THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

# Assessing the Effectiveness of Interventions on Air Quality

**Citation for published version:**
Monks, PS, Allan, JD, Carruthers, D, Carslaw, DC, Fuller, GW, Harrison, RM, Heal, MR, Lewis, AC, Nemitz, E, Reeves, C, Williams, M, Fowler, D, Marner, BB, Williams, A, Moller, S, Maggs, R, Murrells, T, Quincey, P & Willis, P 2020, *Assessing the Effectiveness of Interventions on Air Quality*. Air Quality Expert Group, London. <https://uk-air.defra.gov.uk/library/reports.php?report_id=1004>

**Link:**
Link to publication record in Edinburgh Research Explorer

**Document Version:**
Publisher's PDF, also known as Version of record

**General rights**
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 19. Feb. 2023

AIR QUALITY EXPERT GROUP

# Assessing the Effectiveness of Interventions on Air Quality

Prepared for:

Department for Environment, Food and Rural Affairs;
Scottish Government; Welsh Government;
and Department of Agriculture, Environment and Rural Affairs in Northern Ireland

AIR QUALITY EXPERT GROUP

# Assessing the Effectiveness of Interventions on Air Quality

Prepared for:

Department for Environment, Food and Rural
Affairs; Scottish Government; Welsh Government;
and Department of Agriculture, Environment and
Rural Affairs in Northern Ireland

This is a report from the Air Quality Expert Group to the Department for Environment, Food and Rural Affairs; Scottish Government; Welsh Government; and Department of Agriculture, Environment and Rural Affairs in Northern Ireland, on assessing the effectiveness of interventions on air quality. The information contained within this report represents a review of the understanding and evidence available at the time of writing.

© Crown copyright 2020

Front cover image credit: Maxian/iStock via Getty Images.

United Kingdom air quality information received from the automatic monitoring sites and forecasts may be accessed via the following media:

| | |
|---|---|
| Freephone Air Pollution Information Service | 0800556677 |
| Internet | http://uk-air.defra.gov.uk |

PB 14598

# Terms of Reference

The Air Quality Expert Group (AQEG) is an expert committee of the Department for Environment, Food and Rural Affairs (Defra) and considers current knowledge on air pollution and provides advice on such things as the levels, sources and characteristics of air pollutants in the UK. AQEG reports to Defra's Chief Scientific Adviser, Defra Ministers, Scottish Ministers, the Welsh Government and the Department of the Environment in Northern Ireland (the Government and devolved administrations). Members of the Group are drawn from those with a proven track record in the fields of air pollution research and practice.

AQEG's functions are to:

- Provide advice to, and work collaboratively with, officials and key office holders in Defra and the devolved administrations, other delivery partners and public bodies, and EU and international technical expert groups;
- Report to Defra's Chief Scientific Adviser (CSA): Chairs of expert committees will meet annually with the CSA, and will provide an annual summary of the work of the Committee to the Science Advisory Council (SAC) for Defra's Annual Report. In exception, matters can be escalated to Ministers;
- Support the CSA as appropriate during emergencies;
- Contribute to developing the air quality evidence base by analysing, interpreting and synthesising evidence;
- Provide judgements on the quality and relevance of the evidence base;
- Suggest priority areas for future work, and advise on Defra's implementation of the air quality evidence plan (or equivalent);
- Give advice on current and future levels, trends, sources and characteristics of air pollutants in the UK;
- Provide independent advice and operate in line with the Government's Principles for Scientific Advice and the Code of Practice for Scientific Advisory Committees (CoPSAC).

Expert Committee Members are independent appointments made through open competition, in line with the Office of the Commissioner for Public Appointments (OCPA) guidelines on best practice for making public appointments. Members are expected to act in accord with the principles of public life.

Further information on AQEG can be found on the Group's website at: https://www.gov.uk/government/policy-advisory-groups/air-quality-expert-group.

