update in 2015 estimates had notable changes in primary $NO_2$ for cars and LGVs. The implication of this change on the modelled road emission depends on the traffic composition of the modelled road. Figure 4.21 shows the difference in primary $NO_2$ between 2013 and 2015 NAEI estimates calculated based on 2014 traffic counts in Edinburgh. Since the traffic fleet was dominated by cars and a large increase in primary $NO_2$ for cars was estimated in NAEI 2015, the primary $NO_2$ on most roads increased from 20-25% to 25-30%. The primary $NO_2$ on different major roads was similar, supporting the use of a single value for the city-scale modelling, with an exception for Princes Street where the majority of the traffic was buses. The updated NAEI estimates and Figure 4.19b-d suggest a higher primary $NO_2$ ratio should be used than the ADMS-Urban default value. However, a recent study (Carslaw et al., 2016) suggested the opposite, reporting evidence on decreasing primary $NO_2$ from a peak of around 25% in 2010 to about 15% at the end of 2014 in London. Remote sensing data also revealed that primary $NO_2$/NOx ratio might be too high in the COPERT emission factors especially for Euro 4 and 5 vehicles, where remote sensing data estimated ~26% compared with 46% in COPERT (Carslaw et al., 2016). The introduction of the congestion charging zone and low emission zone in London dramatically changes the distribution of actual vehicle standards in the fleet, which is a caveat in comparing results from studies in London with other UK cities. Nevertheless, there are clearly inconsistencies in the inventory-based and measurement-based primary $NO_2$ ratios. However, the uncertainty in primary $NO_2$ will only affect immediate roadside $NO_2$ concentrations but will have little effect at locations away from road sources where there is sufficient dispersion time for NO + $O_3$ reaction to reach equilibrium.

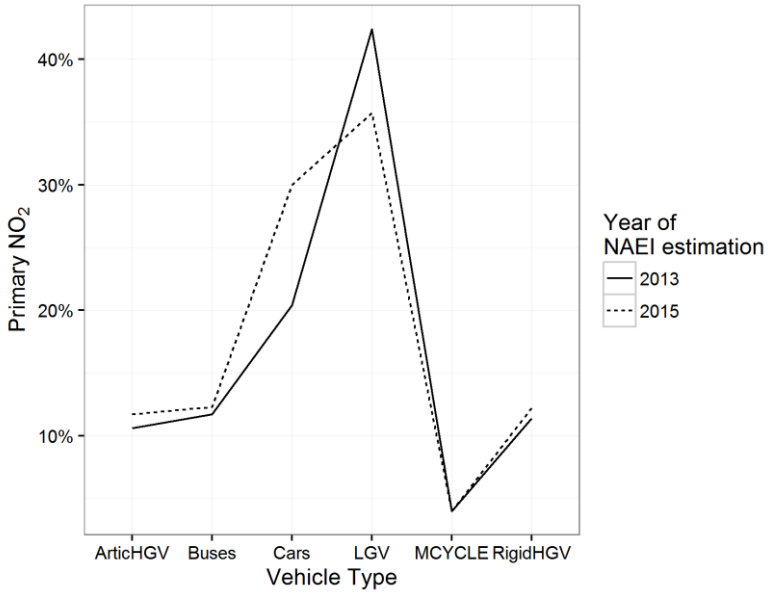



*Figure 4.20 NAEI estimates of primary $NO_2$ to NOx ratios based on 2014 fleet composition.*

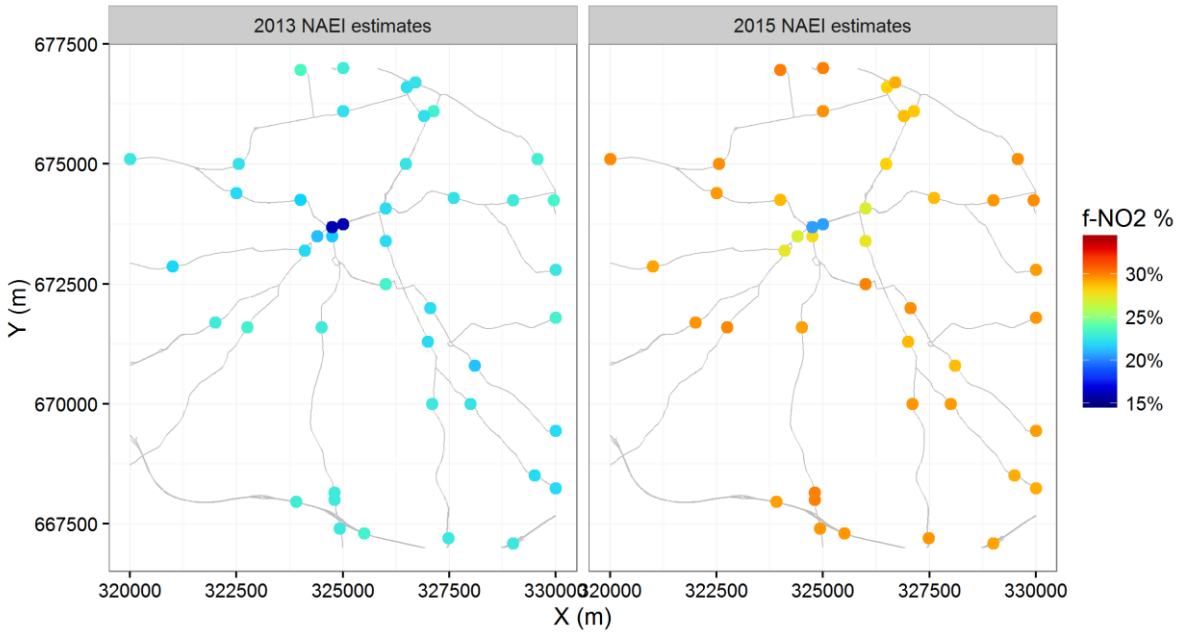



*Figure 4.21 Primary $NO_2$ to $NO_x$ ratios estimated at traffic counting points in Edinburgh using NAEI 2013 and 2015 emissions factors and the traffic count data for 2014.*

## 4.5 Conclusions

Various combinations of emissions and meteorological input datasets have been tested for modelling $NO_2$ concentration in Edinburgh. Differences in emissions from major roads calculated based on bottom-up and top-down approaches were noted as a result of inconsistent vehicle speed used and different emission resolutions (road links or grid cells). The discrepancy in the major road emissions will most likely affect the modelled concentrations at/near roadside but less so at locations away from the road. Minor road emissions were found to be comparable or higher than major road emissions in many urban areas. Modelling minor roads as gridded area sources may falsely increase the background concentration at areas where minor road emissions are dominant local source. It is therefore preferable to model both major and minor roads explicitly as road sources. Without sufficient roadside measurements, it is difficult to assess which particular emission setup gives more accurate estimation. However, using NAEI emission data ensures consistency between major and minor road emissions and is practically easier than preparing DfT data in ADMS-Urban input format.

Selecting the location to measure meteorological parameters representing the urban environment requires great caution. As shown in this study, meteorological data measured simply at a receptor location may be affected by micro-scale climate and may not be suitable to use in the dispersion modelling.

The model setup based on NAEI emissions and Gogarbank meteorological measurements was used in the model validation against reference analyser and PDT measurements. Temporal variability (daily and monthly) of $NO_2$ was predicted well at monitoring sites that were not heavily affected by local effects such as road junctions and bus stops. Temporal variability of $O_3$ was predicted more accurately than for $NO_2$ although only one monitoring site was available for this evaluation. Long-term (seasonal average) spatial variability of modelled $NO_2$ was found to be in good agreement with PDT measurements. Modelled $O_3$ concentrations compared poorly with PDT measurements. However, it was found that the $O_3$ PDT measurement in this study may be affected by non-systematic biases affecting its

accuracy. Validation of modelled $NO_2$ concentrations at a number of locations close to a major road source uncovered that the default primary $NO_2$ to $NO_x$ ratio used in ADMS-Urban might be too low for modelling traffic emissions in 2014. There is no guidance from current literature on what the correct primary $NO_2$ ratio is. However, the uncertainties regarding primary $NO_2$ emissions will only likely affect the immediate roadside concentrations but have little effect on locations a few or tens metres away from the road.

Given the good performance of the ADMS-Urban model at high spatiotemporal resolution, it serves as a useful tool for planning the time-consuming and costly real-world measurements for developing empirical air quality models. One of the challenges in developing and validating the empirical air quality models is the limited monitoring sites. ADMS-Urban is proven to be able to provide a realistic continuous pollution field, where many possible monitoring strategies can be tested and evaluated. The monitoring strategy for building one of the statistical air quality models, land-use regression model, is investigated in Chapter 5.

# References

Cai, Y., Schikowski, T., Adam, M., Buschka, A., Carsin, A.-E., Jacquemin, B., Marcon, A., Sanchez, M., Vierkötter, A., Al-Kanaani, Z., Beelen, R., Birk, M., Brunekreef, B., Cirach, M., Clavel-Chapelon, F., Declercq, C., de Hoogh, K., de Nazelle, A., Ducret-Stich, R.E., Valeria Ferretti, V., Forsberg, B., Gerbase, M.W., Hardy, R., Heinrich, J., Hoek, G., Jarvis, D., Keidel, D., Kuh, D., Nieuwenhuijsen, M.J., Ragettli, M.S., Ranzi, A., Rochat, T., Schindler, C., Sugiri, D., Temam, S., Tsai, M.-Y., Varraso, R., Kauffmann, F., Krämer, U., Sunyer, J., Künzli, N., Probst-Hensch, N., Hansell, A.L., 2014. Cross-sectional associations between air pollution and chronic bronchitis: an ESCAPE meta-analysis across five cohorts. Thorax 69, 1005–1014. doi:10.1136/thoraxjnl-2013-204352

Carslaw, D.C., 2016. The openair Manual. URL https://www.dropbox.com/s/2n7wdyursdul8dk/openairManual.pdf?dl=0 (accessed 19-June-2015).

Carslaw, D.C., Beevers, S.D., 2004. Investigating the potential importance of primary NO2 emissions in a street canyon. Atmos. Environ. 38, 3585–3594. doi:10.1016/j.atmosenv.2004.03.041

Carslaw, D.C., Murrells, T.P., Andersson, J., Keenan, M., 2016. Have vehicle emissions of primary NO2 peaked? Faraday Discuss 189, 439–454. doi:10.1039/C5FD00162E

CERC, 2016a. ADMS-Urban User Guide 4.0 [WWW Document]. URL http://www.cerc.co.uk/environmental-software/user-guides.html (accessed 11-March-2016).

CERC, 2016b. ADMS Technical specifications [WWW Document]. URL http://cerc.co.uk/environmental-software/technical-specifications.html (accessed 17-November-2016).

CERC, 2016c. ADMS Model validation [WWW Document]. URL http://cerc.co.uk/environmental-software/model-validation.html (accessed 17-November-2016).

Dėdelė, A., Miškinytė, A., 2014. Estimation of inter-seasonal differences in NO2 concentrations using a dispersion ADMS-Urban model and measurements. Air Qual. Atmosphere Health 8, 123–133. doi:10.1007/s11869-014-0272-9

Defra, 2015. UK Emission Mapping Methodology 2012. UK. URL http://naei.defra.gov.uk/reports/reports?report_id=854 (accessed 5-October-2016).

DfT, 2015. City of Edinburgh [WWW Document]. Dep. Transp. - Traffic Counts. URL http://www.dft.gov.uk/traffic-counts/cp.php?la=City+of+Edinburgh (accessed 11-March-2015).

ESCAPE, 2010. ESCAPE Exposure assessment manual. URL http://www.escapeproject.eu/manuals/ESCAPE_Exposure-manualv9.pdf

Heal, M.R., Kirby, C., Cape, J.N., 2000. Systematic Biases in Measurement of Urban Nitrogen Dioxide using Passive Diffusion Samplers. Environ. Monit. Assess. 62, 39–54.

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. Atmos. Environ. 42, 7561–7578. doi:10.1016/j.atmosenv.2008.05.057

Kenagy, H.S., Lin, C., Wu, H., Heal, M.R., 2015. Greater nitrogen dioxide concentrations at child versus adult breathing heights close to urban main road kerbside. Air Qual. Atmosphere Health. doi:10.1007/s11869-015-0370-3

Laurent, O., Pedrono, G., Segala, C., Filleul, L., Havard, S., Deguen, S., Schillinger, C., Rivière, E., Bard, D., 2008. Air Pollution, Asthma Attacks, and Socioeconomic Deprivation: A Small-Area Case-Crossover Study. Am. J. Epidemiol. 168, 58–65. doi:10.1093/aje/kwn087

Lin, C., Feng, X., Heal, M.R., 2016. Temporal persistence of intra-urban spatial contrasts in ambient NO2, O3 and Ox in Edinburgh, UK. Atmospheric Pollut. Res. doi:10.1016/j.apr.2016.03.008

Maheswaran, R., Pearson, T., Smeeton, N.C., Beevers, S.D., Campbell, M.J., Wolfe, C.D., 2010. Impact of Outdoor Air Pollution on Survival After Stroke. Stroke 41, 869–877. doi:10.1161/STROKEAHA.109.567743

Moolgavkar, S.H., Roger O. McClellan, Anup Dewanji, Jay Turim, E. Georg Luebeck, Melanie Edwards, 2013. Time-Series Analyses of Air Pollution and Mortality in the United States: A Subsampling Approach. Environ. Health Perspect. 121, 73–78. doi:10.1289/ehp.1104507

NAEI, 2015. UK Emissions Interactive Map [WWW Document]. Natl. Atmospheric Emiss. Inventory. URL http://naei.defra.gov.uk/data/gis-mapping

NAEI, 2013. Primary NO2 Emission Factors for Road Vehicles [WWW Document]. URL http://naei.defra.gov.uk/data/ef-transport (accessed 30-March-2015).

