Development and evaluation of portable passive and real-time measurement systems, and dispersion models, to estimate exposure to traffic-related air pollutants

Nicola Masey

A Thesis presented for the Degree of Doctor of Philosophy

Department of Civil and Environmental Engineering
University of Strathclyde
2018

Abstract

This research developed efficient applications of portable measurement systems to assess human exposure to traffic-related air pollution through direct measurement, and evaluation of exposure models.

Passive NO_2 samplers are deployed at large numbers of sites in epidemiological studies to estimate typical concentrations over 1-4 weeks. I found that deployment time could be reduced to 2 days with limited impact on the accuracy and precision of exposure estimates. This shorter measurement time enabled observation of wind-speed effects leading to overestimation of ambient concentrations by passive samplers. Through development of a post-processing technique and/or inclusion of a membrane I improved sampler accuracy.

Portable sensors can provide detailed estimates of personal exposures to air pollution. Many sensor-based monitors have not been subject to rigorous testing procedures to quantify their accuracy. I observed that the most accurate estimates of concentrations from NO_2 and O_3 sensor-based monitors required regular, intermittent calibration against reference analysers under similar environmental conditions to field measurements. I also found deterioration in BC monitor accuracy and precison when the attenuation of the collection filter exceeded 40 and no improvement in monitor accuracy was observed when filter darkness correction algorithms were applied.

Portable sensors can be used to identify locations with higher concentrations, which may require more detailed monitoring. I established that repeated 6-minute measurements of BC and particle number concentrations estimated similar spatial trends to 1-week NO_2 measurements using passive samplers.

Dispersion models can be used to estimate pollution exposure at multiple locations over a study area. I found that initial user parameterisation in a weather model had limited effect on pollution estimates from a dispersion model. I evaluated a new GIS-based dispersion model (5 x 5 m NO_2 estimates for a 3,500 km^2 area, with model run times of under 10 minutes). I demonstrated that inclusion of discrete street canyon models and geospatial surrogates (accounting for urban morphology) improved model accuracy.

The measurement and modelling evaluation research in this thesis complimented each other by providing efficient ways to directly measure population exposures.

Declaration of Authenticity

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed: Nicola Masey

Date: 17 March 2018

Copyright

Acknowledgements

I would like to thank my supervisors Dr Iain Beverland, Dr Mathew Heal, and Dr Scott Hamilton for their technical advice and support throughout my PhD. I would also like to thank my fellow PhD student Jonathan Gillespie for mentoring me over the last few years and for his valuable insights and discussions on all aspects of my PhD.

Finally I would like to thank my family and friends, for all the emotional support and non-wavering encouragement which has ultimately helped me persevere to complete my PhD.

Contents

1. Introduction

Air pollution refers to any substances introduced in the air which, in high enough concentrations, have an impact on the environment. There is growing literature showing that air pollution exposure leads to detrimental impacts on human health and has been estimated to have caused 3 million deaths in 2012 (World Health Organization, 2016).

The smog in London in 1952, caused by emissions from power stations, transport and household heating, in combination with a period of very stable atmospheric conditions, lead to the mortality rates 40 % higher than for the same period the previous year, and continued to cause elevated rates for several months (M. L. Bell et al., 2004). The Clean Air Act 1956 was introduced to prevent this avoidable-disaster from occurring again, and included laws to prevent smoky fuels being burned in towns and cities, relocation of large industries to outside of urban areas and the use of taller chimneys on industrial sites. This act was the first to tackle the problem of air pollution, and since this time many others have followed, including the Ambient Air Quality Directive (2008/50/EC) in place today, which limits the concentrations of pollutants allowed in ambient air. Generally the trends in air pollutants have decreased over recent years as a result of the legislation introduced (Bigi and Harrison, 2010; Guerreiro et al., 2014).

Today, one of the main contributors to air pollution is from transport emissions (including air, road, rail and sea transport), which have been estimated to account for 21 % of the concentrations in the UK in 2015, up from 15 % in 1990. Ninety-three percent of these traffic emissions are estimated to be attributed to road transport (Department for Transport, 2015). Cleaner engine technology is reducing the concentrations of pollutants in vehicle exhaust fumes (Department for Transport, 2015). However despite this improvement ambient concentrations have not decreased by as much as these emissions reductions would have us to believe, leading to some urban areas being unable to meet the strict limit values for air pollution concentrations imposed by the European Union (for example 86 % of exceedance of EU ambient nitrogen dioxide limit values occurred at traffic sites (Guerreiro et al., 2014)). This could be attributed to an increase in the number of vehicles on the road and the number of vehicle miles travelled, inefficiencies in the emission testing procedure and an increase in the number of diesel vehicles on the road has lead to a slower rate of reduction in traffic-related air pollutants in recent years (Beevers et al., 2012).

Diesel vehicles are making up a larger proportion of the UK transport fleet, with an increase from 7 % to 36 % between 1994 and 2014 (Department for Transport, 2015). The proportion of petrol vehicles on the road is falling, with a 20 % decrease in vehicles

composition to 63 % between 1994 and 2014 (Department for Transport, 2015). Diesel vehicles produce lower emissions of carbon monoxide, but higher oxides of nitrogen and particulate emissions than petrol vehicles (European Environment Agency, 2016).

The test procedure for vehicle emissions has been suggested to be inefficient in determining real-world emissions which have been shown to be seven times greater for oxides of nitrogen in comparison to the legal limits for emissions concentrations for the most recent European Union emissions standards (Degraeuwe and Weiss, 2017; Franco et al., 2014; Williams and Carslaw, 2011). It has recently come to light that certain diesel vehicles had been fitted with a system to detect when an emissions test was being carried out and consequently cheat the system by reducing the vehicles emissions below the limit values under test conditions but not in real-world driving scenarios (Degraeuwe and Weiss, 2017).

The difference in vehicle emissions under test and real-world driving conditions will lead to elevated concentrations of air pollutants in areas in close proximity to traffic sources (Beevers et al., 2012; Brand, 2016). These elevated concentrations could result in greater health impacts on those breathing in this air, and, as previously mentioned, non-compliance with legislation controlling the concentrations of pollutants in ambient air. This failure to meet the limit values has resulted in a public lawsuit which found the UK governments guilty of not doing enough to tackle air pollution (Carrington, 2016).

1.1. Traffic-related air pollutants

The focus on this PhD is traffic-related air pollutants due to their large contribution to the overall ambient air pollutant concentrations. These pollutants can be emitted directly in the exhaust of vehicles, known as primary pollutants, or can be produced when primary pollutants react with other gases in the atmosphere, secondary pollutants. I discuss some of the common traffic-related air pollutants below, but limit these to the pollutants of focus in this thesis: oxides of nitrogen, ozone, and particulate matter. The health effects of the pollutants are also discussed, though it should be noted that the spatial and temporal auto-correlation between different pollutants can lead to confounding, making it difficult to assign health effects to a specific pollutant (Sheppard et al., 2012).

1.1.1. Oxides of nitrogen

Oxides of nitrogen (NO_x) are the collective name for the gaseous pollutants nitric oxide (NO) and nitrogen dioxide (NO_2). NO is formed at high temperatures, such as those in a vehicle engine, when oxygen (O_2) and nitrogen (N_2) gas in the air react with one another, or, to a lesser extent, when nitrogen present in fuel reacts with oxygen in the air (Equations 1 &

2 below). NO is a primary pollutant as it is released directly from vehicle engines. The NO emitted in exhaust fumes quickly reacts with ozone (O_3) in the atmosphere, in the presence of UV light from the sun's rays, to form NO_2, thus NO_2 a secondary pollutant (Equation 3 below). However, nitrogen dioxide can also be classed as a primary pollutant as this is emitted directly from diesel engines during the combustion. Nitrogen dioxide can break down in the presence of light, in other words the reaction only occurs during the day, to produce nitric oxide and ozone (Equation 4 below). This means these three pollutants are in photo equilibrium with one another, with NO and NO_2 concentrations elevated at locations in close-proximity to traffic and O_3 concentrations anti-correlated with these.

$$N_2 + O_2 \rightarrow NO + N \quad \text{[Equation 1]}$$

$$N + O_2 \rightarrow NO + O \quad \text{[Equation 2]}$$

$$NO + O_3 \overset{h\nu}{\rightarrow} NO_2 + O_2 \text{[Equation 3]}$$

$$NO_2 \overset{UV}{\rightarrow} NO + O_3 \quad \text{[Equation 4]}$$

Inhalation of nitrogen dioxide can lead to inflammation of the lungs, which in turn makes a person more susceptible to infection and gives rise to an increase in wheezing, difficulty breathing, hospital admissions and in some cases death (Bayer-Oglesby et al., 2006; Brauer et al., 2002; Mills et al., 2015). Those with pre-existing respiratory conditions, such as asthma, are at higher risk of experiencing these detrimental health effects (Lipsett et al., 1997; Trasande and Thurston, 2005). For the Greater Glasgow area, where the work which will be presented in this thesis was carried out, a reduction of 8 µg/m^3 in NO_2 concentrations was estimated to be able to lead to a reduction in hospital admission by approximately 10 % (Lee et al., 2009).

1.1.2. Ozone

Although not directly emitted in vehicle exhausts, ozone (O_3) is classed as traffic related pollutant as it is involved in reactions with primary pollutants such as oxides of nitrogen as discussed above (Equation 3 above). The natural process of ozone formation involves UV light rays from the sun break down oxygen gas into individual oxygen atoms, which then react with other molecules of oxygen to form ozone (Equations 5 & 6 below). It can also be produced when NO_2 in the air is photolysed, meaning ozone is anti-correlated with nitrogen dioxide (Equation 4 above). As ozone is generated during photochemical reactions, on sunny days ozone concentrations can become elevated quickly.

$$O_2 \overset{UV}{\rightarrow} O + O \quad \text{[Equation 5]}$$

$$O_2 + O \rightarrow O_3 \quad \text{[Equation 6]}$$

When breathed in, ozone can cause oxidisation of lung tissue, leading to irritation and constriction of the airways, shortness in breath, reduced lung functions and can cause increases in death rates (M. Bell et al., 2004; Doherty et al., 2009; Fann et al., 2012). Exposure to O_3 can also exacerbate pre-existing respiratory conditions such as asthma which can lead to increased hospital admissions, especially in children (Gent et al., 2003; McConnell et al., 2002; Petroeschevsky et al., 2001).

1.1.3.Particulate air pollution

Particulate air pollution is produced from both natural (e.g. forest fires) and anthropogenic (e.g. traffic) sources. Particulate matter (PM) includes solid and liquid particles produced during the combustion process and re-suspension of dust from the roads suspended in the air. There are several methods used to measure particulate pollution from road transport, including direct measurement of the particulate matter, a count of the number of particulate particles in a sample of air or measurement of black carbon which is a pollutant produced only by combustion sources. These three classes of particulate pollutant are discussed below in more detail.

1.1.3.1. Particulate Matter

Particulate matter can describe particles produced from a wide variety of sources, including natural (e.g. dust) and anthropogenic (e.g. combustion, industrial) sources. There are different size fractions of PM – commonly measured are PM_{10}, $PM_{2.5}$ and PM_1 - the numbers represents the size of aerodynamic particles being measured, for example $PM_{2.5}$ measures particles in the air with aerodynamic diameter less than 2.5 µm. Those particles produced during combustion have small diameters (less than 0.1 µm) and disperse in a similar manner to gases over large distances from their road sources, while the larger particles (e.g. PM10 produced during re-suspension) do not disperse as far from their source.

When inhaled, particles travel into the lungs and cause respiratory illness, with the smaller particles able to penetrate deeper into the lungs and cause greater impacts (Dockery et al., 1989). These small particles can also be absorbed from the lungs into the circulation, where they can lead to increased local and systemic inflammation which can lead to detrimental health effects including cardiovascular illness (Brook et al., 2010; Dockery et al., 1993; Oberdörster et al., 2005; Pope and Dockery, 2006). Black smoke concentrations (a historic measure of particle concentrations) have been shown to lead to increased all-cause and respiratory mortality in Greater Glasgow (Beverland et al., 2012a, 2012b) while

particulate concentration have been linked to hospital admissions in the same area (Lee et al., 2009). Reductions in particulate matter concentrations can lead to improvements in the respiratory health of children, however there has been no safe threshold concentration of particulate matter where no negative impacts on health have been observed (Bayer-Oglesby et al., 2005).

Another important factor associated with the exposure to particulate matter is the chemical constituents of the particles, which often contain metals that are toxic to human health (Valavanidis et al., 2008). If particles are absorbed into the blood stream these can lead to oxidative stress which can increase the risk of a stroke or cardiovascular illness (including coronary heart disease) (Lodovici and Bigagli, 2011), while a relationship between particular matter and lung cancer has also been demonstrated which is thought to be attributed to metals in the particles leading to DNA damage and promoting inflammation (Raaschou-Nielsen et al., 2016).

1.1.3.2. Particulate Number

Ultra-fine particles (UFP) are also produced from road traffic during the combustion process. Due to their small size (diameters of < 100 nm), direct measurement using mass based methods is difficult as the UFP mass is negligible in comparison to other, larger, particulates emitted from road traffic (Harrison et al., 2010; Kumar et al., 2010; Kumar et al., 2014). In comparison, particles with diameters less than 300 nm (i.e. including UFPs) have been estimated to contribute over 99 % of total particle numbers (Kumar et al., 2009) therefore particle number count has been used to represent UFP concentrations.

There small size causes UFP to behave similarly to gases – namely they can persist in the atmosphere for longer periods of time than larger particle sizes and they can disperse over larger areas than the particulate matter discussed above (Birmili et al., 2013; Constabile et al., 2009).

UFPs can penetrate deep into the lung where they can deposit (Manigrasso and Avino, 2012; Oberdorster, 2000), or alternatively can enter the blood stream where they can be transported to other organs including the brain (Oberdorster et al., 2004), leading to increases in hospital emissions, morbidity and mortality (HEI, 2013; Heusinkveld et al., 2016; Hoek et al., 2010; Li et al., 2003; Pagano et al., 1996). Toxic metals can also be attached to these pollutants which come with additional health problems including causing damage to cells (Canepari et al., 2013; Wiseman and Zereini, 2009).

1.1.3.3. Black Carbon

Particulate matter can be produced by a variety of sources. Recent studies have suggested that black carbon (BC), a constituent of particles, is a more representative metric to measure particles from combustion sources, including diesel road traffic. BC is an organic constituent of particulate matter which is produced during the incomplete combustion of fossil fuels and is one of the main constituents of soot.

The spatial variability of BC has been found to greater than $PM_{2.5}$ as a result of the localised emissions sources (Hoek et al., 2002; Janssen et al., 2008), meaning that fine-scale variations of BC concentrations are likely to be a better indicator of exposure to traffic-related air pollution than the equivalent variations in $PM_{2.5}$ concentrations.

Exposure to BC has been found to be significantly correlated with both all-cause and cardiovascular mortality and also cardiac hospital admissions. These associations were also found to be more pronounced for BC than $PM_{2.5}$, for example estimates of all-cause mortality per 1 $\mu g/m^3$ were 5-15 times higher for BC compared to $PM_{2.5}$ (Janssen et al., 2011).

1.2. Assessing population exposure to air pollution

To minimise the health effects described above governing authorities have set strict limit values for the main pollutants which must be adhered to. The limit values are set at a level at which there has been no observable adverse effects observed. Different concentrations have been set as limit values between the European Union, UK government and Scottish government. The limit values for traffic related air pollutants are shown in Table 1 below (DEFRA, 2007). The compliance with these limit values is assessed at several static monitoring stations located across the UK and funded by local authorities (https://uk-air.defra.gov.uk/networks/).

Table 1: Annual average limit values set by the EU, UK and Scottish Governments to regulate the concentrations of traffic-related air pollutants in ambient air.

Pollutant	EU ($\mu g/m^3$)	UK ($\mu g/m^3$)	Scotland ($\mu g/m^3$)
NO_2 (annual)	40	40	As UK
PM_{10} (annual)	40	40	18
$PM_{2.5}$ (annual)	25	25	12
O_3 (8 hour mean)	120 (not to be exceeded more than 25 days over 3 years)	100 (not to be exceeded more than 10 times in a year)	As UK

Epidemiology studies aim to relate measured pollution concentrations to effects on human health. The exposure assessment methods used in these epidemiology studies, and the techniques commonly used to estimate population exposure to air pollution, as discussed below.

1.2.1. Exposure assessments

Exposure assessment studies are used to determine the health effects identified previously through analysis of an individual or populations estimated exposure to air pollution compared with recorded health data, such as hospital admissions or mortality. Short-term studies identify health effects that occur during the period immediately following times with elevated pollution concentrations (hours, days or weeks), and these usually impact individuals with pre-existing health problems to a greater extent that the general population. Long-term studies (usually over a year) identify chronic health effects within the general population which are attributed to a build-up of exposures leading to illness.

Epidemiology studies utilising available data on the general population do not suffer from any population bias as no specific individuals are chosen to be studied, making the results very generic and representative. However, the large size of these groups means that detailed pollution assessments cannot be carried out for all individuals and concentrations are often based on general trends in atmospheric concentrations and model predicted pollution concentrations.

Cohort studies produce detailed exposure estimates of study participants and can be combined with medical records for example, to study health effects in detail. Three of the most important cohorts to date (Harvard Six Cities, ACS-II and ASHMOG) found relationships between mortality and air pollution for cohorts of between 7,000 and 500,000 people that were followed for between 7 and 15 years (Dockery et al., 1993; Pope et al.,

1995; Abbey et al., 1999). The selection of those within the cohort is crucial and must not be subjected to any bias which could undermine any results observed. The length of time of these is the main challenge in collecting and analysing large volumes of data, and if specific individuals are followed, this can be intrusive (Brunekreef, 2003).

The assignment of pollution concentrations to the study participants above can take two main forms: monitoring or modelling. The former provides detailed information on concentrations at the location at which the measurements were made, for example the home address of participants, however these can be time consuming and costly to set up and maintain. Modelling exposure estimates allows retrospective analysis of health effects and allows estimates for pollution concentrations at all locations in a study area to be made. The common methods used in air pollution monitoring and modelling, and the epidemiology studies in which these have been applied, are discussed below.

1.2.2. Monitoring population exposure

As previously stated, air pollution concentrations are measured real-time at automatic monitoring stations located at several locations over the UK to assess compliance with legislative values. These automatic stations are expensive to set up, run and maintain thus there are only around 100 currently operational in the UK. Historically, the concentrations measured at a single fixed location, such as these automatic stations, have been used in epidemiology studies to assess population exposure in the surrounding area, such as in the Harvard Six Cities, ACS-II and SAPALDIA epidemiology studies (Dockery et al., 1993; Pope et al., 1995; Zemp et al., 1999). In these studies, concentrations from a centrally-located automatic monitoring station within the study areas were used to assess mortality or respiratory illness due to air pollution. The use of a single station to assess exposure is limited as this does not take into account fine-scale spatial variations in pollution concentrations (Baxter et al., 2013). The ASHMOG study (Abbey et al., 1999) improved upon this approach by assigning participant exposures to the average concentration measured at several sites located within the study area instead of using concentrations from a single site only.

Passive samplers can be used to supplement fixed-site monitoring locations (indoor or outdoor) and, due to their low cost, can be deployed readily at a large number of locations, such as the residences of individuals in an exposure study. This allows more detailed information about an individual's exposure to air pollution to be ascertained and takes into account fine-scale variations in pollutant concentrations. They can also be used in personal monitoring, whereby the samplers are deployed on the participant and directly measures the

concentrations they are exposed to. As their name suggests, passive samplers work through diffusion of air and do not require any mains or battery power to operate, however their main disadvantage they only provide time-averaged concentrations for the period they are deployed, which is typically 1 week in duration (Palmes et al., 1976). Additionally the samplers are subject to uncertainties including interferences from chemicals and atmospheric conditions (Cape, 2009), which can lead to less accurate estimates of ambient pollution concentrations and hence population exposures. Concentrations of gaseous air pollutants, such as NO_x and O_3, have been measured using this technique. The use of passive samplers in epidemiology studies is limited to short-term studies with small numbers of participants due to the resource-intensive process associated with the deployment and consequent analysis of passive samplers, making their use for larger (and longer) cohort studies impractical (Brunekreef et al., 1990; Farrow et al., 1997). The use of passive samplers for personal monitoring is limited due to the time/cost restraints and the reliance on participants to engage with the study (including keeping samplers close to themselves and accurately recording time-activity diaries), however this has been successfully used to measure personal exposure of NO_2 (Linaker et al., 2000; Van Roosbroeck et al., 2007). The main use of passive samplers in epidemiology assessments is to obtain detailed spatial information about the concentrations within the study area, which are then used to develop modelled pollution concentrations (these are discussed in more detail below) (Beelen et al., 2013; Brauer et al., 2007, 2002).

The estimation of the personal exposure of a participant at their home address makes the assumption that the individual remains in the same place throughout the study, which can be incorrect such as in the case of working-age participants who can spend a large proportion of their time away from home, for example travelling (Dons et al., 2014). The passive samplers discussed above can be used to give time-averaged exposures to pollution however they do not allow detailed temporal concentrations to be estimated. Developments in technology have allowed the generation of low power, portable, real-time instruments which can be used to provide temporally-resolved personal concentration measurements. These active systems require a pump to pass air into the instrument for measurement, meaning batteries or mains power is required which is a limiting factor. The main advantages of these real time samplers are they are designed to be easily deployed or wearable and provide information on real time pollutant concentrations, similarly to the automatic analysers. Limitations of these real time instruments include their battery life, unknown or vague methods describing concentration derivation and the relatively unknown ability of these instruments to measure ‘true’ ambient concentrations (Lewis and Edwards, 2016; Lewis et al., 2016). Their small

size, however, permits their use in rapid identification of pollution variations within a study area (Gillespie et al., 2017; Wu et al., 2015), citizen-science projects (CITI-SENSE Project, 2015; Snik et al., 2014), and personal monitoring (Montagne et al., 2013; Steinle et al., 2015; Van Roosbroeck et al., 2007). Additionally, they can also be used to obtain measurements for use in air pollution modelling (Deville Cavellin et al., 2016; Dons et al., 2013).

In order to make accurate estimates of an individuals exposure using monitoring, the researcher needs to understand the limitations of each technique and consequently chose the method that suits their requirements. The more accurate the measurement data the less uncertainty in the health effect relationships identified in the epidemiology study (Zeger et al., 1999). The measurement data needs to be representative of the environment in which the individual being assessed is exposed to. For example, the static automatic monitors cannot capture fine-scale spatial information and could lead to a potential underestimation of an individuals exposure as these sites are often located away from direct sources of pollution. The measurement of pollution at multiple fixed locations, for example the use of diffusion tubes at an individuals home address, cannot account for fine scale changes a participant is exposed to during the day (Marshall et al., 2008). The real-time sensors have potential merits in that they can address the need for fine scale spatial resolution, however the accuracy of these instruments is a questionable factor (Snyder et al., 2013). These are likely to require frequent calibrations against reference instruments to ensure their accuracy, however this is a time consuming process and the regularity that the calibrations should be made at in order to maintain high monitor accuracy remains relatively unknown. Additionally, the accuracy of the monitors and the calibration remains relatively unknown when the portable monitors are removed from the reference stations used for calibration, which leads to an unquantifiable uncertainty in the measurements of an individual's exposure. Currently there is no standardised testing regime for these portable real-time monitors, meaning there are many sensors available with limited information available to compare these and to identify the accuracy and uncertainty of the sensors. This can mean individuals are using the monitors without making detailed calibration measurements (which relies on access to an automatic monitoring station which is not publically available) and using the result without the understanding of the confidence of the measurements made, leading to larger errors in the measurement data and consequently more uncertain relationships between exposures and health effects.

1.2.3. Modelling population exposure

Monitoring air pollution provides information about the true pollution concentrations in the study area; however, its cost and time consuming nature means that only a small number of sites can be monitored. Modelling is advantageous as it allows predictions of pollution concentrations to be made over a whole study area, meaning concentrations at specific locations (such as home addresses) can be rapidly obtained for a very large number of participants. However modelling requires computing power, which can often be very intense, complex understanding of fundamental concepts and can also be very demanding in the volume of data required to generate an accurate model. Additional limitations of modelling for use with epidemiology studies include the large volume of data required to generate the models, the model is limited by the quality of the input data, and these models can only estimate outdoor exposures, which is unlikely to be a true estimate of individuals exposure as, in the UK, research has found that an individual spends approximately 1-2 hours (Diffey, 2011).

The models are developed from monitoring within the study area, using fixed sites or a pre-designed network (e.g. using passive samplers or portable instruments to obtain concentration estimates), with the latter being preferable as a larger number of sites over a wider range of pollution environments can be used. The ESCAPE study used a dedicated monitoring campaign for 36 countries in Europe following the same site selection protocol to obtain pollution models, allowing comparisons between the pollution and health effects observed in each country (Beelen et al., 2013).

There are three main modelling processes which are used in population exposure assessments – interpolation of monitoring data using Geographical Information Systems (GIS); Land Use Regression (LUR) modelling; and Dispersion modelling. Interpolation techniques, such as inverse distance weighting, use pollution concentration from fixed monitoring stations and assign concentrations to the surrounding area based on weighted concentrations from close proximity monitoring stations. However, this relatively simple approach does not take into account other factors (such as terrain, or urban morphology) and can become less reliable in areas with limited monitoring data, and its use in epidemiology studies has been in decline since the start of the 21st century (Jerrett et al., 2004). The more recent epidemiology studies have typically used LUR or dispersion modelling, which have become more readily available with improvements in computation power, to model pollution concentrations in the study area. Both these methods use monitoring data as inputs, and use either statistical (LUR) or empirical (dispersion) mathematical technique to allow a more complex and accurate pollution model to be derived. When compared to each other, the

LUR and dispersion models have generally been found to perform similarly (Beelen et al., 2010; de Hoogh et al., 2014; Gulliver et al., 2011; Marshall et al., 2008), with some studies combing the two approaches in order to obtain fine scale, area-specific spatial variation from the LUR in combination with the high temporal resolution provided by the dispersion model (Mölter et al., 2010).

Land use regression modelling was introduced in the SAVIAH project in 1997 (Briggs et al., 1997), and since this time has been adopted in epidemiology studies, including the ESCAPE project (Aguilera et al., 2008; Brauer et al., 2002; Robinson et al., 2015; Wang et al., 2013). Using monitoring sites, the relationship between pollutant concentrations and the surrounding environment are determined, which can then be applied to estimate pollution concentrations elsewhere. Environmental factors often included in LUR models include traffic sources (for example proximity to roads), building heights and land-use type. This is a relatively low cost process, however requires access to large volumes of data (GIS and monitoring) and has limited transferability between study areas, meaning new models should be developed for each specific area (Allen et al., 2011; Poplawski et al., 2008). The combination of LUR and fixed-site real time monitoring has recently demonstrated the potential to include temporal resolution in LUR models (Cordioli et al., 2017).

Dispersion models can be relatively complex as they take into account a wider range of environmental factors (including monitored pollution, emissions and meteorology) when estimating pollution concentrations. The greater volumes of data required introduces uncertainty in the model when, for example assumptions about emissions have to be made when data is unavailable, or when other models are used to provide input to the dispersion model. However, the complexity of this model allows highly resolved spatial and temporal pollution estimates to be determined with less monitoring sites required than LUR modelling. Dispersion models, therefore, have potential for use in epidemiology estimates and can provide models for historic studies (Bellander et al., 2001; Gulliver and Briggs, 2011; Nyberg et al., 2000).

1.3. Thesis Aims and layout

The aims of this thesis are to assess the evaluation and deployment of air pollution measurement systems for the assessment of human exposure to traffic-related air pollution. This PhD develops the literature evaluating the ability of portable sensor technologies and passive samplers to assess spatial and temporal variations in fine-scale pollution concentrations (through calibration, static and mobile monitoring), and evaluates the use of rapid GIS-dispersion models in population exposure assessment.

This thesis is comprised of five self-contained chapters written in the style of manuscripts, and some of these manuscripts have been submitted to journals. A brief summary paragraph is provided at the start of each of the chapters to explain the context of the work, the contribution of any co-authors to the work and the status of any submitted manuscripts at the time of the submission of this PhD thesis. The objectives of the research presented in each of these chapters are to evaluate:

- common passive diffusion samplers used to measure NO_2 and an investigation of the bias factors associated with monitoring of NO_2 concentrations using these samplers.
- performance of portable real-time monitors measuring NO_2, O_3 and BC at an automatic monitoring station and determine the optimal calibration methods for these monitors.
- ability of short-duration monitoring using the portable monitors evaluated above to predict spatial and longer-term temporal trends of NO_2 using peripatetic monitoring.
- impact of user choices during weather modelling on the pollution estimates from a commercial dispersion modelling.
- newly-developed kernel dispersion model (based on open-source software) and investigation into the ability of surrogate variables to estimate the location of, and correct for the influence of, street canyons.

1.4. References

Abbey, D.E., Nishino, N., McDonnell, W.F., Burchette, R.J., Knutsen, S.F., Lawrence Beeson, W., Yang, J.X., 1999. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. Am. J. Respir. Crit. Care Med. 159, 373–382. doi:10.1164/ajrccm.159.2.9806020

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

Allen, R.W., Amram, O., Wheeler, A.J., Brauer, M., 2011. The transferability of NO and NO2 land use regression models between cities and pollutants. Atmos. Environ. 45, 369–378. doi:10.1016/j.atmosenv.2010.10.002

Baxter, L.K., Dionisio, K.L., Burke, J., Ebelt Sarnat, S., Sarnat, J.A., Hodas, N., Rich, D.Q., Turpin, B.J., Jones, R.R., Mannshardt, E., Kumar, N., Beevers, S.D., Ozkaynak, H., 2013. Exposure prediction approaches used in air pollution epidemiology studies: Key findings and future recommendations. J. Expo. Sci. Environ. Epidemiol. 23, 654–659. doi:10.1038/jes.2013.62

Bayer-Oglesby, L., Grize, L., Gassner, M., Takken-Sahli, K., Sennhauser, F.H., Neu, U., Schindler, C., Braun-Fahrlander, C., 2005. Decline of ambient air pollution levels and improved respiratory health in Swiss children. Environ. Health Perspect. 113, 1632–1637.

Bayer-Oglesby, L., Schindler, C., Hazenkamp-von Arx, M.E., Braun-Fahrlander, C., Keidel, D., Rapp, R., Kunzli, N., Braendli, O., Burdet, L., Sally Liu, L.-J., Leuenberger, P., Ackermann-Liebrich, U., the SAPALDIA Team, 2006. Living near Main Streets and Respiratory Symptoms in Adults: The Swiss Cohort Study on Air Pollution and Lung Diseases in Adults. Am. J. Epidemiol. 164, 1190–1198.

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., Voogt, M., Duyzer, J., Zandveld, P., Hoek, G., 2010. Comparison of the performances of land use regression modelling and dispersion modelling in estimating small-scale variations in long-term air pollution concentrations in a Dutch urban area. Atmos. Environ. 44, 4614–4621. doi:10.1016/j.atmosenv.2010.08.005

Beevers, S.D., Westmoreland, E., de Jong, M.C., Williams, M.L., Carslaw, D.C., 2012. Trends in NOx and NO2 emissions from road traffic in Great Britain. Atmos. Environ. 54, 107–116. doi:10.1016/j.atmosenv.2012.02.028

Bell, M., McDermott, A., Zeger, S., Samet, J., Dominici, F., 2004. Ozone and Short-term Mortality in 95 US Urban Communities, 1987-2000. JAMA 292, 2372–2378. doi:10.1001/jama.292.19.2372

Bell, M.L., Davis, D.L., Fletcher, T., 2004. A retrospective assessment of mortality from the London smog episode of 1952: the role of influenza and pollution. Environ. Health Perspect. 112, 6–8.

Bellander, T., Berglind, N., Gustavsson, P., Jonson, T., Nyberg, F., Pershagen, G., Järup, L., 2001. Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. Environ. Health Perspect. 109, 633–639.

Beverland, I.J., Carder, M., Cohen, G.R., Heal, M.R., Agius, R.M., 2012a. Associations between short/medium-term variations in black smoke air pollution and mortality in the Glasgow conurbation, UK. Environ. Int. 50, 1–6. doi:10.1016/j.envint.2012.08.012

Beverland, I.J., Cohen, G.R., Heal, M.R., Carder, M., Yap, C., Robertson, C., Hart, C.L., Agius, R.M., 2012b. A Comparison of Short-term and Long-term Air Pollution Exposure Associations with Mortality in Two Cohorts in Scotland. Environ. Health Perspect. 120, 1280–1285. doi:10.1289/ehp.1104509

Bigi, A., Harrison, R.M., 2010. Analysis of the air pollution climate at a central urban background site. Atmos. Environ. 44, 2004–2012.

Birmili, W., Tomsche, L., Sonntag, A., Opelt, C., Weinhold, K., Nordmann, S., Schmidt, W., 2013. Variability of aerosol particles in the urban atmosphere of Dresden (Germany): effects of spatial scale and particle size. Meteorol. Z. 22, 195–211. doi:10.1127/0941-2948/2013/0395

Brand, C., 2016. Beyond "Dieselgate": Implications of unaccounted and future air pollutant emissions and energy use for cars in the United Kingdom. Energy Policy 97, 1–12. doi:10.1016/j.enpol.2016.06.036

Brauer, M., Hoek, G., Smit, H.A., Jongste, J.C. de, Gerritsen, J., Postma, D.S., Kerkhof, M., Brunekreef, B., 2007. Air pollution and development of asthma, allergy and infections in a birth cohort. Eur. Respir. J. 29, 879–888. doi:10.1183/09031936.00083406

Brauer, M., Hoek, G., Van Vliet, P., Meliefste, K., Fischer, P.H., Wijga, A., Koopman, L.P., Neijens, H.J., Gerritsen, J., Kerkhof, M., Heinrich, J., Bellander, T., Brunekreef, B., 2002. Air pollution from traffic and the development of respiratory infections and asthmatic and allergic symptoms in children. Am. J. Respir. Crit. Care Med. 166, 1092–1098.

Briggs, D., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., Pryl, K., Van Reeuwijk, H., Smallbone, K., Van Der Veen, A., 1997. Mapping urban air pollution using GIS: a regression-based approach. Int. J. Geogr. Inf. Sci. 11, 699–718.

Brook, R.D., Rajagopalan, S., Pope, C.A., III, Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith,

S.C., Jr., Whitsel, L., Kaufman, J.D., Amer Heart Assoc Council, E., Council Kidney Cardiovasc, D., Council Nutr Phys Activity, M., 2010. Particulate Matter Air Pollution and Cardiovascular Disease An Update to the Scientific Statement From the American Heart Association. Circulation 121, 2331–2378. doi:10.1161/CIR.0b013e3181dbece1

Brunekreef, B., 2003. Design of cohort studies for air pollution health effects. J. Toxicol. Environ. Health. 66, 16-19. doi:10.1080/15287390306420

Brunekreef, B., Houthuijs, D., Dijkstra, L., Boleij, J.S.M., 1990. Indoor Nitrogen Dioxide Exposure and Children's Pulmonary Function. J. Air Waste Manag. Assoc. 40, 1252–1256. doi:10.1080/10473289.1990.10466779

Canepari, S., Padella, F., Astolfi, M.L., Marconi, E., Perrino,C., 2013. Elemental concentration in atmospheric particulate matter: estimation of nanoparticle contribution. Aerosol Air Qual. Res. 13, 1619–1629. doi:10.4209/aaqr.2013.03.0081

Cape, J.N., 2009. The Use of Passive Diffusion Tubes for Measuring Concentrations of Nitrogen Dioxide in Air. Crit. Rev. Anal. Chem. 39, 289–310. doi:10.1080/10408340903001375

Carrington, D., 2016. High court rules UK government plans to tackle air pollution are illegal. The Guardian.

CITI-SENSE Project, 2015. Home Page.

Constabile, F., Birmili, W., Klose, S., Tuch, T., Wehner, B., Wiedensohler, A., Franck, U., Konig, K., Sonntag, A., 2009. Spatio-temporal variability and principle components of the particle number size distribution in an urban atmosphere. Atmos. Chem. Phys. 9, 3163–3195. doi:10.5194/acp-9-3163-2009

Cordioli, M., Pironi, C., De Munari, E., Marmiroli, N., Lauriola, P., Ranzi, A., 2017. Combining land use regression models and fixed site monitoring to reconstruct spatiotemporal variability of NO2 concentrations over a wide geographical area. Sci. Total Environ. 574, 1075–1084. doi:10.1016/j.scitotenv.2016.09.089

de Hoogh, K., Korek, M., Vienneau, D., Keuken, M., Kukkonen, J., Nieuwenhuijsen, M.J., Badaloni, C., Beelen, R., Bolignano, A., Cesaroni, G., Pradas, M.C., Cyrys, J., Douros, J., Eeftens, M., Forastiere, F., Forsberg, B., Fuks, K., Gehring, U., Gryparis, A., Gulliver, J., Hansell, A.L., Hoffmann, B., Johansson, C., Jonkers, S., Kangas, L., Katsouyanni, K., Künzli, N., Lanki, T., Memmesheimer, M., Moussiopoulos, N., Modig, L., Pershagen, G., Probst-Hensch, N., Schindler, C., Schikowski, T., Sugiri, D., Teixidó, O., Tsai, M.-Y., Yli-Tuomi, T., Brunekreef, B., Hoek, G., Bellander, T., 2014. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. Environ. Int. 73, 382–392. doi:10.1016/j.envint.2014.08.011

DEFRA, 2007. The Air Quality Strategy for England, Scotland, Wales and Northern Ireland (Volume 1) (No. Cm 7169 NIA 61/06-07).

Degraeuwe, B., Weiss, M., 2017. Does the New European Driving Cycle (NEDC) really fail to capture the NOX emissions of diesel cars in Europe? Environ. Pollut. 222, 234–241. doi:10.1016/j.envpol.2016.12.050

Department for Transport, 2015. Transport Statistics Great Britain: 2015 (No. TSGB03).

Deville Cavellin, L., Weichenthal, S., Tack, R., Ragettli, M.S., Smargiassi, A., Hatzopoulou, M., 2016. Investigating the Use Of Portable Air Pollution Sensors to Capture the Spatial Variability Of Traffic-Related Air Pollution. Environ. Sci. Technol. 50, 313–320. doi:10.1021/acs.est.5b04235

Diffy, B.L., 2011. An overview analysis of the time people spend outdoors. British. J. Dermatology. 164(4), 848-854. doi:10.1111/j.1365-2133.2010.10165x

Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G.J., Speizer, F.E., 1993. An Association between Air Pollution and Mortality in Six U.S. Cities. N. Engl. J. Med. 329, 1753–1759. doi:10.1056/NEJM199312093292401

Dockery, D.W., Speizer, F.E., Stram, D.O., Ware, J.H., Spengler, J.D., Ferris, B.G., 1989. Effects of inhalable particles on respiratory health of children. Am. Rev. Respir. Dis. 139, 587–594. doi:10.1164/ajrccm/139.3.587

Doherty, R.M., Heal, M.R., Wilkinson, P., Pattenden, S., Vieno, M., Armstrong, B., Atkinson, R., Chalabi, Z., Kovats, S., Milojevic, A., Stevenson, D.S., 2009. Current and future climate- and air pollution-mediated impacts on human health. Environ. Health 8. doi:10.1186/1476-069x-8-s1-s8

Dons, E., Kochan, B., Bellemans, T., Wets, G., Panis, L.I., 2014. Modeling Personal Exposure to Air Pollution with AB2C: Environmental Inequality. Procedia Comput. Sci., The 5th International Conference on Ambient Systems, Networks and Technologies (ANT-2014), the 4th International Conference on Sustainable Energy Information Technology (SEIT-2014) 32, 269–276. doi:10.1016/j.procs.2014.05.424

Dons, E., Van Poppel, M., Kochan, B., Wets, G., Int Panis, L., 2013. Modeling temporal and spatial variability of traffic-related air pollution: Hourly land use regression models for black carbon. Atmos. Environ. 74, 237–246.

European Environment Agency, 2016. EMEP/EEA air pollutant emission inventory guidebook (Publication No. 21/2016).

Fann, N., Lamson, A.D., Anenberg, S.C., Wesson, K., Risley, D., Hubbell, B.J., 2012. Estimating the National Public Health Burden Associated with Exposure to Ambient PM2.5 and Ozone. Risk Anal. Int. J. 32, 81–95.

Farrow, A., Greenwood, R., Preece, S., Golding, J., 1997. Nitrogen Dioxide, the Oxides of Nitrogen, and Infants' Health Symptoms. Arch. Environ. Health Int. J. 52, 189–194. doi:10.1080/00039899709602885

Franco, V., Sanchez, F., German, J., Mock, P., 2014. Real-world exhaust emissions from modern diesel cars: A meta-analysis of PEMS emissions data from EU (Euro 6) and US (Tier 2 Bin5/ULEV II) diesel passenger cars. Internation Council on Clean Transportation.

Gent, J.F., Triche, E.W., Holford, T.R., Belanger, K., Bracken, M.B., Beckett, W.S., Leaderer, B.P., 2003. Association of low-level ozone and fine particles with respiratory symptoms in children with asthma. JAMA 290, 1859–1867. doi:10.1001/jama.290.14.1859

Gillespie, J., Masey, N., Heal, M.R., Hamilton, S., Beverland, I.J., 2017. Estimation of spatial patterns of urban air pollution over a 4-week period from repeated 5-min measurements. Atmos. Environ. 150, 295–302. doi:10.1016/j.atmosenv.2016.11.035

Guerreiro, C.B.B., Foltescu, V., de Leeuw, F., 2014. Air quality status and trends in Europe. Atmos. Environ. 98, 376–384. doi:10.1016/j.atmosenv.2014.09.017

Gulliver, J., Briggs, D., 2011. STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. Sci. Total Environ. 409, 2419–2429.

Gulliver, J., de Hoogh, K., Fecht, D., Vienneau, D., Briggs, D., 2011. Comparative assessment of GIS-based methods and metrics for estimating long-term exposures to air pollution. Atmos. Environ. 45, 7072–7080.

Harrison, R.M., Shi, J.P., Xi, S., Khan, A., Mark, D., Kinnersley, R., Yin, J., 2000. Measurement of number, mass and size distributions of particles in the atmosphere. Philos. Trans. R. Soc. Lond. A. 358, 2567–2580. doi: 10.1098/rsta.2000.0669

HEI (Health Effects Institute), 2013. Understanding the health effects of ambient ultrafine particles. Perspectives 3.

Heusinkveld, H.J., Wahle, T., Campbell, A., Westerink, R.H., Tran, L., Johnston, H., Stone, V., Cassee, F.R., Schins, R.P., 2016. Neurodegenerative and neurological disorders by small inhaled particles. Neurotoxicology. 19, 94–106. doi:10.1016/j.neuro.2016.07.007

Hoek, G., Boogaard, H., Knol, A., de Hartog, J., Slottje, P., Ayres, J.G., Borm, P., Brunekreef, B., Donaldson, K., Forastiere, F., Holgate, S., Kreyling, W.G., Nemery, B., Pekkanen, J., Stone, V., Wichmann, H.E., van der Sluijs, J., 2010. Concentration response functions for ultrafine particles and all-cause mortality and hospital admissions: results of a European expert panel elicitation. Environ. Sci. Technol. 44, 476–482. doi:10.1021/es9021393

Hoek, G., Meliefste, K., Cryrs, J., Lewne, M., Bellander, T., Brauer, M., Fischer, P., Gehring, U., Heinrich, J., van Vliet, P., Brunekreef, B., 2002. Spatial variability of fine

particles concentrations in three European areas. Atmos. Environ. 35(25), 4077–4088. doi:10.1016/S1352-2310(02)00297-2

Janssen, N.A.H, Hoek, G., Simic-Lawson, M., Fischer, P., van Bree, L., ten Brink, H., Keuken, M., Atkinson, R.W., Anderson, H.R., Brunekreef, B., Cassee, F.R., 2011. Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM10 and PM2.5. Environ. Health. Perspect. 119(12), 1691-1699. doi:10.1289/ehp.1003369

Janssen, N.A.H, Meliefste, K., Fuchs, O., Wieland, S.K., Cassee, F., Brunekreef, B., Sandstom, T., 2008. High and low volume sampling of particulate matter at sites with different traffic profiles in the Netherlands and Germany: Results from the HEPMEAP study. Atmos. Environ. 42(6), 1110–1120. doi:10.1016/j.atmosenv.2007.10.085

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., Giovis, C., 2004. A review and evaluation of intraurban air pollution exposure models. J. Expo. Sci. Environ. Epidemiol. 15, 185–204. doi:10.1038/sj.jea.7500388

Kumar, P., Morawska, L., Brimili, W., Paasonen, P., Hu, M., Kulmala, M., Harrison, R.M., Norford, L., Britter, R., 2014. Ultrafine particles in cities. Environment International. 66, 1–10. doi: 10.1016/j.envint.2014.01.013

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

Kumar, P., Fennell, P., Hayhurst, A., Britter, R., 2009. Street versus rooftop level concentrations of fine particles in a Cambridge street canyons. Bound-Layer Meteorol. 131, 3–18. doi: 10.1007/s10546-008-9300-3

Lee, D., Ferguson, C., Mitchell, R., 2009. Air pollution and health in Scotland: a multicity study. Biostatistics 10, 409–423.

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

Lewis, A.C., Lee, J.D., Edwards, P.M., Shaw, M.D., Evans, M.J., Moller, S.J., Smith, K.R., Buckley, J.W., Ellis, M., Gillot, S.R., White, A., 2016. Evaluating the performance of low cost chemical sensors for air pollution research. Faraday Discuss. doi:10.1039/C5FD00201J

Li, N., Sioutas, C., Cho, A., Schimitz, D., Mistra, C., Sempf, J., Wang, M., Oberley, T., Friones, J., Nel, A., 2003. Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. Environ. Health. Perspect. 11, 455–460

Linaker, C.H., Chauhan, A.J., Inskip, H.M., Holgate, S.T., Coggon, D., 2000. Personal exposures of children to nitrogen dioxide relative to concentrations in outdoor air. Occup. Environ. Med. 57, 472–476. doi:10.1136/oem.57.7.472

Lipsett, M., Hurley, S., Ostro, B., 1997. Air pollution and emergency room visits for asthma in Santa Clara County, California. Environ. Health Perspect. 105, 216–222.

Lodovici, M., Bigagali, E., 2011. Oxidative stress and air pollution exposure. Journal of Toxicology. 2011. doi: 10.1155/2011/487074

Manigrasso, M., Avino, P., 2012. Fast evolution of urban ultrafine particles: implications for deposition in the human respiratory system. Atmos. Environ. 51, 116–123. doi:10.1016/j.atmosenv.2012.01.039

Marshall, J.D., Nethery, E., Brauer, M., 2008. Within-urban variability in ambient air pollution: Comparison of estimation methods. Atmos. Environ. 42, 1359–1369. doi:10.1016/j.atmosenv.2007.08.012

McConnell, R., Berhane, K., Gilliland, F., London, S.J., Islam, T., Gauderman, W.J., Avol, E., Margolis, H.G., Peters, J.M., 2002. Asthma in exercising children exposed to ozone: a cohort study. Lancet Lond. Engl. 359, 386–391. doi:10.1016/S0140-6736(02)07597-9

Mills, I.C., Atkinson, R.W., Kang, S., Walton, H., Anderson, H.R., 2015. Quantitative systematic review of the associations between short-term exposure to nitrogen dioxide and mortality and hospital admissions. BMJ Open 5, e006946. doi:10.1136/bmjopen-2014-006946

Mölter, A., Lindley, S., de Vocht, F., Simpson, A., Agius, R., 2010. Modelling air pollution for epidemiologic research — Part I: A novel approach combining land use regression and air dispersion. Sci. Total Environ., Special Section: Integrating Water and Agricultural Management Under Climate Change 408, 5862–5869. doi:10.1016/j.scitotenv.2010.08.027

Montagne, D., Hoek, G., Nieuwenhuijsen, M., Lanki, T., Pennanen, A., Portella, M., Meliefste, K., Eeftens, M., Yli-Tuomi, T., Cirach, M., Brunekreef, B., 2013. Agreement of Land Use Regression Models with Personal Exposure Measurements of Particulate Matter and Nitrogen Oxides Air Pollution. Environ. Sci. Technol. 130712144458004. doi:10.1021/es400920a

Nyberg, F., Gustavsson, P., Järup, L., Bellander, T., Berglind, N., Jakobsson, R., Pershagen, G., 2000. Urban air pollution and lung cancer in Stockholm. Epidemiol. Camb. Mass 11, 487–495.

Oberdörster, G., Oberdörster, E., Oberdörster, J., 2005. Nanotoxicology: an emerging discipline evolving from studies of ultrafine particles. Environ. Health Perspect. 113, 823–839.

Oberdörster, G., Sharp, Z., Atudorei, V., Elder, A., Gelein, R., Kreyling, W., Cox, C., 2004. Translocation of inhaled ultrafine particle to the brain. Inhal. Toxicol. 16, 437–445. doi:10/1080/08958370490439597

Oberdörster, G., 2000. Toxicology of ultrafine particles: in vivo studies. Philosoph. Trans. Royal Soc. A. 358, 2719–2740.

Pagano, P., De Zaiacomo, T., Scarcella,E., Bruni, S., Calamosca, M., 1996. Mutagenic activity of total and particle-sized fractions of urban particulate matter. Environ. Sci. Technol. 30, 3512–3516.

Palmes, E.D., Gunnison, A.F., Dimattio, J., Tomczyk, C., 1976. Personal sampler for nitrogen dioxide. Am. Ind. Hyg. Assoc. J. 37, 570–577.

Petroeschevsky, A., Simpson, R.W., Thalib, L., Rutherford, S., 2001. Associations between outdoor air pollution and hospital admissions in Brisbane, Australia. Arch. Environ. Health 56, 37–52. doi:10.1080/00039890109604053

Pope, C.A., Dockery, D.W., 2006. Health effects of fine particulate air pollution: lines that connect. J. Air Waste Manag. Assoc. 1995 56, 709–742.

Pope, C.A., Thun, M.J., Namboodiri, M.M., Dockery, D.W., Evans, J.S., Speizer, F.E., Heath, C.W., 1995. Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of U.S. Adults. Am. J. Respir. Crit. Care Med. 151, 669–674. doi:10.1164/ajrccm/151.3_Pt_1.669

Poplawski, K., Gould, T., Setton, E., Allen, R., Su, J., Larson, T., Henderson, S., Brauer, M., Hystad, P., Lightowlers, C., Keller, P., Cohen, M., Silva, C., Buzzelli, M., 2008. Intercity transferability of land use regression models for estimating ambient concentrations of nitrogen dioxide. J. Expo. Sci. Environ. Epidemiol. 19, 107–117. doi:10.1038/jes.2008.15

Raaschou-Nielsen, O., Beelen, R., Wang, M., Hoek, G., Andersen, Z.J., Hoffman, B., Stafoggia, M., Samoli, E., Weinmayr, G., Dimankopoulou, K., Nieuwenhuijsen, M., Xun, W.W., Fischer, P., Eriksen, K.T., Sorensen, M., Tjonneland, A., Ricceri, F., de Hoogh, K., Key, T., Eeftens, M., Peeters, P.H., Bueno-de-Mesquite, H.B., Meliefste, K., Oftedal, B., Schwarze, P.E., Nafstad, P., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Penell, J., De Faire, U., Korek, M., Pedersen, N., Ostenson, C.G., Pershagen, G., Fratiglioni, L., Concin, H., Nagel, G., Jaensch, A., Ineichen, A, Naccarati, A., Katsoulis, M., Trichpoulou, A., Keuken, M., Jedynska, A., Kooter, I.M., Kokkonen, J., Brunekreef, B., Sokhi, R.S., Katsouyanni, K., Vineis, P., 2016. Particulate matter air pollution components and risk for lung cancer. Environment International. 87, 66–73. doi: 10.1016/j.envint.2-15.11.007

Robinson, O., Basagaña, X., Agier, L., de Castro, M., Hernandez-Ferrer, C., Gonzalez, J.R., Grimalt, J.O., Nieuwenhuijsen, M., Sunyer, J., Slama, R., Vrijheid, M., 2015. The Pregnancy Exposome: Multiple Environmental Exposures in the INMA-Sabadell Birth Cohort. Environ. Sci. Technol. 49, 10632–10641. doi:10.1021/acs.est.5b01782

Sheppard, L., Burnett, R.T., Szpiro, A.A., Kim, S., Jerrett, M., Pope, C.A., Brunekreef, B., 2012. Confounding and exposure measurement error in air pollution epidemiology. Air. Qual. Atmos. Health. 5(2), 203–206. doi:10.1007/s11869-011-0140-9

Snik, F., Rietjens, J.H.H., Apituley, A., Volten, H., Mijling, B., Di Noia, A., Heikamp, S., Heinsbroek, R.C., Hasekamp, O.P., Smit, J.M., Vonk, J., Stam, D.M., van Harten, G., de Boer, J., Keller, C.U., 3187 iSPEX citizen scientists, 2014. Mapping atmospheric aerosols with a citizen science network of smartphone spectropolarimeters. Geophys. Res. Lett. 41, 2014GL061462. doi:10.1002/2014GL061462

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

Trasande, L., Thurston, G.D., 2005. The role of air pollution in asthma and other pediatric morbidities. J. Allergy Clin. Immunol. 115, 689–699. doi:10.1016/j.jaci.2005.01.056

Valavanidis, A., Fiotakis, K., Vlachogianni, T., 2008. Airborne Particulate Matter and Human Health: Toxicological Assessment and Importance of Size and Composition of Particles for Oxidative Damage and Carcinogenic Mechanisms. J. Environ. Sci. Health Part C 26, 339–362. doi:10.1080/10590500802494538

Van Roosbroeck, S., Jacobs, J., Janssen, N.A.H., Oldenwening, M., Hoek, G., Brunekreef, B., 2007. Long-term personal exposure to PM2.5, soot and NOx in children attending schools located near busy roads, a validation study. Atmos. Environ. 41, 3381–3394. doi:10.1016/j.atmosenv.2006.12.023

Wang, M., Beelen, R., Basagaña, X., Becker, T., Cesaroni, G., de Hoogh, K., Dėdelė, A., Declercq, C., Dimakopoulou, K., Eeftens, M., Forastiere, F., Galassi, C., Gražulevičienė, R., Hoffmann, B., Heinrich, J., Iakovides, M., Künzli, N., Korek, M., Lindley, S., Mölter, A., Mosler, G., Madsen, C., Nieuwenhuijsen, M., Phuleria, H., Pedeli, X., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Stempfelet, M., Tsai, M.-Y., Lanki, T., Udvardy, O., Varró, M.J., Wolf, K., Weinmayr, G., Yli-Tuomi, T., Hoek, G., Brunekreef, B., 2013.

Evaluation of Land Use Regression Models for NO2 and Particulate Matter in 20 European Study Areas: The ESCAPE Project. Environ. Sci. Technol. 47, 4357–4364. doi:10.1021/es305129t

Williams, M.L., Carslaw, D.C., 2011. New Directions: Science and policy - Out of step on NOx and NO2? Atmos. Environ. 45, 3911–3912. doi:10.1016/j.atmosenv.2011.04.067

World Health Organization, 2016. Ambient air pollution: A global assessment of exposure and burden of disease. Geneva, Switzerland.

Wiseman, C.L.S., Zereini, F., 2009. Airborne particulate matter, platinum group elements and human health. Sci. Tot. Environ. 407(8), 2493–2500. doi:10.1016/j.scitotenv.2008.12.057

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

Zeger, S.L., Thomas, D., Dominici, F., Samet, J.M., Schwartz, J., Dockery, D., Cohen, A., 1999. Exposure measurement error in time-series studies of air pollution: concepts and consequences. Environ. Health. Perspect. 108(5), 419-426. doi: 10.2307/3454382

Zemp, E., Elsasser, S., Schindler, C., Künzli, N., Perruchoud, A.P., Domenighetti, G., Medici, T., Ackermann-Liebrich, U., Leuenberger, P., Monn, C., Bolognini, G., Bongard, J.-P., Brändli, O., Karrer, W., Keller, R., Schöni, M.H., Tschopp, J.-M., Villiger, B., Zellweger, J.-P., 1999. Long-Term Ambient Air Pollution and Respiratory Symptoms in Adults (SAPALDIA Study). Am. J. Respir. Crit. Care Med. 159, 1257–1266. doi:10.1164/ajrccm.159.4.9807052

2. Description of Measurement Systems

This section describes the principles of the portable measurement systems used in this PhD research. Further details regarding the operational settings used for the real-time instruments are discussed in the individual results chapters.

2.1. Aeroqual S500 NO_2 and O_3 monitors

The Aeroqual S500 monitor is a portable real-time monitor that combines a 'standard' monitor body (containing battery and data storage) with interchangeable sensor heads measuring different gases, allowing the user to create a monitor that meets their required specifications (Aeroqual, 2017). Ambient air is drawn into the sensor head using a fan to ensure constant flow rates, and the air is passed over the sensor and concentrations of the pollutant to be measured are recorded. The technology used in the sensors is dependent on the sensor and the gas to be measured.

The NO_2 sensor used in this research was an electrochemical sensor (ENW2, 0 – 1 ppm range). When the sensor is exposed to NO_2 it is oxidised, leading to reductions in the concentrations of electrons and consequently increases in the current measured by the sensor. The current measured is proportional to the concentrations of NO_2 in the ambient air.

The O_3 sensor used was a gas-sensitive semiconductor sensor (OZU, range 0 – 0.15 ppm). Ozone in the sampled air absorbs onto the surface of the sensor, resulting in a rapid change in the measured conductivity and increased resistance. The change in resistance is proportional to the concentrations of O_3 in the ambient air (Williams et al., 2013).

2.2. microAeth AE51 black carbon monitor

The microAeth AE51 is a portable, real-time monitor that uses aethalometer technology to measure BC concentrations. Black carbon concentrations of between 0 and 1 mg/m^3 (± 0.1 µg) can be measured with the AE51, with a measurement resolution of 0.001 µg (Aethlabs, 2016).

Air is drawn into the instrument and passed through a T60 Teflon-coated borosilicate glass fiber filter, where particles deposit. A 880 nm LED light sources is passed through the centre of the deposition spot and the absorption of light in the spot is measured relative to an adjacent clean spot on the filter (Aethlabs, 2016). The increase in the absorbance of light (the Attenuation, ATN) in the deposition spot as more particles accumulate is used to calculate a concentration of black carbon (Virkkula et al., 2007; Hansen, 1982):

$$BC = \frac{1}{\propto_{abs}} \frac{A}{Q} \frac{\Delta ATN}{\Delta t}$$

Where *BC* is the concentration (µg/m^3), α_{abs} is the mass absorption cross section of BC (m^2/g), *A* is the area of the deposition spot (cm^2), *Q* is the flow of air drawn through the filter (lpm), ΔATN is the change in attenuation with time (Δt).

2.3. TSI CPC3007 particle number count monitor

Particle number count was measured using the TSI Condensation Particle Counter (CPC) 3007. The concentration range is 0 – 100000 particles / cm^3 (± 20 %) with a minimum detectable particle size of 10 nm (TSI, 2012).

The CPC draws ambient air into the instrument where it is passed over through a heated chamber, with walls soaked in isopropyl alcohol. The heat causes the alcohol to evaporate, and then both the ambient air and evaporated alcohol pass to a cooled condensation chamber. The fine particles suspended in the sampled air act as nuclei for the condensation of the alcohol droplets, thus increasing the particle size. These larger particles are then easily counted using an optical detector. The CPC provides a count of the number of particles in the air, but does not provide any information about the original size of the counted particles (TSI, 2012).

2.4. Passive sampling devices

Passive sampling devices have been used to measure NO_2 concentrations in ambient air since the 1970's (Palmes et al., 1976). The measurement of gases is passive, i.e. does not require any active systems to sample the air. The rate of diffusion into the tubes is controlled by Fick's Law:

$$F = -D \frac{\pi d^2}{4l}$$

Where *F* is the sampling rate (m^3/s), *D* is the diffusion coefficient of NO_2 in air (cm^2/s), *d* is the diameter of the sampling device, and *l* is the length (both in cm). This theoretical sampling rate has a number of assumptions, including no chemical reactions occurring within the sampler, no absorption of the gas by the sampler body and that the collection material collects 100 % of the gas (Cape, 2009).

The measurement of NO_2 using the passive samplers has commonly used triethanolamine (TEA) as the collection medium. TEA reacts with NO_2 and, with exception of at very high concentrations of ambient NO_2, gives 100 % capture of the gas (Cape, 2009).

The initial design of the passive sampler was a tube, which was successfully used to measure average NO_2 concentrations over a period of weeks. These hollow tubes were approximately 7 cm long, capped at one end which contained grids coated with TEA and left open at the other end. Ambient air diffuses up the length of tubes where NO_2 in the air is captured by the coated grids located in the caps. The most commonly used diffusion tube is the Palmes tube (Palmes et al., 1976) and these tubes are used in the UK to provide measurements to demonstrate compliance with the National Air Quality Objectives (DEFRA, 2016).

Recent developments to passive sampler designs has aimed to reduce the length of sampling time required through shortening of the diffusion tube path, increasing the surface area of sampling and hence increasing the uptake rate of gas by the sampler. Badge-type samplers, including the Ogawa sampler used in this research (www.ogawausa.com), have shorter path lengths of 1-2 cm, therefore faster sampling rates and thus can be used to measure exposures over shorter periods of time such as 24-hours (Sather et al., 2006).

Developments to passive sampler design have change the sampling design from axial sampling (as in the tube and badge samplers) to radial sampling. Radial samplers, such as the Radiello sampler (Cocheo et al., 1996), have the largest diffusion surface of the types of passive sampler designs as the collection medium is stored in a cartridge in the centre of the sampler, allowing diffusion to occur through a larger surface area.

2.4.1. Laboratory preparation and analysis of passive sampling devices

After exposure the diffusion samplers are analysed in a laboratory to determine the exposure-average concentration of NO_2 measured by the samplers. The most commonly used analysis procedure is based on colorimetry whereby the intensity of a coloured dye is used to determine the ambient concentrations of NO_2 measured by the samplers. Briefly, the trapped gaseous NO_2 molecules on the collection medium are extracted into water as nitrite ions. Acidified sulphanilamide and NEDA solutions react with the nitrite ions to create a pink azo-dye which has a maximum absorbance at a wavelength of 540 nm. The intensity of the azo-dye absorbance is linearly related to the mass of nitrite, and consequently using sampler geometries can be used to calculate the average concentrations of ambient NO_2 the samplers were exposed to.

The laboratory analysis procedures followed during this research for the analysis of the passive sampling devices is discussed below. Two separate methods were followed – the Palmes analysis method followed statutory guidance (DEFRA, 2008) while the Ogawa analysis followed manufacturer guidance (Ogawa, 2006).

2.4.1.1. Palmes diffusion tubes

Preparation of Palmes tubes:

1) A solution of 1:1 TEA:acetone was prepared on each occasion tubes were being prepared. The solution was thoroughly mixed in a beaker to ensure the TEA fully dissolved in the acetone.

2) The required number of stainless collection grids was added to the beaker using tweezers and ensuring no contact between skin and the grids (which can lead to contamination of the grids from chemicals such as nitrite present on the skin).

3) The grids were immersed in the solution for a few minutes, before transferring to tissue paper on the bench. Excess solution was removed from the grids by gently pressing another layer of tissue on top of the grids.

4) Two grids were placed in the cap of the Palmes diffusion tube using tweezers. The tubes were then assembled, ensuring the caps of the tubes were securely in place to ensure air could not enter and contaminate the tubes.

5) Between preparation and tube exposure the tubes were sealed in double-bags in a refrigerator.

6) For each batch of tubes prepared, spare tubes were also prepared to act as laboratory blanks. These tubes were prepared as above, and stored in the fridge without exposure. These tubes were then analysed (see below) along with the exposed tubes to ensure no contamination occurred during the preparation process.

Exposure of the Palmes tubes:

1) The cap of the tube containing the collection grids should be labelled with a unique identifier.

2) The tubes were attached vertically to lampposts (or similar), with the cap not containing the grids facing down.

3) The bottom cap of the tube was then removed, and the time at which the cap was removed was accurately recorded.

4) Two tubes (minimum) were deployed at each location to provide an indication of the precision measurements.

5) After the required exposure time, the tubes should be recapped and the time accurately recorded. The capped tubes should be secure to ensure no contamination.

6) Exposed tubes were stored in the refrigerator between retrieval and analysis.

Analysis of the Palmes tubes:

1) Standard nitrite solutions were prepared to act as calibration standards. Two standard solutions were prepared following the method in steps 2 – 4 below. These were then used to create two sets of calibration standards (steps 10 - 14) which were then analysed by UV to obtain calibration equations. If the calibrations for the two bulk solutions were coincident then the bulk was accepted – if the calibrations were not coincident the process was repeated until this was achieved.

2) Analytical grade solid sodium nitrite was dried in a 110 °C oven overnight, followed by cooling in a dessicator.

3) 1.500 g of the dried sodium nitrite was accurately weighed into a 1L volumetric flask. To the flask, 0.09 g of sodium hydroxide and 1 mL chloroform was added. The flask was agitated to ensure all the solids dissolved, before making up to the mark with deionised water. This bulk solution was 1000 mg/L nitrite.

4) A 1 mg/L nitrite solution was then prepared by taking 1 mL of the bulk solution and adding to 1 L of deionised water.

The bulk solutions can be stored in the fridge for a few months when not in use.

5) A sulphanilamide solution was prepared by dissolving 20 g sulphanilamide and 50 mL orthophosphoric acid (88 %) in a 1 L volumetric flask, made up to the mark with deionised water.

The sulphanilamide solution can be stored in the fridge for a few months when not in use.

6) A NEDA (N-1-naphthyl ethylene diamine dihydrochloride) solution was preared by dissolving 0.7 g NEDA in a 500 mL volumetric flask made up to the mark with deionised water.

The NEDA solution can be stored in the fridge for approximately 1 month.

The analysis of the exposed tubes, blank tubes and calibration standards all follow the method below:

8) Diffusion tubes placed with caps containing grids facing down and upper caps removed.

9) 1.5 mL deionised water was added to each tube and the tubes recapped securely. The tubes were then agitated and left to stand for 30 minutes.

10) Six additional empty tubes (clean tubes containing no grids) should be prepared for calibration standards, and the following reagents added:

a) Tubes A and B: 1.5 mL deionised water only

b) Tube C: 1.125 mL deionised water and 0.375 mL of 1 mg/L nitrite standard prepared in step 4 above (leading to a 375 ng nitrite / tube)

c) Tube D: 0.750 mL deionised water and 0.750 mL of 1 mg/L nitrite standard (750 ng nitrite / tube)

d) Tube E: 0.375 mL deionised water and 1.125 mL of 1 mg/L nitrite standard (1125 ng nitrite / tube)

e) Tube F: 1.500 mL of 1 mg/L nitrite standard (1500 ng nitrite / tube)

11) After the exposed tubes have sat for 30 minutes, remove the caps from both the exposed tubes and the calibration tubes. Add 1.50 mL sulphanilamide solution (prepared in step 5) to all tubes, quickly followed by 0.15 mL NEDA solution (prepared in step 6).

12) Replace the end caps for all tubes and shake briefly. Leave for 30 minutes to allow the pink azo dye to develop in colour.

13) After 30 minutes transfer the solutions to the cuvettes used for UV analsyis. Place one of the 0 ng nitrite / tube solution into the 'blank' space in the dual-beam spectrometer if using. Measure the absorbance of all of the tubes (including calibrations and blanks) and record the absorbance. The absorbance of all tubes should be measured twice – once at the start and end of the sequence of absorbance measurements.

14) Construct a calibration graph using the average absorbance (start and end) measured for each of the calibration samples (B – F).

15) Calculate the ambient concentration of NO_2 from the measured absorbance for each sampler using:

$$C = \frac{QL}{DAt}$$

Where C is the average ambient concentration of NO_2 during the sampler exposure (ng/cm^3); Q is the mass of nitrite measured in the tube determined from the calibration graph (ng); L is the length of the diffusion tube (Palmes tubes = 7.1 cm); D is the diffusion coefficient of NO_2 in air (0.151 cm^2/s); A is the cross sectional area of the tube (0.916 cm^2 for Palmes tubes); and t is the exposure time of the tube (s).

Cleaning of Palmes samplers:

The Palmes sampler parts can all be reused. The sampler caps and bodies should be scrubbed then soaked in Decon-50 detergent solution overnight, followed by air drying in a clean environment on a tissue-paper covered bench and covering the parts with clean tissue paper until dry.

The collection grids should be rinsed and soaked in Decon-50 detergent solutions. A fresh Decon-50 solution should then be added to the grids, and these should be agitated in an ultrasonic bath for a minimum of 30 minutes. The grids should then be transferred to a clean beaker using tweezers and dried in an oven overnight, followed by cooling in a dessicator.

The grids should be stored in an air-tight container, while the caps and tubes should be stored in double-bags.

2.4.1.2. Ogawa badge samplers

Preparation of Ogawa samplers:

1) The collection medium for the Ogawa samplers is collection pads available to purchase from the manufacturer. These pads are supplied pre-coated in TEA solution.
2) The sampler should be assembled in a clean environment using tweezers and not handling the sampler parts with bare hands. The Ogawa sampler has 2 identical chambers which should be prepared simultaneously. The collection pads should be inserted between the two stainless steel meshes of the Ogawa sampler.
3) The prepared Ogawa samplers were stored in air-tight opaque plastic containers in the fridge between preparation and exposure.
4) Laboratory blanks (prepared but not exposed) should be prepared at the same time and stored in the fridge until the sampler laboratory analysis. These act as controls to determine no contamination occurred during sampler preparation.

Exposure of Ogawa samplers:

5) The containers of the samplers were removed from the fridge approximately 12 hours prior to sampler exposure and allowed to equilibrate to room temperature. The containers were then transported to site.
6) The samplers were mounted on lampposts (or similar) using the support clip available to purchase from the manufacturer.
7) The samplers were removed from their containers and attached to the clip. An opaque weather shelter (available to purchase from the manufacturer) was placed over the exposed sampler in an attempt to minimise wind-speed effects. The time of sampler exposure was accurately noted.
8) After the required exposure time, the samplers were removed from the site, placed in the air-tight opaque containers and the time accurately recorded.
9) Between exposure and analysis the samplers were stored in their containers in the fridge.

Analysis of Ogawa samplers:

10) A standard stock solution of 1000 mg/L nitrite was prepared following steps 1 – 3 of the Palmes analysis procedure. A working nitrite solution was prepared by taking 1 mL of the standard stock and making up to 100 mL with deionised water.
11) A sulphanilamide solution was prepared by dissolving 80 g of sulphanilamide in 200 mL of phosphoric acid and made up to 100 mL with deionised water.
12) A NEDA solution was prepared by dissolving 0.56 g of NEDA in 1000 mL deionised water.
13) Remove the collection pads and surrounding stainless steel meshes from the exposed and blank Ogawa samplers and place into extraction vials containing 8 mL deionised water. Seal and shake the extraction vials then leave for 30 minutes.
14) Prepare calibration solutions by taking 0, 0, 2, 4, 6, 8 and 10 mL of the working solution and making up to 100 mL with deionised water (giving calibration solutions of 0 – 1 µg/mL nitrite). Place 8 mL of each calibration solution into a new extraction vial.
15) After 30 minutes place all vials (including calibrations) into the refrigerator to cool (approximately 15 minutes).
16) Prepare the colour-producing reagent. Mix sulphanilamide and NEDA solutions in a 10: 1 ratio.
17) Once the vials have cooled, add 2 mL of colour producing reagent to each vial. Shake the vials and place back into the fridge for 30 minutes.
18) Remove vials from the fridge and allow to equilibrate to room temperature.
19) Transfer the solution from each vial to a cuvette and measure the absorbance at 545 nm. The absorbance should be measured twice – one at the start and end of the sequence of sample vials.
20) Create a calibration slope by dividing the product of the concentrations of the calibration standards (µg/ml) by the measured absorbance for each calibration sample minus the blank by the sum of the concentrations of calibration standards squared.
21) The concentration of the exposed samples is calculated using

$$C = \frac{W * V * 1000 * \propto}{t}$$

Where C is the concentration of NO_2 (ppb); W is the mass (ng) of nitrite collected by the sampler (obtained by dividing the measured absorbance by the calibration slope); V is the extraction volume (8 mL); α is a conversion coefficient dependent on temperature and humidity; t is the exposure time (minutes).
The conversion coefficient α is calculated as follows:

$$\propto = \frac{10000}{(0.677 * [P] * [RH]) + (2.009 * [T]) + 89.8}$$

Where *[RH]* is the average relative humidity during sampler the exposure (%); *[T]* is the average temperature during sampler exposure (°C); and *[P]* is:

$$[P] = \left\{\frac{2P_N}{P_T + P_N}\right\}^{2/3}$$

Where P_N is the water vapour pressure in mm Hg at 20°C (17.353); and P_T is the vapour pressure of water at the ambient temperature measured during the study (constant provided for a range of ambient temperatures in the manufacturer manual).

Cleaning of Ogawa samplers:

The sampler parts of the Ogawa samplers can be reused after thorough cleaning. The samplers should be deconstructed, rinsed then left to soak in Decon-50 solution. The parts should then be placed on a tissue-paper covered bench under more tissue-paper and left to air dry. The collection pads were not reused.

2.5. References

Aeroqual Ltd, 2017. Portable Air Quality Monitors Brochure. https://d2pwrbx99jwry6.cloudfront.net/wp-content/uploads/Aeroqual-Portable-Monitors-Long-Brochure-LR.pdf (accessed 27/12/17)

Aethlabs, 2016. microAeth® AE51 Specification Sheet Revision 4. https://aethlabs.com/microaeth/ae51/tech-specs (accessed 26/12/17)

Cape, J.N., 2009. The use of passive diffusion tubes for measuring concentrations of nitrogen dioxide in air. Critical Reviews in Analytical Chemistry. 39(4), 289-310. doi:10.1080/10408340903001375

Cocheo, V., Boaretto, C., Sacco, P., 1996. High Uptake Rate Radial Diffusive Sampler Suitable for Both Solvent and Thermal Desorption. American Industrial Hygiene Association Journal. 57(10), 897 – 904. doi:10.1080/1542811969101440

DEFRA, 2016. Technical Guidance LAQM (TG16). https://laqm.defra.gov.uk/technical-guidance/ (accessed 03/01/2018)

DEFRA, 2008. Diffusion tubes for ambient NO_2 monitoring: practical guidance for laboratories and users. Report number: ED48673043. Available from: https://laqm.defra.gov.uk/documents/0802141004_NO2_WG_PracticalGuidance_Issue1a.pdf (accessed 04/01/18)

Hansen, A.D.A., Rosen, H., Novakov, T., 1982. The Aethalometer – an instrument for the real-time measurement of optical absorption by aerosol particles. Sci. Tot. Environ. 36, 191-196. doi:10.1016/048-9697(84)90265-1

Ogawa USA, 2006. NO, NO_2, NO_x and SO_2 sampling protocol using the Ogawa sampler V6.06. Available from: http://ogawausa.com/wp-content/uploads/2017/11/prono-noxno2so206_206_1117.pdf (accessed 04/01/18)

Palmes, E.D., Gunnison, A.F., DiMattio, J., Tomczyk, C., 1976. Personal sampler for nitrogen dioxide. American Industrial Hygiene Association Journal. 37, 570-577. doi: 10.1080/0002889768507522

Sather, M.E., Slonecker, E.T., Kronmiller, K.G., Williams, D.D., Daughtrey, H., Mathew, J., 2006. Evaluation of short-term Ogawa passive, photolytic, and federal reference method sampling devices for nitrogen oxides in El Paso and Houston, Texas. J. Environ. Monit., 8(5), 558-563. doi:10.1039/b601113f

TSI, 2012. Hand-held condensation particle counter Model 3007 Specification Sheet. http://www.tsi.com/uploadedFiles/_Site_Root/Products/Literature/Spec_Sheets/3007_5001117_A4.pdf (accessed 26/12/17)

Virkkula, A., Makela, T., Hilamo, R., Yli-Tuomi, T., Hirsikko, A., Hameri., K., Koponen, I., 2007. A simple procedure for correcting loading effects of Aethalometer data. J.Air & Waste Manage. Assoc., 57, 1214-1222. doi: 10.3155/1047-3289.57.10.1214.

Williams, D.E., Henshaw, G.S., Bart, M., Laing, G., Wagner, J., Naisbitt S., Salmond, J.A., 2013. Validation of low-cost ozone measurement instruments suitable for use in an air-quality monitoring network. Meas. Sci. Technol. 24, 065803. doi: 10.1088/0957-0233/24/6/065803

3. Evaluation of passive sampling devices to measured nitrogen dioxide concentrations

This chapter is comprised of two manuscripts, which are described below:

Influence of wind-speed on short-duration NO_2 measurements using Palmes and Ogawa passive diffusion samplers

In this work we investigated the influence of the common bias factors on passive diffusion samplers at an urban background location. The main finding of this work was that wind speed has a linear relationship with the sampler uptake rates, and in order to obtain accurate pollution estimates the effect of wind speed needs to be corrected for.

N. Masey designed the experiment, carried out the field measurements and laboratory analysis, analysed the data and produced the manuscript. J. Gillespie discussed the experimental design, helped with field measurements and provided discussion about the data analysis and editorial comments. S. Hamilton, M. Heal and I. Beverland provided assistance with data analysis and editorial comments.

This manuscript was published in Atmospheric Environment in April 2017 (DOI: http://doi.org/10.1016/j.atmosenv.2017.04.008)

Minimising the impact of wind-speed effects on NO_2 passive diffusion samplers through sampler modification

The second paper evaluates the relationship between wind speed and uptake rate identified above at a second, roadside, site and additionally tests designs of passive samplers thought to minimise the effects of wind speed through modification of sampler geometry.

N. Masey designed the experiment, was involved with the field measurements and laboratory analysis, and carried out the data analysis and produced the manuscript. F. Sutherland and S. Grainger assisted with field and laboratory analysis and were involved with the interpretation during data analysis. I. Beverland assisted with experimental planning and provided feedback on the manuscript. S. Hamilton and M. Heal provided discussion about the data analysis and manuscript editing.

This paper has been formatted to meet Atmospheric Environment guidelines.

Influence of wind-speed on short-duration NO_2 measurements using Palmes and Ogawa passive diffusion samplers

Nicola Masey[1], *Jonathan Gillespie*[1], *Mathew R. Heal*[2], *Scott Hamilton*[3], *Iain J. Beverland*[1*]

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]School of Chemistry, Joseph Black Building, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

[3]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

*CORRESPONDING AUTHOR: Dr Iain J. Beverland, email: iain.beverland@strath.ac.uk; Tel: +44 141 548 3202

Research highlights:

- A variant on the standard Ogawa passive diffusion sampler was also tested
- Wind-induced turbulence increased passive sampler uptake rates.
- In general, relationships between wind speed and uptake rates were linear.
- Estimation of wind speed variations allowed correction of wind speed effects.
- Wind protection of samplers appears necessary when wind speed variations are unknown.

Abstract

We assessed the precision and accuracy of nitrogen dioxide (NO_2) concentrations over 2-day, 3-day and 7-day exposure periods measured with the following types of passive diffusion samplers: standard (open) Palmes tubes; standard Ogawa samplers with commercially-prepared Ogawa absorbent pads (Ogawa[S]); and modified Ogawa samplers with absorbent-impregnated stainless steel meshes normally used in Palmes tubes (Ogawa[P]). We deployed these passive samplers close to the inlet of a chemiluminescence NO_2 analyser at an urban background site in Glasgow, UK over 32 discrete measurement periods. Duplicate relative standard deviation was < 7% for all passive samplers. The Ogawa[P], Ogawa[S] and Palmes samplers explained 93%, 87% and 58% of temporal variation in analyser concentrations respectively. Uptake rates for Palmes and Ogawa[S] samplers were positively and linearly associated with wind-speed ($P < 0.01$ and $P < 0.05$ respectively). Computation of adjusted uptake rates using average wind-speed observed during each sampling period increased the variation in analyser concentrations explained by

Palmes and Ogawa[S] estimates to 90% and 92% respectively, suggesting that measurements can be corrected for shortening of diffusion path lengths due to wind-speed to improve the accuracy of estimates of short-term NO_2 exposure. Monitoring situations where it is difficult to reliably estimate wind-speed variations, e.g. across multiple sites with different unknown exposures to local winds, and personal exposure monitoring, are likely to benefit from protection of these sampling devices from the effects of wind, for example by use of a mesh or membrane across the open end. The uptake rate of Ogawa[P] samplers was not associated with wind-speed resulting in a high correlation between estimated concentrations and observed analyser concentrations. The use of Palmes meshes in Ogawa[P] samplers reduced the cost of sampler preparation and removed uncertainty associated with the unknown manufacturing process for the commercially-prepared collection pads.

Keywords Air pollution; passive samplers; uptake rates; Palmes; Ogawa

1. Introduction

Exposure to nitrogen dioxide (NO_2) has been associated with adverse effects on human health, including cardio-respiratory illness, hospital admissions and mortality (World Health Organization, 2013). Passive samplers, because of their relatively low cost and simplicity of deployment, have been used in studies to estimate outdoor NO_2 concentrations over large geographical areas (Cyrys et al., 2012; Gillespie et al., 2017, 2016; Lewné et al., 2004) and for indoor and personal exposure (Yu et al., 2008). However the temporal resolution of passive samplers is limited, and a number of potential issues may affect their accuracy and precision (Cape, 2009).

Passive samplers can be grouped into tube and badge designs (Cape, 2009; Tang et al., 2001; Yu et al., 2008). Tube samplers, including the commonly-used Palmes sampler (Palmes et al., 1976), have relatively long diffusion paths and low uptake rates, and are typically used for measuring concentrations over 1 - 5 week periods. Badge samplers, including Ogawa samplers (http://ogawausa.com/), have relatively short path lengths and higher uptake rates facilitating measurement of relatively low NO_2 concentrations and exposures over shorter intervals (e.g. 24 hours).

Palmes samplers have often been observed to overestimate concentrations measured by automatic analysers during co-location studies (Cape, 2009). Possible reasons for such overestimations include: wind-speed induced turbulence effectively shortening the diffusion path; and chemical reactions within the diffusion path that result in misrepresentation of external photochemical conditions. Chamber and wind-tunnel studies have reported positive

associations between wind-speed and sampler uptake rates for open tube samplers (Buzica et al., 2005; Martin et al., 2014; Plaisance et al., 2004). Field studies using wind shelters and/or protective meshes or membranes to minimise turbulence within the diffusion path have noted reduced sampler overestimation and higher correlations between sampler and analyser measurements (Bush et al., 2001; Martin et al., 2014; Plaisance et al., 2004). Concentrations of NO_2 determined using UV-transmitting quartz tubes have shown smaller overestimations than standard acrylic tubes suggesting reduction in NO_2 photolysis in the former (Heal et al., 1999, 2000), although this effect has been observed to only occur in summer (Kirby et al., 2000). Collectively there is limited observational evidence regarding whether a single factor is the dominant cause of the overestimations in NO_2 concentrations frequently observed by Palmes tubes, or if they are the consequence of multiple site-specific environmental factors.

The order-of-magnitude shorter diffusion path length in Ogawa samplers is anticipated to limit within-tube chemical effects. The manufacturer of the Ogawa samplers provides shelters to minimise the effect of wind on the relatively large face area of the badges. Temperature, absolute humidity, and wind-speed have been observed to be positively associated with the uptake rate of Ogawa samplers (Hagenbjörk-Gustafsson et al., 2010).

Our study deployed standard Palmes and Ogawa samplers (the latter hereafter referred to as Ogawa[S]) at an urban background monitoring site in Glasgow, UK. We also prepared and deployed modified Ogawa samplers with absorbent-impregnated stainless steel meshes normally used in Palmes tubes (hereafter referred to as Ogawa[P]) to: (a) eliminate one of the technical differences in our comparison between Palmes and Ogawa samplers (namely the difference between collection medium used in the two techniques); (b) reduce operational costs; and (c) eliminate uncertainty in our scientific description of the methods associated with the undisclosed preparation methods of the pre-coated collection pad sold by the manufacturer. The samplers were deployed for 2-day and 3-day exposure times; and 7-day exposure times matched to consecutive 2-day and 3-day periods. Potential influences of meteorological and atmospheric composition factors on agreement between sampler and reference chemiluminescence analyser measurements were investigated. The 2-day and 3-day deployments increased the range of meteorological and chemical conditions sampled by avoiding the reduction in variation of these conditions resulting from longer-term averaging, and allowed evaluation of the precision and accuracy of the samplers at these short exposure periods.

2. Methods

Duplicate Palmes, Ogawa[S] and Ogawa[P] samplers were deployed next to the inlet to reference gas analysers at the Townhead air quality monitoring site in Glasgow, UK (latitude: 55.866°, longitude: –4.243°) (Figure 1). The Ogawa[S] and Ogawa[P] samplers were deployed under shelters purchased from the manufacturer. The Palmes tubes were unsheltered, as is standard practice. The Townhead site is in an urban background location, approximately 1 km north of the city centre and 122 m from the nearest road. Concentrations of NO_2 and O_3 were measured as hourly averages using a Teledyne-API 200A NO_x chemiluminescent analyser and a Thermo Scientific Model 49i O_3 analyser respectively. Both analysers undergo regular Quality Assurance and Quality Control assessments as part of the UK Automatic Urban and Rural Network (www.scottishairquality.co.uk). The concentrations measured by the analysers at Townhead were averaged to the same time periods as the passive sampler exposures.

(a) (b)

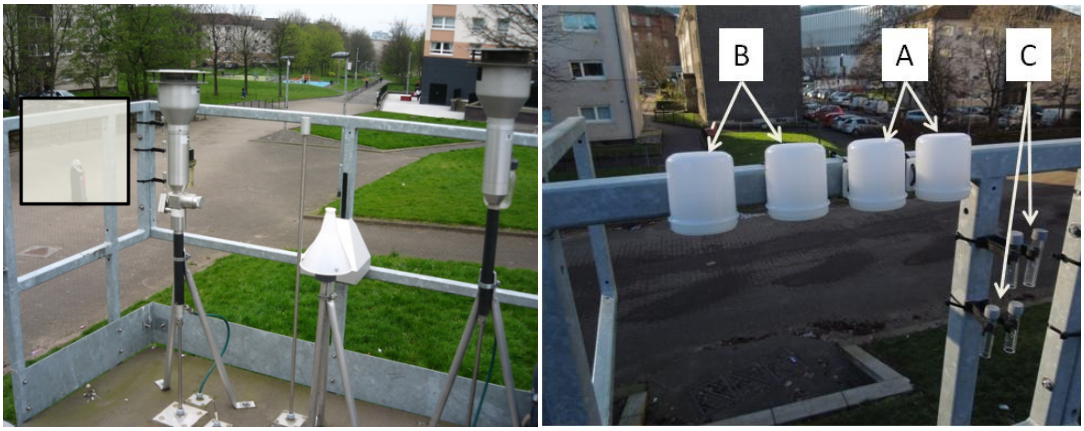


Figure 1: Location of passive NO_2 samplers on the roof of Townhead automatic monitoring site, Glasgow. (a) Location of passive samplers (highlighted by grey square) close to the inlets of the automatic analyser. (b) Deployment of samplers: A = Ogawa[S]; B = Ogawa[P]; C = Palmes. Temperature and relative humidity were measured using HOBO sensors, located on the railing opposite the passive samplers.

The passive samplers were mounted on a railing at a common height approximately 1 m horizontal distance from the analyser inlets. Temperature and relative humidity were measured at 1-minute intervals using an Onset HOBO U23 Pro v2 External Data logger located on the railing under a solar radiation shield, from which hourly-average data were calculated. Hourly wind-speed data was obtained from the nearest Meteorological Office site at Glasgow airport (approximately 10 km west of the Townhead site). Both Glasgow airport and Townhead monitoring sites are in open locations, therefore the wind-speeds

recorded at Glasgow airport were anticipated to provide a reasonably reliable indication of relative temporal changes in wind-speed at Townhead.

We made passive sampler measurements over 32 discrete periods, ranging in duration from 2 days to 8 days, between February 2015 and October 2015. Five of the week-long passive sampler exposure periods were contemporaneous with cumulative 2, 2 and 3 day exposure periods (Supplementary Information Table S1). The 2 and 3-day exposure times provided a wider range of field conditions to test the passive samplers as the impact of averaging over longer exposures was reduced. The ranges (and relative standard deviations) of exposure-averaged wind-speed for all 2 & 3-day and all 7-day exposure periods were 1.1 – 8.2 m/s (46%) and 2.1 – 5.4 m/s (29%) respectively.

We prepared Palmes samplers with two stainless-steel collection meshes dipped in a 1:1 triethanolamine (TEA):acetone solution (Heal, 2008). We assembled Ogawa[S] samplers using pre-coated TEA collection pads purchased from the manufacturer, and Ogawa[P] samplers using two Palmes meshes prepared using the same process as the Palmes samplers. Samplers were prepared in 8 batches and were stored in a refrigerator in sealed bags (Palmes) and containers (Ogawa) before deployment, and between deployment and laboratory analysis.

We analysed the 8 batches of samplers separately. Nitrite collected by Palmes samplers was extracted into aqueous solution and quantified by the Saltzman reaction and colorimetric absorption at 540 nm. The nitrite mass was converted to ambient concentration of NO_2 using the diffusion coefficient for NO_2 in air, the internal length and cross-sectional area of the tube, and the exposure time (Targa and Loader, 2008). The nitrite mass collected by Ogawa[S] samplers and calculation of ambient NO_2 were determined using aqueous extraction and colorimetric absorption following the manufacturer's protocol (Ogawa, 2006). The laboratory determination of nitrite mass collected by Ogawa[P] samplers followed the method in Targa and Loader (2008), while the subsequent conversion to ambient NO_2 used the Ogawa protocol. The volumetric mixing ratios (ppb) of NO_2 calculated from the Ogawa protocol were converted to gravimetric units ($\mu g/m^3$) using a factor of 1.9125 (conversion at 20 °C and 1013 mb) to match the reporting conditions of the reference analyser concentrations.

Duplicate laboratory blanks (for both Ogawa absorbent pads and Palmes coated meshes) were prepared with each batch of samplers and stored in the refrigerator during sampler deployments. The blanks were analysed in an identical manner to the samples and the average masses of nitrite in the blanks were subtracted from the appropriate samples prior to calculation of sampler NO_2 concentration.

The masses of nitrite collected by the Ogawa[P] sampler were lower than those collected by the Ogawa[S] sampler. Linear regression indicated that Ogawa[P] nitrite = 0.62 × Ogawa[S] nitrite (R^2 = 0.93) (Figure S1). The slope of this regression line was of similar magnitude to the ratio of the area of the Palmes mesh to the area of the Ogawa[S] collection pad (0.916 cm^2 / 1.65 cm^2 = 0.56). Therefore the lower nitrite masses collected by the Ogawa[P] samplers was consistent with this difference in collection areas between the Ogawa[S] pad and Palmes mesh.

An empirical passive sampler uptake rate (UR) for each exposure was calculated from measured nitrite mass as follows:

$$UR\ (cm^3\ min^{-1}) = \left[\frac{nitrite\ mass\ (ng)}{([NO_2]\ (ng\ m^{-3}) \times t\ (min))}\right] \times 10^6\ (cm^3\ m^{-3}) \quad (1)$$

where $[NO_2]$ = exposure-averaged analyser NO_2 concentration, and t = exposure duration. Associations between meteorological variables and uptake rates were examined using linear regression.

3. Results and Discussion

3.1. Precision and limits of detection

The average Limit of Detection (LoD), calculated as three times the standard deviation of the blank concentration plus average blank concentration using the shortest exposure time in the calculation, was 1.1 µg/m^3 for the Ogawa samplers and 10.1 µg/m^3 for the Palmes sampler (Table 1). These LoDs were substantially lower than the minimum concentrations measured by the samplers.

The mean relative standard deviation (RSD) for duplicate measurements was < 7% for all samplers and exposure durations (Table 1, Figure S2). These mean RSD values for our measurements were within the 2.8% to 11.0% range of published statistics for exposure periods of 1-week or greater (Bush et al., 2001; Buzica et al., 2005; Heal et al., 1999a, 1999, 2000; Kirby et al., 2000; van Reeuwijk et al., 1998; Vardoulakis et al., 2009), and were lower than the mean RSD of 12.6% reported by Heal et al. (1999a) for 2-day indoor Palmes measurements (this was the only study using Palmes tube exposures of less than 3-days that we located). Our good precision data highlight the potential for passive sampling to provide greater temporal resolution than the time periods normally used. In the remainder of this paper the duplicate mean is used as the NO_2 concentration for a given sampler and exposure period.

Table 1: Limits of detection (LoD) and duplicate relative standard deviation (RSD) statistics for the three passive sampler types and their root mean square error (RMSE), normalised mean bias (NMB), regression forced through the origin (and 95 % confidence interval) and R^2 statistics with respect to the analyser measurements. The RSD, NMB and RMSE statistics are presented for all studies, for exposures of 3 days and less ($t \leq 3$), and for exposures of greater than 3 days ($t > 3$). Statistics are presented for sampler NO_2 concentrations before and after correction of uptake rates (second and third panels respectively – see text for details of correction procedure).

	Palmes	Ogawa[S]	Ogawa[P]
LoD All ($\mu g/m^3$)	10.8	1.1	1.1
Mean duplicate RSD All (%)	5.6	2.5	5.9
Mean duplicate RSD $t \leq 3$ (%)	6.0	2.2	6.4
Mean duplicate RSD $t > 3$ (%)	4.7	3.0	4.5
Data using theoretical uptake rates:			
RMSE sampler *vs.* analyser All ($\mu g/m^3$)	9.6	4.8	8.9
RMSE sampler *vs.* analyser $t \leq 3$ ($\mu g/m^3$)	11.1	5.6	8.9
RMSE sampler *vs.* analyser $t > 3$ ($\mu g/m^3$)	4.9	2.3	8.8
NMB sampler *vs.* analyser All	0.28	0.11	–0.33
NMB sampler *vs.* analyser $t \leq 3$	0.30	0.17	–0.33
NMB sampler *vs.* analyser $t > 3$	0.20	–0.04	–0.32
Regression slope (95 % CI)	1.25 (1.15-1.35)	1.11 (1.06-1.17)	0.67 (0.64-0.69)
R^2 sampler *vs.* analyser All	0.59	0.87	0.93
Data using empirical uptake rates:			
RMSE sampler *vs.* analyser All ($\mu g/m^3$)	3.9	3.4	2.7
RMSE sampler *vs.* analyser $t \leq 3$ ($\mu g/m^3$)	4.5	3.2	3.1
RMSE sampler *vs.* analyser $t > 3$ ($\mu g/m^3$)	2.4	3.7	1.1
NMB sampler *vs.* analyser All	0.03	0.02	0.02
NMB sampler *vs.* analyser $t \leq 3$	0.05	0.08	0.03
NMB sampler *vs.* analyser $t > 3$	–0.01	–0.11	–0.02
Regression slope (95 % CI)	1.06 (1.00-1.09)	1.05 (1.00 – 1.09)	1.03 (0.99 – 1.07)
R^2 sampler *vs.* analyser All	0.90	0.92	0.93

3.2. Sampler accuracy

Using the measured nitrite masses and standard protocols for computation of atmospheric concentrations described in the Methods section, concentrations derived from Palmes and Ogawa[S] samplers were generally higher than reference analyser observations, and concentrations from Ogawa[P] samplers were generally lower than the analyser observations (Figure 2, Table 1). Our observation of Palmes sampler overestimation of analyser concentrations was consistent with the previous studies reviewed in Section 1. Closer

inspection of the Ogawa[S] data showed that while the 2 and 3-day Ogawa[S] exposures overestimated analyser concentrations, the 7-day exposures underestimated, as has been noted previously (Hagenbjörk-Gustafsson et al., 2010; Mukerjee et al., 2008; Sather et al., 2006, 2007; van Reeuwijk et al., 1998). Overall, however, Ogawa[S] samplers had the lowest Root Mean Square Errors (RMSE) and Normalised Mean Bias (NMB) values, while the Palmes samplers had the largest deviations from the analyser (Table 1). The Ogawa[P], Ogawa[S] and Palmes samplers explained 93%, 87% and 58% of temporal variation in analyser concentrations respectively. Preparation of the Ogawa[P] meshes according to standard procedures (Heal, 2008; Targa and Loader, 2008) eliminated scientific uncertainty associated with the unspecified preparation method for the commercially-available collection pads used in Ogawa[S] samplers, and reduced the cost of sampler preparation.