# Membership

## Chair

**Professor Paul Monks (until October 2019)**
University of Leicester
**Professor Alastair Lewis (from October 2019)**
National Centre for Atmospheric Science, University of York

## Members

**Dr James Allan**
National Centre for Atmospheric Science, University of Manchester

**Dr David Carruthers**
Cambridge Environmental Research Consultants

**Dr David Carslaw**
Ricardo Energy and Environment and University of York

**Dr Gary Fuller**
King's College London

**Professor Roy Harrison OBE**
University of Birmingham

**Professor Mat Heal**
University of Edinburgh

**Professor Alastair Lewis (until October 2019)**
National Centre for Atmospheric Science, University of York

**Dr Eiko Nemitz**
Centre for Ecology & Hydrology

**Professor Claire Reeves**
University of East Anglia

**Professor Martin Williams**

King's College London

## *Ad hoc* members

**Professor David Fowler CBE**

Formerly Centre for Ecology and Hydrology

**Dr Ben Marner**

Air Quality Consultants

**Dr Andrew Williams**

University of Chester

## *Ex officio* members

**Dr Sarah Moller**

National Centre for Atmospheric Science, University of York and Senior Research Fellow, Department for Environment, Food and Rural Affairs Systems Research Programme

Central Management and Control Unit of the automatic urban and rural networks: **Dr Richard Maggs**, Bureau Veritas

National Atmospheric Emissions Inventory: **Dr Tim Murrells**, Ricardo Energy and Environment

Non-automatic hydrocarbon monitoring networks and metals monitoring network: **Dr Paul Quincey**, National Physical Laboratory

Quality Assurance and Quality Control of the automatic urban network and the non-automatic monitoring networks: **Dr Paul Willis**, Ricardo Energy and Environment

## Assessors and observers

**Roger Herbert**

Welsh Government

**Barry McCauley**

Department of the Environment in Northern Ireland

**Andrew Taylor**

Scottish Government

**Alison Gowers**

Public Health England

## Secretariat

**Shaun Brace**

Department for Environment, Food and Rural Affairs

**Dr Mohamed Ghalaieny**

Department for Environment, Food and Rural Affairs

**Michelle Brailey-Balster**

Department for Environment, Food and Rural Affairs

Previously:

**Dr Ailsa Stroud**

Department for Environment, Food and Rural Affairs

## Contents

# Executive Summary

In the air quality context, interventions cover a wide range of actions from ‘deliberate’ measures to reduce air pollution to those primarily aimed at other outcomes, but which can indirectly affect air pollution. Air quality interventions span a wide range of situations, spatial scales from (e.g. a single road through to continental scales) and temporal scales from the short-term closure of a road to decadal (or longer) changes. The principal focus of this report is on local scale interventions, such as those that might reasonably be considered by local authorities.

Understanding the impact interventions have on air quality is highly desirable because of the need to quantify the outcome on air quality and health i.e. relate a policy aimed at improving air quality to a robust understanding of the outcome.

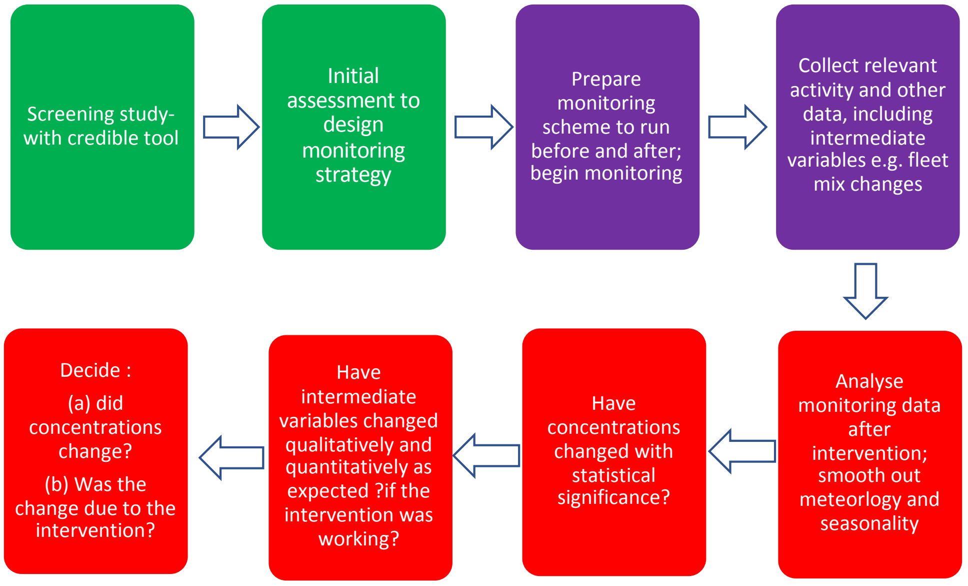
The assessment of interventions can be challenging for several reasons. These challenges include the common situation where interventions rarely occur in isolation from other changes that affect air quality, and the difficulty in detecting and quantifying changes if the interventions are small. Indeed, not every intervention is detectable in terms of quantifying changes in pollutant concentrations or health outcomes, even using sophisticated analysis techniques.