NOAA, 2016. Integrated Surface Database (ISD) [WWW Document]. URL https://www.ncdc.noaa.gov/isd (accessed 18-November-2016).

Oke, T., 2006. Instruments and observing methods report no. 81, Initial guidance to obtain representative meteorological observations at urban sites. World meteorological organization. URL http://ww.instesre.org/GCCE/UrbanMeteorologicalObservations.pdf (accessed 1-November-2016).

Samoli, E., Atkinson, R.W., Analitis, A., Fuller, G.W., Green, D.C., Mudway, I., Anderson, H.R., Kelly, F.J., 2016. Associations of short-term exposure to traffic-related air pollution with cardiovascular and respiratory hospital admissions in London, UK. Occup. Environ. Med. 73, 300–307. doi:10.1136/oemed-2015-103136

Smith, J.D., Mitsakou, C., Kitwiroon, N., Barratt, B.M., Walton, H.A., Taylor, J.G., Anderson, H.R., Kelly, F.J., Beevers, S.D., 2016. London Hybrid Exposure Model: Improving Human Exposure Estimates to NO2 and PM2.5 in an Urban Setting. Environ. Sci. Technol. 50, 11760–11768. doi:10.1021/acs.est.6b01817

Weissert, L.F., Salmond, J.A., Miskell, G., Alavi-Shoshtari, M., Grange, S.K., Henshaw, G.S., Williams, D.E., 2017. Use of a dense monitoring network of low-cost instruments to observe local changes in the diurnal ozone cycles as marine air passes over a geographically isolated urban centre. Sci. Total Environ. 575, 67–78. doi:10.1016/j.scitotenv.2016.09.229

# Chapter 5 Evaluation of Land Use Regression Models for Estimating Residential $NO_2$ Concentration

*This chapter is based on a research paper published in 'Atmospheric Environment' (Wu, H., Reis, S., Lin, C., Heal, M.R., 2017. Effect of monitoring network design on land use regression models for estimating residential NO2 concentration. Atmos. Environ. 149, 24–33. doi:10.1016/j.atmosenv.2016.11.014). I undertook the experimental design, data analysis and drafting of the manuscript. Dr Mathew Heal and Dr Stefan Reis gave advice on the presentation of results and manuscript editing. Dr Chun Lin undertook the passive diffusion tube measurements which were used for model validation.*

## 5.1 Introduction

The assessment of long-term exposure to air pollution for epidemiological and health burden studies has been a challenge because of the high spatial variation of pollutant concentration in the urban environment, particularly $NO_2$ (Briggs, 2005; Jerrett et al., 2005). Over the years, land use regression (LUR) modelling has demonstrated better or equivalent performance to other geostatistical methods (Hoek et al., 2008), and therefore has become popular in health studies to estimate long-term exposure to ambient $NO_2$ (Beelen et al., 2014; Jerrett et al., 2009). LUR modelling is a stepwise multiple regression method that regresses the pollutant concentration at the measurement sites against the land-use variables within buffer areas around the measurement sites (Jerrett et al., 2007). The derived empirical relationship between pollutant concentration and surrounding land use is then applied to un-sampled locations to provide a spatially-resolved seasonal or annual average pollution field.

The selection of monitoring sites to build the LUR model has been identified as one of the factors affecting the quality of the LUR model, but a rigorous method to determine the number and distribution of monitoring sites is lacking (Hoek et al., 2008). One study (Kanaroglou et al., 2005) aimed to develop a formal method to locate air quality monitors for LUR model development. However, the method has been rarely applied due to its complexity and the extensive prior knowledge required

on the population and pollutant distributions. A few studies (Basagaña et al., 2012; Johnson et al., 2010; Wang et al., 2012) evaluated the effect of number of monitoring sites on LUR model performance, but the effect of the distribution of monitoring sites remains to be investigated.

Validation of an LUR model has always been limited to the measurements available in a monitoring campaign (Hoek et al., 2008). The ultimate goal of exposure assessment is to accurately predict the exposure of hundreds or thousands of study subjects, but validation of an LUR model at this level through measurements is practically impossible. However, with the use of a dispersion model it is possible to simulate a pseudo-measured concentration at every residential address, which can then be compared with an LUR-model estimated concentration to assess the validity of the latter.

The aim of this study was to evaluate a large suite of LUR models built from different monitoring network designs and to validate the LUR models using dispersion modelled concentration at each home address. This modelling study used as its basis the city of Edinburgh (population ~460,000) in the east of Scotland, UK (55.94° N, 3.18° W). The outcome of the evaluation is to recommend sampling strategies and to highlight how particular monitoring network designs may lead to potential exposure misclassification.

## 5.2 Method

The evaluation of the performance of the LUR models constructed from different monitoring networks was carried out in four stages. An overview of the methodology is presented first, with details described in subsequent sub-sections. A schematic of the overall workflow is shown in Figure 5.1.

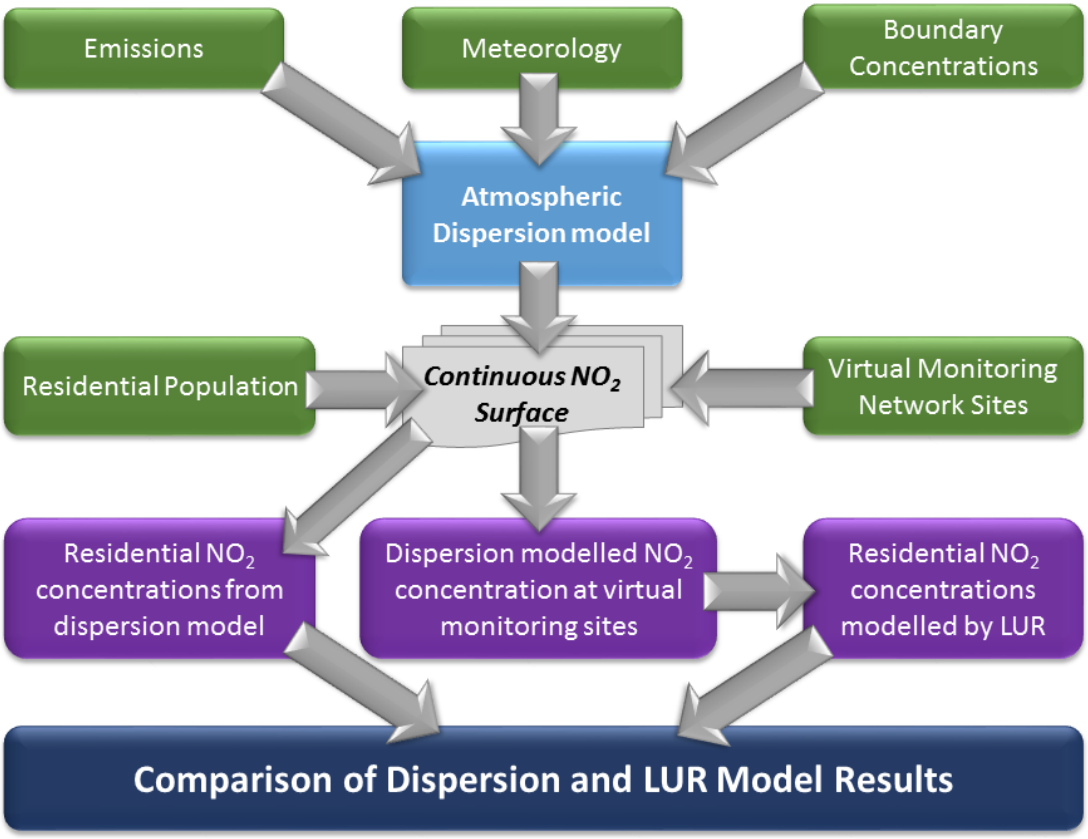



*Figure 5.1 Workflow for evaluating the performance of the LUR models by using dispersion model output.*

The ADMS-Urban model v3.4 (CERC, 2016) was used to simulate $NO_2$ concentrations for each of the population home addresses (centre points of residential buildings) in a 5 km × 5 km study area in Edinburgh (Figure 5.2). This area covers the commercial (city centre) and residential areas of the city and encompasses 7,445 residential buildings housing a total population of 144,715. The dispersion modelled

$NO_2$ concentration is considered to be the reference, on which the subsequent LUR model development and validation are based. Next, three different types of monitoring networks were designed (comprising different numbers of monitoring sites) based on household density and proximity to road. The $NO_2$ concentration at each monitoring site was modelled with ADMS-Urban using the same setup as the modelling of residential $NO_2$ concentration. The third stage was to develop a separate LUR model for each monitoring network, which was then applied to residential address to provide an LUR-model estimate. Finally, the LUR-model-estimated residential concentration was compared with the dispersion modelled residential concentration. The extent of agreement between the two indicates the performance of the LUR model and, in turn, the performance of the monitoring network from which the LUR model was constructed.

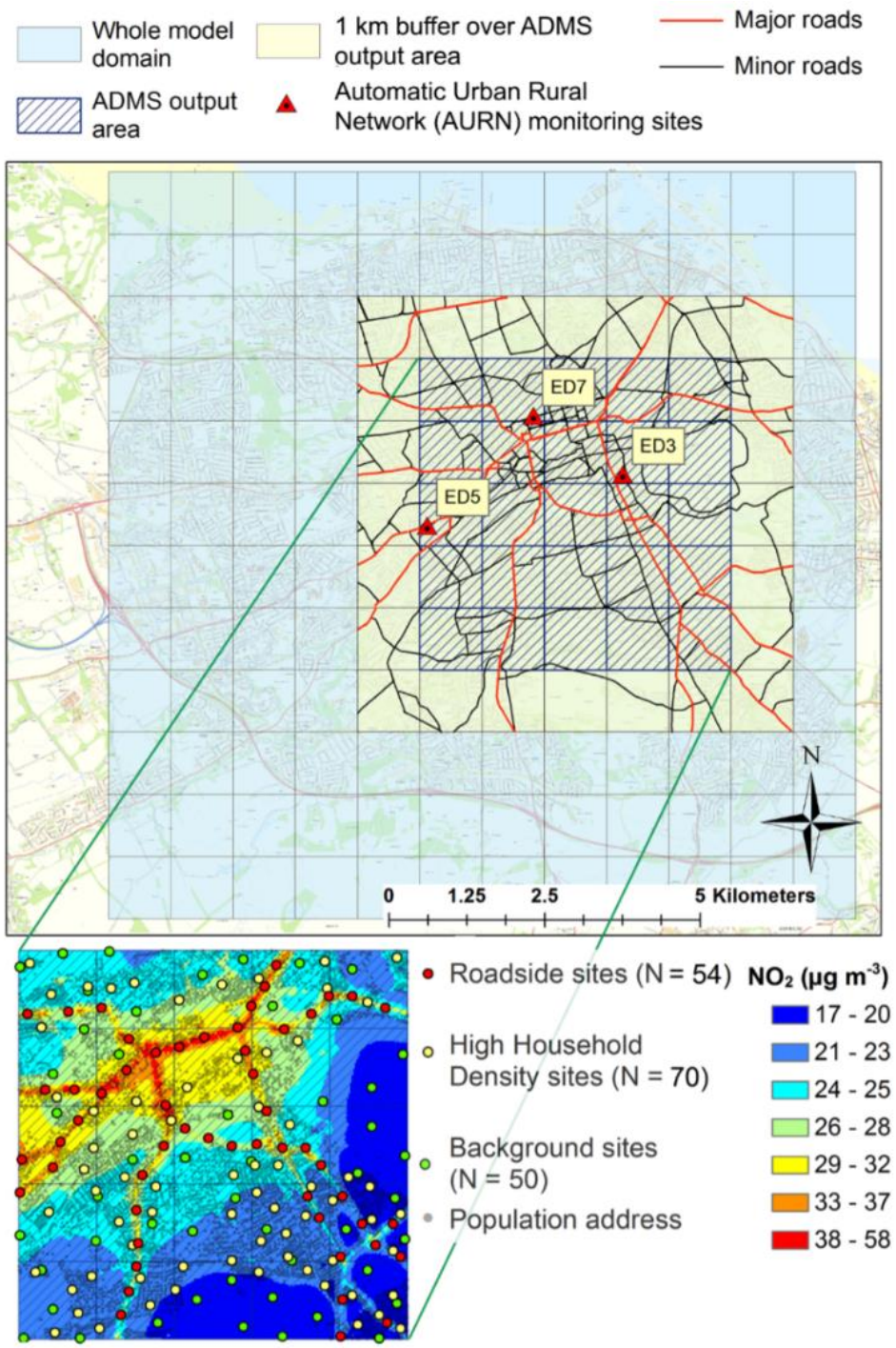



*Figure 5.2 The modelling domain for the city of Edinburgh. The inset map shows the locations of all potential monitoring sites. The underlying contour plot shows the monthly average dispersion model concentration of $NO_2$ for April 2012.*

### 5.2.1 Stage 1 – Dispersion modelling of residential $NO_2$ concentration

#### 5.2.1.1 Data preparation

All the Geographic Information System (GIS) data, including buildings and road networks for the City of Edinburgh, were obtained from EDINA Digimap Ordnance Survey Service (Ordnance Survey, 2015). Annual average daily flow (AADF) of traffic for the major roads in 2013 were downloaded from the Department for Transport (DfT, 2015). The population of each postcode area for the 2011 census was distributed to the buildings within the polygon area based on the volume of the building (building polygon area × building height). The centre of the building polygon with assigned population was used as the home address.