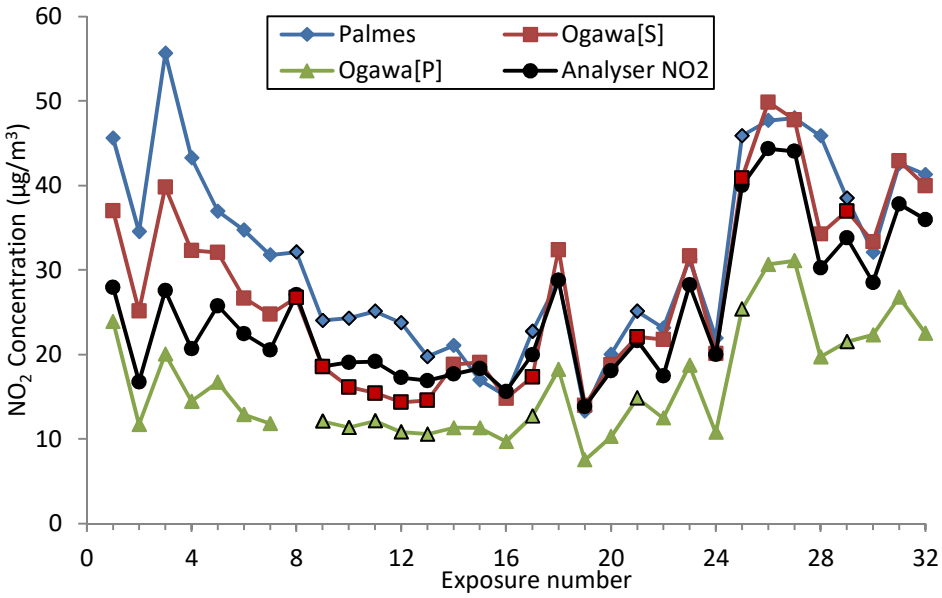


Figure 2: Concentrations of NO_2 measured by automatic analyser and by Palmes, Ogawa[S] and Ogawa[P] samplers (prior to correction for observed wind-speed effects) for 32 separate exposure periods. Symbols with a dark border are measurements from exposures of greater than 3 days.

3.3. Long-term vs. cumulative short-term measurements

The time-weighted weekly average from the three cumulative short-term exposures showed close agreement with the corresponding simultaneous 1-week exposure (Figure S3). Our results are consistent with Heal et al. (1999a) who did not observe significant differences between cumulative 2 or 3-day and 1-week indoor exposure measurements. In contrast, significantly lower concentrations have been observed between 1-month

measurements and cumulative 1 or 2 week measurements in parallel, which have been attributed to loss in the long term of absorbed NO_2 from the TEA absorbent (Bush et al., 2001; Heal et al., 1999a, 1999, 2000; Kirby et al., 2000).

3.4. Correction of sampler uptake rates by wind-speed

We examined correlations between empirical sampler uptake rate, calculated using Equation 1, and the following meteorological and atmospheric composition variables (Table 2) previously reported to be associated with passive sampler uptake rate: air temperature; relative and absolute humidity (Cape, 2009; Hagenbjörk-Gustafsson et al., 2010); wind-speed (Bush et al., 2001; Buzica et al., 2005; Martin et al., 2014; Plaisance et al., 2004); atmospheric NO_2 concentration; and an atmospheric chemistry metric representing the potential for within-tube formation of additional NO_2 [ratio of analyser NO_2 plus minimum of analyser NO or O_3, to analyser NO_2 (Ratio(Min(NO,O_3)+NO_2)/NO_2)] (Heal et al., 1999).

Palmes sampler uptake rates were correlated with temperature and absolute humidity in the opposite direction to that expected from the literature (Cape, 2009) (Table 2). This likely resulted from confounding by negative correlations between temperature and wind-speed, and between absolute humidity and wind-speed (Figure S4). The empirical Ogawa[S] sampler uptake rate was also correlated with temperature in the opposite direction to that expected (Hagenbjörk-Gustafsson et al., 2010). Similarly, the positive correlations between Ogawa[S] and Ogawa[P] sampler uptake rates and relative humidity were contradictory to expectations (Cape, 2009; Hagenbjörk-Gustafsson et al., 2010). The Ogawa[P] sampler was the only sampler whose uptake rate was significantly correlated with ambient NO2 concentrations – the reasons for this correlation are unclear.

Table 2: Pearson correlation coefficients for bivariate relationships between passive sampler empirical uptake rates and exposure-averaged observed meteorological and atmospheric composition variables.

	Palmes_UR	Ogawa[S]_UR	Ogawa[P]_UR
T	–0.69**	–0.48**	–0.22
RH	0.32	0.51**	0.46**
AH	–0.62**	–0.31	–0.04
WS	0.78**	0.40*	0.04
NO_2	–0.15	0.20	0.41*
Ratio(Min(NO,O_3)+NO_2)/NO_2	–0.39*	–0.04	0.18

*Variables are exposure means of: air temperature (T, °C); relative humidity (RH, %); absolute humidity (AH, g/m^3); wind-speed (WS, m/s); analyser NO_2 (NO_2, ppb), ratio of analyser NO_2 plus minimum of analyser NO (ppb) or O_3 (ppb) to analyser NO_2 (Ratio(Min(NO,O_3)+NO_2)/NO_2); uptake rates (UR) for Palmes, Ogawa[S] & Ogawa[P] samplers. *Correlation coefficient significant at $P < 0.05$; **Correlation coefficient significant at $P < 0.01$.*

The correlations between the empirical Palmes and Ogawa[S] uptake rates and wind-speed were significant. The general pattern of increased sampler uptake rate as wind-speed increased (Figure 3a) is consistent with increasing wind-speeds reducing effective diffusion path length (and hence increasing effective uptake rate) because of induced turbulence, and consistent also with similar patterns noted in previous field measurements (Bush et al., 2001; Buzica et al., 2005; Martin et al., 2014; Plaisance et al., 2004). The reduced, but still positive, correlation between the Ogawa[S] sampler uptake rate and wind-speed (Figure 3c) suggested that the wind shelter provided by the manufacturer might not fully protect the sampler from wind effects. On the other hand, there was negligible correlation between Ogawa[P] sampler uptake rate and wind-speed (Figure 3e), despite both types of sampler being deployed under the same type of shelter.

For wind-speeds > 2.3 m/s, the empirical uptake rates for the Palmes sampler were greater than the theoretical uptake rate of 1.2 cm^3/min calculated from sampler geometry (Figure 3a), resulting in large overestimations of NO_2 concentrations at these wind-speeds (Figure 3b). The empirical uptake rates calculated for the Ogawa[S] sampler (Figure 3c) were mostly smaller than the theoretical uptake rate of 12.1 cm^3/min calculated from sampler geometry (Tang et al., 2014) but larger than the uptake rate of 9.3 cm^3/min calculated using manufacturer conversion factors (conversion factor of 56 ppb min/ng for NO_2 at 20 ºC and 70 % RH – see Supplementary Information for details).

From the above analyses of our observations, wind-speed was the variable with the most consistent and pronounced correlation with sampler uptake rate in the anticipated direction (Table 2). Therefore, we corrected sampler uptake rate for wind-speed influence following a similar approach to that used in an evaluation study of Ogawa samplers in Sweden (Hagenbjörk-Gustafsson et al., 2010). We calculated the linear regression between empirical uptake rate (computed from Equation 1) and exposure-period-average wind-speed across the set of measurement periods (Figures 3a, 3c, 3e). This regression line was used to correct the uptake rate using the measured average wind-speed for each exposure period, and the corrected uptake rate then used to correct the estimates of atmospheric concentrations from measured nitrite mass and exposure time by rearrangement of Equation 1. We then compared the corrected NO_2 concentrations to analyser measurements using linear regression lines forced through the origin (Figures 3b, 3d, 3f).. Following reasoning given by Martin et al. (2014) and Pfeffer et al. (2010) we forced the regression lines through the origin because of insignificant laboratory blank concentrations for our passive samplers and incomplete temporal coverage by the reference analyser (resulting from analyser calibration and/or maintenance activities).

After we corrected uptake rates for wind-speed we observed a 30% increase in the explained variation in the bivariate relationship between the corrected Palmes concentrations and analyser concentrations (Figure 3b). We noted a smaller (~ 5%) increase in explained variation for Ogawa[S] after wind-speed correction (Figure 3d). The absence of substantial correlation between the Ogawa[P] sampler uptake rate and wind-speed (Figure 3e) meant that correction was only appropriate by adjusting the uptake rate to the average empirical uptake rate of 5.9 cm^3/min, which was approximately half of the theoretical Ogawa uptake rate because of the approximately factor 2 smaller Ogawa[P] sampler NO_2 capture surface area noted in Section 2.

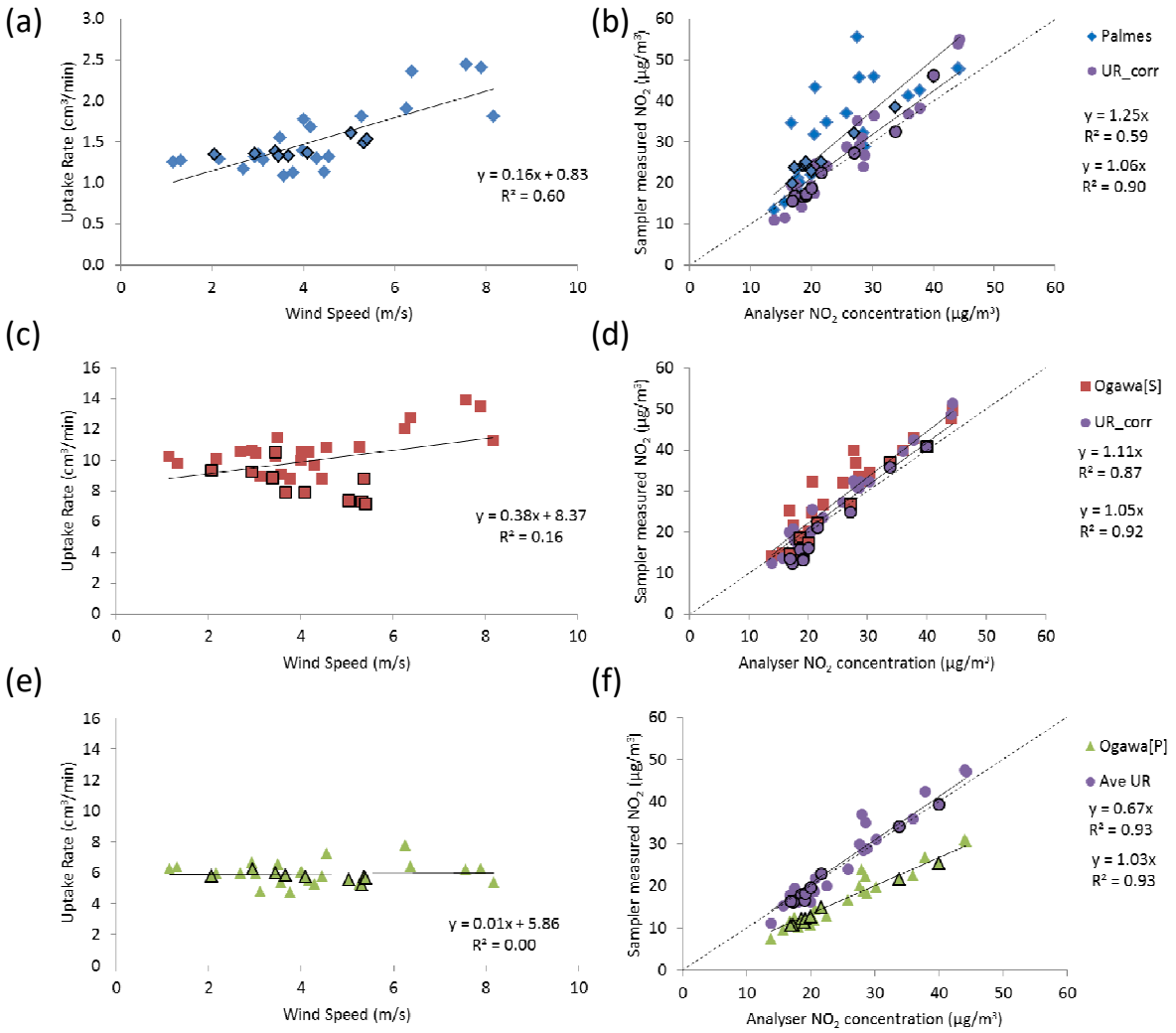


Figure 3: Correction of NO_2 passive sampler measurements using average wind-speed measured during each exposure period (Palmes and Ogawa[S]) and using an average uptake rate (5.9 cm^3/min) for Ogawa[P] due to the lack of a relationship with wind-speed. The uptake rate for each sampler was calculated from the nitrite mass collected by individual samplers divided by the product of exposure time and average analyser concentration and plotted against wind-speed (Graphs: a; c; e). The linear regression line was used to derive a corrected uptake rate for the samplers for the wind-speed during a given exposure period. The corrected uptake rate was then used to calculate a corrected passive sampler NO_2 concentration which is plotted against the automatic analyser concentration, alongside the concentration of NO_2 calculated using the standard method for each sampler (Graphs: b; d; f). Symbols with a dark border are measurements from exposures of greater than 3 days.

The above observations highlight the importance of evaluation of samplers under field conditions and suggest that it may be beneficial to calibrate uptake rates for the Palmes and

Ogawa samplers to the conditions under which measurements are made. Our study used wind-speed data from a single meteorological station, at an open site at the edge of the urban conurbation containing the pollution monitoring site, to correct the passive sampler data. Ideally we would have made wind speed measurements at the pollution monitoring site, however we anticipate the meteorological station data provided a reasonably reliable indication of relative temporal changes in wind-speed at the similarly open monitoring site, and any systematic difference in wind-speed between the two locations would have been accommodated by the empirical nature of the corrections that we developed. We anticipate the correlation between sampler uptake rates and wind speed will be apparent at other locations with similar outlooks to Townhead (i.e. relatively open urban background locations). The method identified in this work could be readily transferred to these other areas, however the linear regression equations identified are likely to be dependent on the magnitude and range of wind speeds at other locations. Future work should investigate the transferability of the method presented in this paper at other locations and/or different seasons.

In some exposure assessment situations the influence of wind-speed can be dealt with by the type of calibration we describe in this paper using measured or estimated wind-speeds (e.g. where passive diffusion samplers are used to record relative temporal changes in gas concentrations at a single site).

In other situations the correction required may be more difficult to implement, for example when the objective is to compare pollution concentrations at multiple sites with different exposure to prevailing wind conditions. In this latter type of situation it would be necessary to record wind measurements at the multiple sites or use a weather model to estimate the wind-speeds to allow site-specific correction of the sampler uptake rates. If the passive samplers were to be used over multiple sites without reliable information on wind speed variations, or in personal sampling to monitor in individuals with different levels and types of physical activity patterns (*e.g.* comparison of pollution exposures in cyclists, pedestrians and car drivers where the passive sampling device will have different relative speeds compared to the surrounding air), our observations suggest that it appears necessary to use field calibrated passive sampling devices with some form of modification to prevent wind-induced turbulence within the diffusion path. Previous studies have attempted to reduce the effect of wind-speed on sampler precision and accuracy through use of shelters over sampler inlets (Bush et al., 2001; Martin et al., 2014; Plaisance et al., 2004). However, shelters may increase the risk of vandalism by their conspicuous appearance; air under the shelter may be of a different composition to ambient air (Kirby et al., 2000); and for personal

sampling it may be difficult to expose the samplers under a shelter. The use of a mesh or membrane across the open end of the diffusion sampler is an alternative, and perhaps more practical, modification to the samplers to reduce the effect of wind-speed turbulence on sampler precision and accuracy.

4. Conclusions

We used standard (open) Palmes tubes, standard Ogawa[S] samplers with commercially-available absorbent pads, and modified Ogawa[P] samplers with TEA-impregnated meshes normally used in Palmes tubes to measure NO_2 at an urban background automatic monitoring site in Glasgow for exposure periods ranging from 2 days to 1-week.

Duplicate relative standard deviation was < 7% for all passive samplers for both short ($t \leq$ 3 days) and long ($t >$ 3 days) exposures demonstrating good potential for application of passive NO_2 sampling at finer temporal resolution than the time periods commonly used.

The Ogawa[P], Ogawa[S] and Palmes samplers explained 93%, 87% and 58% of temporal changes in analyser concentrations respectively. Palmes and Ogawa[S] sampler uptake rates were positively and linearly correlated with wind-speed, which enabled empirical correction of the uptake rates and subsequent re-estimation of corrected NO_2 concentrations. After these corrections the Palmes and Ogawa[S] sampler estimates explained a larger proportion (additional 30% and 5% respectively) of variation in analyser concentration, with regression lines closer to 1:1. The Ogawa[S] sampler uptake rate was similarly, but less markedly, influenced by wind-speed. Our observations suggest that if Palmes and Ogawa[S] samplers are exposed in windy environments (e.g. > 2 m/s) field calibrated uptake rates appear to be necessary to account for the effect of wind-speed on sampler concentration estimates.

The Ogawa[P] uptake rate was not correlated with wind-speed and explained a slightly higher proportion of variation in analyser concentrations than the Ogawa[S] sampler. After adjustment of individual Ogawa[P] uptake rates to the average observed Ogawa[P] uptake rate (5.9 cm^3/min) the regression between the Ogawa[P] and analyser NO_2 measurements was closer to the 1:1 line and maintained a high R^2 value (R^2 = 0.93). Therefore, the use of Palmes meshes in Ogawa samplers was a successful adaptation of the Ogawa sampler providing a reduction in cost of sampler preparation with specified preparation protocols. Further field-testing will help to establish if this observed average uptake rate for the modified Ogawa[P] sampler allows accurate estimation of analyser concentrations at other times and locations.

This research has highlighted that passive samplers require field evaluation at automatic pollution monitoring station to calibrate uptake rates to environmental conditions. In

particular Palmes and Ogawa[S] sampling uptake rates were substantially influenced by wind-speed and we have suggested a method to correct the sampler uptake rates when estimates of wind-speed variations are available. Monitoring situations where it is difficult to reliably estimate wind-speed variations, e.g. across multiple sites with different unknown exposures to local winds; or in personal exposure monitoring; are likely to benefit from protection of the sampling device from the effects of wind, for example by placing a mesh or membrane across the open end prior to field calibration. In light of these findings we would recommend the use of the Palmes sampler (after correction for wind-speed) or the Ogawa[P] sampler to obtain the most accurate estimates of ambient NO_2 concentrations for short-duration exposures.

Acknowledgements

Nicola Masey is funded through a UK Natural Environment Research Council CASE PhD studentship (NE/K007319/1), with support from Ricardo Energy and Environment. Jonathan Gillespie is funded through an Engineering and Physical Sciences Research Council Doctoral Training Grant (EPSRC DTG EP/L505080/1 and EP/K503174/1) studentship, with support from the University of Strathclyde and Ricardo Energy and Environment. We acknowledge access to the AURN measurement data, which were obtained from uk-air.defra.gov.uk and are subject to Crown 2014 copyright, Defra, licenced under the Open Government Licence (OGL). The research data associated with this paper are available at: http://dx.doi.org/10.15129/47838538-8f01-41c8-87f4-b70a07888233.

References

Bush, T., Stevenson, K., Moorcroft, S., Smith, S., 2001. Validation of nitrogen dioxide diffusion tube methodology in the UK. Atmos. Environ. 35, 289–296. doi:10.1016/S1352-2310(00)00172-2

Buzica, D., Gerboles, M., Amantini, L., Ballesta, P.P., De Saeger, E., 2005. Modelling of the uptake rate of the nitrogen dioxide Palmes diffusive sampler based on the effect of environmental factors. J. Environ. Monit. 7, 169. doi:10.1039/b411474d

Cape, J.N., 2009. The Use of Passive Diffusion Tubes for Measuring Concentrations of Nitrogen Dioxide in Air. Crit. Rev. Anal. Chem. 39, 289–310. doi:10.1080/10408340903001375

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

Gillespie, J., Beverland, I.J., Hamilton, S., Padmanabhan, S., 2016. Development, Evaluation, and Comparison of Land Use Regression Modeling Methods to Estimate Residential Exposure to Nitrogen Dioxide in a Cohort Study. Environ. Sci. Technol. 50, 11085–11093. doi:10.1021/acs.est.6b02089

Gillespie, J., Masey, N., Heal, M.R., Hamilton, S., Beverland, I.J., 2017. Estimation of spatial patterns of urban air pollution over a 4-week period from repeated 5-min measurements. Atmos. Environ. 150, 295–302. doi:10.1016/j.atmosenv.2016.11.035

Hagenbjörk-Gustafsson, A., Tornevi, A., Forsberg, B., Eriksson, K., 2010. Field validation of the Ogawa diffusive sampler for NO2 and NOx in a cold climate. J. Environ. Monit. 12, 1315. doi:10.1039/b924615k

Heal, M.R., 2008. The effect of absorbent grid preparation method on precision and accuracy of ambient nitrogen dioxide measurements using Palmes passive diffusion tubes. J. Environ. Monit. 10, 1363–1369. doi:10.1039/b811230d

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. doi:10.1023/a:1006249016103

Heal, M.R., O'Donoghue, M.A., Aguis, R.M., Beverland, I.J., 1999a. Application of passive diffusion tubes to short-term indoor and personal exposure measurement of NO2. Environ. Int. 25, 3–8. doi:10.1016/s0160-4120(98)00092-0

Heal, M.R., O'Donoghue, M.A., Cape, J.N., 1999. Overestimation of urban nitrogen dioxide by passive diffusion tubes: a comparative exposure and model study. Atmos. Environ. 33, 513–524. doi:10.1016/S1352-2310(98)00290-8

Kirby, C., Fox, M., Waterhouse, J., Drye, T., 2000. Influence of environmental parameters on the accuracy of nitrogen dioxide passive diffusion tubes for ambient measurement. J. Environ. Monit. 3, 150–158. doi:10.1039/b007839p

Lewné, M., Cyrys, J., Meliefste, K., Hoek, G., Brauer, M., Fischer, P., Gehring, U., Heinrich, J., Brunekreef, B., Bellander, T., 2004. Spatial variation in nitrogen dioxide in three European areas. Sci. Total Environ. 332, 217–230. doi:10.1016/j.scitotenv.2004.04.014

Martin, N.A., Helmore, J.J., White, S., Barker Snook, I.L., Parish, A., Gates, L.S., 2014. Measurement of nitrogen dioxide diffusive sampling rates for Palmes diffusion tubes using a controlled atmosphere test facility (CATFAC). Atmos. Environ. 94, 529–537. doi:10.1016/j.atmosenv.2014.05.064

Mukerjee, S., Oliver, K.D., Seila, R.L., Jr, H.H.J., Croghan, C., Jr, E.H.D., Neas, L.M., Smith, L.A., 2008. Field comparison of passive air samplers with reference monitors for ambient volatile organic compounds and nitrogen dioxide under week-long integrals. J. Environ. Monit. 11, 220–227. doi:10.1039/B809588D

Ogawa, 2006. NO, NOs, NOx and SO2 Sampling Protocol Using the Ogawa Sampler [WWW Document]. URL http://ogawausa.com/wp-content/uploads/2014/04/prono-noxno2so206.pdf

Palmes, E.D., Gunnison, A.F., Dimattio, J., Tomczyk, C., 1976. Personal sampler for nitrogen dioxide. Am. Ind. Hyg. Assoc. J. 37, 570–577.

Pfeffer, U., Zang, T., Rumpf, E.M., Zang, S., 2010. Calibration of diffusive samplers for nitrogen dioxide using the reference method–Evaluation of measurement uncertainty. Gefahrstoffe Reinhalt. Luft 11, 500–506.

Plaisance, H., Plechocki-Minguy, A., Garcia-Fouque, S., Galloo, J.C., 2004. Influence of meteorological factors on the NO2 measurements by passive diffusion tube. Atmos. Environ. 38, 573–580. doi:10.1016/j.atmosenv.2003.09.073

Sather, M.E., Slonecker, Et., Mathew, J., Daughtrey, H., Williams, D.D., 2007. Evaluation of ogawa passive sampling devices as an alternative measurement method for the nitrogen dioxide annual standard in El Paso, Texas. Environ. Monit. Assess. 124, 211–221. doi:http://dx.doi.org/10.1007/s10661-006-9219-4

Sather, M.E., Terrence Slonecker, E., Kronmiller, K.G., Williams, D.D., Daughtrey, H., Mathew, J., 2006. Evaluation of short-term Ogawa passive, photolytic, and federal reference method sampling devices for nitrogen oxides in El Paso and Houston, Texas. J. Environ. Monit. 8, 558. doi:10.1039/b601113f

Tang, M.J., Cox, R.A., Kalberer, M., 2014. Compilation and evaluation of gas phase diffusion coefficients of reactive trace gases in the atmosphere: volume 1. Inorganic compounds. Atmospheric Chem. Phys. 14, 9233–9247. doi:10.5194/acp-14-9233-2014

Tang, Y.S., Cape, J.N., Sutton, M.A., 2001. Development and types of passive samplers for monitoring atmospheric NO2 and NH3 concentrations. ScientificWorldJournal 1, 513–529.

Targa, J., Loader, A., 2008. Diffusion Tubes for Ambient NO2 Monitoring: A Practical Guide (No. AEA/ENV/R/2504 – Issue 1a). AEA Energy and Environment.

van Reeuwijk, H., Fischer, P.H., Harssema, H., Briggs, D.J., Smallbone, K., Lebret, E., 1998. Field Comparison of two NO2 Passive Samplers to Assess Spatial Variation. Environ. Monit. Assess. 50, 37–51. doi:http://dx.doi.org/10.1023/A:1005703722232

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.atomsenv.2009.02.048

World Health Organization, 2013. Review of evidence on health aspects of air pollution – REVIHAAP project: final technical report.

Yu, C.H., Morandi, M.T., Weisel, C.P., 2008. Passive dosimeters for nitrogen dioxide in personal/indoor air sampling: A review. J. Expo. Sci. Environ. Epidemiol. 18, 441–451. doi:10.1038/jes.2008.22

Supplementary information:

Influence of wind-speed on short-duration NO_2 measurements using Palmes and Ogawa passive diffusion samplers

Nicola Masey[1], *Jonathan Gillespie*[1], *Mathew R. Heal*[2], *Scott Hamilton*[3], *Iain J. Beverland*[1*]

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]School of Chemistry, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

[3]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

*CORRESPONDING AUTHOR: Dr Iain J. Beverland, Department of Civil and Environmental Engineering, University of Strathclyde, 505F James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK; email: iain.beverland@strath.ac.uk; Tel: +44 141 548 3202

Table S1: Details of passive sampler deployments between February and October 2015.

Exposure number	Sampler exposure dates	Sampler exposure time (days)	Exposure details
1	25 Feb – 27 Feb 2015	2	Measurements made consecutively
2	27 Feb – 02 Mar 2015	3	
3	02 Mar – 04 Mar 2015	2	
4	04 Mar – 06 Mar 2015	2	
5	25 Mar – 27 Mar 2015	2	Measurements made consecutively. Ogawa[P] sampler not available for exposure number 8
6	27 Mar – 30 Mar 2015	3	
7	30 Mar – 1 Apr 2015	2	
8	1 Apr – 07 Apr 2015	6	
9	01 May – 7 May 2015	6	Measurements made consecutively
10	7 May – 15 May 2015	8	
11	15 May – 21 May 2015	6	
12	21 May – 28 May 2015	7	
13	6 Jul – 13 Jul 2015	7	Simultaneous measurements of 7-day and cumulative 2- and 3-day exposures of samplers
14	6 Jul – 8 Jul 2015	2	
15	8 Jul – 10 Jul 2015	2	
16	10 Jul – 13 Jul 2015	3	
17	27 Jul – 3 Aug 2015	7	Simultaneous measurements of 7-day and cumulative 2- and 3-day exposures of samplers
18	27 Jul – 29 Jul 2015	2	
19	29 Jul – 31 Jul 2015	2	
20	31 Jul – 3 Aug 2015	3	
21	10 Aug – 17 Aug 2015	7	Simultaneous measurements of 7-day and cumulative 2- and 3-day exposures of samplers
22	10 Aug – 12 Aug 2015	2	
23	12 Aug – 14 Aug 2015	2	
24	14 Aug – 17 Aug 2015	3	
25	14 Oct – 21 Oct 2015	7	Simultaneous measurements of 7-day and cumulative 2- and 3-day exposures of samplers
26	14 Oct – 16 Oct 2015	2	
27	16 Oct – 19 Oct 2015	3	
28	19 Oct – 21 Oct 2015	2	
29	26 Oct – 02 Nov 2015	7	Simultaneous measurements of 7-day and cumulative 2- and 3-day exposures of samplers
30	26 Oct – 28 Oct 2015	2	
31	28 Oct – 30 Oct 2015	2	
32	30 Oct – 2 Nov 2015	3	

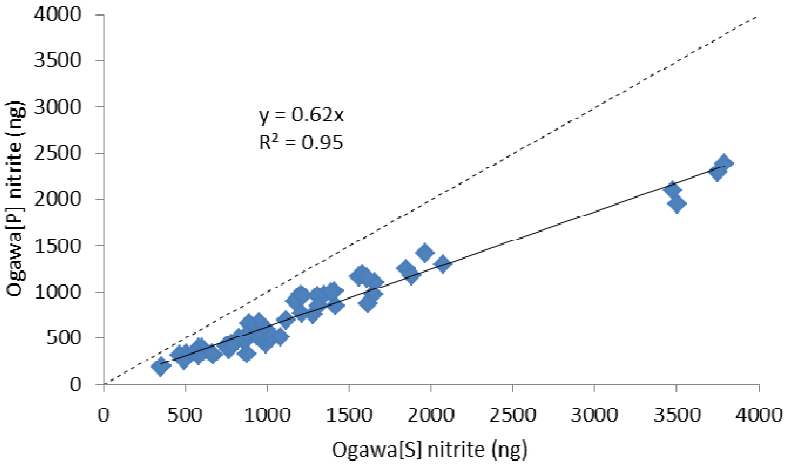


Figure S1: Relationship between mass of nitrite collected by the Ogawa[S] and Ogawa[P] samplers (95 % confidence intervals of slope = 0.60-0.64).

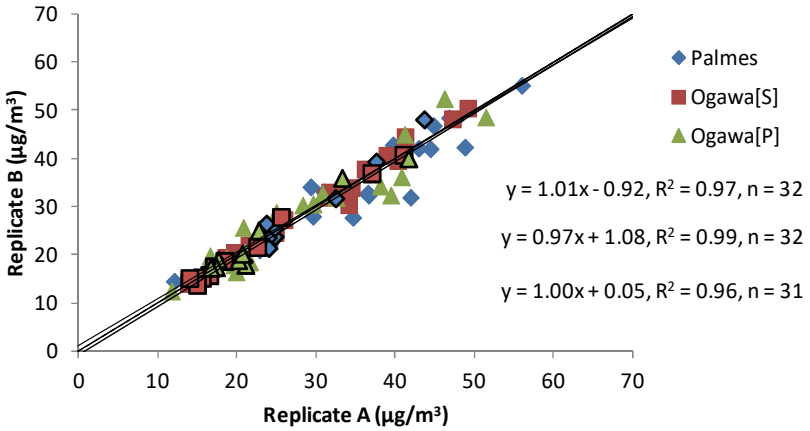


Figure S2: Scatter plot of duplicate Palmes, Ogawa[S] and Ogawa[P] sampler deployments. Symbols with a dark border are measurements from exposures of greater than 3 days. Confidence intervals (95 %) of the equations were: Palmes slope = 0.90-1.12 and intercept = -2.98 – 3.90; Ogawa[S] slope = 0.93-1.01 and intercept = -0.64-1.64; and Ogawa[P] slope = 0.89-1.11 and intercept = -3.02-2.76.

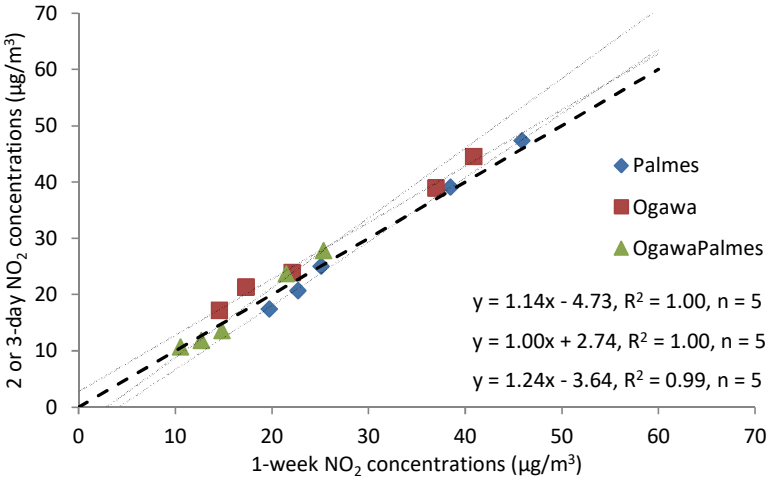


Figure S3: Linear regression between the NO_2 concentrations derived from simultaneous 7-day and cumulative 2 & 3-day exposures. The 95 % confidence intervals for the regressions were: Palmes slope = 1.03 – 1.24 and intercept = -8.08 - -1.27; Ogawa slope = 0.85 – 1.15 and intercept = -1.47 – 7.13; and OgawaPalmes slope = 0.96 – 1.51 and intercept = -8.39 – 1.42.

Calculation of Ogawa[S] uptake rate using the manufacturer supplied concentration conversion coefficient in the Ogawa sampler protocol:

Sampler uptake rate can be defined (from equation (1) in this paper) as:

$$UR\ (cm^3\ min^{-1}) = \frac{nitrite\ mass\ (ng)}{\left([NO_2]\ (ng\ m^{-3}) \times t\ (min)\right)} \times 10^6\ (cm^3\ m^{-3}) \qquad (S1)$$

$[NO_2]$ concentration is defined in volumetric units in the Ogawa analytical protocol (http://ogawausa.com/protocols/) as:

$$[NO_2]\ (ppb) = \frac{\alpha_{NO_2}\ (ppb\ min\ \ ng^{-1}) \times nitrite\ mass\ (ng)}{t\ (min)} \qquad (S2)$$

where α_{NO2} is the concentration conversion coefficient provided in the protocol, and *t* is the duration of the exposure.

A standard conversion factor (CF) can be derived from the gas laws to convert concentrations from volumetric to gravimetric units, hence:

$$[NO_2]\ (ng\ m^{-3}) = \frac{CF\ (ng\ m^{-3}\ ppb^{-1}) \times \alpha_{NO_2}\ (ppb\ min\ ng^{-1}) \times nitrite\ mass\ (ng)}{t\ (min)} \qquad (S3)$$

Substituting $[NO_2]$ in equation (S1) with the expression in equation (S3), gives:

$$UR\ (cm^3\ min^{-1}) = \frac{1}{\left(CF\ (ng\ m^{-3}\ ppb^{-1}) \times \alpha_{NO_2}\ (ppb\ min\ \ ng^{-1})\right)} \times 10^6\ (cm^3\ m^{-3}) \qquad (S4)$$

Using a volumetric to gravimetric conversion factor (*CF*) = 1910 ng m^{-3} ppb^{-1} and a concentration conversion coefficient (αNO_2) = 56 ppb min ng^{-1} (from Ogawa analytical protocol for 20 ºC and 70 % RH) gives *UR(calculated from Ogawa protocol)* = 9.3 cm^3 min^{-1}.

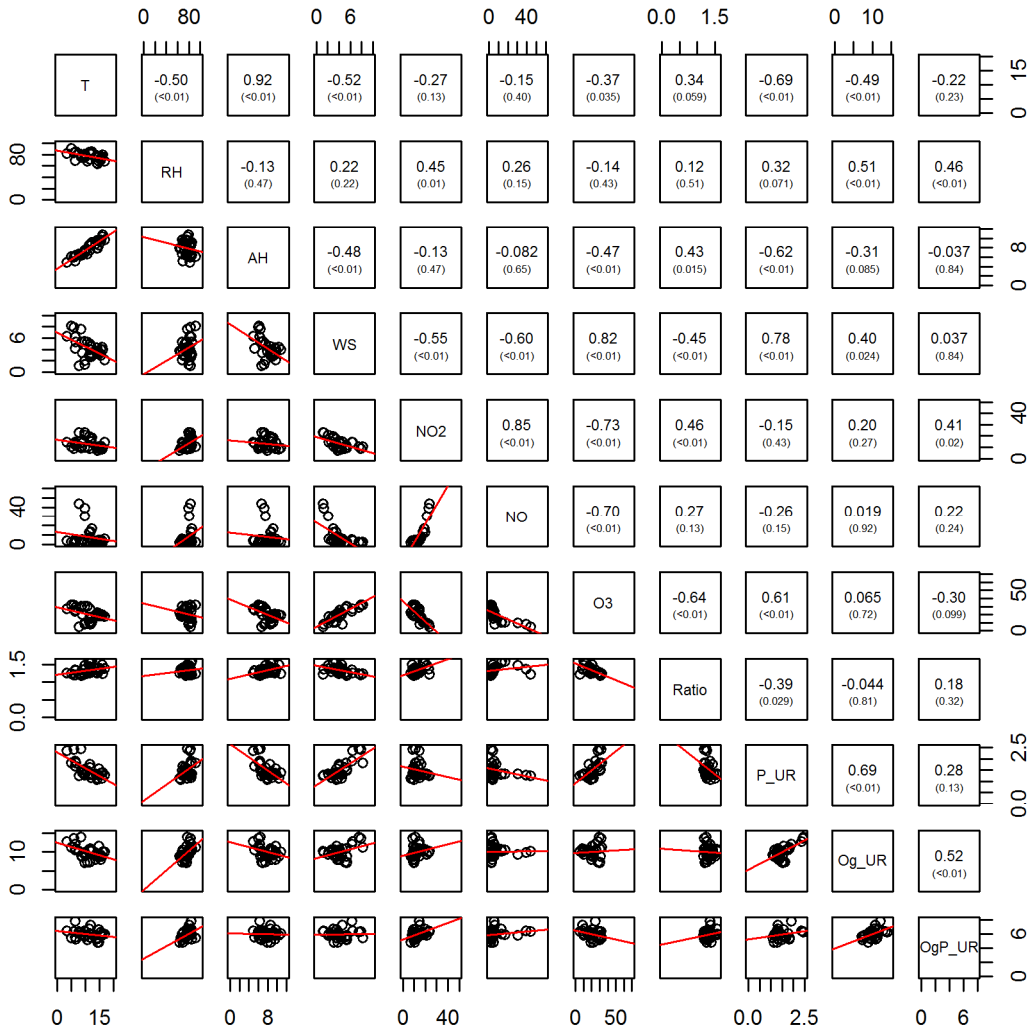


Figure S4: Pearson correlation coefficient matrix for bivariate relationships between observed meteorological and atmospheric composition variables and empirical passive sampler uptake rates. Variables are exposure means of: air temperature (*T*, °C); relative humidity (RH, %); absolute humidity (AH, g/m^3); wind-speed (WS, m/s); analyser NO_2 (NO_2, ppb), analyser NO (NO, ppb), analyser O_3 (O_3, ppb); ratio of analyser NO_2 plus minimum of analyser NO or O_3, to analyser NO_2 (Ratio(Min(NO,O_3)+NO_2)/NO_2 – abbreviated to Ratio); uptake rates for Palmes (P_UR), Ogawa[S] (Og_UR) and Ogawa[P] (OgP_UR) samplers. The *P*-value for the correlation is shown in brackets below the correlation coefficient.

Minimising the impact of wind-speed effects on NO_2 passive diffusion samplers through sampler modifications

Nicola Masey[1], Fiona Sutherland[1], Samuel Grainger[1], Scott Hamilton[2], Mathew R. Heal[3], Iain J. Beverland[1]*

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

[3]School of Chemistry, Joseph Black Building, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

*CORRESPONDING AUTHOR: Dr Iain J. Beverland, email: iain.beverland@strath.ac.uk; Tel: +44 141 548 3202

Research Highlights:

- Palmes NO_2 tubes and modifications tested at two reference analyser sites
- Membranes over tube entrances and post-processing for wind-speed effects tested
- Post-processing for wind-speed improved sampler accuracy (NMB 0.27 *vs.* 0.13)
- Inclusion of membrane most effective (NMB = 0.17 *vs.* 0.00)
- Palmes samplers should be deployed with membranes during field trials

Abstract

Passive samplers measuring NO_2 were deployed at an urban background (Townhead) and roadside (High Street) reference analyser sites in Glasgow, UK, for 30 discrete periods. Samplers were deployed for 2-3 days on each occasion, and concurrent meteorological measurements were made at Townhead during each study. We compared concentrations measured by the Ogawa sampler and a Palmes sampler, Palmes sampler with a membrane over the open end to minimise wind-speed effects and a half-length Palmes sampler including membrane (anticipated to minimise both wind-speed effects and any impact of within-tube chemistry). The Palmes sampler overestimated concentrations at both sites, however greater overestimations were observed at the most exposed Townhead site. We tested a post-processing correction method to take wind-speed effects on uptake rates into account, which was found to improve the agreement between the standard Palmes sampler and reference analyser NO_2 concentrations (NMB = 0.27 *vs.* 0.13 (Townhead) and 0.08 *vs.* -0.06 (High Street)). Inclusion of a membrane over the Palmes tube entrance improved the

sampler accuracy (NMB = 0.17 *vs.* 0.00 for both sites data combined), and removed the significant impact of wind-speed on the Palmes sampler. The short Palmes sampler suffered from significant wind-speed effects (anticipated to be due to the shorter sampler path length) however still improved upon the standard Palmes sampler (NMB = 0.17 *vs.* -0.05 for both sites). The Ogawa sampler provided accurate estimates of the reference analyser concentrations (NMB = -0.03 for both sites). We suggest that Palmes samplers should be deployed with a membrane over their exposed entrance to minimise the effects of wind-speed and consequently improve the sampler ability to estimate ambient NO_2 concentrations.

Keywords: Air pollution; Palmes; Ogawa; uptake rate; passive sampling; wind-speed

1. Introduction

Passive samplers have been widely used to measure nitrogen dioxide (NO_2) over large study areas and numbers of sites, for example during epidemiology studies (Cyrys et al., 2012; Gillespie et al., 2017, 2016; Lewné et al., 2004) and for indoor or personal monitoring (Yu et al., 2008). The low-cost, small-size and simple deployment of these samplers make them well suited for deployments in large numbers and study areas. However passive samplers are only able to provide time-averaged concentrations over the exposure period and previous studies have highlighted factors that can influence sampler accuracy and precision (Cape, 2009).

The main factors thought to influence passive sampler accuracy are wind-speed effects and within-tube chemistry. Reactions within the diffusion path of the samplers can lead to the formation of additional NO_2, which is not present in ambient conditions, which can lead to samplers overestimating ambient concentrations (Heal et al., 2000, 1999; Kirby et al., 2000):

$$NO + O_3 \rightarrow NO_2$$

In the atmosphere the reaction above normally occurs at the same time as NO_2 is broken down in UV light to NO and O, however when the reaction above occurs in the passive sampler the NO_2 generated cannot be broken down as UV light cannot penetrate through the tube material. This leads to increased concentrations of NO_2 in the sampler in comparison to those in the ambient air outside the tube, and consequently sampler overestimation.

This effect is present in samplers with longer path lengths, for example Palmes passive diffusion tubes, while shorter path length badge samplers (such as the Ogawa sampler) do not suffer from these effects due to shorter residence times of air within the samplers (Brown, 2000). Wind-speed effects, which shorten the effective path length of the samplers

through generation of turbulence at the entrance to the sampler, can also lead to passive sampler overestimation. Previous studies have tried to minimise the impact of wind-speed on passive samplers through use of protective shelters (Gerboles et al., 2006; Plaisance et al., 2004), membranes covering the entrance to the tube (Gerboles et al., 2005; Martin et al., 2014) and correction of sampler uptake-rates post deployment (Masey et al., 2017). Shelters have reduced the bias in concentrations measured by Palmes diffusion tubes, however these are conspicuous (limiting their use in field studies due to making them more prone to theft) and have previously been suggested to have a different composition of air under the shelter compared to ambient concentrations (Kirby et al., 2000) therefore these will not be discussed further. Post-processing sampler uptake rates for wind-speed effects has advantages in that it is a simple approach requiring no modifications to the sampler design and this can be back-extrapolated; however this requires information about wind-speed at each measurement site (which is unlikely to be available for large numbers of sites) or relies on the use of a single meteorological site which may not be representative of all locations within the study area (Masey et al., 2017). The inclusion of a membrane over the tube entrance is simple in practice and removes the requirement for detailed wind-speed information for each site, however it is not possible to determine how much the membrane reduces the uptake rate. Consequently this cannot be calculated from first principles and hence calibrations are required before these modified samplers can be used in the field.

This work compares Palmes samplers, Palmes samplers modified using membranes and post-processing Palmes uptake rates to account for wind-speed effects at an urban background and a roadside reference analyser site. We also deployed shorter Palmes samplers and Ogawa samplers at the site to identify any impact of within-tube chemistry. The impact of using meteorological data collected from the nearest weather station and on-site during the sampler investigation was also investigated. A large proportion of the published literature have carried out investigations under controlled laboratory conditions (Buzica et al., 2005; Martin et al., 2014; Plaisance et al., 2004) or have used wind-speed data collected at meteorological stations distant to the measurement site (Hagenbjörk-Gustafsson et al., 2010; Masey et al., 2017; Vardoulakis et al., 2009). Both of these are not necessarily representative of the true conditions experienced by a passive sampler at the study location, and little work has investigated sampler bias using meteorology collected on-site.

2. Methods

Passive diffusion samplers were deployed at two automatic monitoring stations in Glasgow, UK, for 30 discrete deployment periods. The automatic monitoring stations were part of the UK Automatic Urban and Rural Network (AURN) and reference analysers at these sites undergo regular detailed quality assurance and quality control procedures (DEFRA, 2017). Townhead (latitude 55.866°, longitude -4.23°) is an urban background station located in central Glasgow 122 m from the nearest road. High Street (latitude 55.861°, longitude -4.238°) is a roadside station located 5.5 m from the kerb and ~640 m South-East of Townhead. Concentrations of NO_2 were measured at both sites using a Teledyne-API 200A NO_x chemiluminescent analyser and concentrations of O_3 were measured (at Townhead only) using a Thermo Scientific Model 49i O_3 analyser. Study-average concentrations for the site were calculated from the hourly-average concentrations, available to download from www.scottishairquality.co.uk. The Townhead station has a large roof space on which instruments can be securely deployed, while High Street is a smaller site with limited space for affixing monitoring instruments on the cages protecting the analyser inlets.

Hourly wind-speed measurements were available to download from https://mesonet.agron.iastate.edu/request/download.phtml?network=GB__ASOS for Glasgow Airport, the closest weather station to the sites. Five minute temperature and relative humidity measurements were recorded at Townhead using an Onset HOBO U23 Pro V2 External data logger deployed on the railing at Townhead under a solar radiation shield. We measured wind-speed every minute at Townhead using a WindSonic anemometer (Gill Instruments Ltd, Hampshire, UK) located on the railing. These meteorological measurements could not be made at High Street due to the limited space and unsecured access to the site. Study-average meteorological values were calculated for each deployment period to allow assessment of the variables in relation to passive sampler accuracy.

A Kestrel® 5500 Weather Meter was deployed at each site for the duration of passive sampler deployment periods to provide a snapshot of the meteorological conditions (including temperature, relative humidity and wind-speed) at the site. In practice, this gave approximately 10-15 minutes of 1-second measurements at each site for each site visit. This provided a brief comparison of the conditions at Townhead and High Street to allow estimates of longer duration trends at the sites to be made.

(a) (b)

Figure 1: Deployment of the passive diffusion samplers at (a) Townhead and (b) High Street. Townhead samplers were mounted on a railing on the roof of the site located ~1 m from the analyser inlets, while High Street samplers were located on the cage on the roof of the site which also contained the analyser inlets.

Four designs of passive samplers were deployed concurrently at these sites for 2-3 day exposures, with Townhead visited first on each study date (Table S1). The samplers were located on a railing on the roof of the Townhead AURN station, approximately 1 m from the analyser inlets (Figure 1a). At High Street the samplers were deployed on the cage housing the analyser inlets on the roof of the monitoring station (Figure 1b). At each site the samplers were deployed in duplicate. For studies 21 and 29 we were unable to access the roof of High Street to retrieve the samplers due to ladders being unavailable, therefore no data was collected at High Street for these studies and consequently studies 20 and 28 samplers were deployed for two consecutive deployment periods.

The designs of samplers tested were: standard Palmes diffusion tubes (referred to as Palmes); Palmes diffusion tubes with an amorphous polyethylene membrane covering the open end of the tube (referred to as Palmes[M]); a half-length diffusion tube (traditionally used to measure Ammonia), which was slightly opaque, deployed with a membrane over the end (referred to as Palmes[S]); and a modified Ogawa sampler using Palmes meshes as the

collection medium instead of the commercially available Ogawa collection pads (referred to as Ogawa[P]). The variations on the Palmes tube were purchased from Gradko International Limited (www.gradko.com) while the Ogawa sampler parts were purchased from Ogawa & Company USA (www.ogawausa.com). The Ogawa[P] sampler was deployed under a weather shelter, available from Ogawa USA, to minimise wind-induced turbulence. Those samplers utilising a membrane over the tube entrance were assembled on-site to ensure the exposure time of the tube was similar to the standard Palmes sampler – sealed tubes were taken to the site and, at the time of deployment, the tubes were uncapped and the membrane fitted (care was taken not to touch the membrane by hand to prevent transfer of any nitrite present on the technicians hands to the membrane).

The passive samplers were prepared in-house at the University of Strathclyde, with one week of samplers prepared at a time. For each week of samplers, two laboratory blanks were also prepared to act as control samples that would not be exposed. Between preparation and deployment, and retrieval and analysis, the samplers were stored in sealed double-bags in the fridge. The meshes used to collected ambient nitrite were prepared using the 1:1 TEA:acetone dipping method, with two coated meshes assigned to each sampler (Heal, 2008). The analysis of the samplers was also carried out in-house, using the Saltzman-Greiss reagent and UV analysis for all four sampler types (Targa and Loader, 2008). Laboratory analysis of the samplers was carried out each week, and the laboratory blank prepared each week was also analysed during the procedure. The mass of nitrite collected by the blank samplers was subtracted from that collected by the exposed sampler prior to calculation of ambient NO_2 concentrations.

The tubes and membranes were reusable – after each use the tubes and caps were soaked in Deacon solution overnight and then were covered and left to air dry. The membranes and grids were put in Deacon solution and agitated in an ultra-sonic bath for 30 minutes. The membranes were then left to air-dry with the tubes and caps, while the meshes were dried in an oven overnight.

The ambient concentration of NO_2 was calculated using the method in Targa and Loader (2008) for the Palmes samplers and using the Ogawa protocol (2006) for the Ogawa[P] samplers. The concentrations measured by the Palmes[M] and Palmes[S] could not be calculated empirically due to a lack of information about the diffusion area and associated path length when an amorphous polyethylene membrane was used (Martin et al., 2014). The uptake rate of each of the samplers was empirically calculated from the mass of nitrite collected (Equation 1), which was used to estimate the pollution concentrations measured by the Palmes[M] and Palmes[S] samplers. Additionally, linear regression was used to

investigate the relationship between uptake rates and the meteorological and chemical conditions at the site.

$$UR\ (cm^3\ min^{-1}) = \left[\frac{nitrite\ mass\ (ng)}{([NO_2]\ (ng\ m^{-3}) \times t\ (min))}\right] \times 10^6\ (cm^3\ m^{-3}) \quad (1)$$

Where $[NO_2]$ is the analyser exposure-averaged concentrations and *t* is the exposure time.

3. Results & Discussion

3.1. Precision and limit of detection

The limit of detection (LoD) of the analysis procedure was calculated using three times the standard deviation of the blank concentrations, plus the blank concentration for the shortest exposure time in order to obtain the worst case scenario LoD. The study-average limit of detection for each passive sampler was 7.4 µg/m^3, 17.9 µg/m^3, 13.9 µg/m^3 and 3.3 µg/m^3 for Palmes, Palmes[M], Palmes[S] and Ogawa[P] respectively (Table 1). The reasons for the large difference between the LoD of the Palmes and Palmes[M] samplers is unclear as both tubes were stored and analysed similarly. As stated above, the mass of nitrite calculated for the blank samplers was subtracted from the exposed sampler nitrite mass. The Palmes, Palmes[S], Ogawa[P], and the majority of the Palmes[M] sampler measured concentrations for all studies were above the LoD.

The relative standard deviation (RSD) for duplicate measurements was below 15 % for all samplers, with the Ogawa[P] samplers having the most similar duplicate concentrations (8.1 %) and the Palmes[M] samplers showing the largest variation between duplicates (14.7 %) (Table 1). The Palmes RSD values are slightly poorer than reported from our previous Palmes study (Masey et al., 2017) however these still fall within the range of RSD values reported for weekly exposed Palmes samplers (Bush et al., 2001; Buzica et al., 2005; Heal et al., 2000, 1999, 1999a; Kirby et al., 2000; van Reeuwijk et al., 1998; Vardoulakis et al., 2009). We observed that the membranes over the end of the tubes did not, in all instances, fit snugly over the exposed end and could lead to slight movement in the membrane, which could be a reason for the higher RSD values for the tube samplers with the membrane. In the remainder of this paper the mean concentration from the duplicate samplers for each deployment is reported.

Table 1: Limits of detection (LoD) and duplicate relative standard deviation (RSD) statistics for the passive samplers and (a) their root mean square error (RMSE, in µg/m^3), normalised mean bias (NMB) and coefficient of determination (R^2) with respect to the analyser measurements at Townhead (Th) and High Street (HS) stations.

	Palmes	Palmes[M]	Palmes[S]	Ogawa[P]
LoD All (µg/m^3)	7.4	17.9	13.9	3.3
Mean duplicate RSD (Th) (%)	11.7	14.7	14.2	8.1
Mean duplicate RSD (HS) (%)	10.4	11.9	9.8	8.1
Data using standard analysis methods:				
RMSE sampler *vs.* analyser (Th)	13.6	5.8	11.9	10.6
RMSE sampler *vs.* analyser (HS)	10.6	7.9	12.0	12.5
RMSE sampler *vs.* analyser (All)	12.3	6.9	12.0	11.5
NMB sampler *vs.* analyser (Th)	0.27	0.02	0.04	0.01
NMB sampler *vs.* analyser (HS)	0.08	-0.01	-0.12	-0.06
NMB sampler *vs.* analyser (All)	0.17	0.00	-0.05	-0.03
Regression slope	1.22	1.03	0.97	0.94
[95 % C.I.] (Th)	[1.13, 1.32]	[0.97, 1.08]	[0.86, 1.08]	[0.84, 1.04]
R^2 sampler *vs.* analyser (Th)	0.57	0.89	0.12	-0.03
Regression slope	1.03	0.97	0.85	0.92
[95 % C.I.] (HS)	[0.94, 1.11]	[0.91, 1.04]	[0.78, 0.93]	[0.83, 1.02]
R^2 sampler *vs.* analyser (HS)	0.36	0.78	0.51	0.52

3.2. Sampler accuracy

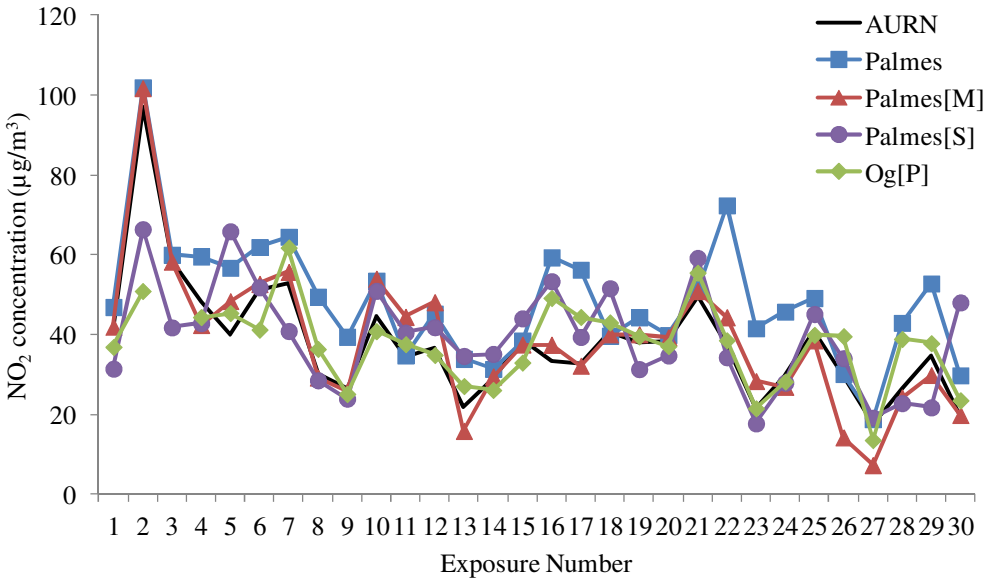
The concentrations for the Palmes and Ogawa[P] samplers were calculated as described in Section 2, while the concentrations for the Palmes[M] and Palmes[S] samplers were calculated using the study-average uptake rates for each site (Townhead Palmes[M] = 0.92 cm^3/min, Palmes[S] = 1.46 cm^3/min; High Street Palmes[M] = 0.88 cm^3/min, Palmes[S] = 1.20 cm^3/min). Our calculated uptake rate is lower than that estimated for Palmes[M] under controlled laboratory conditions (1.11 cm^3/min) (Martin et al., 2014), highlighting the importance of calibrating the samplers under similar conditions to which they will be exposed in the field. The higher uptake rates by the Palmes[S] sampler compared to the Palmes[M] sampler is not unexpected as the shorter tube will reduce the path length and hence increase the uptake rate. The Palmes[M] sampler measured more consistent nitrite masses vs. reference analyser NO_2 concentrations than the Palmes sampler (as represented by the higher R^2 value) (Figure S2).

The concentrations measured by the passive samplers generally followed the temporal trends in analyser concentrations at both sites (Figure 2). Scatter plots showing the correlation between study-average NO_2 concentrations measured by the passive samplers and reference analyser at each site were generated (Figure 3). The regression lines in Figure 3 were forced through the origin following the reasoning of Martin et al. (2014) and Pfeffer et

al. (2010) of insignificant laboratory blanks and incomplete temporal measurements by the analyser (as a result of e.g. maintenance or calibration activities). Despite the difference in the site classification (Townhead = urban background, High Street = roadside) the concentrations measured at the sites were more similar than anticipated (Figure S1), however the sampler accuracies at the two sites were different. Both sites measured an elevated concentration (approximately 100 $\mu g/m^3$ NO_2) during study 2, which coincided with a period of low temperatures (study-average temperature measured at Townhead = -1.2 °C) and low wind-speeds (study-average wind-speed measured at Townhead = 0.39 m/s). During this study the diffusion tubes at Townhead had a visual layer of frost present, however we have retained these measurements in the analysis to test the extreme conditions.

The Palmes sampler overestimated NO_2 concentrations measured at both sites, however this was much more pronounced for Townhead (slope = 1.22 *vs.* 1.03 for Townhead and High Street respectively) (Figures 2 and 3). The overestimation of NO_2 concentrations by the Palmes samplers was consistent with published literature (reviewed in Section 1), with similar overestimations reported by Kirby et al. (2000) and Gerboles et al. (2005, 2006). The regression line and correlation coefficient between the Palmes measurements at Townhead were similar between this and our previous studies (slope = 1.22 *vs.* 1.25, R^2 = 0.57 *vs.* 0.59 (current *vs.* previous study)) despite the different study (with slightly different exposure times) and different season (Masey et al., 2017). The regression slopes for Palmes[M] and Palmes[S] were closer to one than the standard Palmes sampler at Townhead, however for High Street the Palmes[S] sampler regression slope was further from one (Figure 3). The Ogawa[P] sampler produced the most similar regression equations between the two sites (slope = 0.94 and 0.92 at Townhead and High Street). These slopes were closer to one than reported for a previous study at the same site (slope = 0.67) (Masey et al., 2017). The root mean square error (RMSE) and normalised mean bias (NMB) values were smallest for the Palmes[M] sampler and largest for the Palmes sampler (NMB = 0.00, -0.03, -0.05 and 0.17 for the Palmes[M], Ogawa[P], Palmes[S] and Palmes samplers respectively) (Table 1).

(a)



(b)

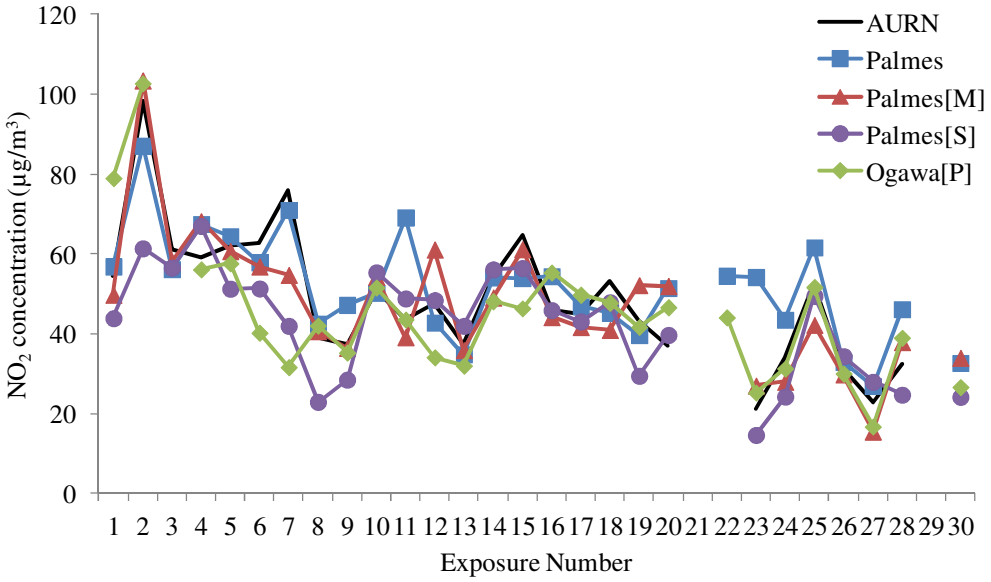
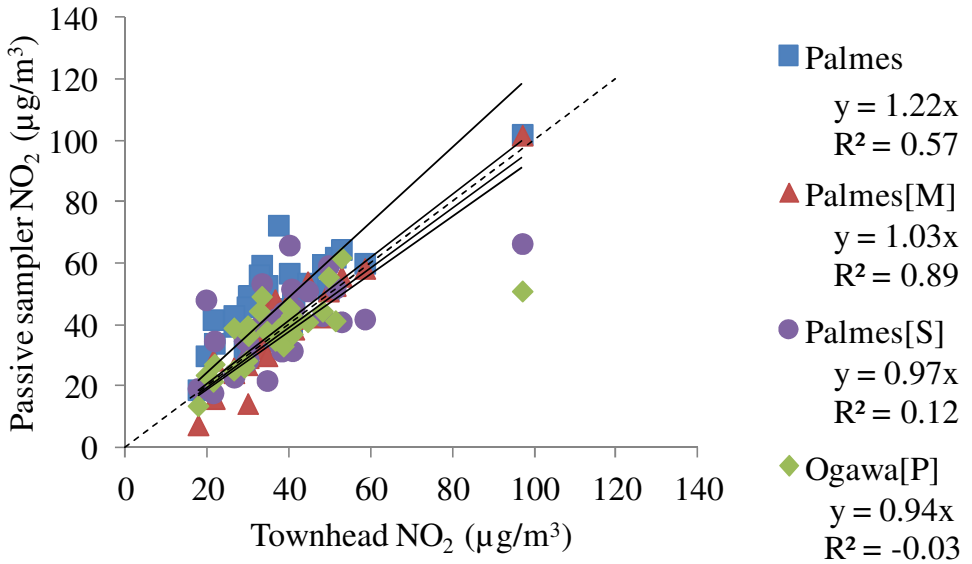


Figure 2: Time series plots showing exposure-averaged NO_2 concentrations measured by each of the samplers and the automatic analyser located at (a) Townhead and (b) High Street. The Palmes and Ogawa[P] concentrations were calculated empirically, while the Palmes[M] and Palmes[S] concentrations were calculated using the study-average uptake rates at each site.

(a)



(b)

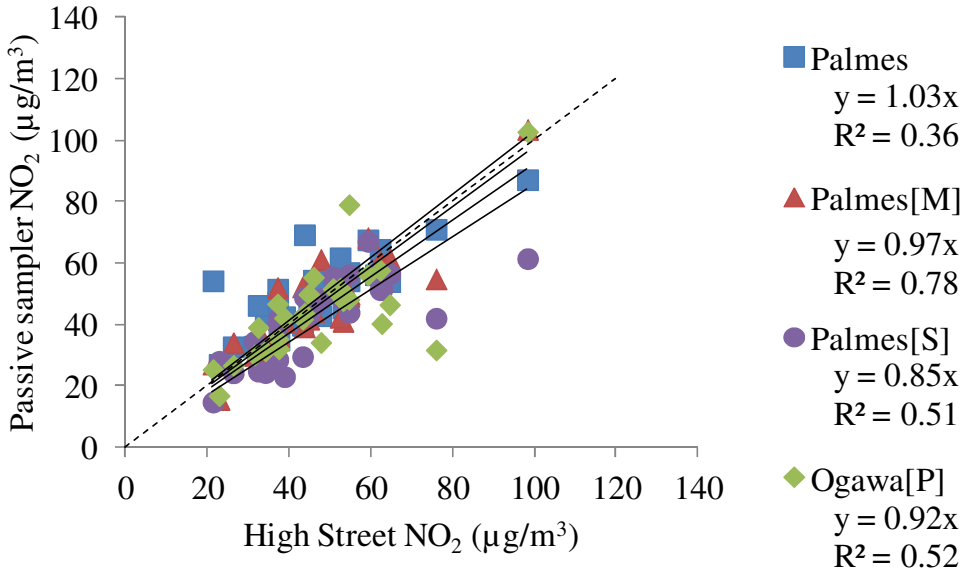


Figure 3: Scatter plots showing regression between passive samplers and analyser exposure-averaged NO_2 concentrations at (a) Townhead and (b) High Street. Palmes and Ogawa[P] were calculated empirically while the Palmes[M] and Palmes[S] samplers were calculated using the study average uptake rates.

3.3. Comparison of on-site and meteorological station wind-speed measurements

The hourly-average wind-speed measurements at Townhead and at the meteorological station located at Glasgow Airport (over the whole study period) were highly correlated, with the Airport measurements being higher as a result of the site being more exposed (Townhead wind-speed = 0.27*Airport wind-speed + 0.26, R^2 = 0.68, n = 1681 h) (Figure 4a). The study-average wind-speed measurements at the two sites were also highly

correlated with similar regression coefficients (Townhead wind-speed = 0.30*Airport wind-speed + 0.15, R^2 = 0.80, n = 30) (Figure 4b). In our previous study we utilised wind-speed measurements at the Airport to correct the uptake rate of the Palmes samplers on the assumption that the Airport data provided an indication of relative temporal trends in wind-speed at the site, and the observations in Figures 4a and 4b support this assumption.

The 'spot' measurements made during the visits to the sites to change the diffusion samplers were averaged to provide a study estimate of wind-speed by averaging the measurements made during sampler deployment and retrieval dates. The spot measurements had moderate correlation with the study-average wind-speeds measured at Glasgow Airport (R^2 = 0.45 and 0.42 for Townhead and High Street) with similar regression coefficients between the spot measurements and study-average Townhead wind-speed measurements (Townhead kestrel wind-speed = 0.24*Airport wind-speed - 0.04, n = 30) (Figure 4b). The spot measurements at High Street showed less variation in wind-speed over the study than Townhead, with wind-speeds at the former site approximately 40 % of the wind-speeds recorded at the latter (Figure 4c).

Previous studies (Hagenbjörk-Gustafsson et al., 2010; Masey et al., 2017; Vardoulakis et al., 2009) have used meteorological information obtained from the nearest station when evaluating passive sampler accuracy and bias factors. We have demonstrated that if meteorological information is collected at sites with similar urban conditions to the pollution site (such as Glasgow Airport and Townhead) then measurements from the meteorological station can be used to provide a reasonable estimate of the 'true' conditions at the pollution site. However, the use of a single weather station is not necessarily representative of all locations (such as High Street) and, if detailed meteorological information is required, weather measurements should be made on-site.

(a)

(b)

(c)

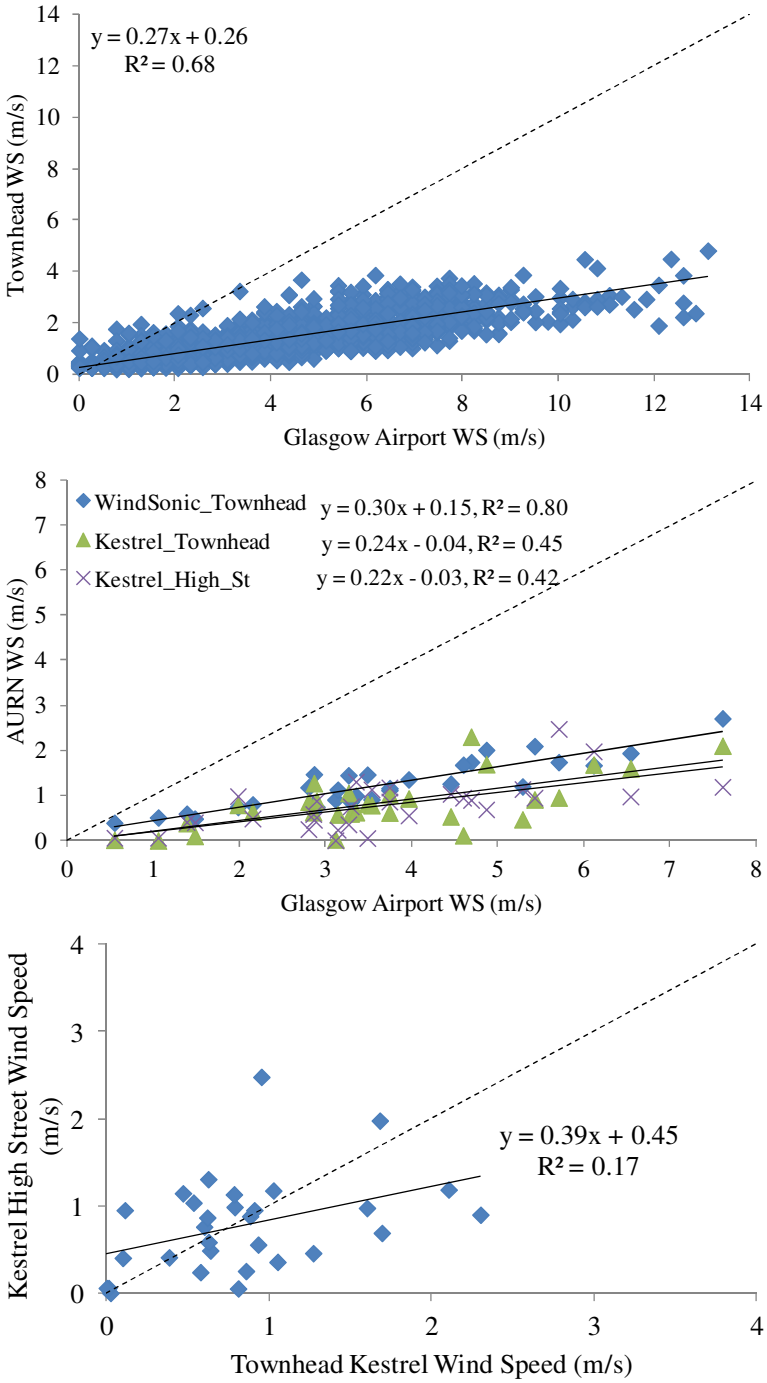


Figure 4: (a) Correlation between hourly-average wind-speed measurements made concurrently at Glasgow Airport and at Townhead for all deployment periods combined (n = 1681). (b) Correlation between wind-speed measurements made at Glasgow Airport and those measurements made by the wind sonic (Townhead) and the kestrel measurements at Townhead and High Street during the kestrel peripatetic measurement periods. Kestrel measurements are the average of those made at the start and end of deployment periods (n = 30). (c) correlation between wind-speed measurements made using the kestrel instrument at Townhead and High Street (average of those made at start and end of deployment periods).

3.4. Correction of Palmes uptake rates for wind-speed

In our previous study we proposed a method to post-process the concentrations measured by the Palmes sampler for the effects of wind-speed (Masey et al., 2017). We evaluated three correction methods following the method in our previous paper:

a) The equation relating Palmes uptake rate (UR) and wind-speed (measured at Glasgow Airport) from our previous publication was applied to the present study to test the transferability of the derived equation (referred to as '*UR_prev*').

$$UR\ (cm^3/min) = 0.16*Airport\ WS\ (m/s) + 0.83,\ R^2 = 0.60$$

b) Using the method from our previous paper, we derived new equations to relate wind-speed measured at Glasgow Airport to the measured uptake rates of the Palmes sampler at Townhead (referred to as '*UR_Airport*') (Figure 5a).

$$UR\ (cm^3/min) = 0.08*Airport\ WS\ (m/s) + 1.24,\ R^2 = 0.16$$

c) The method from the previous paper was used to derive a new equation between wind-speed measured at Townhead to the measured uptake rates of the Palmes sampler at Townhead (referred to as '*UR_Townhead*') (Figure 5b).

$$UR\ (cm^3/min) = 0.30*Townhead\ WS\ (m/s) + 1.17,\ R^2 = 0.22$$

The relationship between UR and wind-speed measured at the Airport was weaker than demonstrated during our previous study, which could be attributable to generally lower wind-speeds and fewer values above 6 m/s wind-speed, or the different meteorological conditions changing the correlations with other meteorological or chemical variables (discussed in Section 3.5 below).

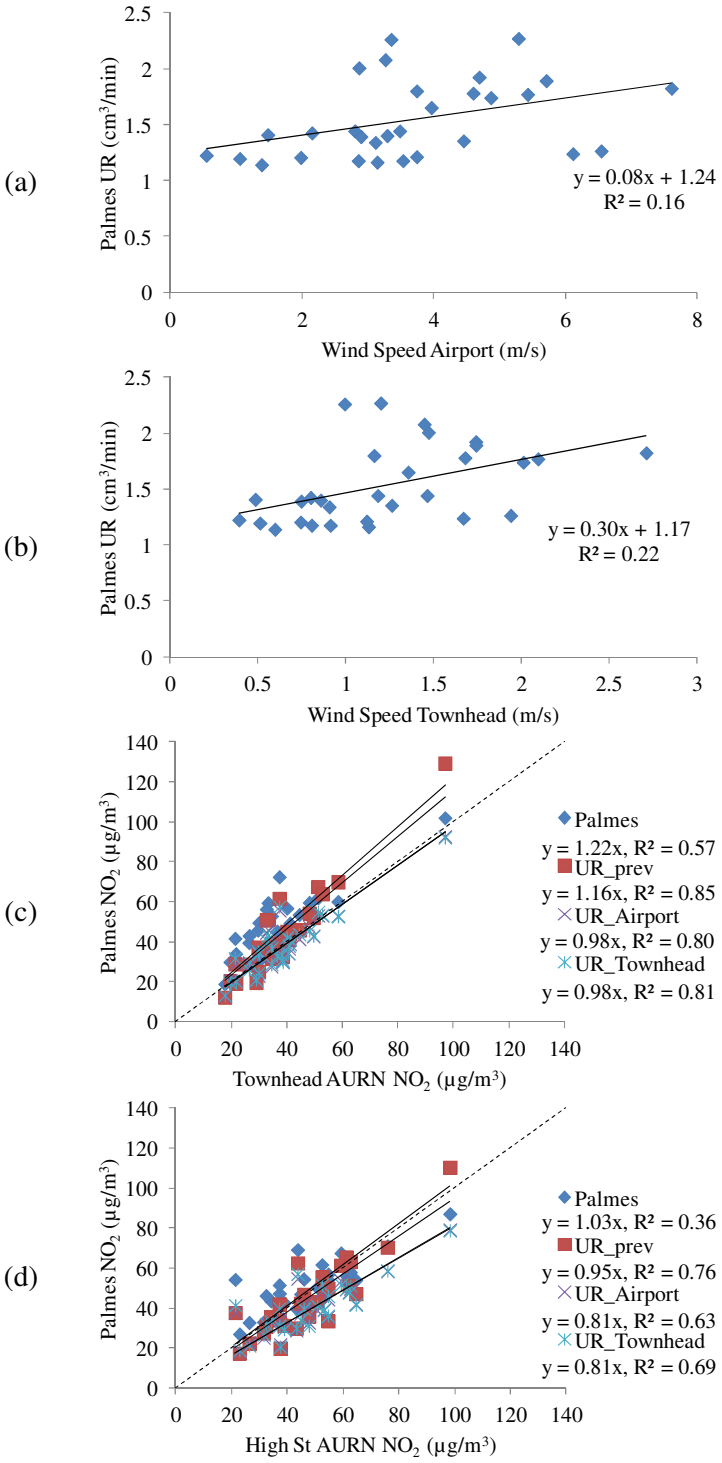


Figure 5: (a) UR *vs.* wind-speed (WS) measured at the Airport (Townhead data only); (b) UR *vs.* WS measured at Townhead (Townhead data only); and AURN *vs.* Palmes concentrations for standard Palmes, and Palmes UR corrected for WS (using Airport (*UR_Airport*) and Townhead (*UR_Townhead*) WS) and using the equation derived for Airport wind-speeds from our previous study (UR (cm^3/min) = 0.16*WS_Airport (m/s) + 0.83, R^2 = 0.60 (*UR_prev*)) applied to (c) Townhead and (d) High Street locations.

The three uptake-rate correction equations were applied to the Palmes measurements at Townhead (Figure 5c) and High Street (Figure 5d). The regression lines were forced through the origin for the reasons previously discussed. At Townhead, correcting the Palmes sampler uptake rate for the effects of wind-speed improved the regression slope and statistics, regardless of the correction equation applied. This suggests that the correction derived during our previous study is transferable temporally at this site, and can improve the accuracy of Palmes diffusion tubes. At High Street, however, only *UR_prev* slightly improved the RMSE and NMB values compared to the standard Palmes sampler, yet the standard Palmes sampler had the regression slope closest to one (Table 2). This suggests that the impact of wind-speed effects on the uptake rates is lower at High Street compared to Townhead. As discussed above, the difference in NO_2 concentration between the two sites is relatively small, leading us to suggest that a change in the meteorological conditions between the sites is the most likely cause of the difference. Based on the spot measurements made at the two sites we believe that wind-speed is the most likely cause of the difference (Figure 4c and Figure S3). These results show the method presented in our previous paper is transferrable to different study areas providing a new regression between wind-speed and uptake rate is calculated based on the typical wind-speeds experienced at the site. In reality this may mean multiple UR vs. wind-speed equations if measurements are to be made at exposed sites and those within streets for example. Further work could aim to make more measurements at a wider range of concentrations and wind-speeds to determine if a single equation could be provided for all realistic wind-speeds or identify the wind-speed at which changes in this relationship occurs.

Table 2: Regression slope (and 95 % confidence interval) and statistics describing the relationship between NO_2 concentrations measured by the analyser and the Palmes samplers after correction for wind-speed effects at (a) Townhead and (b) High Street. Corrections made for wind-speed using equation from previous study (UR_prev) (Masey et al., 2017); equation derived from Airport wind-speed (UR_Airport); and equation derived from Townhead wind-speed (UR_Townhead). Statistics provided are: Root mean square error (RMSE), coefficient of determination (R^2) and normalised mean bias (NMB).

	Palmes	Palmes (UR_prev)	Palmes (UR_Airport)	Palmes (UR_Townhead)
Measured WS range (m/s)	-	0.5 - 8	0.5 - 8	0.5 – 3
(a) Townhead				
Regression slope	1.22	1.16	0.98	0.98
[95 % C.I.]	[1.13, 1.32]	[1.08, 1.24]	[0.92, 1.04]	[0.92, 1.04]
R^2 sampler *vs.* analyser	0.57	0.85	0.80	0.81
RMSE sampler *vs.* analyser (µg/m^3)	13.60	10.53	6.65	6.41
NMB	0.27	0.13	-0.01	-0.01
(b) High Street				
Regression slope	1.03	0.95	0.81	0.81
[95 % C.I.]	[0.94, 1.11]	[0.87, 1.02]	[0.75, 0.87]	[0.75, 0.88]
R^2 sampler *vs.* analyser	0.36	0.76	0.63	0.69
RMSE sampler *vs.* analyser (µg/m^3)	10.61	9.79	11.95	12.37
NMB	0.08	-0.06	-0.17	-0.17

3.5. Uptake rates relationship with chemical and meteorological variables

The average uptake rate for the Palmes, Palmes[M], Palmes[S] and Ogawa[P] samplers were 1.4 cm^3/min, 0.90 cm^3/min, 1.3 cm^3/min and 5.7 cm^3/min respectively. The order of the uptake rates was as anticipated with the Ogawa[P] badge sampler, which has the shortest path length, having the fastest uptake rate, and the Palmes[M] sampler uptake rate being lower than the open Palmes tube. The Palmes[S] sampler uptake rate is similar to the open Palmes tube as the shorter path length reduces the impact of the reduced surface area as a result of inclusion of the membrane.

We investigated the relationship between the uptake rates calculated for all four samplers, and the chemical and meteorological conditions measured at the sites (Table 3). The correlation between Airport wind-speed and Palmes UR was lower than was observed in our previous study (Masey et al., 2017)(this study r = 0.40, previous study r = 0.78 for Townhead), while the correlation with temperature was larger in this study and in the opposite direction (this study r = 0.62, previous study r = -0.69 for Townhead) (Table 3). The strong confounding between temperature and wind-speed was present in both our

studies, however in the current study lower wind-speeds coincided with low temperatures (r = 0.40, $P < 0.05$) (Figure S4), while previously the opposite was true (r = -0.52, $P < 0.01$) (Masey et al., 2017). The previous study was carried out during spring/summer (February to November) while the current work was carried out November to February and this difference in seasonality of the work could be the cause of the difference in temperature - wind-speed relationship observed. The Palmes[M] sampler UR were not correlated with wind-speed at Townhead, suggesting that the membrane provides an effective barrier to turbulent eddy formulation and consequently the impact of wind-speed on uptake rate is minimised. This is consistent with findings from a chamber study published by Martin et al. (2014) for longer (28 day) exposure periods. The Palmes[S] sampler shows a significant relationship with wind-speeds measured at Townhead (r = 0.42, $P < 0.05$) however not with wind-speed measured at the Airport. The greater relationship between UR and wind-speed in the Palmes[S] compared to the Palmes[M] sampler could arise from the shorter path length associated with the Palmes[S] sampler which makes these more susceptible to turbulent effects. The Ogawa[P] UR showed significant correlations with temperature, absolute humidity (r = 0.52 and 0.55 respectively, $P < 0.01$) and NO_2 (r = 0.55, $P < 0.05$) at Townhead. Our previous study found no impact of temperature on the Ogawa[P] samplers (Masey et al., 2017), however low temperatures (and low absolute humidities) have previously been shown to impact Ogawa sampler uptake rate (Hagenbjörk-Gustafsson et al., 2010). This was suggested to be due to lower water content of air at low temperatures which reduces the reaction of TEA with NO_2 (Hagenbjörk-Gustafsson et al., 2010). The temperatures measured in our current study were lower than our previous work which suggests that this effect has less impact with increasing temperature.

At High Street, however, the relationships identified above were not apparent. The Palmes[M], Palmes[S] and Ogawa[P] sampler uptake rates were not significantly correlated with any of the variables tested, while the Palmes UR were only significantly correlated with High Street NO_2 (Table 3). This correlation could suggest that overestimations by the Palmes samplers could be attributed to within-tube chemistry, whereby the rapid changes in NO concentrations (from local traffic sources) mean that reactions occur between NO and O_3 within the diffusion tube, leading to the formation of excess NO_2 within the tube. The negative correlation between Palmes sampler bias and analyser NO_2 has been report by Heal et al. (1999), who suggested that at high concentrations the overestimation of ambient NO_2 by the diffusion tubes is less important and this relationship be attributed to either NO_2 contributing a larger amount to NO_x as NO_2 concentrations increase, or O_3 concentrations are lower at higher NO_x concentrations so there is less impact of NO to NO_2 conversion within

tubes. However, we anticipate that, if bias in diffusion tubes was solely attributable to within-tube reactions, the Palmes[M] would also suffer from these effects, which is not the case. Additionally, if within-tube chemistry was the sole cause we would expect the Palmes sampler concentrations to be close to those predicted by the analyser at the urban background site, where higher ambient O_3 concentrations and an absence of local sources of NO would suggest that the majority of NO_x is made up of NO_2 and hence the diffusion tubes should measure analyser concentrations accurately (Heal et al., 1999).

When both High Street and Townhead were considered together, the variable with the most significant correlation with UR was analyser NO_2 for all except the Palmes[M] sampler, while the Palmes[M] UR was most correlated with absolute humidity. However as previously discussed, the Ogawa[P] and Palmes[S] samplers have shorter path lengths then the Palmes samplers and consequently were not anticipated to suffer from within-tube chemical reactions due to the short residence times of the gases within the samplers. Therefore, we suggest that the observed correlation between analyser NO_2 and sampler uptake rates was a result of confounding of analyser NO_2 concentrations with another variable, with wind-speed having the highest correlation (r = -0.79, P < 0.01 for airport wind-speeds) (Figure S4).

Table 3: Pearson correlation coefficients for relationships between passive sampler uptake rates and exposure-averaged observed meteorological and chemical conditions. No O_3 monitors at High Street therefore the ratio between (Min(NO, O_3)+NO_2)/NO_2 cannot be calculated. The highest correlations for each sampler at each site highlighted in bold.

Townhead (n = 30)	Palmes	Palmes[M]	Palmes[S]	Ogawa[P]
T Townhead	**0.62****	0.38*	0.33	0.52**
RH Townhead	0.04	0.27	0.05	0.25
Abs RH Townhead	0.60**	**0.42***	0.33	**0.55****
WS_Airport	0.40*	-0.28	0.30	0.21
WS_Townhead	0.47**	-0.25	**0.42***	0.34
WS_Kestrel	0.27	-0.31	0.34	0.29
NO_2 Townhead	-0.42*	0.23	-0.39*	-0.45*
Ratio(Min(NO, O_3)+NO_2)/NO_2 Townhead	0.01	0.09	-0.18	0.07
High Street (n = 27)	Palmes	Palmes[M]	Palmes[S]	Ogawa[P]
T Townhead	0.37	**0.35**	-0.07	0.21
RH Townhead	-0.15	-0.09	**-0.31**	-0.03
Abs RH Townhead	0.31	0.31	-0.17	0.23
WS_Airport	0.29	0.14	0.20	0.03
WS_Townhead	0.10	0.21	0.18	0.06
WS_Kestrel	0.38	0.09	0.12	0.08
NO_2_High Street	**-0.55****	-0.23	-0.29	**-0.28**
Combined (n = 57)	Palmes	Palmes[M]	Palmes[S]	Ogawa[P]
T Townhead	0.47**	0.36**	0.17	0.37**
RH Townhead	-0.05	0.11	-0.05	0.12
Abs RH Townhead	0.43**	**0.37****	0.15	0.39**
WS_Airport	0.33*	-0.09	0.24	0.12
WS_Townhead	0.26*	-0.05	0.31*	0.20
WS_Kestrel	0.33*	-0.13	0.27*	0.21
NO_2_AURN	**-0.53****	-0.02	**-0.40****	**-0.40****

*Variables are exposure means of: air temperature (T, °C); relative humidity (RH, %); absolute humidity (AH, g/m^3); wind-speed (WS, m/s); analyser NO_2 (NO_2, ppb), ratio of analyser NO_2 plus minimum of analyser NO (ppb) or O_3 (ppb) to analyser NO_2 (Ratio(Min(NO,O_3)+NO_2)/NO_2); uptake rates (UR) for Palmes, Ogawa[S] & Ogawa[P] samplers. *Correlation coefficient significant at $P < 0.05$; **Correlation coefficient significant at $P < 0.01$.*

Based on the findings above, we suggest that the predominant cause of overestimation by the Palmes sampler is attributed to wind-speed. At the more exposed Townhead urban background site, with higher wind-speed measurements, the overestimation of NO_2 concentrations by the Palmes sampler was larger than observed at the High Street roadside site. The application of post-processing techniques to correct Palmes uptake rates for wind-speed effects improve regression coefficients and statistics at the former site, while at the latter the improvement was limited. Modification of the design of the Palmes sampler to

reduce turbulent effects, through inclusion of a membrane over the exposed tube end, improved regression coefficients at statistics at Townhead however at High Street the opposite was true. This suggests that there may be compromises to make in terms of sampler accuracy if a membrane is included for sites anticipated to have low wind-speeds (High Street wind-speeds were approximately 40 % of those measured at Townhead during the spot measurements, suggesting the whole study average wind-speed at High Street would be approximately 0.5 m/s), however the consistency (represented by R^2) of the measurements made with the membrane in place was greater.

4. Conclusions

Four types of passive sampler were co-located with NO_2 reference analysers at an urban background (Townhead) and a roadside (High Street) sites for 2-3 day exposures over 30 consecutive periods. The samplers tested were: a standard Palmes tube; a Palmes tube with an amorphous polyethylene membrane to minimise the impact of wind-speed (Palmes[M]); a half-length Palmes tube (anticipated to prevent any within-tube chemistry) with membrane (Palmes[S]); and an in-house prepared Ogawa sampler (Ogawa[P]). We measured temperature, relative humidity and wind-speed at the urban background continuously over the deployment periods, and additionally made short duration measurements using a hand-held weather station during the site visits to change the passive samplers.

The overestimation of reference analyser concentrations by the Palmes samplers were larger at the urban background site compared to the roadside site, which could be attributed to higher wind-speeds recorded at the site. We applied a method to post-process the collected NO_2 concentrations from the Palmes samplers for wind-speed effects on sampler uptake rate, which improved the regression equation and statistics between sampler and reference NO_2 (NMB = 0.27 *vs.* 0.13 (Townhead) and 0.08 *vs.* -0.06 (High Street)).

Using a membrane over the open end of the Palmes sampler removed the relationship between sampler uptake rate and wind-speed observed for the standard Palmes tube and improved the statistics describing the sampler and reference analyser concentrations (NMB = 0.17 (Palmes) *vs.* 0.00 (Palmes[M]) for both sites combined). However, the Palmes[S] sampler showed significant relationship with wind-speed, suggesting that the shorter tube length is more influenced by winds-speed, and consequently had larger bias statistics than the Palmes[M] sampler (NMB = -0.05 (Palmes[S] for both sites combined). The regression coefficients for the Palmes[M] and Palmes[S] samplers were poorer than the standard Palmes at High Street (Regression slope = 1.03, 0.97 and 0.85 for Palmes, Palmes[M] and Palmes[S] respectively). The wind-speed measurements at High Street were approximately

40 % lower than those recorded at Townhead and showed less variation throughout the study period. At low (approximately <0.5 m/s) wind-speeds the membrane acts as a barrier to diffusion and reduces the uptake rates below the standard samplers, however the scatter in the collected concentrations was much lower when the membrane was included (R^2 = 0.36 *vs*. 0.78 *vs.* 0.51 for Palmes, Palmes[M] and Palmes[S] respectively at High Street). The in-house Ogawa[P] sampler provided a suitable alternative to the Palmes sampler with less impact of wind-speed (NMB = 0.17 *vs.* -0.03 for both sites combined).

We suggest that for passive sampler deployments when wind-speed information is unavailable (due to, for example, a large number of sites or mobile monitoring) the use of an amorphous polyethylene membrane will reduce turbulent effects and produce more accurate estimates of pollution concentrations than the standard Palmes samplers.

5. References

Brown, R.H., 2000. Monitoring the ambient environment with diffusive samplers: theory and practical considerations. J. Environ. Monit. 2, 1–9. doi:10.1039/A906404D

Bush, T., Stevenson, K., Moorcroft, S., Smith, S., 2001. Validation of nitrogen dioxide diffusion tube methodology in the UK. Atmos. Environ. 35, 289–296. doi:10.1016/S1352-2310(00)00172-2

Buzica, D., Gerboles, M., Amantini, L., Ballesta, P.P., De Saeger, E., 2005. Modelling of the uptake rate of the nitrogen dioxide Palmes diffusive sampler based on the effect of environmental factors. J. Environ. Monit. 7, 169. doi:10.1039/b411474d

Cape, J.N., 2009. The Use of Passive Diffusion Tubes for Measuring Concentrations of Nitrogen Dioxide in Air. Crit. Rev. Anal. Chem. 39, 289–310. doi:10.1080/10408340903001375

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

DEFRA, 2017. The Data Verification and Ratification Process [WWW Document]. URL https://uk-air.defra.gov.uk/assets/documents/The_Data_Verification_and_Ratification_Process.pdf (accessed 4.1.17).