The work in the report suggests the ‘accountability chain’ model provides a useful way to assess the impact of an intervention from a change in activity through to potential beneficial health effects i.e. activity → emissions → concentrations → health outcomes.

From an analysis of a wide range of case studies the following needs have been identified:

1. The availability of appropriate activity data that would be expected to change because of an intervention.
2. The availability of well-sited ambient air quality measurement data increases the potential for data analysis and the application of statistical methods to quantify changes to the concentrations of pollutants.
3. Air quality models can have an important role to play. Although most intervention analyses considered in this report do not use air quality models, they provide the means of assessing the likely impact that interventions have on air quality and allow scenarios to be considered.
4. When interventions are assessed based on changes in concentrations of pollutants, it is important to take account of meteorological variation. Meteorological variation can easily mask or emphasise changes in concentrations resulting from changes in emissions. Methods exist to account for or ‘remove’ the influence of meteorology.

The report suggests the following flow methodology to conduct an analysis of an intervention:



**Colour code: Green: prior to intervention Purple: during intervention Red: After intervention**

The report gives clear recommendations that detail the need for pragmatic advice for practitioners, the assessment of causality against confounding factors and the counter-factual, enhanced statistical approaches and future siting criteria for air quality monitoring.

The following recommendations have been identified.

- The design of the assessment of an intervention should be considered at the planning stage, with the practitioners receiving pragmatic advice on the process. To date, the information relating to the analysis of interventions often resides in academic journals and is not easily accessible by most practitioners.
- Many areas of policy should seek to incorporate the approaches and thinking of intervention analysis, e.g. in determining the efficacy or otherwise of technologies that offer potential reductions to emissions of different pollutants.
- An analysis of the intermediate steps which the intervention is designed to influence should be carried out, to assess the causality of the relationship between the intervention and any measured concentration changes. For example, if the intervention is to change a pattern of fuel use, has fuel consumption changed as intended? Furthermore, there is also a need to take account of other changes that may also have affected the assessment of an intervention and to develop an understanding of whether any changes in concentrations specifically resulting from an intervention can reasonably be quantified.
- Results from the implementation of local plans to mitigate air pollution should be pooled to derive a statistically more robust overall assessment of local measures.
- Where new air quality monitoring sites are planned, their relevance for assessing interventions should be considered.

- Intervention analysis is an active area of research that continues to evolve and be refined. Defra should retain a watching brief in this area to understand new developments and promote good practice.

# 1. Introduction

## 1.1. Aims of the report

This report reviews methods that have been used to evaluate the effectiveness of interventions intended to reduce air pollution and associated health effects. The intention is to explain the different approaches that have been used and to highlight the requirements for undertaking a robust evaluation. The report does not provide information on the effectiveness of different types of intervention that might be implemented. Nonetheless, the overview of the approaches used, and their strengths and limitations, should prove helpful to organisations considering the implementation of interventions and wishing to evaluate their effectiveness.