#### 5.2.1.2 ADMS-Urban setup

The model domain (12 km × 12 km) covered most of the City of Edinburgh in which all the emissions of $NO_x$ and NMVOC (non-methane volatile organic compound) were modelled (Figure 5.2). Within this larger domain, a 5 km × 5 km subset was chosen to output the concentration at each home address. To allow the receptors on the edges of the inner domain to be modelled smoothly, a 1 km buffer zone was added around the 5 km × 5 km output area. Within the buffer zone all the major and minor roads were explicitly modelled as road sources, whereas emissions outside the buffer zone were modelled as a 1 km × 1 km gridded area source. $NO_x$ and NMVOC emissions were downloaded from the UK National Atmospheric Emissions Inventory (NAEI) for 2012 (NAEI, 2015) with a resolution of 1 $km^2$. Road emissions were calculated by dividing the total emissions for the major or minor road subsector by the total length of the corresponding roads within each 1 $km^2$ grid. For grids in which road emissions were explicitly modelled, the road emissions were subtracted from the grid total emission. Measured meteorological data for the model, including wind speed/direction, cloud cover and temperature, were obtained from a WMO station to the west of the model domain (Gogarbank: 55.93° N, 3.35° W) (Met Office, 2012). An urban canopy file was prepared to account for the variation in the vertical profiles of wind speed and turbulence caused by the presence of buildings. Background concentrations of hourly-average $NO_2$, $NO_x$ and $O_3$ were obtained from a rural

national-network monitoring station to the south of the model domain (Bush Estate: 55.86° N, 3.21° W). For 2012, 0.8% and 23% of $O_3$ and $NO_x$ measurements, respectively, were missing. These were replaced by the average concentration for that particular hour over the whole year. Monthly average concentration was calculated at each receptor in ADMS-Urban, from which the annual-average concentration was calculated as the metric in the subsequent analysis.

### 5.2.2 Stage 2 – Sampling network design

Three different types of sampling networks were investigated. The aim in the selection of monitoring sites was to investigate how network selection criteria and number of sites influence the representation of the spatial variation of $NO_2$ at the residential home addresses. Specifically, the exposure study area was first disaggregated into 25 m × 25 m grid cells. The following GIS variables were then calculated for the centroid of each grid cell: total home addresses within a 100 m buffer (HH100) and distance to major/minor road edge (MJRDDIST/MNRDDIST). Three types of monitoring sites were then defined:

- High household density sites (HH sites): centroid of the cells with HH100 falling in the top 10% of all the 25 m grid cells;
- Roadside sites: centroid of the cells with MJRDDIST between 0 – 5 m;
- Background sites: centroid of the cells with both MJRDDIST and MNRDDIST > 200 m.

A subset of each of the total set of HH and roadside sites was then randomly selected, subject to a minimum distance of 300 m between any pair of sites, to form two pools of potential HH and roadside monitoring sites to be used in the sampling networks. The locations of these sets of monitoring sites are shown in Figure 5.2. The purpose of adding a minimum distance constraint to the random selection was to ensure that potential network sites were distributed across the range of localities in the study area. A third subset of sites was randomly selected from the total set of background sites, but with a minimum distance constraint of 500 m, since the background concentrations in this modelling study are mainly determined by the gridded emissions which have a resolution of 1 $km^2$ rather than by the road network

which has a finer resolution. Due to the minimum distance constraint added to the random selection, the number of monitoring sites of each type from which a monitoring network could subsequently be selected comprised 54, 70 and 50 for roadside, HH and background sites, respectively (Figure 5.2).

From these potential monitoring sites, the following three types of sampling networks were designed by randomly selecting different numbers of sites from each type of monitoring site.

- Household density based network (*HH* network): randomly selecting from the HH sites only;
- Proximity to road based network (*Road* network): randomly selecting equal numbers of roadside sites and background sites;
- *Mixed* network: randomly selecting equal numbers of roadside sites and HH sites.

Eleven different numbers of monitoring sites were tested for each type of network design ranging from 10 to 60 (in steps of $N = 5$). Random sampling of each number of monitoring sites was repeated 30 times to obtain a statistical distribution of a particular network configuration, resulting in 990 unique networks (3 network designs × 11 network sizes × 30 random replications). Table 5.1 summarises the configurations of all the networks examined. As a further network sensitivity test, different proportions of roadside and HH sites within the *Mixed* network were investigated to evaluate the effect of network composition in estimating residential $NO_2$ concentration. Table 5.2 summarises the different network compositions investigated.

*Table 5.1 Summary of the configurations of the different networks evaluated. Each combination of type and size of network comprised sites randomly selected from the pool of potential monitoring sites of that designation, whose locations are shown in Figure 5.2. The random selection for each combination was repeated 30 times to give 990 unique networks evaluated in total.*

| **Total number of monitoring sites in each network** | ***HH* network** | ***Road* network** | | ***Mixed* network** | |
|---|---|---|---|---|---|
| | High Household density sites | Roadside sites | Background sites | Roadside sites | High Household density sites |
| **10** | 10 | 5 | 5 | 5 | 5 |
| **15** | 15 | 8 | 7 | 8 | 7 |
| **20** | 20 | 10 | 10 | 10 | 10 |
| **25** | 25 | 13 | 12 | 13 | 12 |
| **30** | 30 | 15 | 15 | 15 | 15 |
| **35** | 35 | 18 | 17 | 18 | 17 |
| **40** | 40 | 20 | 20 | 20 | 20 |
| **45** | 45 | 23 | 22 | 23 | 22 |
| **50** | 50 | 25 | 25 | 25 | 25 |
| **55** | 55 | 28 | 27 | 28 | 27 |
| **60** | 60 | 30 | 30 | 30 | 30 |

*Table 5.2 Summary of the configurations of Mixed networks constructed from different proportions of roadside (Rd) and high household density (HH) sites. Each combination of type and size of the Mixed network comprised sites randomly selected from the pool of potential Rd and HH monitoring sites. The random selection for each combination was repeated 30 times to provide statistics on the variability associated with a particular network configuration.*

| **Percentage of Roadside sites in the mixed network (%)** | **Network size = 20** | | **Network size = 30** | | **Network size = 40** | |
|---|---|---|---|---|---|---|
| | Rd sites | HH sites | Rd sites | HH sites | Rd sites | HH sites |
| **0** | 0 | 20 | 0 | 30 | 0 | 40 |
| **10** | 2 | 18 | 3 | 27 | 4 | 36 |
| **30** | 6 | 14 | 9 | 21 | 12 | 28 |
| **50** | 10 | 10 | 15 | 15 | 20 | 20 |
| **70** | 14 | 6 | 21 | 9 | 28 | 12 |
| **90** | 18 | 2 | 27 | 3 | 36 | 4 |
| **100** | 20 | 0 | 0 | 30 | 40 | 0 |

The *HH* networks were designed in the anticipation that such networks would more accurately estimate concentrations at most residential addresses. However this sampling design might under-predict concentrations for a small fraction of population who live close to roads. The *Road* networks, being a mixture of roadside and background sites, should capture the greatest $NO_2$ variation in the study area; this is the network site selection design used in many monitoring campaigns (Beelen et al., 2013). The *Mixed* networks of roadside and HH sites aimed to capture similar spatial variation of $NO_2$ as the *Road* network, but also to represent where most of the population live. This sampling design resembles the concept of a formal methodology for locating monitoring sites (Kanaroglou et al., 2005), namely locating monitors where the expected pollution spatial variability and density of the study subjects are high. Unlike the formal methodology, however, the sampling design here does not require prior knowledge of the pollutant concentration surface, therefore the application of this sampling design is less restricted.

### 5.2.3 Stage 3 – LUR modelling

#### 5.2.3.1 Predictor variables

A total of 15 predictor variables were selected for model development (Table 5.3). These variables were chosen based on prior knowledge that they may correlate with the input emissions in ADMS-Urban and their inclusion in previous LUR models for $NO_2$ (Beelen et al., 2013). As shown in Figure 5.3, $NO_x$ emissions for each of the 1 $km^2$ grids in the study area are mostly dominated by road transport and combustion in commercial/residential sectors. The total road length, population counts and building plan area within a buffer radius are considered to reflect these emissions. In addition, in some areas, $NO_x$ emissions from 'other' transport (most likely resulting from railways) are also significant. Therefore total railway length within a buffer was also included as a predictor variable. Since the emissions apart from major roads and some minor roads were modelled as 1 $km^2$ grid sources in ADMS-Urban, the buffer radii for the relevant predictor variables were chosen to be comparable with the resolution of input emissions, namely 0.5 and 1 km (Table 5.3). The rest of the predictor variables attempt to account for the increase in $NO_2$ concentration close to road sources (Table 5.3).

*Table 5.3 Predictor variables with buffer sizes and a priori defined directions of effect on $NO_2$ concentration.*

| **Predictor variables** | **Abbreviation** | **Unit** | **Buffer radius (m)** | **Direction of effect** |
|---|---|---|---|---|
| Population counts | POP | n | 500, 1000 | + |
| Building plan area | BA | $m^2$ | 500, 1000 | + |
| Total major and minor road length | RDLEN | m | 500, 1000 | + |
| Total railway length | RAILLEN | m | 500, 1000 | + |
| Distance to the nearest major road (inverse distance and inverse distance squared) | INVMJRDDIST, INVMJRDDIST2 | $m^{-1}$, $m^{-2}$ | NA | + |
| Distance to the nearest road (inverse distance and inverse distance squared) | INVDIST, INVDIST2 | $m^{-1}$, $m^{-2}$ | NA | + |
| Traffic volume on nearest major road | ALLTRAF | $veh.day^{-1}$ | NA | + |
| Product of traffic intensity on nearest major road and inverse distance to the nearest major road and inverse distance squared | TRAFDIST, TRAFDIST2 | $veh.day^{-1}$ $m^{-1}$, $m^{-2}$ | NA | + |

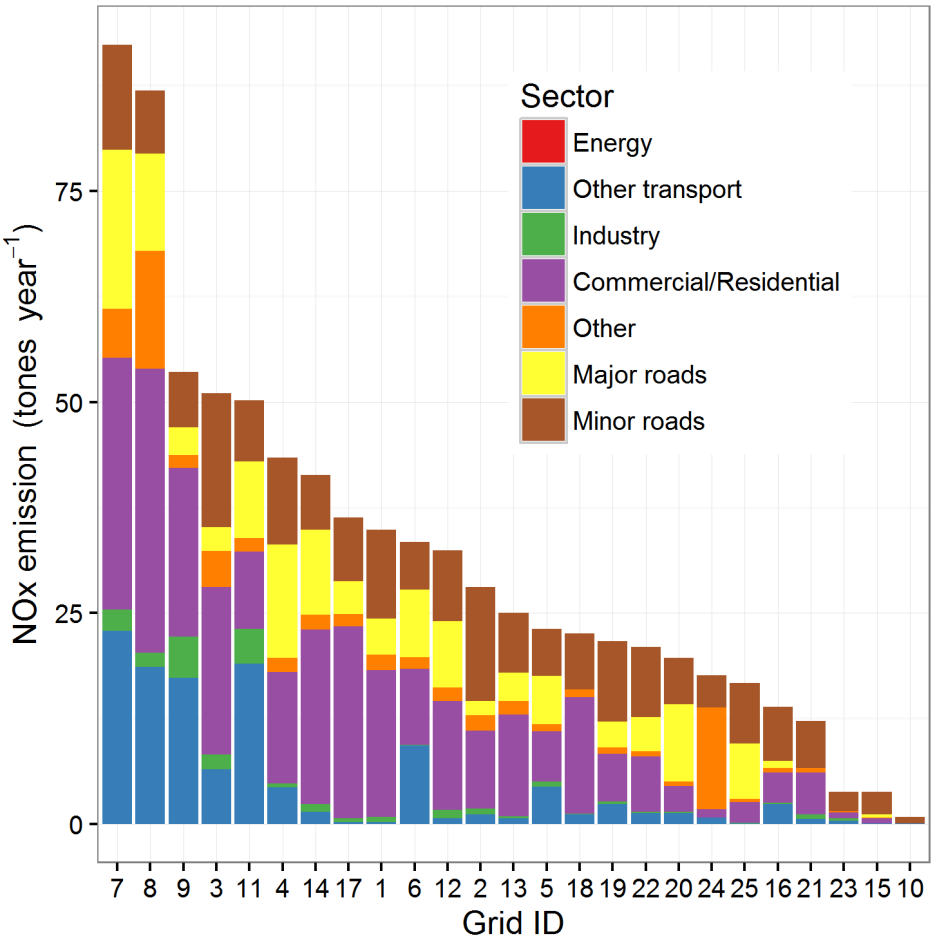



*Figure 5.3 Annual $NO_x$ emissions for the twenty-five 1 km× 1 km grids in the study area apportioned by source. The grid ID count goes left to right from top left to bottom right in Figure 5.2.*

### 5.2.3.2 LUR model development and diagnostics

The development of the LUR models followed the method used in the ESCAPE project (Beelen et al., 2013). The method is a supervised forward stepwise procedure which aims to maximise the adjusted $R^2$ of the model while also ensuring that the included variables are associated with coefficients with pre-defined directions (Table 5.3).