Gerboles, M., Buzica, D., Amantini, L., 2005. Modification of the Palmes diffusion tube and semi-empirical modelling of the uptake rate for monitoring nitrogen dioxide. Atmos. Environ. 39, 2579–2592. doi:10.1016/j.atmosenv.2005.01.012

Gerboles, M., Buzica, D., Amantini, L., Lagler, F., 2006. Laboratory and field comparison of measurements obtained using the available diffusive samplers for ozone and nitrogen dioxide in ambient air. J. Environ. Monit. 8, 112–119. doi:10.1039/B511271K

Gillespie, J., Beverland, I.J., Hamilton, S., Padmanabhan, S., 2016. Development, Evaluation, and Comparison of Land Use Regression Modeling Methods to Estimate Residential Exposure to Nitrogen Dioxide in a Cohort Study. Environ. Sci. Technol. 50, 11085–11093. doi:10.1021/acs.est.6b02089

Gillespie, J., Masey, N., Heal, M.R., Hamilton, S., Beverland, I.J., 2017. Estimation of spatial patterns of urban air pollution over a 4-week period from repeated 5-min measurements. Atmos. Environ. 150, 295–302. doi:10.1016/j.atmosenv.2016.11.035

Hagenbjörk-Gustafsson, A., Tornevi, A., Forsberg, B., Eriksson, K., 2010. Field validation of the Ogawa diffusive sampler for NO2 and NOx in a cold climate. J. Environ. Monit. 12, 1315. doi:10.1039/b924615k

Heal, M.R., 2008. The effect of absorbent grid preparation method on precision and accuracy of ambient nitrogen dioxide measurements using Palmes passive diffusion tubes. J. Environ. Monit. 10, 1363–1369. doi:10.1039/b811230d

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. doi:10.1023/a:1006249016103

Heal, M.R., O'Donoghue, M.A., Aguis, R.M., Beverland, I.J., 1999a. Application of passive diffusion tubes to short-term indoor and personal exposure measurement of NO2. Environ. Int. 25, 3–8. doi:10.1016/s0160-4120(98)00092-0

Heal, M.R., O'Donoghue, M.A., Cape, J.N., 1999. Overestimation of urban nitrogen dioxide by passive diffusion tubes: a comparative exposure and model study. Atmos. Environ. 33, 513–524. doi:10.1016/S1352-2310(98)00290-8

Kirby, C., Fox, M., Waterhouse, J., Drye, T., 2000. Influence of environmental parameters on the accuracy of nitrogen dioxide passive diffusion tubes for ambient measurement. J. Environ. Monit. 3, 150–158. doi:10.1039/b007839p

Lewné, M., Cyrys, J., Meliefste, K., Hoek, G., Brauer, M., Fischer, P., Gehring, U., Heinrich, J., Brunekreef, B., Bellander, T., 2004. Spatial variation in nitrogen dioxide in three European areas. Sci. Total Environ. 332, 217–230. doi:10.1016/j.scitotenv.2004.04.014

Martin, N.A., Helmore, J.J., White, S., Barker Snook, I.L., Parish, A., Gates, L.S., 2014. Measurement of nitrogen dioxide diffusive sampling rates for Palmes diffusion tubes using a controlled atmosphere test facility (CATFAC). Atmos. Environ. 94, 529–537. doi:10.1016/j.atmosenv.2014.05.064

Masey, N., Gillespie, J., Heal, M.R., Hamilton, S., Beverland, I.J., 2017. Influence of wind-speed on short duration NO2 measurements using Palmes and Ogawa passive diffusion samplers. Atmos. Environ. doi:http://www.sciencedirect.com/science/article/pii/S1352231017302467

Ogawa, 2006. NO, NOs, NOx and SO2 Sampling Protocol Using the Ogawa Sampler [WWW Document]. URL http://ogawausa.com/wp-content/uploads/2014/04/prono-noxno2so206.pdf

Pfeffer, U., Zang, T., Rumpf, E.M., Zang, S., 2010. Calibration of diffusive samplers for nitrogen dioxide using the reference method–Evaluation of measurement uncertainty. Gefahrstoffe–Reinhaltung Luft 70, 500–506.

Plaisance, H., Plechocki-Minguy, A., Garcia-Fouque, S., Galloo, J.C., 2004. Influence of meteorological factors on the NO2 measurements by passive diffusion tube. Atmos. Environ. 38, 573–580. doi:10.1016/j.atmosenv.2003.09.073

Targa, J., Loader, A., 2008. Diffusion Tubes for Ambient NO2 Monitoring: A Practical Guide (No. AEA/ENV/R/2504 – Issue 1a). AEA Energy and Environment.

van Reeuwijk, H., Fischer, P.H., Harssema, H., Briggs, D.J., Smallbone, K., Lebret, E., 1998. Field Comparison of two NO2 Passive Samplers to Assess Spatial Variation. Environ. Monit. Assess. 50, 37–51. doi:http://dx.doi.org/10.1023/A:1005703722232

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.atomsenv.2009.02.048

Yu, C.H., Morandi, M.T., Weisel, C.P., 2008. Passive dosimeters for nitrogen dioxide in personal/indoor air sampling: A review. J. Expo. Sci. Environ. Epidemiol. 18, 441–451. doi:10.1038/jes.2008.22

Supplementary Information

Minimising the impact of wind-speed effects on NO_2 passive diffusion samplers through sampler modifications

Nicola Masey[1], *Fiona Sutherland*[1], *Samuel Grainger*[1], *Scott Hamilton*[2], *Mathew R. Heal*[3], *Iain J. Beverland*[1*]

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

[3]School of Chemistry, Joseph Black Building, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

*CORRESPONDING AUTHOR: Dr Iain J. Beverland, email: iain.beverland@strath.ac.uk; Tel: +44 141 548 3202

Table S1: Exposure dates for the passive samplers. Sites were visited on foot, and on each occasion Townhead was visited first (approximately 30 minutes between Townhead and High Street deployment times). On some occasions we were unable to access High Street due to ladders being unavailable – on these dates the samplers were consequently exposed for 2 deployment periods.

Exposure number	Exposure dates	Exposure details
1	21 Nov – 23 Nov 2016	
2	23 Nov – 25 Nov 2016	Townhead diffusion tubes were frosty
3	25 Nov – 28 Nov 2016	
4	28 Nov – 30 Nov 2016	
5	30 Nov – 2 Dec 2016	
6	2 Dec – 5 Dec 2016	
7	5 Dec – 7 Dec 2016	
8	7 Dec – 9 Dec 2016	
9	9 Dec – 12 Dec 2016	
10	12 Dec – 14 Dec 2016	
11	14 Dec – 16 Dec 2016	
12	16 Dec – 19 Dec 2016	
13	9 Jan – 11 Jan 2017	
14	11 Jan – 13 Jan 2017	
15	13 Jan – 16 Jan 2017	
16	16 Jan – 18 Jan 2017	
17	18 Jan – 20 Jan 2017	
18	20 Jan – 23 Jan 2017	
19	23 Jan – 25 Jan 2017	
20	25 Jan – 27 Jan 2017	
21	27 Jan – 30 Jan 2017	No access to High Street on 27th January
22	30 Jan – 1 Feb 2017	High Street analyser off
23	1 Feb – 3 Feb 2017	
24	3 Feb – 6 Feb 2017	
25	6 Feb – 8 Feb 2017	
26	8 Feb – 10 Feb 2017	
27	10 Feb – 13 Feb 2017	
28	13 Feb – 15 Feb 2017	
29	15 Feb – 17 Feb 2017	No access to High Street on 15th February
30	17 Feb – 20 Feb 2017	

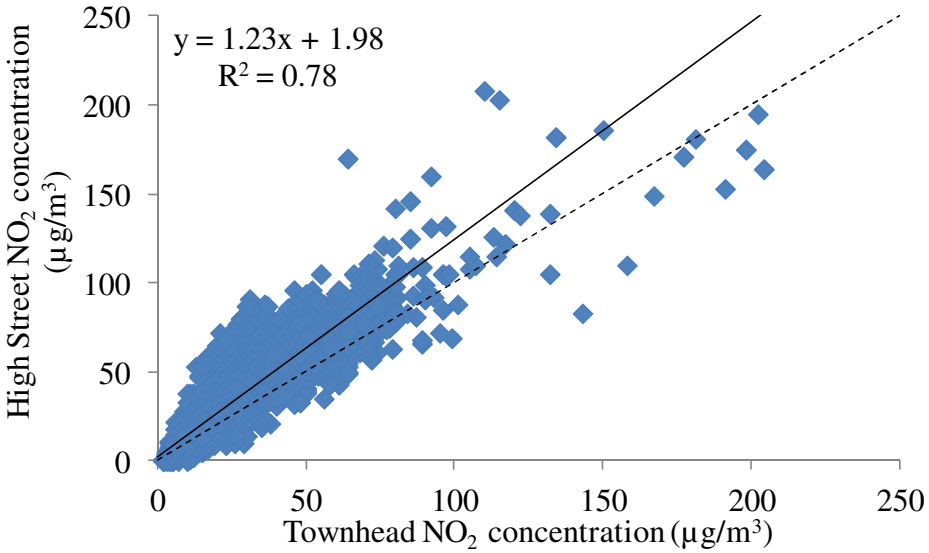


Figure S1: Scatter plots showing hourly average NO_2 concentration measured by the analysers located at Townhead and High Street AURN stations between 20th November 2016 and 22nd February 2017 inclusive. The solid line represents the reduced major axis trend line, while the dashed line represents 1:1 concentrations.

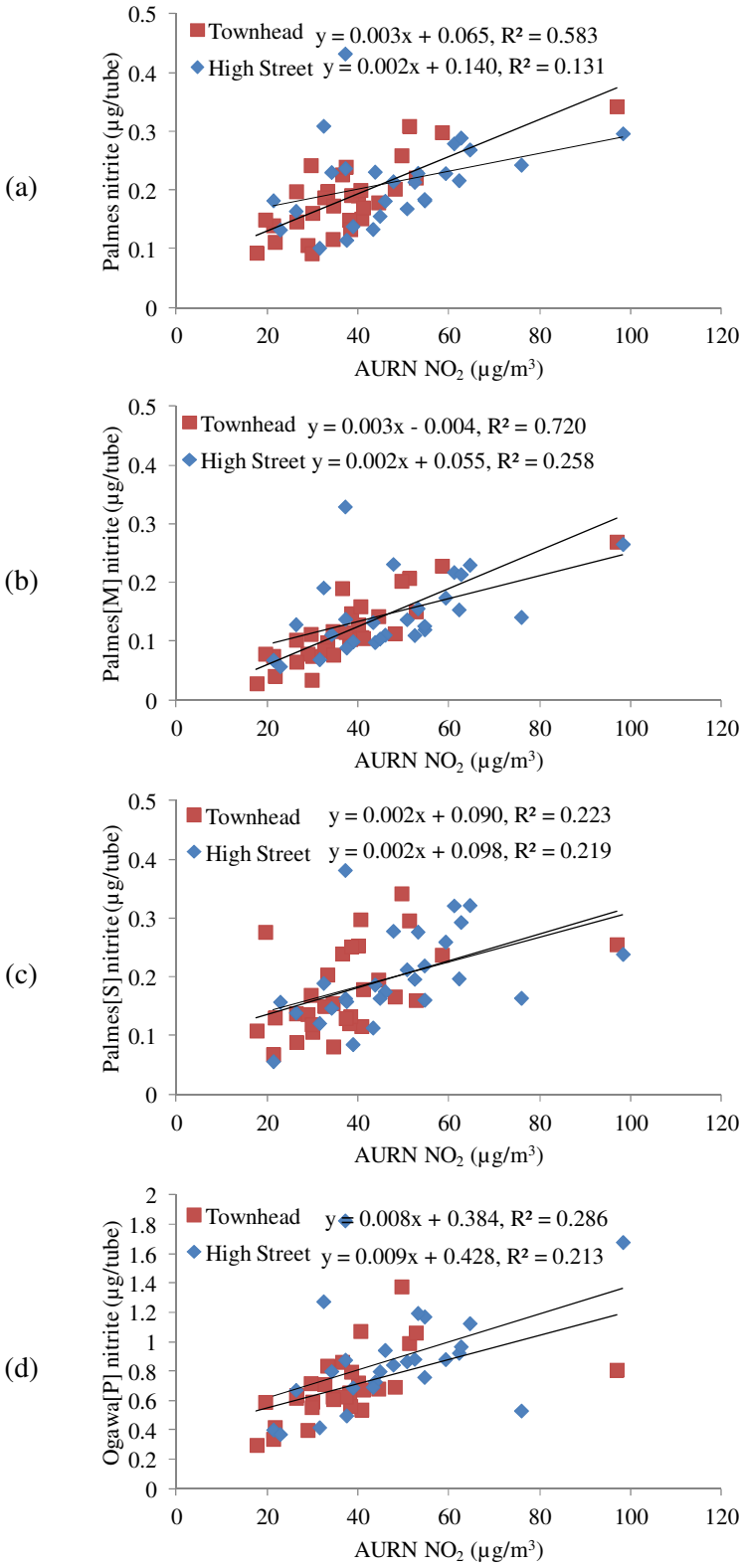


Figure S2: Mass of nitrite collected by (a) Palmes, (b) Palmes[M], (c) Palmes[S] and (d) Ogawa[P] samplers against corresponding study-average NO_2 concentrations measured by the analyser. Data is shown for Townhead and High Street on each plot.

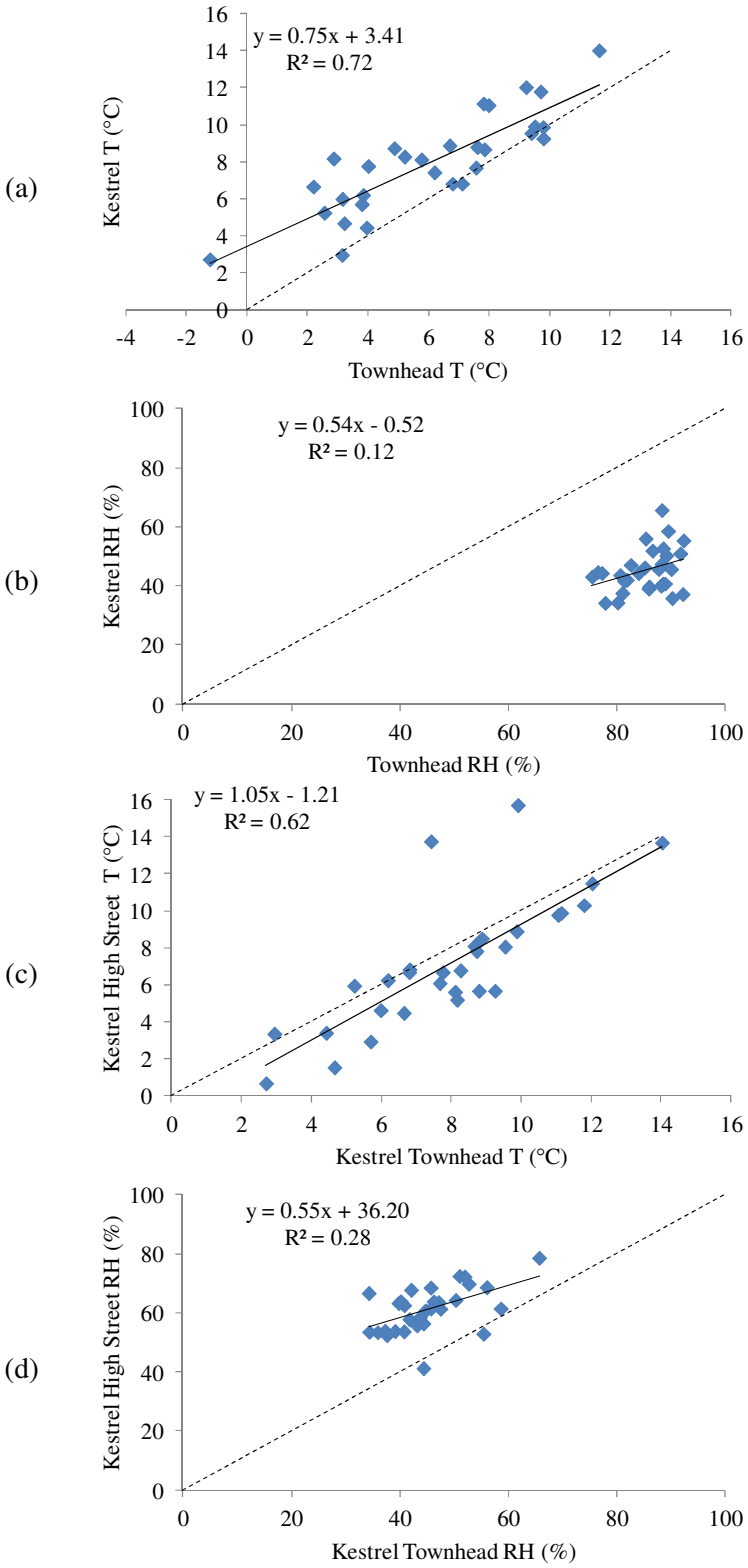


Figure S3: correlation between study-average spot measurements of (a) temperature and (b) relative humidity made by the kestrel instrument at Townhead and the static meteorological measurements made at the site; and between the spot measurements of (c) temperature and (d) relative humidity made at Townhead and High Street.

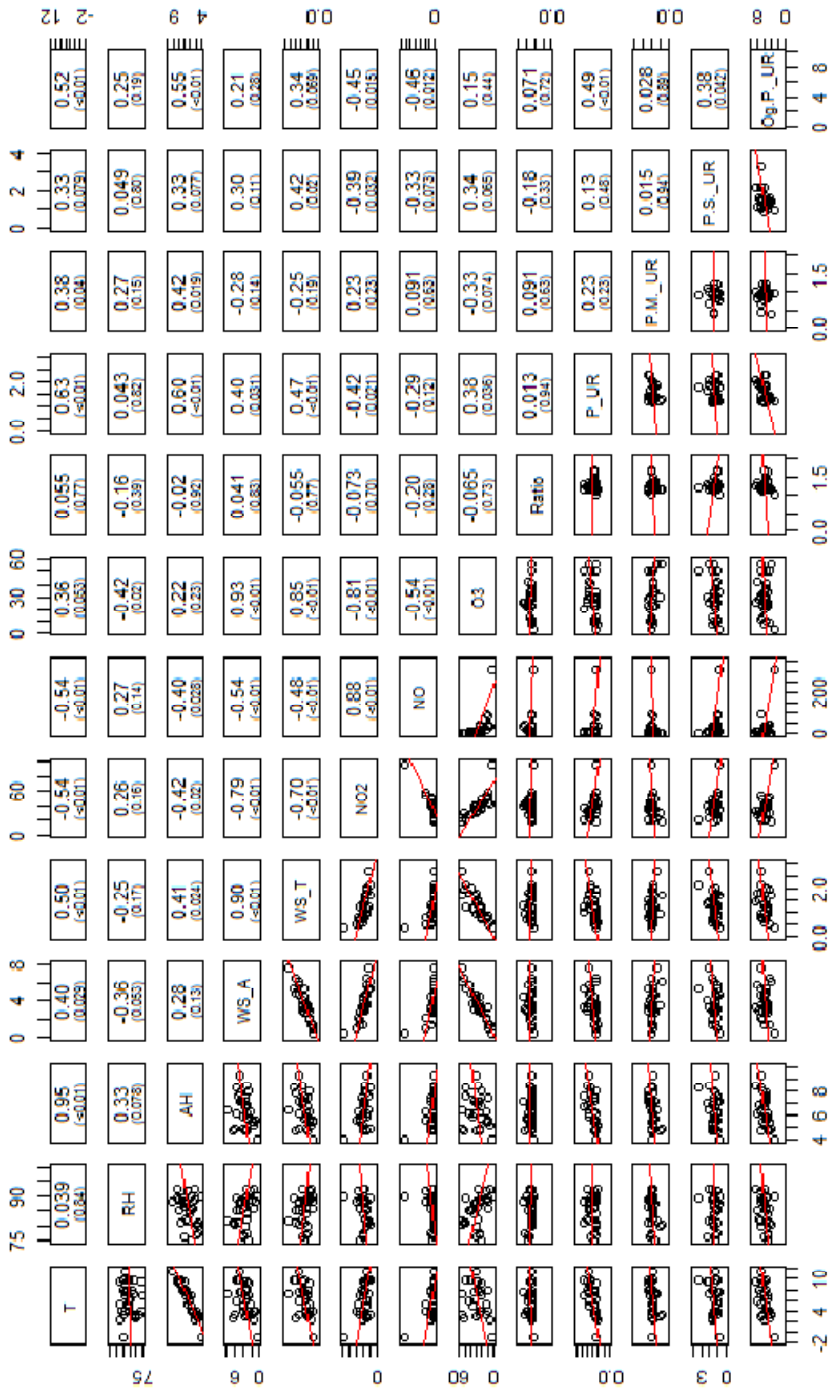


Figure S4: Correlation matrix showing Pearson correlation coefficients between meteorological and chemical variables and the passive sampler uptake rates at Townhead only. Variables are exposure-averaged: T = temperature (°C); RH = relative humidity (%); AH = absolute humidity (g/m^3); WS_A = wind-speed at the airport (m/s); WS_T = wind-speed measured at Townhead (m/s); NO_2, NO and O_3 correspond to concentrations measured by automatic analyser at Towhead ($\mu g/m^3$); Ratio is the min($(NO, O_3)+NO_2)/NO_2$ ($\mu g/m^3$); and UR corresponds to sampler uptake rates (P = Palmes, P.M = Palmes[M], P.S = Palmes[S] and Og.P = Og[P])(cm^3/min). The *P*-value for the correlation is shown in brackets below the correlation coefficient.

4. Calibration of portable real-time air pollution monitors

This chapter is comprised of two manuscripts, which are discussed below in more detail:

Temporal changes in field calibration relationships for Aeroqual S500 O_3 and NO_2 sensor-based monitors

This paper examines the optimal calibration procedure for Aeroqual S500 sensor-based monitors measuring NO_2 and O_3 at an urban background station in Glasgow, UK. Calibration of these monitors has previously been demonstrated to improve concentration estimate; however no work has investigated the impact of the time or duration of the calibration. Appendix A: Evaluation of Little Environmental Observatory (LEO) sensors describes the evaluation of a different portable real-time sensor (also measuring NO_2 and O_3) and highlights the difference in accuracy between different portable monitors and shows the accuracy of the Aeroqual monitors to be superior to the LEO monitors.

N. Masey, the main author, designed the experiment, carried out the practical work and associated data analysis and wrote the paper. J. Gillespie provided advice about experimental design, helped with data collection and provided discussion about the data analysis. E. Ezani helped with data collection. C. Lin and H. Wu provided discussion about data analysis and interpretation. S. Hamilton, M. Heal and I. Beverland provided advice on data analysis and comments during preparation of the manuscripts.

This paper has been formatted to meet the criteria of Environmental Science and Technology.

Consistency of urban background black carbon concentration measurements by portable AE51 and reference AE22 Aethalometers: Effect of corrections for filter loading

Aethalometers measuring black carbon have been shown to suffer from filter loading effects which lead to these monitors overestimating true concentrations without correction for these effects. We investigate two commonly applied correction methods on the accuracy of microAeth AE51 portable Aethalometers at an urban background station in Glasgow, UK.

N. Masey, the main author, designed the experiment, carried out the practical work and associated data analysis and wrote the paper. J. Gillespie provided advice about experimental design, helped with data collection and provided discussion about the data analysis. E. Ezani helped with data collection. S. Hamilton, M. Heal and I. Beverland provided advice on data analysis and comments during preparation of the manuscripts.

This paper has been formatted with the aim of submission of the manuscript to the Journal of Exposure Science and Environmental Epidemiology.

Temporal changes in field calibration relationships for Aeroqual S500 O_3 and NO_2 sensor-based monitors

Nicola Masey[1], Jonathan Gillespie[1], Eliani Ezani[1,2], Chun Lin[3], Hao Wu[3], Neil S Ferguson[1], Scott Hamilton[4], Mathew R. Heal[3], Iain J. Beverland[1]*

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]Department of Environmental and Occupational Health, Faculty of Medicine and Health Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, MALAYSIA

[3]School of Chemistry, Joseph Black Building, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

[4]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

***CORRESPONDING AUTHOR:** Dr Iain J. Beverland, email: iain.beverland@strath.ac.uk

Abstract:

Sensor-based monitors are increasingly used to measure air pollutant concentrations, but require calibration under field conditions. We made intermittent comparisons (6 times over 6-month period) between ozone and nitrogen dioxide concentrations measured by Aeroqual gas-sensitive semiconductor (O_3) and electrochemical (NO_2) sensors (two of each) and reference analysers in the UK Automatic Urban and Rural Network. Each deployment period was split into equal (n = 48 h) training and test datasets, to derive and test calibration equations respectively. We observed significant bivariate linear relationships between *Aeroqual O_3* and *Reference O_3* concentrations, and significant multiple linear relationships between *Aeroqual NO_2* and both *Reference NO_2* and *Aeroqual O_3* concentrations. Changes in monitor responses over time resulted in relatively inaccurate concentrations estimates (*cf.* reference concentrations) from calibration equations derived in the first training period and applied to subsequent *test* deployments (e.g. RMSE = 47.2 µg m^{-3} (n = 286) for a dataset of all *test* periods combined, for one of the two monitor pairs). Substantial improvements in accuracy of estimated concentrations were achieved by combination of repeated intermittent training data into a single calibration dataset (RMSE = 8.5 µg m^{-3} for same *test* dataset described above). This latter approach to field calibration is recommended.

1. Introduction

The concentrations of gaseous air pollutants with deleterious health effects, including nitrogen dioxide (NO_2) and ozone (O_3), are monitored to estimate human exposure and to assess compliance with legislation and guidelines. Monitoring is usually conducted at static, automatic monitoring stations that record concentrations at high temporal resolution. However, the cost of these stations, together with practical considerations regarding suitable sites, invariably limits the spatial coverage of automatic networks. Wider geographical networks of passive diffusion samplers (PDS) can mitigate restrictions on spatial coverage[1,2]; however PDS provide limited temporal information, and can be subject to measurement inaccuracies associated with changing meteorological[3] and atmospheric chemistry conditions.[4]

Battery-powered real-time hand-held sensor-based monitors for air pollutants are continually being developed, which have potential to supplement data from existing monitoring networks.[5] Hand-held monitors usually have lower capital costs than automatic analysers, meaning that more could be made available for deployment, potentially increasing the spatial resolution of measurement networks with high temporal resolution concentration measurements.[6–9] The use of such instruments in mobile and personal monitoring has been reported.[8,10]

The reference analysers in automatic monitoring stations are usually subject to documented quality control and assurance (QC/QA) procedures that aim to ensure the recorded pollutant concentrations are within specified ranges of accuracy and precision. The control of uncertainty requires resource and effort, so it is not unreasonable to anticipate that outputs from lower-cost, portable monitors are also subject to uncertainties, which may also vary with time. Thus, it is important that portable monitors are subject to calibration checks at least as much as reference analysers. An important additional potential limitation of gas sensor-based monitors is their cross-sensitivity to other pollutants and/or to changing environmental conditions.[11,12] For example, the responses of some electrochemical sensors have been shown to be susceptible to variations in temperature or relative humidity.[8,13–15] Sensors for NO_2 have been shown to be cross-sensitive to O_3, meaning that both pollutants must be measured simultaneously to allow correction of the NO_2 sensor response.[8–10,16–18] Consequently, for accurate estimation of ambient concentrations, sensor-based monitors require field calibration under representative ambient conditions.

In this study we evaluated the responses of two pairs of Aeroqual S500 O_3 and NO_2 monitors over time and repeated exposure to outdoor conditions, by deployment adjacent to reference gas analysers. The monitor deployments were designed to be representative of

their likely use in practical field measurements: namely repeated cycles of calibration and monitoring (involving switching monitors on and off, removal from a reference site and usage in other field measurements at different locations and mobile monitoring, and return to the reference site). Additionally, we investigated the effects of combination and timing of field calibration of the monitors on the accuracy of calibrated estimates. Our analyses provide insight to the refinement and use of field calibration relationships for these Aeroqual monitors in mobile measurements across geographical areas with widely varying pollution microclimates.

2. Methods

We evaluated two pairs of Aeroqual (www.aeroqual.com) S500 O_3 and NO_2 monitors at the UK Automatic Urban and Rural Network (AURN) Townhead urban background site in Glasgow city centre (55.866 ºN, 4.244 ºW). Hourly-average concentration data from the reference analysers (API200A chemiluminescence analyser for NO_2 and Thermo 49i photometric analyser for O_3) at this site were downloaded from www.scottishairquality.co.uk. All AURN measurements were subject to documented national QC/QA procedures.[19]

We deployed the Aeroqual monitors within ventilated waterproof enclosures provided by Aeroqual attached to the galvanised steel safety railings surrounding the roof of the AURN monitoring cabin (Supporting Information Figure S1). The NO_2 monitors contained electrochemical sensors (ENW2, range 0 – 1 ppm) and are referred to as *NO_2_1* and *NO_2_2*. The O_3 monitors contained gas-sensitive semiconductor sensors (OZU2, range 0 – 0.15 ppm) and are referred to as *O_3_3* and *O_3_4*. Monitors *NO_2_1* and *O_3_3* were located next to one another on the eastern railing while monitors *NO_2_2* and *O_3_4* were located on the western railing. Mains power was available allowing the monitors to operate continuously. We set the monitors to record gravimetric concentrations (µg/m^3) at 1-minute intervals, prior to computation of hourly-average concentrations for comparison with the analysers.

Six separate monitor co-location deployment periods were undertaken intermittently over 6 months (November 2015 – May 2016) (Table S1). We truncated each deployment period to the first 96 h of field deployment to simplify comparison between periods.

3. Results

3.1. Calibration of Aeroqual O_3 monitors

Aeroqual O_3 measurements closely followed temporal trends in the Reference O_3 concentrations, with *O_3_3* generally over-predicting and *O_3_4* generally under-predicting the analyser concentrations (Figure 1).

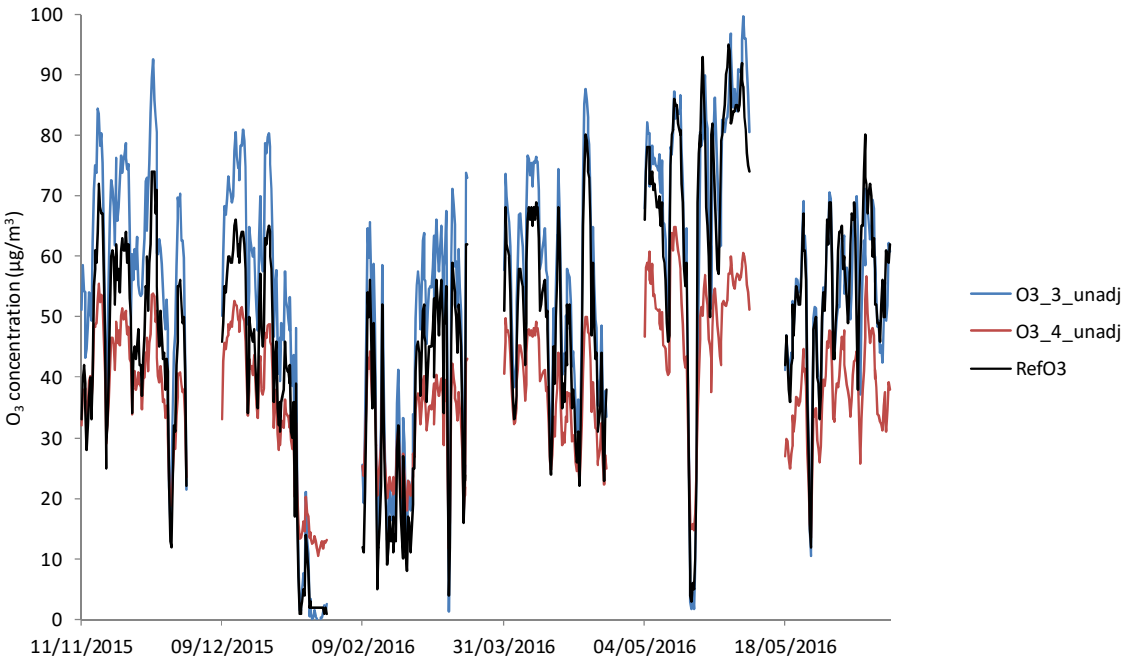


Figure 1: Time series of unadjusted hourly-averaged O_3 concentrations measured by the two O_3 Aeroqual monitors and by the reference analyser for the full duration of each deployment period. Different deployment periods are separated by gaps in time series. The coefficient of determination between the concentrations measured by duplicate Aeroqual O_3 instruments was 0.90.

We calibrated the response of the O_3 monitors by calculating Ordinary Least Squares (OLS) regression equations between unadjusted Aeroqual O_3 measurements and O_3 concentrations measured by the reference analyser. We compared three methods of calibration - in each method the first 48 h for each deployment period was used as 'training' data to generate calibration equations, and the second 48 h was used as 'test' data to evaluate the accuracy of the calibrated predictions.

1) The first calibration method used the training data [0-48 h] for each deployment period to correct the test data [49-96 h] for the same deployment period i.e. a *unique* calibration

for each deployment period (referred to as '*Aq_corr_u*'). This method represented calibrations computed at close time intervals to the measurements being corrected.

2) The second calibration method used a combination of training data from *all* deployment periods to derive a single global calibration equation for each monitor, which was then applied to the training data for each period (referred to as *'Aq_corr_a'*). This method of calibration, involving interspersed intervals during the extended set of deployments, had a larger number of data points in the calibration, spread over a longer time period, and thus incorporated a larger range of pollution conditions.
3) The third calibration method used the calibration equation derived from the training data from the first study period (*November*) to correct the test data from all of the subsequent studies (referred to as *'Aq_corr_N'*). This method represented a 'one-off pre-measurement campaign' calibration, and was included to assess how the accuracy and precision of a single calibration might deteriorate over time during extended field measurements.

The calibration equations are summarised in Table 1 and are shown in Figure S3. Similar to the findings observed by Lin et al.[16] we observed limited and inconsistent effects of temperature and relative humidity on the calibration regression equations across the monitors and deployments (data not shown) and therefore these meteorological variables were not considered further in our analyses.

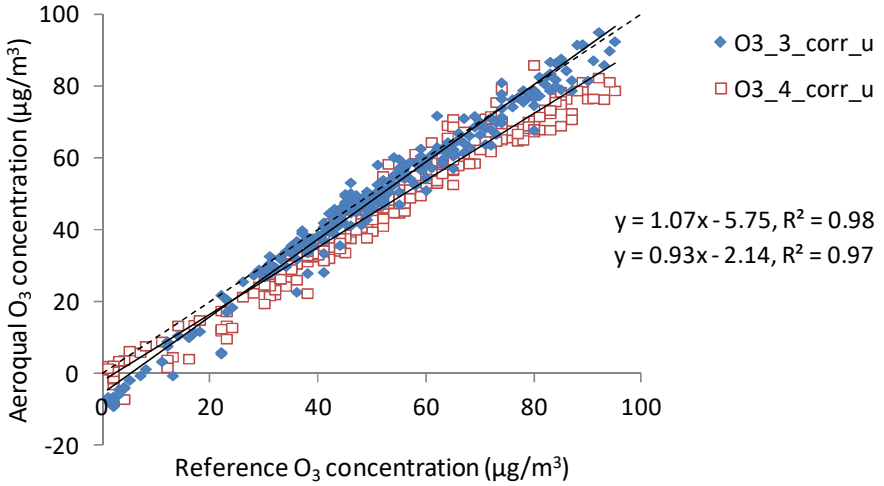
The OLS calibration equations calculated for each of the three methods were used to adjust the test data. We observed similar overall temporal patterns in the time series of adjusted Aeroqual and Reference O_3 concentrations (adjusted concentration estimates shown for both *training* and *test* data in Figure S4). The Reference O_3 and adjusted Aeroqual O_3 concentrations deviated for a short period at the end of the December deployment, when the reference analyser measured very low O_3 concentrations and the adjusted Aeroqual values were slightly negative. This period corresponded to a winter pollution episode with elevated NO_2 concentrations (Figure S5). None of the three calibration methods accurately adjusted the Aeroqual measurements at these very low O_3 concentrations, presumably as a result of the marked change in pollution concentrations between the training and test data.

All three calibration methods yielded high correlation coefficients ($R^2 > 0.90$) and slopes close to 1 (slope: 0.91 – 1.07) for the evaluation of adjusted test data against analyser-measured O_3 concentrations (Figures 2 and S7). However, the scatter plot for the *Aq_corr_N* calibration (for both Aeroqual units) had a relatively large negative intercept compared to the scatter plots for the other two calibration methods, resulting from underestimation of reference analyser O_3. The root mean square error (RMSE), mean bias (MB) and normalised

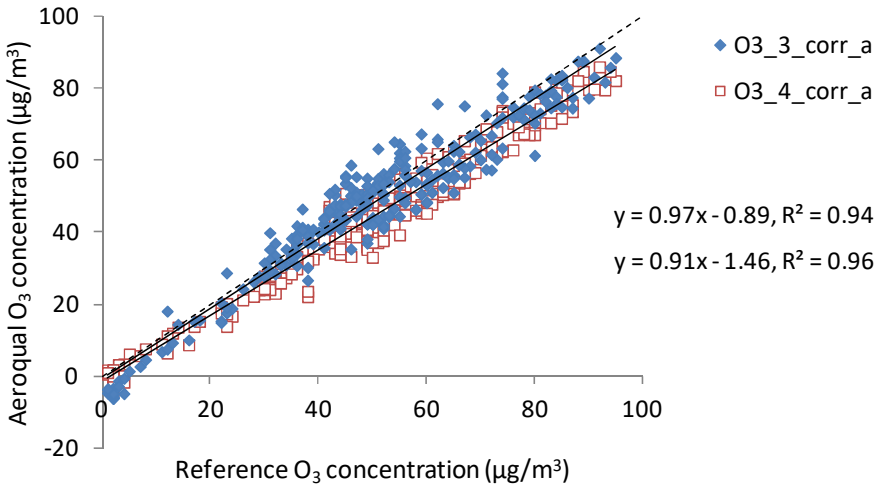
mean bias (NMB) for *Aq_corr_N* were correspondingly larger than for the other two calibration methods (Table 2, Figure S7).

These evaluation statistics indicate that calibrations derived from training data interspersed regularly throughout the deployment periods provide a more accurate estimate analyser concentrations in the test periods than a single calibration at the start of the deployment periods. For these O_3 monitors there were only small differences in R^2, RMSE, MB, and NMB statistics between *Aq_corr_u* and *Aq_corr_a* methods, which were partly monitor dependent. For *O_3_3*, all four statistics for the *Aq_corr_u* method indicated a slightly better fit between adjusted sensor and analyser data. There was a less consistent pattern between calibration methods for the statistics for *O_3_4* although differences were negligible. Hence we conclude that both *Aq_corr_u* and *Aq_corr_a* methods provide useful adjustment to the O_3 monitor data.

(a)



(b)



(c)

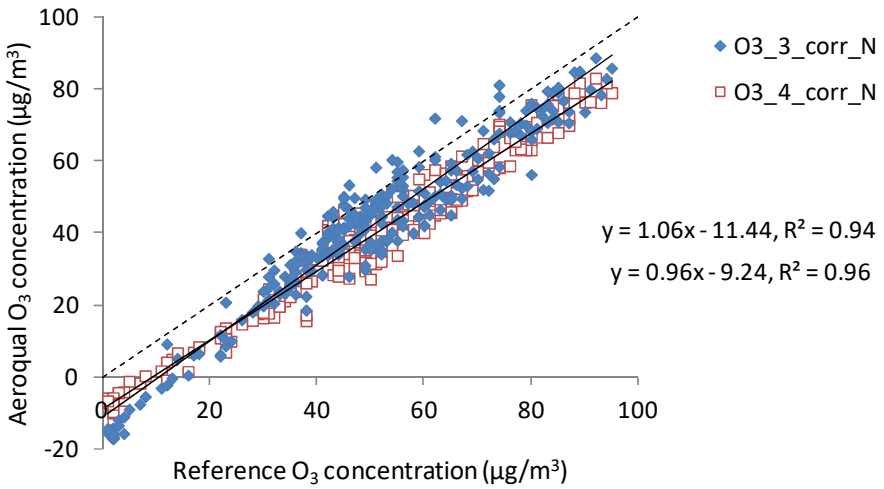


Figure 2: Scatter plots of adjusted hourly Aeroqual O_3 concentration estimates *vs.* reference analyser O_3 concentrations for the test data (i.e. 2nd half of deployment periods only). (a) *O_3_aq* adjusted using calibration derived from training data in each unique deployment period (*O_3_aq_corr_u*). (b) *O_3_aq* adjusted using calibration derived from training data combined from all deployment periods (*O_3_aq_corr_a*). (c) *O_3_aq* adjusted using calibration derived from the first deployment period (November) only (*O_3_aq_corr_N*).

3.2. Calibration of Aeroqual NO_2 monitors

We previously showed that an Aeroqual NO_2 monitor was also sensitive to O_3 concentrations, and demonstrated a linear relationship between *(Aeroqual_NO₂ – Reference_NO₂) vs. Aeroqual_O₃* which we used to correct the Aeroqual NO_2 concentrations.[16] Preliminary investigations of the Aeroqual NO_2 monitors used in the present work also revealed clear sensitivity of output on O_3 concentration. In our previous method,[16] the coefficient for *Reference_NO₂* was effectively constrained to 1, which reflects the expectation that when the Aeroqual sensors are new and recently factory-calibrated their output concentration values should be close to True (i.e. reference) NO_2 concentrations:

$$Aeroqual\ NO_2 - True\ NO_2 = a_1 * Aeroqual\ O_3 + b_1 \quad \text{[Equation 1]}$$

Therefore, we can calibrate Aeroqual NO_2 by rearranging Equation 1:

$$True\ NO_2 = Aeroqual\ NO_2 - a_1 * Aeroqual\ O_3 - b_1 \quad \text{[Equation 2]}$$

A potential shortcoming of this approach is that it does not allow for an Aeroqual NO_2 monitor to have a relationship to True NO_2 (i.e. to reference analyser NO_2) that is not equal to 1; for example if the response of the sensor in the NO_2 monitor has reduced in time since its factory calibration. A more general model for the Aeroqual NO_2 monitor response is therefore based on the following two underlying relationships:

1) Aeroqual O_3 has a linear response to True (reference analyser) O_3, i.e.

$$Aeroqual\ O_3 = m_1 * True\ O_3 + c_1 \quad \text{[Equation 3]}$$

2) Aeroqual NO_2 has linear responses to both NO_2 and O_3, and with its response to O_3 being different to the response of the Aeroqual O_3 monitor to O_3, i.e.

$$Aeroqual\ NO_2 = m_2 * True\ NO_2 + m_3 * True\ O_3 + c_2 \quad \text{[Equation 4]}$$

Substituting for True O_3 from Equation 3 into Equation 4 yields:

$$Aeroqual\ NO_2 = m_2 * True\ NO_2 + \frac{m_3}{m_1}(Aeroqual\ O_3 - c_1) + c_2 \quad \text{[Equation 5]}$$

Rearrangement of Equation 5 into a similar form to Equation 2 gives:

$$True\ NO_2 = \frac{1}{m_2} * Aeroqual\ NO_2 - \frac{m_3}{m_1 * m_2} * Aeroqual\ O_3 - \left(\frac{c_2}{m_2} - \frac{c_1 * m_3}{m_1 * m_2}\right) \quad \text{[Equation 6]}$$

where $\frac{m_3}{m_1 * m_2}$ and $(\frac{c_2}{m_2} - \frac{c_1 * m_3}{m_1 * m_2})$ are constants.

Equation 6 is very similar to Equation 2 produced in the Lin et al. method, except that there is now a coefficient for the Aeroqual NO_2 term that may not be equal to 1. To account for this coefficient for Aeroqual NO_2 in practice, a multiple linear regression was fitted to the Aeroqual NO_2 values to derive three fitted parameters:

$$Aeroqual\ NO_2 = k_1 * True\ NO_2 + k_2 * Aeroqual\ O_3 + k_3 \qquad \text{[Equation 7]}$$

The best estimates for k_1, k_2 and k_3 were then used to calibrate Aeroqual NO_2 concentrations as follows:

$$True\ NO_2 = \frac{1}{k_1}(Aeroqual\ NO_2 - k_2 * Aeroqual\ O_3 - k_3) \qquad \text{[Equation 8]}$$

As for the O_3 monitors, the impact of temperature and relative humidity on NO_2 calibration regression equations was limited and inconsistent and was therefore not considered further (data not shown).

The Aeroqual monitors were arbitrarily paired in this correction (reflecting the arbitrary pairing when a pair of monitors is used for field measurements), with *NO_2_1* corrected using *O_3_3* and *NO_2_2* corrected using *O_3_4*. The three different calibration methods described in Section 3.1 were used; leading to Aeroqual NO_2 concentrations adjusted using the *Aq_corr_u*, *Aq_corr_a* and *Aq_corr_N* selections of training data (Table 1).

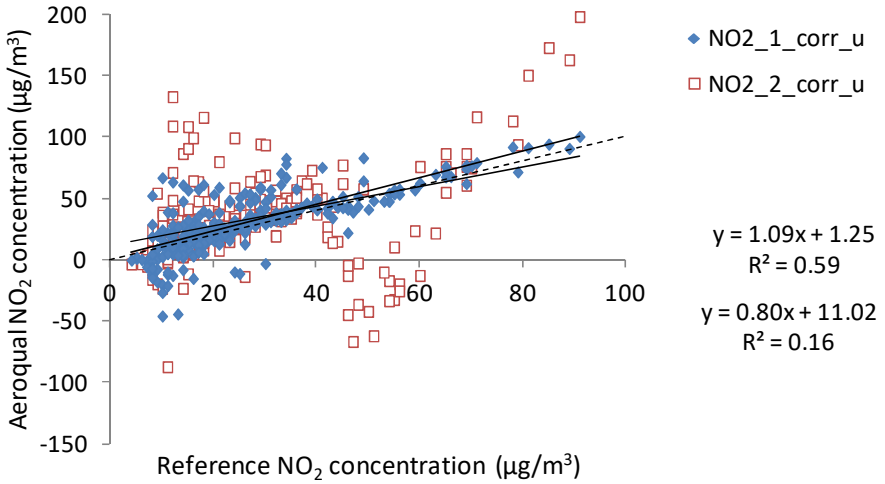
Table 1: OLS equations for calibration equations for each training period (equivalent to *aq_corr_u*) and a global calibration for a combination of 'All' training periods (*aq_corr_a*) for the two pairs of Aeroqual O_3 and NO_2 monitors. The range of concentrations measured during each study period in shown in Table S1.

Pollutant	Study period	O_3_3	O_3_4
O_3	Nov	0.94*Ref_O_3 + 16.11, R^2 = 0.85	0.53*Ref_O_3 + 16.73, R^2 = 0.95
	Dec	1.07*Ref_O_3 + 9.34, R^2 = 0.96	0.59*Ref_O_3 + 12.19, R^2 = 0.90
	Feb	1.08*Ref_O_3 + 5.41, R^2 = 0.98	0.51*Ref_O_3 + 15.27, R^2= 0.96
	April	1.06*Ref_O_3 + 3.61, R^2 = 0.99	0.53*Ref_O_3 + 13.12, R^2 = 0.94
	May1	1.05*Ref_O_3 – 0.07, R^2 = 0.99	0.60*Ref_O_3 + 11.20, R^2 = 0.98
	May2	1.00*Ref_O_3 +0.83, R^2 = 0.99	0.60*Ref_O_3 + 5.41, R^2 = 0.95
	O_3_corr_all	1.02*Ref_O_3 + 6.26, R^2 = 0.93	0.56*Ref_O_3 + 12.66, R^2 = 0.91

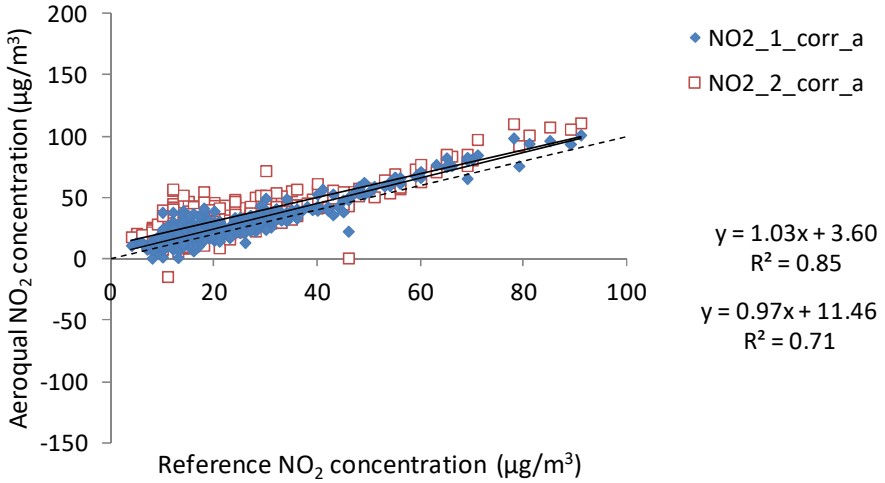
	Study period	NO_2_1	NO_2_2
NO_2	Nov	0.10*Ref_NO_2 + 0.17*O_3_3 + 8.99, R^2 = 0.34	0.38*Ref_NO_2 + 0.84*O_3_4 + 53.78, R^2 = 0.50
	Dec	0.32*Ref_NO_2 + 0.28*O_3_3 - 3.29, R^2 = 0.33	0.07*Ref_NO_2 + 0.19*O_3_4 + 87.99, R^2 = 0.02
	Feb	0.51*Ref _NO_2 + 0.49*O_3_3 – 18.95, R^2 = 0.92	0.35*Ref_NO_2 + 1.02*O_3_4 + 51.98, R^2 = 0.63
	April	0.50*Ref_NO_2 + 0.48*O_3_3 – 20.24, R^2 = 0.72	0.15*Ref_NO_2 + 0.53*O_3_4 + 70.82, R^2 = 0.36
	May1	0.54*Ref_NO_2 + 0.57*O_3_3 – 23.28, R^2 = 0.91	0.38*Ref_NO_2 + 1.05*O_3_4 + 47.73, R^2 = 0.93
	May2	0.17*Ref_NO_2 + 0.27*O_3_3 + 2.32, R^2 = 0.63	0.15*Ref_NO_2 + 0.66*O_3_4 + 71.21, R^2 = 0.74
	NO_2_corr_all	0.42*Ref_NO_2 + 0.43*O_3_3 – 13.19, R^2 = 0.67	0.33*Ref_NO_2 + 0.81*O_3_4 + 58.00, R^2 = 0.62

The adjusted Aeroqual NO_2 and Reference NO_2 concentrations were well correlated temporally, with the exception of some concentrations adjusted using *Aq_corr_u* (Figure S5). The *Aq_corr_u* (*NO_2_1* only), *Aq_corr_a* and *Aq_corr_N* (*NO_2_2* only) calibration methods resulted in clearly defined linear relationships between adjusted Aeroqual and Reference NO_2 concentrations with R^2 values ranging from 0.70 to 0.85 (Figure 3 & Table 2). The regression slopes were closest to 1, R^2 values highest, and RMSE values lowest for the *Aq_corr_a* selection of training data, for both Aeroqual NO_2 monitors (Figures 3 & S7 and Table 2). The period of deviation between Aeroqual *NO_2_2_corr_u* and Reference NO_2 is from the December deployment, during which there was a very poor training data regression fit with small coefficients for *Reference_NO_2* and *O_3_4* and a large offset (Table 1 & Figure S6). The cause of the poor calibration regression during this particular deployment is not known but may be due to transient errors in one or other of the Aeroqual monitor measurements in this period.

(a)



(b)



(c)

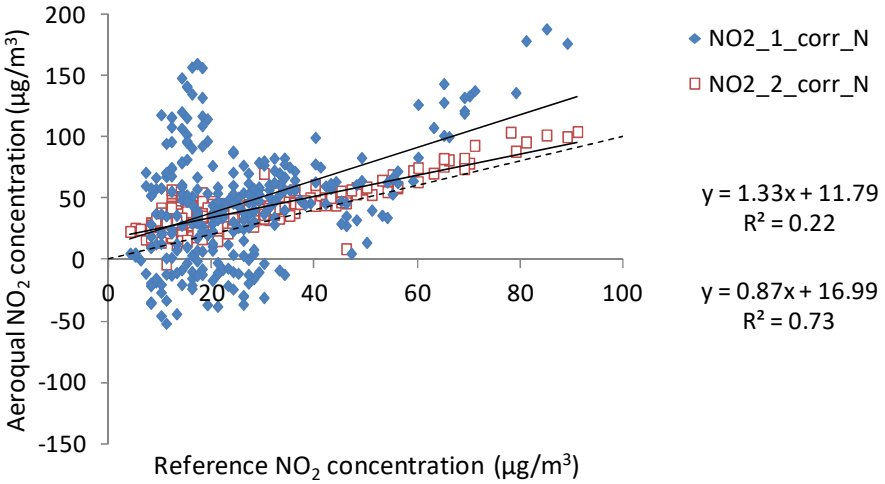


Figure 3: Scatter plots of adjusted hourly Aeroqual NO_2 concentration estimates *vs.* reference analyser NO_2 concentrations for the test data (i.e. 2nd half of deployment periods only). (a) Aeroqual data adjusted using calibration derived from training data in each unique deployment period (*aq_corr_u*). (b) Aeroqual data adjusted using calibration derived from training data combined from all deployment periods (*aq_corr_a*). (c) Aeroqual data adjusted using calibration derived from the first deployment period (November) only (*aq_corr_N*).

Table 2: OLS linear regression parameters (and 95% confidence intervals), coefficient of determination and summary statistics for adjusted Aeroqual NO_2 and O_3 concentrations compared to reference analyser concentrations. The calibration adjustments applied were: the second 48 h of each deployment period corrected using the unique calibration derived using the first 48 h of that deployment period (*Aq_corr_u*); the second 48 h of each deployment period corrected using the calibration derived when the first 48 h of data from all studies were combined (*Aq_corr_a*); and the second 48 h of each deployment period corrected using the calibration derived from the first 48 h in the first (November) study only (*Aq_corr_N*).

Monitor	Correction	Slope [95 % C.I.]	Intercept /µgm^{-3} [95 % C.I.]	R^2	RMSE / µg m^{-3}	MB / µg m^{-3}	NMB	***n***
O_3_3	*O_3_corr_u*	1.07 [1.06, 1.09]	-5.75 [-6.77, -4.72]	0.98	4.59	-2.10	-0.04	286
	O_3_corr_a	0.97 [0.94, 1.00]	-0.89 [-2.50, 0.72]	0.94	6.25	-2.17	-2.17	286
	O_3_corr_N	1.06 [1.03, 1.09]	-11.44 [-13.19, -9.68]	0.94	10.72	-8.54	-8.54	286
O_3_4	*O_3_corr_u*	0.93 [0.91, 0.95]	-2.14 [-3.20, -1.07]	0.97	6.94	-5.56	-0.11	286
	O_3_corr_a	0.91 [0.89, 0.93]	-1.46 [-2.57, -0.35]	0.96	7.22	-5.68	-0.12	286
	O_3_corr_N	0.96 [0.94, 0.99]	-9.24 [-10.41, -8.07]	0.96	11.80	-10.99	-0.23	286
NO_2_1	*NO_2_corr_u*	1.09 [0.98, 1.20]	1.25 [-2.06, 4.56]	0.59	15.75	3.65	0.14	286
	NO_2_corr_a	1.03 [0.98, 1.08]	3.60 [2.03, 5.17]	0.85	8.50	4.45	0.17	286
	NO_2_corr_N	1.33 [1.04, 1.63]	11.79 [2.62, 20.96]	0.22	47.23	20.47	0.78	286
NO_2_2	*NO_2_corr_u*	0.80 [0.58, 1.02]	11.02 [4.29, 17.75]	0.16	31.69	5.79	0.22	286
	NO_2_corr_a	0.97 [0.89, 1.04]	11.46 [9.24, 13.68]	0.71	14.73	10.59	0.41	286
	NO_2_corr_N	0.87 [0.80, 0.93]	16.97 [15.09, 18.88]	0.73	16.21	13.47	0.52	286

3.3. Temporal changes in Aeroqual O_3 and NO_2 monitor responses

We calculated the difference between unadjusted Aeroqual O_3 and Reference O_3 concentrations to assess drift in the output of the Aeroqual O_3 sensors during this study (Figure 4a). The difference generally declined over the 6 months of measurements. This decline may have resulted from deterioration in the sensitivity of the O_3 monitors as the sensors age. However, the calibration procedures overcome these effects to enable reasonably accurate representation of reference analyser concentrations.

We also plotted the difference between Aeroqual and Reference NO_2 concentrations for Aeroqual concentrations adjusted using *Aq_corr_u* and *Aq_corr_a* training data selections (Figures 4b and 4c). The residual concentrations varied more using the *Aq_corr_u* calibration method than for the *Aq_corr_a* method, with periods of large residual values corresponding to deployments with lower correlations between reference and Aeroqual NO_2 concentrations in the data (Table 1). The residual NO_2 concentrations showed no trend across the deployment periods for the *Aq_corr_u* correction method, as this calibration procedure partly corrects for the temporal drift in the Aeroqual O_3 sensors identified above, while for the *Aq_corr_a* correction method the residuals generally increase within each individual deployment and across all the deployments.

Daily averages of the hourly residuals for the Aeroqual O_3 and NO_2 measurements adjusted using the first set of training data (*Aq_corr_N*) were calculated for all of the available study data (including data for deployments > 96 h in length) (Figure S8). The daily O_3 residuals became more negative during each deployment; however at the start of some deployment the residuals appeared to revert to close to zero, as measured for the first residual values in the *Aq_corr_N* calibration period in November 2015. This may be indicative of changes in the O_3 sensor response after the monitor has been turned on, and of repeated instances of this effect for all of the individual deployments. The Aeroqual NO_2 sensors exhibit a general increase in the daily residuals during each deployment that is likely to have resulted from the use of the Aeroqual O_3 data in the calibration of the Aeroqual NO_2 sensor.

Overall, these observations suggest that it is beneficial to make calibration and field measurements after consistent time periods following monitor start-up (as was done here).

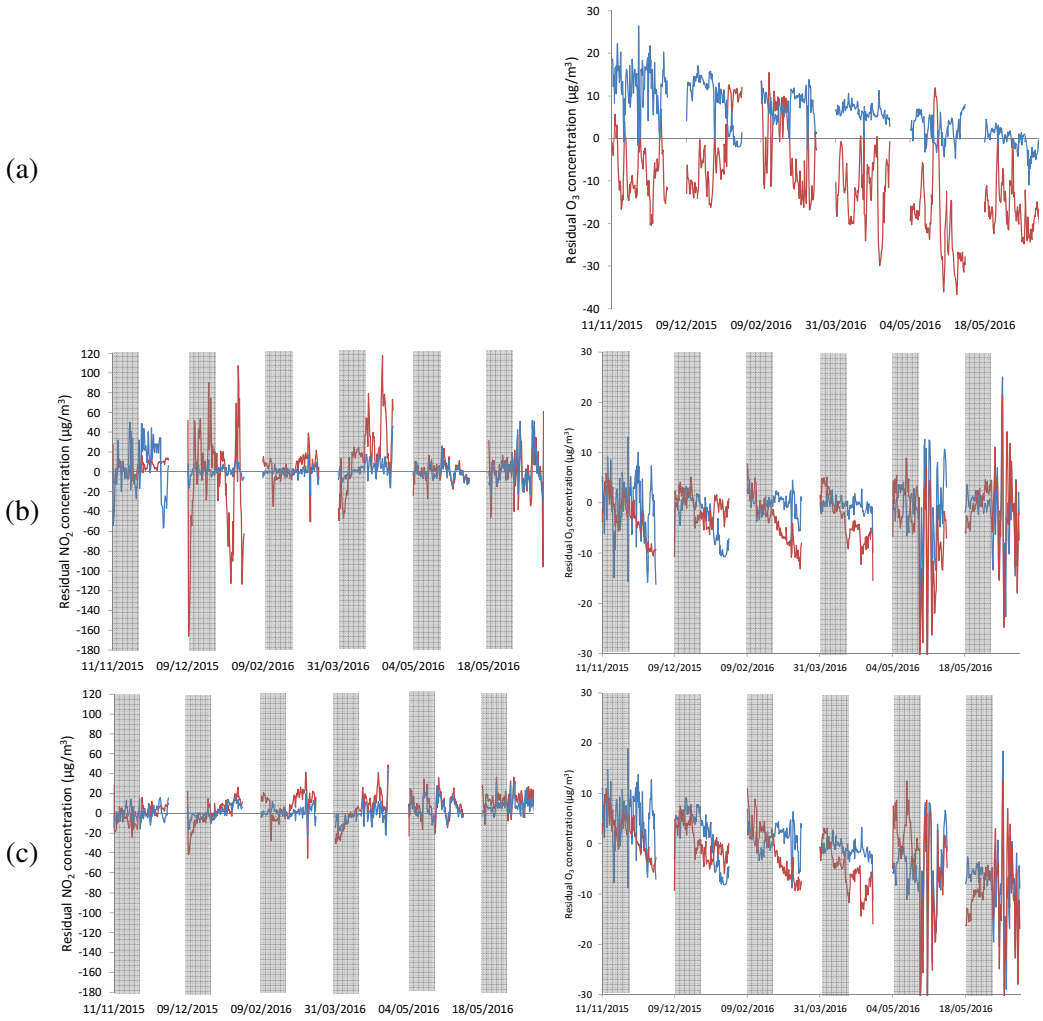


Figure 4: Residual plots of (a) unadjusted Aeroqual O_3 concentrations – reference O_3 concentrations; (b) *Aq_corr_u* – reference concentration and (c) *Aq_corr_a* – reference concentration over the 96-h deployments for NO_2 (left) and O_3 (right). Both training and test data sets are included (training data indicated by grey boxes). Different deployment periods are separated by gaps in time series. Residual plots of the unadjusted Aeroqual NO_2 concentrations are not provided due to the highly variable results which highlight the requirements for calibration.

4. Discussion

As a result of technological developments portable monitors are increasingly used to measure ambient pollution concentrations over extended geographical areas, and for mobile and personal monitoring. Field calibration procedures are necessary to improve and quantify the accuracy of estimates of ambient pollution concentrations from portable monitors. In this study we investigated changes in calibration relationships for two pairs of Aeroqual O_3 and NO_2 monitors *vs.* reference analysers over a period of 6 months as the sensors

approached the manufacturer's specified 2-year lifetime. We developed a multiple linear regression calibration method that allowed the Aeroqual NO_2 sensor to have a regression slope with 'true' NO_2 that is not equal to 1 (as was assumed in the method of Lin et al.[16]), due for example to sensor aging, in addition to allowance for a non-unity response to O_3. The inclusion of O_3 in the calibration of NO_2 concentrations measured by portable sensors has been previously reported in the literature – both O_3 and NO_2 invoke a similar response at the monitoring electrode and thus the sensor measures both gases simultaneously. Some studies have reported this effect to result in the sensor measuring the sum of ambient NO_2 and O_3 concentrations.[8,21,22,23] The use of multiple linear regression compared to linear regression to correct other portable NO_2 sensors has been observed to produce higher correlation between calibrated sensor estimates and reference NO_2 concentrations.[15]

Correction of the Aeroqual O_3 and NO_2 monitor estimates using each of the calibration methods tested in this work resulted in improved sensor accuracy and higher correlation with reference analyser measurements, compared to the unadjusted concentrations from the monitors. This highlights the importance of calibration of monitors in the sorts of urban pollution environments where they are to be used, as has been adopted in previous studies.[9,10,16] One recent study[20] used uncalibrated concentrations from the Aeroqual monitors to measure NO_2 and O_3 in indoor office environments with low O_3 concentrations which may have removed the need to adjust the Aeroqual NO_2 monitors for O_3 cross-sensitivity.

Our observations suggest that a single calibration study may not be able accurately to represent the relationship between monitor and reference analyser over an extended period, for example because of unidentified measurement problems during a particular calibration deployment period. In contrast, we have observed that combination of several short calibration deployments over a period of 6 months into a single calibration dataset provided adjusted Aeroqual concentration estimates that agreed more closely with the reference analyser concentrations, provided that for the NO_2 monitors concurrent O_3 monitor measurements were also used as input to the calibration relationship. We observed differences in Aeroqual NO_2 calibrations over shorter time periods (e.g. within 1 month), which may be attributable to the age of the sensors or differences in the pollutant composition at the sites during the calibration periods. However, during a winter pollution episode with elevated NO_2 concentrations none of the calibration methods tested for the O_3 monitor could accurately represent the analyser concentrations, highlighting the importance of (when possible) calibrating the Aeroqual monitors using pollution concentrations similar to those anticipated during intended field application.

In conclusion, our observations indicate that the benefit [in terms of improved accuracy of calibrated estimates] of calibration averaged by pooling data from a number of calibration periods over a period of a few months (increased size of calibration dataset and potentially increased concentration range) outweighed the disadvantage arising from any monitor drift throughout the overall measurement period. It is possible that the combination of several calibration periods interspersed within the overall measurement period mitigated the risk of only relying on calibration data from short periods when the range of concentrations may have been limited or otherwise unrepresentative. Therefore this pooling of interspersed calibration data is recommended as a pragmatic approach to calibration of Aeroqual sensor-based monitors for field deployment. Although the sensors used in this work were approaching the 2-year working lifetime specified by the manufacturer (Aeroqual, 2015 personal communication), our observations suggest that frequent intermittent calibrations as presented above would appear to allow the sensors to remain useable beyond this specified lifetime.

Supporting Information

The supporting information contains details of the study dates and site information, precision of Aeroqual O_3 monitors, discussion of justification of multiple linear regression approach to calibrate Aeroqual NO_2 monitor concentrations, time series plots of adjusted Aeroqual NO_2 and O_3 concentrations, scatter plots of Aeroqual vs. reference analyser NO_2 concentrations for each deployment period, calibration statistics for the methods, temporal change in calibration over the 6 month study period and investigation into daily NO_2 and O_3 calibrations derived using alternate days.

Acknowledgements

Nicola Masey is funded through a UK Natural Environment Research Council CASE PhD studentship (NE/K007319/1), with support from Ricardo Energy and Environment. Jonathan Gillespie is funded through an Engineering and Physical Sciences Research Council Doctoral Training Grant (EPSRC DTG EP/L505080/1 and EP/K503174/1) studentship, with support from the University of Strathclyde and Ricardo Energy and Environment. Eliani Ezani is funded by the Ministry of Higher Education Malaysia (KPT(BS)860126295394). Chun Lin is funded through NERC/Innovate UK grant NE/N007352/1. Hao Wu is funded by the University of Edinburgh and the NERC Centre for Ecology & Hydrology (NERC CEH project number NEC04544). We acknowledge access to the AURN measurement data, which were obtained from uk-air.defra.gov.uk and are subject

 The research data associated with this paper are available at: http://dx.doi.org/10.15129/a5820b14-367f-47f3-b0b9-beabcb347355.

References

(1) Gillespie, J.; Masey, N.; Heal, M. R.; Hamilton, S.; Beverland, I. J. Estimation of spatial patterns of urban air pollution over a 4-week period from repeated 5-min measurements. *Atmos. Environ.* **2017**, *150*, 295–302 DOI: 10.1016/j.atmosenv.2016.11.035.

(2) Gillespie, J.; Beverland, I. J.; Hamilton, S.; Padmanabhan, S. Development, Evaluation, and Comparison of Land Use Regression Modeling Methods to Estimate Residential Exposure to Nitrogen Dioxide in a Cohort Study. *Environ. Sci. Technol.* **2016**, *50* (20), 11085–11093 DOI: 10.1021/acs.est.6b02089.

(3) Masey, N.; Gillespie, J.; Heal, M. R.; Hamilton, S.; Beverland, I. J. Influence of wind-speed on short duration NO2 measurements using Palmes and Ogawa passive diffusion samplers. *Atmos. Environ.* **2017** DOI: http://doi.org/10.1016/j.atmosenv.2017.04.008.

(4) Heal, M. R.; O'Donoghue, M. A.; Cape, J. N. Overestimation of urban nitrogen dioxide by passive diffusion tubes: a comparative exposure and model study. *Atmos. Environ.* **1999**, *33* (4), 513–524 DOI: 10.1016/S1352-2310(98)00290-8.

(5) 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. The Changing Paradigm of Air Pollution Monitoring. *Environ. Sci. Technol.* **2013**, *47* (20), 11369–11377 DOI: 10.1021/es4022602.

(6) Bart, M.; Williams, D. E.; Ainslie, B.; McKendry, I.; Salmond, J.; Grange, S. K.; Alavi-Shoshtari, M.; Steyn, D.; Henshaw, G. S. High Density Ozone Monitoring Using Gas Sensitive Semi-Conductor Sensors in the Lower Fraser Valley, British Columbia. *Environ. Sci. Technol.* **2014**, *48* (7), 3970–3977 DOI: 10.1021/es404610t.

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

(8) 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.; et al. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmos. Environ.* **2013**, *70*, 186–203 DOI: 10.1016/j.atmosenv.2012.11.060.

(9) Wang, A.; Fallah-Shorshani, M.; Xu, J.; Hatzopoulou, M. Characterizing near-road air pollution using local-scale emission and dispersion models and validation against in-situ measurements. *Atmos. Environ.* **2016**, *142*, 452–464 DOI: 10.1016/j.atmosenv.2016.08.020.

(10) Deville Cavellin, L.; Weichenthal, S.; Tack, R.; Ragettli, M. S.; Smargiassi, A.; Hatzopoulou, M. Investigating the Use Of Portable Air Pollution Sensors to Capture the Spatial Variability Of Traffic-Related Air Pollution. *Environ. Sci. Technol.* **2016**, *50* (1), 313–320 DOI: 10.1021/acs.est.5b04235.

(11) Lewis, A.; Edwards, P. Validate personal air-pollution sensors. *Nature* **2016**, *535* (7610), 29–31 DOI: 10.1038/535029a.

(12) Lewis, A. C.; Lee, J. D.; Edwards, P. M.; Shaw, M. D.; Evans, M. J.; Moller, S. J.; Smith, K. R.; Buckley, J. W.; Ellis, M.; Gillot, S. R.; et al. Evaluating the performance of low cost chemical sensors for air pollution research. *Faraday Discuss.* **2016**, *189* (0), 85–103 DOI: 10.1039/C5FD00201J.

(13) Masson, N.; Piedrahita, R.; Hannigan, M. Quantification Method for Electrolytic Sensors in Long-Term Monitoring of Ambient Air Quality. *Sensors* **2015**, *15* (10), 27283–27302 DOI: 10.3390/s151027283.

(14) Pang, X.; Shaw, M. D.; Lewis, A. C.; Carpenter, L. J.; Batchellier, T. Electrochemical ozone sensors: A miniaturised alternative for ozone measurements in laboratory experiments and air-quality monitoring. *Sens. Actuators B Chem.* **2017**, *240*, 829–837 DOI: 10.1016/j.snb.2016.09.020.

(15) Spinelle, L.; Gerboles, M.; Villani, M. G.; Aleixandre, M.; Bonavitacola, F. Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide. *Sens. Actuators B Chem.* **2015**, *215*, 249–257 DOI: 10.1016/j.snb.2015.03.031.

(16) Lin, C.; Gillespie, J.; Schuder, M. D.; Duberstein, W.; Beverland, I. J.; Heal, M. R. Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide. *Atmos. Environ.* **2015**, *100*, 111–116.

(17) Duvall, R. M.; Long, R. W.; Beaver, M. R.; Kronmiller, K. G.; Wheeler, M. L.; Szykman, J. J. Performance Evaluation and Community Application of Low-Cost Sensors for Ozone and Nitrogen Dioxide. *Sensors* **2016**, *16* (10), 1698 DOI: 10.3390/s16101698.

(18) Jiao, W.; Hagler, G.; Williams, R.; Sharpe, R.; Brown, R.; Garver, D.; Judge, R.; Caudill, M.; Rickard, J.; Davis, M.; et al. Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. *Atmos Meas Tech* **2016**, *9* (11), 5281–5292 DOI: 10.5194/amt-9-5281-2016.

(19) DEFRA. The Data Verification and Ratification Process https://uk-air.defra.gov.uk/assets/documents/The_Data_Verification_and_Ratification_Process.pdf (accessed Apr 1, 2017).

(20) Szigeti, T.; Dunster, C.; Cattaneo, A.; Spinazzè, A.; Mandin, C.; Le Ponner, E.; de Oliveira Fernandes, E.; Ventura, G.; Saraga, D. E.; Sakellaris, I. A.; et al. Spatial and temporal variation of particulate matter characteristics within office buildings — The OFFICAIR study. *Sci. Total Environ.* **2017**, *587–588*, 59–67 DOI: 10.1016/j.scitotenv.2017.01.013.

(21) Hossain, M.; Saffell, J.; Baron, R. Differentiating NO_2 and O3 at low cost air quality amperometric gas sensors. *ACS Sensors*. **2016**, 1(11), *1291 – 1294*. DOI:10.1021/acssensors.6b00603

(22) Williams, D.E.; Salmon, J.; Yung, Y.F.; Akaji, J.;, Wright, B.; Wilson, J.; Henshaw, G.S.; Brett Wells, D.; Ding, G.; Wagner, J.; Laing, G. Development of low cost ozone and nitrogen dioxide measurement instruments suitable use in an air quality monitoring network. *IEEE Sensors*. **2009**, vols 1 – 3. DOI: 10.1109/ICSENS.2009.5398568

(23) Spinelle, L.; Gerboles, M.; Aleixandre, M. Performance evaluation of amperometric sensors for the monitoring of O3 and NO_2 in ambient air at ppb level. *Procedia Eng*. **2015**, 120, 480 – 483. DOI:10.1016/j.proeng.2015.08.676

Supporting Information

Temporal changes in field calibration relationships for Aeroqual S500 O_3 and NO_2 sensor-based monitors

Nicola Masey[1], Jonathan Gillespie[1], Eliani Ezani[1,2], Chun Lin[3], Hao Wu[3], Neil S Ferguson[1], Scott Hamilton[4], Mathew R. Heal[3], Iain J. Beverland[1]*

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]Department of Environmental and Occupational Health, Faculty of Medicine and Health Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, MALAYSIA

[3]School of Chemistry, Joseph Black Building, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

[4]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

Contents:

Pages: 21

Figures: 13

Tables**:** 3

1. Deployment details

The four Aeroqual S500 monitors (two NO_2, two O_3) were deployed on the roof of the Defra (Department for the Environment, Food and Rural Affairs) Automatic Urban and Rural Network monitoring station at Townhead, Glasgow (Figure S1). Monitors *NO_2_1* and *O_3_3* were located in the enclosures on the right hand (eastern) side of the roof. Monitors *NO_2_2* and *O_3_4* were located in the enclosures on the left hand (western) side of the roof. The dates of the six deployments at Townhead are given in Table S1. The site is classified as urban background.

Figure S1: Deployment of Aeroqual monitors on the roof of Glasgow Townhead AURN station. The monitors were deployed within ventilated waterproof enclosures provided by Aeroqual attached to the galvanised steel safety railings surrounding the roof of the AURN monitoring cabin.

Table S1: Dates of 96-hour deployments of Aeroqual monitors at Glasgow Townhead AURN site and the range concentrations of NO_2 and O_3 measured by the automatic analysers located at Townhead during the training (first 48 hour) and test (second 48 hour) periods.

Study Name	Dates of Study	Range of measured NO_2 concentrations Training [Test] ($\mu g/m^3$)	Range of measured O_3 concentrations Training [Test] ($\mu g/m^3$)
November	11/11/2015 – 15/11/2015	4 – 41 [8 – 49]	25 – 72 [12 – 74]
December	09/12/2015 – 13/12/2015	4 – 24 [25 – 91]	34 – 66 [1 – 46]
February	09/02/2016 - 15/02/2016	13 – 74 [15 – 79]	5 – 56 [4 – 62]
April	31/03/2016 – 04/04/2016	8 – 37 [9 – 45]	24 – 69 [22 – 80]
May1	04/05/2016 – 08/05/2016	6 – 61 [7 – 35]	3 – 86 [41 – 95]
May2	18/05/2016 – 22/05/2016	7 – 54 [4 – 21]	12 – 69 [38 – 80]

2. Aeroqual O_3 calibration

The observations from the two Aeroqual O_3 monitors were highly correlated ($R^2 > 0.90$) but with *O_3_4* reporting concentrations approximately half the magnitude of those measured by *O_3_3* and with a significant positive offset (Supporting Information Figure S2).

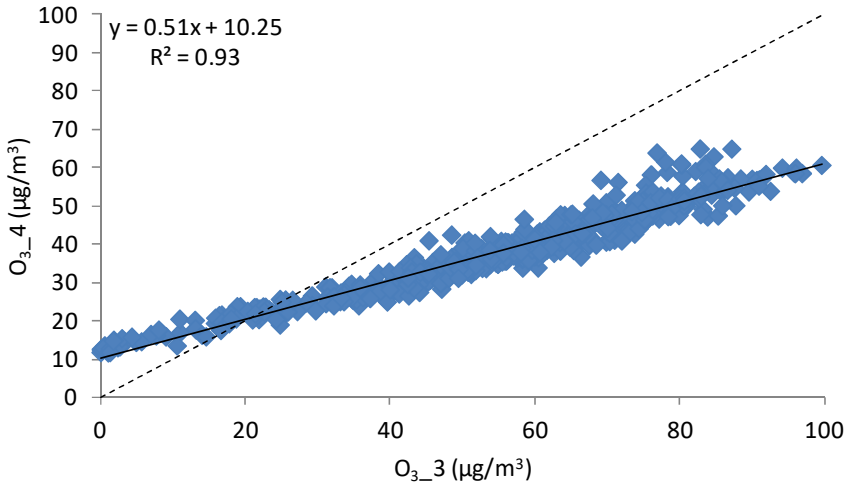


Figure S2: Scatter plot of unadjusted hourly data from the two O_3 Aeroqual monitors.

Scatter plots of hourly Aeroqual O_3 vs. reference analyser O_3, using the first 48 h of data in each deployment period only (training data), were used to generate calibration equations to adjusted the Aeroqual O_3 concentrations. The equations derived for the three calibration methods tested in this work (and described in the main paper) are shown in Figure S3.

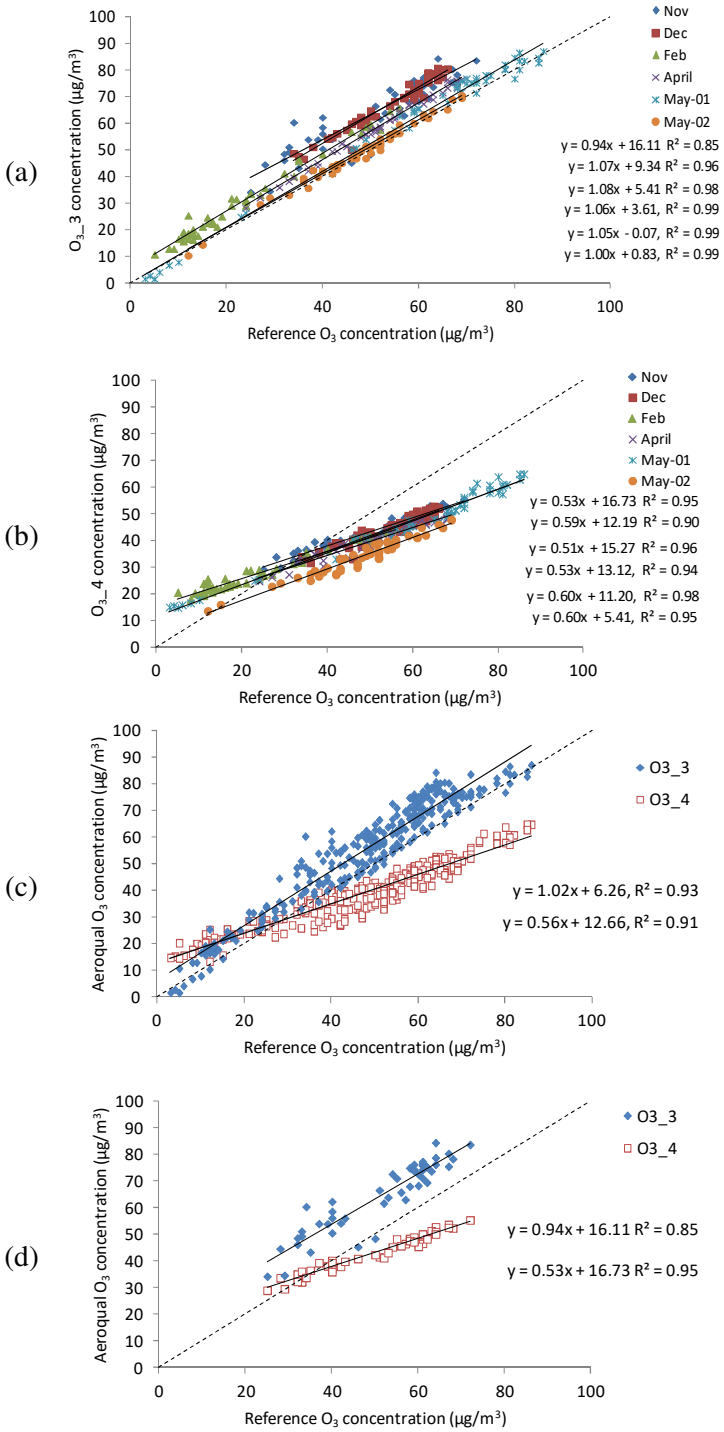


Figure S3: Scatter plots of hourly Aeroqual O_3 vs. reference analyser O_3: (a) and (b) calibrations using training data in each unique deployment period (*O_3_aq_corr_u*) (panels (a) and (b) show data for Aeroqual monitors *O_3_3* and *O_3_4* separately); (c) calibrations using training data combined from all deployment periods (*O_3_aq_corr_a*), for monitors *O_3_3* and *O_3_4* separately; (d) calibrations using the first deployment period (November) only (*O_3_aq_corr_N*), for monitors *O_3_3* and *O_3_4* separately.

The calibration equations shown in Figure S3 were used to adjust the test data (second 48 h period in the deployments). Figure S4 shows time series plots for the adjusted Aeroqual O_3 (*aq_corr_u*, *aq_corr_a* and *aq_corr_N*) and reference analyser O_3 concentrations.

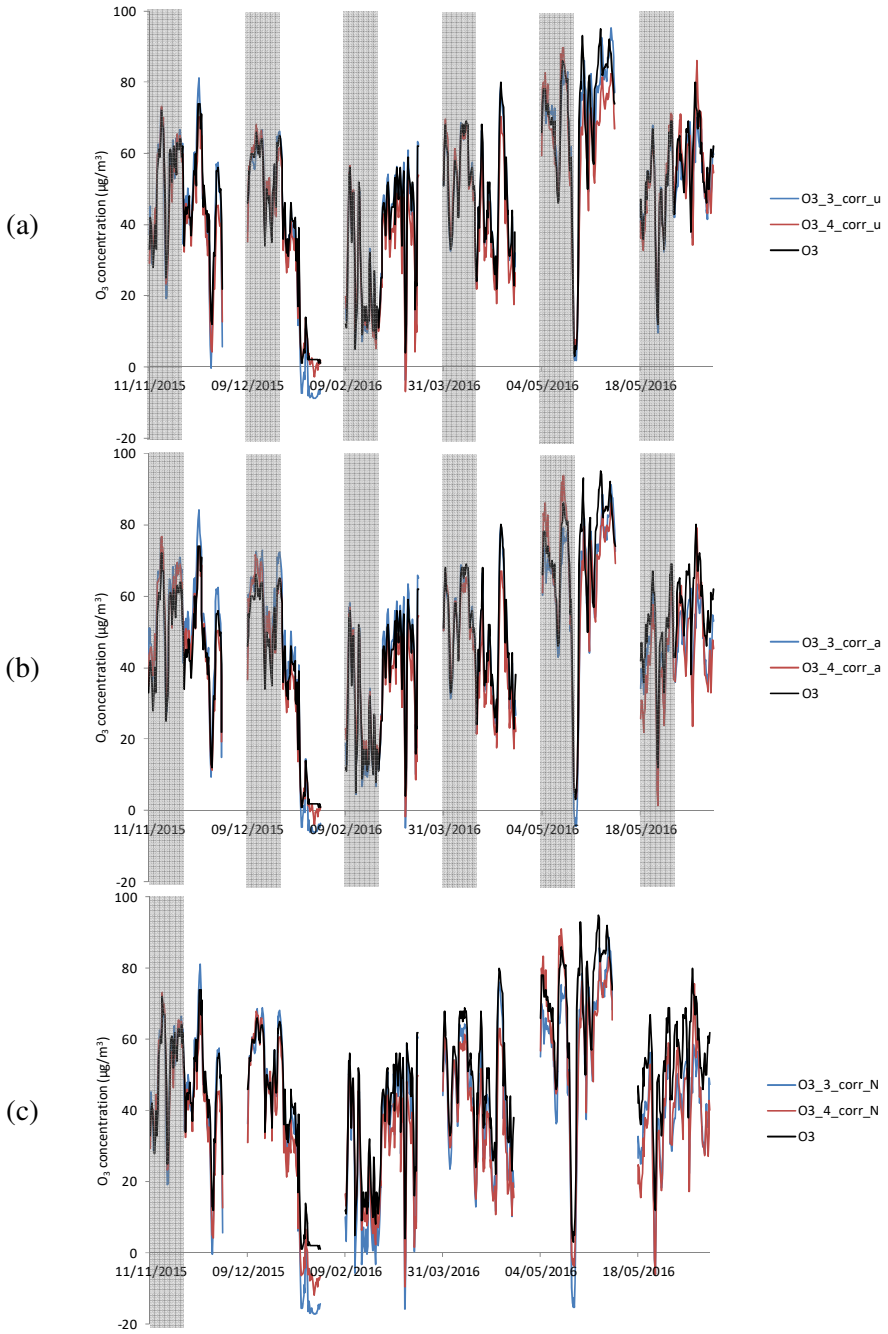


Figure S4: Time series of adjusted Aeroqual O_3 concentrations and reference analyser O_3 concentrations using: (a) training data from each unique deployment period (*aq_corr_u*); (b) training data from all deployment periods (*aq_corr_a*); (c) training data from the first (November) deployment period (*aq_corr_N*). Both training and test data sets are included (training data indicated by grey boxes). Different deployment periods are separated by gaps in time series.

3. Aeroqual NO_2 calibrations

Time series for adjusted Aeroqual concentrations and reference analyser NO_2 concentrations are shown in Figure S5 for each of the three methods of using the test datasets that were tested (*aq_corr_u*, *aq_corr_a* and *aq_corr_N*).

The time series in Figure S5 show there are periods where the adjusted Aeroqual NO_2 concentrations show deviations from those concentrations measured by the reference analyser. Scatter plots of adjusted hourly Aeroqual NO_2 concentration estimates *vs.* reference analyser NO_2 concentrations for the test data (i.e. 2nd half of deployment periods only) from each deployment show these deviations are more pronounced in the concentrations adjusted using *aq_corr_u* (Figure S6).

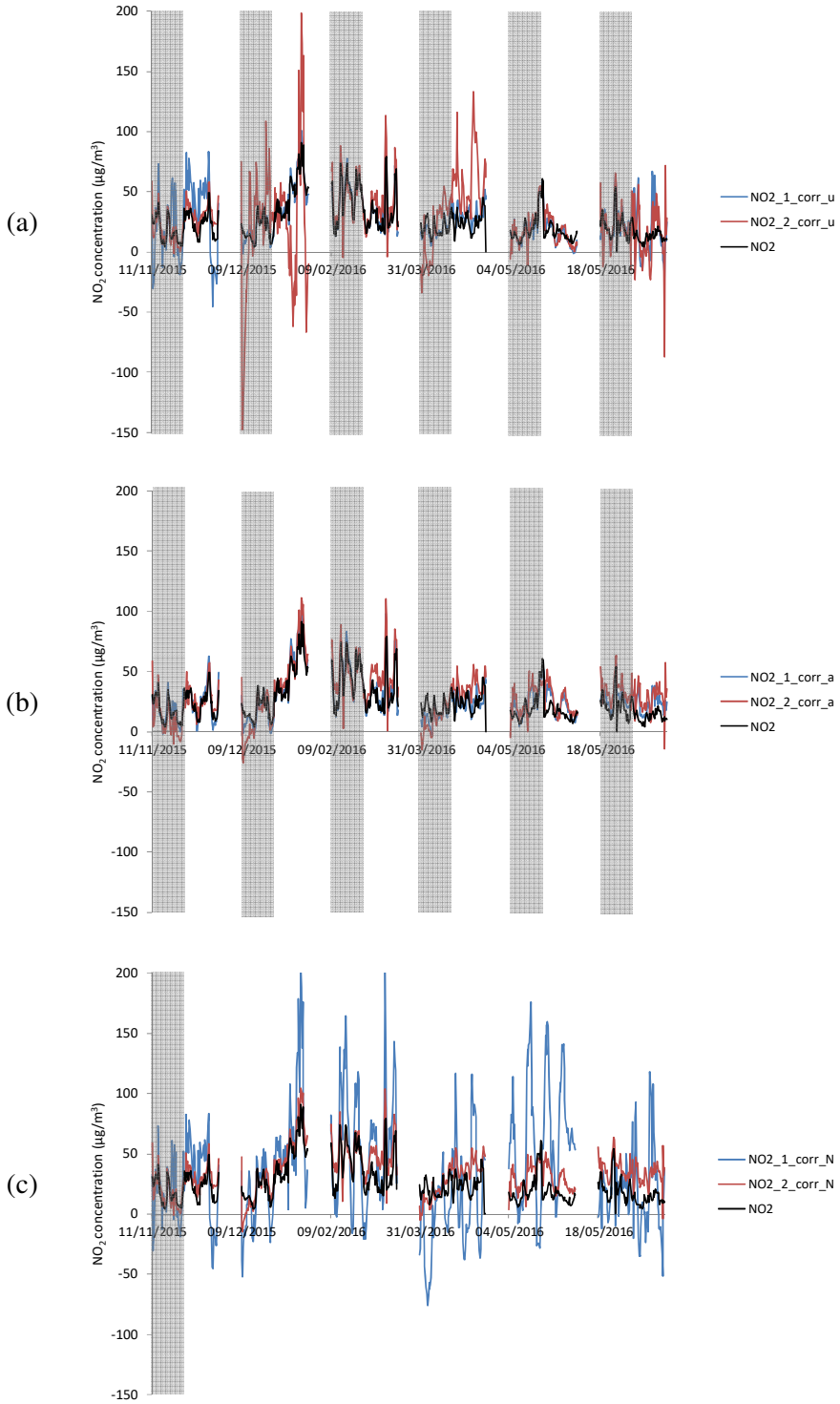


Figure S5: Time series of adjusted hourly Aeroqual NO_2 concentrations and reference analyser NO_2 concentrations using the following calibrations: (a) training data from each unique deployment period (*aq_corr_u*); (b) training data from all deployment periods (*aq_corr_a*); (c) training data from the first (November) deployment period (*aq_corr_N*). Both training and test data sets are included (training data indicated by grey boxes). Different deployment periods are separated by gaps in time series.

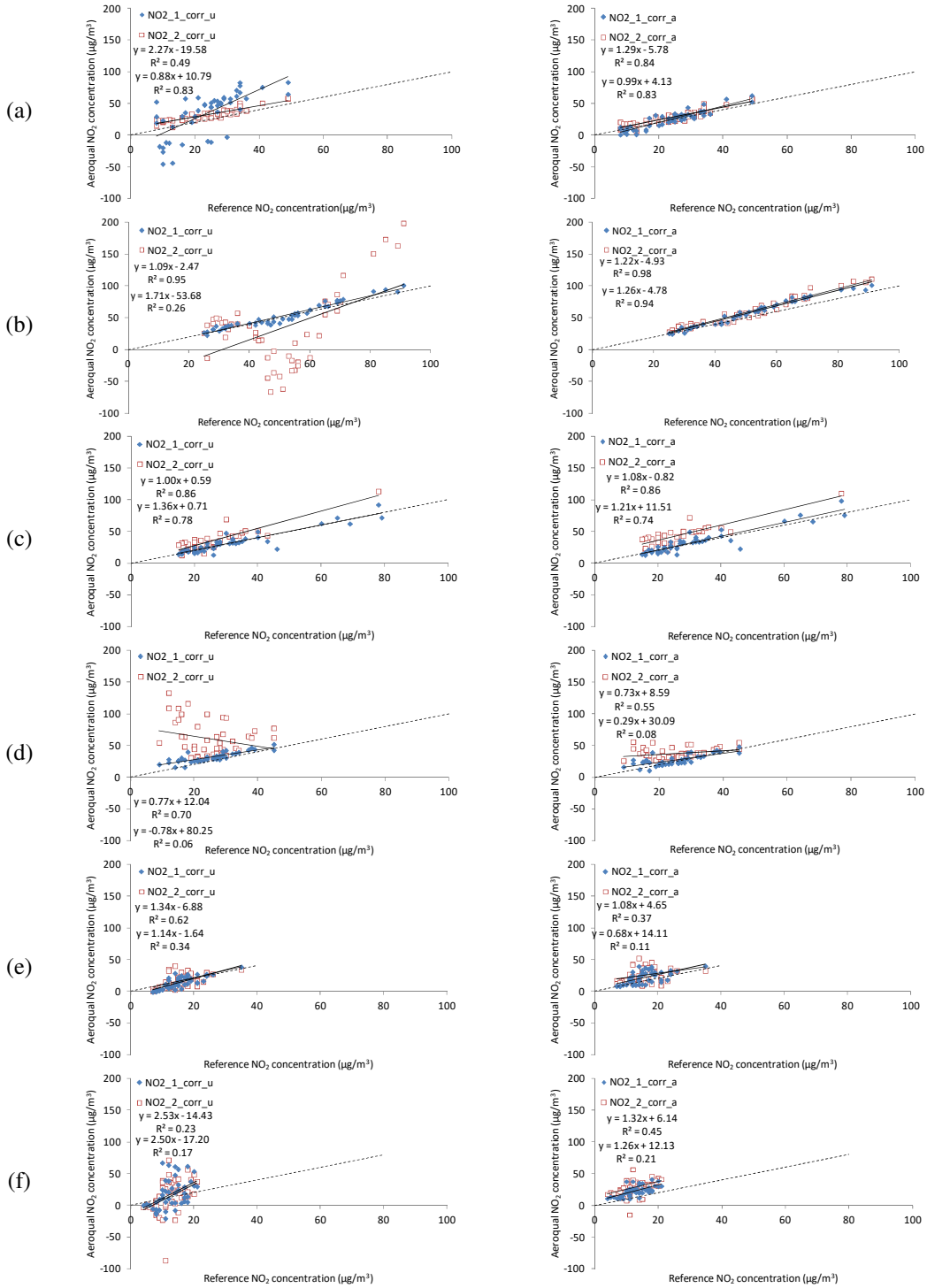


Figure S6: Scatter plots of adjusted hourly Aeroqual NO_2 concentration estimates *vs.* reference analyser NO_2 concentrations for the test data (i.e. 2nd half of deployment periods only). Left hand plots show Aeroqual data adjusted using calibration derived from training data in each unique deployment period (*aq_corr_u*), while right hand plots show Aeroqual data adjusted using calibration derived from training data combined from all deployment periods (*aq_corr_a*). Plots (a) to (f) represent the six deployment periods: November, December, February, April, May1 and May2 respectively.

4. Calibration statistics

The normalised mean bias (NMB) and coefficient of determination (R^2) calculated for the adjusted data for each test period and each Aeroqual monitor were plotted to examine if there were trends (Figure S7).

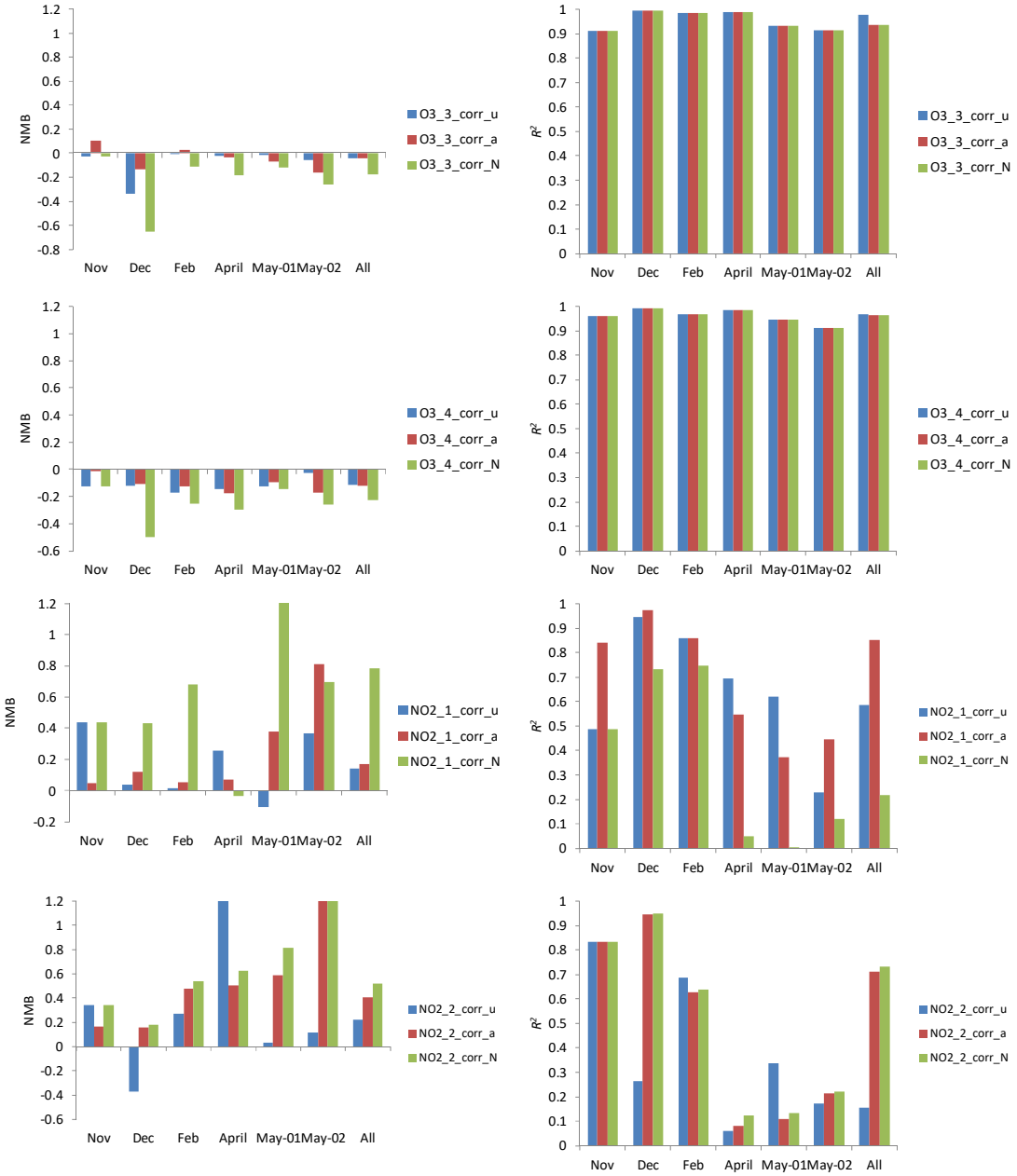


Figure S7: Plots of NMB (left) and R^2 (right) values for each of the test deployment periods for each Aeroqual monitor (top to bottom: O_3_3, O_3_4, NO_2_1 and NO_2_2) against their respective reference analyser.

5. Temporal changes in calibration over 6 months

To examine the long-term trend in validity of a calibration from the time the underlying data were collected, we plotted the difference between Aeroqual and reference O_3 and NO_2 concentrations, the former adjusted using *aq_corr_N* (Figure S8).

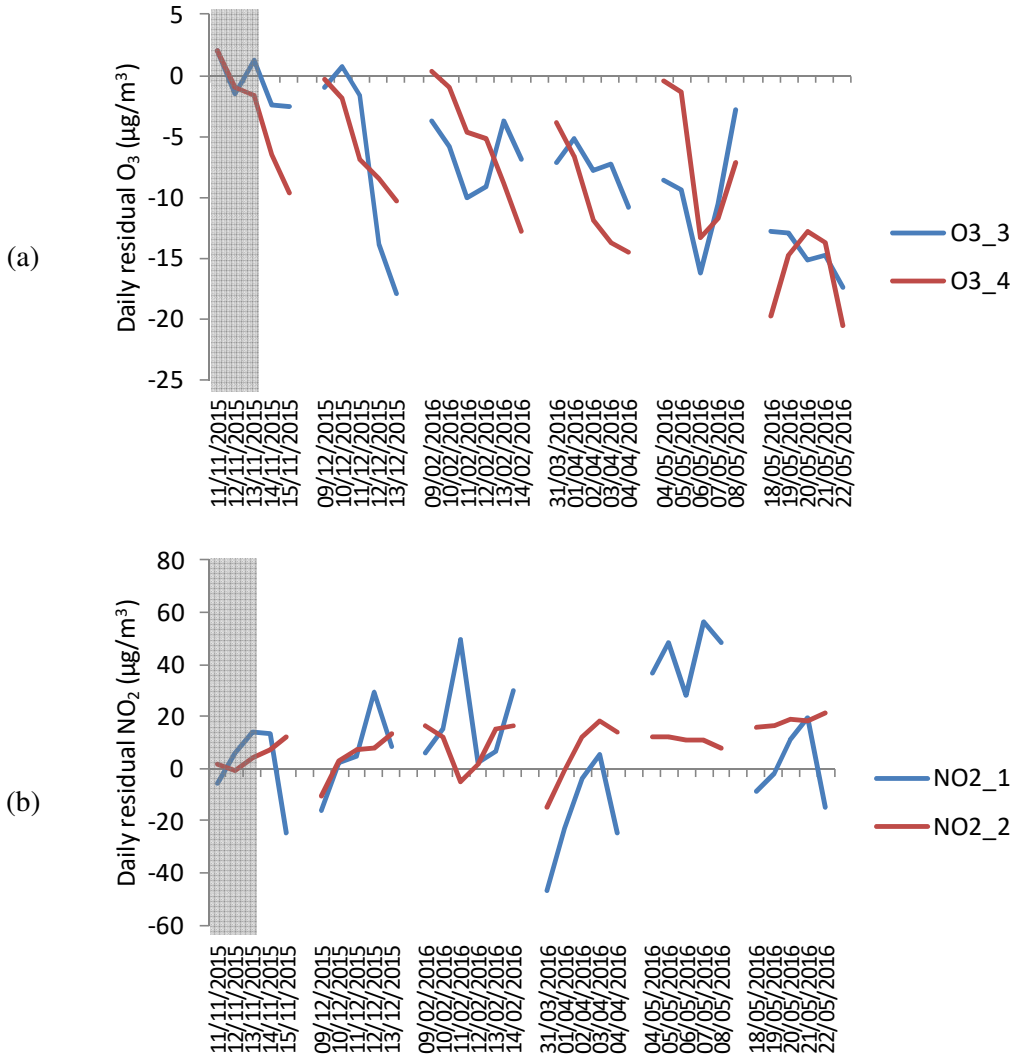


Figure S8: Daily average difference between hourly *Aq_corr_N* adjusted Aeroqual monitor and reference analyser concentrations for each measurement day for (a) O_3 and (b) NO_2. The *Aq_corr_N* calibration period (November) is indicated by grey boxes. All data measured for each deployment period, before truncation to the first 96 h, are shown. Different deployment periods are separated by gaps in time series.

6. Daily NO_2 and O_3 calibrations derived using alternate days

We investigated the effect of short-term (< 2 weeks) temporal drift on the calibration of the Aeroqual monitors by applying 24-h calibration procedures in the first of the deployment periods in May 2016. This deployment period, which extended over 10 days of measurements, was the longest in the study. The 24-h calibration calculations located test and training data close in time to one another, minimising the extent of drift in monitor responses over time, to establish if calibrations improved when the smallest practical length of time between calibration and measurements is used. However the smaller number of data points (24) and reduced concentration ranges may limit the performance of the calibration and consequently result in poorer corrected data.

The observations were split into training and test datasets, with alternative days assigned to each category (training data = 5, 7, 9, 11, and 13 May; test data = 6, 8, 10, 12 and 14 May). Three calibration methods were investigated: (a) individual test days adjusted using calibrations computed from the preceding training day i.e. alternative day correction (*Aq_corr_alt*); (b) all test days adjusted using calibrations computed from training data from all training days combined (*Aq_corr_all_alt*); and (c) all test days adjusted using calibrations computed from the first day of training data [5 May 2016] (*Aq_corr_first*). The regression equations calculated for each of these calibration procedures (Figure S9 and Table S2) were different from the 48 h calibrations for the first deployment period in May 2016 (Figure S3 and main article Table 1), with the 24-h calibrations generally having smaller slope and R^2 values.

We noted similar temporal variations in adjusted Aeroqual and Reference O_3 concentrations for all three calibration methods (Figures S10). However there were periods when the adjusted Aeroqual O_3 concentrations deviated from the reference analyser concentration. These deviations were most apparent when the O_3 data was adjusted using the *Aq_corr_first* training data selection method (Figure S11). The time series for adjusted Aeroqual NO_2 concentrations showed deviations from the reference analyser, the most pronounced of these was during the 9-10th May 2016 which corresponded to a period with calibration equations with low coefficients of determination (Figure S12, Table S2).

The highest R^2 values for the comparisons of *adjusted Aeroqual O_3 vs. reference O_3* concentrations occurred for the *Aq_corr_all_alt* and *Aq_corr_first* calibration methods (Figure 11, Table S3). The *Aq_corr_all_alt* method resulted in adjusted Aeroqual O_3 concentrations that underestimated analyser concentrations at low concentrations (Figure S11) with a regression slope and intercept significantly different from 1 and 0 respectively (Table S3). The *Aq_corr_first* method produced a regression line closer to 1:1 and the

smallest RMSE values (Figure S11, Table S3). The potential benefit of using shorter (*Aq_corr_alt*) calibrations to address the issue of monitor drift appeared to have been outweighed by fewer data and narrower concentration ranges, which hampered interpretation of the overall patterns in these data.

The scatter plots of *adjusted NO_2* test data *vs. reference analyser NO_2* were relatively consistent between the *Aq_corr_all_alt* and *Aq_corr_first* calibration methods (Figure S12). The large positive (> 200 µg/m^3) and negative (< 200 µg/m^3) *NO_2_2_corr_alt* concentrations, the result of the data corrected using the 9th May calibration, contribute to the slope value less than one, large negative intercept and correlation of 0 observed in this calibration. For both Aeroqual NO_2 monitors the RMSE values were highest for the *NO_2_aq_corr_alt* calibration while the *NO_2_aq_corr_all_alt* and *NO_2_aq_corr_first* calibrations RMSE values were similar to one another (range: 13.7 – 16.2 µg/m^3 (Table S3)).

It would be laborious to make calibration measurements the day before every set of field measurements, especially for a long-term trial, and our observations suggest that, for a 10-day study such as this, there is no benefit to making and applying calibrations on every alternate day. However, the combination of several calibration periods interspersed within the overall measurement period (as outlined in section 3.1 of the main article) would appear to mitigate the risk of only relying on calibration data from a short period when the range of concentrations may be limited or otherwise unrepresentative.

It was difficult to discern trends in the daily residual plots between Aeroqual and reference analyser concentrations for O_3 and NO_2, (Figure S13) indicating that on the short time scales of a single deployment period the apparent drift of the sensor response may sometimes be less evident.

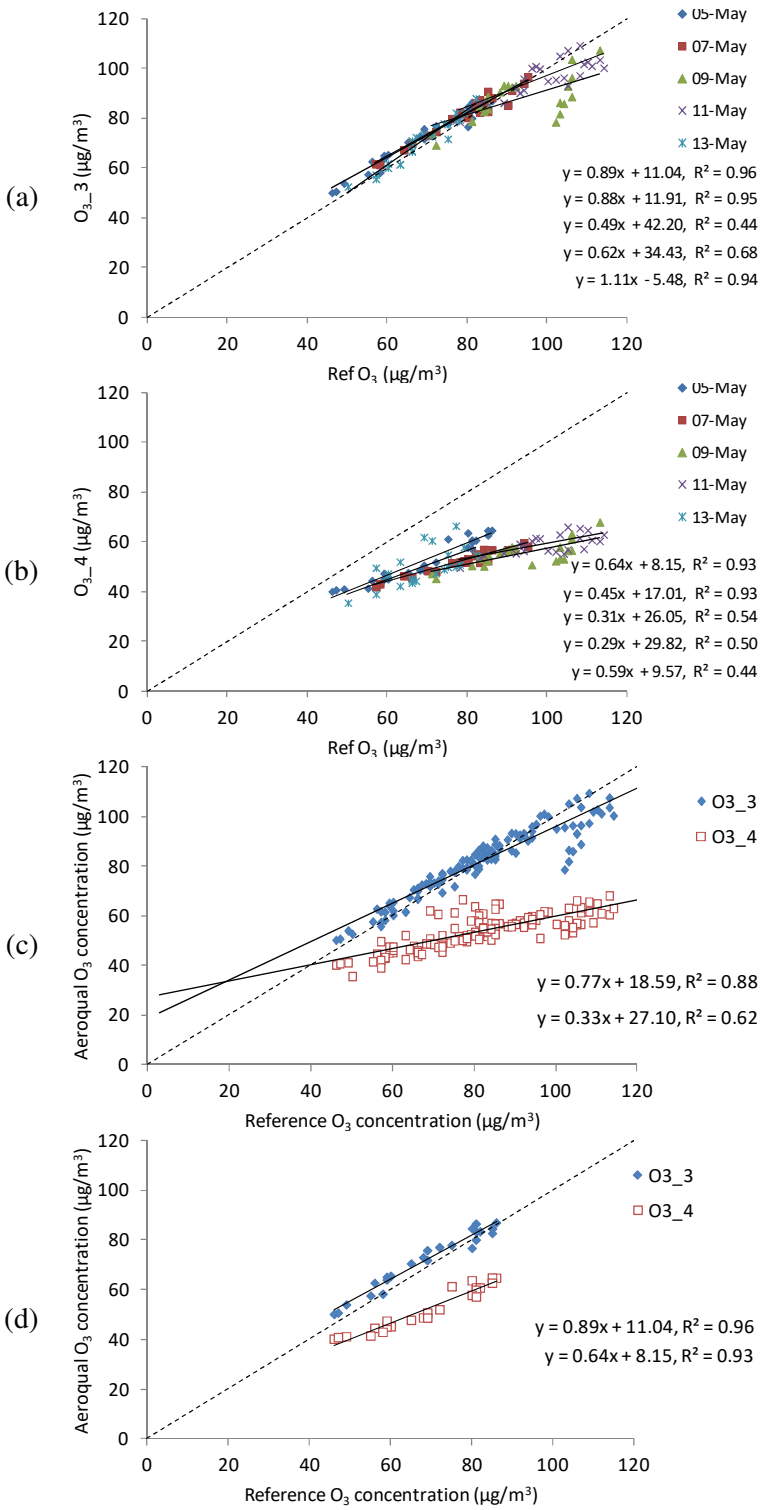
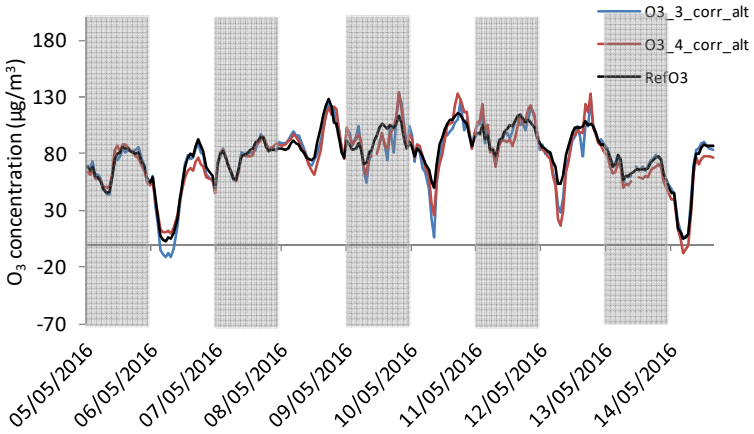
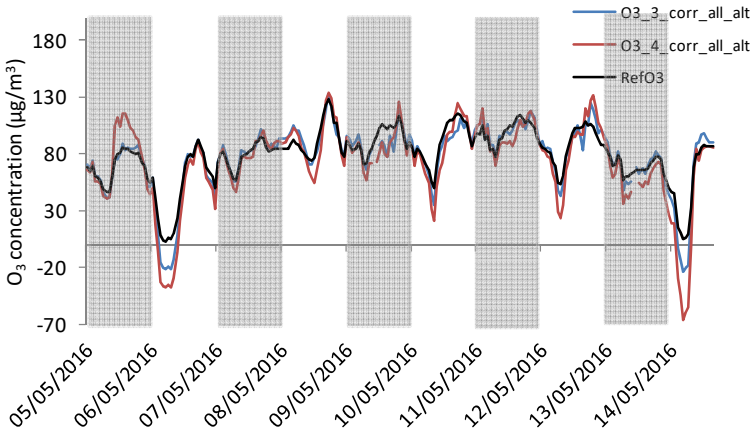


Figure S9: Aeroqual O_3 vs. reference analyser O_3 calibration lines calculated from the first study period in May: (a) alternate days used as training data (O_3_3); (b) alternate days used as training data (O_3_4); (c) all alternate days training data combined; (d) training data from first alternate day only.

(a)



(b)



(c)

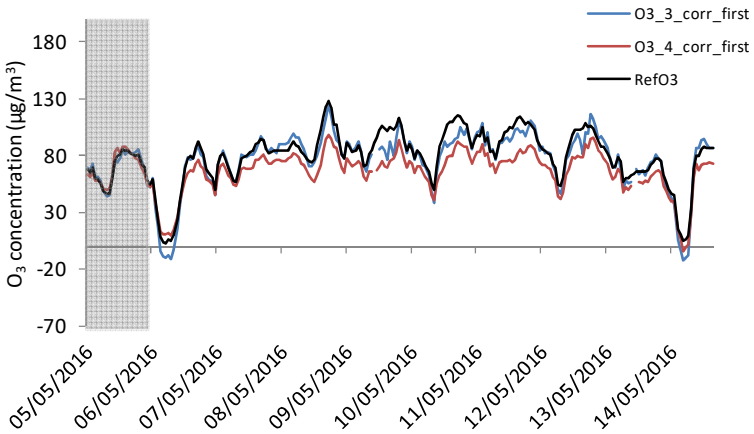


Figure S10: Time series, from the first study in May, of adjusted Aeroqual O_3 concentrations and reference analyser O_3 concentrations using the following calibrations: (a) training data from alternate days training data (*aq_corr_alt*); (b) training data from all alternate days training data (*aq_corr_all_alt*); (c) training data from first alternate day training data (*aq_corr_first*). Both training and test data sets are included (training data indicated by grey boxes).

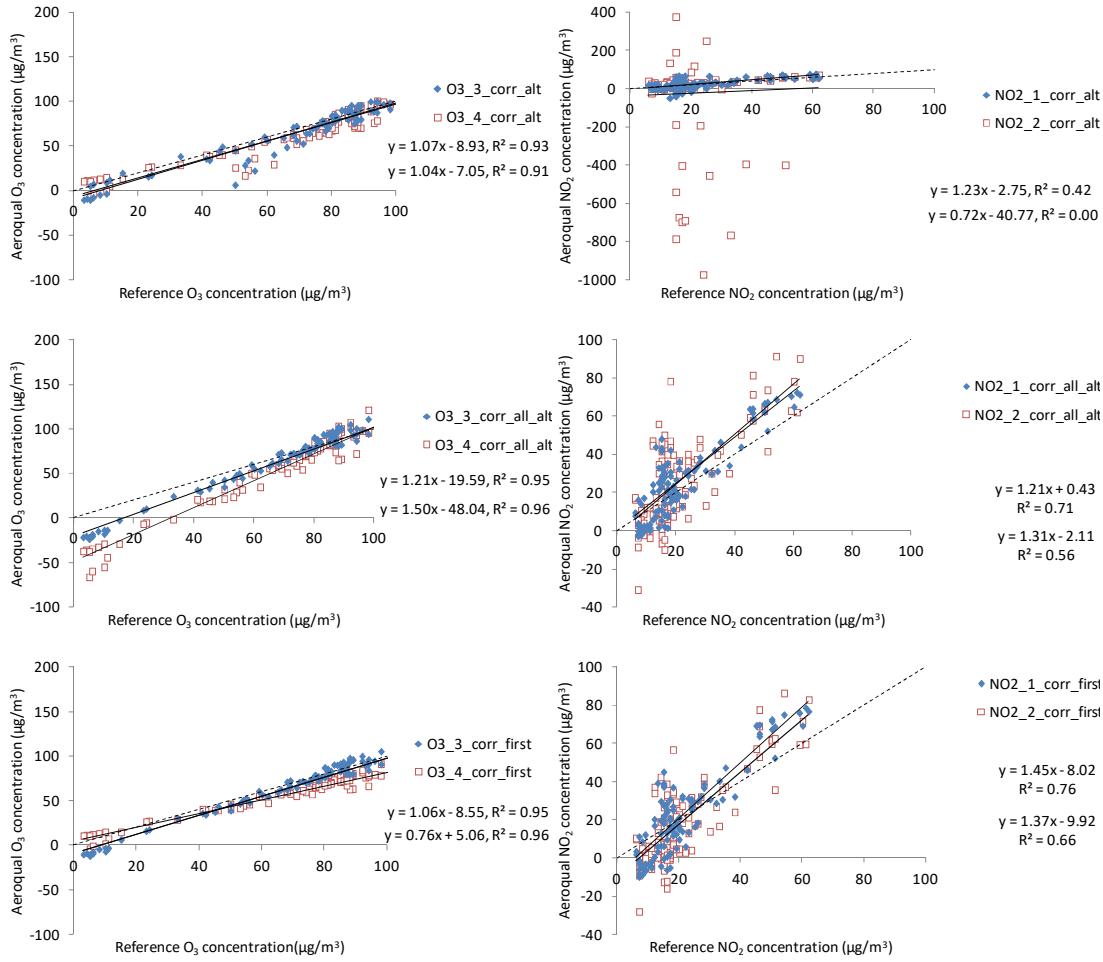


Figure S11: Scatter plots for short-term adjusted *Aeroqual O_3 vs. reference analyser O_3* concentrations (left) and adjusted *Aeroqual NO_2 vs. reference analyser NO_2* concentrations (right) using three calibration procedures (*aq_corr_alt*, *aq_corr_all_alt*, *aq_corr_first*). Only the test data (for 6, 8, 10, 12 and 14 May 2016) is shown in these Figures.

(a)

(b)

(c)

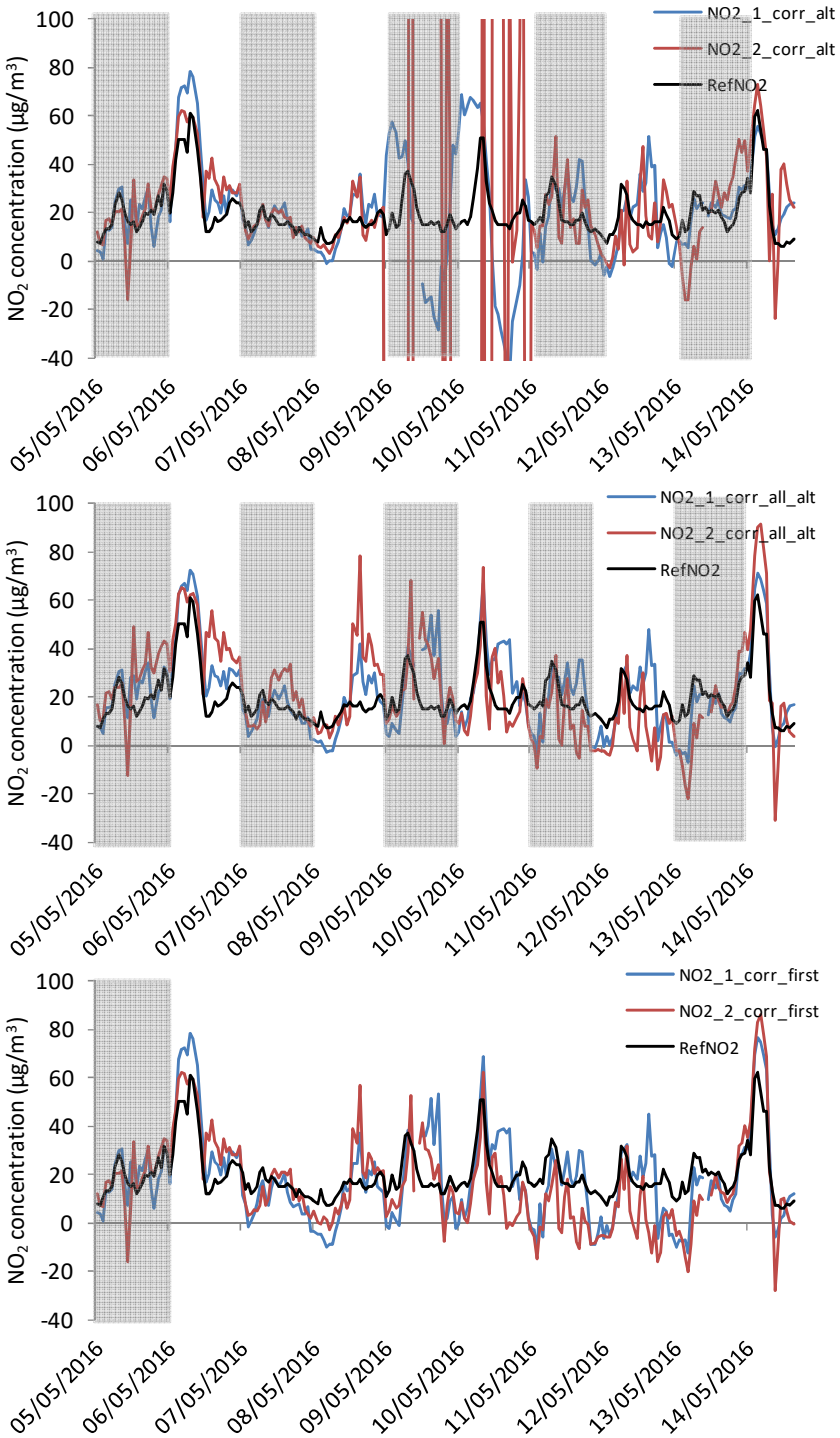
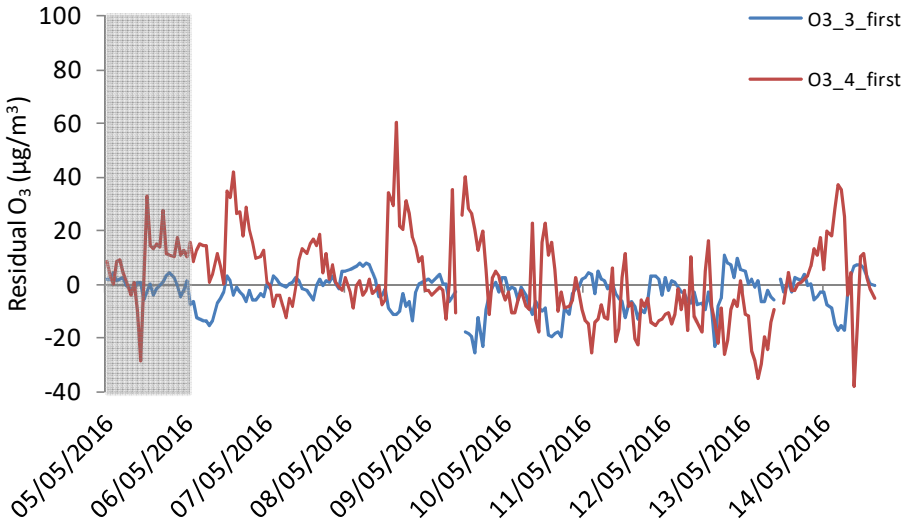


Figure S12: Time series, from the first period in May, of adjusted Aeroqual NO_2 concentrations and reference analyser NO_2 concentrations using the following calibrations: (a) training data from alternate days training data (*aq_corr_alt*); (b) training data from all alternate days training data (*aq_corr_all_alt*); (c) training data from first alternate day training data (*aq_corr_first*). Both training and test data sets are included (training data indicated by grey boxes).

(a)



(b)

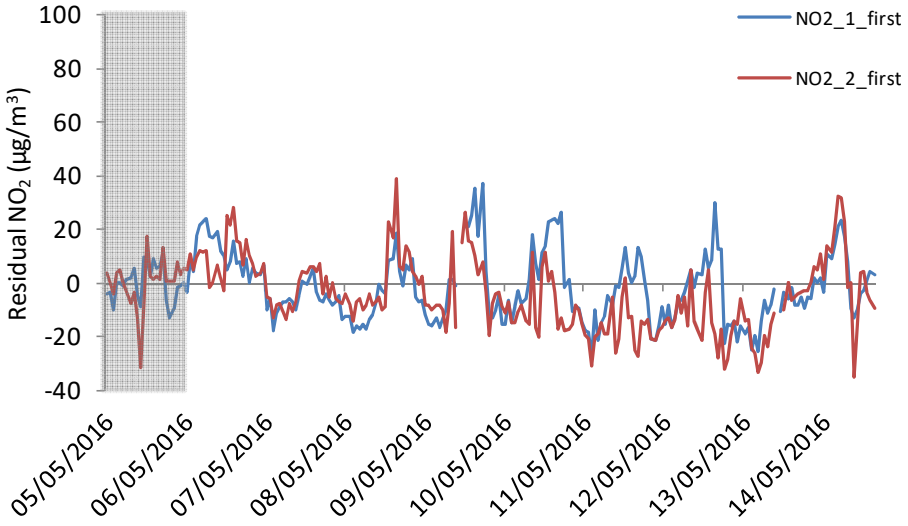


Figure S13: Time series of hourly differences between Aeroqual monitor concentration after correction using *Aq_corr_first* and respective reference analyser concentration for (a) O_3 and (b) NO_2 (training periods highlighted in grey).

Table S2: OLS equations for calibration equations for alternate 24-h calibration equations for the training periods for the short-term study using data from the first May study (*aq_corr_alt*) and a combination of these alternate days (*aq_corr_all_alt*) for the two pairs of Aeroqual O_3 and NO_2 monitors.

Pollutant	Study period	O_3_3	O_3_4
O_3			
	05/05/16	0.89*Ref_O_3 + 11.04, R^2= 0.96	0.64*Ref_O_3 + 8.12, R^2= 0.93
	07/05/16	0.88*Ref_O_3 + 11.91, R^2= 0.95	0.45*Ref_O_3 + 17.01, R^2= 0.93
	09/05/16	0.49*Ref_O_3 + 42.20, R^2= 0.44	0.31*Ref_O_3 + 26.05, R^2= 0.54
	11/05/16	0.62*Ref_O_3 + 34.43, R^2= 0.68	0.29*Ref_O_3 + 29.82, R^2= 0.50
	13/05/16	1.11*Ref_O_3 – 5.48, R^2= 0.94	0.59*Ref_O_3 + 9.57, R^2= 0.44
	O_3_corr_all_alt	0.77*Ref_O_3 + 18.59, R^2= 0.88	0.33*Ref_O_3 + 27.10, R^2= 0.62
	Study period	NO_2_1	NO_2_2
NO_2			
	05/05/16	0.44*Ref_NO_2 + 0.64*O_3_3 – 26.50, R^2 = 0.84	0.34*Ref_NO_2 + 1.08*O_3_4 + 47.05, R^2 = 0.88
	07/05/16	0.56*Ref_NO_2 + 0.67*O_3_3 – 33.65, R^2 = 0.85	0.60*Ref_NO_2 + 1.43*O_3_4 + 23.21, R^2 = 0.91
	09/05/16	-0.22*Ref_NO_2 - 0.08*O_3_3 + 49.43, R^2 = 0.04	0.01*Ref_NO_2 + 0.53*O_3_4 + 82.87, R^2 = 0.31
.	11/05/16	0.38*Ref_NO_2 + 0.48*O_3_3 – 12.31, R^2 = 0.22	0.22*Ref_NO_2 + 0.72*O_3_4 + 64.58, R^2 = 0.52
	13/05/16	0.61*Ref_NO_2 + 0.52*O_3_3 – 25.33, R^2 = 0.35	0.34*Ref_NO_2 + 1.08*O_3_4 + 47.05, R^2 = 0.88
	NO_2_corr_all_alt	0.46*Ref_NO_2 + 0.60*O_3_3 – 25.14, R^2 = 0.69	0.21*Ref_NO_2 + 0.67*O_3_4 + 65.36, R^2 = 0.24

Table S3: OLS linear regression parameters (and 95% confidence intervals), coefficient of determination and summary statistics for corrected NO_2 concentrations compared to the reference analyser NO_2 concentrations. The corrections considered for each study were: alternate 24 hour periods used to correct the next 24 hours (*Aq_corr_alt*); all of the alternate 24 hour periods combined into a single data set and used to correct all of the remaining 24 hour periods (*Aq_corr_all_alt*); and the first 24 hour period used to correct all of the remaining alternate 24 hour periods (*Aq_corr_first*). These analyses were restricted to the first deployment period in May 2016.

Monitor	Correction	Slope [95 % C.I.]	Intercept [95 % C.I.] / µg m^{-3}	R^2	RMSE / µg m^{-3}	MB / µg m^{-3}	NMB	*n*
O_3_3	*O_3_corr_alt*	1.07 [1.01, 1.12]	-8.93 [-13.47, -4.39]	0.93	9.99	-3.26	-0.04	117
	O_3_corr_all_alt	1.21 [1.16, 1.27]	-19.59 [-23.82, -15.36]	0.95	10.64	-3.55	-0.05	117
	O_3_corr_first	1.06 [1.01, 1.10]	-8.55 [-12.24, 4.86]	0.95	8.17	-4.21	-0.05	117
O_3_4	*O_3_corr_alt*	1.04 [0.98, 1.10]	-7.05 [-12.19, -1.92]	0.91	10.77	-4.05	-0.05	117
	O_3_corr_all_alt	1.50 [1.44, 1.56]	-48.04 [-53.03, -43.04]	0.96	20.24	-11.04	-0.14	117
	O_3_corr_first	0.76 [0.73, 0.80]	5.06 [2.52, 7.60]	0.96	15.50	-12.93	-0.17	117
NO_2_1	*NO_2_corr_alt*	1.23 [0.96, 1.50]	-2.75 [-9.68, 4.18]	0.42	19.26	3.44019	0.161505	117
	NO_2_corr_all_alt	1.21 [1.06, 1.35]	0.43 [-3.25, 4.11]	0.71	11.47	4.598973	0.215905	117
	NO_2_corr_first	1.45 [1.30, 1.60]	-8.02 [-11.92, -4.12]	0.76	12.42	1.424792	0.066889	117
NO_2_2	*NO_2_corr_alt*	0.72 [-2.32, 3.76]	-40.77 [-118.21, 36.67]	0.00	223.01	-48.80	-2.29	117
	NO_2_corr_all_alt	1.31 [1.09, 1.53]	-2.11 [-7.71, 3.49]	0.56	16.21	4.19	0.20	117
	NO_2_corr_first	1.37 [1.18, 1.55]	-9.92 [-14.62, -5.21]	0.66	13.70	-1.99	-0.09	117

7. References

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.

Consistency of urban background black carbon concentration measurements by portable AE51 and reference AE22 Aethalometers: Effect of corrections for filter loading

*Nicola Masey[1], Eliani Ezani[1,2], Jonathan Gillespie[1], Chun Lin[3], Scott Hamilton[4], Mathew R. Heal[3], Iain J. Beverland[1]**

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]Department of Environmental and Occupational Health, Faculty of Medicine and Health Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, MALAYSIA

[3]School of Chemistry, Joseph Black Building, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

[4]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

*CORRESPONDING AUTHOR: Dr Iain J. Beverland,
Department of Civil and Environmental Engineering, University of Strathclyde, 505F James Weir Building, 75 Montrose Street,
Glasgow, G1 1XJ, UK;
Email: iain.beverland@strath.ac.uk; Tel: +44 141 548 3202

Abstract

We collected over 1000 hourly-averaged black carbon (BC) measurements from co-located duplicate portable AethLab AE51 Aethalometers and a government reference Magee Scientific AE22 Aethalometer (latter adjusted for filter darkening effects using a standard procedure) at an urban background site in Glasgow, UK. AE51 and reference AE22 Aethalometer concentrations were highly correlated ($R^2 \geq 0.75$) for 3 of the 4 deployment periods. Application of two previously-reported methods for correction of underestimation of concentrations by AE51 monitors associated with filter loading generally overestimated [corrected] reference AE22 Aethalometer concentrations (average normalised mean bias = 0.22 and 0.09 for both instruments for two correction methods *cf.* 0.00 for unadjusted data across the full range of measurements). At ATN values > 40, RMSE increased for both unadjusted and adjusted BC measurements (e.g. RMSE for unadjusted measurements was 0.16 and 0.48 μg/m^3 below and above ATN of 40 respectively). Our observations suggest that AE51 monitors may not require correction in environments with low concentrations.

Keywords: air pollution; black carbon; Aethalometer; attenuation; filter loading

1. Introduction

Black carbon (BC) is a constituent of airborne particulate matter (PM) produced during incomplete combustion of carbon-based fuels. The health effects associated with exposure to PM and BC include respiratory and cardiovascular diseases[1,2], including associations between health outcomes and proximity to roads where BC concentrations are frequently elevated.[3–5] In the UK, BC concentrations are measured continuously at 14 sites in a nationally-coordinated network (https://uk-air.defra.gov.uk) using a mains-powered rack-mounted Aethalometer (AE22, Magee Scientific, CA, USA). The number of BC measurements at these fixed sites is limited because of the high installation, equipment and maintenance costs and because BC does not need to be measured under current UK air quality compliance legislation.

Battery-powered, hand-held Aethalometers have also been developed to measure real-time concentrations of BC. The small size and light weight of these monitors make them suitable for use in a variety of applications including static locations in a network[6–8], mobile monitoring[9–11] and/or personal monitoring.[12–15] In this study, we evaluated two microAeth AE51 portable BC Aethalometers (AethLabs, San Francisco, CA, USA).

The Aethalometers measure the concentration of BC using optical absorption techniques. Air is sampled through a filter and the concentration of BC is estimated by comparing the attenuation (ATN) of light passing through the particles deposited on the filter to that passing through an unloaded reference point on the same filter. However, the relationship between ATN and BC loading is not linear at higher attenuation values. Different methods to correct Aethalometer data to account for these filter loading effects have been proposed[16,17] but few studies have compared the correction algorithms to determine which provides the most accurate correction.[18]

Another potential source of error for the portable AE51 Aethalometers is rapid change in temperature and/or relative humidity, such as those experienced during personal monitoring, which can lead to large spurious positive and negative readings because of water condensation on the monitor filters or optics.[19] This source of error may be of limited importance for reference Aethalometers situated within air-conditioned cabins and sampling air in fixed locations. Cai et al. (2013) reduced the impact of condensation errors on the AE51 monitor through use of a diffusion dryer on the inlet. They also noted that drier efficiency improved when the monitors were used as personal monitors, which was attributed to heating (and thus reduction of relative humidity) of the sampled air.

Instrument noise from mechanical shocks, again manifested as large spurious positive and negative readings, have also been observed during mobile monitoring.[9] Cushioning the monitors, e.g. using foam padding, has been used to minimise the vibrations experienced by the monitor during mobile measurements[10], together with correction algorithms to remove these erroneous concentration peaks.[9,20]

The aim of our study was to compare the efficacy of methods commonly used to correct BC measurements from portable AE51 Aethalometers and to establish the ATN value at which filters require to be changed to maintain consistent AE51 Aethalometer accuracy. We build on the study of Good et al.[18], who evaluated correction algorithms for filter darkening, by evaluating some of these correction methods during repeated Aethalometer field deployments. Good et al.[18] used an online photoacoustic extinctiometer (PAX Droplet Measurement Technologies, Boulder, CO, USA) as a reference instrument under controlled laboratory conditions. In contrast, we deployed duplicate AE51 Aethalometers close to the inlet of an AE22 'reference' Aethalometer at an urban background site in the city of Glasgow, UK, for approximately 50 days of co-located measurements over a 5 month period. The static outdoor deployment of the AE51 Aethalometers avoided mechanical vibrations, and was anticipated not to have been unduly affected by abrupt temperature or relative humidity changes, enabling focus on field based assessment of filter loading effects on the agreement between measurements from portable and reference Aethalometer instruments.

2. Methods

2.1. Site description; AE22 Aethalometer operation

Measurements were made at the Glasgow Townhead monitoring site, an urban background location in central Glasgow (55.866 ºN, 4.244 ºW) that is part of the UK black carbon monitoring network. Hourly-averaged reference BC measurements made at this site using a Magee Scientific AE22 Aethalometer are publicly available (https://uk-air.defra.gov.uk/networks/network-info?view=ukbsn). AE22 instrument operation and data ratification are subject to national QA/QC protocols.[21] Unadjusted data at 5-min averaging were also obtained from personal communication with King's College London to evaluate the accuracy of the AE51 Aethalometers at higher temporal resolution.[22] The hourly-averaged reference concentration data are corrected prior to publication for filter darkening effects using a standard correction procedure based on Virkkula et al.[17,23] The unadjusted 5-min AE22 concentration data were corrected in a similar way prior to use in our analyses (see Supplementary Information).

2.2. AE51 Aethalometer set up and operation

Two microAeth AE51 Aethalometers (https://aethlabs.com/), subsequently named 'BC1204' and 'BC1303', were deployed in waterproof boxes on the roof of the monitoring station and sampled ambient air through a 1 m length of tubing supplied by the manufacturer. A conductive asbestos sampling inlet (SKC Ltd, UK) was connected to the inlet of the tubing as a rain shield to prevent water ingress. The flow rate of the AE51 Aethalometers was set to 50 mL/min and data were recorded each minute.

The AE51 Aethalometers were co-located at the monitoring station on 4 occasions in 2016 (Table 1): April (94 hours – BC1303 only), May (726 hours), July (68 hours) and August (334 hours). During each deployment the site was visited approximately every 5 days to download the AE51 data and to change filters. The AE51 filter attenuation did not exceed a value of 65 during these deployment periods.

The 1-min data collected by the AE51 Aethalometers were averaged to produce 5-min and hourly averaged concentrations for comparison with reference AE22 BC concentrations.

Table 1: Dates and descriptive statistics for co-located AE51 and reference AE22 BC Aethalometer deployments at the Glasgow Townhead network monitoring site.

Study name	Start date	Duration (hours)	Min hourly concentration ($\mu g/m^3$)	Max hourly concentration ($\mu g/m^3$)	Mean hourly concentration ($\mu g/m^3$)	Median hourly concentration ($\mu g/m^3$)
April	31/03/2016	94	0.10	1.70	0.61	0.60
May	29/04/2016	651	0.00	3.80	0.57	0.50
July	01/07/2016	68	0.10	0.70	0.37	0.40
August	08/08/2016	318	0.00	4.20	0.75	0.60

2.3. Correction of AE51 black carbon data

Prior to filter darkening correction procedures we smoothed the AE51 data to minimise the number of negative values in the dataset using Optimized Noise Algorithm (ONA) software from the AethLabs website (https://aethlabs.com/dashboard).[20] We set the change in attenuation value used to average BC concentration data in this ONA method to 0.05.

We then applied one of two alternative filter darkening correction procedures to the ONA-adjusted BC data from the portable AE51 Aethalometers to account for potential underestimation of BC as the darkness of the filter increased:

(a) Correction procedure published by Kirchstetter and Novak[16] using the standard coefficients (subsequently referred to as *K&N*):

$$BC_{corrected} = BC_{ONA}((0.88 \times exp^{(-ATN/100)}) + 0.12)^{-1}$$

where BC_{ONA} is the AE51 concentration after ONA correction (outlined above) and *ATN* is the attenuation of the AE51 filter. We used the ATN values directly from the AE51 monitors as clean filters were used on each occasion, in comparison to Good et al.[18] who used the percent change in attenuation between the start and end of the study to account for the use of preloaded filters.

(b) Correction procedure developed by Virkkula et al.[17] based on the concept that increasing ATN results in a linear underestimation of the correct BC concentration:

$$BC_{correct} = (1 + k * ATN)BC_{measured}$$

We estimated *k* from the linear regression slope between the ratio of AE22 Aethalometer concentrations divided by AE51 BC_{ONA} concentrations *vs.* corresponding AE51 *ATN* values with a fixed intercept value of 1.[24] When all available data were included in the regression analyses the values of *k* calculated were 0.0011 and 0.0045 for 1 h average data, and -0.0014 and -0.0019 5-min average data, for BC1204 and BC1303 monitors respectively. Negative *k* values imply that non-corrected BC concentrations overestimate the 'true' BC concentrations and have been suggested to occur in summer months when the ratio of black carbon to aerosol volume concentrations are lower.[17] The range of *k* values that we calculated was consistent with estimates calculated in other studies (e.g. k = -0.0039 [24] and k = 0.0033 [25]) and slightly lower than one recent study (k = 0.01 [26]), however these previous studies have been carried out a roadside sites in comparison to the urban background site in our work. We then corrected BC_{ONA} values as follows:

$$BC_{corrected} = (1 + k * ATN)BC_{ONA}$$

3. Results

3.1. Precision and accuracy of AE51 Aethalometers

For the full dataset, unadjusted concentrations measured by the two duplicate AE51 Aethalometers were highly correlated (R^2 = 0.93) with both instruments reporting similar values (slope = 1.05, intercept = 0.00, n = 1053) (Figure 1b). Similarly high correlations between duplicate AE51 Aethalometers have been reported in other studies, e.g. R^2 values > 0.95 between 13 co-located monitors.[12]

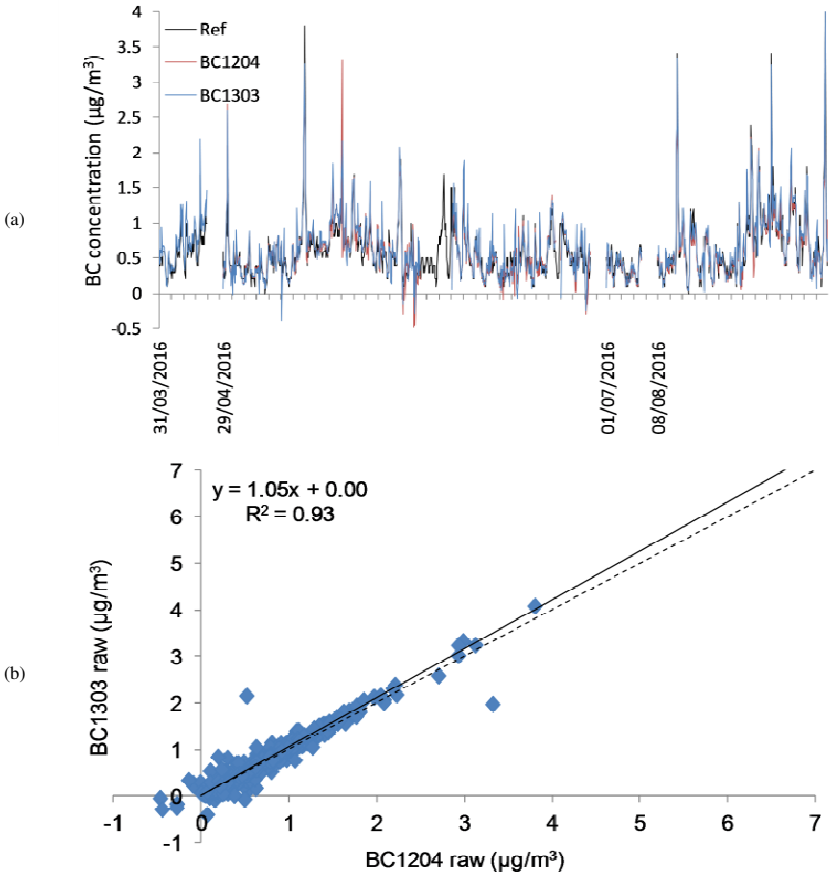


Figure 1: (a) Time series of adjusted hourly-average BC concentrations from reference AE22 Aethalometer and unadjusted concentrations measured by two AE51 BC Aethalometers. Gaps in plot separate non-consecutive time series. (b) Scatter plot with reduced major axis (RMA) regression of hourly-average unadjusted concentrations measured by duplicate AE51 BC Aethalometers.

Time series of hourly-average unadjusted concentrations from the duplicate AE51 Aethalometers and the reference AE22 Aethalometer showed very similar temporal patterns (Figure 1a). During the deployments in April, May and August the correlation coefficients between each AE51 and AE22 Aethalometer were greater than 0.75 (left panels of Figure 2). A previous study at a BC network monitoring site in Birmingham, UK reported high correlation (R^2 = 90) between AE51 and reference AE22 Aethalometer concentrations over a similar range of reference analyser BC concentrations (0 – 5 µg/m^3)[27] to the BC concentrations observed during our deployments. The August deployment had the most accurate AE51 measurements (NMB = 0.04 for BC1303) (Table 2). The lower R^2 values observed during our deployment in July (Table 1, Figure 2c) may have been the result of relatively low BC concentrations (< 1 µg/m^3) and the correspondingly limited range of

concentrations. However, the AE51 measurements during April showed the largest deviations from the AE22 Aethalometer concentrations (NMB = 0.19 for BC1303) (Table 2).

There was more scatter in the 5-min average concentrations measured by the AE51 and the AE22 Aethalometers, leading to lower coefficients of determination (R^2 = 0.08 – 0.58) than for comparisons using hourly averages (R^2 = 0.58 – 0.92) (Table 2). Despite the lower coefficients for 5-min average BC concentrations the NMB values were similar between the two temporal averages (e.g. NMB = −0.14 (hour) *vs.* −0.15 (5-min) for BC1204 during the July study).

Table 2: OLS regression equations [and 95% confidence intervals] between (a) hourly-average and (b) 5-min average AE51 Aethalometer (unadjusted) and reference AE22 Aethalometer BC concentrations for each deployment period. Descriptive statistics: coefficient of determination (R^2), root mean square error (RMSE), mean bias (MB), normalised mean bias (NMB) and number of hourly observations (n). Monitor BC1204 was not deployed in April 2016.

Monitor #	Study	Slope [95 % C.I.]	Intercept [95 % C.I.] (μg m^{-3})	R^2	RMSE (μg m^{-3})	MB (μg m^{-3})	NMB	n
(a) Hourly								
1204	May	0.98 [0.94, 1.03]	0.04 [0.01, 0.06]	0.76	0.19	0.03	0.05	651
	Jul	0.76 [0.64, 0.88]	0.04 [-0.01, 0.09]	0.70	0.10	-0.05	-0.14	68
	Aug	0.86 [0.83, 0.88]	0.08 [0.05, 0.10]	0.92	0.16	-0.03	-0.04	318
1303	April	1.18 [1.09, 1.27]	0.01 [-0.05, 0.07]	0.88	0.17	0.12	0.19	94
	May	1.00 [0.96, 1.05]	0.04 [0.01, 0.07]	0.75	0.21	0.04	0.08	648
	Jul	0.79 [0.63, 0.97]	0.06 [-0.01, 0.12]	0.58	0.11	-0.02	-0.05	68
	Aug	0.92 [0.89, 0.95]	0.09 [0.06, 0.12]	0.91	0.16	0.03	0.04	318
(b) 5-min								
1204	May	0.80 [0.78, 0.82]	0.14 [0.13, 0.15]	0.51	0.34	0.02	0.04	7642
	Jul	0.18 [0.14, 0.23]	0.25 [0.23, 0.27]	0.08	0.38	-0.06	-0.15	791
	Aug	0.59 [0.58, 0.61]	0.26 [0.25, 0.28]	0.58	0.46	-0.04	-0.06	3934
1303	April	0.83 [0.78, 0.89]	0.22 [0.18, 0.26]	0.43	0.37	0.12	0.19	1105
	May	0.81 [0.79, 0.83]	0.14 [0.13, 0.16]	0.51	0.34	0.04	0.07	7588
	Jul	0.21 [0.16, 0.26]	0.27 [0.25, 0.30]	0.09	0.38	-0.03	-0.07	791
	Aug	0.64 [0.62, 0.66]	0.29 [0.27, 0.31]	0.52	0.50	0.02	0.02	3934

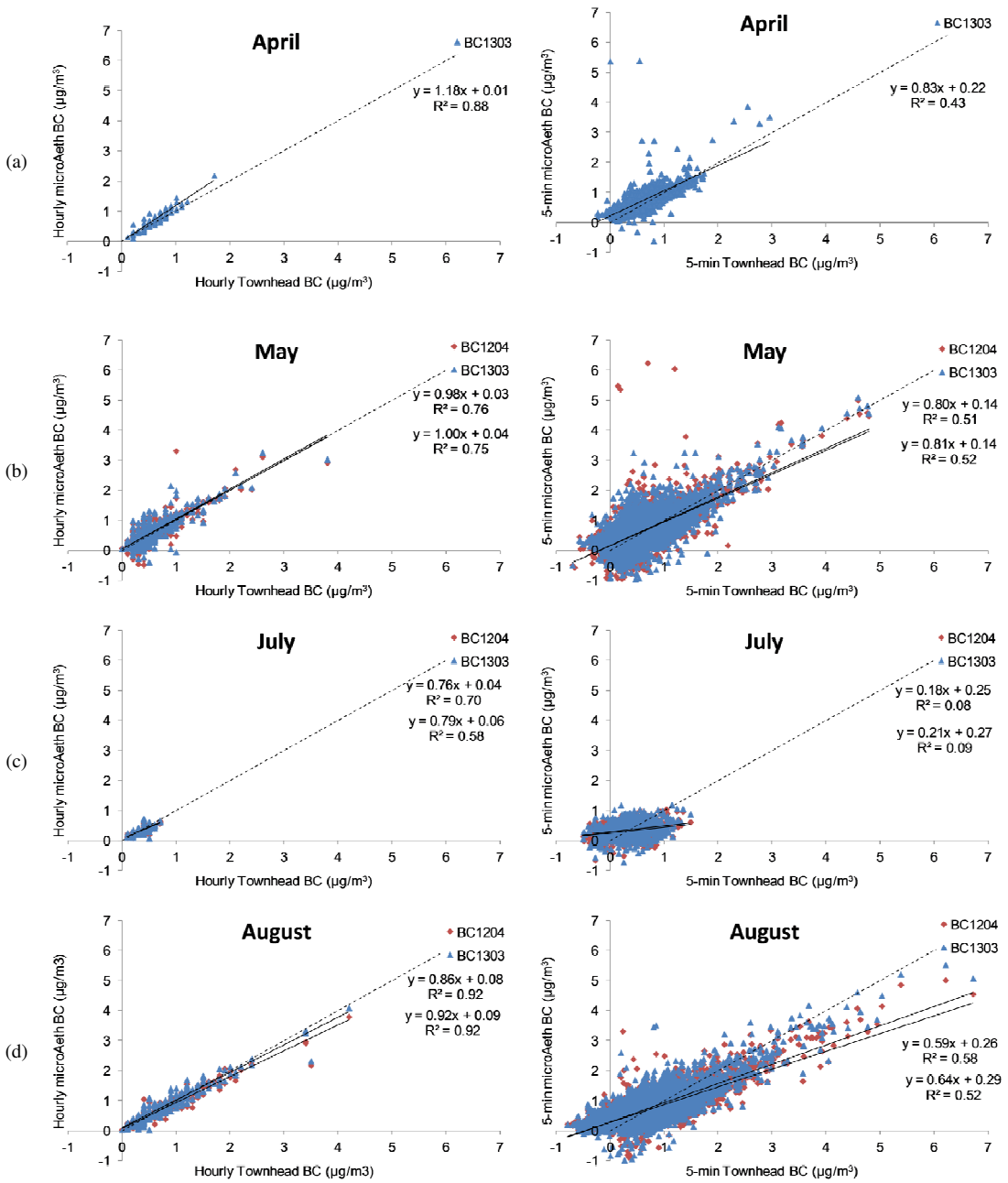


Figure 2: Scatter plots of unadjusted BC concentrations from AE51 BC Aethalometers *vs.* adjusted BC concentrations measured by the AE22 reference Aethalometer (graphs on left hand side show hourly averaged data and graphs on right hand side show 5-min average data) for each individual deployment: (a) April; (b) May; (c) July; and (d) August. Solid lines are ordinary least squares (OLS) regressions and dashed lines are the 1:1 line.

3.2. Effect of filter loading corrections

The correlation between the unadjusted AE51 and reference AE22 Aethalometer concentrations was higher for the hourly averaging compared to the 5-min averaging (Figure 3a).

The ONA-correction had very little impact on the regression lines, or regression statistics for either averaging resolution compared to the unadjusted concentrations (Figure 3b and Table 3). The ONA-algorithm is designed to remove noise in the AE51 Aethalometer data, such as that caused by vibrations when walking. In this work the Aethalometers were deployed statically, so the small impact of the ONA-algorithm is not unexpected. A similar observation was noted by Cheng and Lin[24] for a static AE51 Aethalometer deployed at a roadside location with higher concentrations than in this present work.

The *K&N* corrected AE51 Aethalometer data increases the regression slopes for both temporal resolutions, with higher correlation between the AE51 hourly data and the AE22 Aethalometer (R^2 = 0.80 *vs.* 0.42 for 1 h and 5-min averaging respectively) (Figure 3c). Both AE51 hourly-data and 5-min data overestimate the concentrations measured by the reference AE22 Aethalometer (NMB = 0.16 *vs.* 0.15 for hourly and 5-min average BC 1204 concentrations) (Table 3).

Similarly, the Virkkula et al. correction increased the regression slope between the AE51 and AE22 Aethalometers (Figure 3d). The accuracy of the AE51 monitors was similar for hourly averaging compared to 5-min averaging (NMB = 0.02 *vs.* -0.03 for BC1204 for hourly and 5-min averaging) (Table 3). For the hourly-averaged data the corrected regression slopes are closer to unity than for the *K&N* correction; however for 5-min averaging the *K&N* correction slope is closer to unity. There was no impact on the R^2 values from unadjusted to Virkkula et al. corrected concentrations however the NMB values were larger after correction (NMB = −0.002 (unadjusted) *vs.* 0.02 (Virkkula et al.)) (Table 3).

The *K&N* corrected hourly AE51 Aethalometer concentrations have the largest RMSE and NMB for both hourly and 5-min concentrations (NMB = 0.16 for BC1204), while the unadjusted and ONA-corrected data have the lowest values (NMB = 0.00 for BC1204) (Table 3). The *K&N* equation overcorrects the AE51 Aethalometer concentrations for filter loading effects while the Virkkula et al. correction overcorrects the concentrations by a lesser amount (NMB = 0.02 for BC1204) but still shows lower correlation with the reference Aethalometer concentrations than the unadjusted or ONA-corrected data.

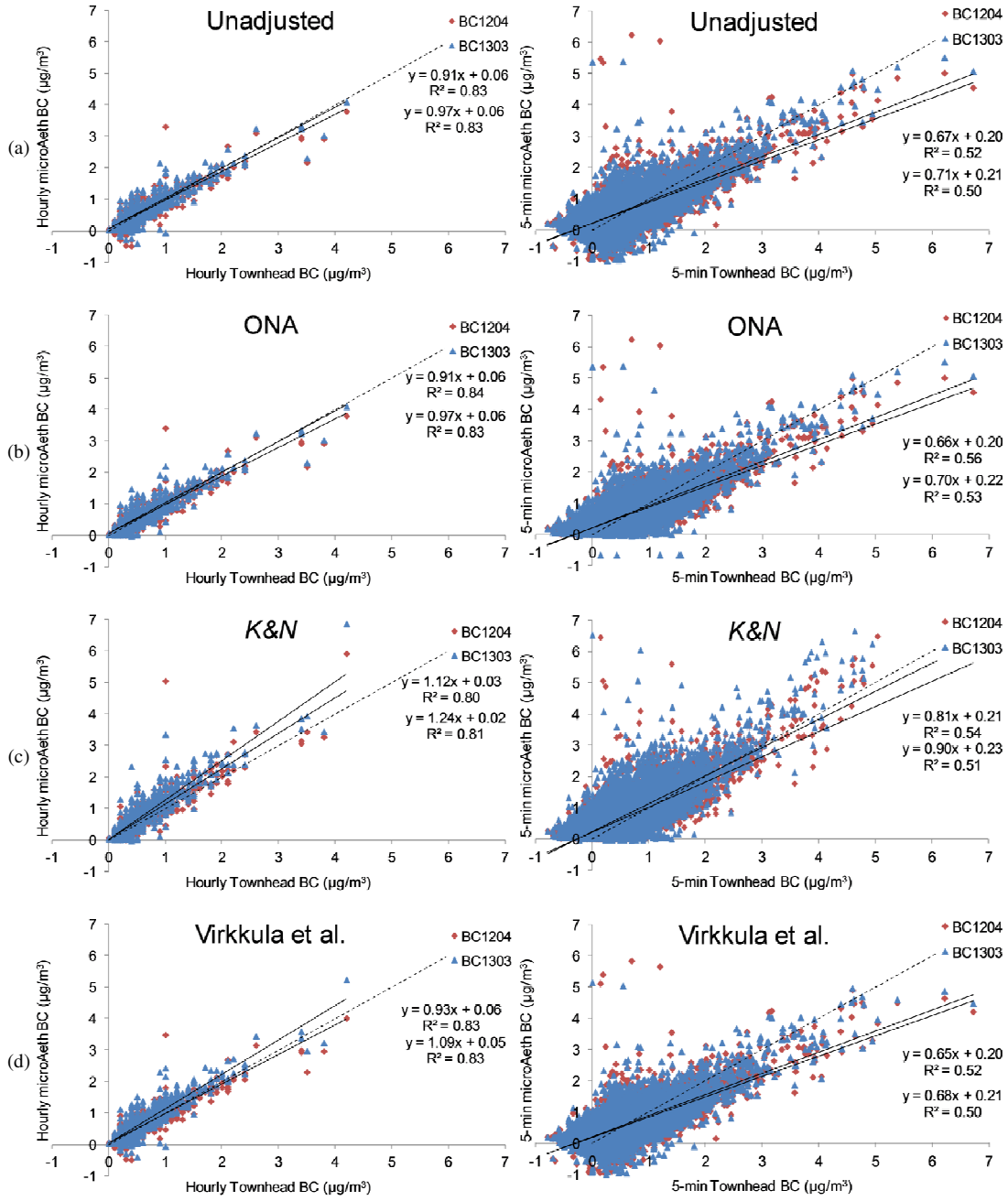


Figure 3: Scatter plots of BC concentrations from AE51 BC Aethalometers *vs.* adjusted BC concentrations measured by the AE22 reference Aethalometer (graphs on left hand side show hourly averaged data and graphs on right hand side show 5-min average data), for all deployment periods combined, for the following correction methods: (a) no corrections to AE51 data; (b) noise corrected using ONA; (c) *K&N* corrected; and (d) Virkkula et al. corrected (using *k* value derived from all studies [hourly or 5 min measurements]).

Table 3: OLS regression equations [95% confidence intervals] between the (a) hourly average and (b) 5-min average AE51 Aethalometer and reference AE22 Aethalometer BC concentrations for each filter loading correction method. Descriptive statistics: coefficient of determination (R^2), root mean square error (RMSE), mean bias (MB), normalised mean bias (NMB) and number of hourly observations (*n*).

Monitor #	Correction	Slope [95 % C.I.]	Intercept [95 % C.I.] (μg m^{-3})	R^2	RMSE (μg m^{-3})	MB (μg m^{-3})	NMB	*n*
(a) Hourly								
1204	Unadjusted	0.91 [0.88, 0.93]	0.06 [0.04, 0.08]	0.83	0.18	0.00	0.00	1038
	ONA	0.91 [0.89, 0.93]	0.06 [0.04, 0.08]	0.84	0.17	0.00	0.00	1038
	K&N	1.12 [1.08, 1.15]	0.03 [0.00, 0.05]	0.80	0.26	0.10	0.16	1038
	Virkkula et al.	0.89 [0.87, 0.92]	0.05 [0.04, 0.07]	0.83	0.18	0.01	0.02	1038
1303	Unadjusted	0.97 [0.94, 0.99]	0.06 [0.04, 0.08]	0.83	0.19	0.04	0.07	1129
	ONA	0.97 [0.94, 0.99]	0.06 [0.04, 0.08]	0.83	0.18	0.04	0.07	1129
	K&N	1.24 [1.21, 1.28]	0.02 [0.00, 0.05]	0.81	0.32	0.17	0.28	1129
	Virkkula et al.	0.77 [0.75, 0.79]	0.07 [0.05, 0.08]	0.83	0.23	0.10	0.16	1129
(b) 5-min								
1204	Unadjusted	0.67 [0.66, 0.68]	0.20 [0.19, 0.21]	0.52	0.38	-0.00	-0.00	12367
	ONA	0.66 [0.65, 0.67]	0.20 [0.20, 0.21]	0.56	0.36	-0.00	-0.00	12367
	K&N	0.81 [0.79, 0.82]	0.21 [0.20, 0.22]	0.54	0.42	0.09	0.15	12367
	Virkkula et al.	0.65 [0.64, 0.66]	0.20 [0.19, 0.21]	0.53	0.38	-0.02	-0.03	12367
1303	Unadjusted	0.71 [0.70, 0.72]	0.21 [0.20, 0.22]	0.50	0.40	0.03	0.06	13417
	ONA	0.71 [0.69, 0.72]	0.22 [0.21, 0.22]	0.53	0.38	0.04	0.06	13417
	K&N	0.90 [0.88, 0.91]	0.23 [0.22, 0.24]	0.51	0.49	0.17	0.27	13417
	Virkkula et al.	0.68 [0.66, 0.69]	0.21 [0.20, 0.22]	0.50	0.39	0.01	0.02	13417

3.3. Effect of ATN values on AE51 corrections

Different studies have recommended a range of ATN values (from 40 to 80) at which the filter in the portable AE51 Aethalometer should be changed.[13,17,24] We investigated the difference between the AE51 and reference AE22 Aethalometer concentrations as a function of ATN (Figure 4) to establish the range of ATN values at which the AE51 Aethalometer can reproduce the AE22 Aethalometer reference BC concentrations and if there is a limit to the ATN range under which the correction investigated can be successfully applied.

At low (< 30) ATN values the unadjusted, ONA-corrected and Virkkula et al. corrected BC concentrations residuals were all close to 0 (Figures 4a, 4b & 4d). The *K&N* corrected BC concentrations in this ATN range showed an increase in residual concentrations with increasing ATN range (Figure 4c).

The RMSE values between the AE51 and AE22 concentrations were similar (~0.18 μg/m^3) for ATN values up to 40 for the unadjusted, ONA-corrected and Virkkula et al. corrected BC concentrations (Figure 5, Table S1). Above ATN values of 40 the RMSE increased for these three methods, with the increase more pronounced for BC1204 (RMSE > 0.4 μg/m^3). The RMSE for the *K&N* corrected BC concentrations were similar to the other

methods for ATN < 10, however above this ATN value the *K&N* corrected RMSE increased and was much greater than for the other methods (Figure 5, Table S1).

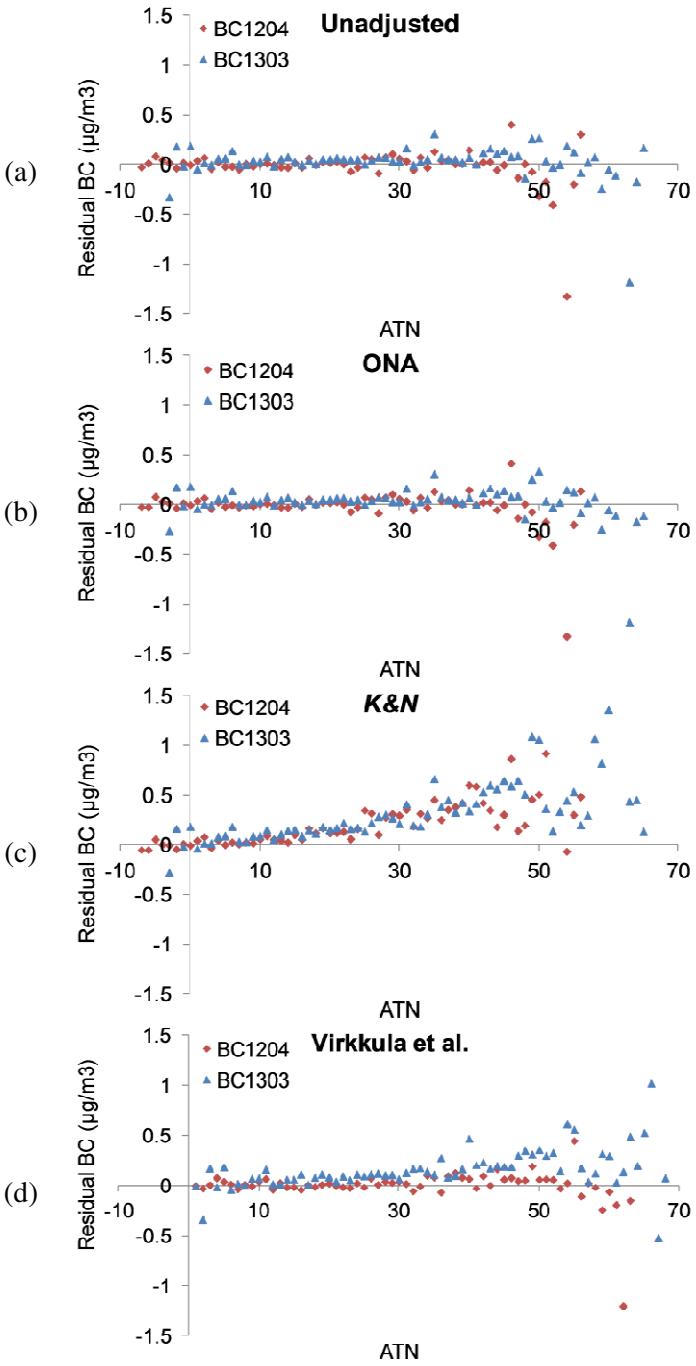


Figure 4: Scatter plots of the difference between the hourly BC concentrations measured by the AE51 Aethalometers and by the AE22 reference Aethalometer *vs.* ATN values measured by AE51 Aethalometers. Plots are shown for: (a) unadjusted AE51 measurements; (b) ONA-corrected AE51 measurements; (c) K&N corrected AE51 measurements; and (d) Virkkula et al. corrected AE51 measurements; with all study data included. The average ratio for each ATN integer value is shown for ease of visualisation.

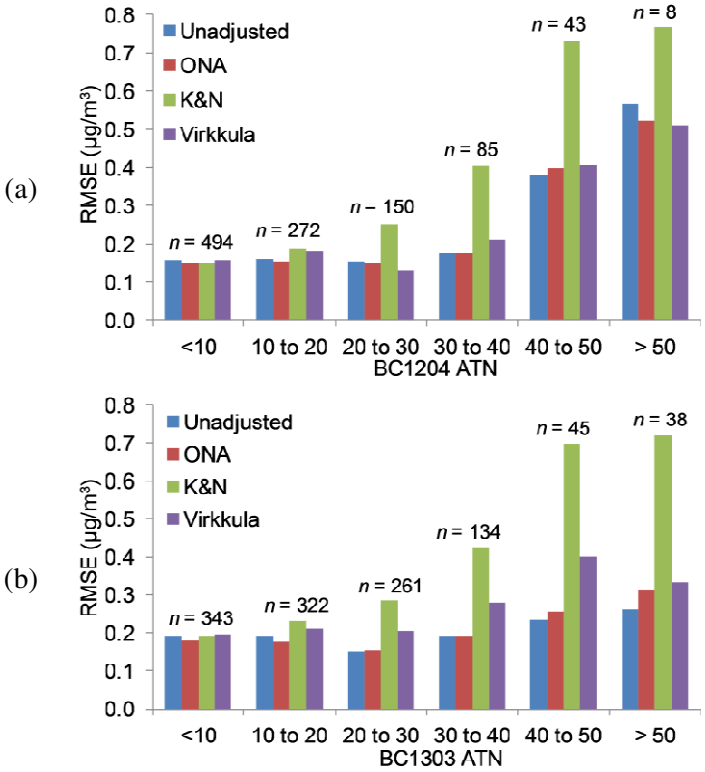


Figure 5: Root mean square error (RMSE) between hourly BC concentrations measured by the (a) BC1204 and (b) BC1303 AE51 Aethalometers and the AE22 reference Aethalometer. The RMSE values are shown for ATN values < 10; 10 – 20; 20 – 30; 30 – 40; 40 – 50; and > 50. The number of hourly values contributing to each group of RMSE values is given above each group.

4. Discussion

Portable AE51 Aethalometers have been increasingly used in mobile and personal monitoring. However, there have been limited field evaluations of the AE51 monitors under conditions representative of personal monitoring. Our evaluation of the AE51 monitor at 5-min and hourly averaged concentrations showed the higher temporal resolution concentrations to have more scatter, and consequently lower R^2 values, than the hourly evaluations. A similar observation was observed for AE51 monitors evaluated at a site with a larger concentration range, with higher scatter present at concentrations <2 µg/m^3.[24] High temporal resolutions are required for mobile and personal monitoring in order to record rapidly changing pollution environments, and this suggests the AE51 portable Aethalometer may be subject to greater noise at these high temporal resolutions.

At lower concentrations (< ~1.5 µg/m^3) the concentrations measured by the AE51 Aethalometers showed larger deviations from the reference Aethalometer than at higher concentrations. The use of AE51 for personal monitoring may be subject to varying error

according to the pollution environment being monitored. For example, greater errors may be observed at lower BC concentrations such as the 0.5 μg/m^3 encountered by Williams and Knibbs[15] during indoor personal monitoring. Smaller errors may be seen in high-pollution scenarios such as the BC concentrations of 4-6 μg/m^3 reported by Dons et al.[12] during personal monitoring of commuting.

The correlation between the AE51 and AE22 reference Aethalometer ranged from 0.58 to 0.92 when hourly averaging was used. Higher R^2 values (> 0.75) between AE51 monitor and Thermo Multi-Angle Absorption Photometer (Carusso, MAAP) BC concentrations have been reported for 10-min average concentrations when the minimum concentration was 1.5 μg/m^3.[28] The lower correlations observed in our study could be attributed to a difference in the reference method used or, as discussed above, greater noise in the AE51 monitor measurements at these low concentrations and high temporal resolutions. A study evaluating the AE51 against a similar reference instrument to used in this work found the R2 between the instruments at an urban background site to be > 0.9.[27] As previously discussed the lower concentrations of BC measured in this work by the reference analyser may be in part cause of the lower R2 values in this study, with lower values occurring when the range and absolute concentration of BC measured by the analyser was small.

We compared the correction algorithms to account for filter shadowing effects in the Aethalometer under ambient pollution conditions. The AE51 Aethalometer concentrations after correction for filter darkening overestimated the reference Aethalometer concentrations for *K&N* correction and both the *K&N* and Virkkula et al. correction had larger errors than the unadjusted BC concentrations. The underestimation of BC using the Virkkula et al. correction has been reported previously in a chamber experiment comparing BC concentrations measured using a loaded and unloaded filter.[18] Correction using the *K&N* formula has previously been shown to underestimate the filter loading effects when a clean filter was used, but overestimate the loading effects when pre-loaded filters were used, under controlled chamber-experiments.[18] We suggest that at the lower concentrations in this study, or under static monitoring, the *K&N* correction should be avoided. The deviations between portable AE51and reference AE22 Aethalometer concentrations were minimal at ATN < 40 for unadjusted and Virkkula et al. adjusted BC concentrations. This suggests that, under the conditions of this work (static monitoring of low ambient concentrations using low flow rates), correction algorithms do not need to be applied at these low attenuation values. This finding is in contrast to those by Good et al.[18], who stated that corrections for filter shadowing effects should be applied over the range of ATN values tested (0 – 125).

Above ATN values of 40 the BC residuals and RMSE increased for the unadjusted and Virkkula et al. corrected BC concentrations. Good et al.[18] showed that an ATN value of 80 should not be exceeded in order to maintain accurate BC measurements using the AE51 Aethalometer, and also noted some divergence in the measurements for ATN values greater than 60. Our work suggests that changing the filter at a value of 60 could still lead to underestimation by the AE51 Aethalometer at this higher attenuation value, and that the filters should be changed before ATN reaches 40 in order to ensure filter darkening effects do not adversely affect the AE51 Aethalometer accuracy.

In summary, we compared field measurements of BC concentrations at an urban background site using 2 portable AE51 Aethalometers and 1 reference AE22 Aethalometer. After correction for filter loading corrected AE51 Aethalometer concentrations generally overestimated reference AE22 Aethalometer concentrations (hourly NMB = 0.00 and 0.10 for unadjusted and adjusted BC1204 concentrations respectively). Our observations appear to suggest that AE51 Aethalometer measurements may not require correction for filter loading to maintain consistency with reference AE22 concentrations when AE51 ATN values are less than 40.

Acknowledgements

Nicola Masey is funded through a UK Natural Environment Research Council (NERC) CASE PhD studentship (NE/K007319/1), with support from Ricardo Energy and Environment. Eliani Ezani is funded by the Ministry of Higher Education Malaysia (KPT(BS)860126295394). Jonathan Gillespie is funded through an Engineering and Physical Sciences Research Council Doctoral Training Grant (EPSRC DTG EP/L505080/1 and EP/K503174/1) studentship, with support from the University of Strathclyde and Ricardo Energy and Environment. Chun Lin is funded through NERC/Innovate UK grant NE/N007352/1. We acknowledge access to the reference black carbon measurement data, which were obtained from uk-air.defra.gov.uk and are subject to Crown 2014 copyright, Defra, licenced under the Open Government Licence (OGL). We thank Kings College London for providing access to the 5-min temporal resolution black carbon concentrations.

Supplementary Information is available at Journal of Exposure Science and Environmental Epidemiology's website.

The authors declare no conflict of interest.

References

1. World Health Organization. Air quality guidelines. Global update 2005. Particulate matter, ozone, nitrogen dioxide and sulfur dioxide [Internet]. Copenhagen, Denmark; 2006 [cited 2017 Feb 27]. Available from: http://www.euro.who.int/__data/assets/pdf_file/0005/78638/E90038.pdf.

2. World Health Organization. Review of evidence on health aspects of air pollution – REVIHAAP Project: Technical Report [Internet]. Copenhagen, Denmark; 2013 [cited 2017 Feb 27]. Available from: http://www.euro.who.int/__data/assets/pdf_file/0004/193108/REVIHAAP-Final-technical-report-final-version.pdf.

3. Grahame TJ, Klemm R, Schlesinger RB. Public health and components of particulate matter: The changing assessment of black carbon. J Air Waste Manag Assoc. 2014;64(6):620–60.

4. Janssen N, Hoek G, Simic-Lawson M, Fischer P, van Bree L, ten Brink H, et al. Black carbon as an additional indicator of the Adverse Health Effects of Airborne Particles Compared with PM10 and PM2. 5. Environ Health Perspect. 2011;119(12):1691–9.

5. Janssen N, Gerlofs-Nijland M, Lanki T, Salonen R, Cassee F, Hoek G, et al. Health effects of black carbon [Internet]. World Health Organisation; 2012 [cited 2017 Jan 19]. Available from: http://www.euro.who.int/__data/assets/pdf_file/0004/162535/e96541.pdf

6. Gillespie J, Masey N, Heal MR, Hamilton S, Beverland IJ. Estimation of spatial patterns of urban air pollution over a 4-week period from repeated 5-min measurements. Atmos Environ. 2017 Feb;150:295–302.

7. Montagne DR, Hoek G, Klompmaker JO, Wang M, Meliefste K, Brunekreef B. Land Use Regression Models for Ultrafine Particles and Black Carbon Based on Short-Term Monitoring Predict Past Spatial Variation. Environ Sci Technol. 2015 Jul 21;49(14):8712–20.

8. Weichenthal S, Farrell W, Goldberg M, Joseph L, Hatzopoulou M. Characterizing the impact of traffic and the built environment on near-road ultrafine particle and black carbon concentrations. Environ Res. 2014 Jul;132:305–10.

9. Apte JS, Kirchstetter TW, Reich AH, Deshpande SJ, Kaushik G, Chel A, et al. Concentrations of fine, ultrafine, and black carbon particles in auto-rickshaws in New Delhi, India. Atmos Environ. 2011 Aug;45:4470–80.

10. Hankey S, Marshall JD. On-bicycle exposure to particulate air pollution: Particle number, black carbon, PM2.5, and particle size. Atmos Environ. 2015 Dec;122:65–73.

11. Van den Bossche J, Peters J, Verwaeren J, Botteldooren D, Theunis J, De Baets B. Mobile monitoring for mapping spatial variation in urban air quality: Development and validation of a methodology based on an extensive dataset. Atmos Environ. 2015 Mar;105:148–61.

12. Dons E, Int Panis L, Van Poppel M, Theunis J, Wets G. Personal exposure to Black Carbon in transport microenvironments. Atmos Environ. 2012 Aug;55:392–8.

13. Dons E, Temmerman P, Van Poppel M, Bellemans T, Wets G, Int Panis L. Street characteristics and traffic factors determining road users' exposure to black carbon. Sci Total Environ. 2013 Mar;447:72–9.

14. Dons E, Van Poppel M, Kochan B, Wets G, Int Panis L. Implementation and validation of a modeling framework to assess personal exposure to black carbon. Environ Int. 2014 Jan;62:64–71.

15. Williams RD, Knibbs LD. Daily personal exposure to black carbon: A pilot study. Atmos Environ. 2016 May;132:296–9.

16. Kirchstetter TW, Novakov T. Controlled generation of black carbon particles from a diffusion flame and applications in evaluating black carbon measurement methods. Atmos Environ. 2007 Mar;41(9):1874–88.

17. Virkkula A, Mäkelä T, Hillamo R, Yli-Tuomi T, Hirsikko A, Hämeri K, et al. A Simple Procedure for Correcting Loading Effects of Aethalometer Data. J Air Waste Manag Assoc. 2007 Oct 1;57(10):1214–22.

18. Good N, Mölter A, Peel JL, Volckens J. An accurate filter loading correction is essential for assessing personal exposure to black carbon using an Aethalometer. J Expo Sci Environ Epidemiol [Internet]. 2016 Dec 21 [cited 2017 Jan 18]; Available from: http://www.nature.com/jes/journal/vaop/ncurrent/abs/jes201671a.html

19. Cai J, Yan B, Kinney PL, Perzanowski MS, Jung K-H, Li T, et al. Optimization Approaches to Ameliorate Humidity and Vibration Related Issues Using the MicroAeth Black Carbon Monitor for Personal Exposure Measurement. Aerosol Sci Technol. 2013 Nov 1;47(11):1196–204.

20. Hagler GSW, Yelverton TLB, Vedantham R, Hansen ADA, Turner JR. Post-processing Method to Reduce Noise while Preserving High Time Resolution in Aethalometer Real-time Black Carbon Data. Aerosol Air Qual Res. 2011 Oct;11:539–46.

21. Butterfield D, Beccaceci S, Quincey P, Sweeney B, Lilley A, Bradshaw C, et al. 2014 Annual Report for the UK Black Carbon Network [Internet]. 2015 Jul. Report No.: NPL Report AS 97. Available from: https://uk-air.defra.gov.uk/library/reports?report_id=844

22. Environmental Research Group, King's College London, Font A. High Temporal Resolution BC and ATN data for Glasgow Townhead AURN - Email. 2016.

23. Butterfield D, Beccaceci S, Quincey P, Sweeney B, Lilley A, Bradshaw C, et al. 2015 Annual Report for the UK Black Carbon Network [Internet]. National Physical Laboratory; 2016 Jun. Available from: https://uk-air.defra.gov.uk/library/reports?report_id=920

24. Cheng Y-H. Real-Time Performance of the microAeth® AE51 and the Effects of Aerosol Loading on Its Measurement Results at a Traffic Site. Aerosol Air Qual Res [Internet]. 2013 [cited 2014 Feb 25]; Available from: http://www.aaqr.org/Doi.php?id=23_AAQR-12-12-OA-0371&v=13&i=6&m=12&y=2013

25. Cheng Y-H, Liao C-W, Liu Z-S, Tsai C-J, Hsi H-C. A size-segregation method for monitoring the diurnal characteristics of atmospheric black carbon size distribution at urban traffic sites. Atmos Environ. 2014 Jun;90:78–86.

26. Morales Betancourt R, Galvis B, Balachandran S, Ramos-Bonilla JP, Sarmiento OL, Gallo-Murcia SM, et al. Exposure to fine particulate, black carbon, and particle number concentration in transportation microenvironments. Atmos Environ. 2017 May;157:135–45.

27. Delgado-Saborit JM. Use of real-time sensors to characterise human exposures to combustion related pollutants. J Environ Monit. 2012;14(7):1824.

28. Viana M, Rivas I, Reche C, Fonseca AS, Pérez N, Querol X, et al. Field comparison of portable and stationary instruments for outdoor urban air exposure assessments. Atmos Environ. 2015 Dec;123, Part A:220–8.

Supplementary Information

Consistency of urban background black carbon concentration measurements by portable AE51 and reference AE22 Aethalometers: Effect of corrections for filter loading

*Nicola Masey[1], Eliani Ezani[1,2], Jonathan Gillespie[1], Chun Lin[3], Scott Hamilton[4], Mathew R. Heal[3], Iain J. Beverland[1]**

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]Department of Environmental and Occupational Health, Faculty of Medicine and Health Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, MALAYSIA

[3]School of Chemistry, Joseph Black Building, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

[4]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

*CORRESPONDING AUTHOR: Dr Iain J. Beverland, Department of Civil and Environmental Engineering, University of Strathclyde, 505F James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK;
Email: iain.beverland@strath.ac.uk; Tel: +44 141 548 3202

The 5-min data provided by King's College London for the Glasgow Townhead AE22 reference Aethalometer was the raw BC concentrations prior to correction using the Virkkula et al. (2007) correction as specified in the Annual Reports for the UK Black Carbon Network.[1] Figure S1 compares the hourly averages of these 5-min data, with and without this correction, against the hourly averages downloaded from the UK-air website.

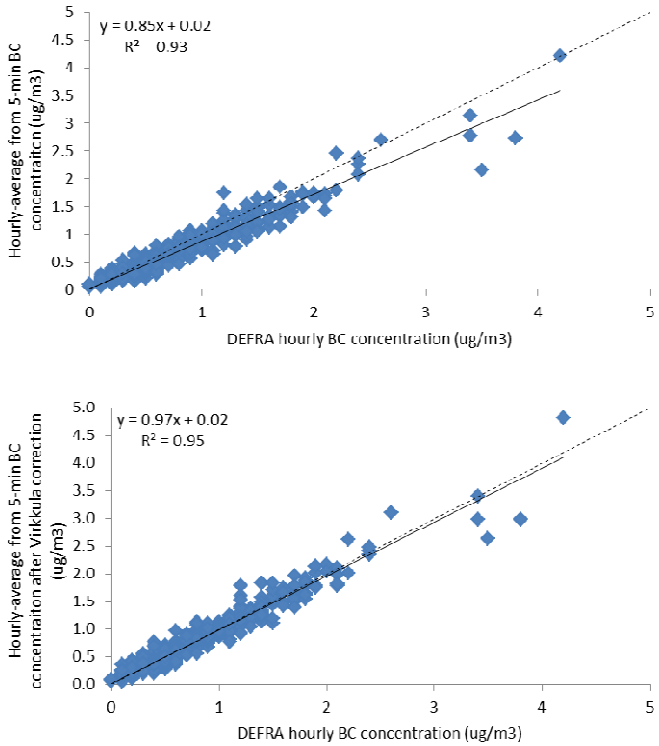


Figure S1: Scatter plots between hourly-average concentrations derived from the 5-min concentrations *vs.* hourly-average BC concentration downloaded from the UK air quality database website. Fig (a) plots hourly averages computed using unadjusted 5-min data. Fig (b) plots hourly averages computed using 5-min data after correction using Virkkula et al. correction. Solid lines are ordinary least squares (OLS) regressions and dashed lines are 1:1 lines.

Table S1: Descriptive statistics between hourly average BC concentrations measured by AE51 ((a) BC1204 and (b) BC1303) and AE22 Aethalometers split by ATN values measured by the AE51 instruments.

		ATN					
		< 10	10 - 20	20 - 30	30 - 40	40 – 50	> 50
(a) BC1204							
n		*494*	*272*	*150*	*85*	*43*	*8*
RMSE	Unadjusted	0.16	0.16	0.15	0.17	0.38	0.57
($\mu g/m^3$)	ONA	0.15	0.15	0.15	0.17	0.40	0.52
	K&N	0.15	0.19	0.25	0.40	0.73	0.77
	Virkkula	0.16	0.18	0.13	0.21	0.41	0.51
MB	Unadjusted	0.00	0.00	0.01	0.03	0.05	-0.25
($\mu g/m^3$)	ONA	0.00	0.00	0.01	0.03	0.05	-0.30
	K&N	0.01	0.08	0.16	0.33	0.39	0.58
	Virkkula	0.00	0.00	0.01	0.07	0.08	-0.16
NMB	Unadjusted	-0.01	0.00	0.01	0.04	0.07	-0.14
	ONA	-0.01	0.00	0.01	0.04	0.07	-0.14
	K&N	0.02	0.14	0.25	0.41	0.56	0.31
	Virkkula	0.00	0.00	0.02	0.09	0.12	-0.09
(b) BC1303							
n		*343*	*322*	*261*	*134*	*45*	*38*
RMSE	Unadjusted	0.19	0.19	0.15	0.19	0.23	0.26
($\mu g/m^3$)	ONA	0.18	0.18	0.15	0.19	0.26	0.31
	K&N	0.19	0.23	0.29	0.42	0.70	0.72
	Virkkula	0.20	0.21	0.20	0.28	0.40	0.33
MB	Unadjusted	0.03	0.0.4	0.05	0.06	0.11	-0.02
($\mu g/m^3$)	ONA	0.03	0.04	0.05	0.06	0.11	-0.04
	K&N	0.05	0.12	0.20	0.31	0.61	0.43
	Virkkula	0.04	0.08	0.12	0.17	0.32	0.18
NMB	Unadjusted	0.05	0.06	0.08	0.09	0.11	-0.03
	ONA	0.05	0.06	0.08	0.09	0.12	-0.05
	K&N	0.10	0.21	0.34	0.46	0.64	0.55
	Virkkula	0.07	0.13	0.20	0.26	0.34	0.23

References

1 Butterfield D, Beccaceci S, Quincey P, Sweeney B, Lilley A, Bradshaw C *et al.* 2015 Annual Report for the UK Black Carbon Network. National Physical Laboratory, 2016https://uk-air.defra.gov.uk/library/reports?report_id=920.

5. Spatial variations in urban concentrations of NO_2, O_3, BC and PN: interpretation of repeated co-located measurements of concentrations measured over 6-minute and 1-week averaging periods

The ability of the portable sensors evaluated in Chapter 3 to estimate spatial variations in pollution concentrations over an urban area in Glasgow were investigated. Additionally, the ability of the peripatetic measurements to estimate longer duration concentrations, derived from study-average and longer-term static measurements, was investigated. This was a follow-on study to our published paper (Appendix B), and we developed upon the published work by investigating a larger number of portable monitors (measuring both gaseous and particulate concentrations) and in a different study area.

The post-processing of the black carbon data in this work used the Kirschetter and Novac process, which (in the previous chapter) I showed to suffer from larger errors than some of the other correction algorithms available. However, in order to allow comparison between this study and our previously published work (to evaluate transferability of the previous findings to different locations and at a different time) the same method as presented in the published paper was used. The research evaluating the different post-processing algorithms presented in Chapter 4 was carried out after the publication of our paper therefore could not be used in the experimental design.

In addition, this research used standard Palmes tubes in the measurement of weekly NO_2 concentrations (with no modification of design or uptake rate to account for the wind-speed effects demonstrated in Chapter 3). The research presented in Chapter 2 was carried out after the design of the research and collection of data presented in this chapter. We anticipate that any wind-speed effects experienced by these tubes will be minimal as it was shown in Chapter 3 that the wind-speed effects were much lesser for tubes that were exposed for 1 week (as is the case in this work).

N. Masey was responsible for experimental design, data collection, data analysis and preparation of the manuscript. J. Gillespie provided input on the design of the experiment and data analysis, and helped with field measurements. E. Ezani assisted with field measurements. S. Hamilton, M. Heal and I. Beverland provided advice about data analysis and editorial comments on the manuscript.

This manuscript has been formatted with the intention of submission to Atmospheric Environment.

Spatial variations in urban concentrations of NO_2, O_3, BC and PN: interpretation of repeated co-located measurements of concentrations over 6-min and 1-week averaging periods

Nicola Masey[1], Eliani Ezani[1,2], Jonathan Gillespie[1], Scott Hamilton[3], Mathew R. Heal[4], Iain J Beverland[1]*

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]Department of Environmental and Occupational Health, Faculty of Medicine and Health Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, MALAYSIA

[3]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

[4]School of Chemistry, Joseph Black Building, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

*CORRESPONDING AUTHOR: Dr Iain J. Beverland, Department of Civil and Environmental Engineering, University of Strathclyde, 505F James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK; email: iain.beverland@strath.ac.uk; Tel: +44 141 548 3202

Research highlights:

- Similar spatial pattern in repeated 1-week NO_2 and 6-min NO_2, BC & PN concentrations
- Relatively high correlation between 6-min BC, PN and NO_2 concentrations
- Highest correlations between measurements during morning 'rush hour' periods
- Overall-average BC and NO_2 concentrations highly correlated ($R^2 \geq 0.70$)

Abstract

We measured black carbon (BC), particle number (PN), nitrogen dioxide (NO_2) and ozone (O_3) concentrations at 10 sites over a 0.5 x 0.5 km area in Glasgow, UK, using hand-held monitors. Measurements were made for 6-minute periods at three times of day (AM, Noon and PM) during one day at the start and end of each week over 6 consecutive weeks. We used passive diffusion tubes (PDTs) to measure NO_2 concentrations for the 6 x 1-week periods between monitor measurements at the same sites. Similar spatial patterns were noted in the average of 6-min BC, PN and NO_2 measurements at the start and end of each week and 1-week PDT NO_2 measurements. Average 6-min BC, PN and NO_2 concentrations, over

all weeks at each site, were highly correlated, especially during morning measurements (R^2 = 0.94 [BC *vs.* PN], 0.77 [BC *vs.* NO_2], 0.71 [PN *vs.* NO_2]). Over the 6-week period average 6-min BC and average 1-week NO_2 were highly correlated ($R^2 \geq 0.70$ for all times of day) suggesting that it may be possible to use repeated short-term BC measurements to estimate longer-term NO_2 concentrations.