First, all variables were individually regressed against the $NO_2$ concentrations in that monitoring network. The variable with the highest adjusted $R^2$ and a coefficient with pre-defined direction formed the initial model. Second, the remaining variables were

successively added to the start model and the change in adjusted $R^2$ recorded. The variable resulting in the highest increase in adjusted $R^2$ was added to the model if: (i) the increase in adjusted $R^2$ was greater than 1%; and (ii) the coefficients of this variable and the variables already in the model conformed to the pre-defined direction. The selection process was continued until no variable fulfilled the above criteria. At the final step, variables with *p*-value greater than 0.1 were subsequently removed from the model starting from the variable with the highest *p*-value.

Diagnostic tests were performed on the final model. Multicollinearity in the variables was checked using Variance Inflation Factor (VIF). Predictors with high VIF value (> 3) were excluded from the model one at a time starting with the variable with the highest VIF. Potential influential observations were investigated using Cook's D value. An influential observation (indicated by a Cook's D > 1) was generally caused by including a variable with extreme values or many zero values. A sensitivity test was therefore conducted on a model with an influential observation problem by fitting a new model without using the observation with Cook's D > 1. If the change in the coefficient for that variable was large (over 100% of the coefficient derived from using all the observations), a new LUR model was developed following the above procedure but excluding that specific variable from the outset.

For the LUR model validation, leave-one-out-cross-validation (LOOCV) was used to assess the generalisability of the LUR model. LOOCV uses the variables in the final model to develop a regression model using $N - 1$ observations ($N$ = total number of observations in a monitoring network), which was then applied to the leave-out site. The procedure was repeated $N$ times at which point all the predicted concentrations are compared with the observations to test the validity of the model within the dataset. Values of $R^2$ and Root Mean Squared Error (RMSE) calculated from LOOCV were used to assess the LUR model's capability to predict the concentrations within a monitoring network.

### 5.2.4 Stage 4 – Evaluation of LUR model's capability at estimating simulated $NO_2$ concentrations at residential addresses

This aspect of LUR model evaluation compares the LUR modelled concentration at residential address with that modelled by ADMS-Urban. In essence this is similar to the concept of hold-out validation (HV) in a regression model validation, where the training data and testing data are completely independent. However, the validation dataset is based on ADMS-Urban output and is of constant size and much larger (7,445 residential addresses) than the traditional HV validations based on measurement data. In this context, the evaluation results not only reflect the performance of the LUR model but also indicate the relative effectiveness of the underlying monitoring sites used to build the LUR model. $R^2$, RMSE and Mean Bias (MB) were used here to evaluate the LUR modelled concentration for all population addresses and for different concentration ranges.

All GIS calculations were conducted in the Feature Manipulation Engine (FME) (Safe Software Inc., 2015). Statistical analyses were conducted in R software (R Core Team, 2015).

## 5.3 Results

### 5.3.1 ADMS-Urban model validation

ADMS-Urban was evaluated against measurements taken by both reference chemiluminescence analyser and passive diffusion tube (PDT). Comparison between the modelled annual average concentration of 2012 and the measurement by reference analyser at three monitoring stations in the study area showed that the bias was small at urban background (ED3) and minor roadside (ED7) (Table 5.4). The relatively large underestimation at major roadside (ED5) could be associated with the known issue of under-reporting of $NO_x$ from diesel vehicles (Carslaw and Rhys-Tyler, 2013).

*Table 5.4 Dispersion model versus measured annual average $NO_2$ concentration in 2012. ED3, ED5 and ED7 are three real-time monitoring stations (shown in Figure 5.2) located at urban background, major roadside and minor roadside, respectively.*

| | **Modelled $NO_2$ (μg $m^{-3}$)** | **Measured $NO_2$ (μg $m^{-3}$)** | **Model bias** |
|---|---|---|---|
| **ED3** | 24.8 | 24.1 | 2.8% |
| **ED5** | 35.4 | 39.1 | –9.5% |
| **ED7** | 27.5 | 28.1 | –2.0% |

A network of 30 PDT sampling sites covering a range of urban topographies (e.g. urban background, roadside, street canyon and traffic junction) within this study area were deployed weekly for 6 weeks during summer and winter periods of 2013/2014. Details of the site locations and characteristics can be found in (Lin et al., 2016). Seasonal average concentration (i.e. mean of 6 weekly average $NO_2$ concentrations) was compared with ADMS-Urban output. Sites containing more than one week's missing weekly $NO_2$ were excluded from the model evaluation. Some PDTs were located next to bus stops or at road junctions, where additional emissions from buses and queueing traffic are considered to be great but not modelled in the current model setup. These sites therefore do not reflect the general predictive ability of the ADMS-Urban model and were also excluded from the evaluation. Figure 5.4 shows the relationship between modelled and PDT-measured $NO_2$ concentrations during different seasons. Overall the model underestimated $NO_2$ concentrations compared to the PDT measurements. However the spatial variation in the measured $NO_2$ was explained very well by the model ($R^2$ = 73% and 77% for summer and winter, respectively) and was comparable to a previous ADMS model evaluation study (Dėdelė and Miškinytė, 2014). This indicates that although there is bias between modelled and PDT-measured $NO_2$ concentration the spatial pattern predicted by the model is consistent with the measurements. The bias could result from both the errors in the model and the errors in the PDT measurements. Large discrepancy (55% for summer and 82% for winter) between PDT measurement and reference analyser was observed during the deployment period at one co-location site (Table 5.5). This partly explains the general underestimation in the modelled $NO_2$ compared to the PDT measurements.

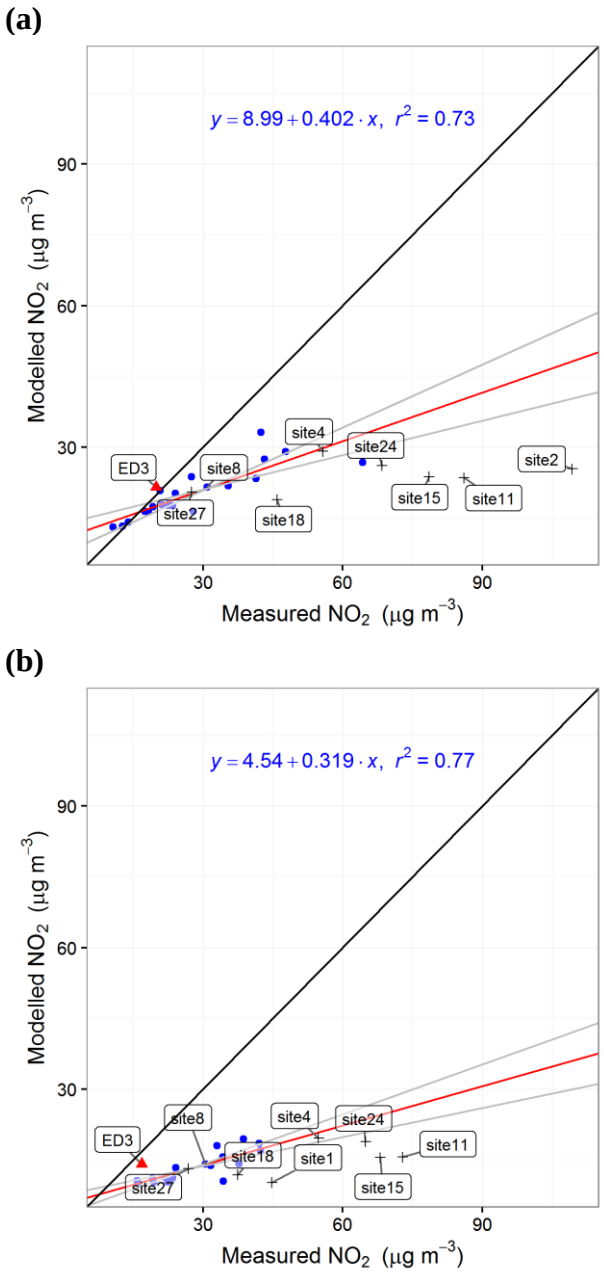



*Figure 5.4 Comparison of modelled and measured $NO_2$ concentration for summer (a) and winter (b) seasons. The cross markers denote the sites excluded from regression analysis due to special local effects as described in the text. Site 8 PDTs were co-located with reference analyser at ED3 marked by the red triangle. This site is also marked on Figure 5.2. The red and grey lines represent the linear fit and 95% confidence interval. The solid black line shows the 1:1 relationship.*

*Table 5.5 Comparison of seasonal average $NO_2$ concentrations measured by reference analyser and passive diffusion tubes (PDTs).*

| **Summer seasonal mean (µg $m^{-3}$)** | | |
|---|---|---|
| **Reference measured** | PDT measured | PDT overestimation |
| 19.9 | 30.9 | 55% |
| **Winter seasonal mean (µg $m^{-3}$)** | | |
| 16.8 | 30.5 | 82% |

Given the good agreement between the model and real-time analyser measurements at the urban background and minor roadside monitor locations, and the very good capture of spatial pattern indicated by the dense PDT network, it can be deduced that the dispersion model here fulfils the purpose of this study; that is, to simulate a realistic pollution surface of $NO_2$ for the evaluation of the LUR model validity and of the monitoring sites used to build the LUR model.

### 5.3.2 Evaluation of the LUR models constructed from different monitoring networks

The distributions of $NO_2$ concentrations at the locations of each type of monitoring site, and at all the population addresses, are summarised in Figure 5.5. Consistent with the expectations underpinning the network design principles, Figure 5.5 shows that a *Road* network (roadside sites + background sites) is likely to cover the whole range of concentration across the modelled domain, whereas a *HH* network (only HH sites) matches most closely the interquartile range of residential $NO_2$ concentration.

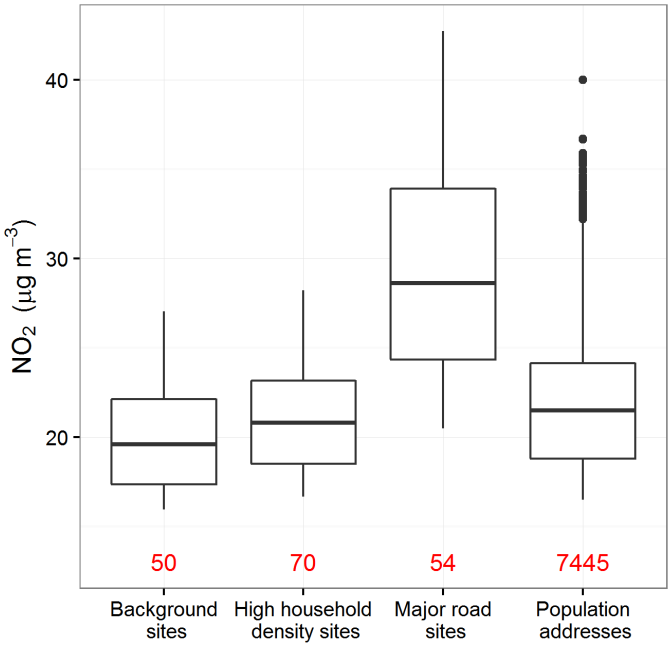



*Figure 5.5 Distributions of annual average $NO_2$ concentrations for the different types of potential monitoring sites and for all the population addresses. The length of the box denotes the inter-quartile range (IQR). The upper and lower whisker extend to the highest and lowest concentrations that are still within 1.5 × IQR of the upper and lower quartile. The number of data points contributing to each summary is shown beneath each box plot.*

Figure 5.6 summarises the following statistics evaluating LUR model performance as a function of network design and size: (i) the percentage of variance explained within the data used to build the LUR model (LUR $R^2$); (ii) the ability of the LUR model to predict the observed concentrations at the virtual monitoring sites (LOOCV $R^2$ and LOOCV RMSE); and (iii) the effectiveness of the monitoring networks at predicting concentrations at all the residential addresses (Residential $R^2$ and Residential RMSE). Figure 5.6 shows that LUR $R^2$ and LOOCV $R^2$ slightly decreased with increasing network size, while LOOCV RMSE slightly increased. In contrast, the effectiveness of the monitoring networks at predicting residential $NO_2$ concentration improved with increasing network size as shown by the increasing Residential $R^2$ and decreasing Residential RMSE (Figure 5.6). The improvement in the prediction of residential concentration (Residential $R^2$ and RMSE) was, however, insignificant between LUR models constructed with >30 monitoring sites, as indicated by the overlap of inter-quartile range of the statistic calculated from 30 random repetition.

The fact that the LOOCV $R^2$ was significantly higher than the Residential $R^2$ across the network size for *Road* and *Mixed* networks (Figure 5.6a) suggests that using LOOCV to evaluate the LUR model's predictive ability might be overly optimistic. The contrast between the performance of the LUR model and its ability to predict residential $NO_2$ concentration was especially large for the *Road* network design (comprising a mixture of roadside and background sites), and for the other network designs when there were only 10 or 15 monitoring sites (Figure 5.6a). The most effective type of monitoring network was the *Mixed* network, as indicated by the highest Residential $R^2$ limit and lowest Residential RMSE limit. The variability of Residential $R^2$ and RMSE in the 30 random repetitions of each network configuration (whiskers in Figure 5.6) decreased with increasing network size, suggesting that larger number of monitoring sites better capture the actual relationship between predictor variables and $NO_2$ concentration, hence less between-LUR-model variabilities.