Keywords: air pollution; hand-held monitors; passive samplers; peripatetic monitoring

1. Introduction

Detrimental health effects are associated with exposure to air pollution (World Health Organization, 2013), thus pollutants are monitored and their concentrations assessed against legislative limits. Historically, monitoring has taken place at government funded automatic monitoring stations, which provide real-time but spatially limited pollutant concentrations. These automatic networks can be supplemented using passive monitoring systems, which can be deployed at higher spatial density due to their low cost and simple siting requirements (Gillespie et al., 2016, 2017), but whose measurements are at considerably lower temporal resolution than the automatic monitors. Population exposures to pollution have been estimated using concentrations measured at the nearest monitoring station or through pollution modelling; however these can only provide estimates of an individual's exposure at limited spatial and temporal resolutions.

Portable real-time monitors are continually being developed to help address the need for high spatially and temporally resolved monitoring networks and population exposure estimates. These have been used to measure mobile pollution concentrations, via deployments on platforms of bicycles or vehicles (Hankey and Marshall, 2015; Van den Bossche et al., 2015), or statically when deployed for short periods at selected locations within a study area (Deville Cavellin et al., 2016; Norris and Larson, 1999). The data from these monitors have been used to model concentrations over an extended study area (Montagne et al., 2015) or in personal exposure assessments (Delgado-Saborit, 2012; Dons et al., 2013). To date, only limited study of the relationships between hand-held monitors measuring different pollutants or the representativeness of these relatively-short duration monitor measurements of longer-term concentrations has been reported (Beckerman et al., 2008; Durant et al., 2014; Gillespie et al., 2017; Klompmaker et al., 2015; Riley et al., 2016).

In an earlier study we investigated the relationship between short-duration peripatetic measurements of black carbon and particle number, and longer-term average nitrogen dioxide concentrations in central Glasgow, UK (Gillespie et al., 2017). Here we extend our

earlier work through study of a different geographical area (to identify the transferability of our previous findings), inclusion of short-term gaseous (in addition to particle) pollutant measurements, and investigation if the time of day when short-duration measurements are made influences observed bivariate relationships between different pollutants and averaging periods.

2. Methods

2.1. Study Design

The study was carried out in the city centre of Glasgow, UK (55.87° N, 4.26° W. The study measurements were made around the University of Strathclyde campus, which encompasses Cathedral Street and George Street, which have annual average daily flows of approximately 20,000 and 9,500 vehicles day^{-1} respectively (Figure 1).

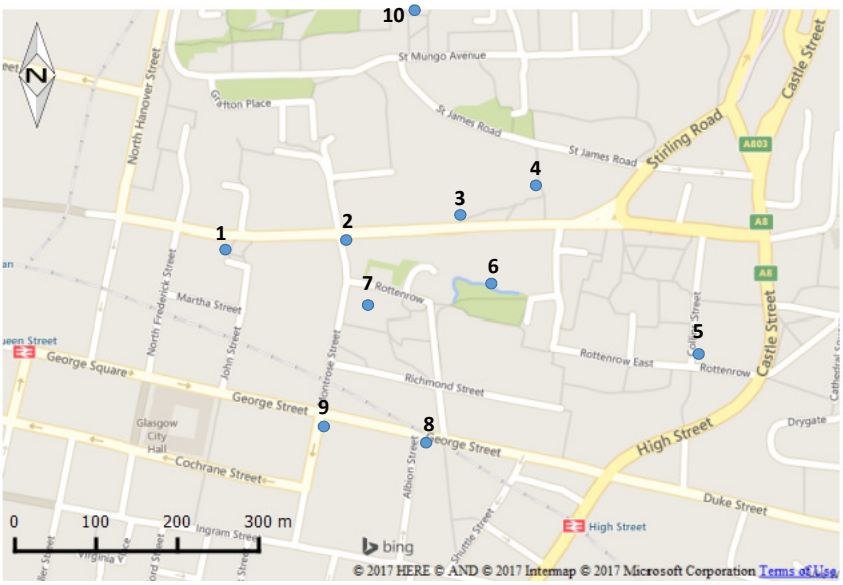


Figure 1: Location of sites in central Glasgow where 6-minute monitor and 1-week PDT measurements were made. Site 10 is located at the Townhead monitoring station, which is part of the UK Automatic Urban and Rural (AURN) Network (http://www.scottishairquality.co.uk/latest/index?site_id=GLKP). Sites were visited in numerical order in repeated sets of measurements over approximately 1.5 hour periods. Sites 2, 3, 8 and 9 were located within 2 m of major roads.

Measurements were made during six consecutive 1-week study periods between 27th October 2015 and 8th December 2015. Weekly-exposed passive diffusion tubes (PDTs) for NO_2 were located at 10 sites anticipated to have different pollution environments (Figure 1). Sites 2, 3, 8 and 9 were located within 2 m of the major roads of Cathedral and George

Streets, while Site 5 was located within 2 m of a side road with lower traffic flows. Site 4 was located ~35 m from Cathedral Street with an area of grass separating the site from the road, while Site 1 was set ~15 m from Cathedral Street separated by a hedge and a wall. Sites 6 and 7 were located in park areas not in direct proximity to traffic sources. Site 10 was located within 10 m of the Glasgow Townhead monitoring station that is part of the UK Automatic Urban and Rural Network and which contains real-time analysers reporting hourly-average concentrations of NO_2, O_3 and BC (DEFRA, 2017). All aspects of the operation and data ratification of these analysers are subject to published network-wide QA/QC procedures (DEFRA, 2017). The AURN station is located in an urban background location with no roads in close proximity.

On the Tuesday of each week, peripatetic measurements of PN, BC, NO_2 and O_3 were made at each PDT site during morning (AM), lunchtime (Noon) and afternoon (PM) periods by walking around all sites with portable monitors. Tuesday was selected as representative of 'typical' traffic flows during the week. The AM and PM measurements took place during peak times with anticipated higher traffic flows (0800 – 0930 and 1600 – 1730), while Noon measurements were off-peak (1300 – 1430). A static, 6-minute 'spot' measurement of NO_2, O_3, BC and PN was made at each site on each weekly visit, giving 7 repeat peripatetic measurements at each site over the full study. The portable monitors were: for NO_2, an Aeroqual S500 with electrochemical sensor (ENW2, range 0 – 1 ppm); for O_3, an Aeroqual S500 with gas-sensitive semi-conductor sensor (OZU2, range 0 – 0.15 ppm), for BC, an Aethlabs microAeth AE51 portable Aethalometer; and for PN, a TSI Condensation Particle Counter (CPC 3007). The clocks of the monitors were synchronised before each measurement occasion. Sites were visited in numerical order on each measurement occasion, and the times at each site were accurately recorded.

In the rest of this paper, the measurements made on each of the 7 spot measurement days will be referred to using '*w/c X*' which corresponds to the week commencing *X*, where *X* is the week number of the study. The concentrations measured during consecutive days of peripatetic measurements were averaged to provide a single concentration for each site and week, referred to as '*Week X*', where *X* corresponds to the number of the start of the week (e.g. *Week 1* refers to the average concentrations from *w/c 1* and *w/c 2*).

2.2. Preparation, deployment and analysis of passive samplers

Static measurements of NO_2 were measured at each site using Palmes-style PDTs. The tubes were prepared in-house, using the dipping method to coat the stainless steel collection meshes with 1:1 TEA:acetone solution (Heal, 2008), for each week of the study. Duplicate

PDTs were deployed on a lamppost (~2 m height) at each site, using foam to provide a gap between the post and the PDTs, and secured using cable ties. This setup was selected to minimise the likelihood of theft of the PDTs as the sites in this work are located on busy streets. The PDTs were exposed for 1 week, to give six weekly measurements. Between preparation and deployment, and retrieval and analysis, the PDTs were stored in sealed bags in a fridge. Two laboratory blank PDTs were prepared and stored in the fridge unexposed for the duration of each study, then analysed with the field PDTs.

Laboratory analysis of the PDTs was carried out in-house using the Griess-Saltzman colorimetric method following guidelines provided by DEFRA (2008). Analysis was carried out for each week of PDT exposure as close to the date of retrieval as possible. The average NO_2 concentration during the PDT exposure, C_0 (ng/cm^3) was determined using the equation:

$$C_0 = \frac{QL}{DAt}$$

where Q is the nitrite mass measured in each PDT (ng), L and A are the length and area of the PDT (7.1 cm and 0.916 cm^2 respectively), D is the diffusion coefficient of NO_2 in air (0.151 cm^2/s), and t is the exposure time (s).

The limit of detection was below the concentrations of field samplers for all study weeks (study-average limit of detection (concentration of blank + 3*standard deviation of blanks) = 6.9 µg/m^3).

2.3. Measurements using portable monitors

The BC monitor was carried inside a backpack and sampled air through 1 m of conductive tubing attached to the shoulder strap of the backpack. The inlet was shielded using a conducting asbestos sampling inlet (SKC Ltd, UK) to prevent water ingress. Data was recorded every second and air was sampled at a flow rate of 150 mL/min. The filters in the BC monitor were changed each week. The raw BC data was post-processed using the Optimised Noise-Reduction Algorithm (ONA) on the Aethlabs website (https://www.aethlabs.com/dashboard). This algorithm, based on the equations published by Hagler et al. (2011), corrects the BC data for spurious peaks introduced by mechanical shock and vibrations by averaging the concentrations over a user-defined range of change in filter attenuation value (we used ΔATN = 0.01). The BC data was further processed to account for filter darkening effects, which leads to an underestimation of BC concentrations at high filter attenuation values, using the method described by Kirchstetter and Novakov (2007):

$$BC = BC_0(0.88Tr + 0.12)^{-1}$$

where $Tr = \exp\left(-\frac{ATN}{100}\right)$, BC_0 is the concentration measured by the instrument after ONA smoothing, and *ATN* is the filter attenuation value.

The PN monitor, which recorded data every second, was carried by hand as it requires to be kept horizontal to prevent solvent running into the detector. The isopropyl alcohol cartridge, used to grow the particles via condensation before they are counted, was replenished after each use. The PN data was downloaded from the monitor and was not post-processed prior to use.

The two Aeroqual monitors (NO_2 and O_3) were located in external mesh pockets on either side of the backpack, and thus sampled air at approximately chest height. Data was recorded at the highest temporal resolution of an instantaneous value every 1 minute (i.e. not an average of concentrations measured throughout the minute). Previous studies have identified that the Aeroqual monitors require calibration, particularly the NO_2 monitor which has cross-interference from O_3 (Lin et al., 2015; Masey et al., Submitted). We deployed the Aeroqual monitors at repeated intermittent intervals at Townhead between November 2015 and May 2016 to establish calibration equations to correct the mobile Aeroqual monitor concentrations (Table 1). The use of a global calibration equation derived from repeated calibration deployments has previously been shown to yield better agreement of adjusted concentrations with ambient measurements (Masey et al., Submitted). Therefore, the Aeroqual O_3 monitor concentrations were adjusted using a global ordinary least-squares calibration equation between *Aeroqual O_3 vs. Reference O_3* concentrations derived from the amalgamated calibration deployments. The Aeroqual NO_2 monitor measurements were corrected for their cross-sensitivity to O_3 concentrations using a global multi-linear regression calibration equation between the *Aeroqual NO_2 vs. Reference NO_2 + Aeroqual O_3* (Masey et al., Submitted). It should be noted that the pairing of the instruments in this research was different from that presented in Chapter 4 of this thesis as a result of one of the O_3 instruments being broken. As a result, the calibration equations presented in Table 1 below are different from those presented in the earlier chapter as a result in the change of Aeroqual NO_2-O_3 pairings.

Table 1: Dates of calibration deployments of Aeroqual instruments at the Townhead AURN site. The concentrations collected during deployments were amalgamated into a single data set and a global calibration equation derived for O_{3_aq} and NO_{2_aq} (following the procedure detailed in (Masey et al., Submitted)).

Calibration study	Date
November	11/11/2015 – 15/11/2015
December	09/12/2015 – 13/12/2015
February	09/02/2016 – 15/02/2016
April	31/03/2016 – 04/04/2016
May1	04/05/2016 – 08/05/2016
May2	18/05/2016 – 22/05/2016
Global calibration equations	
$NO_{2\_aq} = 0.33*NO_{2\_ref} + 0.68*O_{3\_aq} - 12.34$	
$O_{3\_aq} = 0.53*O_{3\_ref} + 12.53$	

None of the portable monitors used in this work are waterproof, meaning additional precautions had to be taken during adverse weather. On those days, the backpack housing the BC and Aeroqual monitors was worn on the front and was shielded from the rain using an umbrella. The PN monitor was also held in the shelter of the umbrella.

2.4. Corrections for temporal variations in background concentrations

To allow comparison between measurements made each week, the concentrations measured at each site were corrected for the background concentrations measured by the portable monitors at site 10 (Townhead) using a 'difference' method (Gillespie et al., 2017; Klompmaker et al., 2015). Briefly, the average concentration for each pollutant at site 10 was calculated and the difference between the study-average concentration and the concentration measured at site 10 each week was added to the concentrations measured at the other sites.

During w/c 5 the NO_2 and O_3 monitors failed at site 10 so no data was available for background correction using the method above. The concentration measured by the analyser at Townhead was used to estimate the concentration that would have been measured during the spot measurement if the instruments were operational, and these estimates of site 10 peripatetic concentrations were used (see Section 3.1 below).

3. Results and Discussion

3.1. Agreement between peripatetic measurements and automatic analyser concentrations

The weekly NO_2 concentrations measured by the PDT at the Townhead AURN station (site 10) were highly correlated with the exposure-averaged concentrations from the reference NO_2 analyser at this site (R^2 = 0.77, n = 6), although there was overestimation by the PDTs (Figure 2a). The higher correlation than reported previously at this site may be attributed to longer exposures and higher NO_2 concentrations (Masey et al., 2017).

The concentrations measured during site 10 peripatetic measurements were compared to the analyser concentrations measured during the 2-hour peripatetic measurement period. However note that no measure of reference PN was available at the Townhead AURN site. Linear relationships between the peripatetic and analyser concentrations were identified, with R^2 values of 0.72, 0.41 and 0.83 for BC, NO_2 and O_3 respectively (Figure 2b to 2d, n = 21). The correlation between the microAeth portable Aethalometer and reference BC concentrations was of similar magnitude to previously published co-location studies, whose R^2 values ranged from 0.75 to 0.90 (Cheng, 2013; Delgado-Saborit, 2012; Viana et al., 2015), and to co-location of this BC monitor at Townhead (R^2 = 0.80) (Masey et al., In preparation). The correlation between measurements from the Aeroqual monitors and reference analysers was slightly lower than published co-locations, which have R^2 values > 0.60 for NO_2 and R^2 > 0.90 for O_3 (Delgado-Saborit, 2012; Lin et al., 2015; Masey et al., Submitted). The large (~ 22 µg/m^3) intercept in the O_3 data suggests that the calibration applied overestimates the lower concentrations, which could be attributed to the long (6-month) temporal span of the data used to derive the O_3 calibration. It has previously been suggested that using a global calibration is better than a single dataset due to the wider range of pollution concentrations anticipated to be covered and the greater number of points (Masey et al., Submitted).

The high correlations between the 2-hour average analyser and the 6-minute (adjusted) peripatetic concentrations give confidence that the portable real-time instruments provide an estimate of the true values.

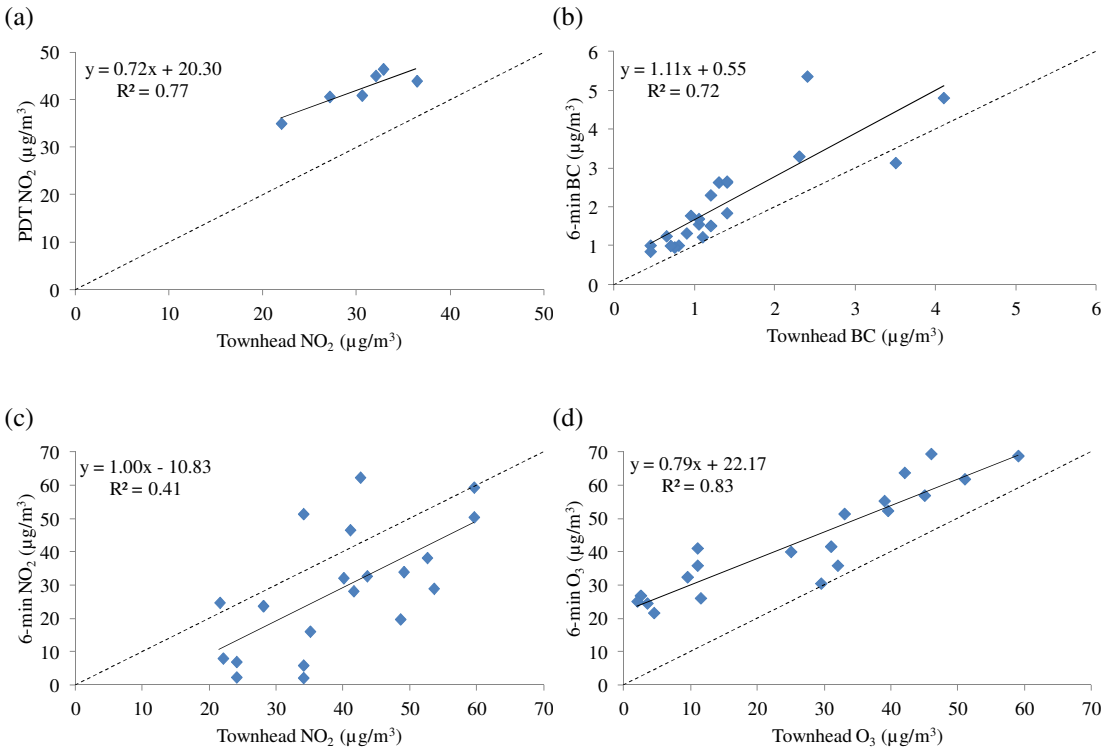


Figure 2: Relationship between measurements made at site 10 (Townhead) during this study and near-simultaneous measurements from AURN reference pollution monitors. (a) 1-week NO_2 measured using PDTs *vs.* 1-week chemiluminescence analyser (n = 6). (b) 6-min BC *vs.* reference AE22 Aethalometer (n = 21). (c) 6-min NO_2 *vs.* chemiluminescence analyser (n = 21 (AM, Noon & PM for 7 measurement periods). (d) 6-min O_3 *vs.* reference analyser (n = 21). AURN reference monitor measurements were selected from 2-hour 'window' during which the 6-min measurements were made. The solid lines are ordinary least-squares regression lines, and the dashed lines are 1:1 lines.

3.2. Temporal variations in pollution concentrations

3.2.1. Weekly variations

Similar spatial variations in pollution concentrations were observed each week, with high NO_2, BC and PN concentrations and low O_3 concentrations at roadside sites (Supplementary Information Figure S1). The stability of NO_2 and O_3 concentrations over the duration of our study were less pronounced than those published by Lin et al. (2016), who demonstrated relatively stable concentrations over consecutive study weeks, which could be attributed to the longer time-averaging of the measurements in the latter study. The similarity in BC and PN spatial trends between weeks was similar to our previously published work (Gillespie et al., 2017). During *w/c 6* the lower concentration sites had negative BC and PN

concentrations, which was a result of elevated particle concentrations during the background measurements at site 10.

The correlation between weekly and study-average peripatetic measurements was used to estimate the ability of the peripatetic measurements to predict longer-term concentration trends (Table S1 and Table 2). The average correlation between the weekly and study-average peripatetic BC measurements was $R^2 \geq 0.60$ for all times of day (weekly range: 0.01 to 0.96) (Table 2). The correlation between the peripatetic and study-average PN concentrations was slightly poorer than for the BC measurements ($R^2 = 0.09 – 0.91$), with the peripatetic PN measurements less able to represent the study-average concentrations during the off-peak measurements (average $R^2 = 0.35$). The average correlation between peripatetic and study-average O_3 concentrations decreased as the day went on and consequently the opposite was true for NO_2 as these pollutants are anti-correlated. For all hand-held instruments, the concentrations that were the average of the three periods throughout the day showed best agreement with the study-averaged concentrations (Table 2), which is also highlighted by the similarities in daily-average concentrations (Figure S1). This suggests that short, static measurements using these portable monitors can be used to give an indication of the pollution environment in an area and generally provide representative information about the spatial trends; however the concentrations recorded by the monitors should only be used as an indication as there are large variations between the concentrations each week. The hand-held instruments could be used to help identify monitoring sites for longer–term measurements, which could provide better indications of the 'true' pollution concentrations at the sites.

The weekly PDT NO_2 measurements were highly correlated with the study-average concentrations ($R^2 > 0.7$ for all weeks, average $R^2 = 0.88$) (Table S1), indicating the ability of the weekly PDT NO_2 measurements to predict longer-duration concentrations.

Table 2: Average coefficient of determination between concentrations measured each individual week (*w/c x*) and the overall study-average concentrations (n = 7 weeks for BC, PN, O_3 & NO_2 and n = 6 for PDT NO_2).

Pollutant	AM R^2 *(min – max)*	Noon R^2 *(min – max)*	PM R^2 *(min - max)*	Ave R^2 *(min – max)*
BC	0.60 (0.24 – 0.80)	0.65 (0.37 – 0.96)	0.61 (0.01 - 0.92)	0.79 (0.56 – 0.94)
PN	0.58 (0.11 – 0.85)	0.35 (0.09 – 0.68)	0.52 (0.17 - 0.73)	0.62 (0.30 – 0.91)
O_3	0.66 (0.38 – 0.94)	0.53 (0.09 – 0.80)	0.46 (0.32 - 0.66)	0.74 (0.50 – 0.92)
NO_2	0.42 (0.05 – 0.77)	0.92 (0.81 – 0.96)	0.93 (0.86 - 0.98)	0.91 (0.86 – 0.98)
PDT NO_2	-	-	-	0.88 (0.72 – 0.97)

3.2.2. Daily variations

Regardless of the time of day of the measurements, similar spatial trends were observed between the study-average pollutant concentrations measured at each site (Figure 3). The spatial trends of BC, PN and NO_2 were broadly similar to those measured by the PDTs, with higher concentrations present at roadside sites (e.g. site 2) and lower concentrations at sites further from roads (e.g. site 6). The measurements made during lunchtime periods generally had lower concentrations at each site for NO_2, BC and PN, which we anticipate is due to fewer vehicles on the road (Department for Transport statistics, 2015; Wu et al., 2015). The opposite was the case for O_3 which generally had higher concentrations during the lunchtime periods when less nitric oxide, emitted in traffic exhausts, was present to react photo chemically with O_3. Measured BC concentrations were similar between morning and afternoon periods, with the exception of sites 1-3, located in close proximity to a busy road, which had elevated concentrations during the afternoon. In comparison, the PN concentrations were largest during the morning measurements at sites 1-3 (Figure 3), which could suggest an additional source of PN compared to the BC measurements in the morning, such as a greater number of buses on the road, localised industrial sources or a larger number of diesel cars. The NO_2 measurements were similar between the noon and afternoon measurements, with elevated concentrations at site 1 in comparison to those NO_2 concentrations measured during the morning measurements and the BC and PN concentrations. The similarities in the spatial trends for measurements made at different times of day suggest that spatial information about pollution concentrations can be accurately gathered independent of the measurement time.

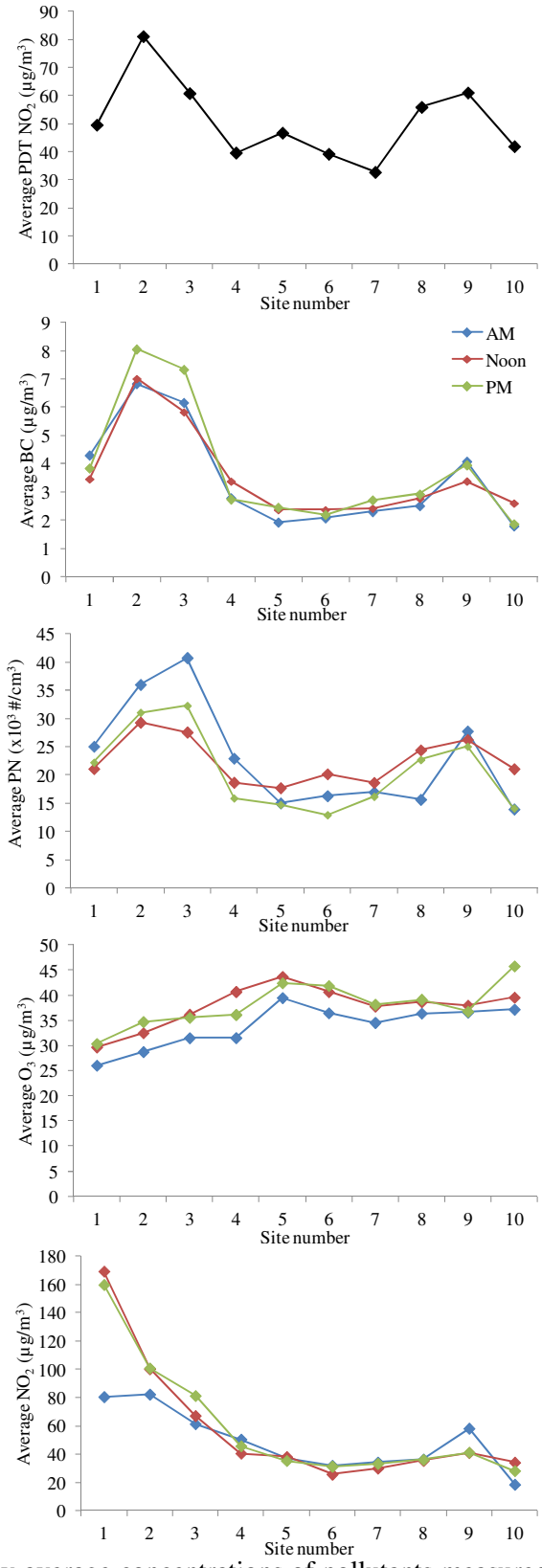


Figure 3: Study-average concentrations of pollutants measured at each site during the 7 peripatetic BC, PN, NO_2 and O_3 measurements and 6 weekly passive sampler NO_2 concentrations. Concentrations are shown for AM, Noon and PM measurements separately.

The ratio of within-site to between-site variance was calculated for each of AM, Noon, PM and daily-average measurements and compared to previously published values (Table 3) (Gillespie et al., 2017; Klompmaker et al., 2015). Our ratios of BC (all times) and PN (AM, PM and Average) concentrations were lower than previously published results, including those from our previous study which used the same instruments (Gillespie et al., 2017). The Noon measurements for PN showed much larger ratios which could, in part, be attributed to the very low concentrations of PN recorded during week 6 which, as previously discussed, was due to elevated concentrations at the background site (Figure S1). Those measurements made during peak times showed lower ratios for the particle metrics than the Noon studies, however the opposite was true for the gaseous pollutants. For all measurements the ratios for the daily-average concentrations had the lowest values, suggesting that in order to get the best estimate of long-term pollution concentrations from peripatetic measurements, measurements should be made at a variety of times throughout the day and a daily-average calculated. For a large number of sites this may not be practical due to time restraints, in which case particle measurements made during rush-hour are more representative of longer-term concentrations, while the opposite is true for gaseous pollutants. Our previous study also observed lower than anticipated variance ratios, based on the short-duration of the spot measurements, which we previously suggested could be due to the small size of the study area and short measurement period (Gillespie et al., 2017). However, the measurements in the current study were made at a different location to our previous work, and over a slightly longer duration, suggesting the temporal-stability of the pollutants at these short-time scales may be greater than those measurements made over longer durations.

Table 3: Ratio of within-site to between-site variation (after adjustment for temporal variation) from this study compared with other published studies (based on table published by Klompmaker et al. (2015) and Gillespie et al. (2016)). Na = not available.

Study (Within:Between)		BC	PN	O_3	NO_2	PDT NO_2
This study – AM		0.13	0.33	1.25	0.69	-
This study – Noon		0.20	1.88	0.35	0.04	-
This study – PM		0.15	0.22	0.91	0.06	-
This study – av		0.07	0.19	0.17	0.03	0.04
Gillespie et al. (2016)		0.21	0.77	-	-	0.05
Klompmaker et al. (2015)	MUSiC	2.44	2.17	-	-	-
	ESCAPE	0.09	-	-	-	-
	RUPIOH	-	0.31	-	-	-
	VE3SPA	0.69	Na	-	-	-

3.3. Correlations between mobile measurements of air pollutants

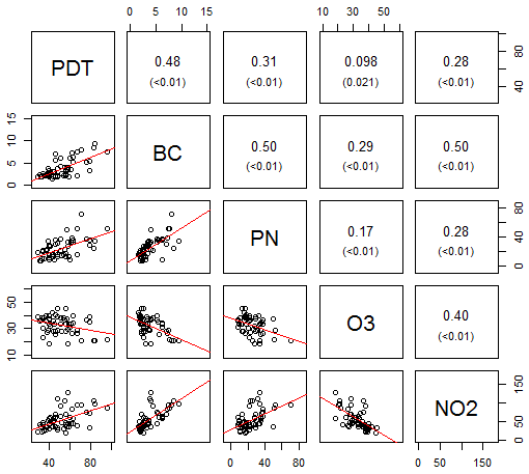
Greater correlations were observed between BC and PN during the afternoon and 'all' measurements in comparison to the morning and noon measurements (range R^2 = 0.50, 0.46, 0.73 and 0.84 for AM, Noon, PM and All) (Figure 4). The correlation between NO_2 and O_3 was similar between the AM, Noon and PM measurements (range R^2 = 0.40 – 0.41), with the highest correlation when the average of these measurement times was used (R^2 = 0.62). The correlation between the particle metrics and NO_2 was significant, with the exception of the PN noon measurements, with the highest R^2 values observed during the morning measurements (R^2 = 0.50 and R^2 = 0.28 for BC-NO_2 and PN-NO_2 respectively). The R^2 value between BC and O_3 have been reported to be 0.29 (Riley et al., 2016), which was similar to our AM measurements (R^2 = 0.29), however our correlations were much lower during Noon (R^2 = 0.15) and non-significant during PM (R^2 = 0.04).

Less scatter around the regression line, and consequently stronger correlations, were observed when study-average pollutant concentrations for each site were used (Figure 5). The correlation between BC and PN was consistently high (R^2 > 0.65) for all times of day and suggested that BC concentrations could provide an estimate of the particle number at a given location. Significant correlations of a similar magnitude (R^2 = 0.49 – 0.77) between BC and PN have been published by Wu et al. (2015), however some studies report lower correlation coefficients between these pollutants as a result of differences in the sources of the pollutants such as wood-burning in rural areas (R^2 = 0.08 – 0.86 reported by Gramsch et al.(2014)) and changes in fleet compositions (proportion of diesel / petrol cars, HGV and buses for example) and time of day measurements were made in urban areas (Reche et al., 2011). The BC concentrations also provided an estimate of NO_2 concentrations, which was more pronounced during the morning measurements (R^2 = 0.77). An outlier in the lunchtime and afternoon measurements degraded the correlations - excluding this outlier increased the R^2 values > 0.90 (data not shown). The R^2 values from our AM measurements were of similar magnitude to those reported by Beckerman et al. (2008) for 10 minute pooled measurements at peak time for BC-PN and BC-O_3 but lower for BC-NO_2, PN-O_3, and PN-NO_2 which could be attributed to the difference in study size, site number and the difference in the data collection (Beckerman used less portable instruments mounted in a vehicle). The BC-O_3 correlation exceeded that reported by Riley et al. (2016) during the morning and noon measurements. The correlation between NO_2 and O_3 was greater than 0.60 for each time of day, with the highest R^2 (0.83) when a daily average concentration was calculated from the

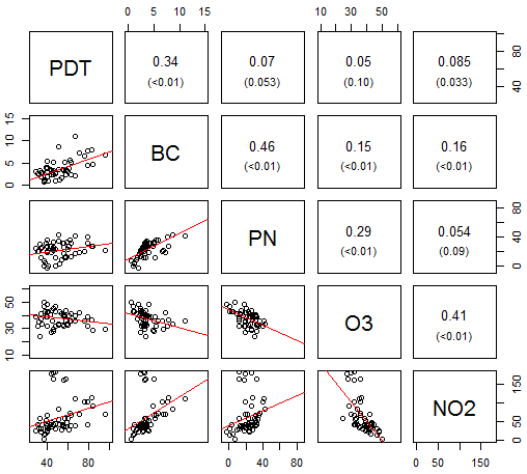
AM, Noon and PM measurements, however this large correlation was anticipated as O_3 is used to correct for the NO_2 monitor cross-sensitivity to O_3.

The above observations suggest that it may be possible to estimate the concentrations of PN and, to a lesser extent, NO_2 from BC measurements. However, the robustness of this estimation may be dependent on the time of day the measurements, with generally more accurate estimates during rush hour periods. The more repeat measurements made, the greater the ability to explain a higher proportion of the variance in a different pollutants concentration. The results presented in this research may be transferred to other urban locations anticipated to experience meteorological conditions, fleet compositions and emissions from domestic and industrial sources similar to Glasgow. Changes in fleet compositions or emission sources will likely change the concentrations of pollutants emitted and the proportions of these pollutants to one another in the ambient air will thus change, leading to changes in the correlations observed (Reche et al., 2011). Changes in meteorological conditions could change the dispersion of these pollutants from their sources thus changing their ratios in the ambient air. The results presented in this research are unlikely to be transferable to rural locations where the predominant sources of these emissions are not traffic-related. However for all cases above the methodology presented would be suitable to obtain estimates of these pollutant relationships under different conditions, which can help identify the key drivers influencing the relationships of these pollutants.

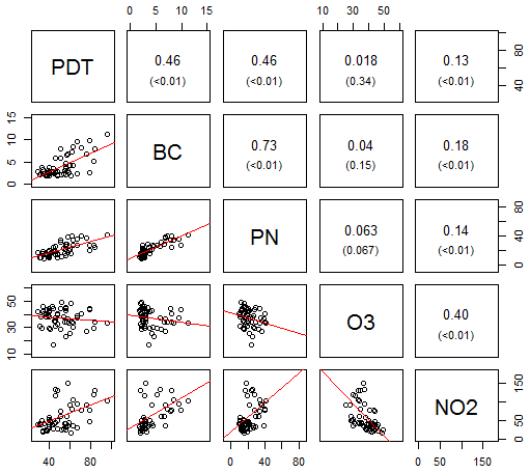
Concentrations during all AM periods



Concentrations during all Noon periods



Concentrations during all PM periods



Daily average concentrations during all days

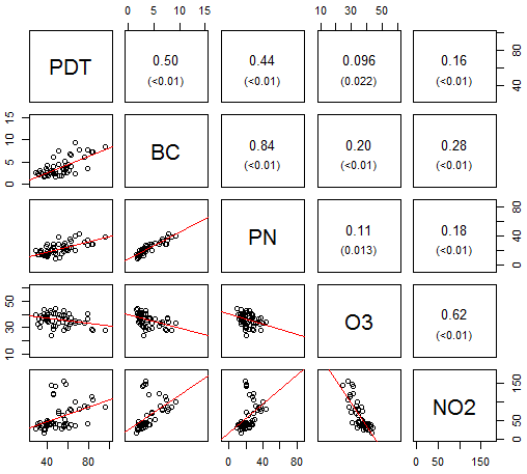
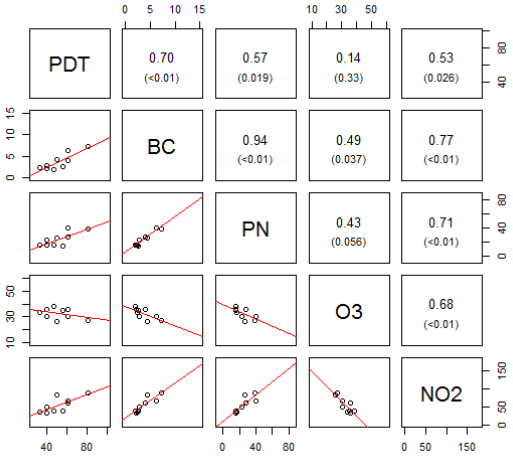
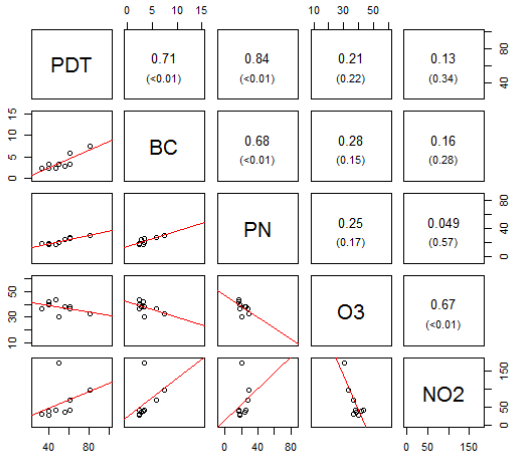


Figure 4: Correlation between average of 6-min measurements (BC, PN, O_3 and NO_2) made at the start and end of each week and 1-week passive sampler (PDT) concentrations for all measurement periods and sites (n = 60), subdivided by the time of day during which the 6-min measurements were made. The coefficient of determination (R^2) and p value for the relationships are shown. Measurements have been adjusted by 'difference' method using temporal variations observed at site 10 (outlined in detail in text).

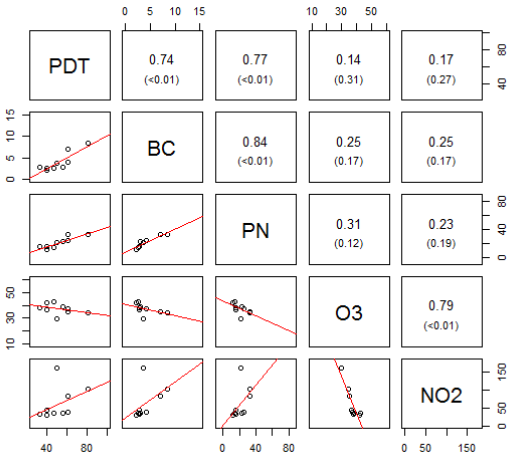
Average concentrations over all AM periods



Average concentrations over all Noon periods



Average concentrations over all PM periods



Average concentrations over all periods

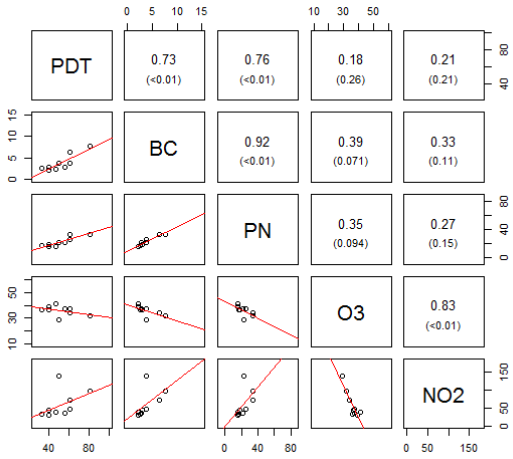


Figure 5: Correlation between 6-min and 1-week PDT measurements at each site (n = 10) averaged across all measurement periods, categorised by different times of day. The coefficients of determination (R^2) and p values for the relationships are also shown. The variables in the graph are described in the caption of Figure 4.

3.4. Peripatetic vs. weekly passive sampler concentrations

When all measurement periods are considered, the particle metrics had the strongest correlation with the longer-term PDT NO_2 concentrations (Figure 4). These correlations were greatest for measurements made during the peak times. The correlations between BC and PDT NO_2 concentrations in this work were greater than those reported by Riley et al. (2016) for winter measurements (R^2 = 0.38). We identified similar correlations between the BC, PN and PDT NO_2 concentrations for the weekly studies to our previous study: this study AM R^2 (BC, PN) = 0.48, 0.31; previous study R^2 (BC, PN) = 0.50, 0.24 (measurements also

made during morning rush-hour) (Gillespie et al., 2017). The slope and intercept between PDT NO_2 and BC in our previous study was larger than in the present study (slope = 9.0 vs. 7.1, intercept = 26.4 vs. 25.6 (previous vs. present)), while for PN the slope was larger while the intercept was smaller in our previous study (slope = 1.9 vs. 1.3, intercept = 15.0 vs. 19.8 (previous vs. present)) (Table S2). The broadly similar regression values and correlations between our two studies suggest that there is a relationship between these short-duration particle measurements made using the portable monitors and longer-term concentrations of NO_2 measured using passive diffusion samplers.

The correlations between the study-average peripatetic and PDT NO_2 concentrations were greater than when study weeks were considered independently, highlighting the importance of repeat measurements (Figure 5). The greater the number of repeat measurements peripatetic measurements combined the larger the correlation between the peripatetic and PDT measurements (R^2 = 0.71 *vs.* 0.59 *vs.* 0.52 when seven, four and two repeat BC measurements were considered – Figure S2). The correlation between BC and PDT NO_2 was similar between the AM, Noon and PM measurements (R^2 range 0.60 – 0.74), while for PN the correlation was more variable (R^2 range 0.57 – 0.84). These correlations were greater than those reported in our previous work (R^2 = 0.75 (BC-PDT NO_2) and R^2 = 0.33 (PN-PDT NO_2) for non-background adjusted concentrations (Gillespie et al., 2017). This suggests that the number of repeat peripatetic measurements was more important for deriving representative estimates of longer duration concentrations than the time of day during which the measurements were made. The correlation between these particle metrics and NO_2 measured using passive samplers were within the range of values reported from the ESCAPE and SAPALDIA studies, which co-located NO_2 and $PM_{2.5}$ measurements for 14 days (R^2 ranging from 0.21 to 0.79) (Durant et al., 2014; Eeftens et al., 2015), and, for the SAPALDIA study only, R^2 range 0.47 – 0.82 for NO_2 and PN (Eeftens et al., 2015). This suggests that longer-term concentrations of NO_2 can be estimated using short-duration (6-minute) particle measurements, with BC providing more accurate estimates than PN concentrations.

The O_3 and PDT NO_2 measurements were anti-correlated, with the highest correlations occurring during the Noon studies when O_3 concentrations are higher (R^2 = 0.14 - 0.21) (Figure 5). This correlation, however, was poorer than the results of Riley et al. (2016), which could be attributed to fewer sites and the different sampling methods used between the two studies. The correlation between peripatetic and PDT NO_2 concentrations was improved when the study-average concentrations are used compared to the all study data, however the correlations were not significant (Figure 5). There was an elevated spot NO_2 concentration

during the noon and PM studies at site 1 (Figure 3) which was inconsistent with the spatial trends of the BC, PN and PDT NO_2 concentrations. The Aeroqual monitors were left running in a laboratory between the AM, Noon and PM measurements and we suggest the elevated concentrations could be a result of rapid changes in exposure conditions when the Aeroqual monitors were moved from indoor to outdoor environment. Removing this outlier at site 1 increased the R^2 values to greater than 0.65 (data not shown). A limitation of this type of short-duration, peripatetic measurement is that a single outlier, such as in this example, can have a large influence on the results. We suggest that the Aeroqual monitors require time to equilibrate when moving between indoor and outdoor environments in order to prevent the occurrence of these non-representative measurements.

The correlation between the short peripatetic measurements and the longer-term NO_2 concentrations suggest that short-duration monitoring of BC and PN could be used as an indicative measure of longer-term NO_2 pollution concentrations in an area. However, peripatetic NO_2 and O_3 concentrations were not significantly correlated with PDT NO_2 concentrations and could not be used to estimate these longer-term concentrations. The short-duration monitoring tested in this work highlights the ability of these portable monitors to estimate spatial trends in pollution concentrations and these could be used to identify candidate sites for further monitoring.

4. Conclusions

Portable, real-time monitors measuring BC, PN, NO_2 and O_3, were used to make short duration (6-minute) peripatetic measurements at 10 sites around Glasgow city centre. We additionally made longer duration (1-week) measurements of NO_2 at the sites using passive diffusion tubes (PDTs). Measurements were made during morning peak time (AM), mid-day (Noon) and evening peak time (PM) to assess if time of day of the measurement has an impact on the relationships.

The peripatetic measurements made next to an urban background automatic monitoring station showed linear relationships with reference analyser concentrations with correlations of similar magnitude to previously published co-location studies. After accounting for changes in background concentrations, the spatial trends measured each week were similar, with roadside sites having elevated BC, PN and NO_2 concentrations in comparison to sites located further from traffic sources. The spatial trends measured by the portable BC, PN and NO_2 monitors were broadly consistent to those concentrations measured by the week-exposed NO_2 PDTs, showing the ability of peripatetic measurements to provide estimates of spatial trends over a geographical area.

The correlations between study-average peripatetic concentrations were highest during the AM measurements, with R^2 values of 0.94 (BC-PN), 0.49 (BC-O_3), 0.77 (BC-NO_2), 0.43 (PN-O_3), and 0.71 (PN-NO_2). This suggests that estimates of PN and NO_2 concentrations can be derived from BC measurements, with the most accurate estimates being measured in during the morning peak time.

The correlation between study-average peripatetic and passive longer-duration NO_2 concentrations poor (R^2 range 0.13 to 0.53), which was suggested to be due to an elevated NO_2 concentration at site 1 during the peripatetic measurements, highlighting the limitation of this short-duration measurements which can be influenced by atypical pollution occurrences during measurements. The correlation between weekly PDT NO_2 and peripatetic O_3 concentrations was not significant. The study-average peripatetic BC and PN measurements were correlated with weekly-average NO_2 concentrations from PDTs, regardless of measurement time (R^2 = 0.70 – 0.74 for BC, and R^2 = 0.57 – 0.84 for PN), suggesting that particle metrics can be used as an indicator of longer term NO_2 concentrations. Using many repeat measurements during peripatetic monitoring is more important than the time of day during which the measurements were made in order to obtain the best characterisation of the pollution trends at a site and minimising the impact of any atypical concentration measurements.

Acknowledgements

Nicola Masey is funded through a UK Natural Environment Research Council CASE PhD studentship (NE/K007319/1), with support from Ricardo Energy and Environment. Eliani Ezani is funded by the Ministry of Higher Education Malaysia (KPT(BS)860126295394). Jonathan Gillespie is funded through an Engineering and Physical Sciences Research Council Doctoral Training Grant (EPSRC DTG EP/L505080/1 and EP/K503174/1) studentship, with support from the University of Strathclyde and Ricardo Energy and Environment.

References

Beckerman, B., Jerrett, M., Brook, J.R., Verma, D.K., Arain, M.A., Finkelstein, M.M., 2008. Correlation of nitrogen dioxide with other traffic pollutants near a major expressway. Atmos. Environ. 42, 275–290.

Cheng, Y.-H., 2013. Real-Time Performance of the microAeth® AE51 and the Effects of Aerosol Loading on Its Measurement Results at a Traffic Site. Aerosol Air Qual. Res. doi:10.4209/aaqr.2012.12.0371

DEFRA, 2017. The Data Verification and Ratification Process [WWW Document]. URL https://uk-air.defra.gov.uk/assets/documents/The_Data_Verification_and_Ratification_Process.pdf (accessed 4.1.17).

Delgado-Saborit, J.M., 2012. Use of real-time sensors to characterise human exposures to combustion related pollutants. J. Environ. Monit. 14, 1824. doi:10.1039/c2em10996d

Department for Transport statistics, 2015. Average annual daily flow and temporal traffic distributions (TRA03): Table TRA0308 - Car traffic distribution on all roads by time of day in Great Britain, 2015.

Deville Cavellin, L., Weichenthal, S., Tack, R., Ragettli, M.S., Smargiassi, A., Hatzopoulou, M., 2016. Investigating the Use Of Portable Air Pollution Sensors to Capture the Spatial Variability Of Traffic-Related Air Pollution. Environ. Sci. Technol. 50, 313–320. doi:10.1021/acs.est.5b04235

Dons, E., Temmerman, P., Van Poppel, M., Bellemans, T., Wets, G., Int Panis, L., 2013. Street characteristics and traffic factors determining road users' exposure to black carbon. Sci. Total Environ. 447, 72–79.

Durant, J.L., Beelen, R., Eeftens, M., Meliefste, K., Cyrys, J., Heinrich, J., Bellander, T., Lewné, M., Brunekreef, B., Hoek, G., 2014. Comparison of ambient airborne PM2.5, PM2.5 absorbance and nitrogen dioxide ratios measured in 1999 and 2009 in three areas in Europe. Sci. Total Environ. 487, 290–298. doi:10.1016/j.scitotenv.2014.04.019

Eeftens, M., Phuleria, H.C., Meier, R., Aguilera, I., Corradi, E., Davey, M., Ducret-Stich, R., Fierz, M., Gehrig, R., Ineichen, A., Keidel, D., Probst-Hensch, N., Ragettli, M.S., Schindler, C., Künzli, N., Tsai, M.-Y., 2015. Spatial and temporal variability of ultrafine particles, NO2, PM2.5, PM2.5 absorbance, PM10 and PMcoarse in Swiss study areas. Atmos. Environ. 111, 60–70. doi:10.1016/j.atmosenv.2015.03.031

Gillespie, J., Beverland, I.J., Hamilton, S., Padmanabhan, S., 2016. Development, Evaluation, and Comparison of Land Use Regression Modeling Methods to Estimate Residential Exposure to Nitrogen Dioxide in a Cohort Study. Environ. Sci. Technol. 50, 11085–11093. doi:10.1021/acs.est.6b02089

Gillespie, J., Masey, N., Heal, M.R., Hamilton, S., Beverland, I.J., 2017. Estimation of spatial patterns of urban air pollution over a 4-week period from repeated 5-min measurements. Atmos. Environ. 150, 295–302. doi:10.1016/j.atmosenv.2016.11.035

Gramsch, E., Reyes, F., Oyola, P., Rubio, M.A., Lopez, G., Perez, P., Martinez, M., 2014. Particle size distribution and its relationship to black carbon in two urban and one rural site

in Santiago de Chile. J. Air Waste Manag. Assoc. 64(7), 785-796. doi:10.1080/10962247.2014.890141

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

Hankey, S., Marshall, J.D., 2015. On-bicycle exposure to particulate air pollution: Particle number, black carbon, PM2.5, and particle size. Atmos. Environ. 122, 65–73. doi:10.1016/j.atmosenv.2015.09.025

Heal, M.R., 2008. The effect of absorbent grid preparation method on precision and accuracy of ambient nitrogen dioxide measurements using Palmes passive diffusion tubes. J. Environ. Monit. 10, 1363–1369. doi:10.1039/b811230d

Kirchstetter, T.W., Novakov, T., 2007. Controlled generation of black carbon particles from a diffusion flame and applications in evaluating black carbon measurement methods. Atmos. Environ. 41, 1874–1888. doi:10.1016/j.atmosenv.2006.10.067

Klompmaker, J.O., Montagne, D.R., Meliefste, K., Hoek, G., Brunekreef, B., 2015. Spatial variation of ultrafine particles and black carbon in two cities: Results from a short-term measurement campaign. Sci. Total Environ. 508, 266–275. doi:10.1016/j.scitotenv.2014.11.088

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

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.

Masey, N., Ezani, E., Gillespie, J., Lin, C., Heal, M.R., Hamilton, S., Beverland, I.J., In preparation. Impact of filter shadowing correction algorithms on agreement of portable AE51 micro-aethalometer and reference AE22 aethalometer black carbon measurements.

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

Masey, N., Gillespie, J., Heal, M.R., Hamilton, S., Beverland, I.J., 2017. Influence of wind-speed on short duration NO2 measurements using Palmes and Ogawa passive diffusion samplers. Atmos. Environ. doi:http://doi.org/10.1016/j.atmosenv.2017.04.008

Montagne, D.R., Hoek, G., Klompmaker, J.O., Wang, M., Meliefste, K., Brunekreef, B., 2015. Land Use Regression Models for Ultrafine Particles and Black Carbon Based on Short-Term Monitoring Predict Past Spatial Variation. Environ. Sci. Technol. 49, 8712–8720. doi:10.1021/es505791g

Norris, G., Larson, T., 1999. Spatial and temporal measurements of NO2 in an urban area using continuous mobile monitoring and passive samplers. J. Expo. Anal. Environ. Epidemiol. 9, 586–593. doi:10.1038/sj.jea.7500063

Reche, C., Querol, X., Alastuey, A., Viana, M., Pey, J., Moreno, T., Rodriguez, S., Gonzalez, Y., Fernandez-Camacho, R., Sanchez de la Campa, A.M., de la Rosa, J., Dall-Osto, M., Prevot, 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. Atmos. Chem. Phys. 11, 6207-6227. doi:10.5194/acp-11-6207-2011

Riley, E.A., Schaal, L., Sasakura, M., Crampton, R., Gould, T.R., Hartin, K., Sheppard, L., Larson, T., Simpson, C.D., Yost, M.G., 2016. Correlations between short-term mobile monitoring and long-term passive sampler measurements of traffic-related air pollution. Atmos. Environ. 132, 229–239. doi:10.1016/j.atmosenv.2016.03.001

Targa, J., Loader, A., 2008. Diffusion Tubes for Ambient NO2 Monitoring: A Practical Guide (No. AEA/ENV/R/2504 – Issue 1a). AEA Energy and Environment.

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., Rivas, I., Reche, C., Fonseca, A.S., Pérez, N., Querol, X., Alastuey, A., Álvarez-Pedrerol, M., Sunyer, J., 2015. Field comparison of portable and stationary instruments for outdoor urban air exposure assessments. Atmos. Environ. 123, Part A, 220–228. doi:10.1016/j.atmosenv.2015.10.076

World Health Organization, 2013. Review of evidence on health aspects of air pollution – REVIHAAP project: final technical report.

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

Supplementary information

Spatial variations in urban concentrations of NO_2, O_3, BC and PN: interpretation of repeated co-located measurements of concentrations over 6-min and 1-week averaging periods

Nicola Masey[1], Eliani Ezani[1,2], Jonathan Gillespie[1], Scott Hamilton[3], Mathew R. Heal[4], Iain J Beverland[1]*

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]Department of Environmental and Occupational Health, Faculty of Medicine and Health Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, MALAYSIA

[3]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

[4]School of Chemistry, Joseph Black Building, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

*CORRESPONDING AUTHOR: Dr Iain J. Beverland, Department of Civil and Environmental Engineering, University of Strathclyde, 505F James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK; email: iain.beverland@strath.ac.uk; Tel: +44 141 548 3202

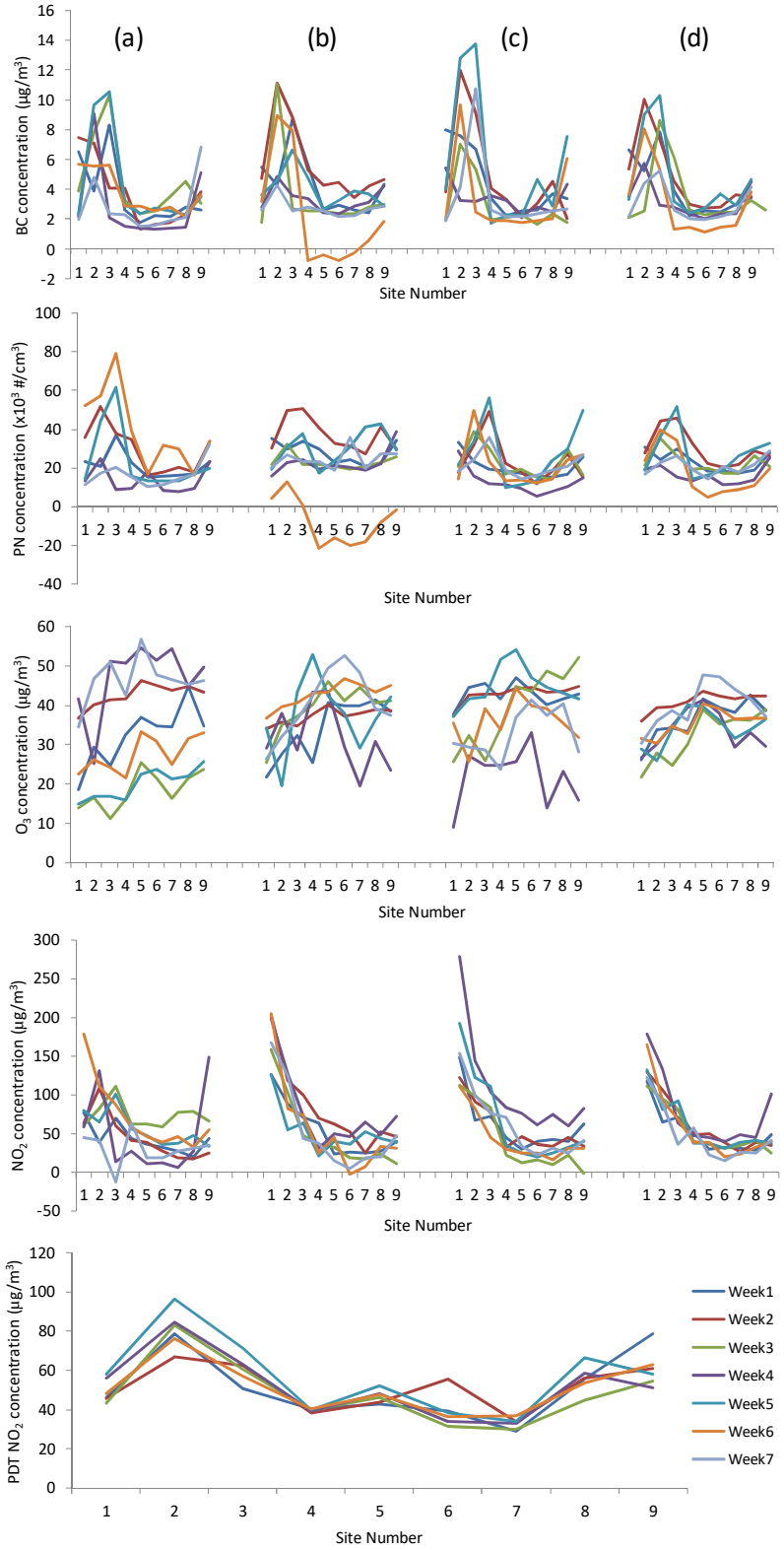


Figure S1: 6-min concentrations measurements measured during each week categorised by time of day (a) AM, (b) Noon (c) PM (d) average of all three periods in a given day. Concentrations are shown after adjustment for temporal variation using the 'difference' method (explained within main text).

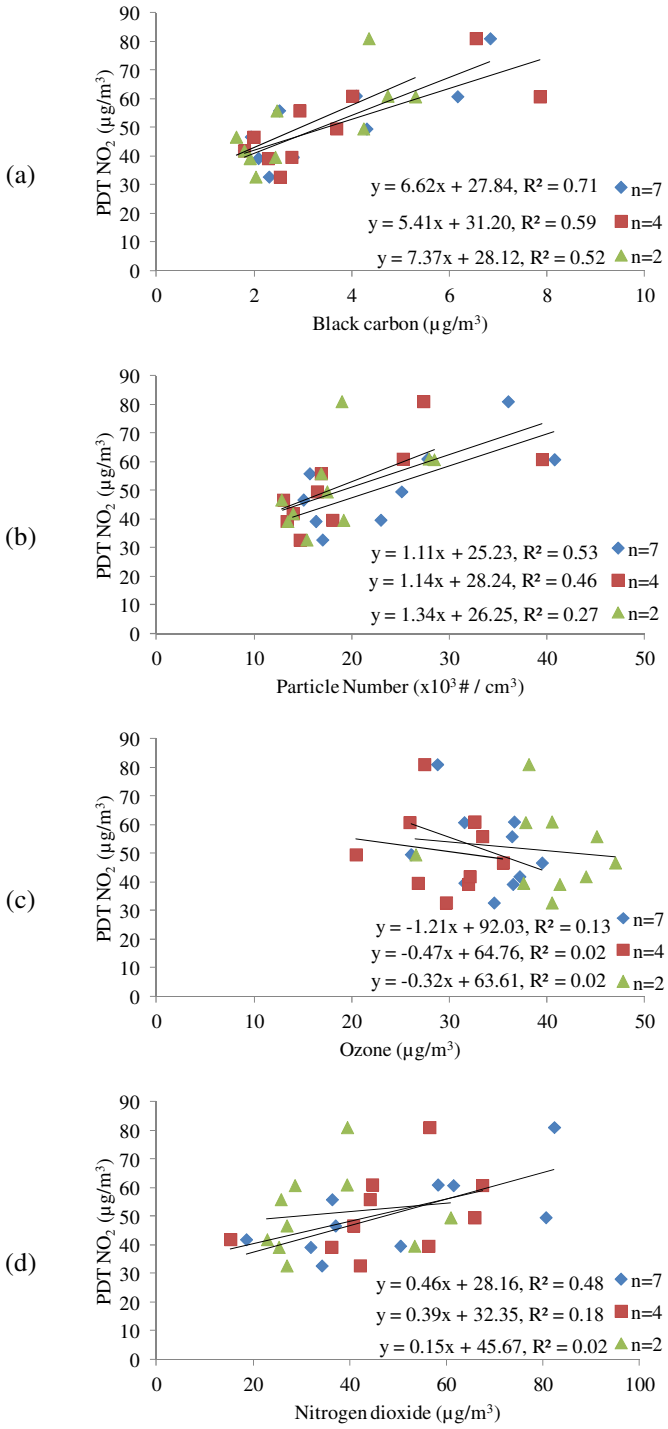


Figure S2: Scatter plots showing concentrations measured by 1-week PDT NO_2 measurements and 6-minute average concentrations of (a) BC, (b) PN, (c) O_3 and (d) NO_2 when seven (n=7), four (n=4) and two (n=2) weeks were used in the averaging process. Data is shown for the morning measurements, however results were observed for the other times of day.

Table S1: Correlations and reduced major axis regressions between concentrations measured each week (*w/c x*) and the study-average concentrations (all) for each pollutant (n = 9). Data shown for AM, noon and PM periods along with the average concentration measured during the different times of the day. Na = not available.

	AM			Noon			PM			Average		
W/C x	R^2	Slope (95 % CI)	Intercept (95% CI)	R^2	Slope (95 % CI)	Intercept (95% CI)	R^2	Slope (95 % CI)	Intercept (95% CI)	R^2	Slope (95 % CI)	Intercept (95% CI)
BC(all) vs. BC(w/c 1)	0.51	1.2 (0.4, 3.3)	-0.8 (-8.6, 2.2)	0.37	1.2 (0.1, 5.4)	-0.0 (-15.6, 3.8)	0.61	1.1 (0.5, 2.5)	0.1 (-5.6, 2.5)	0.59	1.1 (0.5, 2.5)	0.1 (-5.2, 2.4)
BC(all) vs. BC(w/c 2)	0.58	1.3 (0.5, 3.0)	-0.9 (-7.3, 1.8)	0.96	1.6 (1.3, 1.8)	-0.0 (-1.1, 0.9)	0.85	1.6 (1.1, 2.3)	-1.4 (-4.2, 0.7)	0.92	1.3 (1.0, 1.7)	-0.3 (-1.8, 0.9)
BC(all) vs. BC(w/c 3)	0.73	1.5 (0.8, 2.6)	-0.8 (-4.9, 1.7)	0.59	1.7 (0.7, 3.7)	-2.7 (-10.0, 1.0)	0.87	0.9 (0.6, 1.3)	-0.6 (-2.1, 0.5)	0.87	1.2 (0.8, 1.7)	-0.8 (-2.7, 0.5)
BC(all) vs. BC(w/c 4)	0.54	1.4 (0.5, 3.5)	-2.3 (-10.0, 1.0)	0.63	0.5 (0.2, 1.0)	1.5 (-0.5, 2.5)	0.01	0.2 (Na, Na)	2.8 (Na, Na)	0.56	0.6 (0.2, 1.5)	0.8 (-2.5, 2.3)
BC(all) vs. BC(w/c 5)	0.80	1.8 (1.2, 3.0)	-2.4 (-6.6, 0.1)	0.47	0.7 (0.2, 2.0)	1.5 (-3.3, 3.4)	0.92	2.1 (1.6, 2.8)	-2.9 (-5.5, 2.1)	0.90	1.6 (1.2, 2.1)	-1.2 (-3.4, 0.4)
BC(all) vs. BC(w/c 6)	0.80	0.8 (0.5, 1.3)	0.8 (-1.1,1.9)	0.92	2.3 (1.7, 3.0)	-6.2 (-8.8, -4.2)	0.48	1.2 (0.4, 3.7)	-1.7 (-11.6, 1.9)	0.94	1.3 (1.0, 1.6)	-1.7 (-3.0, -0.7)
BC(all) vs. BC(w/c 7)	0.24	0.9 (-0.2, -83.4)	-0.5 (308.8, 3.6)	0.62	0.4 (0.2, 0.8)	1.4 (-0.0, 2.2)	0.53	1.2 (0.4, 3.0)	-1.5 (-8.7, 1.9)	0.76	0.7 (0.4, 1.1)	0.5 (-1.3, 1.5)
CPC(all) vs. CPC(w/c 1)	0.73	0.7 (0.4, 1.2)	4.4 (-8.2, 12.1)	0.25	1.4 (-0.1, -7.9)	-3.5 (208.0, 29.6)	0.17	0.7 (-0.6, -2.7)	5.3 (77.6, 34.1)	0.50	0.8 (0.3, 2.7)	4.3 (-36.9, 16.8)
CPC(all) vs. CPC(w/c 2)	0.77	1.3 (0.8, 2.2)	-1.4 (-22.1, 11.4)	0.45	2.1 (0.6, 7.7)	-9.9(-136.9, 24.4)	0.61	1.5 (0.6, 3.2)	-9.3 (-44.6, 10.1)	0.80	1.5 (0.9, 2.3)	-3.1 (-22.4, 9.4)
CPC(all) vs. CPC(w/c 3)	-	No data	-	0.58	0.8 (0.3, 1.8)	3.9 (-18.5, 16.0)	0.64	1.2 (0.6, 2.5)	-3.0 (-30.6, 10.8)	0.64	0.9 (0.4, 1.7)	3.1 (-16.2, 14.1)
CPC(all) vs. CPC(w/c 4)	0.11	0.9 (Na, Na)	-6.7 (Na, Na)	0.28	1.2 (0.3, 5.8)	-4.4 (-108.1, 29.3)	0.16	0.7 (-0.7, -1032.9)	-1.8 (22188.7, 28.6)	0.30	0.7 (-0.1, 4.4)	1.5 (-83.7, 17.5)
CPC(all) vs. CPC(w/c 5)	0.82	1.8 (1.2, 2.9)	-20.8 (-45.2, -4.9)	0.14	2.00 (Na, Na)	-14.8 (Na, Na)	0.73	2.3 (1.3, 4.1)	-22.0 (-60.9, 0.2)	0.73	1.8 (1.0, 3.1)	-14.2 (-43.9, 4.7)
CPC(all) vs. CPC(w/c 6)	0.85	2.2 (1.4, 3.2)	-12.1 (-36.2, 5.0)	0.68	2.8 (1.4, 5.4)	-71.4 (-130.1, -40.0)	0.58	1.6 (0.6, 3.6)	-13.3 (-55.8, 7.8)	0.91	2.0 (1.5, 2.6)	-26.3 (-40.2, -15.5)
CPC(all) vs. CPC(w/c 7)	0.21	0.6 (-0.3, 33.1)	2.3 (-780.6, 23.5)	0.09	0.8 (Na, Na)	6.2 (Na, Na)	0.73	0.9 (0.5, 1.5)	2.6 (-11.0, 11.3)	0.48	0.7 (0.2, 1.8)	6.0 (-19.7, 17.2)
O_3(all) vs. O_3(w/c 1)	0.67	1.69 (0.82, 3.26)	-24.38 (-76.93, 4.94)	0.45	1.95 (0.69, 10.17)	-38.97 (-347.67, 8.20)	0.32	0.75 (0.06, 7.85)	14.73 (-249.61, 40.66)	0.81	1.30 (0.83, 2.05)	-10.63 (-37.60, 6.33)
O_3(all) vs. O_3(w/c 2)	0.94	0.70 (0.55, 0.86)	19.26 (13.25, 24.09)	0.69	0.48 (0.25, 0.92)	19.39 (2.63, 27.75)	0.66	0.61 (0.30, 1.22)	20.27 (-2.60, 31.91)	0.92	0.58 (0.45, 0.75)	20.05 (13.89, 24.81)
O_3(all) vs. O_3(w/c 3)	0.68	1.11 (0.57, 2.18)	-18.63 (-54.50, -0.77)	0.80	1.40 (0.87, 2.24)	-13.58 (-45.10, 6.49)	0.48	2.93 (1.09, 12.11)	-69.91 (-411.69, -1.36)	0.76	1.65 (1.00, 2.91)	-27.27 (-72.90, -3.67)
O_3(all) vs. O_3(w/c 4)	0.38	2.09 (0.35, 10.39)	-22.83 (-300.90, 35.31)	0.09	2.52 (Na, Na)	-62.79 (Na, Na)	0.37	2.03 (0.36, 12.22)	-53.86 (-433.40, 8.43)	0.50	1.38 (0.53, 4.67)	-16.06 (-135.18, 14.60)
O_3(all) vs. O_3(w/c 5)	0.77	0.92 (0.56, 1.55)	-10.78 (-32.11, 1.19)	0.32	2.09 (0.06, 14.91)	-40.79 (-522.39, 35.35)	0.50	1.45 (0.50, 4.39)	-9.09 (-118.58, 26.08)	0.56	1.23 (0.52, 3.28)	-10.23 (-84.22, 15.42)
O_3(all) vs. O_3(w/c 6)	0.68	1.13 (0.61, 2.31)	-10.15 (-49.69, 7.04)	0.63	0.74 (0.35, 1.60)	14.94 (-17.29, 29.68)	0.32	1.38 (0.05, 10.87)	-15.44 (-368.80, 34.37)	0.83	0.86 (0.57, 1.32)	4.54 (-12.39, 14.98)
O_3(all) vs. O_3(w/c 7)	0.50	1.30 (0.37, 3.31)	2.98 (-64.32, 34.11)	0.76	2.01 (1.20, 3.44)	-35.02 (-88.68, -4.52)	0.45	1.79 (0.59, 7.78)	-33.65 (-257.02, 10.97)	0.81	1.46 (0.93, 2.33)	-12.81 (-44.31, 6.20)
NO_2(all) vs. NO_2(w/c 1)	0.49	0.90 (0.24, 2.42)	-4.14 (-83.91, 30.18)	0.89	0.77 (0.56, 1.07)	7.38 (-10.72, 20.21)	0.89	0.81 (0.59, 1.11)	9.70 (-9.09, 23.92)	0.92	0.80 (0.61, 1.04)	5.76 (-8.30, 16.67)
NO_2(all) vs. NO_2(w/c 2)	0.71	1.43 (0.74, 2.55)	-30.69 (-89.53, 5.89)	0.95	1.15 (0.94, 1.40)	11.26 (-4.23, 24.12)	0.94	0.74 (0.60, 0.93)	11.05 (-0.52, 20.21)	0.94	1.00 (0.80, 1.27)	2.15 (-13.20, 14.27)
NO_2(all) vs. NO_2(w/c 3)	0.05	0.46 (Na, Na)	49.22 (Na, Na)	0.96	1.06 (0.89, 1.27)	-14.78 (-27.34, -4.31)	0.86	1.00 (0.70, 1.46)	-21.29 (-50.15, -2.24)	0.89	0.89 (0.65, 1.24)	2.74 (-17.64, 16.99)
NO_2(all) vs. NO_2(w/c 4)	0.42	2.76 (0.63, 11.34)	-96.33 (-546.99, 15.26)	0.93	1.12 (0.88, 1.44)	10.55 (-8.68, 25.68)	0.93	1.58 (1.24, 2.00)	8.03 (-18.40, 29.23)	0.86	1.38 (0.96, 2.00)	-2.72 (-39.29, 22.03)
NO_2(all) vs. NO_2(w/c 5)	0.41	1.08 (0.16, 3.81)	-0.52 (-144.15, 47.54)	0.81	0.63 (0.40, 0.98)	14.38 (-6.44, 28.62)	0.98	1.39 (1.23, 1.57)	-20.49 (-31.73, -10.58)	0.91	0.97 (0.75, 1.31)	0.61 (-18.22, 14.71)
NO_2(all) vs. NO_2(w/c 6)	0.77	2.33 (1.34, 3.81)	-49.63 (-127.18, 2.05)	0.95	1.33 (1.09, 1.62)	-25.34 (-43.06, -10.59)	0.95	0.73 (0.59, 0.89)	-1.47 (-11.87, 7.01)	0.98	1.29 (1.14, 1.45)	-18.07 (-27.80, -9.39)
NO_2(all) vs. NO_2(w/c 7)	0.06	0.47 (Na, Na)	4.92 (Na, Na)	0.93	1.19 (0.93, 1.52)	-19.14 (-39.30, -3.55)	0.94	1.00 (0.80, 1.25)	-1.30 (-17.10, 11.38)	0.89	0.99 (0.72, 1.37)	-10.16 (-32.09, 5.80)
PDT(all) vs. PDT (w/c 1)										0.81	1.2 (0.8, 1.9)	-10.5 (-47.1, 11.3)
PDT(all) vs. PDT (w/c 2)										0.72	0.8 (0.5 1.5)	10.0 (-26.6, 28.2)
PDT(all) vs. PDT (w/c 3)										0.93	1.1 (0.9, 1.4)	-9.1 (-25.0, 3.3)
PDT(all) vs. PDT (w/c 4)										0.91	1.1 (0.8, 1.5)	-5.1 (-23.9, 8.9)
PDT(all) vs. PDT (w/c 5)										0.94	1.3 (1.1, 1.7)	-12.0 (-29.5, 1.9)
PDT(all) vs. PDT (w/c 6)										0.97	0.9 (0.8, 1.0)	5.2 (-3.0, 12.0)

Table S2: Reduced major axis regression between weekly NO_2 concentrations (PDTs) and the weekly-average peripatetic measurements of BC, PN, O_3 and NO_2 made at each site. '*Week x*' refers to the average of the peripatetic BC, PN, NO_2 and O_3 measurements made at the start and end of week x , while for PDT it refers to the start week the PDTs were deployed (n = 9 for each week). '*All*' refers to when the weekly measurements were treated as a single data set (n = 54) and '*site*' refers to when the study-average concentrations were used in the regression (n = 9). Na = not available.

Pollutants	AM			Noon			PM			Average		
	R^2	Slope (95 % CI)	Intercept (95% CI)	R^2	Slope (95 % CI)	Intercept (95% CI)	R^2	Slope (95 % CI)	Intercept (95% CI)	R^2	Slope (95 % CI)	Intercept (95% CI)
PDT vs. BC (*week 1*)	0.11	0.13 (Na, Na)	-3.11 (Na, Na)	0.18	0.12 (-0.05, -0.22)	-1.14 (16.28, 7.79)	0.22	0.16 (-0.03, -0.64)	-3.29 (37.72, 6.13)	0.20	0.15 (0.00,-0.24)	-3.23 (16.70, 4.55)
PDT vs. BC (*week 2*)	0.3	0.19 (0.02, -6.48)	-5.89 (338.66, 3.07)	0.41	0.21 (0.04, 0.80)	-6.34 (-36.85, 2.70)	0.38	0.23 (0.04, 1.28)	-7.80 (-62.11, 1.89)	0.41	0.20 (0.04, 0.88)	-6.00 (-41.01, 2.08)
PDT vs. BC (*week 3*)	0.87	0.14 (0.10, 0.20)	-2.91 (-5.78, -0.96)	0.76	0.11 (0.06, 0.18)	-1.88 (-5.54, 0.26)	0.82	0.06 (0.04, 0.10)	0.18 (-1.53, 1.22)	0.90	0.10 (0.07, 0.13)	-1.28 (-2.94, -0.07)
PDT vs. BC (*week 4*)	0.69	0.17 (0.09, 0.33)	-5.10 (-13.70, -1.06)	0.37	0.05 (0.01, 0.40)	0.84 (-17.00, 3.08)	0.55	0.16 (0.07, 0.47)	-3.64 (-19.87, 0.98)	0.64	0.12 (0.06, 0.27)	-2.33 (-10.22, 0.78)
PDT vs. BC (*week 5*)	0.58	0.12 (0.05, 0.33)	-2.83 (-14.53, 0.90)	0.64	0.12 (0.06, 0.28)	-3.94 (-12.99, -0.42)	0.67	0.18 (0.09, 0.37)	-5.58 (-16.56, -0.73)	0.68	0.14 (0.07, 0.28)	-3.85 (-12.29, -0.29)
PDT vs. BC (*week 6*)	0.69	0.10 (0.06, 0.21)	-2.04 (-7.56, 0.36)	0.7	0.16 (0.09, 0.31)	-5.89 (-16.50, -2.07)	0.67	0.17 (0.09, 0.35)	-5.07 (-14.46, -1.18)	0.77	0.13 (0.08, 0.23)	-3.79 (-8.55, -1.13)
PDT vs. BC (*all*)	0.48	0.14 (0.11, 0.19)	-3.59 (-6.20, -1.76)	0.34	0.13 (0.08, 0.18)	-2.83 (-5.89, -0.60)	0.46	0.17 (0.12, 0.23)	-4.59 (-7.82, -2.35)	0.50	0.14 (0.10, 0.18)	-3.31 (-5.69, -1.60)
PDT vs. BC (*site*)	0.7	0.13 (0.07, 0.26)	-3.21 (-9.97, -0.10)	0.71	0.12 (0.07, 0.24)	-2.73 (-8.52, 0.11)	0.74	0.15 (0.09, 0.28)	-3.95 (-10.48, -0.62)	0.73	0.14 (0.08, 0.25)	-3.223 (-9.24, -0.23)
PDT vs.PN (*week 1*)	0.28	0.55 (0.06, -4.51)	-2.73 (256.66, 22.36)	0.29	0.31 (-0.02, 3.45)	16.54 (-144.01, 33.57)	0.18	0.30 (-0.25, -1.54)	6.46 (100.92, 34.81)	0.27	0.39 (0.00, -6.41)	6.75 (355.20, 26.88)
PDT vs. PN (*week 2*)	0.23	1.03 (-0.22, -6.13)	-23.58 (345.75, 41.02)	0.44	0.48 (0.12, 1.66)	5.42 (-55.55, 23.96)	0.36	0.82 (0.13, 6.49)	-19.14 (-312.03, 16.47)	0.43	0.62 (0.15, 2.42)	-5.67 (-98.48, 18.59)
PDT vs. PN (*week 3*)	0.53	0.45 (0.19, 1.56)	-8.19 (-61.60, 4.36)	0.44	0.28 (0.08, 1.20)	9.52 (-34.97, 19.02)	0.60	0.38 (0.17, 0.94)	-0.67 (-27.77, 9.24)	0.89	0.27 (0.20, 0.38)	6.39 (1.29, 10.03)
PDT vs. PN (*week 4*)	0.65	0.64 (0.33, 1.42)	-14.45 (-55.20, 1.36)	0.03	0.75 (Na, Na)	-12.36 (Na, Na)	0.37	0.65 (0.17, 11.02)	-13.65 (-551.83, 11.25)	0.48	0.46 (0.16, 1.73)	-2.04 (-67.83, 13.38)
PDT vs. PN (*week 5*)	0.37	0.94 (0.18, 6.80)	-22.15 (-356.48, 21.61)	0.62	0.41 (0.19, 0.90)	-11.66 (-40.08, 0.69)	0.63	0.68 (0.35, 1.61)	-14.07 (-67.47, 4.73)	0.62	0.60 (0.29, 1.38)	-11.70 (-56..23, 6.25)
PDT vs. PN (*week 6*)	0.28	0.83 (-0.05, 25.58)	-14.16 (-1279.51, 30.54)	0.69	0.51 (0.27, 0.98)	-17.32 (-41.68, -5.16)	0.84	0.61 (0.42, 0.92)	-9.90 (-25.45, 0.24)	0.66	0.60 (0.30, 1.25)	-11.29 (-44.62, 3.89)
PDT vs. PN (*all*)	0.31	0.77 (0.49, 1.16)	-15.28 (-35.67, -0.91)	0.07	0.86 (0.28, 4.65)	-22.49 (-219.13, 7.38)	0.46	0.61 (0.45, 0.84)	-10.10 (-21.96, -2.00)	0.44	0.52 (0.38, 0.72)	-4.38 (-14.83, 3.06)
PDT vs. PN (*site*)	0.57	0.72 (0.33, 2.07)	-12.67 (-83.12, 7.42)	0.84	0.32 (0.22, 0.50)	5.18 (-3.80, 10.59)	0.77	0.53 (0.33, 0.92)	-5.91 (-25.98, 4.74)	0.76	0.49 (0.30, 0.86)	-2.46 (-22.06, 7.37)
PDT vs. O_3 (*week 1*)	0.00	-0.07 (Na, Na)	40.99 (Na. Na)	0.01	-0.40 (Na, Na)	56.07 (Na, Na)	0.05	0.07 (Na, Na)	39.17 (Na, Na)	0.00	-0.04 (Na, Na)	40.88 (Na, Na)
PDT vs. O_3 (*week 2*)	0.00	-0.17 (Na. Na)	39.15 (Na, Na)	0.07	-0.22 (Na, Na)	49.56 (Na, Na)	0.01	-0.44 (Na, Na)	63.74 (Na, Na)	0.02	-0.35 (Na, Na)	54.58 (Na, Na)
PDT vs. O_3 (*week 3*)	0.48	-0.34 (-1.06, -0.10)	49.30 (37.71, 83.93)	0.00	0.16 (Na, Na)	27.69 (Na, Na)	0.04	-0.36 (Na, Na)	47.92 (Na, Na)	0.13	-0.27 (Na, Na)	45.85 (Na, Na)
PDT vs. O_3 (*week 4*)	0.65	-0.36 (-.074, -0.18)	52.12 (42.74, 72.16)	0.04	-0.38 (Na, Na)	54.41 (Na., Na)	0.03	-0.46 (Na, Na)	57.22 (Na, Na)	0.25	-0.34 (1.33, -0.03)	51.26 (35.34, -35.12)
PDT vs. O_3 (*week 5*)	0.02	-0.57 (Na, Na)	56.49 (Na, Na)	0.38	-0.26 (-1.18, -0.04)	55.10 (42.34, 107.45)	0.41	-0.23 (-0.95, -0.04)	53.51 (42.97, 94.88)	0.39	-0.19 (-0.94, -0.04)	45.39 (36.82, 88.36)
PDT vs. O_3 (*week 6*)	0.02	0.16 (Na, Na)	28.87 (Na, Na)	0.36	-0.43 (-2.38, -0.06)	63.36 (44.41, 163.43)	0.37	-0.43 (-6.24, -0.11)	56.70 (40.19, 353.25)	0.20	-0.35 (0.92, 0.09)	55.73 (33.01, -9.29)
PDT vs. O_3 (*all*)	0.10	-0.52 (-1.68, -0.22)	59.41 (43.78, 119.78)	0.05	-0.38 (-3.09, -0.05)	57.59 (40.43, 197.34)	0.02	-0.28 (6.56, 0.32)	51.29 (20.66, -303.04)	0.10	-0.30 (-0.85, -0.11)	51.36 (41.43, 79.72)
PDT vs. O_3 (*site*)	0.14	-0.35 (Na, Na)	50.44 (Na, Na)	0.21	-0.28 (0.83, 0.07)	52.29 (34.11, -5.44)	0.14	-0.26 (Na, Na)	50.68 (Na, Na)	0.18	-0.26 (0.49, 0.14)	49.11 (28.26, 10.40)
PDT vs. NO_2 (*week 1*)	0.13	1.39 (Na, Na)	-27.28 (Na, Na)	0.04	1.78 (Na, Na)	-23.18 (Na, Na)	0.04	1.68 (Na, Na)	-27.02 (Na, Na)	0.06	1.59 (Na, Na)	-24.57 (Na, Na)
PDT vs. NO_2 (*week 2*)	0.30	1.67 (0.09, -208.91)	-27.38 (10839.91, 54.10)	0.02	2.71 (Na, Na)	-74.56 (Na, Na)	0.10	3.73 (Na, Na)	-142.72 (Na, Na)	0.08	3.45 (Na, Na)	-119.80 (Na, Na)
PDT vs. NO_2 (*week 3*)	0.64	1.81 (0.93, 4.15)	-26.43 (-139.25, 16.12)	0.13	3.71 (Na, Na)	-114.56 (Na, Na)	0.13	3.98 (Na, Na)	-117.50 (Na, Na)	0.28	2.65 (0.27, -17.44)	-61.26 (908.18, 53.43)
PDT vs. NO_2 (*week 4*)	0.57	1.87 (0.84, 5.16)	-45.01 (-215.73, 8.78)	0.19	2.35 (-1.24, -7.81)	-56.07 (471.67, 130.33)	0.24	4.12 (-0.26, -23.03)	-127.26 (1282.77, 100.47)	0.34	2.61 (0.42, 36.24)	-67.13 (-1813.52, 46.70)
PDT vs. NO_2 (*week 5*)	0.23	1.84 (-0.05, -5.53)	-40.95 (379.80, 67.02)	0.13	2.22 (Na, Na)	-72.76 (Na, Na)	0.31	2.54 (0.37, -109.56)	-89.41 (6312.05, 34.37)	0.23	2.16 (-0.26, -8.46)	-65.18 (541.27, 73.02)
PDT vs. NO_2 (*week 6*)	0.06	2.18 (Na, Na)	-60.09 (Na, Na)	0.15	4.07 (Na, Na)	-153.81 (Na, Na)	0.14	3.04 (Na, Na)	-102.44 (Na, Na)	0.13	2.95 (Na., Na)	-98.07 (Na, Na)
PDT vs. NO_2 (*all*)	0.28	1.75 (1.10, 2.83)	-35.51 (-91.36, -1.84)	0.08	3.88 (1.54, 16.08)	-139.30 (-772.05, -17.79)	0.13	2.62 (1.08, 5.52)	-73.06 (-223.52, 7.21)	0.16	2.89 (1.57, 6.30)	-89.88 (-266.82, -21.17)
PDT vs. NO_2 (*site*)	0.53	1.56 (0.66, 5.06)	-25.94 (-207.35, 20.94)	0.13	3.59 (Na, Na)	-124.24 (Na, Na)	0.17	3.58 (-0.91, -3.84)	-122.70 (262.31, 110.36)	0.21	2.81 (-0.22, -6.81)	-85.66 (413.34, 71.70)

6. Influence of Weather Research Forecasting model parameterization on predictions from an air pollution dispersion model

The use of models to estimate population exposure to air pollution has been widely used in epidemiology studies. The inclusion of meteorology in these models allows higher temporal resolution and more accurate estimates to be made. Meteorological measurements can be obtained from the nearest weather station; however these can often be located far from the study area in question. The Weather Research Forecasting (WRF) model can be used to estimate meteorology in a study area that lacks a nearby weather station in order to allow these models to be run. We investigate if the WRF set up used impacts the outcome of a dispersion model.

N. Masey designed the experiment, ran the models, carried out data analysis and produced the manuscript. S. Hamilton assisted with experimental design, provided support for model running and data analysis; and both S. Hamilton and I. Beverland provided editorial comments on the manuscript.

This manuscript has been formatted with the intention to submit to Atmospheric Environment.

Influence of WRF model parameterization on predictions from an air pollution dispersion model

Nicola Masey[1], Iain J. Beverland[1], Scott Hamilton[2]

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

Highlights:

- WRF over predicted temperature and under predicted wind speed observations
- Fine-scale (< 1.5 km) horizontal WRF resolution did not improve model evaluation statistics
- No detrimental effect on dispersion model accuracy using fewer vertical levels
- WRF suitable AERMOD input to model dispersion from 10 m stacks
- Addition of urban canopy model had limited effect on modelled concentrations

Glossary:

WRF (Weather Research Forecasting), UCM (Urban Canopy Model), MMIF (Mesoscale Model Interface Programme)

Abstract

The Weather Research Forecasting (WRF) model has been used to estimate meteorological information for air pollution models in the absence of observations. We evaluated the effects of changing the configuration of a WRF model on air pollution concentrations predicted in a hypothetical dispersion study in Glasgow, UK; including the effects of horizontal grid resolution, the number of vertical levels and inclusion of an Urban Canopy Model (UCM). The WRF predicted temperature, wind speed and wind direction measurements were compared to observations at two meteorological stations in Glasgow to establish the accuracy of the WRF model. Predictions from an AERMOD pollution dispersion model using WRF predicted meteorological variables were compared against the same model run with observed meteorological variables. Reduction of the horizontal grid from 3 x 3 km to 1.3 x 1.3 km had only small effects on the agreement between modelled and observed meteorological variables (Temperature mean bias -1.22 °C (3 km), -1.40 °C (1.3 km)). Reducing the number of vertical levels from 45 to 20 had minimal effect on modelled wind speed and temperature; however deviations in wind direction were observed

for fewer vertical levels. Despite these less accurate wind direction estimates, dispersion models using fewer vertical levels estimated similar monthly-average concentrations to the model using 45 levels suggesting that, for pollution modelling, the number of levels can be reduced, minimising run times and file sizes (e.g. approximately 30 % saving in time). Adding a UCM to the WRF model reduced the accuracy of temperature (-1.22 (no UCM) *vs.* -1.55 (UCM) °C), improved the wind speed estimates (1.08 (no UCM) vs. 0.63 (UCM) m/s) but had little effect on the AERMOD dispersion model (MB = 0.01 $\mu g/m^3$ for 10 and 50 m stacks for WRF with and without UCM). The concentrations estimates from the AERMOD dispersion model using WRF as the input meteorological data was correlated with concentrations estimated by AERMOD using observation meteorological input data (Index of Agreement = 0.90) for 10 m stack height. However, concentrations estimated for 50 m and 100 m stack heights using WRF input to AERMOD were approximately 1.5 times greater than equivalent predictions using observed meteorological data.

Keywords: WRF, AERMOD, Air pollution, sensitivity analysis, dispersion model.

1. Introduction

Air quality dispersion models have been used to estimate public exposure to air pollution to help identify the health effects of air pollution (Batterman et al., 2014; de Hoogh et al., 2014). Detailed surface and vertical meteorological data is important to enable accurate model estimates of the dispersion of pollutants (Tiwari et al., 2013). However, meteorological data is often only available at a small number of observation stations which can be located a distance from the area to be modelled and often lack vertical measurements. Recent studies have used weather models to estimate meteorological conditions over study areas rather than relying on nearby weather stations (Appel et al., 2010; Baker et al., 2013; Borge et al., 2008). Weather models provide advantages of finer spatial resolution, upper level weather estimates, and capability to predict short-term forecasts.

In previous studies the Weather Research Forecasting model (WRF) has provided accurate predictions of surface wind speed/direction (Tartakovsky et al., 2015) and temperatures (Carbonell et al., 2013). However Carbonell et al. (2013) noted that WRF overestimated wind speeds and wind direction estimates deviated from observations at several meteorology stations. An Urban Canopy Model (UCM) can be added to the WRF model to improve simulation of convective and turbulent processes associated with urban land use, however the parameters assigned to the UCM have to be derived from available land use data (Lee et al., 2011).

Relatively few studies have investigated the effect of WRF model set up on the predictions of pollution models. Changing the Planetary Boundary Layer (PBL) scheme used in the WRF model influenced the dispersion and concentrations of a pollution plume estimated using the FLEXPART dispersion model (Madala et al., 2015). A tracer study evaluating the WRF model combined with a chemistry pollution model (WRF-CHEM) optimised the WRF set up to increase the accuracy of temperature, wind speed, and pollution estimates (Saide et al., 2011).

We investigated the effect of initial WRF model set up on estimates from the AERMOD dispersion model, including horizontal and vertical resolution, and inclusion of a UCM. Firstly we examined the agreement between WRF predictions and observed data at two meteorological stations in Glasgow, UK. We then examined the effect of WRF set up on AERMOD pollution estimates, to establish if pollution modelling has different WRF requirements and acceptance criteria to weather modelling.

2. Methods

2.1. WRF model set up and data

WRF-Environmental Modelling System (EMS) (v 3.4.1.14.16) was used to run the WRF model. This uses code from WRF model (Skamarock et al., 2008) in combination with scripts to acquire boundary conditions to make modelling more user-friendly while retaining the same modelling accuracy as the WRF model (Qiu et al., 2015). The settings used for the WRF model are listed in Table 1. Three domains were used in the work - the largest over the UK (27 km grid), the second over Scotland (9 km grid) and a third over Greater Glasgow (Figure 1). The model setup of domain three was changed in this work to investigate the effect of horizontal and vertical resolutions, and inclusion of a UCM. The ratio of the grid size of domain two to domain three was changed from 3 to 5 to 7 to vary the horizontal resolution in domain three (leading to horizontal resolutions of 3 km, 1.8 km and 1.3 km). The number of vertical levels tested was 45, 30 and 20 (the height of these was automatically assigned by the model). The UCM model requires additional user information, such as average building heights for different land use classes, and the user specified values are given in Supplementary Information Table S1. Some default values were used (due to better values being unknown) while other values were identified using Geographical Information Systems and CORINE land cover data (available from https://eip.ceh.ac.uk/).

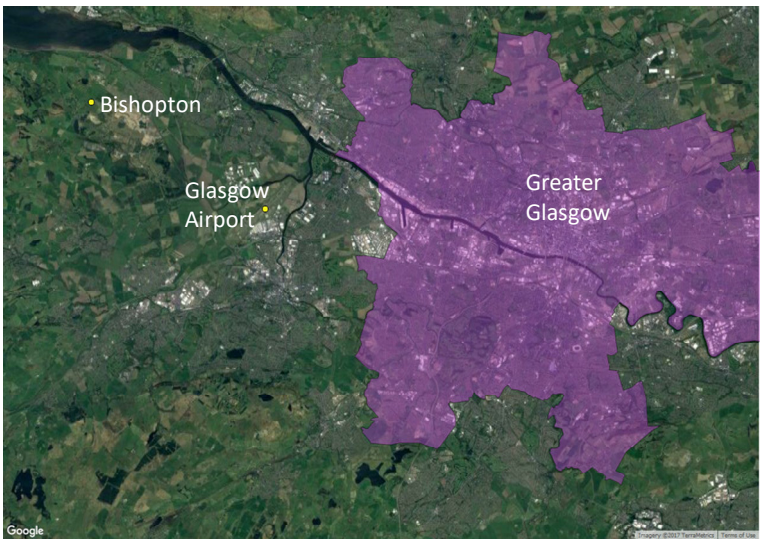


Figure 1: Location of WRF domain three over Glasgow (resolution varied from 3 x 3 km to 1.8 x 1.8 km to 1.3 x 1.3 km). The locations of Bishopton (latitude 55.90°, longitude -4.53°) and Glasgow Airport (latitude 55.87°, longitude -4.43°) meteorological stations in relation to the city of Glasgow are also shown.

Table 1: Details of setup of WRF models.

	Experiment							
Parameter	Hor3	Hor5	Hor7	Vert45	Vert30	Vert20	UCM	**References**
Boundary Data	Climate Forecast System Reanalysis and Reforecast							Saha et al. (2010)
PBL	Mellor-Yamada-Janjic							Janjic (1994)
Microphysics scheme	Lin et al.							Lin et al. (1983)
Cumulus scheme	Kain Fritsh (domains 1 and 2 only)							Kain (2004)
Long wave radiation	Rapid Radiative Transfer Model							Mlawer et al. (1997)
Short wave radiation	Dudhia scheme							Dudhia (1988)
Surface Physics	Noah Land Surface Model							Chen and Dudhia (2001)
Land Use Data	United States Geographical Survey							Grossman-Clarke et al. (2005)
Horizontal resolution (km)	3x3	1.8x1.8	1.3x1.3	3x3	3x3	3x3	3x3	
Vertical Resolution (levels)	45	45	45	45	30	20	45	
UCM	No						Yes	

Hourly WRF model outputs were extracted at Glasgow Airport (Latitude 55.872°, Longitude -4.433°) and Bishopton (Latitude 55.900°, Longitude -4.533°) meteorological

stations using Integrated Data Viewer (IDV) software (Unidata and UCAR, 2015). The WRF predicted temperature, wind speed and wind direction were compared to the observation data (downloaded from http://mesonet.agron.iastate.edu/sites/locate.php for Glasgow Airport and from http://www.ncdc.noaa.gov/ for Bishopton).

The model was run for June 2015 as a wide range of meteorological conditions were observed at the sites in this month (temperature range: 3.0 – 25.5 °C, wind speed: 0.3 – 13.4 m/s). The model was run for 8 days before reinitialisation, as shorter WRF model runs have been shown previously to perform better than a single longer run (Mohan and Sati, 2016). The first 24-hour period of each model overlapped the previous model to allow a day for model to spin up before using the modelled data.

2.2. Dispersion model setup

The dispersion model used in this work was AERMOD, and its' associated meteorological pre-processor AERMET (USEPA, 2015). This model uses meteorological observations to estimate dispersion curves and predict concentrations at receptors. AERMOD was run for flat terrain with no nearby buildings to influence downwash to simplify the model. A hypothetical stack was located at the Glasgow Airport meteorological station with the following parameters: emission rate 1 g/s; stack exit temperature 150 degrees; exit velocity 15 m/s; and diameter 1 m. The stack height was varied from 10 m to 50 m to 100 m to determine if the height of the stack influences the choice of WRF model set up required.

Receptors were located every 50 m up to a distance of 2 km from the stack, giving a total of 6561 receptor locations, to compare the models at a localised scale. Concentrations of a passive tracer were reported hourly and then a monthly-average pollutant concentration was calculated for each receptor. The passive tracer can be thought to represent chemically inert pollutants such as carbon monoxide or particulate matter. The concentrations predicted when WRF generated meteorology (for the different set ups discussed above), referred to as WRF-AERMOD, was used was compared to those predicted when meteorological observations were used, referred to as Observed-AERMOD. The meteorological data used for Observed-AERMOD was from Glasgow Airport for surface observations, and upper air data was obtained from the radiosonde station at Albemarle (data available from http://esrl.noaa.gov/raobs/). In order to get the WRF data in an AERMET readable format for WRF-AERMOD, the MMIF (Mesoscale Model Interface) programme was used. MMIF can also be used to transform the WRF data into data directly readable by AERMOD,

however this is not advised by the EPA for air quality modelling (U.S. Environmental Protection Agency, 2015).

2.3. Model evaluation

The metrics used to assess model performance against the observation data were Mean Bias (MB), Root Mean Square Error (RMSE) and Index of Agreement (IoA):

$$\mathrm{MB} = \frac{1}{\mathrm{n}}\sum_{j=1}^{J}\sum_{i=1}^{I}\left(P_j^i - O_j^i\right)$$

$$RMSE = \left[\frac{1}{n}\sum_{j=1}^{J}\sum_{i=1}^{I}\left(P_j^i - O_j^i\right)^2\right]^{1/2}$$

$$IOA = 1 - \left[\frac{n * RMSE^2}{\sum_{j=1}^{J}\sum_{i=1}^{I}\left|P_j^i - M_o\right| + \left|O_j^i - M_o\right|}\right]$$

Where: O_j^i is individual observed quantity at site i and time j, n is the number of observations (i.e. the number of sites multiplied by the number of time periods); P_j^i is the individual predicted quantity at site i and time j; and M_o is the mean observation which is calculated from all sites with valid data within region and for given time period:

$$M_o = \frac{1}{n}\sum_{j=1}^{J}\sum_{i=1}^{I} O_j^i$$

Model statistics were calculated using the R package 'openair' (Carslaw and Ropkins, 2012).

3. Results and discussion

3.1. Agreement between WRF predictions and meteorological observations

The temperatures predicted by the WRF model generally under predicted temperature and over predicted wind speeds at the weather stations (Figures 2 & 3 (Glasgow Airport), and Supplementary Information Figures S2 & 3 (Bishopton)). At low wind speeds (~0.5 m/s) this effect is less pronounced and the over estimation increased with greater wind speeds (Figure 3). These general trends in WRF accuracy for temperature and wind speed have been commonly reported in the literature (Borge et al., 2008; Mohan and Sati, 2016; Ritter et al., 2013). The WRF estimated wind direction is different from the observed data, especially at the airport site (Figures 3 & S4). The WRF modelled wind roses for Bishopton and the Airport sites are similar to each other, whereas the observed wind roses are quite different, despite the reasonably close proximity of the two sites (approximately 7 km from

each other). Quartile-Quartile (Q-Q) plots comparing the observations at the two sites show they have similar measurements for temperature and wind speed yet large deviations in their measured wind directions (Supplementary Information Figure S1). This suggests that there may be some shielding of the site at the airport, leading to wind directions predominately from the North West and very different wind directions to those measured at Bishopton. The WRF models meet the IoA acceptance criteria proposed by Emery and Tai (2001) for all three meteorological variables tested, regardless of the setup (Table 2). This is not the case for the MB and RMSE value, which often do not meet the criteria. The use of IoA alone to assess model accuracy is not appropriate.

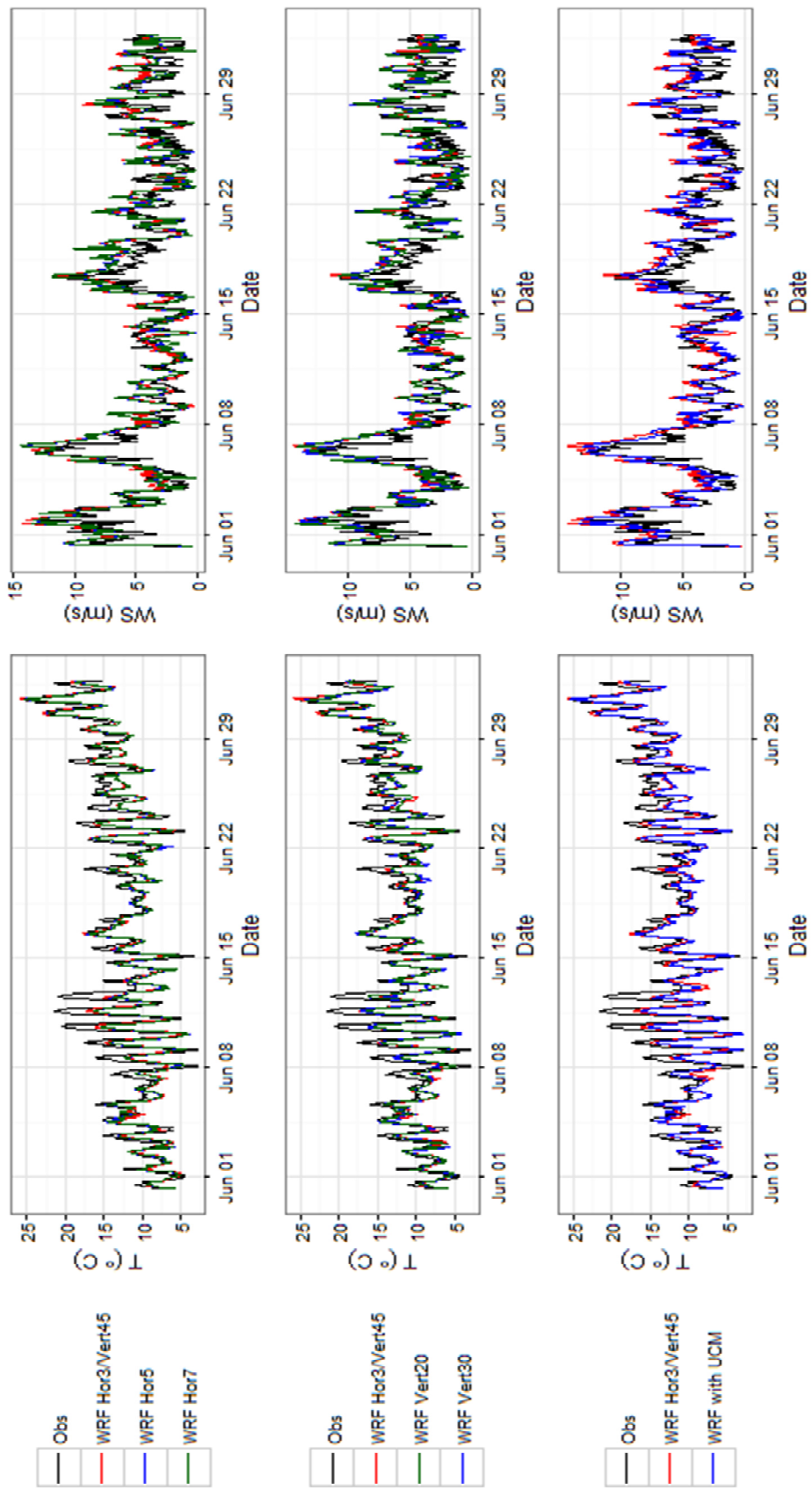


Figure 2: Time series of WRF predictions and observed data for temperature (left) and wind speed (right) at Glasgow Airport. Top: effect of horizontal resolution, middle: effect of vertical resolution and bottom: inclusion of UCM.

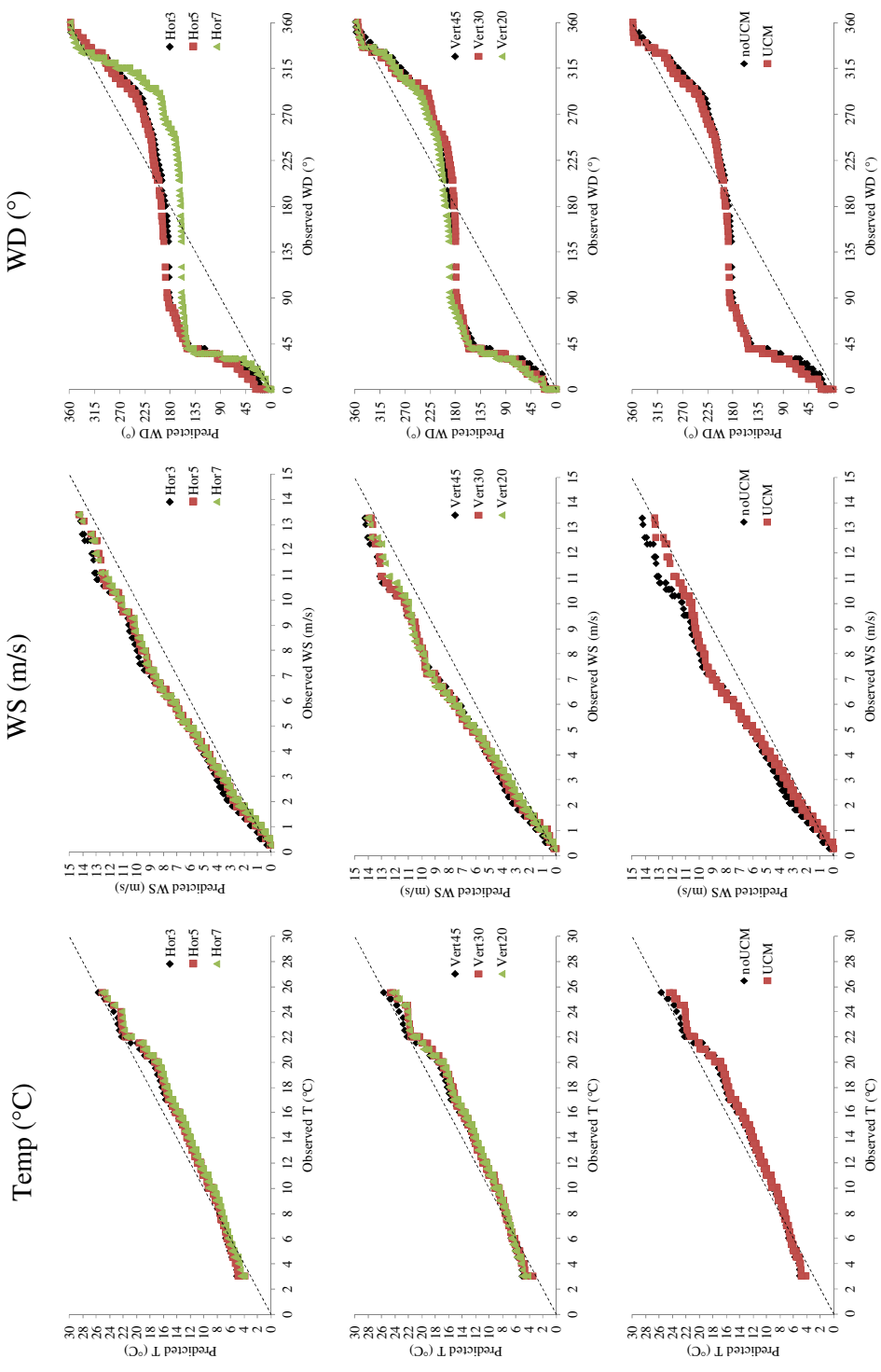


Figure 3: Q-Q plots of predicted and observed temperature, wind speed and wind direction at Glasgow Airport for each of the WRF model setups. The dashed line denotes 1:1 relationship between predicted and observed meteorological values.

Table 2: Evaluation statistics for each of the WRF model setups compared to observations of temperature (T), wind speed (WS) and wind direction (WD) at the two meteorological stations. The Mean Bias (MB), Root Mean Square Error (RMSE) and Index of Agreement (IoA) statistics are shown.

Variable	Site	Statistic	Hor3/Vert45	Hor5	Hor7	Vert30	Vert20	UCM	Acceptance Criteria[1]
T (°C)	Airport	MB	-1.22	-1.30	-1.40	-1.33	-1.40	-1.55	< 0.5
		RMSE	2.10	2.16	2.21	2.17	2.26	2.28	NA
		IoA	0.92	0.92	0.92	0.92	0.91	0.91	> 0.8
	Bishopton	MB	-1.23	-1.03	-1.00	-1.11	-1.12	-1.16	
		RMSE	2.19	2.15	2.10	2.20	2.13	2.14	
		IoA	0.91	0.92	0.92	0.91	0.92	0.92	
WS (m/s)	Airport	MB	1.08	0.85	0.85	0.97	0.89	0.63	< 0.5
		RMSE	1.90	1.84	1.90	1.89	1.92	1.61	< 2
		IoA	0.88	0.88	0.88	0.88	0.87	0.91	> 0.6
	Bishopton	MB	1.19	1.14	1.18	1.35	1.26	1.23	
		RMSE	2.15	2.09	2.14	2.29	2.23	2.11	
		IoA	0.81	0.81	0.81	0.80	0.80	0.82	
WD (°)	Airport	MB	-10.68	-5.24	-31.86	-29.55	-24.64	-28.45	< 10
		RMSE	119.39	117.75	129.71	131.31	135.67	134.02	NA
		IoA	0.52	0.53	0.46	0.42	0.36	0.41	NA
	Bishopton	MB	-1.49	2.35	-18.04	-26.33	-20.73	-21.15	
		RMSE	72.91	71.81	125.62	121.96	124.10	121.98	
		IoA	0.77	0.77	0.30	0.37	0.30	0.34	

[1.] *Emery and Tai (2001) acceptance criteria values*

3.1.1. Effect of Horizontal Resolution on WRF meteorological estimates

Increasing the horizontal resolution resulted in an increase in run time (approximately twofold and threefold increase when increasing the resolution of the finest grid from 3 km to 1.8 km and 3 km to 1.3 km respectively) and the size of the files produced by the WRF model (Table 3). The most resolved horizontal grid produced temperature estimates that were furthest from the observed values at Glasgow Airport; however the IoA values were uninfluenced (Figure 3, Table 2). At Bishopton, however, the MB and RMSE values were reduced when more resolved horizontal grids were used, however no discernible difference is observed in the Q-Q plots (Supplementary Information Figure S3 and Table 2). The land use data may be inaccurate at the finer scales resulting in lower model accuracy at higher resolutions which may be more pronounced for the Airport site, located in mixed urban-rural location, in comparison to Bishopton, located in a predominantly rural location (De Meij and Vinuesa, 2014; Zhang et al., 2011). Regardless of the site and horizontal resolution used, the temperature MB values do not meet the acceptance criteria of Emery and Tai (2001).

The 3 km and 1.3 km grids performed similarly for wind speed, while the 1.8 km had the lowest error values (Figures 3 & S3 and Table 2). No change in IoA values was observed with changing the horizontal resolution (Table 2). The RMSE acceptance criterion was met at the airport site however neither site met the MB criteria.

The Q-Q plots showed little change in the airport wind direction when the resolution was increased from 3 km to 1.8 km, however using a horizontal grid of 1.3 km showed the largest deviations between the modelled and observed wind directions (Figures 3 and S3). As a result, the error values improve from 3 km to 1.8 km, and are poorest for 1.3 km (Table 2). The noticeably poorer accuracy for wind directions for the most highly resolved grid has been reported previously in the literature by Wu et al. (2008).

Changing the horizontal resolution from coarse grids to intermediate grids (e.g. 36 km to 12 km) has been shown to improve model accuracy for temperature, wind speed and wind direction, however further increasing the grid resolution to e.g. 4 km generally has much less of an impact (Tartakovsky et al., 2015; URS Corporation, 2008; Wu et al., 2008). Higher horizontal grid resolutions are required if weather models are to be used for e.g. complex terrain or urban forecasting (Baklanov et al., 2002). Saide et al. (2011) found no influence of increasing the resolution beyond 4 km while Gibbs et al. (2011) found poorer model performance with increased resolution from 4 k to 1 km.

Our results suggest that there may be a small improvement in the predictive power of the WRF model when the horizontal resolution is increased from 3 km to 1.8 km, however further increasing the resolution to 1.3 km had little or a negative effect on model accuracy.

This suggests that, at these detailed scales, the horizontal resolution has little effect on model accuracy and supports published findings that the associated increases in model run times and cost with higher resolution are not necessary (Kain et al., 2008; Schwartz et al., 2009).

Table 3: Approximate size and model run times per week (note time dependent on other processes running and on computer specifications) for WRF models.

	Hor3	**Hor5**	**Hor7**	**Vert30**	**Vert20**	**UCM**
Approximate run time (per week)	3 hours	7 hours	10 hours	2 hours	1 hours	5 hours
File size final model (per week)	6.9 Gb	7.2 Gb	7.4 Gb	6.4 Gb	6.0 Gb	8.0 Gb

3.1.2. Effect of Vertical Resolution on WRF meteorological estimates

Changing the number of vertical levels had little impact on the Q-Q plots however reducing the number of vertical levels lead to an increase in MB and RMSE and a reduction in IoA at the airport (Figure 3 and Table 2). Fewer vertical levels may not be able to accurately reproduce the conditions in the boundary layer resulting in poorer model accuracy. However, at Bishopton the reduction in vertical levels to 20 improved the bias and IoA values (Table 2). The fewer vertical levels used in this instance may average out the effect of interactions with land use data meaning if the land use is inaccurate fewer vertical levels could perform better.

Using fewer vertical levels lead to a reduction in MB, but an increase in RMSE, values at the airport (Table 2, Figure 3). At Bishopton the opposite was true, where fewer vertical levels resulted in increased bias values (Table 2). The RMSE values at the airport meet the Emery and Tai acceptance criteria, but for the Bishopton site the bias criteria are not met.

The greatest similarity between the measured and modelled wind roses observed when 45 vertical levels were used (Figures 3 & S3). There were deviations between the WRF predicted and observation wind rose plots when the number of vertical levels was reduced (Figures 4 & S4). The RMSE values for both sites increased and IoA decreased when fewer vertical levels were used (Table 2). The MB acceptance criterion is only met for wind direction estimates at Bishopton using 45 vertical levels.

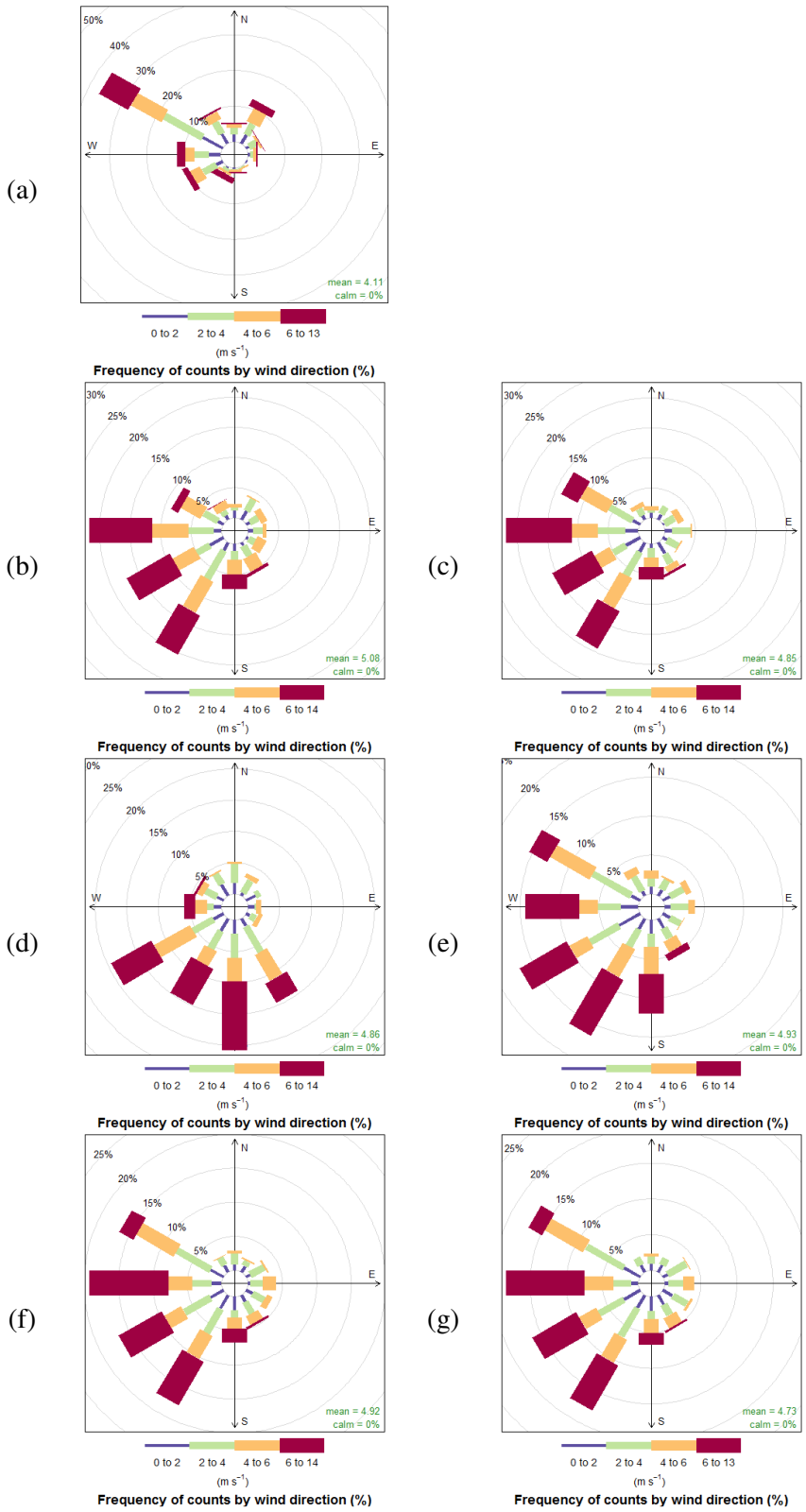


Figure 4: Wind rose plots for the observation at the meteorological stations and the predictions from the different setups of the WRF model. Roses are shown for Glasgow Airport for (a) Observation, (b) Hor3, (c) Hor5, (d) Hor7, (e) Vert30, (f) Vert20 and (g) UCM.

A similar study by Saide et al. (2011) focusing on temperature and wind speed measurements found little difference between WRF model results run with 44 and 39 levels and consequently the lower resolution was used in the remainder of their study. Similar findings were published by Floors et al. (2013) who found no improvement in agreement between the vertical wind speed profile predicted by WRF and measured by LIDAR measurements when the number of vertical levels was increased from 41 to 63.

Our study found reducing the number of vertical levels used in the model resulted in higher bias values and lower IoA values. This suggests that reducing the number of vertical levels in order to reduce run times and minimise output files (Table 3) is likely to compromise the accuracy of the model. It is possible the number of vertical levels could be reduced below 45 if the optimal heights of the layers were determined; however, the influence of vertical level height was not examined in this work.

3.1.3. Effect of UCM on WRF meteorological predictions

Adding the UCM to the WRF model resulted in an increase in the model run time from ~3 to 5 hours and increased the size of the model by approximately 1 Gb (Table 3). The Q-Q plots for these relationships show little visual difference when the UCM was included (Figures 3 & S3). At the Airport, including the UCM increased the temperature bias values and reduces the IoA values, while the opposite was true for the Bishopton site (Table 2). The difference in the results between the two sites was attributed to different site types - Lee et al. (2011) compared results from a WRF model run with and without a UCM and, for commercial/industrial sites, found the MB values to decrease but the RMSE to increase during the day. However, for rural sites adding the UCM resulted in a reduction in both MB and RMSE.

At the airport, adding the UCM reduced the wind speeds predicted by the model, bringing the Q-Q plot closer to 1:1 and reducing the bias values (Figure 3 and Table 2). At Bishopton, however, the Q-Q plot showed little improvement when the UCM was added however a slight increase in MB values was observed (Figure S3, Table 2). The lowering of wind speeds has been reported by Kim et al. (2015) for urban sites in Paris and Holt and Pullen (2007) in New York when a UCM was included in the WRF model. The Q-Q and wind rose plots including the UCM were poorer than those for the WRF model without the UCM (Figures 3 & 4 and Figures S3 & S4), and the UCM model had larger MB and RMSE values and lower IoA values (Table 2).

The inclusion of the UCM model generally lowered the accuracy of the WRF model at these sites. The literature illustrates the improvement in model performance when the UCM

is added, however most of this work has focussed on urban locations in large cities. Therefore, the UCM should be run only in urban areas as this work has shown the impact of the UCM to be negative for the rural or semi-rural location of the sites tested.

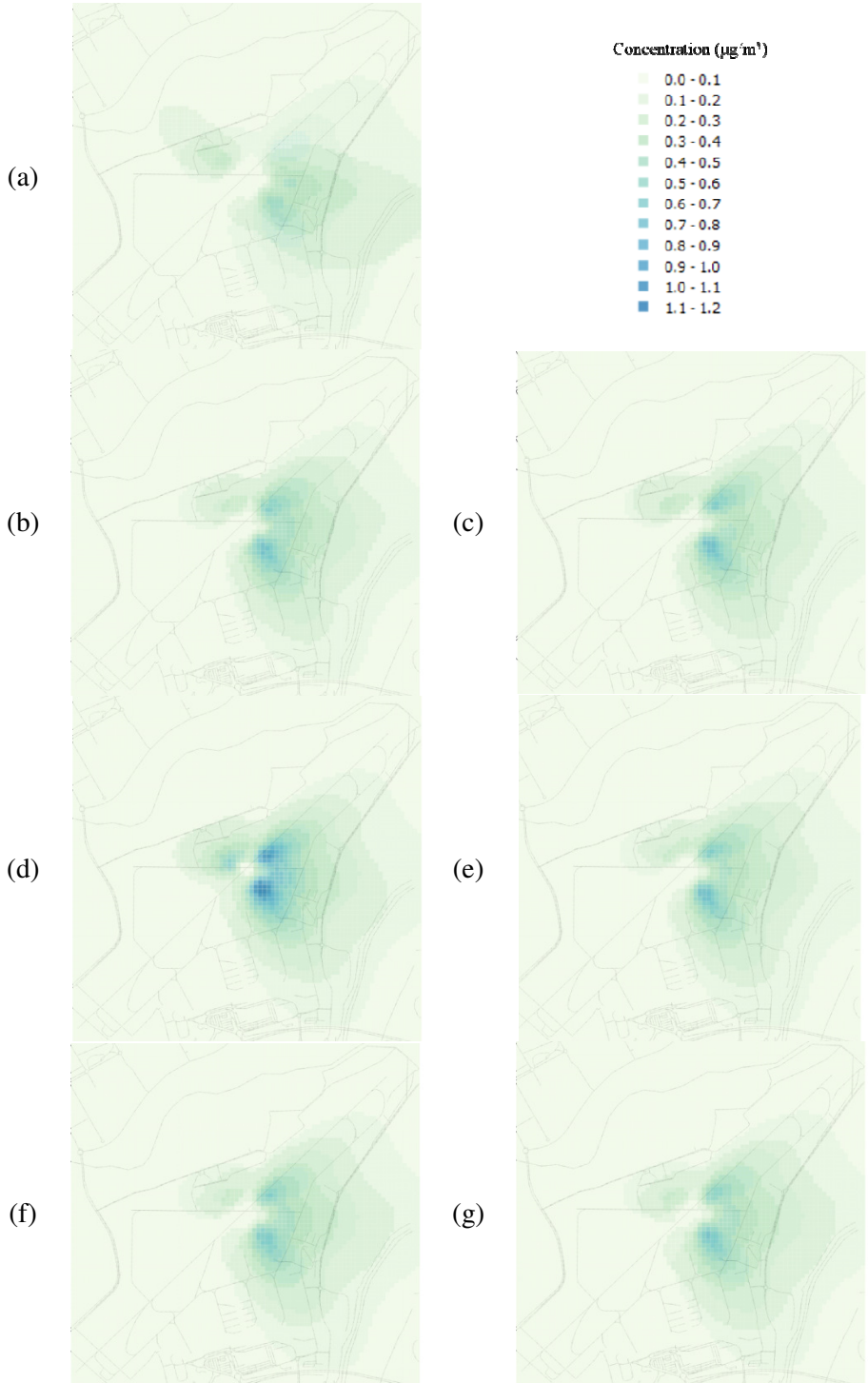


Figure 5: Monthly average dispersion plume concentrations (µg/m^3) for a 50 m stack, located in the centre of the study area, emitting at 1 g/s predicted by: (a) Observed-AERMOD; (b) WRF-AERMOD (Hor3); (c) WRF-AERMOD (Hor5); (d) WRF-AERMOD (Hor7); (e) WRF-AERMOD (Vert30); WRF-AERMOD (Vert20); and WRF-AERMOD (UCM). The road network in the study area is shown in grey.

3.2. Dispersion modelling using WRF and meteorological inputs

The dispersion plumes estimated using WRF-AERMOD were more dispersed with a higher concentration of pollution closer to the source, regardless of WRF set up, than those predicted using the Observed-AERMOD (Figures 5, S5 & S6). We showed above that the estimates of wind speed using WRF over predicted the observation data, which could lead to the dispersion plume bending over and depositing pollutants closer to the stack. The wind directions estimated using WRF showed deviations from the observation data which could explain the larger plume widths in WRF-AERMOD compared to Observed-AERMOD. The underestimation of temperature by WRF may influence the AERMOD model, as lower temperatures lead to lower Monin-Obukhov lengths which in turn increase the surface friction velocity and convective scale velocity, however this effect is expected to be negligible (Kesarkar et al., 2007). Table 4 provides summary statistics for the concentrations predicted by each AERMOD model and model evaluation statistics.

Table 4: Model statistics for the monthly average pollution concentrations predicted at the 6561 receptor sites from AERMOD for each of (a) 10 m stack, (b) 50 m stack and (c) 100 m stack heights. The percentiles, max and min concentrations in the model are provided along with MB, RMSE, and IoA for AERMOD estimates using WRF input and observation meteorological input (the latter is assumed to represent the true concentrations).

Statistic	Obs-AMOD	Hor3-AMOD	Hor5-AMOD	Hor7-AMOD	Vert30-AMOD	Vert20-AMOD	UCM-AMOD
(a) 10 m stack							
25th percentile	0.03	0.02	0.02	0.02	0.02	0.03	0.03
50th percentile	0.06	0.05	0.05	0.05	0.05	0.06	0.06
75th percentile	0.15	0.10	0.11	0.14	0.11	0.12	0.11
Min	0	0	0	0	0	0	0
Max	6.77	6.97	6.95	12.14	7.12	6.58	6.85
MB	-	-0.03	-0.03	-0.02	-0.03	-0.03	-0.03
RMSE	-	0.21	0.22	0.29	0.22	0.20	0.20
IoA	-	0.91	0.90	0.88	0.90	0.91	0.91
(b) 50 m stack							
25th percentile	0.02	0.02	0.02	0.02	0.02	0.02	0.02
50th percentile	0.04	0.04	0.04	0.04	0.04	0.04	0.04
75th percentile	0.07	0.08	0.08	0.09	0.08	0.08	0.08
Min	0	0	0	0	0	0	0
Max	0.46	0.75	0.76	1.20	0.76	0.71	0.74
MB	-	0.01	0.01	0.03	0.01	0.01	0.01
RMSE	-	0.06	0.06	0.10	0.06	0.06	0.06
IoA	-	0.85	0.85	0.73	0.84	0.86	0.86
(c) 100 m stack							
25th percentile	0.01	0.01	0.01	0.02	0.01	0.01	0.01
50th percentile	0.02	0.03	0.03	0.04	0.02	0.03	0.03
75th percentile	0.04	0.06	0.06	0.08	0.06	0.06	0.06
Min	0	0	0	0	0	0	0
Max	0.13	0.22	0.23	0.39	0.23	0.21	0.23
MB	-	0.01	0.01	0.03	0.01	0.02	0.02
RMSE	-	0.03	0.03	0.06	0.03	0.03	0.03
IoA	-	0.79	0.78	0.59	0.79	0.78	0.80

3.2.1. Effect of horizontal resolution on WRF-AERMOD pollution estimates

The size and shape of the plume predicted by the WRF-AERMOD using different horizontal grid resolutions was similar for the 10 m and 50 m stack (Figure 5 & S5). Some deviations in plume shape with changing WRF horizontal resolution are apparent for the 100 m stack, with the finest horizontal grid producing different predictions than the other models (Figure S6). The plume pattern predicted using Observed-AERMOD showed a longer but generally narrower plume than the WRF-AERMOD estimates (Figures 5, S5 & S6). The average concentrations predicted at each site using WRF-AERMOD (Hor3 and Hor5 horizontal resolution) were close similar to concentrations predicted using Observed-AERMOD for the 10 m stack height, while for the finest horizontal grid WRF-AERMOD over predicted the concentrations from Observed-AERMOD above concentrations of ~1.5

$\mu g/m^3$ (Figure 6a). For stack heights of 50 m and 100 m the WRF-AERMOD model over predicted the concentrations estimated by Observed-AERMOD (Figures 6b & 6c).

Increasing the stack height resulted in larger differences between concentrations estimated using WRF-AERMOD and Observed-AERMOD, leading to larger RMSE and lower IoA values (Figure 6 and Table 4). Irrespective of the stack height modelled, the coarsest horizontal grid WRF-AERMOD model was most similar to the Observed-AERMOD concentrations. This suggests that, similarly to the comparison between WRF and observed meteorological data, the associated increase in run time and file size for increased horizontal resolution is not required as it does not necessarily improve model accuracy. Saide et al. (2011) found some improvement in predicted WRF-CHEM carbon monoxide concentrations increasing from 6 km to 2 km grid resolution, but moving to 667 m had little impact. Similar tests using the MM5 weather model as input for the CMAQ pollution model found the effect of horizontal resolution to be pollutant dependent (Wu et al., 2008). Ozone predictions were not influenced by horizontal resolution in the winter, however in the summer months the larger grid sizes produced pollution estimates closer to observations. However, for $PM_{2.5}$ the finer model produced the best agreement with the observations. Another study using similar scale grids to this work found no influence on ozone model accuracy with changing resolution (Shrestha et al., 2009).

This work shows that for small stacks, e.g. 10 m, WRF modelled weather data can be used instead of observation data as these produce similar concentrations data. However, for the larger stacks tested (50 m and 100 m) the pollution concentrations estimated using WRF as input weather data were greater than the concentrations predicted using the observations meteorological data (approximately 1.5 times larger for the 100 m stack) suggesting that observation data should be used in preference to WRF for taller stacks. If observation data is not available, using WRF will provide an over-estimate of the modelled concentrations with meteorological data.

3.2.2. Effect of vertical resolution on WRF-AERMOD pollution estimates

Changing the WRF vertical resolution had little effect on the final plume shape and concentrations predicted by WRF-AERMOD (Figures 5, S5 and S6). There was also limited effect of changing the number of vertical levels on the Q-Q plots between WRF-AERMOD and Observed-AERMOD concentrations (Figure 6). Table 4 also shows very little change in MB, RMSE or IoA values when the number of vertical levels used in the WRF set up is changed for all three stack heights tested. This suggests that it may be possible to reduce the number of vertical levels from 45 to 20, saving run time, without compromising dispersion

model accuracy. As discussed previously, the height of the vertical levels may be important and should be investigated further to determine if there is an optimal height for the levels in order to get improved agreement between the modelled and measured concentrations for the taller stacks. Saide et al. (2011) found reducing the number of levels from 44 to 39 had little effect on carbon monoxide concentrations however a marked jump in performance was found when 28 levels were used. Teixeira et al. (2015) found that increasing the number of vertical levels from 30 to 60 resulted in an improvement in the ability of the model to predict dust storms, however if more levels were used no further improvement or some deterioration in model accuracy was found.

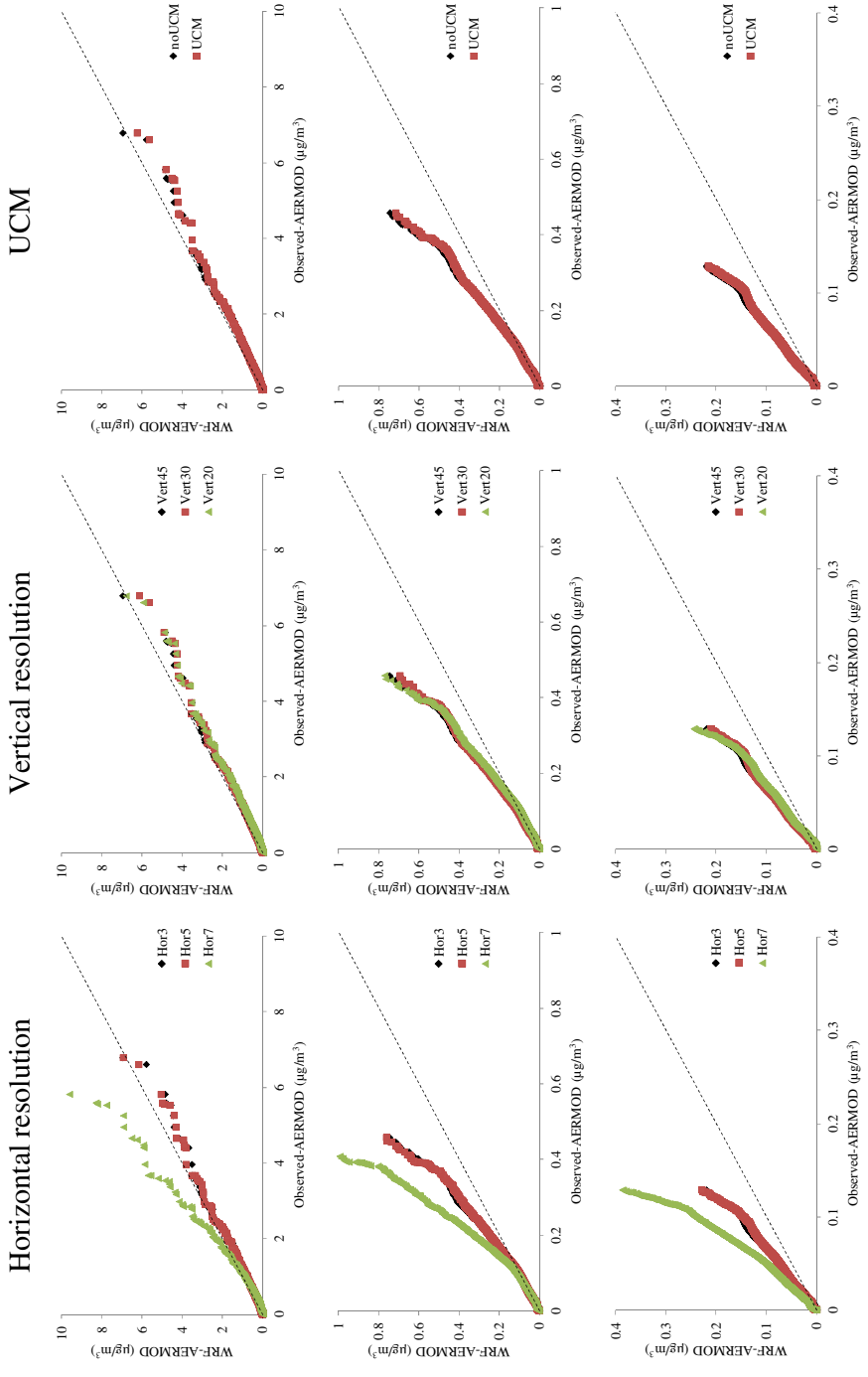


Figure 6: Q-Q plots showing correlation between monthly pollution estimates using Observed-AERMOD and WRF-AERMOD (for varying horizontal and vertical resolution, and inclusion of the UCM in WRF model) for each receptor location (n = 6561). The plots for each stack height are shown: (a) 10 m: (b) 50 m; and (c) 100 m. Black solid line represents 1:1 relationship between Observed-AERMOD and WRF-AERMOD concentrations.

3.2.3. Effect of UCM on WRF-AERMOD pollution estimates

Adding the UCM had little impact on the size or shape of the plume predicted by WRF-AERMOD (Figures 5, S5 & S6) nor was there a visible change in the Q-Q plot comparing pollution estimates between WRF-AERMOD and Observed-AERMOD (Figure 6). Adding the UCM resulted in no change in the mean bias values and slight improvements in RMSE and IoA (improved by 0.01) compared to when the WRF-AERMOD model was run without the UCM (Table 4). The small improvement gained when the UCM is added does not outweigh the additional run time and information required for the UCM (Table 3). De Meij et al. (2015) tested WRF-CHEM using more detailed land use data in the UCM and found that the more resolved land use data gave better model agreement with PM_{10} concentrations, and similar model accuracy for carbon monoxide however these results were found to vary between areas and for different time periods. The lack of improvement when the UCM is added in this work may be due to the rural location and flat terrain of the Airport site.

4. Conclusions

We investigated the effect of initial WRF set up on the predictions of meteorological and pollution dispersion modelling. We compared WRF estimates of temperature, wind speed and wind direction to observations at two meteorological stations in Glasgow, UK. The WRF model generally underestimated temperature, overestimated wind speed and had a pronounced pattern of discrepancy with wind direction measurements at the two meteorological stations. Increasing the horizontal resolution from 3 x 3 km to 1.8 x 1.8 km and 1.3 x 1.3 km had little effect on temperature and wind speed predictions. However, the wind direction predictions from the 1.3 x 1.3 km horizontal resolution model were more discrepant with observations than equivalent predictions from the coarser resolution horizontal WRF models. The apparent loss of accuracy when a finer-scale horizontal resolution was used illustrated that the associated increased in model run time and file size provided no additional benefit for model predictions. Reducing the number of vertical levels used in the WRF simulation had little effect on the WRF temperature and wind speed estimates, but larger deviations between observed and modelled wind direction were observed when fewer vertical levels were used. Reducing the number of vertical levels from the default 45 to 30 approximately reduced the run time by 30 % and reduced the model output file size by 500 Mb. Adding the UCM detrimentally impacted the ability of the model to predict wind direction (Index of Agreement reduced from 0.77 to 0.34 with inclusion of the UCM) at the weather stations which were located in a rural environment. This suggests that if WRF if to be run in rural areas the UCM should not be included.

The plume shapes predicted when WRF estimates were used as meteorological input to AERMOD (WRF-AERMOD) were wider and showed higher concentrations closer to the source than equivalent AERMOD concentration predictions when observed meteorological data was used as input (Observed-AERMOD). Increasing the horizontal resolution of the WRF model used for WRF-AERMOD input did not improve the correlation between WRF-AERMOD and Observed-AERMOD estimated concentrations (Table 4). The closest pollution estimates between the WRF-AERMOD and Observed-AERMOD concentrations was for the coarsest WRF grid for the stack heights investigated. The pollution estimates when the number of vertical levels in the WRF model was changed from 45 to 30 to 20 showed negligible changes suggesting that if WRF is to be used for pollution modelling then number of vertical levels can be reduced to 20 without detrimental effects on predictions of passive tracer concentrations. Similarly, adding the UCM to WRF had negligible effect on the WRF-AERMOD pollution estimates compared to the WRF-AERMOD model run without the UCM. Therefore, the associated time and file size required to run the UCM does not appear necessary for pollution modelling in rural or semi-urban locations. WRF can be used instead of observed meteorological data as input to AERMOD to predict concentrations from small stack heights (10 m). In contrast, the WRF-AERMOD concentrations predicted for stack heights of 50 m and 100 m were greater than Observed-AERMOD concentrations by a factor of approximately 1.5.

This study highlights that WRF can be used as input to dispersion modelling and found that changing the WRF setup had only small effects on dispersion model predictions for 10 m stack height. In the absence of meteorological observations WRF could act as an appropriate surrogate but the concentrations predicted for taller heights may be overestimated.

Funding Sources

Nicola Masey is funded through a National Environmental Research Council CASE PhD studentship (NE/K007319/1) with industrial support from Ricardo Energy and Environment.

References

Appel, K.W., Roselle, S.J., Gilliam, R.C., Pleim, J.E., 2010. Sensitivity of the Community Multiscale Air Quality (CMAQ) model v4.7 results for the eastern United States to MM5 and WRF meteorological drivers. Geosci. Model Dev. 3, 169–188.

Baker, K.R., Misenis, C., Obland, M.D., Ferrare, R.A., Scarino, A.J., Kelly, J.T., 2013. Evaluation of surface and upper air fine scale WRF meteorological modeling of the May and

June 2010 CalNex period in California. Atmos. Environ. 80, 299–309. doi:10.1016/j.atmosenv.2013.08.006

Baklanov, A., Rasmussen, A., Fay, B., Berge, E., Finardi, S., 2002. Potential and Shortcomings of Numerical Weather Prediction Models in Providing Meteorological Data for Urban Air Pollution Forecasting. Water Air Soil Pollut. Focus 2, 43–60. doi:10.1023/A:1021394126149

Batterman, S., Ganguly, R., Isakov, V., Burke, J., Arunachalam, S., Snyder, M., Robins, T., Lewis, T., 2014. Dispersion Modeling of Traffic-Related Air Pollutant Exposures and Health Effects Among Children with Asthma in Detroit, Michigan. Transp. Res. Rec. 2452, 105–112. doi:10.3141/2452-13

Borge, R., Alexandrov, V., José del Vas, J., Lumbreras, J., Rodríguez, E., 2008. A comprehensive sensitivity analysis of the WRF model for air quality applications over the Iberian Peninsula. Atmos. Environ. 42, 8560–8574. doi:10.1016/j.atmosenv.2008.08.032

Carbonell, L.T., Mastrapa, G.C., Rodriguez, Y.F., Escudero, L.A., Gacita, M.S., Morlot, A.B., Montejo, I.B., Ruiz, E.M., Rivas, S.P., 2013. Assessment of the Weather Research and Forecasting model implementation in Cuba addressed to diagnostic air quality modeling. Atmospheric Pollut. Res. 4, 64–74. doi:10.5094/APR.2013.007

Carslaw, D., 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

Chen, F., Dudhia, J., 2001. Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. Mon. Weather Rev. 129, 569–585. doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2

de Hoogh, K., Korek, M., Vienneau, D., Keuken, M., Kukkonen, J., Nieuwenhuijsen, M.J., Badaloni, C., Beelen, R., Bolignano, A., Cesaroni, G., Pradas, M.C., Cyrys, J., Douros, J., Eeftens, M., Forastiere, F., Forsberg, B., Fuks, K., Gehring, U., Gryparis, A., Gulliver, J., Hansell, A.L., Hoffmann, B., Johansson, C., Jonkers, S., Kangas, L., Katsouyanni, K., Künzli, N., Lanki, T., Memmesheimer, M., Moussiopoulos, N., Modig, L., Pershagen, G., Probst-Hensch, N., Schindler, C., Schikowski, T., Sugiri, D., Teixidó, O., Tsai, M.-Y., Yli-Tuomi, T., Brunekreef, B., Hoek, G., Bellander, T., 2014. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. Environ. Int. 73, 382–392. doi:10.1016/j.envint.2014.08.011

De Meij, A., Bossioli, E., Penard, C., Vinuesa, J.F., Price, I., 2015. The effect of SRTM and Corine Land Cover data on calculated gas and PM10 concentrations in WRF-Chem. Atmos. Environ. 101, 177–193. doi:10.1016/j.atmosenv.2014.11.033

De Meij, A., Vinuesa, J.F., 2014. Impact of SRTM and Corine Land Cover data on meteorological parameters using WRF. Atmospheric Res. 143, 351–370. doi:10.1016/j.atmosres.2014.03.004

Dudhia, J., 1988. Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. J. Atmospheric Sci. 46, 3077–3107. doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2

Emery, C., Tai, E., Yarwood, G., 2001. Enhanced Meteorological Modeling And Performance Evaluation For Two Texas Ozone Episodes. Prepared for The Texas Natural Resource Conservation Commission.

Floors, R., Vincent, C.L., Gryning, S.-E., Peña, A., Batchvarova, E., 2013. The Wind Profile in the Coastal Boundary Layer: Wind Lidar Measurements and Numerical Modelling. Bound.-Layer Meteorol. 147, 469–491. doi:10.1007/s10546-012-9791-9

Gibbs, J.A., Fedorovich, E., van Eijk, A.M.J., 2011. Evaluating Weather Research and Forecasting (WRF) Model Predictions of Turbulent Flow Parameters in a Dry Convective Boundary Layer. J. Appl. Meteorol. Climatol. 50, 2429–2444. doi:10.1175/2011JAMC2661.1

Grossman-Clarke, S., Zehnder, J., Stefanov, W., Liu, Y., Zoldak, M., 2005. Urban Modifications in a Mesoscale Meteorological Model and the Effects on Near-Surface Variables in an Arid Metropolitan Region. J. Appl. Meteorol. 44, 1281–1297. doi:10.1175/JAM2286.1

Holt, T., Pullen, J., 2007. Urban canopy modeling of the New York City metropolitan area: A comparison and validation of single- and multilayer parameterizations. Mon. Weather Rev. 135, 1906–1930. doi:10.1175/MWR3372.1

Janjic, Z., 1994. The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. Mon. Weather Rev. 122, 927–945. doi:10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2

Kain, J.S., 2004. The Kain–Fritsch Convective Parameterization: An Update. J. Appl. Meteorol. 43, 170–181. doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2

Kain, J.S., Weiss, S.J., Bright, D.R., Baldwin, M.E., Levit, J.J., Carbin, G.W., Schwartz, C.S., Weisman, M.L., Droegemeier, K.K., Weber, D.B., Thomas, K.W., 2008. Some Practical Considerations Regarding Horizontal Resolution in the First Generation of Operational Convection-Allowing NWP. Weather Forecast. 23, 931–952. doi:10.1175/WAF2007106.1

Kesarkar, A.P., Dalvi, M., Kaginalkar, A., Ojha, A., 2007. Coupling of the Weather Research and Forecasting Model with AERMOD for pollutant dispersion modeling. A case

study for PM10 dispersion over Pune, India. Atmos. Environ. 41, 1976–1988. doi:10.1016/j.atmosenv.2006.10.042

Kim, Y., Sartelet, K., Raut, J.-C., Chazette, P., 2015. Influence of an urban canopy model and PBL schemes on vertical mixing for air quality modeling over Greater Paris. Atmos. Environ. 107, 289–306. doi:10.1016/j.atmosenv.2015.02.011

Lee, S.-H., Kim, S.-W., Angevine, W.M., Bianco, L., McKeen, S.A., Senff, C.J., Trainer, M., Tucker, S.C., Zamora, R.J., 2011. Evaluation of urban surface parameterizations in the WRF model using measurements during the Texas Air Quality Study 2006 field campaign. Atmospheric Chem. Phys. 11, 2127–2143. doi:10.5194/acp-11-2127-2011

Lin, Y., Farley, R., Orville, H., 1983. Bulk Parameterization of the Snow Field in a Cloud Model. J. Clim. Appl. Meteorol. 22, 1065–1092. doi:10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2

Madala, S., Satyanarayana, A.N.V., Srinivas, C.V., Kumar, M., 2015. Mesoscale atmospheric flow-field simulations for air quality modeling over complex terrain region of Ranchi in eastern India using WRF. Atmos. Environ. 107, 315–328. doi:10.1016/j.atmosenv.2015.02.059

Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J., Clough, S.A., 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res. Atmospheres 102, 16663–16682. doi:10.1029/97JD00237

Mohan, M., Sati, A.P., 2016. WRF model performance analysis for a suite of simulation design. Atmospheric Res. 169, 280–291.

Qiu, X., Yang, F., Corbett-Hains, H., Roth, M., 2015. ASHRAE Research Project 1561-RP Procedure to Adjust Observed Climatic Data for Regional or Mesoscale Climatic Variations. Novus Environmental and Klimaat Consulting & Innovation.

Ritter, M., Mueller, M.D., Tsai, M.-Y., Parlow, E., 2013. Air pollution modeling over very complex terrain: An evaluation of WRF-Chem over Switzerland for two 1-year periods. Atmospheric Res. 132, 209–222. doi:10.1016/j.atmosres.2013.05.021

Saha, S., Moorthi, S., Pan, H., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Ganyo, G., Wang, J., Hou, Y., Chuang, H., Hann-Ming, H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Van Del Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R., Rutledge, G., Goldberg, M., 2010. The NCEP Climate Forecast System Reanalysis. Bull. Am. Meteorol. Soc. 91, 1015–1057. doi:10.1175/2010BAMS3001.1

Saide, P.E., Carmichael, G.R., Spak, S.N., Gallardo, L., Osses, A.E., Mena-Carrasco, M.A., Pagowski, M., 2011. Forecasting urban PM10 and PM2.5 pollution episodes in very stable nocturnal conditions and complex terrain using WRF–Chem CO tracer model. Atmos. Environ. 45, 2769–2780. doi:10.1016/j.atmosenv.2011.02.001

Schwartz, C.S., Kain, J.S., Weiss, S.J., Xue, M., Bright, D.R., Kong, F., Thomas, K.W., Levit, J.J., Coniglio, M.C., 2009. Next-Day Convection-Allowing WRF Model Guidance: A Second Look at 2-km versus 4-km Grid Spacing. Mon. Weather Rev. 137, 3351–3372. doi:10.1175/2009MWR2924.1

Shrestha, K.L., Kondo, A., Kaga, A., Inoue, Y., 2009. High-resolution modeling and evaluation of ozone air quality of Osaka using MM5-CMAQ system. J. Environ. Sci. 21, 782–789. doi:10.1016/S1001-0742(08)62341-4

Skamarock, W., Klemp, J., Dudhia, J., O'Gill, D., Barker, D., Duda, M., Huang, X., Wang, W., Powers, J., 2008. A description of the Advanced Research WRF Version 3 (Technical Report No. NCAR/TN-475STR). Mesoscal and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado, USA.

Tartakovsky, D., Stern, E., Broday, D.M., 2015. Evaluation of modeled wind field for dispersion modeling. Atmospheric Res. 166, 150–156. doi:10.1016/j.atmosres.2015.07.004

Teixeira, J.C., Carvalho, A.C., Tuccella, P., Curci, G., Rocha, A., 2015. WRF-chem sensitivity to vertical resolution during a saharan dust event. Phys. Chem. Earth Parts ABC. doi:10.1016/j.pce.2015.04.002

Tiwari, S., Srivastava, A.K., Bisht, D.S., Parmita, P., Srivastava, M.K., Attri, S.D., 2013. Diurnal and seasonal variations of black carbon and PM2.5 over New Delhi, India: Influence of meteorology. Atmospheric Res. 125–126, 50–62. doi:10.1016/j.atmosres.2013.01.011

Unidata, UCAR, 2015. Integrated Data Viewer (IDV) 5.2. doi:http://doi.org/10.5065/D6RN35XM

URS Corporation, 2008. MM5 Performance Evaluation and Quality Assurance Review for the White River Field Office Ozone Assessment. Denver, Colorado, USA.

U.S. Environmental Protection Agency, 2015. Guidance on the use of the Mesoscale Model Interface Program (MMIF) for AERMOD applications (No. EPA-454/B-15-001). U.S. Environmental Protection Agency, North Carolina, US.

USEPA, 2015. AERMOD.

Wu, S.-Y., Krishnan, S., Zhang, Y., Aneja, V., 2008. Modeling atmospheric transport and fate of ammonia in North Carolina—Part I: Evaluation of meteorological and chemical predictions. Atmos. Environ., Agricultural Air Quality: State of the Science (AAQ-2006) 42, 3419–3436. doi:10.1016/j.atmosenv.2007.04.031

Zhang, Y., Cheng, S.-H., Chen, Y.-S., Wang, W.-X., 2011. Application of MM5 in China: Model evaluation, seasonal variations, and sensitivity to horizontal grid resolutions. Atmos. Environ. 45, 3454–3465. doi:10.1016/j.atmosenv.2011.03.019

Supplementary information

Influence of WRF model parameterization on predictions from an air pollution dispersion model

Nicola Masey[1], Iain J. Beverland[1], Scott Hamilton[2]

[1]Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2]Ricardo Energy and Environment, 18 Blythswood Square, Glasgow, G2 4BG, UK

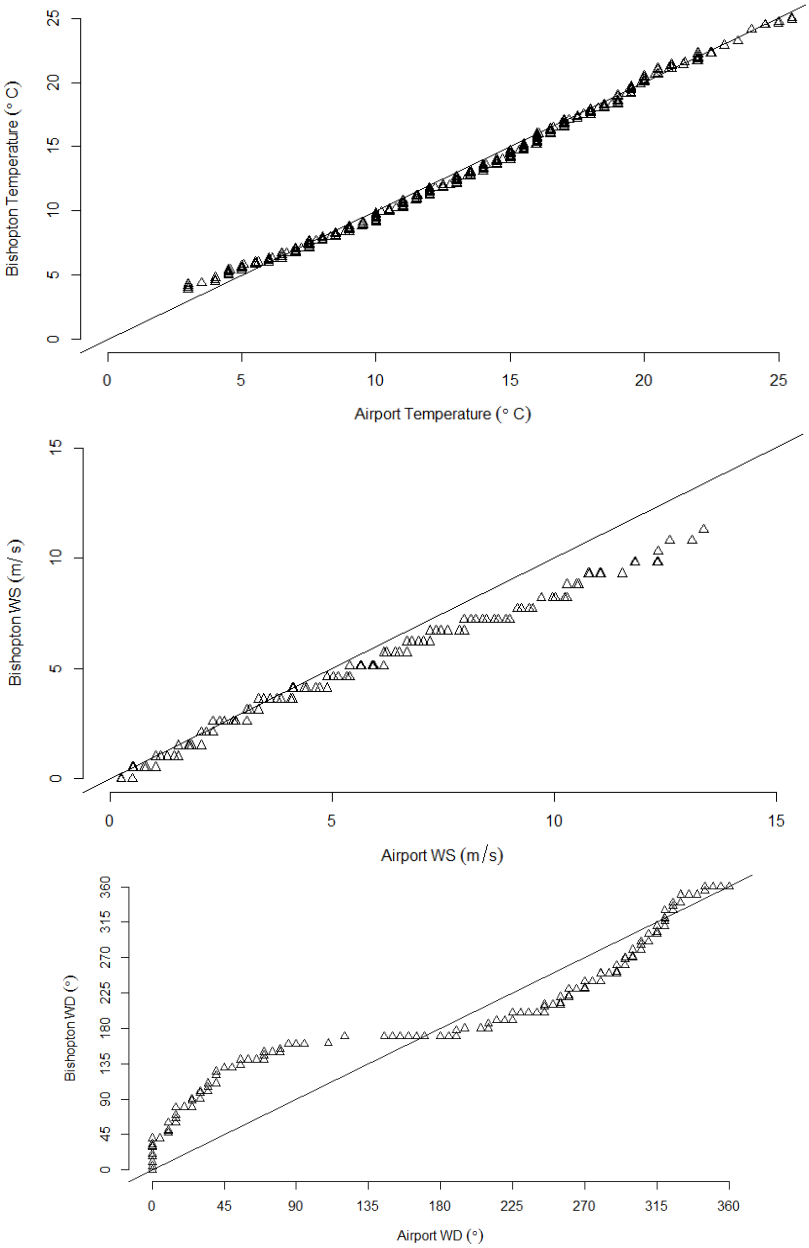


Figure S1: Q-Q plots showing the temperature (top), wind speed (middle) and wind direction (bottom) measured at the Airport and Bishopton site. The solid black line represents the 1:1 relationship between Airport and Bishopton measurements.

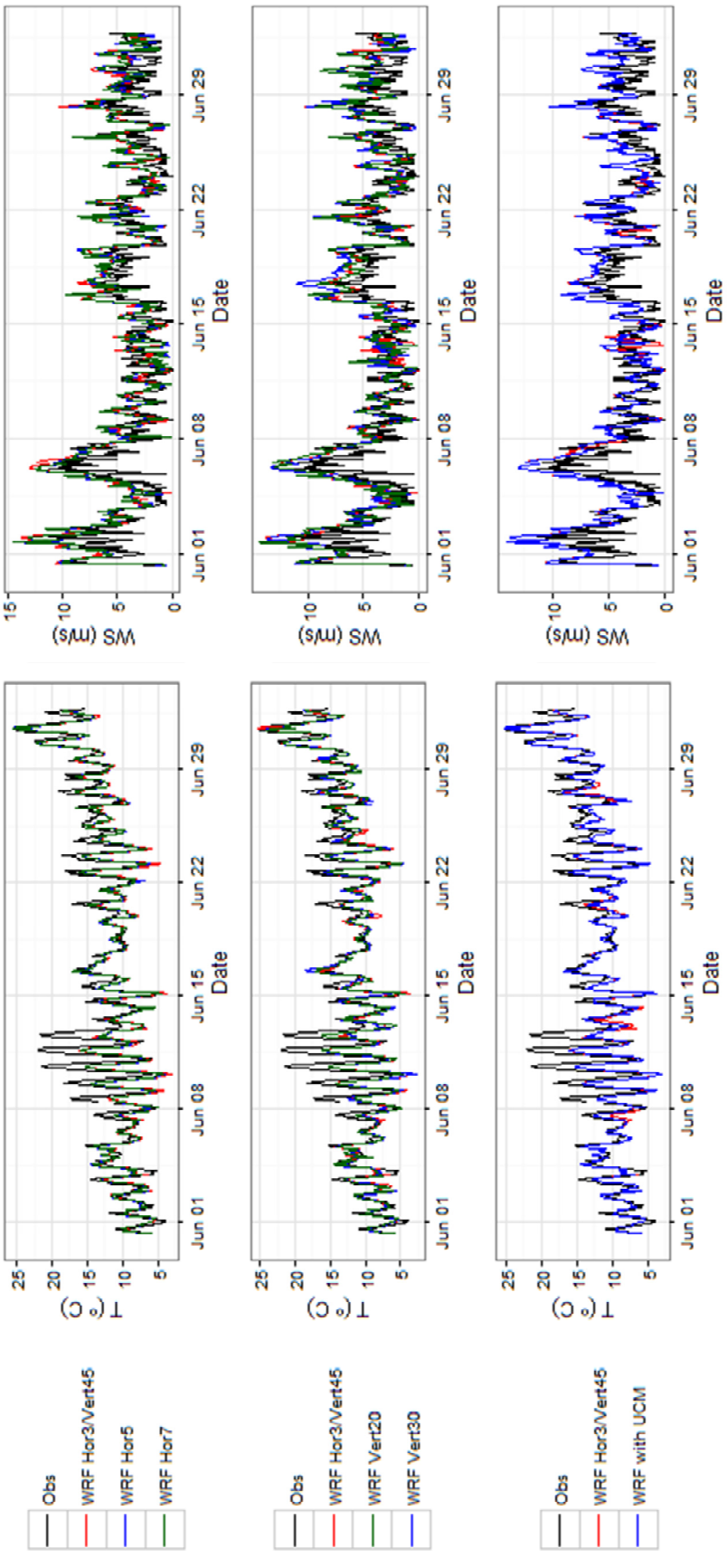


Figure S2: Time series plots of temperature and wind speed at Bishopton. Top graphs show the effect of changing the horizontal resolution on WRF predictions. middle shows vertical resolution influence and bottom shows the effect of including a UCM.

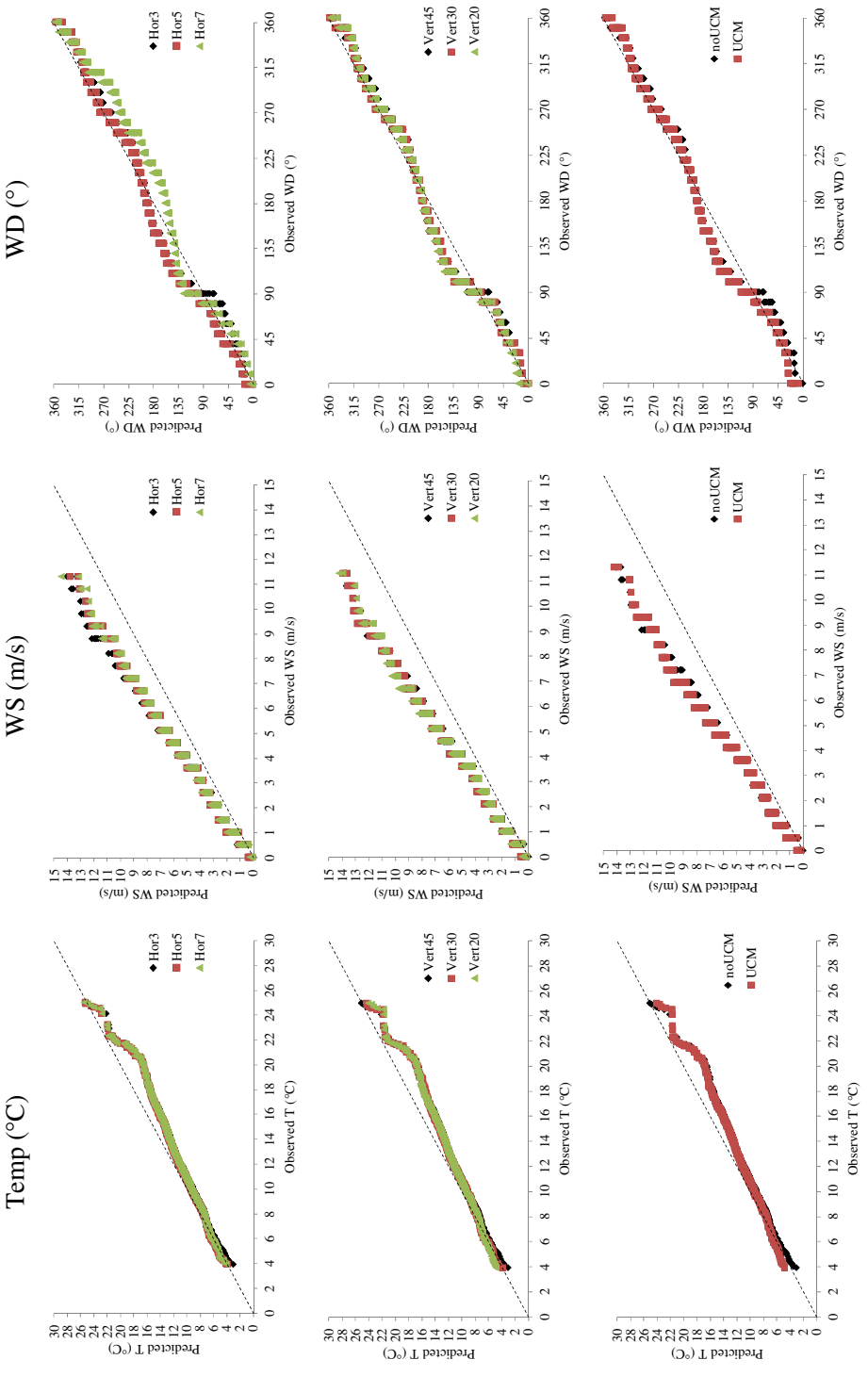


Figure S3: Q-Q plots of predicted and observed temperature, wind speed and wind direction at Bishopton for each of the WRF model setups. The black dashed line denotes a 1:1 relationship between predictions and observations.

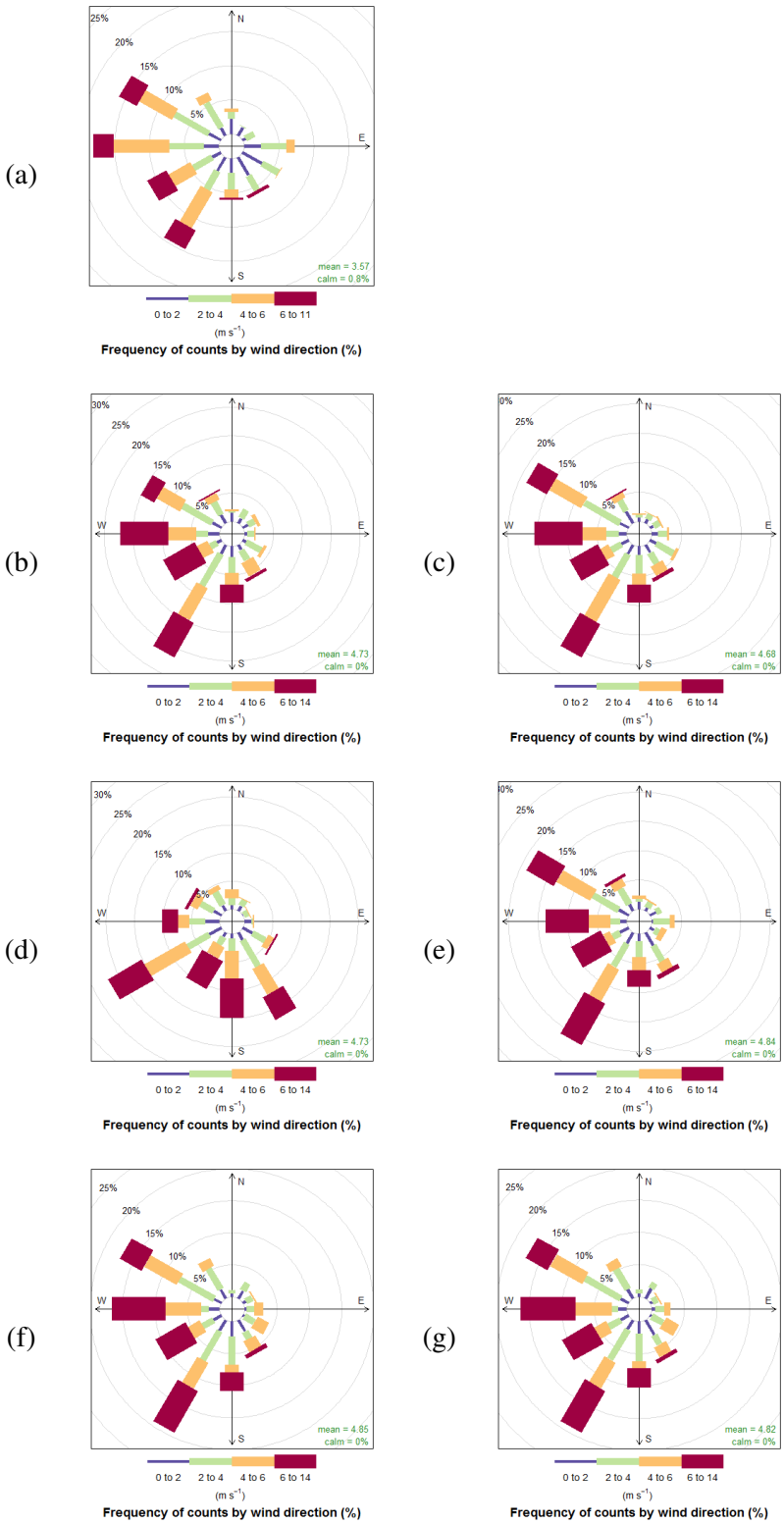


Figure S4: Wind rose plots for Bishopton for the month of June: (a) observed meteorology data; (b) Hor3; (c) Hor5; (d) Hor7; (e) Vert30; (f) Vert20; and (g) UCM.

Figure S5: Monthly average dispersion plume concentrations (µg/m^3) for a 10 m stack, located in the centre of the study area, emitting at 1 g/s predicted by: (a) Observed-AERMOD; (b) WRF-AERMOD (Hor3); (c) WRF-AERMOD (Hor5); (d) WRF-AERMOD (Hor7); (e) WRF-AERMOD (Vert30); WRF-AERMOD (Vert20); and WRF-AERMOD (UCM). The road network in the study area is shown in grey.

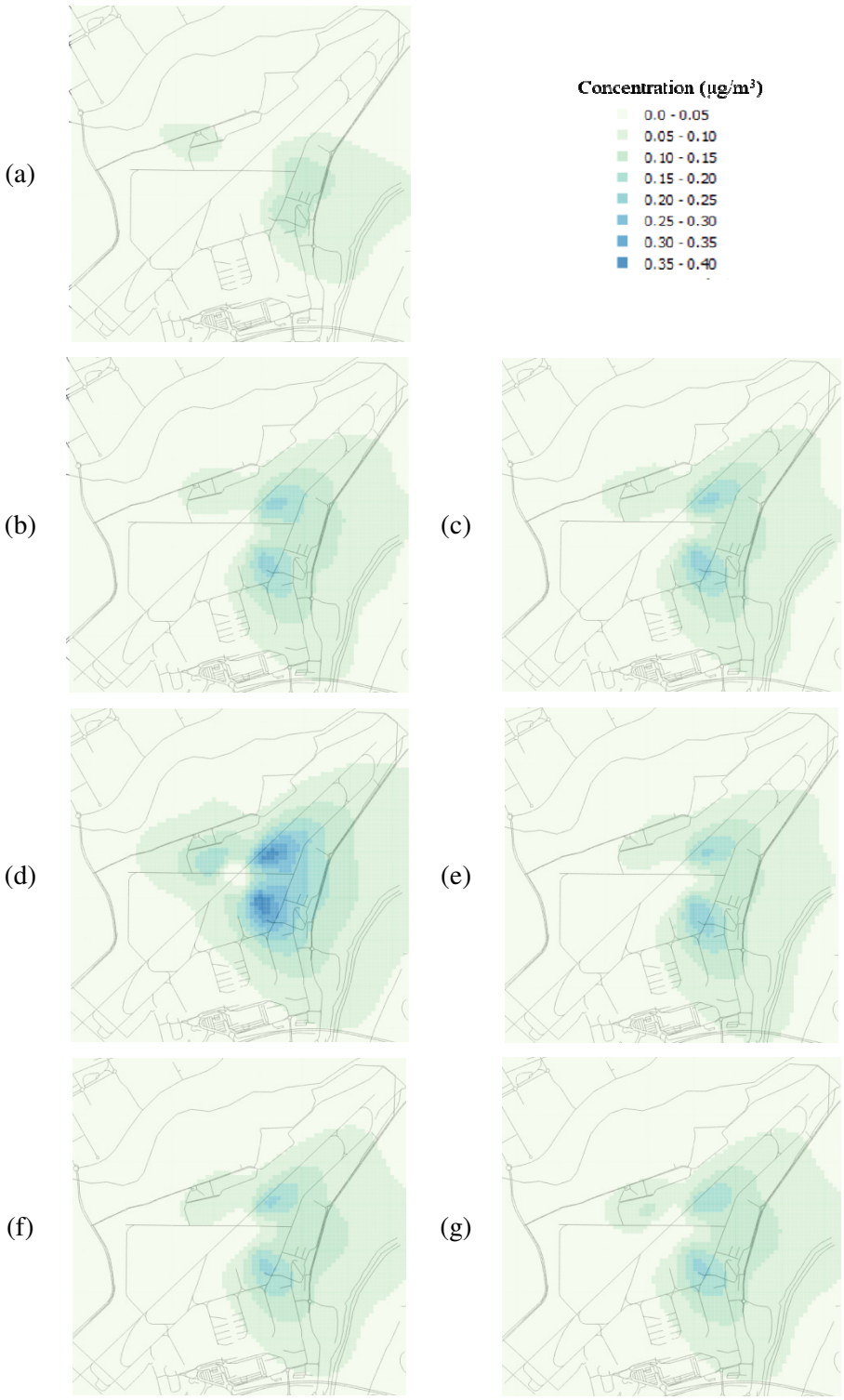


Figure S6: Monthly average dispersion plume concentrations (µg/m^3) for a 100 m stack, located in the centre of the study area, emitting at 1 g/s predicted by: (a) Observed-AERMOD; (b) WRF-AERMOD (Hor3); (c) WRF-AERMOD (Hor5); (d) WRF-AERMOD (Hor7); (e) WRF-AERMOD (Vert30); WRF-AERMOD (Vert20); and WRF-AERMOD (UCM). The road network in the study area is shown in grey.

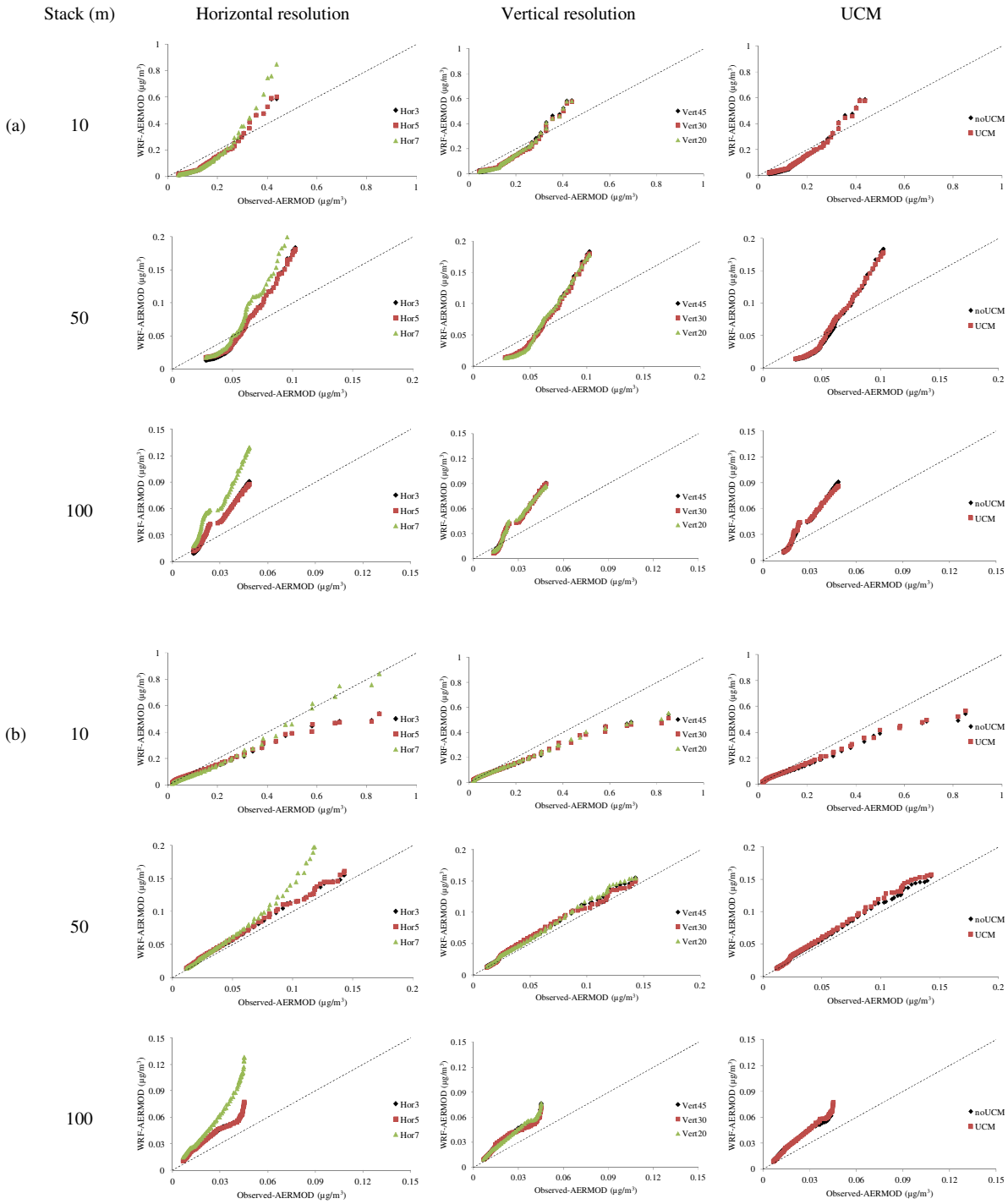


Figure S7: The average concentration (in µg/m^3) at each unique (a) x and (b) y is compared in Q-Q plots for the predicted and observed concentrations (n = 81).

Table S1: Urban Canopy parameters used in the UCM setup.

	High density residential	Low density residential	Commercial
Roof height (m)	17	9	11
SD roof height (m)	10	4	7
Width of roof (=sqrt roof area)	22	16	33
Road width (m)*	15	14	16
Anthropogenic heat (default)	50	20	90
Fraction urban landscape with no natural vegetation (default for commercial and low. Change high from default value of 0.90).	0.85	0.5	0.95
Heat capacity of roof / building wall / ground (default for all)	1E6 / 1E6 / 1.4E6		
Thermal conductivity of roof / wall / ground (default)	0.67 / 0.67 / 0.4004		
Surface albedo of roof / wall / ground (default)	0.2 / 0.2 / 0.2		
Surface emissivity roof / wall / ground (default)	0.9 / 0.9 / 0.95		
Roughness length for momentum over ground / roof (default)	0.01 / 0.01		
Coefficient modifying kanda approach to computing surface layer exchange coefficients (default)	1.29		
Thickness of building wall / building wall layer (default)	0.05 / 0.05		
Thickness of road layer (default)	0.05, 0.25, 0.50 and 0.75		
Boundary condition for roof layer temperature / wall layer temperature / ground layer temperature (default)	1 / 1 / 1		
Lower boundary condition for roof temperature / wall temperature / ground temperature (default)	293 / 293 / 293		
Ch of wall and road (default)	2		
Surface and layer temperatures (default)	1		
Ahoption (default)	0		
Anthropogenic heating profiles (default)	0.16, 0.13, 0.08, 0.07, 0.08, 0.26, 0.67, 0.99, 0.89, 0.79, 0.74, 0.73, 0.75, 0.76, 0.82, 0.90, 1.00, 0.95, 0.68, 0.61, 0.53, 0.35, 0.21, 0.18		

7. Evaluation of the RapidAir dispersion model in London and the use of geospatial surrogates to represent street canyon effects

The use of air pollution models to assess population exposures to air pollutants has advantages over measurements including the ability to be able to estimate concentrations over a large number of sites. However, there are a number of limitations associated with the use of models, including the high costs in terms of licensing and computing requirements which can make their use limited for large study area. RapidAir® is a dispersion model developed by Ricardo Energy & Environment which can produce highly spatially resolved pollution estimates in relatively short computation times through use of modern scientific computing developments. We evaluate the RapidAir model in London, UK, at 107 receptor locations to assess the model performance, and additionally we investigate if the inclusion of models or surrogates to account for street canyon morphologies can be applied to dispersion models for large study areas and the impact the inclusion of these has on modelled concentrations.

This manuscript has been produced as a joint effort between Ricardo Energy & Environment and the University of Strathclyde. The experiment was designed by, model run, data analysed and manuscript written by N. Masey. The model was developed by S. Hamilton and Ricardo Energy & Environment. S. Hamilton provided model code, discussion about experimental design, data analysis and editorial comments on the manuscript. I. Beverland provided editorial comments on the manuscript.

This manuscript has been prepared for submission to Environmental Modelling and Software.

Development and evaluation of the RapidAir® dispersion model, including the use of geospatial surrogates to represent street canyon effects

Nicola Masey[1], *Scott Hamilton*[2*], *Iain Beverland*[1]

[1] Department of Civil and Environmental Engineering, University of Strathclyde, 75 Montrose Street, Glasgow, UK, G1 1XJ

[2]Ricardo Energy & Environment, 18 Blythswood Square, Glasgow, UK, G2 4BG

*CORRESPONDING AUTHOR: Dr Scott Hamilton, Ricardo Energy & Environment, 18 Blythswood Square, Glasgow, UK, G2 4BG; email: scott.hamilton@ricardo.com; Tel: +44 (0)1235 753 716

Research Highlights:

- RapidAir® dispersion model evaluated for annual NO_x and NO_2 in London, UK
- 5 x 5 m resolution model generated for 3,500 km^2 area in under ten minutes
- Annual NO_2 NMB = -0.12 µg/m^3 with larger underestimation at kebside sites (-0.31 µg/m^3)
- Inclusion of canyon models reduced bias in street canyons (NMB -0.15 *vs.* -0.21 µg/m^3)
- Geospatial surrogates used to represent urban morphology over large geographical areas

Abstract

We developed a dispersion model (RapidAir®) to estimate air pollution concentrations at fine spatial resolution over large geographical areas with fast run times. RapidAir® was evaluated by estimating NO_x and NO_2 at 107 locations in London, UK (consisting of 12 reference analyser and 95 passive diffusion tube sites). Concentrations were modelled at 5 m spatial resolution over an area of ~3,500 km^2 in < 10 minutes per run. We included discrete canyon models or geospatial surrogates (sky view factor, hill shading and wind effect) to improve the accuracy of model predictions at kerbside locations. Geospatial surrogates provide alternatives to discrete street canyon models where it is not practical to run canyon models for hundreds to thousands of separate streets within a large city dispersion model (with advantages including: ease of operation; faster run times; and more complete / transparent treatment of building effects).

Keywords: Dispersion modelling; Air pollution; GIS; NO_x; NO_2; street canyon

1. Introduction

The estimation of population exposures to air pollution is increasingly important as numerous studies highlight the detrimental effects of air pollution on human health (World Health Organization, 2013, 2016). The use of air pollution monitors allows direct measurement of ambient concentrations, and the on-going development of portable real-time monitors is providing improvements in temporally resolved concentration estimates (Dons et al., 2012; Spinelle et al., 2017, 2015). However, monitoring only provides concentration estimates at specific locations, whereas it has been observed that pollution concentrations can vary substantially over small areas (Gillespie et al., 2017; Lin et al., 2016). Models can overcome some of the limitations associated with monitoring as concentrations can be estimated at multiple locations within a study area. However, inherent uncertainties within models require to be quantified by comparison of predictions against air pollution measurements.

Two main types of models are commonly used to estimate urban air pollution – land use regression (LUR) models and dispersion models (we do not include discussion of Computational Fluid Dynamics (CFD) models in this paper as CFD models have not been used widely in operational predictions of spatial patterns of urban air pollution due to excessive computational constraints when operating over large geographical areas).

Land use regression (LUR) models use Geographical Information Systems (GIS) to quantify relationships between measured pollutant concentrations and land use variables (including traffic and population), which can then be extrapolated to estimate human exposure to air pollution at fine spatial resolution (Briggs et al., 1997). LUR models have been widely applied in in cohort epidemiological studies (Gillespie et al., 2016; Johnson et al., 2013; Wang et al., 2013) and in personal monitoring studies (Dons et al., 2014a, 2014b). LUR models are frequently used to estimate longer-term (e.g. annual) pollution exposure and often do not take into account the effects of meteorology. Additionally the transfer of LUR models between study areas has been shown have substantial limitations (Gillespie et al., 2016; Mukerjee et al., 2012; Patton et al., 2015). Many regulatory organisations are interested in source apportionment to inform policy on air pollution controls, which requires preparation of spatially accurate multi-source air quality emissions. However, LUR models seldom use direct quantitative estimates of emissions from sources (instead more commonly they assess the effects of receptor proximity to sources) and consequently LUR models have had limited application in air quality management policy development.

Dispersion models simulate atmospheric transport and transformation of air pollutants emitted from sources to allow estimation of concentrations at receptors. The most

commonly used models are based on Gaussian plume concepts. Dispersion models can be used to estimate short term (e.g. hourly) variations in pollution concentrations (Gibson et al., 2013), and to estimate population exposures in cohort studies (Bellander et al., 2001; Nyberg et al., 2000). Additionally, projected emissions estimates (if available) can be used to estimate future concentrations. Commercially available software packages have been developed to simplify user inputs and modelling procedures, however this has often resulted in high license costs (Gulliver and Briggs, 2011), particularly when it is necessary to apply models over large geographical areas. Furthermore, Gaussian dispersion model run-times for large urban area can quickly become prohibitive due the computational demands of calculating concentrations at what can extend to millions of discrete locations. This may necessitate the use of GIS interpolation routines which may introduce other errors into estimated exposures.

Some studies have addressed these challenges to achieve fine spatial and temporal resolution by combining dispersion and LUR models (Beevers et al., 2012a; Korek et al., 2016; Michanowicz et al., 2016; Wilton et al., 2010); and/or including meteorological information within LUR models (Su et al., 2008a; Tan et al., 2016). For example a hybrid GIS-dispersion model (STEMS-AIR) has been developed to enable fine spatial and temporal resolution while minimising run times with readily-available computer software (Gulliver and Briggs, 2011). The STEMS-Air model estimates pollution concentrations from emission sources in 45 degree upwind 'wedge' shaped GIS-buffer areas, scaled by the distance between sources and receptors.

In built-up urban areas air pollution can become trapped in street canyons surrounded by tall buildings, especially if the wind is blowing from a direction perpendicular to the street, leading to recirculation of pollutants within the canyon. As a result, pollution concentrations in street canyons can become elevated and may be underestimated by 'standard' air pollution models. Exposure estimates may be improved by combining additional models that take into account urban topography in such locations with background pollution estimates from Gaussian-based air pollution models. Street canyon models range from complex computational fluid dynamic (CFD) models to simpler empirical (e.g. USEPA STREET box-model (Dabberdt et al., 1973; Johnson et al., 1973)) and semi-empirical models (e.g. Danish Operational Street Pollution Model (OSPM) (Vardoulakis et al., 2003)). Some dispersion models include additional software modules for street canyon effects, however these may increase model run time (Fallah-Shorshani et al., 2017; Jackson et al., 2016).

Geospatial surrogates can be used to estimate the effect of street canyons on air quality in urban locations. Such metrics are commonly used in studies of urban climate where

temperature, and hence comfort levels, are affected by building density and height. For example, sky view factor (SVF, which estimates the percentage of sky that can be observed using a fish-eye lens pointed vertically, with areas with low SVF corresponding to the presence of tall buildings (Carrasco-Hernandez et al., 2015)) has been incorporated into a LUR model to estimate the presence of street canyons (Eeftens et al., 2013). Building height and/or volume information has also been observed to improve the accuracy of LUR model estimates (Gillespie et al., 2016; Su et al., 2008b; Tang et al., 2013). Geospatial surrogates can be readily applied across entire cities in automated processes which are likely to be less susceptible to human error than use of currently available street canyon models, as the latter require user judgement to identify street canyon locations and detailed information (e.g. on traffic flow) for each location. The use of geospatial surrogates also has potential to improve the reproducibility of dispersion model pollution estimates as the number of model design choices is reduced substantially (with corresponding substantial reduction in manpower costs).

In this paper we describe the development and evaluation of a new dispersion model (RapidAir®, Ricardo-AEA Ltd) that uses modern scientific computing methods based on open-source Python libraries (www.python.org). A key motivation for the development of RapidAir was our experience of a lack of a cost-effective operational city-scale dispersion model with convenient run times, which does not require large amounts of manpower to operate. We focused on operational convenience of the modelling process and accuracy of model predictions in a case study and compared our results to results from other published studies which evaluated other modelling codes in a similar study area. The design concept for RapidAir is similar to the STEMS-Air model described by Gulliver and Briggs (2011) with some additional enhancements. RapidAir includes a dispersion model (AERMOD), with detailed treatment of boundary layer meteorology, and street canyon models. Additionally, we investigated the incorporation of geospatial surrogates to represent street canyon effects on spatial variation of pollution climates; and established methods to efficiently post-process the output from fine resolution dispersion models over large geographical areas using these surrogates.

2. Methods

2.1. Study area and receptor locations

We modelled concentrations of oxides of nitrogen (NO_x) in Greater London (urban conurbation approximately bounded by the M25 orbital motorway). Greater London was chosen as the study area because it contains a large network of air pollution monitoring sites, and has detailed open-access traffic and building height data. Additionally this was the study

area used in a previous Department for Environment, Food and Rural Affairs (DEFRA) Urban Model Evaluation exercise, which evaluated several commercially available and industry accepted models (Carslaw, 2011). We modelled annual average NO_x and NO_2 concentrations for 2008, which was the same year as used in the DEFRA study to enable comparison between RapidAir and the models assessed in the DEFRA comparison.

We evaluated the RapidAir model at 107 receptor locations: 12 Automatic Urban and Rural Network (AURN) sites; and 95 passive diffusion tube sites (Figure A1, Table A1). These sites are maintained by the Environmental Research Group, Kings College London (London AURN stations) and local authorities (diffusion tubes). The data collected were subject to national-ratification and detailed QA-QC procedures (DEFRA, 2017; Targa and Loader, 2008) and are available at https://uk-air.defra.gov.uk/ (AURN) and https://data.london.gov.uk/dataset/air-quality-summary-statistics (diffusion tubes). For model evaluation purposes the monitoring sites were designated as receptors classified as kerbside, roadside, suburban and urban background according to their proximity to road traffic: kerbside sites were located within 1 m of a busy road; roadside sites were located within 1 – 5 m of a busy road; suburban sites were located in a residential area on the edge of a urban conurbation; and urban background sites were located in urban areas but were free from the immediate influence of local sources to provide a good indication of background concentrations (DEFRA, 2016).