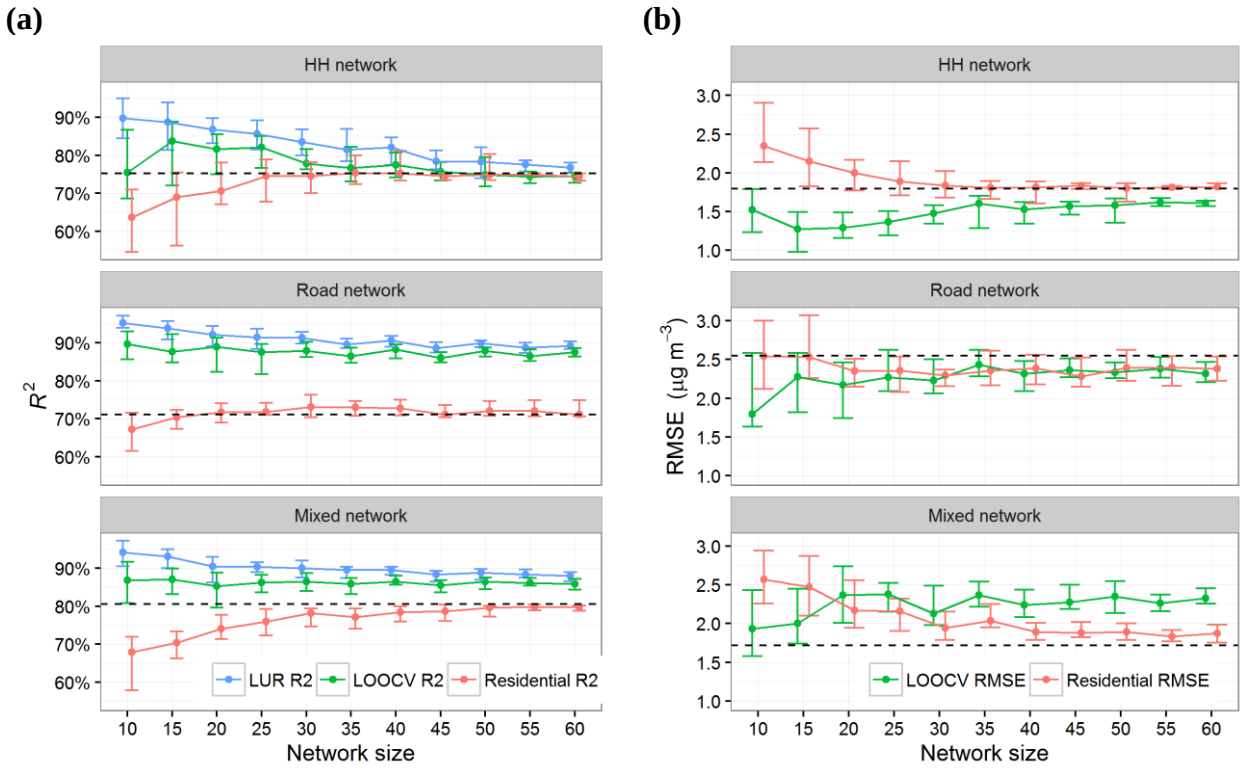



*Figure 5.6 Diagnostic statistics for LUR models as a function of network design and size for simulating network site concentrations (LUR $R^2$ and LOOCV $R^2$ shown in (a), LOOCV RMSE shown in (b)); and for predicting residential $NO_2$ concentration (Residential $R^2$ in (a); Residential RMSE in(b)). The points represent the median of the statistics for the 30 random repetitions of each network configuration. The whiskers extend to 25$^{th}$ and 75$^{th}$ percentiles of the statistics for the 30 random repetitions of each network configuration. The horizontal dashed lines denote the Residential $R^2$ and RMSE if all the potential monitoring sites (70, 104 and 124 for HH, Road and Mixed networks, respectively) are used for calculation.*

The performances of the LUR models in estimating residential concentration within three separate ranges of $NO_2$ concentration are compared in Figure 5.7. At the low end of $NO_2$ concentration (<20 µg m$^{-3}$), RMSE was similar between *Mixed* and *HH* networks, but both *HH* and *Road* networks significantly overestimated (MB) the overall residential $NO_2$ concentration. For $NO_2$ concentrations between 20 and 30 µg m$^{-3}$, the *HH* networks generally underestimated the residential concentration (Figure 5.7b). The most distinctive difference between the three network designs was observed at the high end of $NO_2$ concentration (> 30 µg m$^{-3}$). For these $NO_2$ concentrations, the prediction errors (RMSE) and the extent of overall underestimation (MB) were significantly higher for the *HH* networks (Figure 5.7).

*Mixed* and *Road* networks performed similarly at high $NO_2$ concentration, although they both still, on average, underestimated (Figure 5.7b). Similar to the statistics in Figure 5.6, the variability of estimation errors also reduced with increasing network size. Overall, considering the results shown in Figure 5.6 and Figure 5.7 together, the *Mixed* networks were most effective in estimating residential $NO_2$ concentration when considering both all residential addresses together and subsets of addresses in different ranges of $NO_2$ concentrations.

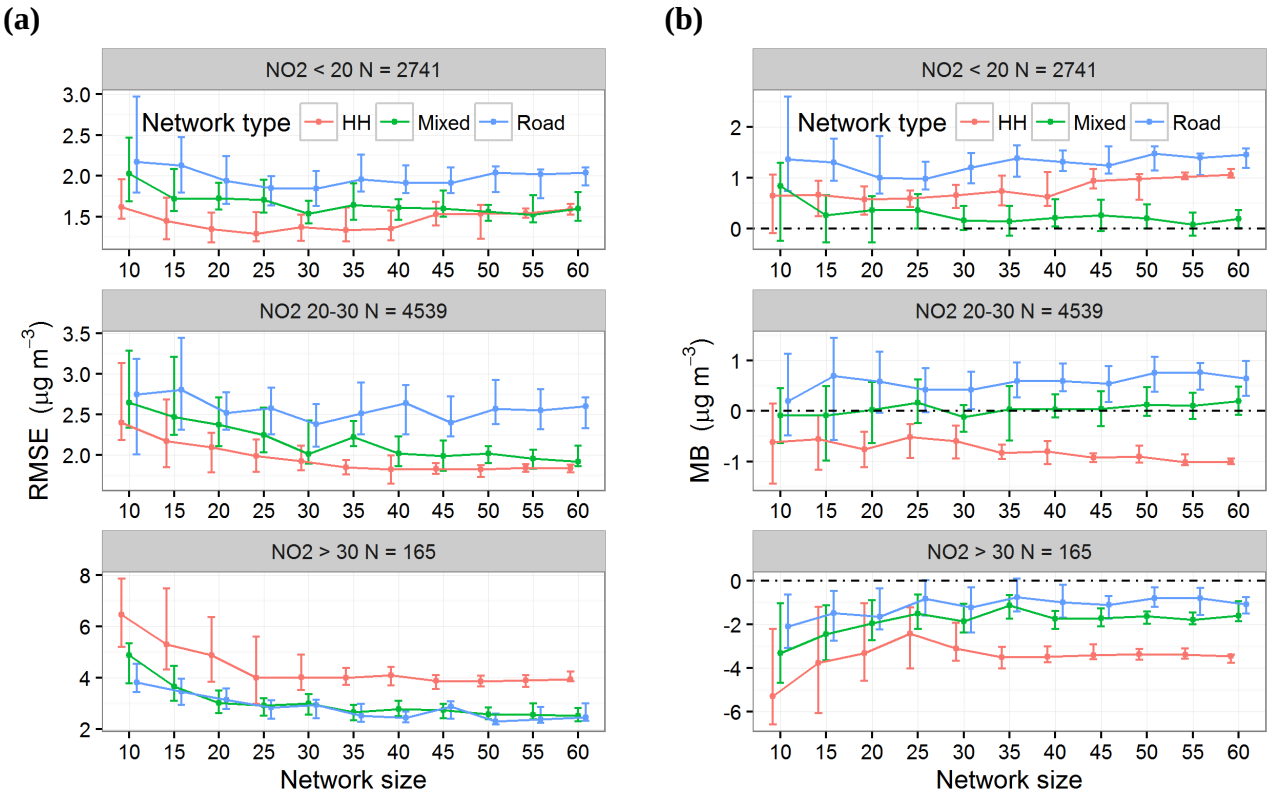



*Figure 5.7 Summary statistics of (a) RMSE and (b) MB in estimating residential concentration for different ranges in $NO_2$ concentration. The whiskers extend to 25$^{th}$ and 75$^{th}$ percentiles of the statistics for the 30 repetitions of each network configuration.*

Figure 5.8 shows the results of the investigation of the different proportions of HH sites and roadside sites within the *Mixed* network on the LUR model predictions of residential $NO_2$ concentrations. There was no significant trend in the $R^2$ values across the different proportions of roadside sites in the *Mixed* network (Figure 5.8a). The prediction error (RMSE) increased with increasing proportions of roadside sites

(Figure 5.8b). The variability of the RMSE resulting from networks consisting only of roadside sites was the largest, whereas the LUR models derived from networks with more HH sites were relatively more precise (Figure 5.8b).

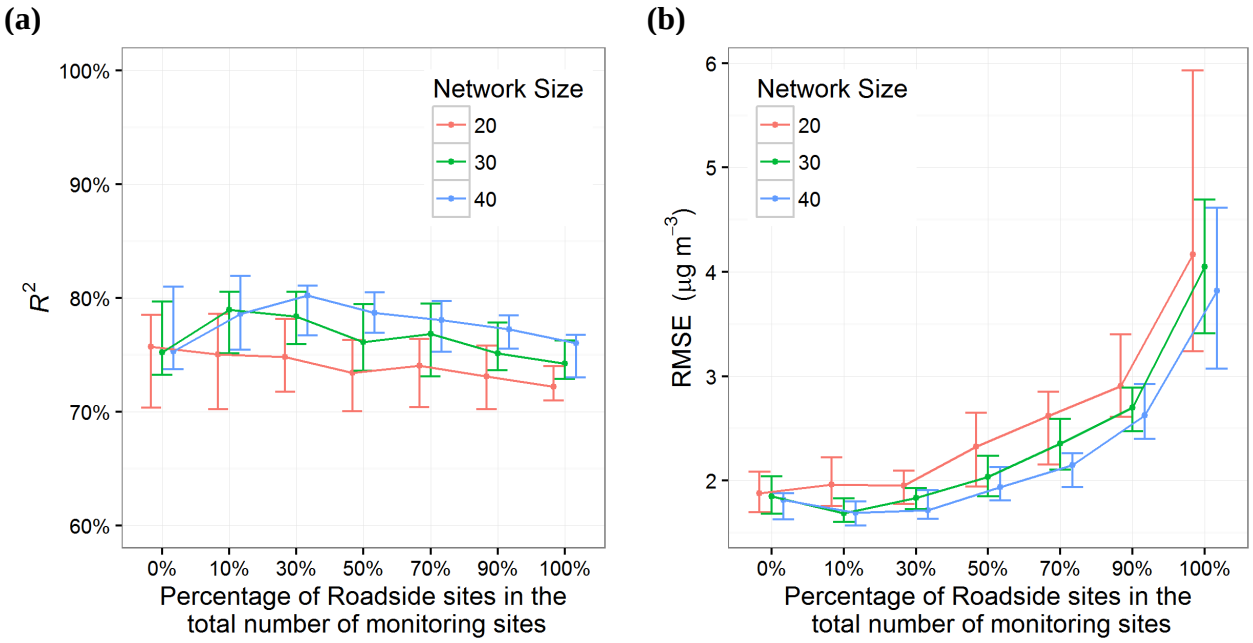



*Figure 5.8 Summary statistics of (a) $R^2$ and (b) RMSE for Mixed networks containing different proportions of roadside sites in predicting residential $NO_2$ concentration. The whiskers extend to 25th and 75th percentiles of the statistics for the 30 repetitions of each network configuration.*

The performance of networks containing different percentages of roadside sites at different $NO_2$ concentration ranges are compared in Figure 5.9. For $NO_2$ concentration <30 µg $m^{-3}$, the RMSE increased with increasing percentage of roadside sites (Figure 5.9a) and the MB suggested overestimation for networks with high roadside site composition (Figure 5.9b). However, at $NO_2$ concentration >30 µg $m^{-3}$, the RMSE decreased with increasing percentage of roadside sites (Figure 5.9a) and higher roadside site composition led to reduced bias (Figure 5.9b). Figure 5.9a shows that LUR models constructed with only roadside sites resulted in high variabilities in the RMSE at $NO_2$ concentration <30 µg $m^{-3}$, whereas LUR models constructed with all HH sites resulted in high variabilities in the RMSE at high level of $NO_2$ concentrations (>30 µg $m^{-3}$).

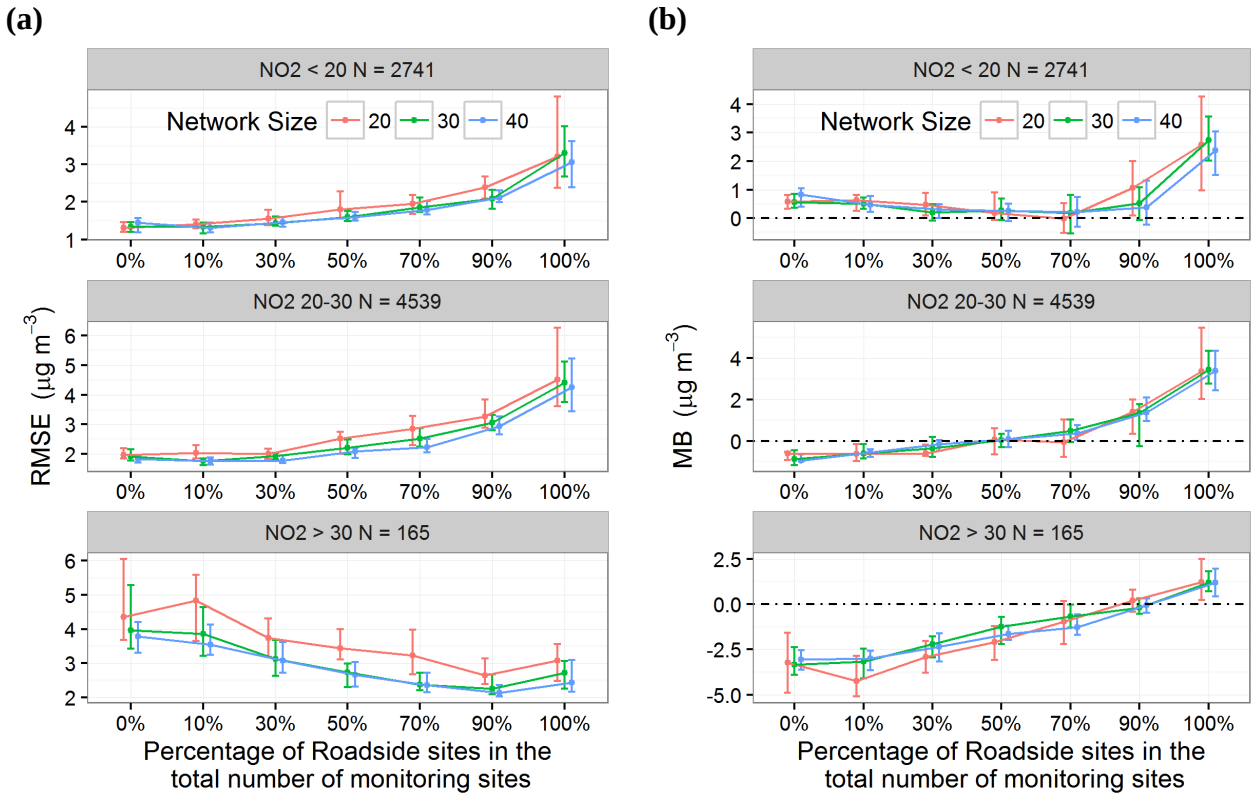



*Figure 5.9 Summary statistics of (a) RMSE and (b) MB in estimating residential concentration in different $NO_2$ concentration ranges for Mixed networks containing different proportions of roadside sites. The whiskers extend to 25$^{th}$ and 75$^{th}$ percentiles of the statistics for the 30 repetitions of each network configuration.*

## 5.4 Discussion

In most LUR studies, LOOCV and/or hold-out validation (dividing monitoring sites into two independent sets for model development and validation) have been used to validate the LUR models. LOOCV tests how well the LUR model predicts the observation within the training dataset. Hold-out validation evaluates the predictive ability of the LUR model at locations that were not used in model development. The latter evaluation is of more interest for an LUR model; in practice, however, there is always a trade-off between building a more robust LUR model using a larger training dataset and giving more power to the validation using a larger validation dataset. A limited number of monitoring sites in many studies makes the division of the dataset even more difficult. Evaluation of LUR models on all potential exposure subjects has been unfeasible in reality. However, this can be achieved by using a dispersion model to provide a realistic spatial field of urban ambient $NO_2$ concentration. Although there may be uncertainties in the dispersion-modelled concentrations, the nature of the errors should be similar at the virtual monitoring sites and at the residential addresses.

As expected, more monitoring sites yielded better estimation of residential $NO_2$ concentration by the LUR model (Figure 5.6). For all three network types, however, the improvement in the estimation was insignificant for networks with more than ~30 monitoring sites. Although the improvement was insignificant, higher number of monitoring sites increased the stability of the developed LUR models as shown by the very small inter-quartile range for the statistics at larger network sizes in Figure 5.6 and Figure 5.7. As the number of monitoring sites increased, the number of unique variables appearing in the LUR models decreased (Table 5.6), indicating that a greater number of monitoring sites was more effective at eliminating insignificant predictor variables. This is consistent with the findings of Basagaña et al. (Basagaña et al., 2012) using actual $NO_2$ measurements. In this work, it was found that ~30 observations are sufficient to capture the spatial variation of the residential $NO_2$ concentrations in a dispersion modelled pollution surface of an urban area of 25 $km^2$, but this number is expected to be larger in reality due to local effects (e.g. street canyon effect and traffic queueing) that were not modelled by ADMS-Urban and for

larger areas than simulated in this study. Basagaña et al. (2012) showed that the improvement of $R^2$ in hold-out validation was minor after ~60 monitoring sites in a study area of 45.7 $km^2$. In a national-wide study in the Netherlands, LUR models constructed with over ~90 monitoring sites seemed to result in similar prediction ability. Collectively these results suggest that a minimum optimal number of monitoring exists but depends on the actual study area.

*Table 5.6 Frequency of the predictor variables appearing in the final LUR models developed from Mixed networks with varying number of monitoring sties. Definitions of the variable abbreviations can be found in Table 5.3. NA indicates not appearing.*

| **Variable Name** | **% Appearance (*N* = 20)** | **% Appearance (*N* = 40)** | **% Appearance (*N* = 60)** |
|---|---|---|---|
| BA1000 | 57 | 80 | 93 |
| INVMJRDDIST | 40 | 70 | 77 |
| RAILLEN1000 | 50 | 60 | 57 |
| RAILLEN500 | 27 | 37 | 43 |
| TRAFDIST | 20 | 13 | 20 |
| RDLEN1000 | 33 | 17 | 7 |
| RDLEN500 | 7 | 3 | 3 |
| TRAFDIST2 | 17 | NA | 3 |
| INVDIST | 17 | 3 | *NA* |
| INVDIST2 | 7 | *NA* | *NA* |
| INVMJRDDIST2 | 3 | 13 | *NA* |
| POP500 | 7 | 3 | *NA* |

In the ESCAPE study (Beelen et al., 2013) and many other LUR studies (Aguilera et al., 2008; Madsen et al., 2007), urban background sites were selected in conjunction with roadside sites to build the LUR model. Urban background sites are usually defined with respect to the distance to road source or traffic activity within a certain buffer, irrespective of the distribution of the exposure study subjects, as was represented by the *Road* network design in this study. Figure 5.6a shows that LUR models derived from such networks were generally poorer at estimating $NO_2$ concentration at residential addresses than LUR models derived from networks with sites selected on the basis of household density. LUR models derived from *Mixed* networks were better at estimating residential $NO_2$ concentration than those derived

from *Road* networks (Figure 5.6) and also gave comparable errors to *Road* network-derived LUR models for estimating concentrations at the high end of the distribution (Figure 5.7). This observation emphasises the importance of characterising both the concentration and population distribution in the study area when designing a monitoring network.

The composition of different types of measurement sites in most monitoring networks used to construct LUR models to date has been rather arbitrary. Some researchers (Cyrys et al., 2012) followed the principle of over-representing the roadside sites with respect to the fraction of addresses close to the roads, as this captures the spatial variation of $NO_2$. In this work, LUR models constructed from networks containing 0 – 30% of roadside sites (compared with 0.2% of addresses within 10 m to the roads in the study area) showed lower estimation errors (Figure 5.8b) compared to other network compositions for all three network sizes tested. When examining the estimated residential concentrations at different $NO_2$ levels, LUR models constructed from networks containing 0 – 30% of roadside sites resulted in larger errors at high $NO_2$ concentrations compared to networks containing higher proportions of roadside sites (Figure 5.9). The results here suggest that a greater proportion of roadside sites in a monitoring network yielded LUR models that better characterised the higher end of the residential $NO_2$ concentration (Figure 5.9) but also introduced greater prediction error considering the population as a whole (Figure 5.8), and vice versa for LUR models derived from networks containing a greater proportion of HH sites. No particular network composition was simultaneously able to provide an LUR model capable of good overall estimation of the residential $NO_2$ concentrations and a good estimation of the higher end concentration. This illustrates the limitation of LUR models to capture the spatial contrast in residential $NO_2$ concentration predicted by the dispersion model.

As a common LUR model evaluation method, the LOOCV $R^2$ statistic was found to overestimate the LUR predictive ability, consistent with the limited number of other studies on the same topic (Basagaña et al., 2012; Johnson et al., 2010; Wang et al., 2012). Collectively, the results from these studies highlight the limited predictability of empirical $NO_2$ LUR models that are highly dependent on the measurement sites.

Dispersion modelling, as demonstrated in this study, is a potentially useful tool to design an effective monitoring network and to better evaluate the LUR models in a way that is otherwise unfeasible in reality.

It is acknowledged that the area of the domain in this study (25 $km^2$) is smaller than in some LUR studies (Aguilera et al., 2008; Fernández-Somoano et al., 2011). This choice was mainly limited by the intensive computational requirement of the dispersion model to calculate concentration at the large number of residential addresses. Clustering addresses with similar characteristics would reduce the calculation time and facilitate dispersion modelling over a larger area. In this study, a dispersion model provided the $NO_2$ concentrations for development of the LUR models and for evaluation of their predictive capabilities. Whilst accepting potential discrepancies between dispersion model and real measurements, this work shows that more comprehensive evaluation of LUR models and their underpinning monitoring networks is needed. Although the LUR models were only evaluated for $NO_2$, results for the effect of number and type of monitoring sites on LUR model performance should be transferable to other traffic-related air pollutants such as black carbon and ultrafine particle number, given their mutual high correlations.

## 5.5 Conclusions

Using a greater number of sites to build an LUR model improved its ability to estimate residential $NO_2$ concentrations. However, improvement in LUR model predictive capability was not significant beyond a certain number of monitoring sites: the predictive capability achieved using ~30 monitoring sites was similar to that achieved using 70 – 100 monitoring sites, but a greater number of monitoring sites tended to decrease imprecision. LUR models constructed from a network design incorporating both high household density areas and roadside sites better characterised the full range of residential concentrations and specifically those with highest concentrations. It is therefore recommended to incorporate monitoring sites representing most of the study subjects when designing of a monitoring network aimed at studying the health effects of air pollutants. The more roadside sites included in a monitoring network used to construct LUR model, the larger the RMSE

for the estimation of residential $NO_2$ concentrations, but the lower the estimation error for high $NO_2$ concentrations. The fact that no particular proportion of roadside sites within the network design estimated well both the overall residential concentration and higher level of $NO_2$ concentrations suggested a lack of spatial contrast in LUR modelled pollution surface. A dispersion model has been shown to be a useful tool for both designing a monitoring network for LUR models and for the evaluation of the LUR models. As a common LUR model evaluation method, the LOOCV $R^2$ statistic was overly optimistic on describing the LUR predictive ability.

# References

Aguilera, I., Sunyer, J., Fernández-Patier, R., Hoek, G., Aguirre-Alfaro, A., Meliefste, K., Bomboi-Mingarro, M.T., Nieuwenhuijsen, M.J., Herce-Garraleta, D., Brunekreef, B., 2008. Estimation of Outdoor NOx, NO2, and BTEX Exposure in a Cohort of Pregnant Women Using Land Use Regression Modeling. Environ. Sci. Technol. 42, 815–821. doi:10.1021/es0715492

Basagaña, X., Rivera, M., Aguilera, I., Agis, D., Bouso, L., Elosua, R., Foraster, M., de Nazelle, A., Nieuwenhuijsen, M., Vila, J., Künzli, N., 2012. Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmos. Environ. 54, 634–642. doi:10.1016/j.atmosenv.2012.01.064

Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M.-Y., Künzli, N., Schikowski, T., Marcon, A., Eriksen, K.T., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Tuomi, T., Declercq, C., Falq, G., Stempfelet, M., Birk, M., Cyrys, J., von Klot, S., Nádor, G., Varró, M.J., Dėdelė, A., Gražulevičienė, R., Mölter, A., Lindley, S., Madsen, C., Cesaroni, G., Ranzi, A., Badaloni, C., Hoffmann, B., Nonnemacher, M., Krämer, U., Kuhlbusch, T., Cirach, M., de Nazelle, A., Nieuwenhuijsen, M., Bellander, T., Korek, M., Olsson, D., Strömgren, M., Dons, E., Jerrett, M., Fischer, P., Wang, M., Brunekreef, B., de Hoogh, K., 2013. Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. Atmos. Environ. 72, 10–23. doi:10.1016/j.atmosenv.2013.02.037

Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z.J., Weinmayr, G., Hoffmann, B., Wolf, K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., Vineis, P., Xun, W.W., Katsouyanni, K., Dimakopoulou, K., Oudin, A., Forsberg, B., Modig, L., Havulinna, A.S., Lanki, T., Turunen, A., Oftedal, B., Nystad, W., Nafstad, P., De Faire, U., Pedersen, N.L., Östenson, C.-G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K.T., Overvad, K., Ellermann, T., Eeftens, M., Peeters, P.H., Meliefste, K., Wang, M., Bueno-de-Mesquita, B., Sugiri, D., Krämer, U., Heinrich, J., de Hoogh, K., Key, T., Peters, A., Hampel, R., Concin, H., Nagel, G., Ineichen, A., Schaffner, E., Probst-Hensch, N., Künzli, N., Schindler, C., Schikowski, T., Adam, M., Phuleria, H., Vilier, A., Clavel-Chapelon, F., Declercq, C., Grioni, S., Krogh, V., Tsai, M.-Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Katsoulis, M., Trichopoulou, A., Brunekreef, B., Hoek, G., 2014. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. The Lancet 383, 785–795. doi:10.1016/S0140-6736(13)62158-3

Briggs, D., 2005. The Role of Gis: Coping With Space (And Time) in Air Pollution Exposure Assessment. J. Toxicol. Environ. Health A 68, 1243–1261. doi:10.1080/15287390590936094

Carslaw, D.C., Rhys-Tyler, G., 2013. New insights from comprehensive on-road measurements of NOx, NO2 and NH3 from vehicle emission remote sensing

in London, UK. Atmos. Environ. 81, 339–347. doi:10.1016/j.atmosenv.2013.09.026

CERC, 2016. ADMS-Urban User Guide 4.0 [WWW Document]. URL http://www.cerc.co.uk/environmental-software/user-guides.html (accessed 11-March-2016).