It was not possible to use exactly the same locations as the DEFRA Urban Model Evaluation (Carslaw, 2011). When we imported the locations used in the DEFRA into a GIS programme some were incorrect, in a few cases up to several kilometres from their true location. We relocated receptors to our best approximation of their true location using aerial photography and street level photographs but small discrepancies in the locations may still persist. This may have affected our evaluation of the accuracy of model predictions at measurement sites, and comparisons of our estimates with the estimates of other groups in this paper.

2.2. Model description

The RapidAir dispersion and decision support model uses open source Python libraries to rapidly estimate concentrations at high spatial resolution (5 m in this study) over extended geographical areas. RapidAir is conceptually similar to the STEMS-Air model published by Gulliver and Briggs (2011), with a number of technical developments. In RapidAir we control AERMOD (and its associated meteorological pre-processor AERMET) in a custom Windows, Linux or MacOS environment written in Python 2.7 (available from

www.python.org). The modelling system makes extensive use of the *numpy, scipy* and *pandas* libraries which enable very efficient scientific computation (partly due to their use of the C programming language).

AERMOD is an internationally recognised air dispersion modelling code and is a mandated USEPA model of choice for road traffic air quality assessments (U.S. Environmental Protection Agency, 2015). The USEPA impose strict guidelines on use of dispersion models and the use of AERMOD is written into law in 'Appendix W' of the Guideline on Air Quality Models (U.S. Environmental Protection Agency, 2017). RapidAir is parameterised using guidance from the USEPA in their 'Hotspot Conformity' guidance (U.S. Environmental Protection Agency, 2015, Appendix J). This sets out methods for dispersion modelling of road traffic emissions - for example prescriptive methods are given for setting release height, initial plume depth and other factors. The combination of using a tried and tested kernel modelling core and parameterising it to international statutory requirements makes RapidAir a robust modelling platform for dispersion modelling for road traffic sources.

RapidAir includes an automated meteorological processor based on AERMET which obtains and processes meteorological data of a format suitable for use in AERMOD. A major advantage of this approach is reproducibility as human error can often lead to problems when working with the AERMOD/AERMET suite. Surface meteorological data was obtained from Heathrow Airport (available from ftp://ftp.ncdc.noaa.gov/pub/data/noaa/) and upper air data from Camborne (available from http://www.esrl.noaa.gov/raobs/) which were the nearest meteorological stations to the study area.

The dispersion model uses a kernel convolution procedure which is similar to algorithms used in image processing software. A moving-window dispersion model plume (the kernel) calculated for a small idealised area source in AERMOD is passed over a road traffic emission raster at the same resolution pixel by pixel so the final city wide model comprises millions of overlapping plumes from the road source emissions. A weighted kernel (55 x 55 cells of size 5 m) was created to characterise concentrations at progressively lower resolution with distance from the source. A theoretical source was located at the centre of the kernel in AERMOD, assigned with a constant emission rate of 1 g/s.

Link based NO_x emissions data were obtained for London in 2008 from the London Atmospheric Emissions Inventory (LAEI) website (https://data.london.gov.uk/dataset/laei-2008) (Figure A1). In this study the link-based emissions from the LAEI were converted to a 5 m raster using the ESRI ArcGIS 'Line Density' tool (ESRI, 2014) though RapidAir has since moved to open source routines for preparing the emissions grid. During the AERMOD

step we rotated the wind direction by 180 degrees to represent the contribution of cells within the kernel to the central cell i.e. the cell in which we are trying to estimate the pollution concentration. This produced a plume which identifies pollution sources that contributed to the central cell and estimated a scaling factor for each source that fell within the plume based on its distance and location to the source.

For each receptor cell (in this case at 5 m resolution) the sum of concentrations falling within the kernel plume, which are by definition weighted by their distance to the source, are written to the centre cell of the concentrations raster. In this way the pollution surface is created by the convolution step iterating over the gridded emission data. This means that model run time is linearly dependent on the spatial resolution of the output number of cells and is unaffected by the number of emissions sources in the domain. This is a key benefit compared with other Gaussian models whose run time is linearly dependent both on resolution/number of receptors and number of sources. Our experience suggests that run times in the order of several days/weeks can be expected for city scale Gaussian models with only a few hundred thousand receptor locations, which are then interpolated to provide continuous pollution surfaces. In contrast, the RapidAir model computes concentrations at > 100 million discrete receptors in less than 10 minutes.

Regional background concentrations calculated by the Pollution Climate Mapping (PCM) model were added to the pollution raster (Figure A2). PCM concentrations (at 1 x 1 km grid resolution) are available to download from https://uk-air.defra.gov.uk/data/modelling-data). The PCM model estimates background concentrations taking into consideration a variety of pollution sources and breakdowns of the sources are provided. We removed those sources attributed to road transport prior to adding the PCM model to the modelled pollution concentrations above to prevent double-counting of traffic related air pollutants.

2.3. Surrogates for street canyons

Building height data were used to calculate simple surrogates that could readily be applied to a study area to indicate those locations that were located within street canyons, and consequently allow the modelled concentrations in these areas to be corrected for resulting urban morphological effects on air pollution. Building height data for London was downloaded from http://buildingheights.emu-analytics.net/ and a 5 m raster created of the maximum building height within each cell. The building height data was derived by the suppliers from national scale LiDAR surveys published by the Environment Agency for England (http://environment.data.gov.uk/ds/survey/index.jsp#/survey). We investigated three surrogates for street canyons (Figure A3):

- Sky view factor (*SVF*) produces a value representing the amount of sky visible from each location when looking vertically up to the sky and using a fish eye lens (a value between 0 and 1, where 1 is all visible sky). The Relief Visualization Toolbox (RVT) (Kokalj et al., 2011; Zaksek et al., 2011) was used to calculate this value using the building height raster as the input and a search radius of 200 m (Eeftens et al., 2013).
- Hill shading (*HS*) is commonly used to identify areas in shade as a result of surrounding topographical features (Zaksek et al., 2011). In our analysis we used wind direction in place of the direction of the sun and the 'shading' identified was anticipated to represent areas of higher concentration on the windward side of the canyon. The Analytical hill-shading option was run within RVT using an elevation angle of 45 degrees (suggested to be most appropriate for steep terrain as is encountered in an urban environment (Kokalj et al., 2013)). We calculated *HS* values for 8 sectors (i.e. every 45 degrees) and averaged the HS values calculated to produce estimated *HS* value over the study area.
- Wind Effect (*WE*) is a module in SAGA GIS (Conrad et al., 2015) which predicts if an area is wind shadowed or exposed, where values below 1 area shadowed and above 1 are exposed (Böhner and Antonić, 2009). As above for *HS*, *WE* values were calculated for 8 sectors and the average if these values were used. A search radius of 200 m was used.

The surrogate values for 5 m buffers around each receptor location were extracted to allow for slight errors in the coordinates of receptor locations (e.g. receptors located 'within' buildings rather than on lampposts on the road).

2.4. Model evaluation

Modelled concentrations of NO_x and NO_2 were extracted from the model outputs at the gird references for the monitoring sites to enable comparison. The R package OpenAir (Carslaw and Ropkins, 2012) was used to generate model evaluation statistics commonly used to evaluate pollution models, including FAC2, mean bias (MB), normalised mean bias (NMB), root mean square error (RMSE), coefficient of efficiency (COE) and index of agreement (IOA) (Carslaw, 2011; Chang and Hannah, 2004; Derwent et al., 2010).

3. Results and Discussion

The baseline RapidAir kernel model (i.e. no urban morphology treatment) highlighted contributions from major roads in London, and Heathrow Airport in the west of the study

area (Figure A4), however systematically underestimated observed NO_x concentrations at the receptor locations. The systematic underestimation of the PDT concentrations by the kernel model may be the result of uncertainties in background concentrations, road traffic emissions or monitoring locations. However, as we used publicly available open source data to generate the model we did not investigate these uncertainties further to ensure reproducibility of the model results and comparability with other groups who also used the same data sets.

It is likely that road traffic NO_x emissions data are underestimated in the inventory we used. These data were prepared by a statutory body (Greater London Authority (GLA)) and are the officially recognised emissions dataset for London. The European Environment Agency's COPERT road traffic emissions model has been observed to under-predict historical NO_x emissions from diesel vehicles in the UK fleet (Carslaw et al., 2011). Given the date of the data used in the model, the fact that evidence for COPERT under predicting traffic NO_x only came to light around 2011, and the fact that COPERT was used by the GLA to make the emissions estimates in 2008, it is likely that reported under-prediction of emissions in the diesel fleet biased the inventory towards under-prediction. We corrected the NO_x kernel model for systematic underestimation bias using the regression equation derived between the modelled and measured concentrations following UK statutory operational guidance provided by DEFRA (2016), which included forcing the intercept through the origin. The receptor locations were split randomly into *training* (n = 63) and *test* (n = 32) data sets, with the latter used as an independent verification data set. The linear regression (using the *training* data) for the model adjustment of the raw model was:

$$PDT\ measured\ NOx = 1.79 * Kernel\ modelled\ NO_x,\ R^2 = 0.88 \quad [1]$$

Where *PDT measured* NO_x and *Kernel modelled* NO_x are concentrations in µg/m^3.

Legislative limit values specified by the European Union and UK government are for NO_2, and not NO_x, therefore we converted RapidAir NO_x concentrations to NO_2 concentrations using the DEFRA NO_x to NO_2 calculator (version 3.2 downloaded from https://laqm.defra.gov.uk/archive/archiveno-calculator.html on 20/3/2017). The calculator was set to use the built-in fleet composition for London (which automatically sets the fraction of NO_x emissions as NO_2 (f-NO_2)) and the average NO_x background concentration over the study area from the PCM model. Estimated NO_2 concentrations were plotted against NO_x concentrations and fitted with a polynomial regression equation (Equation 2 and Figure A5) subsequently applied to the kernel model output to estimate NO_2 concentrations over the study area:

$$NO_2 = -0.0001 * (NO_x)^2 + 0.2737 * NO_x + 18.648, \quad R^2 = 1.00 \qquad [2]$$

where NO_x and NO_2 concentrations are in $\mu g/m^3$.

The calculator uses estimates of regional NO_2, NO_x and O_3 concentrations from the PCM model for the local authority area being modelled. We compared NO_2 conversion estimates for two local authorities within our study area, which had different regional NO_2, NO_x and O_3 concentrations, and found little effect on the NO_x to NO_2 conversion rate (Figure A5).

The concentrations of NO_2 estimated from RapidAir (Figure 1) were compared to NO_2 concentrations measured at the receptor locations described in Table A1. The remainder of this manuscript focuses on the analysis of the model for NO_2 concentrations. Corresponding tables and figures for NO_x model evaluation are provided in Appendix A: Supplementary Information.

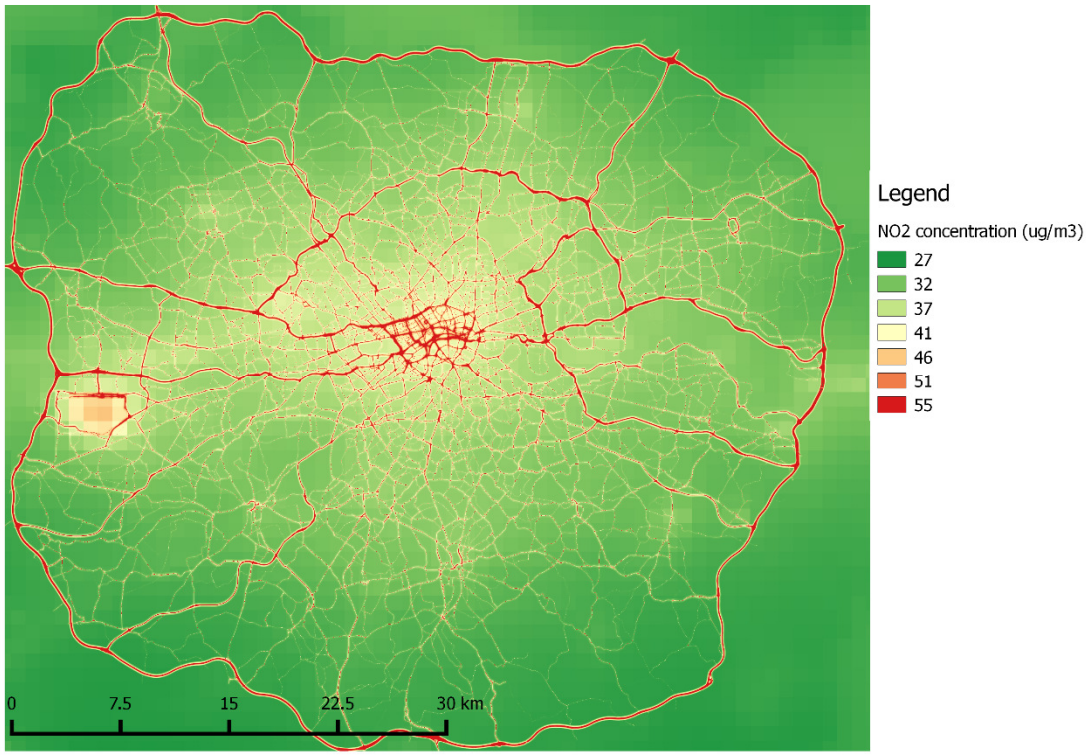


Figure 1: NO_2 concentrations estimated by RapidAir model over the Greater London conurbation for the RapidAir kernel model after correction for systematic biases.

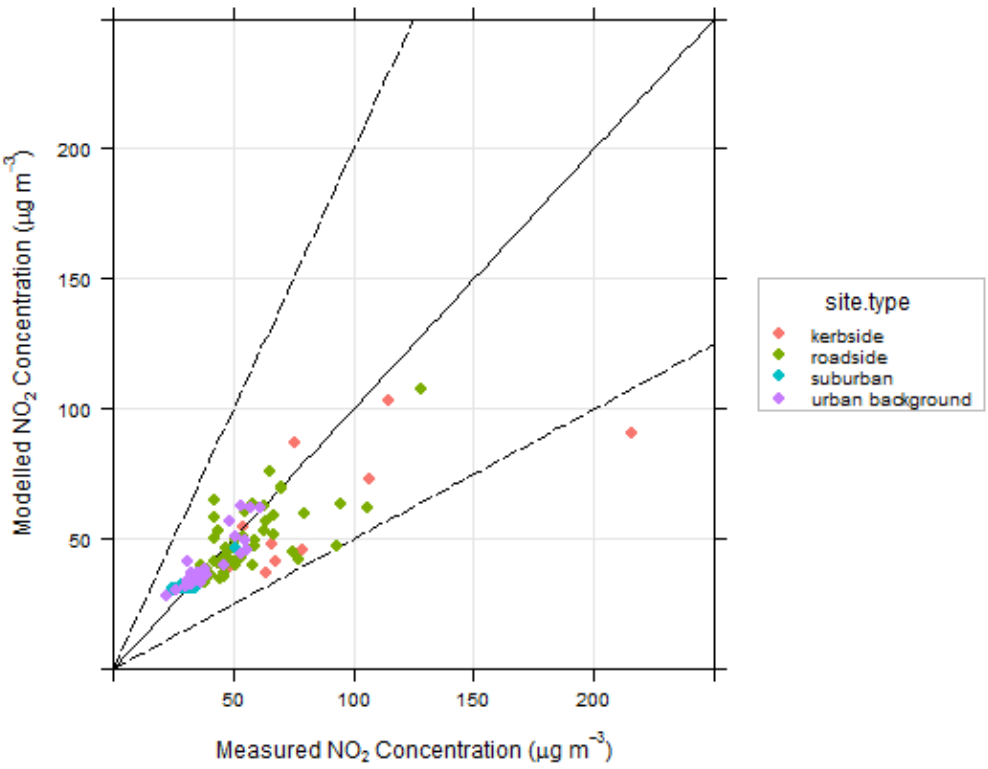


Figure 2: Scatter plot of NO_2 estimated by RapidAir kernel model *vs*. observed concentrations at PDT receptor locations (n = 95). Receptors are colour coded to represent the different site types. Solid line represents 1:1. Dashed lines represent FAC2 values.

3.1. Kernel model evaluation at receptor locations

NO_2 concentrations predicted by RapidAir were similar to measured NO_2 concentrations at PDT receptor locations; however the model underestimated concentrations at some very high concentration kerbside measurement sites (Figure 2, Table 1). Underestimation by RapidAir model could be attributed to urban morphologies (including street canyon effects) or underestimation in the emissions rates used to predict the NO_x concentrations (Beevers et al., 2012b). The correlation between modelled and observed NO_2 concentrations was high (r = 0.78) and of similar magnitude to previous evaluations of dispersion models (e.g. r = 0.74 during an evaluation of NO_x dispersion models carried out during the ESCAPE study (de Hoogh et al., 2014)).

Table 1: NO_2 kernel model evaluation statistics (after adjustment for systematic bias) for PDT receptor locations categorised by site type, and for AURN receptor locations.

Site type	n	FAC2	MB ($\mu g/m^3$)	NMB	RMSE ($\mu g/m^3$)	r	COE	IOA
PDT All	95	0.99	-6.12	-0.12	18.32	0.78	0.43	0.71
PDT Kerbside	10	0.90	-27.49	-0.31	44.88	0.68	0.13	0.56
PDT Roadside	45	1.00	-7.36	-0.13	15.57	0.69	0.23	0.62
PDT Suburban	13	1.00	1.03	0.03	3.46	0.90	0.33	0.67
PDT Urban background	27	1.00	0.43	0.01	5.02	0.88	0.57	0.78
All AURN	9	1.00	-0.55	-0.01	6.93	0.97	0.73	0.86

DEFRA recommend that an air quality model is acceptable for use if more than half of its observations fall within a factor of 2 of the observations (Williams et al., 2011). The NO_2 RapidAir model meets the *FAC2* criterion for all site types, with the lowest FAC2 value calculated for kerbside sites (FAC2 = 0.90) (Table 1). Kerbside concentrations represent the worst-case exposure scenarios that are not representative of population exposures over extended periods, and consequently annual limit values do not apply at these sites (DEFRA, 2016). Similar findings were reported in the DEFRA urban model evaluation exercise for NO_2 which found that *FAC2* values were lower for the kerbside sites than the three other site types tested, however all models met the criteria at the different site types (Carslaw, 2011). One of the groups participating in the DEFRA exercise included queuing emissions (presumably as an additional emission source), which resulted in higher emission rates at some locations than the other models and consequently improved the model agreement (FAC2 = 1.00). We made no attempt to include queuing emission (which are not included in the official release of the LAEI 2008). However, had data been available, RapidAir could have been used to estimate dispersion from queuing traffic. It should be noted that the number of sites tested in the DEFRA model evaluation was less than the number of sites evaluated in our study – DEFRA only included receptor locations where all models predicted however no details were provided about the sites not included in the published report.

Another criterion proposed by DEFRA to identify acceptable models is that *NMB* values should lie between -0.2 and 0.2 (Williams et al., 2011). *NMB* values for RapidAir meet this criterion when all sites were considered together; and for the individual site types, with the exception of the kerbside sties (Table 1). None of the models tested during the DEFRA model evaluation exercise met the *NMB* 'acceptance values' proposed by DEFRA at the kerbside sites. The numbers of models meeting the criteria gets progressively higher for

kerbside, roadside and urban background site classifications – with all models meeting the *NMB* values at urban background locations (Carslaw, 2011).

Model evaluation statistics were calculated for the AURN stations, using annual mean concentrations from hourly measurements (Table 1). The correlation between modelled and measured concentrations was greater for the AURN sites (r = 0.97) compared to the PDT sites (r = 0.78), with lower bias values. It is possible that the added uncertainty associated with PDT measurements can reduce correlation between observed and modelled concentrations. Annual average NO_2 concentrations were similar in magnitude for co-located PDT and AURN measurements (data not shown), however annual average NO_x concentration measured by PDTs were lower than concentrations measured using the AURN analysers (Figure A9). The underestimation of NO_x concentrations measured using diffusion tubes has been attributed to incomplete oxidation of nitric oxide in the tube, which is most pronounced at roadside locations with high nitric oxide concentrations (Jimenez et al., 2011).

3.2. Accounting for street canyon effects in RapidAir

We investigated the inclusion of two techniques within the RapidAir model to describe the effects of street canyons on pollution concentrations. The first technique used geospatial surrogates to account for building morphologies within a study area, and the second applied industry-standard street canyon models to user-defined street canyon geometries. These techniques are discussed in the following sub-sections.

3.2.1. GIS-surrogates for street canyons

We investigated if street canyon surrogates measured at each receptor could be used to estimate, and subsequently correct for, the effects of urban morphology on modelled NO_x concentrations, and NO_x concentrations converted to NO_2 concentrations using the method described above.

The PDT NO_x receptors were split randomly into *training* (n = 63) and *test* (n = 32) datasets with the former used to develop to surrogate-correction equations and the latter used as an independent dataset to test the correction equations derived. A multiple-linear calibration equation was derived between *Unadjusted modelled* NO_x *vs*. both *PDT* NO_x and *Surrogate* for each of the three surrogate values investigated (Table 2a). Applying the calibration equations to the *test* PDT NO_x data resulted in similar coefficients of determination and regression equations to the RapidAir estimates (Table 2b). The correlation between the concentrations and surrogates was unaffected by the surrogate used (R^2 = 0.66).

Table 2: Panel (a) shows the linear regression equations between measured and modelled NO_x for the training data (n = 63) for the baseline model, and the baseline model after inclusion of the surrogates. These equations were used to adjust the test data, and the linear regression between the measured and modelled NO_x after this adjustment for linear bias is shown in panel (b). Where non-significant intercepts were calculated, the regression lines have been forced through the origin.

(a)	Surrogate	Measured NO_x	R^2
	RapidAir	1.79*RapidAir_NO_x	0.88
	SVF	1.55*RapidAir_NO_x – 49.53*SVF + 53.61	0.66
	WE	1.62*RapidAir_NO_x – 83.56*WE + 93.97	0.66
	HS	1.61*RapidAir_NO_x – 78.14*HS + 88.73	0.66
(b)	Model	Measured NO_x	R^2
	RapidAir	1.16*RapidAir_NO_x	0.88
	SVF	1.14*RapidAir_NO_x	0.89
	WE	1.13*RapidAir_NO_x	0.88
	HS	1.14*RapidAir_NO_x	0.88

3.2.2. Street canyon models

Of the 95 PDT receptor locations we identified 23 sites that were located within urban street canyons through observations of the urban morphology using GIS and Google Maps Street View (Map data ©2017 Google) (Table A1). Concentrations of NO_x within these street canyons were estimated using two street canyon models: the STREET model (Dabberdt et al., 1973; Johnson et al., 1973) and the AEOLIUS Model (Buckland and Middleton, 1999). CFD models are complex and have long run times therefore are not an operationally feasible solution for large scale model correction for canyon effects, therefore were not considered during this study.

The STREET model estimates pollution concentrations empirically within a street canyon based on the emissions estimates within the canyon, and takes into account vehicle-induced turbulence and entry of air from the top of the canyon. Concentrations were calculated for the windward (C_W) and leeward (C_L) sides of the canyon using equations 3 and 4:

$$CL = \frac{K*Q}{(U+0.5)*\left[(x^2+z^2)^{\frac{1}{2}}+L_0\right]} \quad [3]$$

$$CW = \frac{K*Q*(H-z)}{W*(U+0.5)*H} \quad [4]$$

Where K is a scaling constant (set to 14 here); Q is the emission rate (g/m/s); U is the wind speed (m/s); L_0 is the length of individual vehicles (set to 3 m); W is the width of the

canyon (m); H is the average building height of the canyon (m); x is the distance from emission source to receptor (m); and z is the receptor height (set to 1 m).

The AEOLIUS model was developed by the UK Meteorological Office in the 1990s and was originally developed as a nomogram procedure (Buckland and Middleton, 1999). The scientific basis for the model is presented in a series of papers (Buckland, 1998; Manning et al., 2000; Middleton, 1999, 1998a, 1998b). The AEOLIUS model shares many common features with the Operational Street Pollution Model (OSPM) (Berkowicz, 2000; Hertel and Berkowicz, 1989) which underpins many street canyon models included in commercial road source dispersion models. There are three principal contributions to concentrations estimated by the AEOLIUS model: a direct contribution from the source to the receptor; a recirculating component within a vortex caused by winds flowing across the top of the canyon; and the urban background concentration. The RapidAir model only takes the recirculating component from the canyon model and sums this with the kernel derived concentrations. The AEOLIUS model is written in python 2.7 and the implements the equations as described in the reference Met Office papers. We gratefully acknowledge the work of Professor Bryan Harris at the University of Bath, who previously developed a Mathcad version of the model and who diligently described the model formulations he obtained from the Met Office's original author Dr D. R. Middleton; and much of our work was based on Professor Harris' description of the system of equations which make up AEOLIUS (Harris, 2004) (see Appendix A).

A subset of 383 hours of the annual hourly meteorological data was used in the street canyon models to reduce model run times (Appendix A). The effect of using a subset of meteorological data on computed annual average concentrations compared to the whole dataset was minimal for both canyon models. AEOLIUS was slightly more sensitive to the use of a sampled meteorological record (STREET model: slope = 1.00, intercept = -0.13, R^2 = 1.00; AEOLIUS model: slope = 0.91, intercept = 0.58, R^2 = 0.99) (Figure A6).

The windward and leeward concentrations predicted by each of the street canyon models were averaged on the assumption that over a year concentrations are well mixed within the street canyon. The concentrations predicted within by the canyon model were then added to the baseline NO_x concentrations predicted by the RapidAir model (representing the urban background in the area), and the models corrected for systematic bias following the guidance in DEFRA Technical Guidance 2016 (DEFRA, 2016) (Table 3).

Table 3: Linear adjustment equations to account for systematic bias in kernel model performance. Equations are shown for the kernel model; the kernel model corrected for surrogates; and kernel model including street canyon model. Intercept was not significant for Kernel therefore this has not been included. Corrections following guidance in TG16 (DEFRA, 2016).

Model	Model_NO_x ($\mu g/m^3$)
Kernel	1.79*PDT_NO, R^2 = 0.63
STREET	1.07*PDT_NO_x + 35.97, R^2 = 0.68
AEOLIUS	1.46*PDT_NO_x + 19.13, R^2 = 0.68

3.2.3. RapidAir evaluation after accounting for street canyon effects

At the receptor locations identified as being located within street canyons the underestimation of the receptor concentrations is lowest for the street canyon models, with the surrogates model and kernel models similarly under predicting the concentrations (NO_2 NMB = -0.21 for kernel and surrogates, -0.11 for STREET and -0.15 for AEOLIUS models (n = 23)) (Table 4 (NO_2) and Table A2 (NO_x)). The STREET model predicted higher concentrations than the AEOLIUS model which resulted in the smaller NMB values (Figure A8). The difference in modelled concentrations between the STREET and AEOLIUS models was very small which is similar to previously published findings (Ganguly and Broderick, 2011, 2010; Gualtieri, 2010; Zhu et al., 2015).

When all receptor locations were considered, there was little difference between the pollution concentrations estimated at the receptor locations for the RapidAir model, surrogates and the street canyon models (Figure 3 (NO_2) and Figure A7 (NO_x)). There was also little difference in the model evaluation statistics when the surrogates and street canyon models were included (Table 4 (NO_2) and Table A2 (NO_x)). Inclusion of the street canyon models reduced the NO_2 NMB values compared to the standard kernel model, however inclusion of the surrogates had little impact on NMB values at the kerbside sites (Kernel = -0.31, Surrogates = -0.30, STREET = -0.24 and AEOLIUS = -0.25) (Table 2 (NO_2) and Table A2 (NO_x)). LUR models for NO_2 incorporating *SVF* street canyon surrogates also found little improvement in coefficient of determination values after surrogate inclusion (R^2 = 0.76 vs. 0.78) (Eeftens et al., 2013).

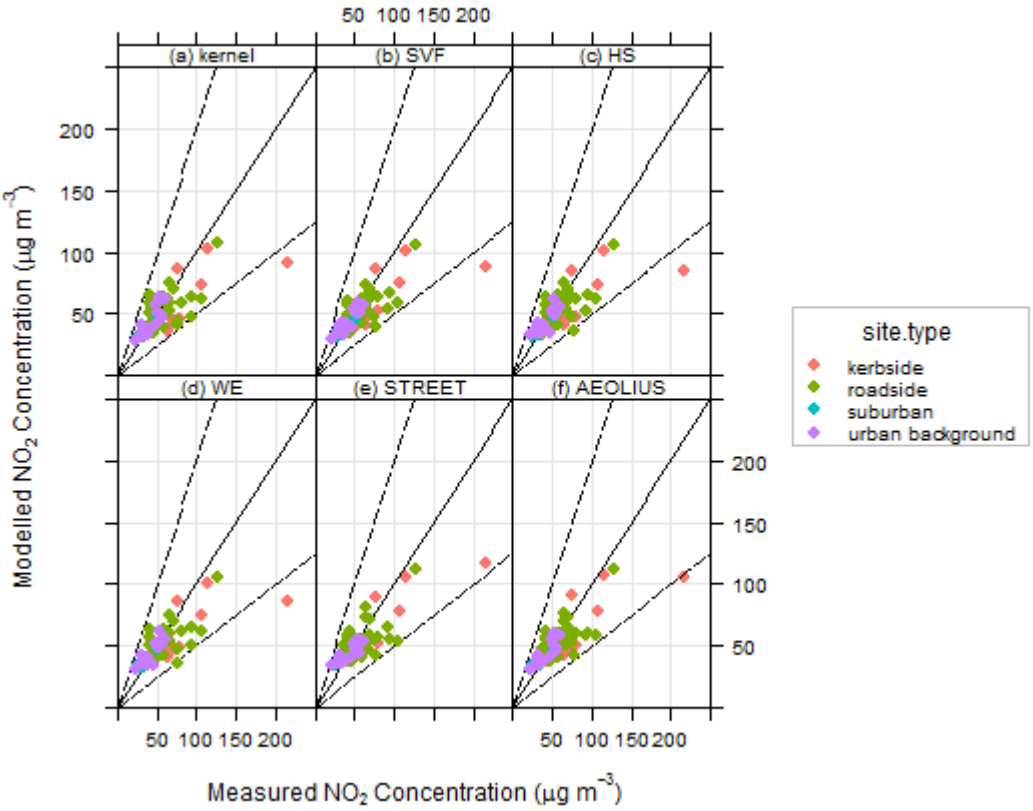


Figure 3: Scatter plot of NO_2 estimated by RapidAir kernel model *vs*. observed concentrations at PDT receptor locations (NO_2) (n = 95): (a) uncorrected concentrations from the base-kernel model; and the kernel model after correction using the surrogates for street canyons: (b) sky view factor (SVF), (c) hill shading (HS), (d) wind effect (WE), (e) STREET canyon model and (f) AEOLIUS canyon model. The solid line represents 1:1, while the dashed lines represent FAC2 values.

Despite the negligible change in model evaluation statistics the canyon models require less adjustment for systematic bias than the kernel model (Table 3). Therefore, when this model is applied to areas of the city which do not have any measurements the model is less likely to be subject to over or under estimation than the standard model which does not attempt to address urban morphology. The smallest systematic adjustments were required for the two separate street canyon models combined with the kernel model (Table 3). For instance the run including the USEPA STREET model required adjustment using the linear regression equation *Adjusted* NO_X *= 1.07 * Modelled NOx + 35.58.* The slope is significantly lower than the regression equation used to correct the raw model (1.07 vs 1.79). The measured concentrations of NO_x in the street canyons were typically very high (up to 520 µg/m^3) so the reduction of the slope required in the adjustment is an acceptable trade-off for

a small increase in the intercept. The results were similar for the AEOLIUS model. The inclusion of the street canyon models is therefore an important step in accounting for urban morphology which can in practice be as influential to concentrations as emission density for air pollution concentrations in an urban setting.

The use of surrogates to account for urban morphology effects, such as street canyons, have advantages over the more traditional street canyon models in that they can be more readily applied to large study areas. Using GIS means that surrogate values can rapidly be calculated across a large study area, and consequently correct the model quickly over the study area. Canyon models, however, require user input to select canyon locations (and therefore cannot be easily computed for large areas), and require additional information about canyon widths, heights, and traffic information such as speed. Additionally, the transition from "built up" to "open" within the city (for example at the boundary between a park and buildings) is treated in a gradual manner- unlike normal street canyon models which impose a hard boundary at the canyon edge which is "smoothed" artificially in a GIS with interpolation routines.

Currently these surrogates do not take wind speed into account which, for annual averages, we anticipate to have little influence on the model accuracy. However, if the surrogates were to be applied to a dispersion model with higher (e.g. hourly) temporal resolution then some modification of the surrogates to account for wind speed effects may be required in order to obtain similar modelled and measured pollution concentrations.

Table 4: Summary model evaluation statistics for annual mean NO_2 at receptor locations: (a) PDT locations (all sites and split by site type); (b) AURN sites only; (c) street canyon sites only. Statistics are given for the bias corrected Kernel only model, the kernel model after correction using the surrogates for street canyons and then bias corrected, and using the street canyon models with bias correction.

Site type	Model	n	FAC2	MB ($\mu g/m^3$)	NMB	RMSE ($\mu g/m^3$)	r	COE	IOA
(a) all sites:									
PDT All	Kernel	95	0.99	-6.12	-0.12	18.32	0.78	0.43	0.71
	SVF	95	0.99	-4.98	-0.09	17.80	0.80	0.45	0.73
	WE	95	0.98	-4.97	-0.09	18.24	0.78	0.44	0.72
	HS	95	0.98	-5.10	-0.10	18.38	0.78	0.43	0.72
	STREET	95	1.00	-4.10	-0.08	16.25	0.84	0.41	0.71
	AEOLIUS	95	0.99	-4.10	-0.08	16.54	0.83	0.42	0.73
PDT Kerbside	Kernel	10	0.90	-27.49	-0.31	44.88	0.68	0.13	0.56
	SVF	10	0.90	-26.24	-0.29	44.28	0.69	0.17	0.59
	WE	10	0.90	-26.62	-0.30	45.10	0.66	0.16	0.58
	HS	10	0.90	-27.21	-0.30	45.64	0.66	0.15	0.58
	STREET	10	1.00	-21.12	-0.24	35.58	0.84	0.45	0.65
	AEOLIUS	10	0.90	-22.59	-0.25	39.03	0.77	0.33	0.62
PDT Roadside	Kernel	45	1.00	-7.36	-0.13	15.57	0.69	0.23	0.62
	SVF	45	1.00	-6.65	-0.12	14.60	0.73	0.30	0.65
	WE	45	0.98	-6.40	-0.11	15.04	0.70	0.30	0.65
	HS	45	0.98	-6.53	-0.11	14.97	0.71	0.29	0.64
	STREET	45	1.00	-6.07	-0.11	15.57	0.66	0.19	0.60
	AEOLIUS	45	1.00	-5.55	-0.10	14.65	0.70	0.25	0.64
PDT Suburban	Kernel	13	1.00	1.03	0.03	3.46	0.90	0.33	0.67
	SVF	13	1.00	2.55	0.08	4.38	0.91	0.16	0.58
	WE	13	1.00	3.20	0.10	4.76	0.90	0.14	0.57
	HS	13	1.00	3.17	0.10	4.87	0.88	0.09	0.55
	STREET	13	1.00	4.86	0.15	6.57	0.90	-0.26	0.37
	AEOLIUS	13	1.00	3.51	0.11	5.15	0.90	0.23	0.54
PDT Urban background	Kernel	27	1.00	0.43	0.01	5.02	0.88	0.57	0.78
	SVF	27	1.00	2.05	0.05	4.89	0.91	0.53	0.77
	WE	27	1.00	1.49	0.04	5.42	0.87	0.48	0.74
	HS	27	1.00	1.48	0.04	5.69	0.85	0.47	0.73
	STREET	27	1.00	1.16	0.03	5.95	0.88	-0.04	0.73
	AEOLIUS	27	1.00	1.49	0.04	5.24	0.88	-0.12	0.76
(b) AURN sites only:									
	Kernel	9	1.00	-0.55	-0.01	6.93	0.97	0.73	0.86
	SVF	9	1.00	-0.54	-0.01	7.31	0.97	0.73	0.86
	WE	9	1.00	-0.23	-0.00	7.50	0.97	0.71	0.86
	HS	9	1.00	-0.80	6.25	7.26	0.97	0.71	0.86
	STREET	9	1.00	2.06	0.04	9.71	0.94	0.45	0.80
	AEOLIUS	9	1.00	2.05	0.04	8.32	0.95	0.50	0.83
(c) street canyon sites only:									
	Kernel	23	0.96	-17.20	-0.23	32.68	0.70	0.20	0.59
	SVF	23	0.96	-15.37	-0.20	31.68	0.72	0.25	0.63
	WE	23	0.96	-16.10	-0.21	32.46	0.70	0.22	0.61
	HS	23	0.96	-16.33	-0.22	32.72	0.69	0.21	0.61
	STREET	23	1.00	-8.62	-0.11	26.23	0.80	0.34	0.64
	AEOLIUS	23	0.96	-11.03	-0.15	28.24	0.76	0.33	0.64

3.3. Advantages and limitations of RapidAir

The central focus of this work is to evaluate an air quality modelling platform aimed at the operational setting where time is often a priority and manpower/computational resources are limited. An example of an operational use of RapidAir is given in Appendix A. The RapidAir model succeeds as an operational air quality model in the context of very large urban areas and as a decision support tool but the efficiency comes with some drawbacks. Therefore, it is appropriate to outline the key benefits and limitations of the approach to enable practitioners to interpret this work in light of their current experiences in running city scale dispersion models.

Clearly a significant benefit with RapidAir is reduced computational burden. Run times of 10 minutes or less for a very large city with > 8 million inhabitants present a significant benefit for the operational modeller and decision makers who require fast but robust analyses. The RapidAir platform allows extremely efficient policy testing and other "what if" model runs for new emission scenarios to be undertaken in a few minutes on a standard office computer which is to our knowledge not possible using existing platforms. The model performance metrics for RapidAir in Table 4 are very similar to those computed for other dispersion modelling systems in the DEFRA inter comparison exercise. For example the RapidAir outputs for kerbside locations in London have NO_2 RMSE values of 35.58 – 45.64 $\mu g/m^3$ (r = 0.66 - 0.84, n = 10) where the models in the inter comparison have RMSE values ranging from 29.39 to 67.09 $\mu g/m^3$ (r = 0.15 - 0.93, n = 7). At roadside locations the RapidAir outputs have NO_2 RMSE values of 14.60 – 15.57 $\mu g/m^3$ (r = 0.66 - 0.73, n = 45) where the models in the inter comparison have RMSE values ranging from 9.94 to 19.69 $\mu g/m^3$. (r = 0.38 - 0.89, n = 30). Some of the variation between RapidAir and the other models will be due to the different number of receptors in each category (which in reality may help or hinder our model performance) but it is impossible for us to match the locations exactly for the reasons explained earlier. The model results also yielded good results for the COE and IOA when compared with the definitions for these metrics provided by Carslaw and Ropkins (2012). Suffice to say the key model metrics for the 2008 model run in London are very similar to standard modelling suites used in the UK and which are used and accepted by DEFRA for use in compliance assessments at the highest level of statutory European air quality reporting.

In our view the potential drawbacks of the model must be balanced against the benefits described above. There may be the suggestion that the kernel based model represents a significantly simplified treatment of urban dispersion compared with models currently in use in the UK which iterate over thousands of receptors and calculate contributions at those

receptors as a function of those sources (with very significant run times). In fact all Gaussian and empirical models are already a greatly simplified picture of reality in urban settings and the methodology in RapidAir does not significantly alter the overall level of simplification compared with the real situation. In any case the model results are compared against pollution measurements as with all other models using the same metrics and the results of that performance assessment are comparable with other platforms. We therefore suggest that in the operational setting these perceived deficiencies in the methodology are greatly outweighed by the benefits and the model is demonstrably fit for purpose for use in air pollution modelling in the urban environment.

The performance statistics for the surrogates for urban morphology are reasonably close to those from the models which treat canyons discretely. Again our focus is on operational modelling where reproducible and efficient workflows are as important as the tools selected for use. Based on this work we would suggest that for compliance assessment RapidAir is used with either the STREET or AEOLIUS model options included as the run times are not significantly impacted by including these models. The model results should be compared with measured concentrations and the modeller may choose the best performing street canyon model for their case. The surrogate models should be used as screening tools and perhaps to spatially delineate locations where the street canyon models should be invoked, which is often difficult for a large and complex urban environment where resources do not permit thorough investigation and spatial treatment of the morphological conditions.

4. Conclusions

We developed a kernel-based dispersion model (RapidAir) combining AERMOD and open-source scientific computing methods to estimate pollution concentrations at fine spatial resolution. Model input data was sourced from public domains to facilitate comparison with pollution estimates by other research groups modelling the same location with the same input data. The RapidAir dispersion model took approximately 7 minutes to model the Greater London conurbation (~ 3,500 km^2) at 5 x 5 m resolution.

We evaluated NO_x and NO_2 model predictions at 107 sites (12 automatic monitoring and 95 passive sampling tube sites) across this conurbation. After correction for systematic under estimation bias in the initial RapidAir model, FAC2 values for modelled concentrations were > 0.90 at the 107 evaluation sites. RMSE values decreased through the site categories: Kerbside, Roadside, Urban Background and Suburban (RMSE = 45, 16, 5 and 4 $\mu g/m^3$ respectively) and the correlations between predicted and observed concentrations were higher for the less built up sites than the Roadside or Kerbside sites (r =

0.68, 0.69, 0.88 and 0.90 for Kerbside, Roadside, Urban Background and Suburban respectively). This finding is consistent with results from other modelling groups participating in the DEFRA inter comparison, whose RMSE and r values ranged from 3-70 $\mu g/m^3$ and 0.15 – 0.93 respectively.

The larger RMSE values at the sites in proximity to traffic sources may have resulted from the presence of street canyons that trap pollutants leading to elevated concentrations – an effect that cannot be described in dispersion models unless urban morphologies are taken into consideration. Correspondingly, we used geospatial surrogates (sky-view factor, hill shading and wind effect) and separate street canyon models (STREET and AEOLIUS) to improve modelled concentrations at roadside sites. The street canyon models increased the model accuracy while the street canyon surrogates had limited effect on model prediction accuracy (NMB kerbside: RapidAir base-kernel = -0.31, surrogates = -0.30, STREET model = -0.24 and AEOLIUS = -0.25). Consequently, the combined models may be anticipated to provide more accurate estimates when extrapolated to locations without monitoring. The geospatial surrogates have potential as simple means of incorporating canyon effects into a large city scale dispersion model. The advantage of using simple geospatial surrogates for street canyons instead of modelling canyons discretely include: reduced run times, smaller user input required and the transition from 'built up' to 'open' environments is treated gradually.

Overall the RapidAir model performance statistics were consistent with those found by other groups in London in 2008.

Acknowledgements

Nicola Masey is funded through a UK Natural Environment Research Council CASE PhD studentship (NE/K007319/1), with support from Ricardo Energy and Environment. We acknowledge access to the AURN and passive sampler measurement data, and the PCM model data, which were obtained from uk-air.defra.gov.uk and are subject to Crown 2014 copyright, Defra, licensed under the Open Government Licence (OGL). RapidAir® is a registered trade mark of Ricardo-AEA Limited, used under permission of Ricardo-AEA Limited.

References

Beevers, S.D., Kitwiroon, N., Williams, M.L., Carslaw, D.C., 2012a. 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

Beevers, S.D., Westmoreland, E., de Jong, M.C., Williams, M.L., Carslaw, D.C., 2012b. Trends in NOx and NO2 emissions from road traffic in Great Britain. Atmos. Environ. 54, 107–116. doi:10.1016/j.atmosenv.2012.02.028

Bellander, T., Berglind, N., Gustavsson, P., Jonson, T., Nyberg, F., Pershagen, G., Järup, L., 2001. Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. Environ. Health Perspect. 109, 633–639.

Berkowicz, R., 2000. OSPM - A Parameterised Street Pollution Model. Environ. Monit. Assess. 65, 323–331. doi:10.1023/A:1006448321977

Böhner, J., Antonić, O., 2009. Chapter 8 Land-Surface Parameters Specific to Topo-Climatology, in: Reuter, T.H. and H.I. (Ed.), Developments in Soil Science, GeomorphometryConcepts, Software, Applications. Elsevier, pp. 195–226. doi:10.1016/S0166-2481(08)00008-1

Briggs, D., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., Pryl, K., Van Reeuwijk, H., Smallbone, K., Van Der Veen, A., 1997. Mapping urban air pollution using GIS: a regression-based approach. Int. J. Geogr. Inf. Sci. 11, 699–718.

Buckland, A.T., 1998. Validation of a Street Canyon Model in Two Cities. Environ. Monit. Assess. 52, 255–267. doi:10.1023/A:1005828128097

Buckland, A.T., Middleton, D.R., 1999. Nomograms for calculating pollution within street canyons. Atmos. Environ. 33, 1017–1036. doi:10.1016/S1352-2310(98)00354-9

Carrasco-Hernandez, R., Smedley, A.R.D., Webb, A.R., 2015. Using urban canyon geometries obtained from Google Street View for atmospheric studies: Potential applications in the calculation of street level total shortwave irradiances. Energy Build. 86, 340–348. doi:10.1016/j.enbuild.2014.10.001

Carslaw, D., 2011. Defra urban model evaluation analysis - Phase 1. King's College London.

Carslaw, D., Beevers, S.D., Westmoreland, E., Williams, M., Tate, J., Murrells, T., Stedman, J., Li, Y., Grice, S., Kent, A., Tsagatakis, I., 2011. Trends in NOx and NO2 emissions and ambient measurements in the UK.

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

Chang, J.C., Hanna, S.R., 2004. Air quality model performance evaluation. Meteorol. Atmospheric Phys. 87, 167–196. doi:10.1007/s00703-003-0070-7

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Boehner, J., 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geosci. Model Dev. 8, 1991–2007. doi:10.5194/gmd-8-1991-2015

Dabberdt, W.F., Ludwig, F.L., Johnson, W.B., 1973. Validation and applications of an urban diffusion model for vehicular pollutants. Atmospheric Environ. 1967 7, 603–618. doi:10.1016/0004-6981(73)90019-X

de Hoogh, K., Korek, M., Vienneau, D., Keuken, M., Kukkonen, J., Nieuwenhuijsen, M.J., Badaloni, C., Beelen, R., Bolignano, A., Cesaroni, G., Pradas, M.C., Cyrys, J., Douros, J., Eeftens, M., Forastiere, F., Forsberg, B., Fuks, K., Gehring, U., Gryparis, A., Gulliver, J., Hansell, A.L., Hoffmann, B., Johansson, C., Jonkers, S., Kangas, L., Katsouyanni, K., Künzli, N., Lanki, T., Memmesheimer, M., Moussiopoulos, N., Modig, L., Pershagen, G., Probst-Hensch, N., Schindler, C., Schikowski, T., Sugiri, D., Teixidó, O., Tsai, M.-Y., Yli-Tuomi, T., Brunekreef, B., Hoek, G., Bellander, T., 2014. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. Environ. Int. 73, 382–392. doi:10.1016/j.envint.2014.08.011

DEFRA, 2017. The Data Verification and Ratification Process [WWW Document]. URL https://uk-air.defra.gov.uk/assets/documents/The_Data_Verification_and_Ratification_Process.pdf (accessed 4.1.17).

DEFRA, 2016. Local Air Quality Management Technical Guidance (TG16) (No. TG16). Defra.

Derwent, D., Fraser, A., Abbott, J., Jenkin, M., Willis, P., Murrells, T., 2010. Evaluating the performance of air quality models (No. Issue 3). Defra.

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., Van Poppel, M., Int Panis, L., De Prins, S., Berghmans, P., Koppen, G., Matheeussen, C., 2014a. Land use regression models as a tool for short, medium and long term exposure to traffic related air pollution. Sci. Total Environ. 476–477, 378–386. doi:10.1016/j.scitotenv.2014.01.025

Dons, E., Van Poppel, M., Kochan, B., Wets, G., Int Panis, L., 2014b. Implementation and validation of a modeling framework to assess personal exposure to black carbon. Environ. Int. 62, 64–71.

Eeftens, M., Beekhuizen, J., Beelen, R., Wang, M., Vermeulen, R., Brunekreef, B., Huss, A., Hoek, G., 2013. Quantifying urban street configuration for improvements in air pollution models. Atmos. Environ. 72, 1–9. doi:10.1016/j.atmosenv.2013.02.007

ESRI, 2014. ArcGIS Desktop Version 10.2.2. Environmental Systems Research Institute, CA, USA.

Fallah-Shorshani, M., Shekarrizfard, M., Hatzopoulou, M., 2017. Integrating a street-canyon model with a regional Gaussian dispersion model for improved characterisation of near-road air pollution. Atmos. Environ. 153, 21–31. doi:10.1016/j.atmosenv.2017.01.006

Ganguly, R., Broderick, B.M., 2011. Application of urban street canyon models for predicting vehicular pollution in an urban area in Dublin, Ireland. Int. J. Environ. Pollut. 44, 71–77. doi:10.1504/IJEP.2011.038404

Ganguly, R., Broderick, B.M., 2010. Estimation of CO concentrations for an urban street canyon in Ireland. Air Qual. Atmosphere Health Dordr. 3, 195–202. doi:http://dx.doi.org/10.1007/s11869-010-0068-5

Gibson, M.D., Kundu, S., Satish, M., 2013. Dispersion model evaluation of PM2.5, NOx and SO2 from point and major line sources in Nova Scotia, Canada using AERMOD Gaussian plume air dispersion model. Atmospheric Pollut. Res. 4, 157–167. doi:10.5094/APR.2013.016

Gillespie, J., Beverland, I.J., Hamilton, S., Padmanabhan, S., 2016. Development, Evaluation, and Comparison of Land Use Regression Modeling Methods to Estimate Residential Exposure to Nitrogen Dioxide in a Cohort Study. Environ. Sci. Technol. 50, 11085–11093. doi:10.1021/acs.est.6b02089

Gillespie, J., Masey, N., Heal, M.R., Hamilton, S., Beverland, I.J., 2017. Estimation of spatial patterns of urban air pollution over a 4-week period from repeated 5-min measurements. Atmos. Environ. 150, 295–302. doi:10.1016/j.atmosenv.2016.11.035

Gualtieri, G., 2010. A Street Canyon Model Intercomparison in Florence, Italy. Water Air Soil Pollut. Dordr. 212, 461–482. doi:http://dx.doi.org/10.1007/s11270-010-0360-x

Gulliver, J., Briggs, D., 2011. STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. Sci. Total Environ. 409, 2419–2429.

Harris, B., 2004. An Implementation of the OSPM/AEOLIUS air pollution model in MATHCAD and its application for a small Wiltshire town [WWW Document]. URL web.onetel.com/~bandmaharris/AeoliusInMathcad1.pdf (accessed 3.20.17).

Hertel, O., Berkowicz, R., 1989. Modelling pollution from trafic in a street canyon: evaluation of data and model development. Rep. DMU LUFT A129. doi:DOI: 10.13140/RG.2.1.2251.6326

Israel, G., 1992. Fact Sheet PEOD-6: Determining Sample Size.

Jackson, M., Hood, C., Johnson, C., Johnson, K., 2016. Calculation of Urban Morphology Parameterisations for London for use with the ADMS-Urban Dispersion Model. Int. J. Adv. Remote Sens. GIS 0, 1678–1687.

Jimenez, A.S., Heal, M.R., Beverland, I.J., 2011. Intercomparison study of NOx passive diffusion tubes with chemiluminescence analysers and evaluation of bias factors. Atmos. Environ. 45, 3062–3068.

Johnson, M., MacNeill, M., Grgicak-Mannion, A., Nethery, E., Xu, X., Dales, R., Rasmussen, P., Wheeler, A., 2013. Development of temporally refined land-use regression models predicting daily household-level air pollution in a panel study of lung function among asthmatic children. J. Expo. Sci. Environ. Epidemiol. 23, 259–267. doi:10.1038/jes.2013.1

Johnson, W.B., Ludwig, F.L., Dabberdt, W.F., Allen, R.J., 1973. An Urban Diffusion Simulation Model For Carbon Monoxide. J. Air Pollut. Control Assoc. 23, 490–498. doi:10.1080/00022470.1973.10469794

Kokalj, Z., Zaksek, K., Ostir, K., 2011. Application of sky-view factor for the visualisation of historic landscape features in lidar-derived relief models. Antiquity 85, 263–273.

Kokalj, Z., Zaksek, K., Ostir, K., Pehani, P., Cotar, K., 2013. Relief Visualization Toolbox version 1.3 - Manual.

Korek, M., Johansson, C., Svensson, N., Lind, T., Beelen, R., Hoek, G., Pershagen, G., Bellander, T., 2016. Can dispersion modeling of air pollution be improved by land-use regression? An example from Stockholm, Sweden. J. Expo. Sci. Environ. Epidemiol. doi:10.1038/jes.2016.40

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

Manning, A.J., Nicholson, K.J., Middleton, D.R., Rafferty, S.C., 2000. Field Study of Wind and Traffic to Test a Street Canyon Pollution Model. Environ. Monit. Assess. 60, 283–313. doi:10.1023/A:1006187301966

Michanowicz, D.R., Shmool, J.L.C., Cambal, L., Tunno, B.J., Gillooly, S., Olson Hunt, M.J., Tripathy, S., Naumoff Shields, K., Clougherty, J.E., 2016. A hybrid land use regression/line-source dispersion model for predicting intra-urban NO2. Transp. Res. Part Transp. Environ. 43, 181–191. doi:10.1016/j.trd.2015.12.007

Middleton, D.R., 1999. Development of AEOLIUS for street canyon screening. Clean Air 29, 155–161.

Middleton, D.R., 1998a. Dispersion Modelling: A Guide for Local Authorities, Met Office Turbulence and Diffusion Note number 241. The Meteorologcal Office, Bracknell, Bercks.

Middleton, D.R., 1998b. A New Box Model to Forecast Urban Air Quality: Boxurb. Environ. Monit. Assess. 52, 315–335. doi:10.1023/A:1005817202196

Mukerjee, S., Smith, L., Neas, L., Norris, G., 2012. Evaluation of Land Use Regression Models for Nitrogen Dioxide and Benzene in Four US Cities. Sci. World J. 2012, e865150. doi:10.1100/2012/865150

Nyberg, F., Gustavsson, P., Järup, L., Bellander, T., Berglind, N., Jakobsson, R., Pershagen, G., 2000. Urban air pollution and lung cancer in Stockholm. Epidemiol. Camb. Mass 11, 487–495.

Patton, A.P., Zamore, W., Naumova, E.N., Levy, J.I., Brugge, D., Durant, J.L., 2015. Transferability and Generalizability of Regression Models of Ultrafine Particles in Urban Neighborhoods in the Boston Area. Environ. Sci. Technol. 49, 6051–6060. doi:10.1021/es5061676

Spinelle, L., Gerboles, M., Villani, M.G., Aleixandre, M., Bonavitacola, F., 2017. Field calibration of a cluster of low-cost commercially available sensors for air quality monitoring. Part B: NO, CO and CO2. Sens. Actuators B Chem. 238, 706–715. doi:10.1016/j.snb.2016.07.036

Spinelle, L., Gerboles, M., Villani, M.G., Aleixandre, M., Bonavitacola, F., 2015. Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide. Sens. Actuators B Chem. 215, 249–257. doi:10.1016/j.snb.2015.03.031

Su, J.G., Brauer, M., Ainslie, B., Steyn, D., Larson, T., Buzzelli, M., 2008a. An innovative land use regression model incorporating meteorology for exposure analysis. Sci. Total Environ. 390, 520–529. doi:10.1016/j.scitotenv.2007.10.032

Su, J.G., Brauer, M., Buzzelli, M., 2008b. Estimating urban morphometry at the neighborhood scale for improvement in modeling long-term average air pollution concentrations. Atmos. Environ. 42, 7884–7893. doi:10.1016/j.atmosenv.2008.07.023

Tan, Y., Dallmann, T.R., Robinson, A.L., Presto, A.A., 2016. Application of plume analysis to build land use regression models from mobile sampling to improve model transferability. Atmos. Environ. 134, 51–60. doi:10.1016/j.atmosenv.2016.03.032

Tang, R., Blangiardo, M., Gulliver, J., 2013. Using Building Heights and Street Configuration to Enhance Intraurban PM10, NOX, and NO2 Land Use Regression Models. Environ. Sci. Technol. 47, 11643–11650. doi:10.1021/es402156g

Targa, J., Loader, A., 2008. Diffusion Tubes for Ambient NO2 Monitoring: A Practical Guide (No. AEA/ENV/R/2504 – Issue 1a). AEA Energy and Environment.

Tejada, J., Punzalan, J., 2012. On the Misuse of Slovin's Formula. Philipp. Stat. 61, 129–136.

U.S. Environmental Protection Agency, 2017. 40 CFR Part 51 Revisions to the Guideline on Air Quality Models: Enhancements to the AERMOD Dispersion Modelling System and Incorporation of Approaches To Address Ozone and Fine Particulate Matter, EPA-HQ-OAR-2015-0310; FRL-9956-23-OAR.

U.S. Environmental Protection Agency, 2015. Transportation Conformity Guidance for Quantitative Hot-spot Analysis in PM2.5 and PM10 Nonattainment and Maintenance Areas (No. EPA-420-B-15-084).

Vardoulakis, S., Fisher, B.E.A., Pericleous, K., Gonzalez-Flesca, N., 2003. Modelling air quality in street canyons: a review. Atmos. Environ. 37, 155–182. doi:10.1016/S1352-2310(02)00857-9

Wang, M., Beelen, R., Basagaña, X., Becker, T., Cesaroni, G., de Hoogh, K., Dėdelė, A., Declercq, C., Dimakopoulou, K., Eeftens, M., Forastiere, F., Galassi, C., Gražulevičienė, R., Hoffmann, B., Heinrich, J., Iakovides, M., Künzli, N., Korek, M., Lindley, S., Mölter, A., Mosler, G., Madsen, C., Nieuwenhuijsen, M., Phuleria, H., Pedeli, X., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Stempfelet, M., Tsai, M.-Y., Lanki, T., Udvardy, O., Varró, M.J., Wolf, K., Weinmayr, G., Yli-Tuomi, T., Hoek, G., Brunekreef, B., 2013. Evaluation of Land Use Regression Models for NO2 and Particulate Matter in 20 European Study Areas: The ESCAPE Project. Environ. Sci. Technol. 47, 4357–4364. doi:10.1021/es305129t

Williams, M., Barrowcliffe, R., Laxen, D., Monks, P., 2011. Review of Air Quality modelling in Defra. Defra.

Wilton, D., Szpiro, A., Gould, T., Larson, T., 2010. Improving spatial concentration estimates for nitrogen oxides using a hybrid meteorological dispersion/land use regression model in Los Angeles, CA and Seattle, WA. Sci. Total Environ. 408, 1120–1130. doi:10.1016/j.scitotenv.2009.11.033

World Health Organization, 2016. Ambient air pollution: A global assessment of exposure and burden of disease. Geneva, Switzerland.

World Health Organization, 2013. Review of evidence on health aspects of air pollution – REVIHAAP Project: Technical Report. Copenhagen, Denmark.

Yamane, T., 1967. Statistics, An Introductory Analysis, 2nd Edition. ed. Harper and Row, New York.

Zaksek, K., Ostir, K., Kokalj, Z., 2011. Sky-View Factor as a Relief Visualization Technique. Remote Sens. 3, 398–415. doi:10.3390/rs3020398

Zhu, G., Zhang, P., Tshukudu, T., Yin, J., Fan, G., Zheng, X., 2015. Forecasting traffic-related nitrogen oxides within a street canyon by combining a genetic algorithm-back

propagation artificial neural network and parametric models. Atmospheric Pollut. Res. 6, 1087–1097. doi:10.1016/j.apr.2015.06.006

Appendix A: Supplementary Information

Evaluation of the RapidAir dispersion model in London and the use of geospatial surrogates to represent street canyon effects

Nicola Masey[1], *Scott Hamilton*[2*], *Iain Beverland*[1]

[1] Department of Civil and Environmental Engineering, University of Strathclyde, 75 Montrose Street, Glasgow, UK, G1 1XJ

[2]Ricardo Energy & Environment, 18 Blythswood Square, Glasgow, UK, G2 4BG

*CORRESPONDING AUTHOR: Dr Scott Hamilton, Ricardo Energy & Environment, 18 Blythswood Square, Glasgow, UK, G2 4BG; email: scott.hamilton@ricardo.com; Tel: +44 (0)1235 753 716

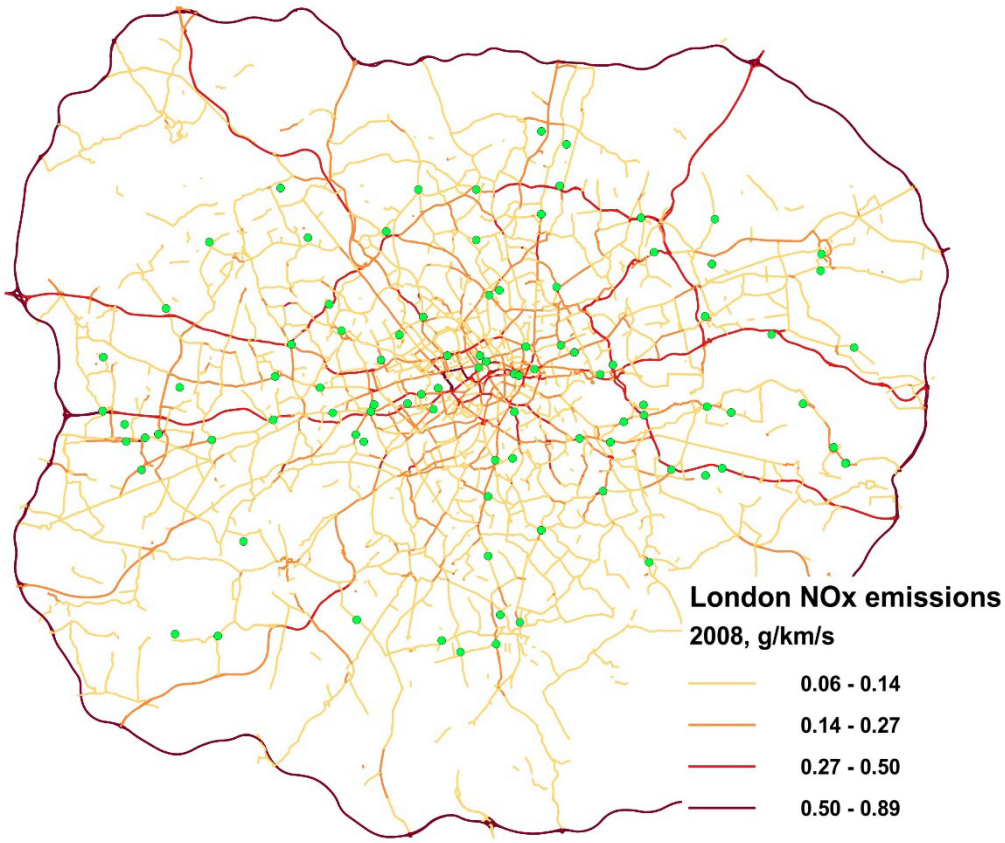


Figure A1: NO_x emissions for each road link in London, generated from LAEI emissions data. The points on the plot represent the location of the receptors (n = 107).

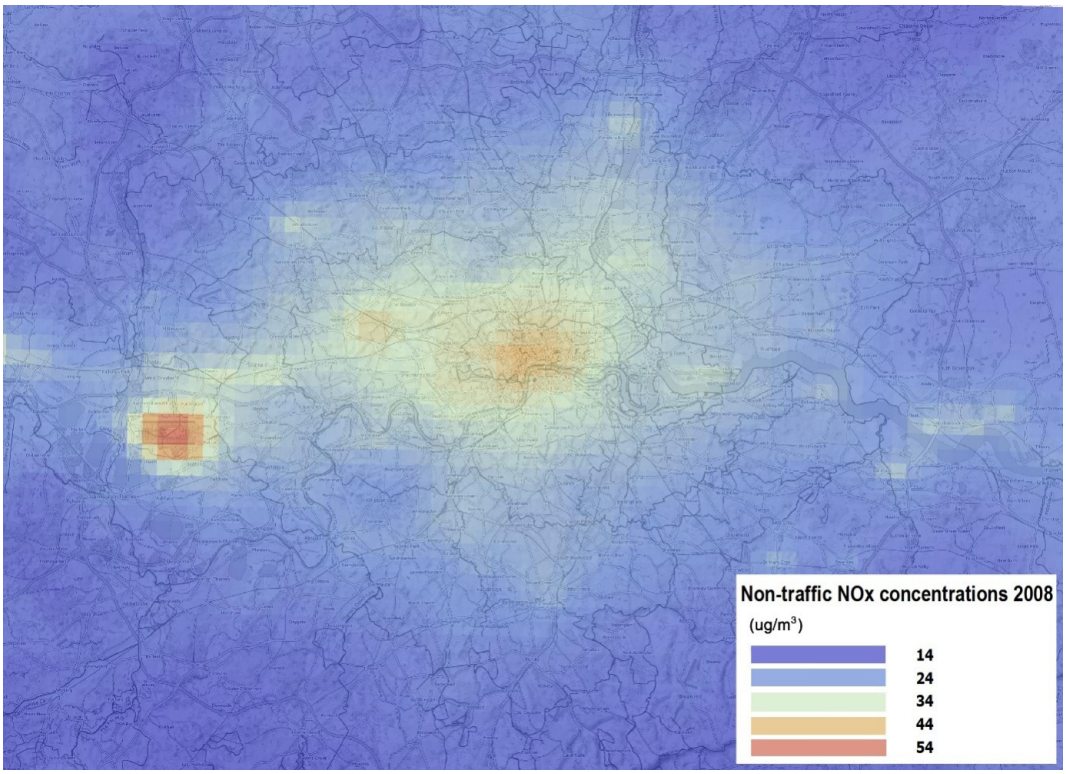
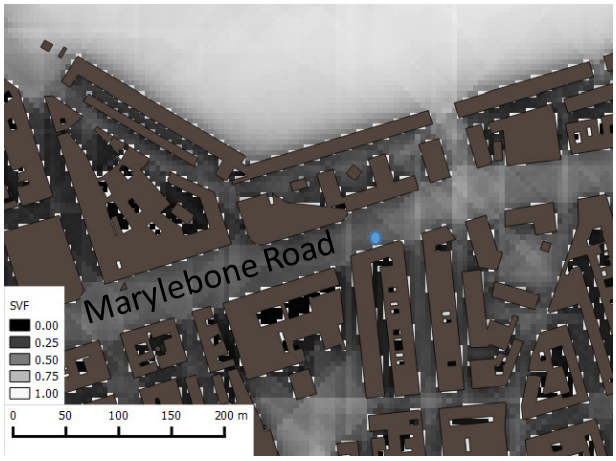
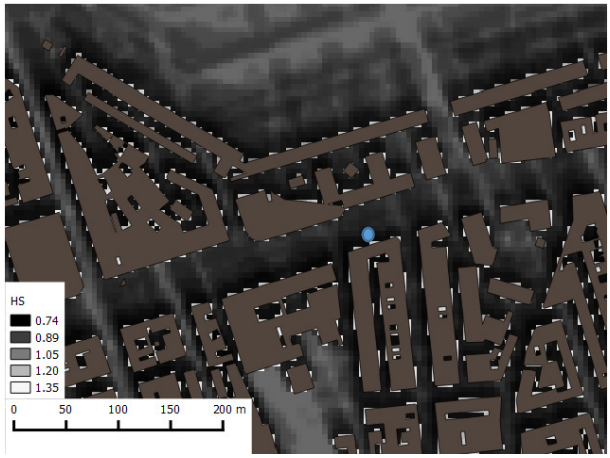


Figure A2: Annual average background concentrations of NO_x from non-road traffic sources in London in 2008 from the DEFRA Pollution Climate Mapping dataset. The main visible emissions source sectors are Heathrow Airport to the west and various domestic, commercial, industrial and other non-road traffic source types around the centre of the city.

(a)



(b)



(c)

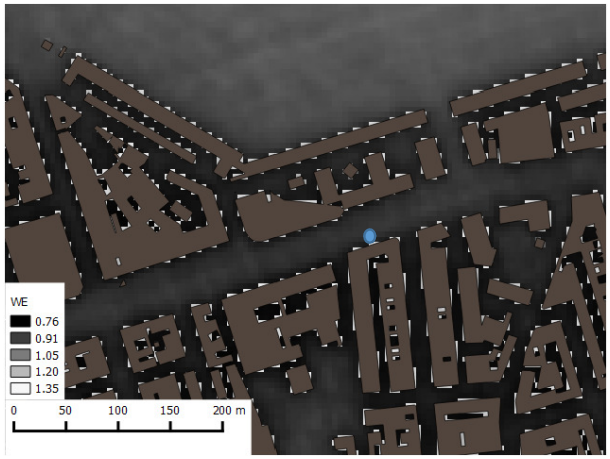
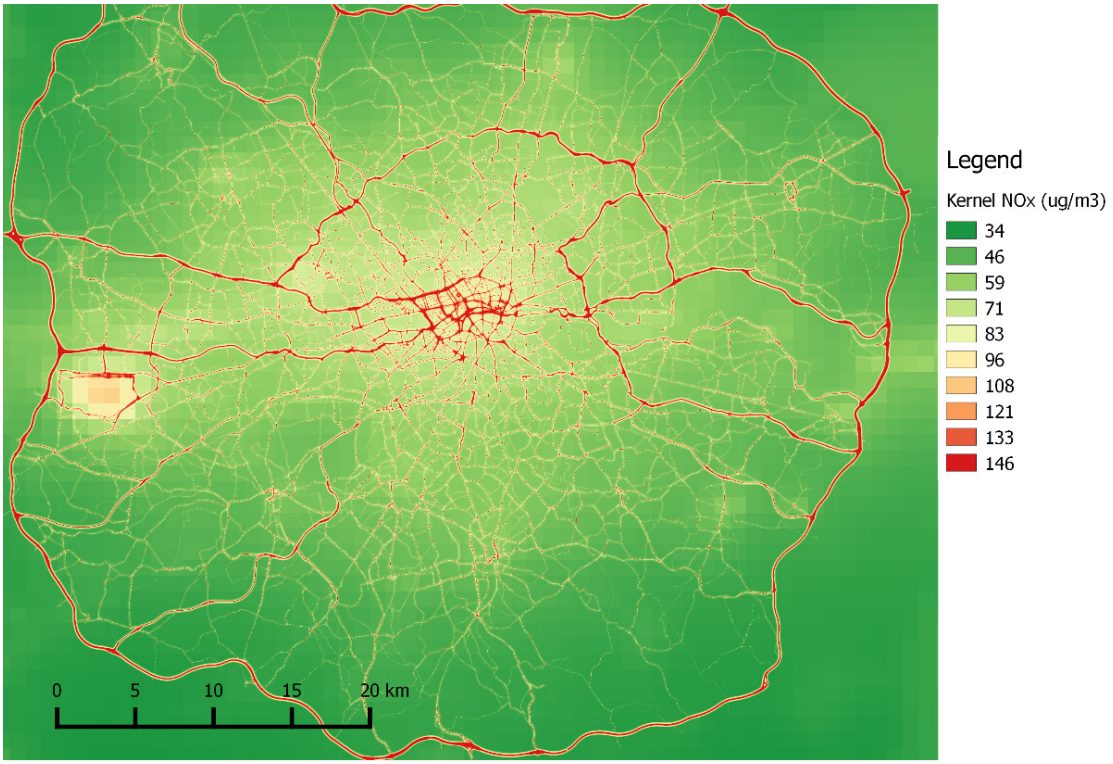


Figure A3: Geospatial surrogate values at Marylebone Road for (a) sky view factor (SVF), (b) hill shading (HS) and (c) wind effect (WE). The point shown represents the PDT monitoring site located at Marylebone Road.

(a)



(b)

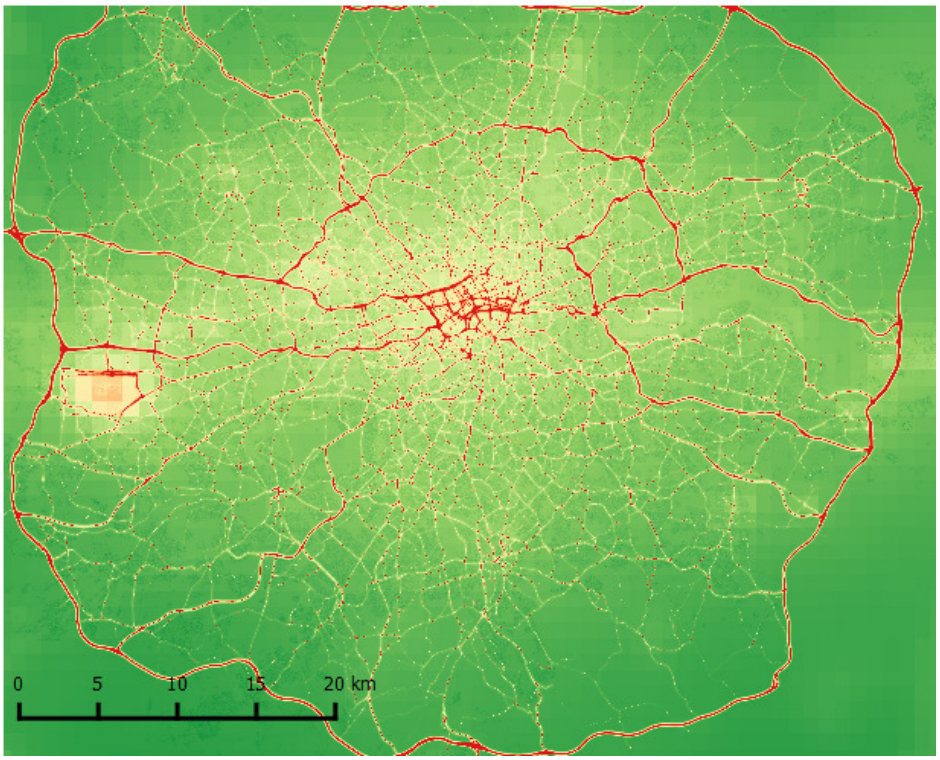
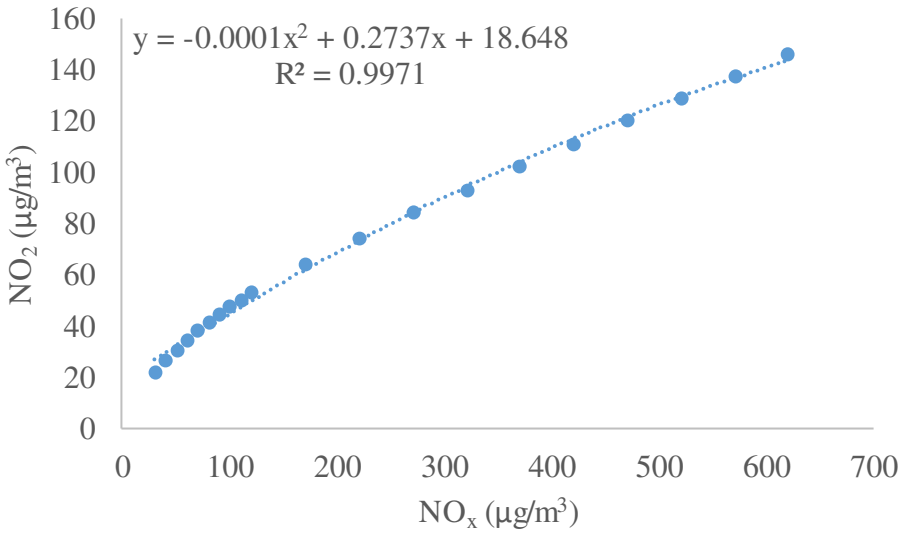


Figure A4: Concentrations of NO_x estimated using RapidAir over the London study area for (a) the Kernel model after correction for systematic biases; and (b) the kernel model + SVF correction, and then adjusted for systematic bias..

(a)



(b)

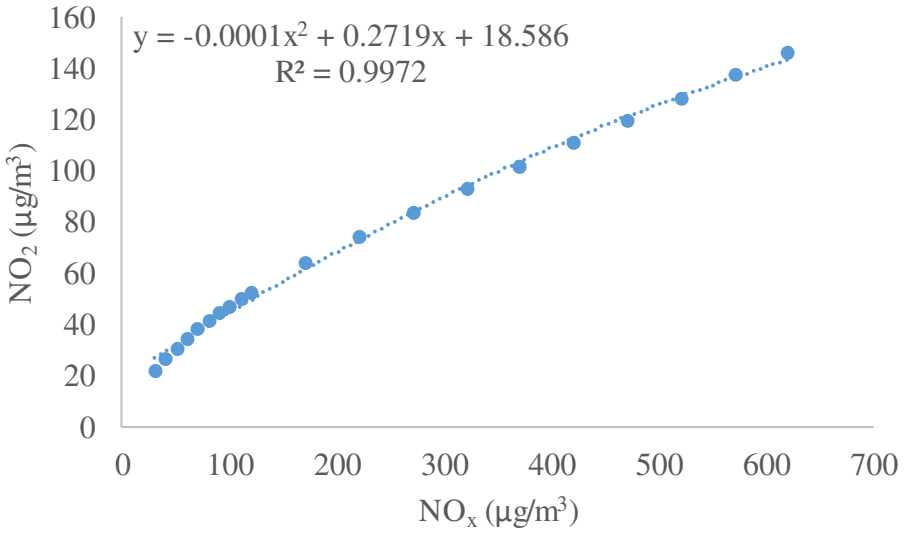


Figure A5: Regression between hypothetical NO_x and NO_2 concentrations based on the DEFRA NO_x to NO_2 calculator (version 3.2, 6 September 2012, available from https://laqm.defra.gov.uk/archive/archiveno-calculator.html). (a) shows the regression when the study district was set to Camden (estimated regional concentrations of NO_2 and O_3 29 and 47 µg/m^3 respectively) and (b) shows when the study district used was Bromley (estimated regional concentrations of NO_2 and O_3 17 and 55 µg/m^3 respectively).

AEOLIUS model parameterisation

The AEOLIUS model is a semi-empirical model to account for street canyon effects, and as such is more complex than the STREET model. Concentrations are calculated for the windward and leeward sides of the road using the equations detailed below (based on equations from Harris (2004). The leeward and windward concentrations described below are only calculated for streets that were perpendicular to the direction of the wind. Concentrations calculated in ppb, and were converted to µg/m^3 by multiplication by 1.91.

Inputs:

Emission rates (Q, µg/m/s); traffic speeds (v_t, mph), traffic density (f, vehicles per hour), % of cars and heavy good vehicles (f_c and f_h respectively), wind speed at roof level (u_r, m/s), street canyon width (w, m), street canyon height (h, m), and angle of street (θ).

Leeward concentrations:

The leeward concentrations = sum(C_{dlee} + C_{rec}) where C_{dlee} is the direct contribution from vehicles and C_{rec} is the pollution associated with recirculation.

Direct contribution (C_{dlee}):

$$Recirculationzone(l_r) = min\left(w, l_v * sin(\theta)\right) \quad \text{(meters)}$$

Where:

$$vortexlength(l_v) = 2 * r * h \quad \text{(meters)}$$

And r = wind speed dependence factor = 1 if $u_r > 2$ m/s and = $u_r/2$ otherwise.

If the recirculation zone is greater than the width of the canyon:

$$C_{dlee} = \sqrt{\frac{\frac{2}{\pi} * Q}{(w * \sigma_w)} * ln\left[\left(\frac{\sigma_w * w}{h_o * u_s}\right) + 1\right]}$$

Where:

σ_w = mechanical turbulence from wind and traffic (m/s) = $\sqrt{(\lambda * u_s)^2 + {\sigma_{wo}}^2}$

λ = constant for removal at the top of the canyon = 0.1

σ_{wo} = traffic-created turbulence (m/s) = $b * \sqrt{\frac{v_t * f_c * s_c + v_t * f_h * s_h}{w}}$

where s_c = mean surface area of cars (4 m^2), s_h = mean surface area of heavy vehicles (16 m^2) and b = aerodynamic constant (0.18)

u_s = wind speed at street level (m/s) = $u_r\left(\frac{ln\left(\frac{h_o}{z_o}\right)}{ln\left(\frac{h}{zo}\right)}\right)\left(1 - d * sin(\theta)\right)$

h_o = effective height of emissions (2 m)

z_o = effective roughness length (0.6 m)

d = model dependence (0.45)

If the recirculation zone is less than the width of the canyon:

$$C_{dlee} = \sqrt{\frac{2}{\pi}}\frac{Q}{(w * \sigma_w)}\left[ln\left[\left(\frac{\sigma_w * d_1}{h_o * u_s}\right)+1\right] + R * ln\left(\frac{h_o + \frac{\sigma_w * d_6}{u_s}}{\frac{\sigma_w * l_r}{u_s} + h_o}\right) + \frac{\sigma_w}{\omega_t}\left[1 - e^{\left(\frac{-\omega_t d_7}{u_s h}\right)}\right]\right]$$

Where:

d_1 (m) = min(w, l_r)

R = max(0, C_{ang})

C_{ang} = cos(2*r* θ)

d_6 (m)= min(max(l_{max}, l_r), x_1)

l_{max} = w/sin(θ)

x_1 = vertical distance (m) at which pollutants can escape canyon = $\frac{u_s(h-h_o)}{\sigma_w}$

ω_t = removal at top of the canyon (m/s) = $\sqrt{(\lambda * u_r)^2 + 0.4(\sigma_{wo})^2}$

d_7 (m) = max(l_{max}, x_1)-x_1

Recirculation contribution (C_{rec}):

$$C_{lee} = \frac{\left[\left(\frac{Q}{w}\right)d_1\right]}{\omega_t * d_2 + \omega_s * d_3}$$

Where

d_2 (m) = min(w, 0.5*l_r)

d_3 (m) = $max\left(0, \frac{2w}{l_r} - 1\right)$ l_s

l_s (m) = $\sqrt{(0.5 * l_r)^2 + h^2}$

ω_s = removal speed at the side of the canyon (m/s) = $\sqrt{{u_s}^2 + {\sigma_{wo}}^2}$

<u>Windward concentrations (C_{dwind})</u>:

Final windward concentrations = C_{dwind} + C_{rec}. C_{dwind} = 0 if $l_r \geq$ w, else:

$$C_{dwind} = \sqrt{\frac{2}{\pi}}\frac{Q}{w * \sigma_w}\left[ln\left(\frac{\sigma_w + d_4}{u_s + h_o} + 1\right) + \frac{\sigma_w}{\omega_t}\left[1 - e^{\left(\frac{-\omega_t d_5}{u_s h}\right)}\right]\right]$$

d_4 (m) = min[(w – l_r), x_1]

d_5 (m) = [max[(w – l_r),x_1]]-x_1

Impact of using a subset of meteorology on street canyon model predictions

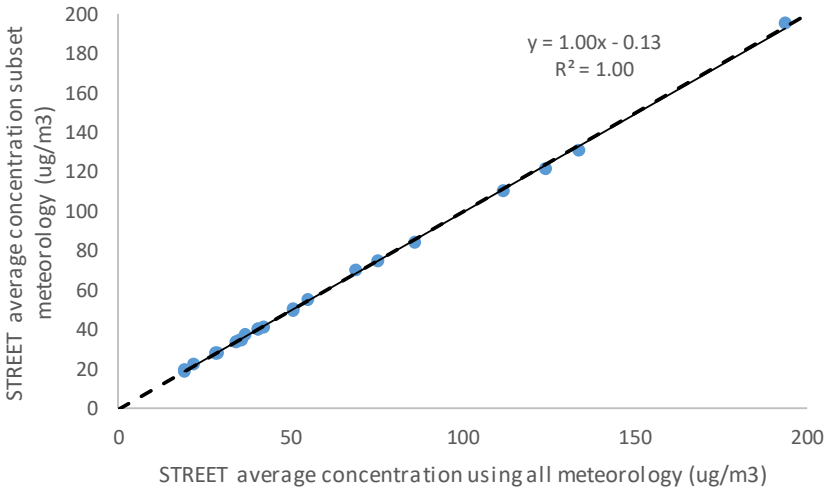
We used a subset of the annual hourly meteorology data in the two street canyon models to improve model run times, without detrimentally impacting the predicted concentrations (Israel, 1992; Tejada and Punzalan, 2012; Yamane, 1967). The number of meteorological hours required to provide a representative sample size (n) was calculated using:

$$n = \frac{N}{1 + N(e^2)}$$

Where N = the number of samples in the whole population (N = 8600 hours) and e is the level of precision (e = 0.05 for a 95 % confidence interval).

This leads to a representative sample size of 383 hours of meteorology data, which was randomly selected from the complete data set.

(a)



(b)

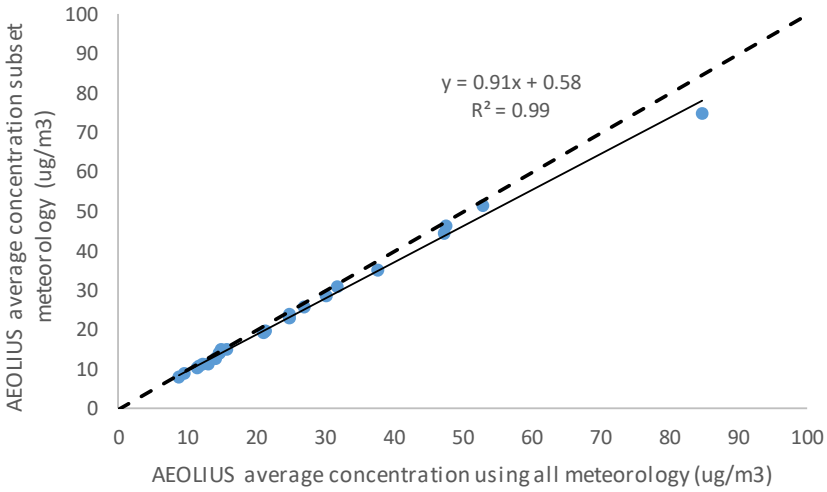


Figure A6: Scatter plots showing the agreement between (a) STREET NO_x and (b) AEOLIUS estimates when a subset of meteorological data was used (y-axis) compared to the whole meteorological data (x-axis).

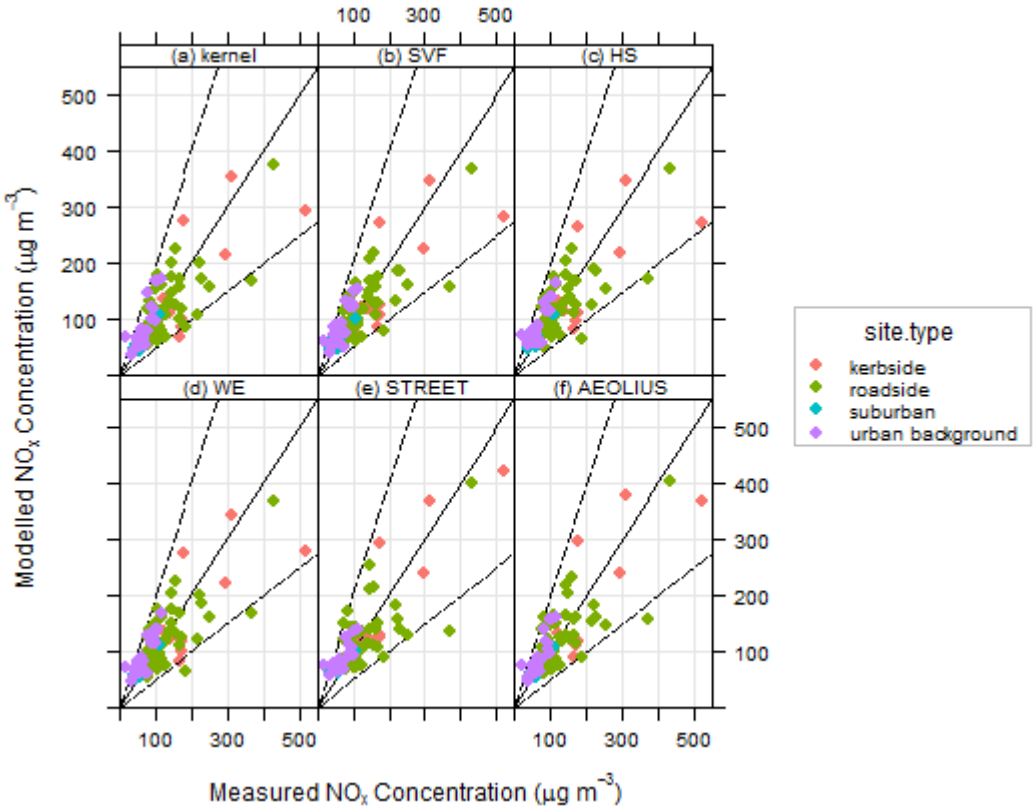


Figure A7: Scatter plots showing the NO_x concentrations measured by the passive diffusion tubes and those estimated by the kernel model: (a) concentrations from the kernel model; and the raw kernel model after correction using the surrogates for street canyons: (b) sky view factor (SVF), (c) hill shading (HS), (d) wind effect (WE), (e) STREET canyon model and (f) AEOLIUS canyon model. The solid line represents 1:1, while the dashed lines represent FAC2 values.

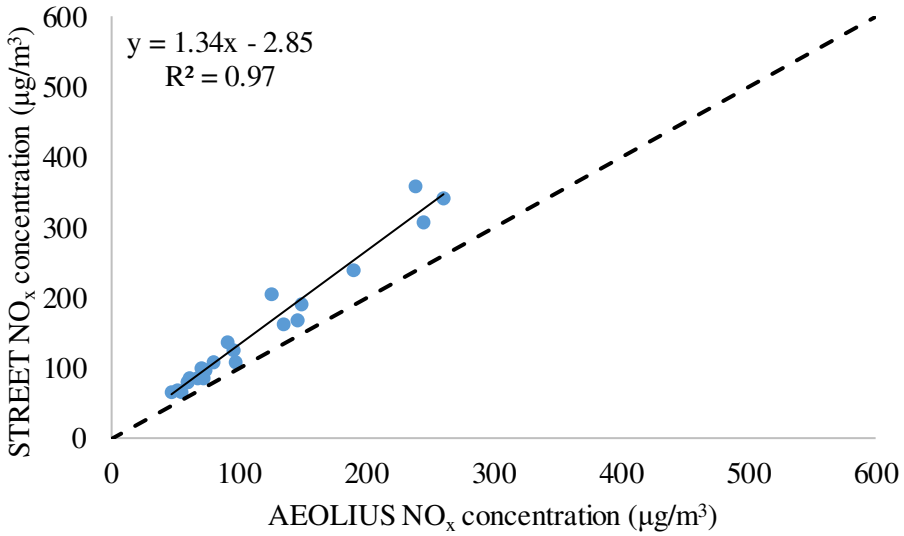


Figure A8: Scatter plot showing a comparison of NO_x concentrations predicted in the street canyon locations by the STREET model and the AEOLIUS model.

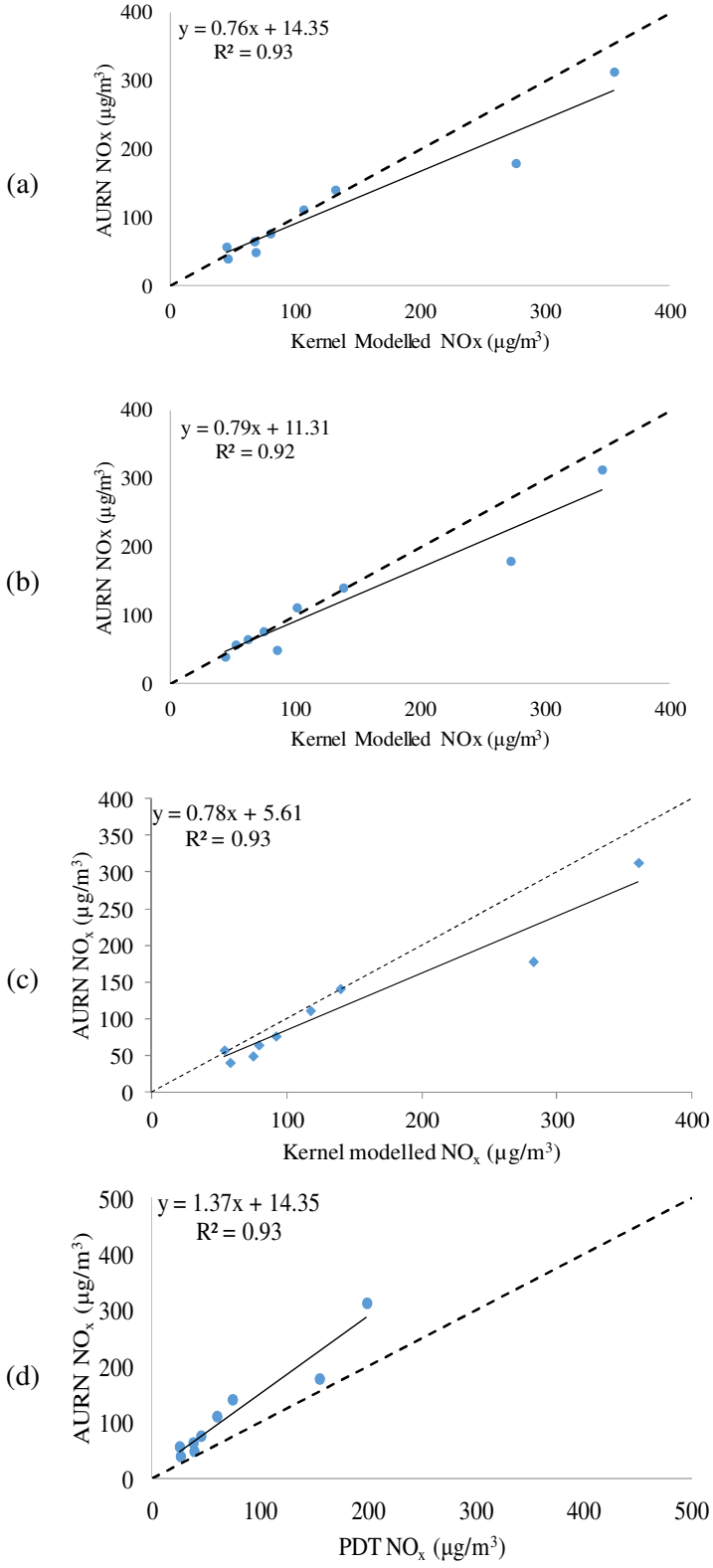


Figure A9: Scatter plots showing NO_x concentrations measured by the AURN stations against concentrations predicted by (a) the adjusted kernel model; (b) SVF kernel; (c) STREET kernel; and (d) co-located PDTs.