Cyrys, J., Eeftens, M., Heinrich, J., Ampe, C., Armengaud, A., Beelen, R., Bellander, T., Beregszaszi, T., Birk, M., Cesaroni, G., Cirach, M., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Dėdelė, A., Dimakopoulou, K., Eriksen, K., Galassi, C., Grąulevičienė, R., Grivas, G., Gruzieva, O., Gustafsson, A.H., Hoffmann, B., Iakovides, M., Ineichen, A., Krämer, U., Lanki, T., Lozano, P., Madsen, C., Meliefste, K., Modig, L., Mölter, A., Mosler, G., Nieuwenhuijsen, M., Nonnemacher, M., Oldenwening, M., Peters, A., Pontet, S., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Sugiri, D., Stephanou, E.G., Taimisto, P., Tsai, M.-Y., Vaskövi, É., Villani, S., Wang, M., Brunekreef, B., Hoek, G., 2012. Variation of NO2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study. Atmos. Environ. 62, 374–390. doi:10.1016/j.atmosenv.2012.07.080

Dėdelė, A., Miškinytė, A., 2014. Estimation of inter-seasonal differences in NO2 concentrations using a dispersion ADMS-Urban model and measurements. Air Qual. Atmosphere Health 8, 123–133. doi:10.1007/s11869-014-0272-9

DfT, 2015. City of Edinburgh [WWW Document]. Dep. Transp. - Traffic Counts. URL http://www.dft.gov.uk/traffic-counts/cp.php?la=City+of+Edinburgh (accessed 11-March-2015).

Fernández-Somoano, A., Estarlich, M., Ballester, F., Fernández-Patier, R., Aguirre-Alfaro, A., Herce-Garraleta, M.D., Tardón, A., 2011. Outdoor NO2 and benzene exposure in the INMA (Environment and Childhood) Asturias cohort (Spain). Atmos. Environ. 45, 5240–5246. doi:10.1016/j.atmosenv.2011.02.010

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. Atmos. Environ. 42, 7561–7578. doi:10.1016/j.atmosenv.2008.05.057

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., Giovis, C., 2005. A review and evaluation of intraurban air pollution exposure models. J. Expo. Sci. Environ. Epidemiol. 15, 185–204. doi:10.1038/sj.jea.7500388

Jerrett, M., Arain, M.A., Kanaroglou, P., Beckerman, B., Crouse, D., Gilbert, N.L., Brook, J.R., Finkelstein, N., Finkelstein, M.M., 2007. Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. J. Toxicol. Environ. Health A 70, 200–212. doi:10.1080/15287390600883018

Jerrett, M., Finkelstein, M.M., Brook, J.R., Arain, M.A., Kanaroglou, P., Stieb, D.M., Gilbert, N.L., Verma, D., Finkelstein, N., Chapman, K.R., Sears, M.R., 2009. A Cohort Study of Traffic-Related Air Pollution and Mortality in Toronto, Ontario, Canada. Environ. Health Perspect. 117, 772–777. doi:10.1289/ehp.11533

Johnson, M., Isakov, V., Touma, J.S., Mukerjee, S., Özkaynak, H., 2010. Evaluation of land-use regression models used to predict air quality concentrations in an

urban area. Atmos. Environ. 44, 3660–3668. doi:10.1016/j.atmosenv.2010.06.041

Kanaroglou, P.S., Jerrett, M., Morrison, J., Beckerman, B., Arain, M.A., Gilbert, N.L., Brook, J.R., 2005. Establishing an air pollution monitoring network for intra-urban population exposure assessment: A location-allocation approach. Atmos. Environ., 12th International Symposium, Transport and Air Pollution12th International Symposium, Transport and Air Pollution 39, 2399–2409. doi:10.1016/j.atmosenv.2004.06.049

Lin, C., Feng, X., Heal, M.R., 2016. Temporal persistence of intra-urban spatial contrasts in ambient NO2, O3 and Ox in Edinburgh, UK. Atmospheric Pollut. Res. doi:10.1016/j.apr.2016.03.008

Madsen, C., Carlsen, K.C.L., Hoek, G., Oftedal, B., Nafstad, P., Meliefste, K., Jacobsen, R., Nystad, W., Carlsen, K.-H., Brunekreef, B., 2007. Modeling the intra-urban variability of outdoor traffic pollution in Oslo, Norway—A GA2LEN project. Atmos. Environ. 41, 7500–7511. doi:10.1016/j.atmosenv.2007.05.039

Met Office, 2012. Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853-current). URL http://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0 (accessed 1-December-2015).

NAEI, 2015. UK Emissions Interactive Map [WWW Document]. Natl. Atmospheric Emiss. Inventory. URL http://naei.defra.gov.uk/data/gis-mapping

Ordnance Survey, 2015. OS MasterMap Topography Layer [Shape geospatial data]. URL http://digimap.edina.ac.uk/

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Safe Software Inc., S.S.I., 2015. FME [WWW Document]. URL http://www.safe.com/fme/fme-desktop/ (accessed 14-October-2015).

Wang, M., Beelen, R., Eeftens, M., Meliefste, K., Hoek, G., Brunekreef, B., 2012. Systematic Evaluation of Land Use Regression Models for NO2. Environ. Sci. Technol. 46, 4481–4489. doi:10.1021/es204183v

# Chapter 6 Conclusions and future work

## 6.1 Conclusions

The aims of this thesis were to explore various measurement and modelling techniques that have proven to be promising for improving current exposure assessment and to demonstrate their use in an urban setting. A number of portable air quality monitors were used to study the intra-urban variability of ultrafine particles (UFP), black carbon (BC) and $PM_{2.5}$ through a series of short-term measurements. Three types of low-cost portable monitors were evaluated in detail against reference instruments. High resolution modelling of $NO_2$ and $O_3$ concentrations in Edinburgh was investigated using ADMS-Urban, making use of UK-wide available input datasets. ADMS-Urban was then used as the basis for the development and evaluation of land-use regression (LUR) modelling which is commonly used in health studies as the exposure assessment method. This chapter integrates and broadens out the key conclusions from each chapter to discuss the capabilities and limitations of the application of the various techniques investigated here on human exposure assessment.

### 6.1.1 Implication for exposure assessment from measurement studies

Any analytical measurement is subject to some sources of uncertainties that include variable response (i.e. variability between duplicate instruments or long-term drift of the same instrument), bias to the true value, interference from other variables and increasing uncertainty when operating close to limit of detection. It needs to be recognised that eliminating or controlling these and other potential sources of errors requires resource and effort. Due to the low-cost nature of emerging portable monitors, such resources and efforts will by design be limited. Therefore, it should be expected that lower-cost monitors are likely to have greater problems and be subject to greater uncertainties than the reference instruments that have high capital and on-going operational costs. A framework for testing and benchmarking the

commercial air-quality monitors against the most accurate monitors is essential to assure fair use of monitors of varying capabilities.

This thesis uncovered some technical and practical challenges facing the evaluation of three types of commercial air quality monitors. These monitors include Aeroqual S500 $NO_2$ and $O_3$ monitors and RTI microPEM monitor for measuring $PM_{2.5}$. None of the three monitor types yielded data immediately comparable to their respective reference instruments. Therefore, some form of calibration is required if the outputs from these monitors are intended to be compared with reference analyser concentrations. The relationship between all three monitor types and their respective reference instruments varies to different degrees with time, which lessens the confidence in applying the calibration equation derived from one deployment to another. Due to the uncertainty in deriving a stable calibration equation to correct the data from the monitors tested here, by definition they cannot be accepted as an "indicative" method (Spinelle et al., 2013).

The portability of these monitors means that they can be used in a mobile fashion, which potentially adds to instability of the measurement by physically moving about. Strictly speaking, evaluation of the data from a portable monitor during a mobile measurement requires comparison with its respective reference instrument whilst on the move, which is, in practice, not feasible. Exposure studies utilising mobile measurements from portable monitors therefore should focus on demonstrating that the relative trend in a given set of mobile measurements is reliable and that the relationship between duplicate monitors is consistent. On this matter, the three types of monitors evaluated in this study are shown to be useful due to their generally good correlation with the reference measurements and relatively stable relationship between duplicate units.

Short-term mobile measurements with portable monitors revealed high intra-urban variability of UFP number concentrations and BC concentrations. Both UFP and BC were highly correlated with traffic counts, which supports the use of traffic variables to explain UFP and BC concentrations in the land-use regression (LUR) models. However, $PM_{2.5}$ mass concentration and particle number concentration in the 0.5 – 2.5 μm size range were mainly driven by synoptic meteorological events that

governed the transport of regional sources. Local traffic also contributed to the $PM_{2.5}$ concentrations, but to a much lesser extent than long-range transport of secondary aerosols. This results in greater temporal variability but a more spatially homogeneous distribution of $PM_{2.5}$ concentrations. Since LUR models are usually derived from a series of weekly or monthly average concentrations to represent the annual average concentration, it is especially important for $PM_{2.5}$ that the representativeness of measurement periods selected to calculate the annual average is checked against available continuous monitoring data.

Particle number concentration (UFPNC) and surface area correlated well with $NO_x$ at both background and roadside localities, but to a varying degree with $NO_2$ because the partition of $NO_2$ in $NO_x$ varies in relation to the distance to traffic source. Although both UFPNC and $NO_x$ can be a good proxy for traffic-related air pollution, the relationship between UFPNC and $NO_x$ varied depending on the particle size distribution. Despite the high traffic-dependency of UFPNC, the daily averages of UFPNC showed good correlation between an urban background site and a roadside site. This suggests that the measurement at the urban background site adequately represents the variation of daily averages of UFPNC at other urban environments, although it may misrepresent the absolute value at locations closer to road sources. This justifies the use of measurement from a fixed monitoring site in time-series studies even for a pollutant with high spatial variability such as UFP.

### 6.1.2 Implications for exposure assessment from modelling studies

ADMS-Urban modelled daily and monthly average $NO_2$ and $O_3$ concentrations showed good correlation with reference measurements using UK-wide available emission and meteorological data at monitoring sites that were not subject to specific local effects, such as queueing traffic or bus stops. Additional measures should be taken to incorporate such local effects for the assessment of local air pollution. Given the large contribution or regional/long-range transport of aerosols to local $PM_{2.5}$ concentrations, it is anticipated that ADMS-Urban, which only uses emissions locally from a study area, will not accurately model the temporal variation of $PM_{2.5}$

concentrations. Modelling of $PM_{2.5}$ particularly requires nesting a local dispersion model in a regional chemical transport model.

The capability of ADMS-Urban at modelling the long-term spatial pattern of $NO_2$ was demonstrated in a comparison against data from a network of diffusion tube measurements. This capability indicates that the ADMS-Urban model can be an effective and economical tool for planning time consuming and costly real-world measurement ahead of time. An example was demonstrated by using ADMS-Urban to design an optimal monitoring network for the development of a LUR model. The results of this study highlighted that the design of a monitoring network should be tailored for the purpose of the LUR model, which can be for compliance purpose (i.e. to quantify pollutant concentrations at roadside where it is most likely to breach the air quality standards) or for exposure assessment purpose (i.e. to quantify pollutant concentrations at the available health data units). It was found that LUR models developed from a monitoring network consisting of monitoring sites located at densely populated areas and roadside sites resulted in better estimation of $NO_2$ concentrations at residential addresses. No particular proportion of roadside sites within the network design estimated well both the overall residential concentration and higher level of $NO_2$ concentrations, suggesting a lack of spatial contrast in the LUR-modelled pollution surface. In addition, the development of LUR models is usually more costly because of the extensive measurement data required. Whilst the dispersion model can provide higher temporal resolution estimate of pollutant concentration than the LUR models and at a lower cost, it is still unlikely to provide detailed personal exposure assessment due to the large uncertainties in the emission data at the resolution relevant to human activity (both in spatial and temporal aspects). In comparison, personal exposure monitoring perhaps is still the best way to quantify individual's exposure although it is limited by the practicality of obtaining measurements from a large enough cohort of individuals for statistical power in health studies.

## 6.2 Future work

Assessment of human exposure to ambient air pollution over the last decade has mainly been limited to assigning the outdoor concentration to the individuals' home addresses. This largely ignores the role of human activity in the determination of a person's exposure. The reason for the rather crude exposure assessment method in the past was mainly due to the lack of population mobility data and the imperfect high spatiotemporal resolution modelling. Future exposure assessment is moving towards a more dynamic approach than assigning exposure to the concentration at some fixed locations. Improving the exposure assessment not only relies on accurately quantifying the pollutant concentration at a given space and time but also on knowing when and where the interaction of air pollution with the person occurs.