Example use case for the RapidAir model- estimating number of buildings in London with levels of NO_2 above the annual mean standard

The speed of the modelling systems allows very efficient decision support analyses. For instance, we can use GIS based zonal statistics on the generated NO_2 concentration surface to determine the average for all buildings, and then the number of buildings in London which exceeded the annual mean NO_2 standard of 40 µg/m^3 in 2008.

The building height layer used in the analysis contains 764,308 building footprints as polygons in Esri Shapefile format. Using the zonal statistics (available from https://docs.qgis.org/2.2/en/docs/user_manual/plugins/plugins_zonal_statistics.html) plugin in the QGIS software we compute the maximum, minimum and average NO_2 concentration from the RapidAir dispersion model for concentration cells which lie within each building polygon (Figure S10). This step is again very computationally efficient and only adds about 5 minutes onto the analysis time. In this instance we use the NO_2 concentration rasters from the linearly adjusted RapidAir model though the process is exactly the same for all of the model outputs. We used the pandas library in python 2.7 to calculate the from the total number of buildings. The average concentration across every building in the city was 33 µg/m^3 (Figure S11).

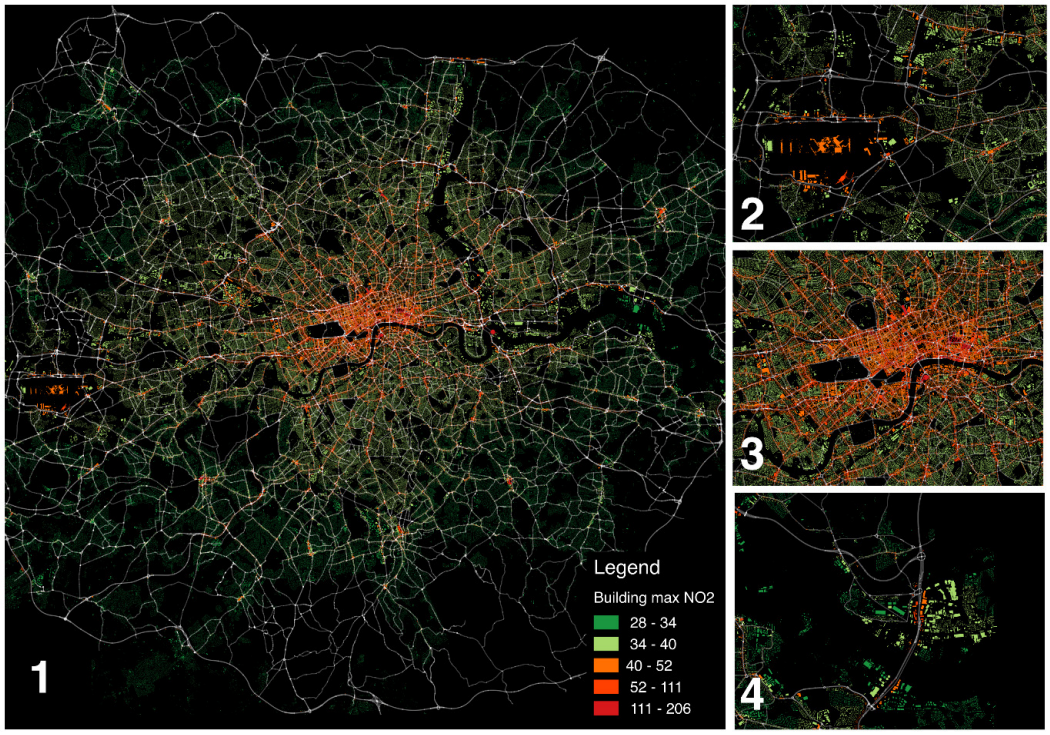


Figure A10: Maximum modelled NO_2 values (µg/m^3) in building footprints in London in 2008. Buildings with orange or red symbology are > 40 µg/m^3 annual mean standard. Box 1 shows the whole city model; 2 shows the area around Heathrow Airport; 3 is the city centre; and 4 is around the Dartford Crossing. Most buildings exceeding the standard are in the city centre with more isolated pockets of higher concentrations around key roads and the airport.

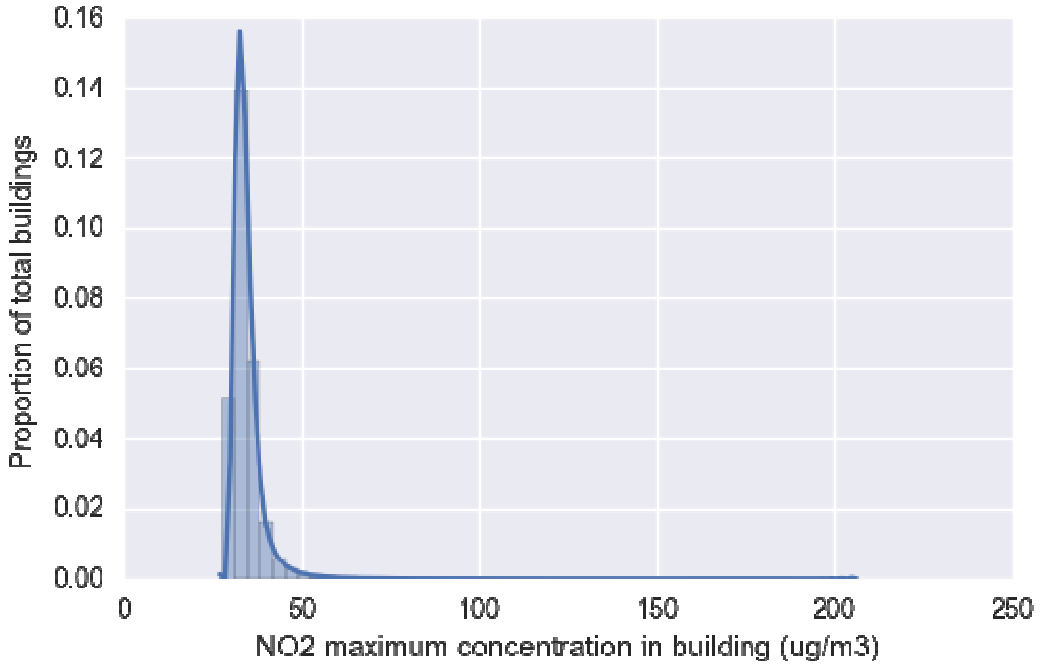


Figure A11: Frequency distribution of maximum annual average (NO_2) concentrations at >700,000 buildings in London. The vast of majority of buildings in the city are below the annual average standard of 40 µg/m^3.

Table A1: Details of receptor used for model evaluation. The type of receptor (PDT or AURN), name of site, site type, coordinates and measured NO_x as NO_2 are provided for each site. The PDT sites marked with an asterisk (*) were used as test data (i.e. they were not used to derive the surrogate correction equations). ^{SC} denotes PDT sites that were identified as street canyon locations.

Receptor	Site name	Site type	Easting	Northing	NO_x as NO_2 ($\mu g/m^3$)
AURN	Camden Kerbside	urban traffic	526633	184390	178
	Haringley Roadside	urban traffic	533894	190707	76
	London Bexley	suburban	551859	176381	57
	London Eltham	suburban	543981	174655	40
	London Harlington	urban industrial	508295	177800	64
	London Hillingdon	urban background	506941	178610	111
	London Marylebone	urban traffic	528126	182015	313
	London N Kensington	urban background	524045	181749	49
	Tower Hamlets	urban traffic	535927	182218	141
PDT	BG1*	suburban	551053	187233	41
	BG2	suburban	548043	183320	57
	BG3	kerbside	543955	184432	102
	BL0	urban background	530123	182014	94
	BN1^{SC}	kerbside	526342	192223	139
	BN2*	urban background	524370	189640	54
	BT1*	suburban	519560	189271	52
	BT4*	roadside	520866	185169	253
	BT6*	roadside	521619	183554	104
	BT7	urban background	525173	183297	53
	BX1	suburban	551864	176379	57
	BX2	suburban	549980	179064	49
	BX7	roadside	552615	175416	106
	BX8*	roadside	552566	175384	84
	BY7	roadside	540518	169324	81
	CD1*^{SC}	kerbside	526629	184391	177
	CD3*^{SC}	roadside	530057	181285	169
	CD4	urban background	530511	181665	104
	CD5	urban background	530511	181665	101
	CR2	roadside	531123	164299	127
	CR4*^{SC}	roadside	532583	165636	95
	CR5*^{SC}	kerbside	530626	169707	173
	CR6	suburban	531369	166096	59
	CT1	urban background	532235	180892	80
	CT3*	urban background	533480	181186	91
	CT6^{SC}	roadside	532527	180789	431
	CY1*	roadside	533901	171290	112
	EA1	urban background	517541	180738	67
	EA2*^{SC}	roadside	520304	180054	127
	EA6	roadside	518537	182708	370
	EA7*	urban background	511677	180071	52
	EL1	urban background	511403	164915	36
	EL2^{SC}	roadside	514024	164792	92

EN1	suburban	533900	195800	47
EN3	urban background	535440	195000	46
EN4	roadside	535056	192470	98
EN5	roadside	529894	192223	171
GB6	roadside	544997	175098	109
GN0*	roadside	544084	178881	101
GN2	urban background	540169	178999	73
GN3*[SC]	roadside	545560	178526	90
GR4	suburban	543978	174655	40
GR5*	roadside	538960	177954	101
GR7	roadside	538141	176710	112
GR8*	roadside	540200	178367	223
GR9*	roadside	541879	175016	101
HF1	roadside	523420	178590	186
HF2*	urban background	523625	179010	61
HG1[SC]	roadside	533891	190707	75
HG2	urban background	529894	189125	49
HI0	suburban	506945	178609	110
HI1	roadside	510835	184916	104
HI2	roadside	506990	181925	72
HI3	roadside	509551	176974	84
HK4	urban background	534830	186234	89
HK6[SC]	roadside	532947	182575	145
HR1	urban background	517877	192314	39
HR2[SC]	roadside	513504	188998	111
HS2*	suburban	510371	177198	66
HS4*[SC]	roadside	521083	178501	165
HS5*	roadside	517423	178070	145
HS6	roadside	513653	176842	130
HS7	urban background	509332	174997	63
HV1*	roadside	553110	182516	80
HV3[SC]	roadside	551105	188261	76
IS2[SC]	roadside	530698	185735	158
IS6*	urban background	531325	186032	58
KC1	urban background	524045	181752	49
KC2	roadside	526524	178965	156
KC3	roadside	527518	179395	229
KC4*[SC]	roadside	527268	178089	217
KC5[SC]	kerbside	525695	178363	295
LB1	roadside	530628	173368	115
LB3	urban background	532137	175701	54
LB4*[SC]	kerbside	531070	175593	519
LH0	urban background	508300	177800	19
LH2	urban background	508393	176742	115
LW1	urban background	537675	173689	101
LW2*[SC]	roadside	536241	176932	148
MY1[SC]	kerbside	528125	182016	312
RB1*	urban background	544377	187647	57
RB3	kerbside	544555	190402	123
RB4*	roadside	540823	188369	95
RB5	roadside	540017	190488	115
RI1[SC]	roadside	522498	177165	83
RI2	suburban	522989	176729	43

SK1	urban background	532240	178561	78
ST3	suburban	527776	164513	51
ST4*[SC]	kerbside	528925	163804	175
ST6[SC]	kerbside	522557	165787	167
TD0	suburban	515115	170778	38
TH1	urban background	537509	180867	56
TH2	roadside	535927	182221	141
TH3	urban background	535100	182664	57
TH4*	roadside	538290	181452	166

Table A2: Summary model evaluation statistics for annual mean NO_x at the test receptor at (a) PDT locations (all sites and split by site type) and (b) AURN stations. Statistics are given for the Kernel model, the kernel model after correction using the surrogates for street canyons and using the street canyon models.

Site type	Model	n	FAC2	MB ($\mu g/m^3$)	NMB	RMSE ($\mu g/m^3$)	r	COE	IOA
(a) PDT All	Kernel	95	0.95	-8.13	-0.07	50.16	0.80	0.41	0.71
	SVF	95	0.97	-3.84	-0.03	48.26	0.81	0.43	0.71
	HS	95	0.97	-4.33	-0.04	49.46	0.80	0.41	0.71
	WE	95	0.96	-3.81	-0.03	49.14	0.80	0.42	0.71
	STREET	95	0.96	0.00	0.00	46.54	0.82	0.41	0.71
	AEOLIUS	95	0.97	0.00	0.00	46.37	0.82	0.42	0.71
PDT Kerbside	Kernel	10	0.80	-46.06	-0.21	96.54	0.71	0.18	0.58
	SVF	10	1.00	-41.46	-0.19	91.68	0.73	0.29	0.64
	HS	10	0.90	-43.13	-0.20	95.49	0.70	0.24	0.62
	WE	10	1.00	-45.65	-0.21	97.17	0.70	0.23	0.61
	STREET	10	1.00	-17.74	-0.08	61.15	0.88	0.45	0.73
	AEOLIUS	10	1.00	-24.43	-0.11	76.41	0.80	0.33	0.66
PDT Roadside	Kernel	45	0.96	-16.50	-0.12	52.20	0.73	0.24	0.62
	SVF	45	0.96	-13.82	-0.10	50.32	0.74	0.32	0.66
	HS	45	0.96	-12.76	-0.09	51.26	0.72	0.29	0.65
	WE	45	0.96	-13.24	-0.10	51.16	0.73	0.28	0.64
	STREET	45	0.93	-11.34	-0.08	57.79	0.63	0.19	0.59
	AEOLIUS	45	0.96	-9.23	-0.07	52.53	0.70	0.25	0.63
PDT Suburban	Kernel	13	1.00	-1.52	-0.03	7.30	0.92	0.47	0.73
	SVF	13	1.00	4.27	0.08	9.74	0.88	0.25	0.62
	HS	13	1.00	6.75	0.12	9.63	0.93	0.34	0.67
	WE	13	1.00	6.63	0.12	10.32	0.90	0.26	0.63
	STREET	13	1.00	13.08	0.24	16.32	0.92	-0.26	0.37
	AEOLIUS	13	1.00	7.91	0.14	11.09	0.92	0.23	0.61
PDT Urban background	Kernel	27	0.96	16.69	0.25	28.91	0.84	0.00	0.50
	SVF	27	0.96	22.83	0.34	28.55	0.87	-0.27	0.36
	HS	27	0.96	20.59	0.31	26.33	0.86	-0.18	0.41
	WE	27	0.96	20.56	0.31	26.39	0.84	-0.19	0.41
	STREET	27	0.96	19.17	0.29	23.25	0.84	-0.04	0.48
	AEOLIUS	27	0.96	20.62	0.31	27.30	0.84	-0.12	0.44
(b)	Kernel	9	1.00	16.54	0.14	36.95	0.96	0.65	0.83
	SVF	9	1.00	16.26	0.14	35.59	0.96	0.68	0.84
	HS	9	1.00	15.16	0.13	34.10	0.96	0.64	0.82
	WE	9	1.00	17.42	0.15	36.15	0.96	0.65	0.82
	STREET	9	1.00	27.09	0.24	47.25	0.94	0.45	0.72
	AEOLIUS	9	1.00	27.49	0.24	47.83	0.95	0.50	0.75
(c)	Kernel	23	0.91	-23.89	-0.13	73.82	0.77	0.27	0.63
	SVF	23	1.00	-16.87	-0.09	69.21	0.79	0.37	0.68
	HS	23	1.00	-20.75	-0.12	73.15	0.77	0.31	0.66
	WE	23	0.96	-19.82	-0.11	72.51	0.77	0.32	0.66
	STREET	23	0.96	12.21	0.07	58.22	0.85	0.34	0.67
	AEOLIUS	23	1.00	2.06	0.01	62.77	0.82	0.33	0.66

8. Conclusions

As a result of detrimental impacts associated with exposure to air pollution, including cardiovascular and respiratory illness, ambient concentrations that members of the public are exposed to need to be assessed. Such assessment involves monitoring or modelling studies, with the former providing measures of ambient concentrations at a limited number of locations, while the latter provides estimates at multiple locations within a study area. It is important that the pollution estimates from these techniques are accurate to avoid exposure misclassification, which may lead to underestimation of health effects associated with actual exposures.

Monitoring by government and local authorities in the UK involves static real-time monitors supplemented by Palmes passive sampling devices. Typically Palmes samplers provide an average concentration over a 4-5 week period, which is appropriate to determine annual average concentrations and thus can be used to estimate longer-term exposures. However, short-duration exposures to pollution can also impact human health. Therefore I deployed Palmes passive samplers for 2-3 day exposures and demonstrated that the precision and accuracy of the samplers were not affected compared with samplers exposed for longer (1 week) durations. This opens up the possibility of using passive samplers for short duration personal monitoring including periods when people are mobile, thus providing more accurate estimates of exposure for those individuals. This can be used by epidemiologists to identify locations or individuals with higher exposures and provide more evidence describing the relationships between exposure and health. Alternatively, the shorter exposure times can be used to rapidly measure a large number of sites within a study area to provide information about the spatial gradients in pollution concentrations. However, I observed that the Palmes samplers overestimated ambient concentrations as a result of wind-speed effects causing turbulence within the diffusion tube, consequently reducing the effective tube length and hence increasing the uptake rate. I established that the relationship between uptake rate and wind-speed could be used to post-process the measurements to minimise this effect. However, this approach relied on wind-speed estimates at the diffusion tube location (which are generally not available for large networks of sites) and would be unlikely to be applicable for personal monitoring when turbulence could be different during different activities e.g. walking *vs.* indoors. Modification of the design of the Palmes sampler, through use of a membrane over the open end of the passive sampler reduced the turbulent effects associated with wind-speed, provided more accurate estimates of ambient concentrations for both the exposed urban background site and less exposed roadside location. Therefore I suggest that

Palmes samplers could be deployed with a membrane over the open end during measurements to minimise wind-speed effects and consequently improve the accuracy of the sampler measurements to measure ambient concentrations under all static pollution measurements. As part of these studies I also deployed badge type Ogawa samplers, whose shorter path length and faster uptake rates make them more suited to shorter exposure durations. After correction of the Palmes sampler for wind-speed effects, the two sampler types produced similar estimates of pollution and therefore there is no necessity to change the regulatory diffusion tubes to badge samplers if shorter duration exposures are required. The passive sampler research was carried out at two monitoring locations in Glasgow that were anticipated to have different pollution concentrations (urban background vs. roadside) and experience different wind-speeds at the site (higher wind-speeds anticipated at the more open urban background locations). The correction of uptake rate for wind-speed had a more positive effect at the urban background location than the roadside site, however this may have been a result of inaccurate estimates of wind-speed at the roadside location (wind-speed measurements at the urban background site were used). Future work should aim to make concurrent wind-speed and passive sampler measurements at the roadside location to determine the 'true' wind-speed at these sites. These results may show that when a site experiences low wind-speeds the correction processes applied may not be necessary.

Recently, portable real-time sensors have been developed to improve upon the temporal limitation of the diffusion samplers. These real-time sensors provide fine temporally resolved (generally 1-minute) concentration estimates while their small size enables use in personal monitoring campaigns. The accuracy of many if not all portable sensors is relatively unknown as a result of limited field-evaluation (in contrast to government real-time reference analysers, which go through rigorous quality assurance and quality control procedures). I demonstrated that calibration of these sensors under representative field conditions is essential in order to obtain accurate pollution estimates at an urban background site. For example, Aeroqual NO_2 sensors required correction for cross-sensitivity to O_3. I also identified that the most accurate pollution estimates from Aeroqual sensors were obtained when frequent, intermittent calibrations were used; and that estimated concentrations from a single calibration at the start of a monitoring campaign were subject to larger errors as a result of instrument drift over the study period. I also evaluated AE51black carbon aethalometers at the same site and found deterioration in accuracy when the attenuation of the collection filter was higher than 40. Application of reported correction algorithms designed to account for darkening AE51 filters did not improve the sensor accuracy above these ATN values in my measurements. In highly polluted environments a

filter attenuation of 40 could be reached quickly and my observations suggest that filters should be changed regularly to ensure black carbon concentrations are estimated accurately. The above calibration measurements were made at a single urban background site that is unlikely to be representative of conditions in more polluted locations. Correspondingly it is not known how well these calibrations represent instrument responses over the wider range of pollutant concentrations observed in urban environments. Future research would benefit from access to a reference site in a more polluted location. If inaccurate calibration equations or procedures are used, the real-time monitors will not report accurate pollution concentrations. If these were to be used for personal monitoring in epidemiology studies, for example, this could lead to personal exposures being underestimated (resulting in individuals at risk of detrimental health effects being missed) or overestimated (resulting in a larger number of 'at-risk' individuals identified). Alternatively if these inaccurately calibrated sensors were used to measure ambient concentration this could lead to pressures on Local Authorities or Governments to improve air quality in the areas where concentrations were overestimated (leading to financial implications for the budgets of the areas in question), or areas with unacceptable air quality may be missed if the sensors underestimate the ambient concentrations.

Additionally, there may be other sources of error when these sensors are used in personal mobile monitoring, for example the impact of turbulence or rapidly changing concentrations is unknown and in practice would be difficult to quantify. To gain a more general and practical understanding of the scale of such errors I used the real-time monitors to make short (6-min) static measurements at sites where I also made measurements of 1-week NO_2 concentrations using passive diffusion tubes (PDT). The short BC measurements followed the general spatial trends in weekly NO_2 concentrations, however the 6-min Aeroqual NO_2 and O_3 concentrations trends were less similar. The discrepancy in spatial pattern between Aeroqual and PDT measurements partly resulted from higher than anticipated NO_2 Aeroqual concentrations (probably associated with lower than anticipated O_3 Aeroqual concentrations) at the first study site, which may have resulted from moving the monitors from an indoor to an outdoor environment. The potential implications for use of the Aeroqual sensors in personal monitoring require further investigation, particularly if sensors require time to adjust to rapidly changing conditions (e.g. temperature and/or concentration). BC measurements appeared unaffected by the indoor-outdoor changes in my measurements, and hence appeared to be useful for provision of rapid estimation of pollution concentrations over a large study area through short duration peripatetic monitoring. In the absence of more detailed field evaluation, pollution estimates from real-time sensors are indicative only.

However, the real-time sensor concentrations could be used to help quickly build up a picture of relative spatial ranges in concentrations in a study area, and identify areas of elevated pollution concentrations that would benefit from more detailed monitoring. This could potentially save Local Authorities money on these costly, long-term, detailed measurement campaigns by ensuring these are only carried out in the required locations. The measurements were made repeatedly in a single urban area and the relationships between pollutants identified will be specific for this area, and may be transferable to other urban areas with similar sources of emissions. The methods described, however, could be readily applied to other study areas to build a rapid picture of the ambient concentrations for that area.

Pollution models can be used to estimate concentrations at multiple locations within a given study area. These normally use estimates of pollution emission and background concentrations to predict concentrations over the study area. To obtain more accurate estimates, meteorology should be taken into consideration which can take the form of measurements from local weather stations. Alternatively weather models can be used to predict local meteorology in the absence of a weather station, and to forecast future meteorological conditions. This allows pollution forecasts to be issued, which in turn can help individuals who are most at risk to health effects from air pollution (e.g. people with pre-existing respiratory conditions) to make informed changes to minimise the impact of elevated concentrations on their health. In my research, the initial user parameterisation of the Weather Research Forecasting model had limited effect on the pollution estimates from a dispersion model (AERMOD) for low level (10 m) sources. Consequently parameters can be selected to minimise run times and allow faster generation of weather data for the models. The optimal parameterisation of the weather model identified in this work is only applicable to low height emissions sources (such as roads) and should not be applied without further validation to taller emission sources (such as industrial stacks). Conventional dispersion models have high costs associated with them (including license, computational and user input costs) restricting the modelling of large study areas. Developments in computing have allowed improvements in computation time, and I evaluated a computationally efficient model for estimation of NO_2 concentrations at 5 x 5 m resolution over a 3,500 km^2 study area in less than 10 minutes. The accuracy of this model (compared to measurements made at 107 receptor sites) was similar to previously published model evaluation studies. These short run times and high spatial resolution concentration estimates open up possibilities to use models to rapidly assess proposed pollution mitigation schemes in areas with high pollution concentrations identified by the mobile monitoring techniques discussed above.

For example, the impact of changing a road layout or the introduction of a low emission zone could be quickly estimated. Inclusion of street canyon models improved the accuracy of model predictions at roadside and kerbside locations; however it is practically challenging to discretely model individual canyons within large study areas. Geospatial surrogates for urban morphology can be used to rapidly identify, and apply correction to, locations falling within canyon locations; however the normalised mean bias errors were lower when the canyons were discretely modelled. Further work could optimise geospatial surrogate parameters, which in turn may improve the accuracy of combined dispersion-surrogate model predictions to match the accuracy of combined dispersion and canyon models. A benefit of dispersion models is that these can be readily transferred between study areas providing the appropriate input data for the individual area is provided. The RapidAir model can be quickly parameterised for each unique study area to allow pollution estimates to be made quickly. In its current release, however, the RapidAir model may not be transferrable to countries whose main sources of emissions are not traffic-related (other predominant sources could include industry or natural sources such as particles from dust storms) – future work aims to further develop the model to allow incorporation of these additional sources to allow the model to be more widely applicable.

In summary, exposure of populations to air pollution can be assessed using measurement and modelling techniques. Newly-developed portable real-time sensors have the potential to provide detailed personal exposure estimates to complement spatial measurements from passive sampling devices, however more research into real-time sensor accuracy is required. The use of portable instruments to provide rapid estimates of spatial patterns in pollution concentrations can be used to help identify locations where more detailed monitoring should be carried out. Using this approach, areas of high pollution concentrations can be identified and mitigation strategies designed to reduce pollution concentrations. The development of rapid high resolution dispersion models will allow the impact of mitigation schemes to be estimated prior to their implementation and will help to identify the most effective and efficient methods to reduce ambient concentrations and population exposures.

Appendix A – Field evaluation of Little Environmental Observatory (LEO) monitors

Abstract:
Four Little Environmental Observatory (LEO) monitors were provided on loan to the University of Strathclyde from the Institute of Occupational Medicine (IOM) during August 2016. The instruments were evaluated by: a co-location study at an automatic monitoring station to assess accuracy of O_3 and NO_2 measurements; and a set of mobile measurements to measure O_3 and NO_2 exposure of an individual researcher in transects moving away from busy roads. In the co-location study we observed that the LEO sensors appear to have pronounced cross sensitivity to O_3 in an unexpected direction (higher observed LEO O_3 resulting in lower observed LEO NO_2). However (unlike our experience with Aeroqual sensors), corrected LEO NO_2 concentrations still had relatively low correlation with analyser NO_2 concentrations. In the transect study concentrations of NO_2 measured by the LEO instruments were lower than the Aeroqual instruments, including many negative concentrations observed by each of the LEO instruments. LEO O_3 concentrations were much higher than Aeroqual O_3 concentrations. In accordance with prior expectations Aeroqual instruments observed slightly lower O_3 concentrations at the road side sites. However this pattern was not apparent in the LEO measurements. Correlations between LEO and Aeroqual NO_2 and O_3measurements were low ($R^2 < 0.1$).

Collocation study at automatic monitoring station

The four LEO instruments were deployed on the roof of the Townhead automatic monitoring station (urban background site north of Glasgow city centre latitude/longitude**:** 55.865782, -4.243631). The instruments were suspended from the roof of a passively-ventilated enclosure to prevent water ingress damage while ensuring that the sensor inlets were not obstructed (Figure 2). Measurements were made by the sensors at 1-min intervals. The mobile phones paired with the LEO's to upload data to a server computer system were sealed in a waterproof box next to the instrument enclosure. LEO instruments were run off mains power available on the roof of the monitoring station. Instruments were deployed between 8-22 August 2016; with site visits approximately every 5 days to check on instrument operation. Hourly pollution data from the Townhead site was available from www.scottishairquality.co.uk and at the time this report was written the data was still provisional (the availability of ratified data is approximately 3 months). Additionally, we deployed a HOBOware temperature and relative humidity sensor at the site during the measurements to record these variables every 5 minutes. The HOBO meteorological

measurements and the LEO sensor measurements were averaged to produce hourly measurements.

Of the 4 LEO sensors deployed data from 2 sensors did not upload to the server so were unavailable for analysis. We are not clear as to the reason why this data was not available on the server as clicking on 'Verify Data' within the Expo App showed data being recorded by the LEO sensor. The remaining 2 sensors for which data was available did not record for the full duration of the deployment – upon returning to the site both the blue and orange lights on the front of the instruments were continuously lit and the sensors required to be reset. Similar to the data upload problem described above - for unknown reasons the remaining 2 sensors did not upload data to the server beyond 14 and 15 August.

The temperature and RH measured by the two LEO sensors and measurements made using the HOBO Ware instrument were in reasonably close agreement (Figure 3 and Figure 4).

The LEO sensors recorded a large number of negative NO_2 concentrations. NO_2 and NO concentrations recorded by the two LEO sensors were highly correlated, with LEO_980266 estimating higher concentrations than LEO_9801275. The correlation for O_3 measurements was much lower (Figure 5). The concentrations of NO_2, O_3 and NO estimated by the LEO sensors were plotted against concentrations measured by the Townhead analysers during the same period (Figure 6). The correlation between the sensor and analyser concentrations was low for all pollutants ($R^2 < 0.15$ in all instances).

Some studies have found NO_2 sensors to be cross-sensitive to ozone and have suggested correction methods to take this into account (e.g. Lin et al. 2015[1]). We followed the process described in this paper and observed that the LEO sensors appear to have pronounced cross sensitivity to O_3 in an unexpected direction (higher observed LEO O_3 resulting in lower observed LEO NO_2). However (unlike our experience with Aeroqual sensors; and despite clearly defined cross sensitivity relationships in the difference between LEO NO_2 and analyser NO_2 vs. LEO O_3), corrected LEO NO_2 concentrations still had relatively low correlation with analyser NO_2 concentrations (Figure 7).

[1] Lin, C., Gillespie, J., Schuder, M.D., Duberstein, W., Beverland, I.J. and Heal, M.R., 2015. Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide. *Atmospheric Environment*, *100*, pp.111-116.

Mobile monitoring study

The mobile monitoring study was designed as part of an MSc dissertation project. Measurements were made using a LEO sensor (LEO_980275 carried using the arm band) alongside Aeroqual NO_2 and Aeroqual O_3 sensors (carried in a backpack). We measured concentrations of NO_2 and O_3 at 13 sites located at different distances from busy roads in the West End of Glasgow (Figure 8). 5-min 'spot' measurements were made at each of the sites during the afternoon of Friday 26 July 2016 and average concentrations at each location calculated. It was anticipated that those sites located closest to the main road would have the highest NO_2 concentrations and the lowest O_3 concentrations. Measurements were repeated twice at each site (Repeat 1: site 1 – 13, Repeat 2: site 13 – 1). A third repeat set of measurements was incomplete as the LEO instrument stopped recording data (Repeat 3: site 1 – 3 only).

The Aeroqual NO_2 instruments are cross-sensitive to O_3 and therefore were corrected following the procedures in Lin et al. (2015). The calibration equations were derived from the deployment of the instruments at Townhead during the same period as discussed in the previous section:

$Aeroqual\_NO_2\_corrected = Aeroqual\_NO_2\_uncorrected – ((1.15*Aeroqual\_O_3) + 35.04)$

Concentrations of NO_2 measured by the LEO instruments were lower than the Aeroqual instruments, including many negative concentrations observed by each of the LEO instruments (Figure 9). LEO O_3 concentrations were much higher than Aeroqual O_3 concentrations. In accordance with prior expectations Aeroqual instruments observed slightly lower O_3 concentrations at the road side sites. However this pattern was not apparent in the LEO measurements. Correlations between LEO and Aeroqual measurements were low for both NO_2 and O_3 (R^2 M1 = 0.03, M2 = 0.00, All = 0.00 for NO_2; R^2 M1 = 0.00, M2 = 0.06, All = 0.03 for O_3).

Figure 2: Deployment of the four LEO instruments (two instruments located back to back) in a passively ventilated enclosure on the roof of Townhead automatic monitoring station.

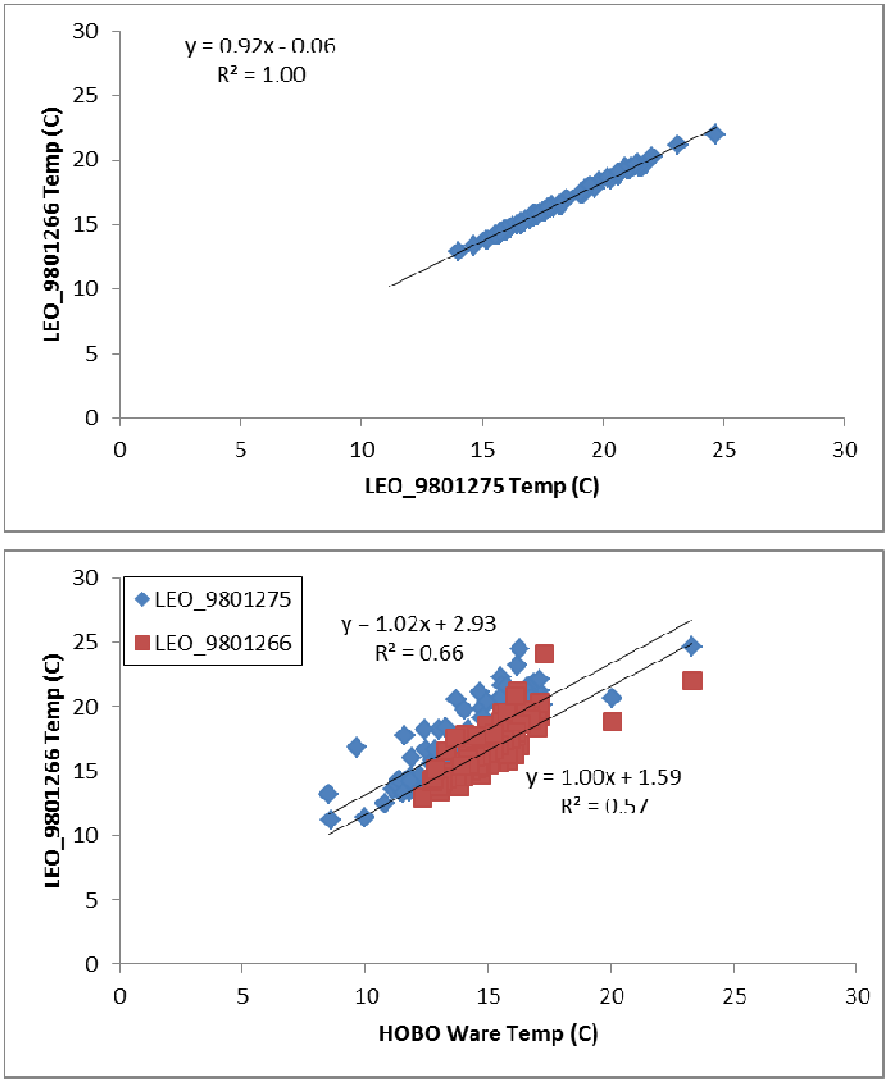


Figure 3: LEO sensor temperature measurements. Top: scatter plot between two LEO instruments; bottom: scatter plot between LEO instruments and HOBO temperature measurements.

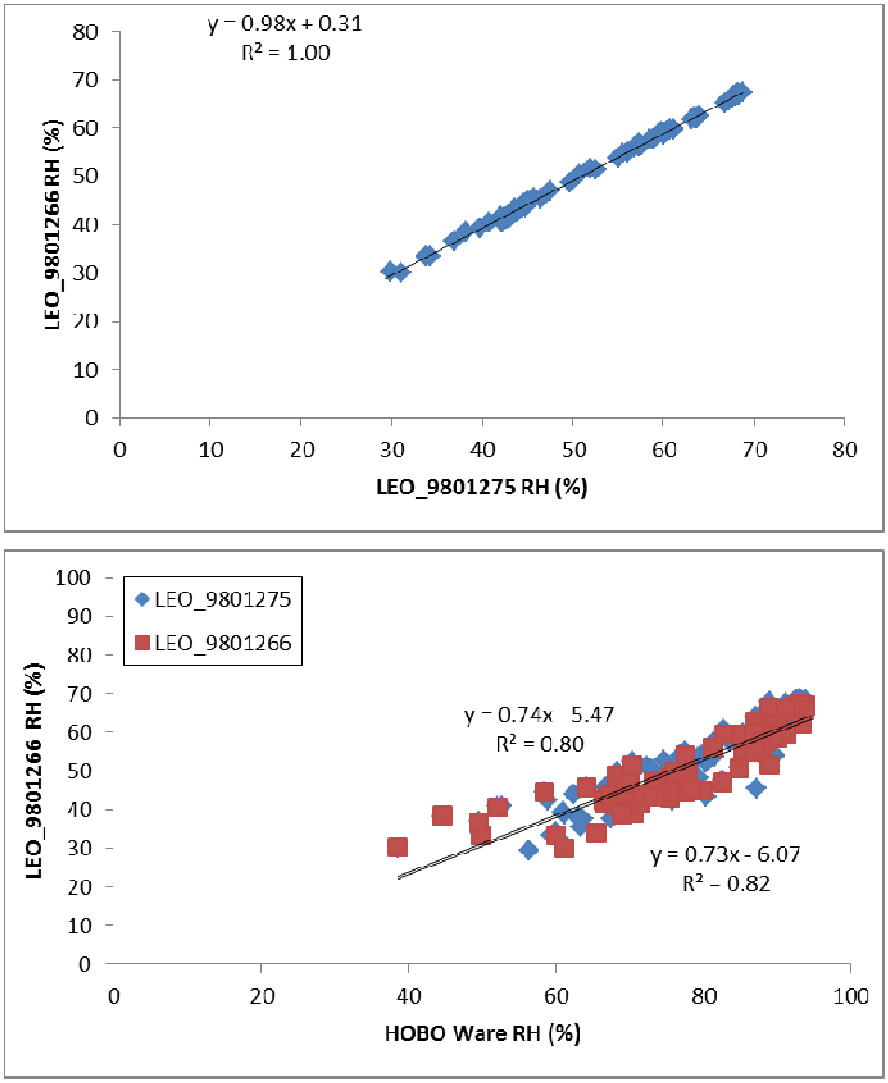


Figure 4: Top: Scatter plot between LEO RH measurements; bottom: scatter plot between LEO and HOBO RH measurements.

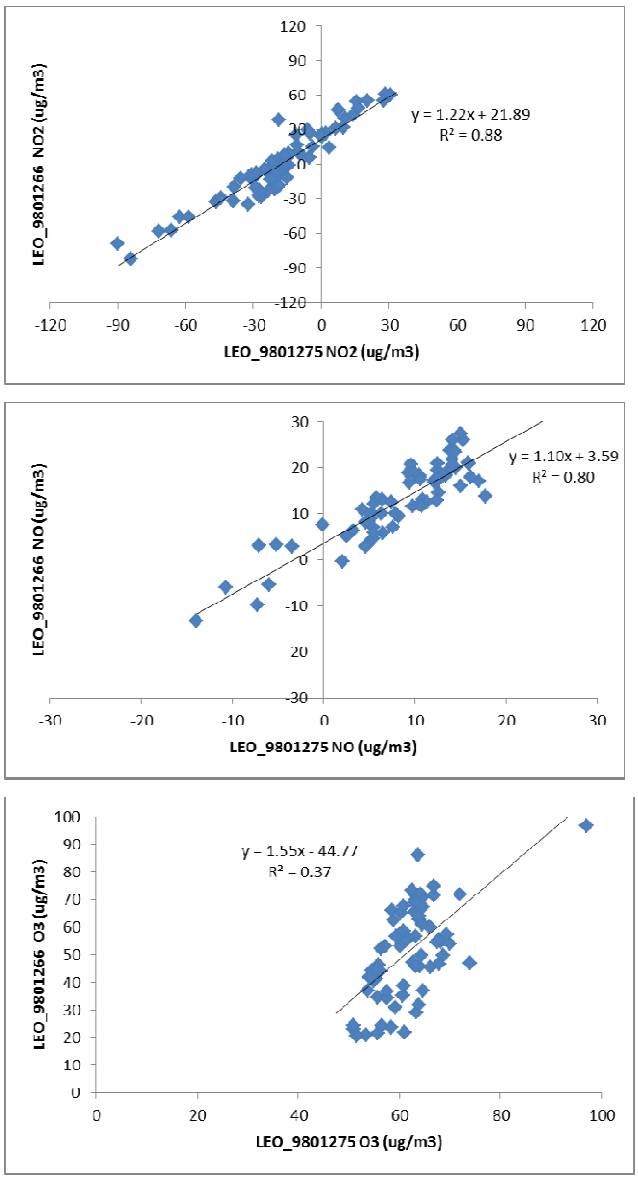


Figure 5: Scatter plots between duplicate LEO sensors for hourly concentrations of (from top to bottom): NO_2, NO and O_3.

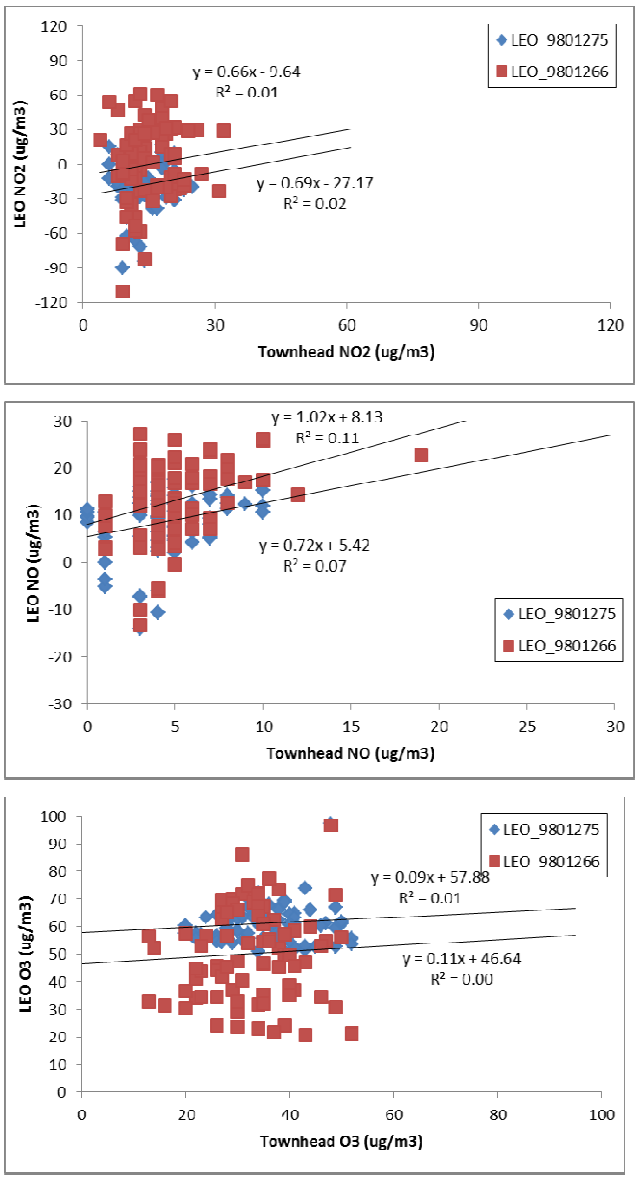


Figure 6: Scatter plots between pollution concentrations measured by LEO sensors and Townhead analysers. Top to bottom: NO_2, NO, O_3.

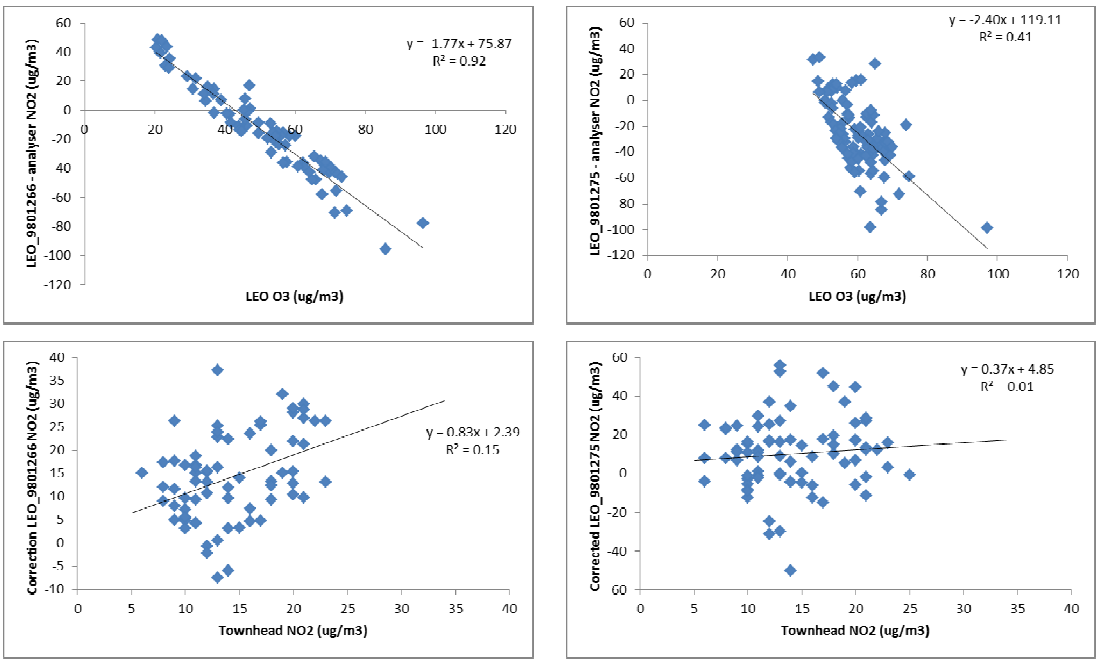


Figure 7: Correction of LEO NO_2 sensor for cross-sensitivity to O_3 (following method from Lin et al. 2015). Top graphs show NO_2 concentrations measured by the LEO sensors minus the concentrations measured by the analyser at Townhead *vs.* O_3 concentrations measured by the LEO sensors. The equation of this line was used to correct the LEO NO_2 concentrations, and this corrected concentration is shown plotted *vs.* Townhead analyser NO_2 in the bottom graphs.

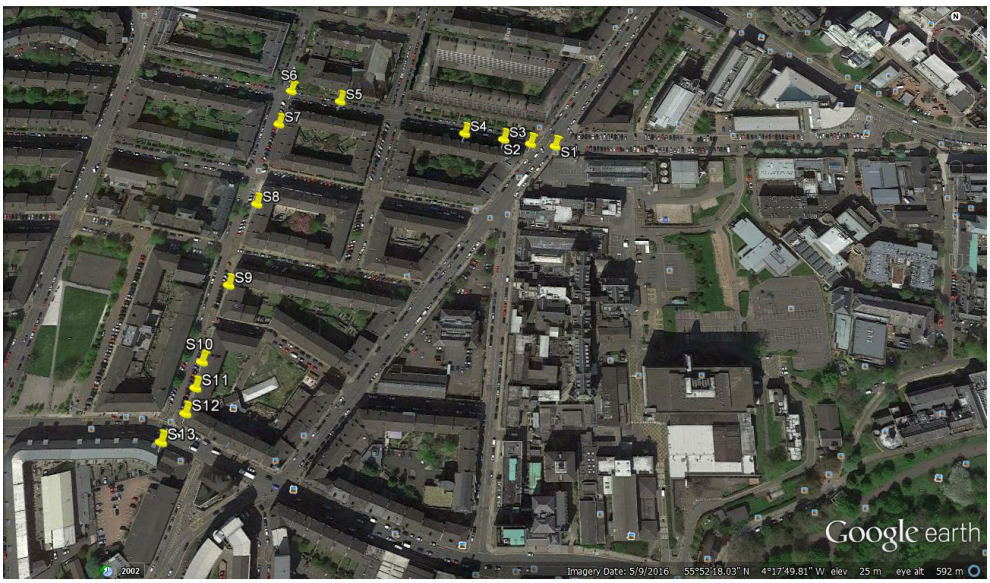


Figure 8: Location of the thirteen sites during the West End mobile measurements. These are located in two transects moving away from busy roads (Byres Road running approximately North-East to South-West and Dumbarton Road running approximately East-West). Five minute spot measurements were made at each site, and two or three repeat measurements were made at each site over the course of one afternoon.

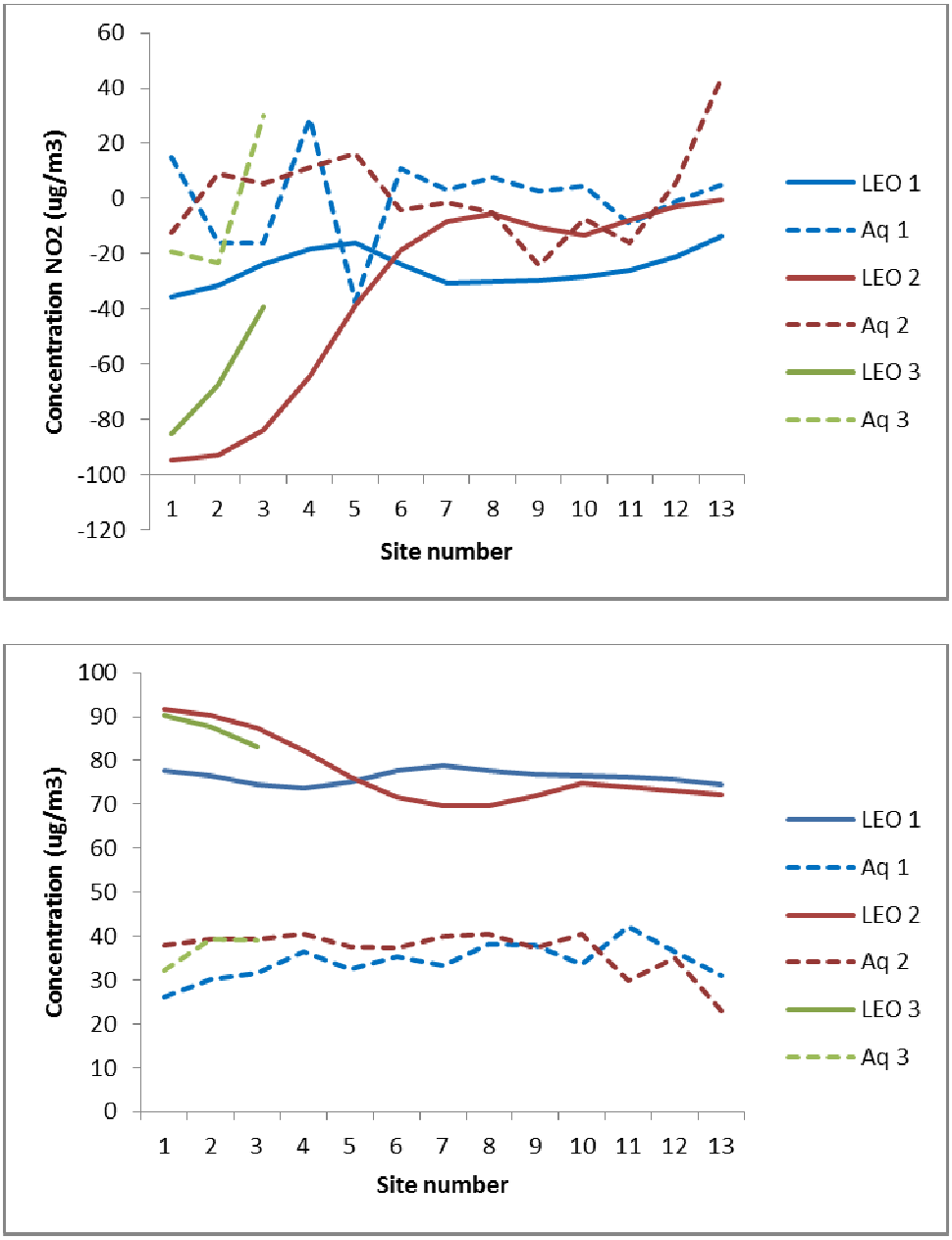


Figure 9: Line plots showing the average LEO and Aeroqual concentration measured during the 5 minute spot measurement at each site in the West End work. Top: NO_2, bottom: O_3.

Appendix B – Technical Note: Estimation of spatial patterns of urban air pollution over a 4-week period from repeated 5-minute measurements

The Technical Note below describes our earlier published study investigating the ability of repeated short duration measurements of black carbon and particle numbers to estimate spatial and temporal variations in nitrogen dioxide concentrations measured using passive samplers. This work provided the basis for the study presented in Chapter 4.

The technical note was published in Atmospheric Environment in 2017 (DOI: 10.1016/j.atmosenv.2016.11.035)

Estimation of spatial patterns of urban air pollution over a 4-week period from repeated 5-minute measurements

Jonathan Gillespie[1], *Nicola Masey*[1], *Mathew R. Heal*[2], *Scott Hamilton*[3], *Iain J. Beverland*[1]*

[1] Department of Civil and Environmental Engineering, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK

[2] School of Chemistry, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

[3] Ricardo Energy and Environment, Blythswood Square, Glasgow G2 4BG

*CORRESPONDING AUTHOR

Dr Iain J. Beverland, Department of Civil and Environmental Engineering, University of Strathclyde, 505F James Weir Building, 75 Montrose Street, Glasgow, G1 1XJ, UK; email: iain.beverland@strath.ac.uk; Tel: +44 141 548 3202

Abstract

Determination of intra-urban spatial variations in air pollutant concentrations for exposure assessment requires substantial time and monitoring equipment. The objective of this study was to establish if short-duration measurements of air pollutants can be used to estimate longer-term pollutant concentrations. We compared 5-min measurements of black carbon (BC) and particle number (PN) concentrations made once per week on 5 occasions, with 4 consecutive 1-week average nitrogen dioxide (NO_2) concentrations at 18 locations at a range of distances from busy roads in Glasgow, UK. 5-min BC and PN measurements (averaged over the two 5-min periods at the start and end of a week) explained 40 - 80%, and 7 - 64% respectively, of spatial variation in the intervening 1-week NO_2 concentrations for individual weeks. Adjustment for variations in background concentrations increased the percentage of explained variation in the bivariate relationship between the full set of NO_2 and BC measurements over the 4-week period from 28% to 50% prior to averaging of repeat measurements. The averages of five 5-min BC and PN measurements made over 5 weeks explained 75% and 33% respectively of the variation in average 1-week NO_2 concentrations over the same period. The relatively high explained variation observed between BC and NO_2 measured on different time scales suggests that, with appropriate steps to correct or average out temporal variations, repeated short-term measurements can be used to provide useful information on longer-term spatial patterns for these traffic-related pollutants.

Keywords: Black carbon; nitrogen dioxide; microaethalometer; air pollution; mobile; particle number.

1 Introduction

A major challenge in quantifying the effect of air pollution on human health is the resource required reliably to measure spatial and temporal variations in pollutant concentrations within urban environments (Hoek et al., 2008). The development of lightweight, lower-power portable monitoring equipment provides new opportunities to design monitoring studies that supplement static monitoring networks by using mobile measurements. Two approaches are possible: continuously mobile monitoring, where the monitoring equipment is moved throughout the duration of the study; and peripatetic monitoring, where mobile equipment is deployed at specific sites for short time periods before moving to another site.

Peripatetic measurements allow collection of observations through a monitoring network over a period of time and over relatively large areas with limited equipment. This approach has been used to monitor air pollution at sequential locations in studies in Canada (Abernethy et al., 2013; Deville Cavellin et al., 2016; Larson et al., 2009), Germany (Merbitz et al., 2012), India (Saraswat et al., 2013), the Netherlands (Klompmaker et al., 2015), Spain (Rivera et al., 2012), Switzerland (Ragettli et al., 2014) and the USA (Riley et al., 2016). A limitation with this approach is the difficulty in accounting for fluctuating background concentrations, although this can be mitigated by using a static background site during the study (Hoek et al., 2008; Klompmaker et al., 2015).

The objective of our study was to examine the quantitative relationships between short-term peripatetic measurements made with handheld equipment and longer-term average spatial air pollutant patterns, to assess if one can be used as a surrogate for the other. We combined peripatetic 5-min 'spot' measurements of black carbon (BC) and particle number (PN) [using portable low-power equipment] with weekly nitrogen dioxide (NO_2) measurement [using passive diffusion tubes (PDT)] over four 1-week periods at 18 sites of varying distance from major roads in Glasgow, UK. Five-minute static measurements were made at each site during deployment and retrieval of the PDTs. The combination of PDT and peripatetic measurements enabled investigation of spatial correlations between different pollutants at different averaging periods.

2 Methods:

2.1 Monitoring plan:

The study was conducted in the city of Glasgow (population ~ 600,000) in the west of Scotland (55.87° N, 4.26° W), for four consecutive weeks beginning on 24 October 2013. Eighteen monitoring sites were selected in a mixed residential and commercial area in the

West End of the city to provide a range of local traffic influence (Figure). The two busiest roads in the study area, Byres Road and Dumbarton Road, have annual average daily flows (AADF) of approximately 10,000 vehicles. A background site (Site 18) in a nearby park provided measurements free from immediate influence of local traffic sources.

Duplicate NO_2 PDTs were located at each site and changed approximately weekly at times that avoided adverse weather to avoid damage to real-time equipment. Therefore, weeks 1 and 2 spanned 8 and 6 days respectively, while weeks 3 and 4 spanned 7 days. During PDT exchange, while stationary at each site, 5-min peripatetic measurements were made using handheld BC and PN instruments (section 2.2). PDT changeovers began around 08:00 local time and took approximately 2.5 hours to complete. All real-time instrument clocks were synchronized prior to measurements. Peripatetic measurements were made during this time of the morning when many people were traveling on roads to get to work to maximise the range of observed concentrations. To reduce the possibility of systematic bias, sites 1 to 17 were visited in opposite order on alternate weeks (starting with site 1 in w/c week 1). Because of its distance from the other sites, the background site, 18, was always visited last. A duplicate PN instrument at the background site provided an indication of changes in background concentration during each measurement period (Supplementary information – Figure S4). A duplicate BC instrument was not available, so for consistency we made background adjustments for both pollutants using 5-min measurements made at site 18 at the end of each measurement period.

In this Technical Note the notation ‘week X’ and abbreviation ‘wX’ refer to PDT measurements throughout week X or to ‘weekly spot’ measurements derived from averaging the mobile measurements made at the beginning and end of week X, while ‘w/c week X’ and abbreviation ‘w/cX’ refer to mobile measurements made at the start of week X only. For example, ‘w/c week 5’ refers to mobile measurements made at the end of the fourth week of the study when PDTs were collected for the final time.

2.2 Instruments and data processing:

BC concentrations were measured using a microaethelometer (Model AE51, Aethlabs, San Francisco, CA) carried in a backpack, with the manufacturer-supplied 1 m conductive plastic tubing inlet mounted on the shoulder strap. BC was recorded at 1-min resolution during w/c week 1, and 1-s resolution during subsequent weeks. This change made it easier to synchronise arrival and departure times at PDT sites with logged data. At 1-s temporal resolution microaethelometers are prone to measurement artifacts (Hagler, 2011). Consequently 1-s BC data were processed using an optimised noise adjustment (ONA)

method (Δ ATN = 0.01) to retain the highest possible temporal resolution (AethLabs, 2013; Hagler, 2011). A second adjustment, to account for non-linear response with increasing BC deposition, was also applied (Apte et al., 2011):

$$BC = BC_o(0.88\ \exp(-\frac{ATN}{100}) + 0.12)^{-1}$$

where BC = adjusted BC concentration, BC_0 = unadjusted BC concentration, and ATN = attenuation value from the instrument. A single filter strip was sufficient for all measurements and gave an ATN value of < 50 at the conclusion of the study. The AE51 instrument was evaluated by deployment next to an AE22 aethalometer used for black carbon measurements at the UK government Automatic Urban and Rural Network (AURN) monitoring site at Townhead, Glasgow (Figure S2)

PN was measured using two handheld condensation particle counters (CPC 3007, TSI Inc., Shoreview, MN). Before each set of measurements the CPCs were checked for zero reading, supplied with fresh isopropyl alcohol, and allowed to warm up for 10 min. Precision of the duplicate PN instruments was assessed by walking them together through urban environments with a similar range of pollutant concentrations to those in this study. Duplicate instruments exhibited a high degree of precision (R^2 = 0.93, Figure S1) and < 2% normalised mean bias between the paired instruments.

Palmes NO_2 PDTs were deployed in duplicate at 2.5 m elevation at each site. PDTs were prepared as a single batch at the beginning of the campaign by dipping stainless steel mesh grids into 50% triethanolamine-acetone solution (Heal, 2008) and stored double bagged in a refrigerator pre and post deployment. Two 'travel' blanks were carried during deployment and retrieval of PDTs, and kept in a laboratory refrigerator during the intervening period. Two 'field' blanks were deployed close to site 13, and two 'laboratory' blanks were kept in the refrigerator during the exposure period. PDTs and blanks were analysed within 3 or 4 days of retrieval using a standard protocol (Targa et al., 2008). Laboratory and travel blanks showed no significant concentration values. The mean relative standard deviation (± 1 sd) for all 70 duplicate PDT measurements was 6.4 (± 6.8)%, comparable with that reported in the literature (*e.g.* Lewne et al., 2004). Four out of 144 (3%) PDTs were lost during measurements, consisting of pairs of duplicate tubes lost in week 1 and week 4 from sites 17 and 16 respectively. Consequently statistics in weeks 1 and 4 were not fully comparable with other weeks, as data from two of the highest concentration sites were missing during these weeks.

We used four approaches to assess if short-term measurements could provide useful information on longer-term spatial trends in pollutant concentrations. Firstly, we compared 5-min BC and PN 'spot' measurements made in each week to the average of all 5-min

measurements for BC and PN over the study. Secondly, we calculated the average of BC and PN spot measurements made at the start and end of each week (subsequently referred to as 'weekly spot' measurements) and compared this average to weekly NO_2 concentrations measured by PDTs throughout the intervening period. Thirdly, we corrected weekly 'spot' BC and weekly NO_2 concentrations to allow for changes in background concentrations measured at site 18 at the end of each measurement period. This was done by using a 'difference' method (Klompmaker et al., 2015) that involved: (a) computation of the overall mean concentration for the full set of 5-min measurements at the background site (site 18) for each pollutant ($C_{ref,ave}$); (b) computation of differences between period specific measurements at the background site ($C_{ref,t}$) and the estimated overall background mean ($C_{ref,ave}$) for each pollutant for each period (t) ($C_{diff,ref,t} = C_{ref,ave} - C_{ref,t}$); (c) correction of the period measurement at each site (x) by addition of the difference calculated in step (b) ($C_{x,t,corrected} = C_{x,t,measured} + C_{diff,ref,t}$). A 'ratio' method of temporal adjustment (Klompmaker et al. 2015) was also examined but found to produce less consistent reduction in within-site/between-site variance ratios (Table 2); therefore most of our analyses with temporal adjustments were focused on the difference method. In a fourth approach we examined the bivariate relationships between estimates of the overall averages of NO_2, BC and PN concentrations for the 4-week period across the 18 sites. Reduced major axis (RMA) regression was used to compare pollutant metrics in the above approaches (Ayers, 2001). One-way analysis of variance (ANOVA) was used to compare within-site (temporal)/between-site (spatial) variance ratios (Klompmaker et al. 2015).

3 Results and discussion

3.1 NO_2, BC and PN by site

Descriptive statistics and discussion of the time series of measurements are provided in the Supplementary Information.

Relatively high NO_2 concentrations were consistently observed across all weeks at sites closest to main roads (sites 1, 6, 7, 16, 17) (Figure). The lowest NO_2 concentration each week was observed at the background site (Site 18), where NO_2 concentration varied markedly between weeks, but was always 2 - 3 times lower than the maximum observed concentration for each week (Figure 2, Table S1). Despite large variations in average NO_2 concentrations between weeks, spatial patterns of relative concentrations across the sites remained consistent from week to week (concentrations were highly correlated between pairs of successive weeks (R^2 = 82%, 88% and 82%)).

5-min averaged 'spot' measurements for BC and PN demonstrated qualitatively similar spatial patterns to NO_2 (Figure). However, the spot measurements showed a less consistent spatial pattern between successive weekly measurements than was observed for NO_2 (R^2 ranges of 29 – 81% and 0 – 29% for BC and PN respectively).

Background concentrations measured at site 18 at the end of each weekly monitoring period were, on average across all weeks, ~65%, 40% and 30% of the mean concentration of all other sites for NO_2, PN and BC respectively suggesting that, of those pollutants measured in this study, BC was the metric with spatial variations that are most influenced by proximity to local traffic sources.

3.2 Longer-term predictions from 5 minute measurements

The linear relationship between 5-min PN measurements and the average of five 5-min spot PN was not significant for 1 out of 5 weeks (Table 1a). Relatively low week-to-week correlations between PN measurements (section 3.1) may have resulted from changes in atmospheric processes that determine the formation of ultrafine particles through changes in meteorology between and within weeks. Meteorological conditions also influence NO_2 concentrations but the influence will be reduced for the 1-week averaged PDT measurements compared with the short-term PN measurements. Correlations between 'weekly spot' PN and 1-week NO_2 concentrations were not significant on 2 of 4 weeks, and explained < 25% of variation in 1-week NO_2 concentrations during all but one week (Table 1b). Consequently the remainder of this Technical Note focuses on the more clearly observed relationships between BC and NO_2.

5-min BC measurements were significantly associated with the average of five 5-min spot measurements taken once per week (average explained variation 55%, range 44 – 87%) (Table 1a). 'Weekly spot' measurements of BC explained between 40% and 80 % (average = 62%) of the variation in 1-week NO_2 concentrations (Table 1b, Figure 3a). The lowest explained variation was observed during week 4 and may have resulted from limited variation in BC concentration and missing data at one of the higher concentration NO_2 sites (Site 16). The regression slope and intercept varied between weeks, with the *y*-axis intercept providing a good approximation of the background NO_2 concentration measured at Site 18 (Table 1). This suggests that (subject to confirmation using observations for larger areas and longer periods) it may be possible to estimate urban background NO_2 concentration using short-term BC measurements alongside weekly NO_2 PDTs.

Correlation between background-adjusted BC and NO_2 was highly significant and explained 50% of the variation in weekly NO_2 concentrations for the full set of

measurements (*cf.* 28% explained variation prior to background adjustment) (Table 1, Figure 3). The overall average of 5-min BC spot measurements over the full study period (5-min measurements repeated 5 times over 4 weeks) explained 75% of the variation in overall average NO_2 concentrations (Table 1, Figure 3c). Averaging selected subsets of repeated BC spot measurements interspersed evenly within the 5 measurement periods (to simulate a lower repeat peripatetic measurement frequency) resulted in a lower percentage of explained variation in overall average NO_2 concentrations (69% and 59% using weeks 1, 3 & 5 and weeks 1 & 5 respectively (Figures 3d & 3e)).

Our results and conclusions are broadly coherent with comparisons of mobile real-time and static passive measurements of traffic-related pollutants in Baltimore, USA using different measurement approaches over different time and geographical scales (Riley et al., 2016). Our findings can also be set in the context of quantitative analyses of within-site/between-site variance ratios for BC and PN peripatetic measurements in the Netherlands (Klompmaker et al. 2015). Our 5-min peripatetic observations have lower within-site/between-site variance ratios (i.e. exhibit more temporal consistency in spatial patterns) than 30-min peripatetic observations in the MUSiC study in the Netherlands (Table 2). The magnitudes of the ratios we observed are relatively close to those observed with 14 day measurement periods in the European ESCAPE project (Table 2). The reasons for the relatively limited temporal variation in the observations in our study are not fully clear, but may be related to the relatively small geographical area and short time period over which measurements were conducted. The relatively limited temporal variation in our measurements are also consistent with temporally persistent spatial variations in NO_2 and O_3 concentrations observed in PDT measurements in the nearby city of Edinburgh, UK (Lin et al., 2016).

Collectively the relatively high correlations observed between NO_2 and BC measurements, and relatively low within-site/between-site variance ratios suggest that short-term measurements with limited repetition are capable of partly characterising pollution concentration gradients in the urban environment. However, some limitations are relevant for consideration. Firstly, in the absence of continuous longer-term measurements of BC to compare with short-duration measurements, our study made use of the relationship between BC and NO_2 to assess the effectiveness of 5-min measurements for estimation of longer-term spatial contrasts. Other studies have shown BC and NO_2 to be highly correlated over extended time periods (Durant et al., 2014). Secondly, each week of mobile measurements was completed in approximately 2.5 hours, around the time of the morning rush hour. Our measurements may have been affected by changing traffic and meteorological conditions.

We attempted to minimise systematic bias by reversing the order in which sites were visited on alternate weeks. Thirdly, there is uncertainty regarding the optimal duration for 'spot' measurements, and whether the sampling period should be the same for all site classifications. We observed limited variation in concentrations during the 5-min 'spot' measurements at sites adjacent to roads with lower traffic flows (Figure S3), where a shorter duration 'spot' measurement may have been sufficient. Conversely, 'spot' measurements made at sites near higher and variable traffic flows were more variable (Figure S3) and may benefit from a longer measurement period. Additionally, when measurements are made in areas where traffic is influenced by local traffic signals, increasing the 'spot' measurement period to encompass the full cycle of signals may be beneficial.

4 Conclusions:

This study compared 5-min 'spot' measurements of black carbon (BC) and particle number (PN) concentrations, measured at weekly intervals at 18 locations in the city of Glasgow, against 1-week measurements of NO_2 concentrations. On average, 5-min BC measurements during individual measurement periods explained 55% variation of the overall average of five 5-min spot measurements taken once per week over the 4-week period. BC measurements of 5-min duration at the beginning and end of weeks explained 40 - 80% of spatial variations in NO_2 during the intervening 1-week periods. Equivalent measurements of PN explained 7 to 64% of 1-week NO_2 spatial variations. After adjusting for changes in background NO_2 and BC concentrations, spot measurements of BC and PN conducted repeatedly over a 4-week period, explained 50% and 24% respectively of the spatial variation in the complete set of corresponding 1-week NO_2 concentrations. The average of 5 replicate 5-min BC and PN spot measurements explained 75% and 33% respectively of the spatial variation in 4-week average NO_2 concentrations. Reducing the number of replicate peripatetic BC measurements to 3 and 2 replicates reduced the percentage of explained variation in spatial variation in 4-week average NO_2 concentrations to 67% and 59% respectively. Collectively these observations (with appropriate allowance for their relatively limited duration and spatial extent) suggest that short-term peripatetic measurements can be used to estimate longer term spatial contrasts in traffic-related air pollution provided that appropriate steps are taken to correct or average out temporal variations.

Figures:

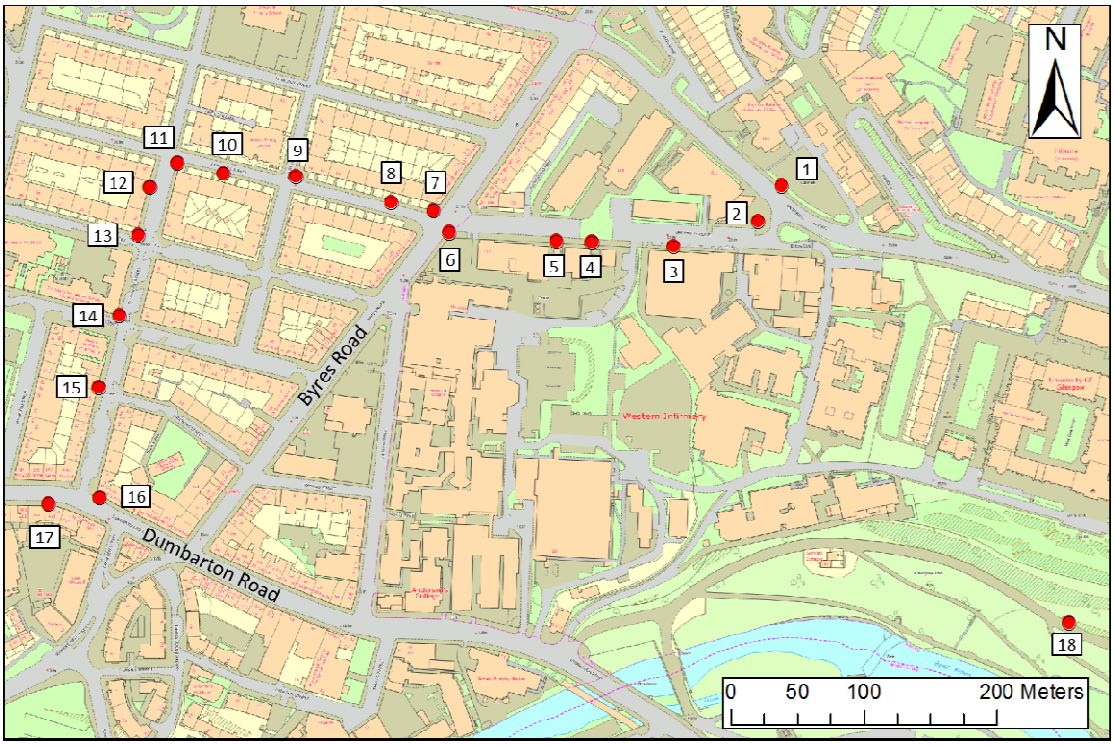


Figure 1: Locations in the West End of Glasgow of NO_2 passive diffusion tubes and 'spot' monitoring locations for BC and PN concentrations.

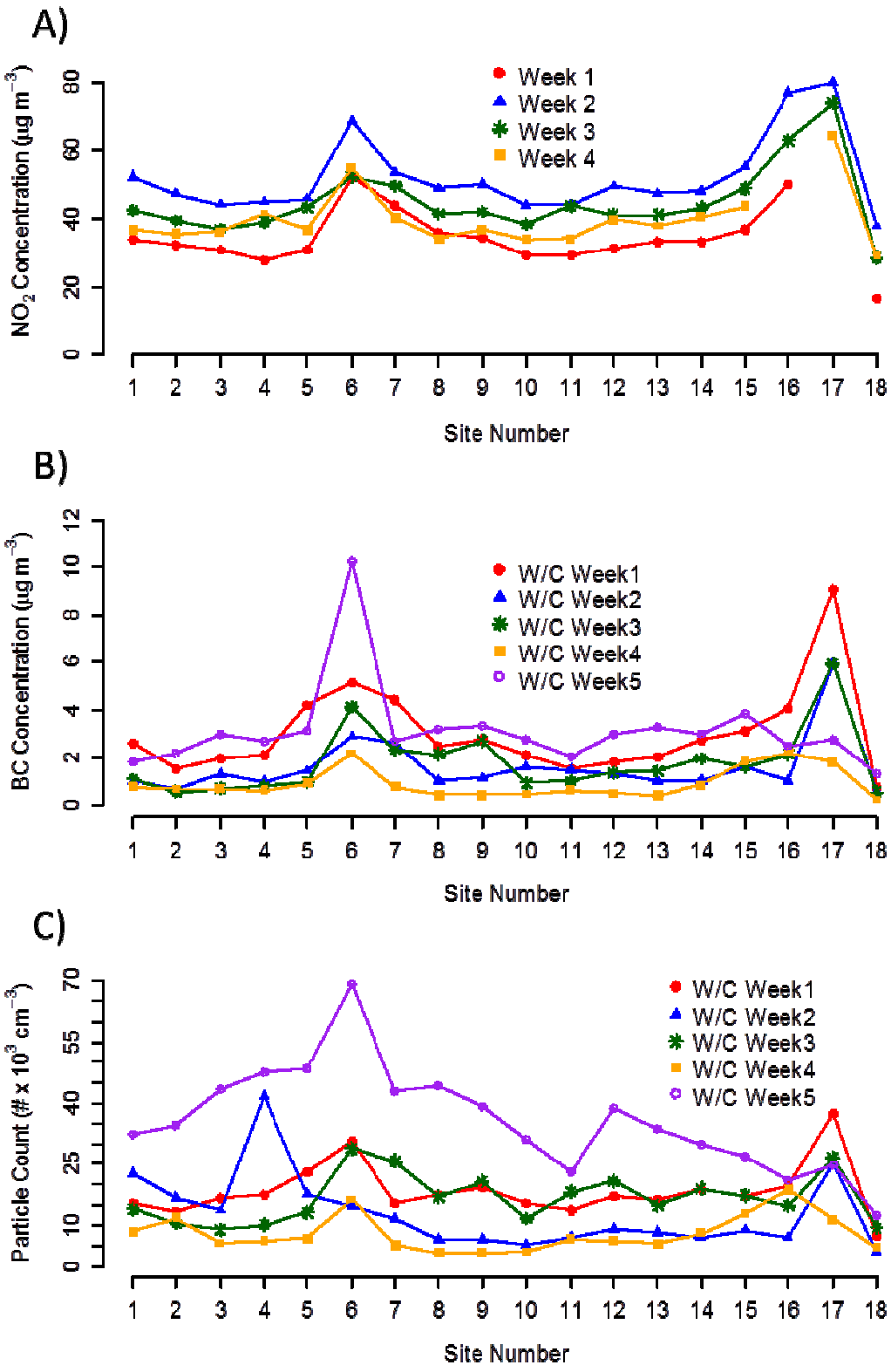


Figure 2: Concentrations of (A) NO_2, (B) BC and (C) PN at each site. NO_2 concentrations are 1-week averages and BC and PN concentration are averages of 5-min 'spot' measurements.

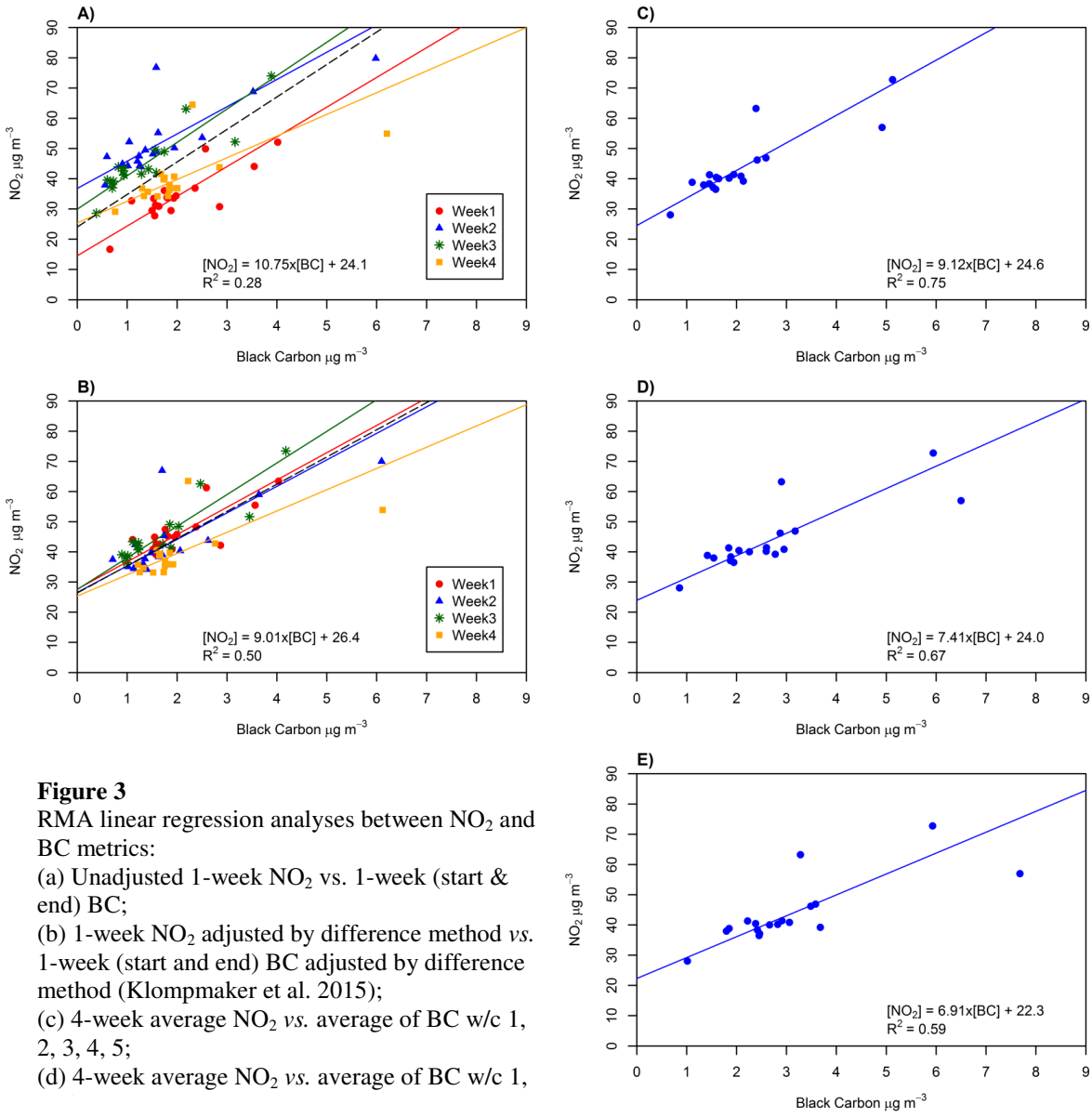


Figure 3
RMA linear regression analyses between NO_2 and BC metrics:
(a) Unadjusted 1-week NO_2 vs. 1-week (start & end) BC;
(b) 1-week NO_2 adjusted by difference method *vs.* 1-week (start and end) BC adjusted by difference method (Klompmaker et al. 2015);
(c) 4-week average NO_2 *vs.* average of BC w/c 1, 2, 3, 4, 5;
(d) 4-week average NO_2 *vs.* average of BC w/c 1,

Table 1 Reduced major axis regression statistics between longer-term average pollution measurements and 5-min BC (μg m^{-3}) and 5-min PN (# x 10^3 cm^{-3}) spot measurements.

Pollutant	*n*	R^2 (*p*)	Slope (95 % CI)	Intercept (95 % CI)	Background $[NO_2]$[1]
(a) Spot metrics for specific weeks *vs.* overall average of spot measurements for all weeks:					
BC(all) *vs.* BC(w/c1)	18	0.84 (1 x10^{-7})	1.49 (1.31- 2.01)	-0.4(-1.2 - 0.27)	-
BC(all) *vs.* BC(w/c2)	18	0.75 (3 x 10^{-6})	1.07 (0.82 - 1.39)	-0.6(-1.3 - -0.1)	-
BC(all) *vs.* BC(w/c3)	18	0.87 (2 x 10^{-8})	1.18 (0.98 - 1.43)	-0.7(-1.2 - -0.3)	-
BC(all) *vs.* BC(w/c4)	18	0.57 (2 x 10^{-4})	0.55 (0.39 – 0.77)	-0.2(-0.7 - 0.1)	-
BC(all) *vs.* BC(w/c5)	18	0.44 (0.003)	1.60 (1.09 – 2.35)	-0.2(-1.8 - 0.9)	-
PN(all) *vs.* PN(w/c1)	18	0.62 (1 x 10^{-4})	1.26(0.92 - 1.74)	-4.7(-13.5 – 1.7)	-
PN(all) *vs.* PN(w/c2)	18	0.38 (0.007)	1.78 (1.19 – 2.68)	-19.9(-36.4 - -8.9)	-
PN(all) *vs.* PN(w/c3)	18	0.31 (0.014)	1.15 (0.75 - 1.75)	-4.4(-15.6 - 2.8)	-
PN(all) *vs.* PN(w/c4)	18	0.16 (0.10)	0.85 (0.53 - 1.35)	-7.6(-17.0 - -1.8)	-
PN(all) *vs.* PN(w/c5)	18	0.61 (1 x 10^{-4})	2.45 (1.77 – 3.40)	-9.4(-26.8 - 3.1)	-
(b) NO_2 for specific weeks *vs.* spot metrics for equivalent weeks:					
NO_2(w1) *vs.* BC(w1)	17	0.67 (6 x10^{-5})	9.8 (7.2-13.4)	14.6 (7.3-19.9)	16.7
NO_2(w2) *vs.* BC(w2)	18	0.61 (1 x 10^{-4})	9.0 (6.5-12.5)	36.8 (31.9-41.1)	37.9
NO_2(w3) *vs.* BC(w3)	18	0.80 (6 x 10^{-7})	11.0 (8.7-14)	30.0 (26.0-33.2)	28.6
NO_2(w4) *vs*. BC(w4)	17	0.40 (6 x 10^{-3})	7.2 (4.7-10.8)	25.4 (18.0-30.3)	29.1
NO_2(all) *vs.* BC(all)	70	0.28 (3 x 10^{-6})	10.7 (8.8 – 13.2)	24.1 (19.8 – 27.6)	
NO_2(ave) *vs.* BC(ave)[2]	18	0.75 (4.0 x 10^{-6})	9.1 (7.0-11.9)	24.6 (18.8-29.0)	
NO_2(w1) *vs.* PN(w1)	17	0.07 (0.31)	1.5 (0.9-2.5)	12 (-2.9-20.9)	-
NO_2(w2) *vs.* PN(w2)	18	0.23 (0.05)	1.2 (1.3-3.3)	20.5 (2.4-32.0)	-
NO_2(w3) *vs.* PN(w3)	18	0.64 (7 x 10^{-5})	2.4 (1.8-3.3)	15 (3.91-23.1)	-
NO_2(w4) *vs*. PN(w4)	17	0.18 (0.09)	1.2 (0.7-1.9)	13.8 (-2.39-23.7)	-
NO_2(all) *vs.* PN(all)	70	0.07 (0.03)	1.5 (1.2 – 1.8)	9.4 (0.7 – 16.2)	-
NO_2(ave) *vs.* PN(ave)[1]	18	0.33 (0.013)	2.0 (1.3-3.1)	6.4 (-13.1-19.2)	-
(c) Background-adjusted NO_2 for specific periods *vs*. background-adjusted spot metrics for equivalent periods[3]:					
NO_2(w1) *vs.* BC(w1)	16	0.60 (4 x10^{-4})	9.0 (6.3 – 12.9)	27.7 (19.6 – 33.4)	-
NO_2(w2) *vs.* BC(w2)	17	0.59 (4 x10^{-4})	8.8 (6.2 – 12.4)	26.6 (19.7 – 31.4)	-
NO_2(w3) *vs.* BC(w3)	17	0.79 (2 x 10^{-6})	10.4 (8.2 – 13.4)	27.6 (22.5 – 31.5)	-
NO_2(w4) *vs*. BC(w4)	16	0.36 (0.014)	7.0 (4.5 – 11.0)	25.4 (17.5 – 30.5)	-

Pollutant	*n*	R^2 (*p*)	Slope (95 % CI)	Intercept (95 % CI)	Background $[NO_2]$[1]
NO_2(all) *vs.* BC(all)	66	0.50 (3 x 10^{-11})	9.0 (7.6-10.8)	26.4 (23.1 – 29.2)	-
NO_2(w1) *vs.* PN(w1)	16	0.01 (0.91)	1.4 (0.8 – 2.4)	22.5 (4.9 – 32.7)	-
NO_2(w2) *vs.* PN(w2)	17	0.16 (0.11)	2.2 (1.4 – 3.6)	7.1 (-15.3 – 21.0)	-
NO_2(w3) *vs.* PN(w3)	17	0.59 (0.003)	2.3 (1.7 – 3.3)	14.4 (1.6 – 23.4)	-
NO_2(w4) *vs*. PN(w4)	16	0.10 (0.24)	1.3 (0.8 – 2.2)	11.5 (-7.7 – 22.9)	-
NO_2(all) *vs.* PN(all)	66	0.24 (4 x10-5)	1.9 (1.5-2.4)	15.0 (8.0 – 20.6)	-

[1] *NO_2 measurements at background site are listed alongside intercepts for non-background-adjusted NO_2 vs. BC regression lines.*

[2] *'ave' represents the average of all NO_2 PDT and BC/PN spot measurements over the full study period at each site.*

[3] *Background adjusted data represent observed concentration for specific period adjusted for temporal changes in concentrations at background (site 18) using the 'difference' method described in Methods Section 2.2 (based on method described by Klompmaker et al., 2015).*

Table 2 Comparison of within:between site variance ratios from this study, and other European studies (MUSiC, ESCAPE, RUPIOH, VE^{3}SPA) summarised by Klompmaker et al. (2015).

Study:	Repeats x duration	NO_2	NO_2(adj[1])	BC	BC(adj[1])	PN	PN(adj[1])
Present study ('ratio'	5 x 5-min BC & PN	0.14	0.31	0.28	0.13	1.17	0.64
Present study ('difference'	5 x 5-min BC & PN	0.14	0.05	0.28	0.21	1.17	0.77
MUSiC	3 x 30- min PN			3.25	2.44	2.21	2.17
ESCAPE	3 x 14-day $PM_{2.5}$ absorbance			0.39	0.09		
RUPIOH	3 x 1-day PN					0.5	0.31
VE^{3}SPA	6 x 4-day $PM_{2.5}$ absorbance			2.55	0.69		

[1] Adj = adjustment for temporal variation. For overview of different methods used to adjust for temporal variation see Klompmaker et al. (2015).

Author Contribution

Iain Beverland, Jonathan Gillespie and Nicola Masey contributed to the design of the study. Jonathan Gillespie, Nicola Masey conducted data collection. The first draft of the manuscript was manuscript was written by Jonathan Gillespie and all other authors contributed to discussions on data analysis and revisions of the paper. All authors have given approval to the final version of the manuscript.

Funding Sources

Jonathan Gillespie is funded is funded through an Engineering and Physical Sciences Research Council Doctoral Training Grant (EPSRC DTG EP/L505080/1 and EP/K503174/1) studentship, with support from the University of Strathclyde and Ricardo Energy and Environment. Nicola Masey is funded through a UK Natural Environment Research Council CASE PhD studentship (NE/K007319/1), with industrial support from Ricardo Energy and Environment. The AE51 instrument was purchased with funding from a NERC multi-institution grant (NE/1007822/1).

Acknowledgements

We would like to thank Stephanie Burns and Gabor Puikovics for their assistance in making field measurements during this study. We acknowledge access to UK government Automatic Urban and Rural Network (AURN) measurement data, which were obtained from uk-air.defra.gov.uk and are subject to Crown 2014 copyright, Defra, licenced under the Open Government Licence (OGL). The research data associated with this paper are available at: http://dx.doi.org/10.15129/bccc477e-2a12-4f3e-a690-333b06bad6c6.

References:

Abernethy, R.C., Allen, R.W., McKendry, I.G., Brauer, M., 2013. A Land Use Regression Model for Ultrafine Particles in Vancouver, Canada. Environ. Sci. Technol. 47, 5217–5225. doi:10.1021/es304495s

AethLabs, 2013. microAeth [WWW Document]. URL https://aethlabs.com/microaeth

Apte, J.S., Kirchstetter, T.W., Reich, A.H., Deshpande, S.J., Kaushik, G., Chel, A., Marshall, J.D., Nazaroff, W.W., 2011. Concentrations of fine, ultrafine, and black carbon particles in auto-rickshaws in New Delhi, India. Atmos. Environ. 45, 4470–4480. doi:10.1016/j.atmosenv.2011.05.028

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

Deville Cavellin, L., Weichenthal, S., Tack, R., Ragettli, M.S., Smargiassi, A., Hatzopoulou, M., 2016. Investigating the Use Of Portable Air Pollution Sensors to Capture the Spatial Variability Of Traffic-Related Air Pollution. Environ. Sci. Technol. 50, 313–320. doi:10.1021/acs.est.5b04235

Durant, J.L., Beelen, R., Eeftens, M., Meliefste, K., Cyrys, J., Heinrich, J., Bellander, T., Lewné, M., Brunekreef, B., Hoek, G., 2014. Comparison of ambient airborne PM2.5, PM2.5 absorbance and nitrogen dioxide ratios measured in 1999 and 2009 in three areas in Europe. Sci. Total Environ. 487, 290–298. doi:10.1016/j.scitotenv.2014.04.019

Hagler, G.S.W., 2011. Post-processing Method to Reduce Noise while Preserving High Time Resolution in Aethalometer Real-time Black Carbon Data. Aerosol Air Qual. Res. doi:10.4209/aaqr.2011.05.0055

Heal, M.R., 2008. The effect of absorbent grid preparation method on precision and accuracy of ambient nitrogen dioxide measurements using Palmes passive diffusion tubes. J. Environ. Monit. 10, 1363–1369. doi:10.1039/b811230d

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

Klompmaker, J.O., Montagne, D.R., Meliefste, K., Hoek, G., Brunekreef, B., 2015. Spatial variation of ultrafine particles and black carbon in two cities: Results from a short-term measurement campaign. Sci. Total Environ. 508, 266–275. doi:10.1016/j.scitotenv.2014.11.088

Larson, T., Henderson, S.B., Brauer, M., 2009. Mobile Monitoring of Particle Light Absorption Coefficient in an Urban Area as a Basis for Land Use Regression. Environ. Sci. Technol. 43, 4672–4678. doi:10.1021/es803068e

Lewne, M., Cyrys, J., Meliefste, K., Hoek, G., Brauer, M., Fischer, P., Gehring, U., Heinrich, J., Brunekreef, B., Bellander, T., 2004. Spatial variation in nitrogen dioxide in three European areas. Sci. Total Environ. 332, 217–230. doi:10.1016/j.scitotenv.2004.04.014

Merbitz, H., Fritz, S., Schneider, C., 2012. Mobile measurements and regression modeling of the spatial particulate matter variability in an urban area. Sci. Total Environ. 438, 389–403. doi:10.1016/j.scitotenv.2012.08.049

Ragettli, M.S., Ducret-Stich, R.E., Foraster, M., Morelli, X., Aguilera, I., Basagaña, X., Corradi, E., Ineichen, A., Tsai, M.-Y., Probst-Hensch, N., Rivera, M., Slama, R., Künzli, N., Phuleria, H.C., 2014. Spatio-temporal variation of urban ultrafine particle number concentrations. Atmos. Environ. 96, 275–283. doi:10.1016/j.atmosenv.2014.07.049

Riley, E.A., Schaal, L., Sasakura, M., Crampton, R., Gould, T.R., Hartin, K., Sheppard, L., Larson, T., Simpson, C.D., Yost, M.G., 2016. Correlations between short-term mobile monitoring and long-term passive sampler measurements of traffic-related air pollution. Atmos. Environ. 132, 229–239. doi:10.1016/j.atmosenv.2016.03.001

Rivera, M., Basagaña, X., Aguilera, I., Agis, D., Bouso, L., Foraster, M., Medina-Ramón, M., Pey, J., Künzli, N., Hoek, G., 2012. Spatial distribution of ultrafine particles in urban settings: A land use regression model. Atmos. Environ. 54, 657–666. doi:10.1016/j.atmosenv.2012.01.058

Saraswat, A., Apte, J.S., Kandlikar, M., Brauer, M., Henderson, S.B., Marshall, J.D., 2013. Spatiotemporal Land Use Regression Models of Fine, Ultrafine, and Black Carbon Particulate Matter in New Delhi, India. Environ. Sci. Technol. 47, 12903–12911. doi:10.1021/es401489h

Targa, J., Loader, A., The Defra Working Group on Harmonisation of Diffusion Tubes, 2008. Diffusion Tubes for Ambient NO2 Monitoring: A Practical Guide.