Attempts have been made to couple local-scale atmospheric dispersion models with regional-scale atmospheric chemical transport models (Beevers et al., 2012; Pepe et al., 2016; Smith et al., 2016). The coupled model is a realistic way to simulate detailed pollutant concentrations within densely populated cities and also to provide pollutant concentrations at peripheral areas of the city at a coarser resolution. In theory a coupled model also has advantage over a local dispersion model by accounting for the contribution from regional sources, which are especially important for modelling $O_3$ and $PM_{2.5}$ concentrations. A few exposure assessment studies that take account of individual travel behaviour have also been emerging in recent years (Dhondt et al., 2012; Smith et al., 2016). These latter studies have shown a significant difference between an exposure assessment assuming that people are always at their home address and an exposure assessment considering population mobility. Since people spend over 90% of their time indoors, an exposure model should also accurately estimate the outdoor to indoor transmission factors for a given air pollutant and take account of significant indoor sources (Smith et al., 2016).

Validating an exposure model that incorporates an individual's activity requires personal exposure monitoring. Considering the large amount of time people spend indoors and the challenges in making mobile measurements with the portable monitors discussed above, it suggests that a monitoring strategy designed with

predominately indoor measurements at participants' homes and workplaces, complemented by targeted measurements in the transport micro-environment, may be preferable and more economical to attempting 24/7 exposure monitoring of study participants.

# References

Beevers, S.D., Kitwiroon, N., Williams, M.L., Carslaw, D.C., 2012. One way coupling of CMAQ and a road source dispersion model for fine scale air pollution predictions. Atmos. Environ. 59, 47–58. doi:10.1016/j.atmosenv.2012.05.034

Dhondt, S., Beckx, C., Degraeuwe, B., Lefebvre, W., Kochan, B., Bellemans, T., Int Panis, L., Macharis, C., Putman, K., 2012. Health impact assessment of air pollution using a dynamic exposure profile: Implications for exposure and health impact estimates. Environ. Impact Assess. Rev. 36, 42–51. doi:10.1016/j.eiar.2012.03.004

Pepe, N., Pirovano, G., Lonati, G., Balzarini, A., Toppetti, A., Riva, G.M., Bedogni, M., 2016. Development and application of a high resolution hybrid modelling system for the evaluation of urban air quality. Atmos. Environ. 141, 297–311. doi:10.1016/j.atmosenv.2016.06.071

Smith, J.D., Mitsakou, C., Kitwiroon, N., Barratt, B.M., Walton, H.A., Taylor, J.G., Anderson, H.R., Kelly, F.J., Beevers, S.D., 2016. London Hybrid Exposure Model: Improving Human Exposure Estimates to NO2 and PM2.5 in an Urban Setting. Environ. Sci. Technol. 50, 11760–11768. doi:10.1021/acs.est.6b01817

Spinelle, L., Aleixandre, M., Gerboles, M., 2013. Protocol of evaluation and calibration of low-cost gas sensors for the monitoring of air pollution. (No. EUR 26112 EN). European Commission, Luxembourg.

# Appendix I

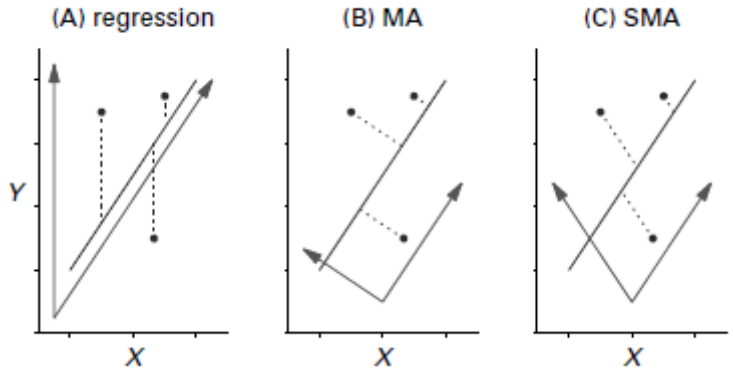
## Regression analysis of bivariate data

A topic that is frequently mentioned in this thesis is the regression analysis of bivariate data. A common regression method, ordinary least squares (OLS) regression, is widely used as the default regression method, such as in Microsoft Excel. This method finds the best linear fit in the bivariate data by minimising the sum of squares of residuals in the *y* direction (Appendix I Figure 1a), assuming that there is error only in the *y* variable (Warton et al., 2006). This method is most suitable in the calibration context (e.g. determination of the relationship between concentration and absorbance using UV spectroscopy) where the error in the standards is anticipated to be much smaller than the error in the instrument response. However applying the OLS regression in data where there are comparable errors in both *x* and *y* data may result in underestimation of the slope and overestimation of the intercept (Davis, 2011).

A method that takes into account the uncertainty in both sets of paired data is the major axis (MA) regression (also known as orthogonal regression). This method finds the best linear fit by minimising the sum of squares of residuals perpendicular to the line (Appendix I Figure 1b) (Warton et al., 2006). This method assumes that the uncertainty in the *x* and *y* variables is similar, which is reasonable when considering two variables on the same scales. This is also the method specified by EC (2010) for testing of equivalent methods for ambient air monitoring.

Regression analysis for bivariate data with different scales is better conducted with standardised major axis (SMA) [also known as reduced major axis (RMA)] regression. The SMA regression standardises the data first, which ensures the variables in *x* and *y* directions are on the same scales. The MA regression is calculated on standardised data, then rescaled to the original axes (Warton et al., 2006). The direction of the residual in SMA is measured as best fit line reflected about the *y* axis, as shown in Appendix I Figure 1c.

Appropriate method is used in this thesis for deriving linear relationship between bivariate datasets based on the suitability of each regression analysis method discussed above. For example the calibration between duplicate instruments or between portable air quality monitors and reference instruments is conducted with MA regression. The relationships between UFP and BC or $NO_x$ are investigated with RMA regression.




*Appendix I Figure 1 The direction in which residuals are measured is (A) vertical for linear regression (B) perpendicular to the line for major axis estimation (C) the fitted line reflected about the Y axis for standardised major axis estimation. Axes are plotted on the same scale. The broken lines indicate residuals, and the arrows represent the fitted and residual axes, which are useful for understanding methods of estimation and inference about these lines. (Source: Warton et al. (2006))*

## Evaluation statistics for pairwise data

This section details the most relevant statistics used throughout the thesis for evaluation of the correlation and agreement between two sets of data. In the following equations, $n$ represents the number of data points, $Y_i$ and $X_i$ represent the $i$th pair of data values, $\bar{Y}$ and $\bar{X}$ are the means of the two variables, and $\sigma_Y$ and $\sigma_X$ are the standard deviations of the two variables.

**Correlation coefficient (*r*)**

The Pearson correlation coefficient *r* measures the strength of the linear relationship between two variables and is defined as:

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{Y_i - \bar{Y}}{\sigma_Y}\right)\left(\frac{X_i - \bar{X}}{\sigma_X}\right)$$

For a perfect positive linear relationship between two variables, $r = 1$. For a perfect negative linear relationship between two variables, $r = -1$. If there is no linear relationship between the variables, $r = 0$.

**Fraction of paired data within a factor of two (FAC2)**

FAC2 quantifies the fraction of paired data that satisfy:

$$0.5 \leq \frac{Y_i}{X_i} \leq 2.0$$

**Mean Bias (MB)**

Mean bias provides an indication of the mean absolute difference between the two variables and is defined as:

$$MB = \frac{1}{n}\sum_{i=1}^{n} Y_i - X_i$$

**Normalised Mean Bias (NMB)**

Normalised mean bias is useful to compare between two sets of paired data of different scales on a relative basis and is defined as:

$$NMB = \frac{\sum_{i=1}^{n} Y_i - X_i}{\sum_{i=1}^{n} X_i}$$

**Root Mean Squared Error (RMSE)**

Root mean squared error quantifies overall how close the magnitudes of two variables are on a pairwise basis and is defined as:

$$RMSE = \left( \frac{\sum_{i=1}^{n} (Y_i - X_i)^2}{n} \right)^{\frac{1}{2}}$$

# References

Davis, J.C., 2011. Statistics and data analysis in geology, 3rd ed. John Wiley & Sons.

EC, 2010. Guide to the demonstration of equivalence of ambient air monitoring methods. European Commission. URL http://ec.europa.eu/environment/air/quality/legislation/pdf/equivalence.pdf

Warton, D.I., Wright, I.J., Falster, D.S., Westoby, M., 2006. Bivariate line-fitting methods for allometry. Biol. Rev. 81, 259–291. doi:10.1017/S1464793106007007

# Appendix II

*Appendix II Table 1 Summary of regression coefficients (with 95% confidence interval) between Aeroqual $O_3$ monitor and reference analyser.*

| Period | Intercept (95% confidence interval) | Slope (95% confidence interval) |
|---|---|---|
| **P1** | 3.1 (0.68, 5.42) | 1.22 (1.18, 1.26) |
| **P2** | 13.97 (11.42, 16.37) | 1.13 (1.07, 1.19) |
| **P3** | 5.9 (3.5, 8.17) | 1.07 (1.01, 1.13) |
| **P4** | -12.29 (-14.98, -9.76) | 1.23 (1.17, 1.29) |
| **P5** | -12.2 (-13.38, -11.05) | 1.1 (1.07, 1.12) |
| **P6** | -20.76 (-22.46, -19.12) | 1.24 (1.2, 1.28) |

*Appendix II Table 2 Summary of regression coefficients (with 95% confidence interval) between duplicate microPEMs.*

| Period | Intercept (95% confidence interval) | Slope (95% confidence interval) |
|---|---|---|
| **RH-corrected microPEM** | | |
| **P1.1** | 0.036 (-0.24,0.30) | 1.23 (1.20, 1.27) |
| **P1.2** | 0.16 (0.025, 0.29) | 1.15 (1.13, 1.18) |
| **P1.3** | -2.64 (-2.78, -2.51) | 1.19 (1.18, 1.21) |
| **P2.1** | 0.90 (0.71, 1.10) | 1.11 (1.10, 1.13) |
| **P2.2** | -0.43 (-0.94, 0.038) | 1.28 (1.21, 1.36) |
| **P3.1** | -3.94 (-4.57, -3.37) | 1.42 (1.30, 1.55) |
| **P3.2** | -2.23 (-2.52, -1.96) | 1.22 (1.16, 1.29) |
| **P3.3** | -0.37 (-0.47, -0.27) | 1.04 (1.02, 1.06) |
| **RH-uncorrected microPEM** | | |
| **P1.3** | -2.74 (-2.88, -2.61) | 1.2 (1.19, 1.22) |
| **P2.1** | 1.47 (1.25, 1.69) | 0.99 (0.98, 1.01) |
| **P2.2** | -0.69 (-1.31, -0.12) | 1.4 (1.32, 1.49) |
| **P3.1** | -3.97 (-4.95, -3.12) | 1.21 (1.08, 1.36) |
| **P3.2** | -1.57 (-1.75, -1.41) | 0.93 (0.9, 0.96) |
| **P3.3** | -0.62 (-0.71, -0.52) | 1.15 (1.13, 1.16) |
| **RH measured by microPEMs** | | |
| **P1.1** | 3.8 (2.73, 4.84) | 0.94 (0.91, 0.96) |
| **P1.2** | 0.94 (0.29, 1.57) | 1.01 (0.99, 1.02) |
| **P1.3** | 1.14 (0.74, 1.54) | 1.03 (1.02, 1.04) |
| **P2.1** | 2.25 (1.22, 3.26) | 0.79 (0.77, 0.81) |
| **P2.2** | -3.25 (-6.04, -0.62) | 1.27 (1.2, 1.33) |
| **P3.1** | 0.69 (0.26, 1.12) | 0.93 (0.92, 0.94) |
| **P3.2** | 0.5 (-0.75, 1.72) | 0.85 (0.83, 0.87) |
| **P3.3** | 1.12 (0.39, 1.83) | 1.14 (1.12, 1.15) |

*Appendix II Table 3 Summary of regression coefficients (with 95% confidence interval) between microPEM 586N and TEOM-FDMS.*

| Period | Intercept (95% confidence interval) | Slope (95% confidence interval) |
|---|---|---|
| **RH-corrected microPEM vs TEOM-FDMS** | | |
| **P1.1** | -1.12 (-2.8, 0.26) | 1.1 (0.91, 1.33) |
| **P1.2** | 3.35 (2.08, 4.51) | 0.24 (0.04, 0.45) |
| **P1.3** | 1.48 (1.1, 1.86) | 0.78 (0.74, 0.82) |
| **P2.1** | 0.87 (0.01, 1.66) | 1.32 (1.22, 1.42) |
| **P2.2** | 1.06 (0.08, 1.88) | 0.96 (0.81, 1.14) |
| **P3.1** | 4 (3.49, 4.47) | 0.27 (0.13, 0.42) |
| **P3.2** | 3.71 (3.08, 4.3) | 0.23 (0.06, 0.41) |
| **P3.3** | -0.22 (-0.65, 0.2) | 0.93 (0.86, 0.99) |
| **RH-uncorrected microPEM vs TEOM-FDMS** | | |
| **P1.3** | 1.04 (0.61, 1.44) | 0.85 (0.82, 0.89) |
| **P2.1** | -0.13 (-1.33, 0.93) | 1.73 (1.59, 1.87) |
| **P2.2** | 0.64 (-0.49, 1.58) | 1.12 (0.95, 1.33) |
| **P3.1** | 4.12 (2.69, 5.15) | 0.66 (0.35, 1.09) |
| **P3.2** | -9.23 (31.41, 0.58) | 4.35 (1.51, -7.44) |
| **P3.3** | -1.46 (-2.15, -0.84) | 1.24 (1.14, 1.34) |