## 1.2. What is an intervention?

Interventions cover a wide range of scenarios from 'deliberate' measures aimed at improving air quality to interventions that may be primarily aimed at other outcomes but which can indirectly affect air pollution. Interventions also span a very wide range of conditions and can cover the small scale (e.g. a single road) through to regional scale (e.g. continental scale). Additionally, interventions can span a very wide range of temporal scales such as from the short-term closure of a road to decadal (or longer) scales. The type of intervention can also cover a wide range of situations from changes resulting in emission reduction due to national or international emissions or air quality legislation, through to decisions made by local authorities that can affect local air quality.

## 1.3. Scope of the report

Given the very wide range of potential definitions of what constitutes an intervention, this report necessarily focuses on a limited set of situations. The principal focus is on local scale interventions, such as those that might reasonably be considered by local authorities. However, it is also useful to consider some evidence from non-local interventions where there is relevance and where valuable information is available.

A narrow view of the metrics used to consider interventions might only consider whether there have been changes to ambient concentrations resulting from an intervention. While such a metric is clearly highly valuable – and is of direct relevance to legislation controlling ambient concentrations, it is a limited view. Interventions need to consider a wider range of practical issues and the ‘gold standard’ quantification of changes to ambient concentrations or impacts on health is not always achievable. For this reason, it is also necessary to consider other indicators of change that might involve surrogate or proxy measures of change – for example, changes to measured or modelled traffic flows, or vehicle fleet composition. For many assessments, it may be impractical (or impossible) to quantify changes in ambient concentrations and a more limited, partial consideration of changes can still be useful. Of importance though, is an understanding of these limitations.

At the heart of these issues is establishing causation: whether a change can be attributed to a known cause. In air pollution, establishing causation can be highly challenging. These challenges include: interventions rarely occur in isolation from other changes to emissions, the effect of meteorology can easily mask or emphasise changes in ambient concentrations and the information available to quantify interventions is rarely compiled with the specific aim of quantifying interventions. However, the careful consideration of interventions can also help to challenge claims made about changes or indeed show that there is in fact no robust evidence that an intervention had a measurable effect.

This report does not consider interventions that aim to reduce concentrations after pollutants enter the atmosphere such as the use of photocatalytic surfaces.

This report covers some of the more practical issues such as considering the types of information that are useful and the expertise necessary. An important part of the report is the consideration of the evidence from published papers and reports on these issues, highlighted through a wide series of case studies.

## 1.4. Assessments for regulatory purposes at the local level

Assessments carried out for regulatory purposes make up a large proportion of all UK assessments of interventions. Assessments have mainly been carried under the following regulatory regimes:

- Local Air Quality Management, in which local authorities are required to report on the impacts of measures contained within their Local Air Quality Action Plans;
- Local Transport/Implementation Plans, in which strategic transport authorities are required to appraise the outcomes of their plans and measures;
- The Design Manual for Roads and Bridges Post Opening Project Evaluation, through which Highways England appraises the effects of Highways interventions;
- The Planning System, through which pre-construction air quality assessments will occasionally be followed by post-construction air quality monitoring.

There are also requirements for local authorities to report to Defra on the effects of projects for which they have received grant funding. Very often, these assessments fail to reach robust conclusions with regard to observed effects and their causation. There are many reasons for this, but better study design, along with a more realistic view of the changes that might be identified, could improve the usefulness of many of these studies. In particular, many of the principles used in research studies of interventions, as outlined in this report, might prove useful for regulatory assessments.

## 1.5. Overview of assessing interventions

A useful way in which to consider the impact of interventions is to think of the 'chain' of steps from a change in activity through to potential health effects. This sequence of steps is often called an 'accountability chain' and has been popularised by the Health Effects Institute (HEI) in the US (Greenbaum and Org, 2017).

The schematic shown in Figure 1 provides an overview of some of the steps that could be involved in the assessment of interventions; in this case for an example based on a road traffic intervention. The shaded boxes in Figure 1 show the main steps involved from the original policy, which results in a change in activity, *Δactivity*, a change in emissions, *Δemissions*, a change in concentrations, *Δconcentration* and finally some health change, *Δhealth*.

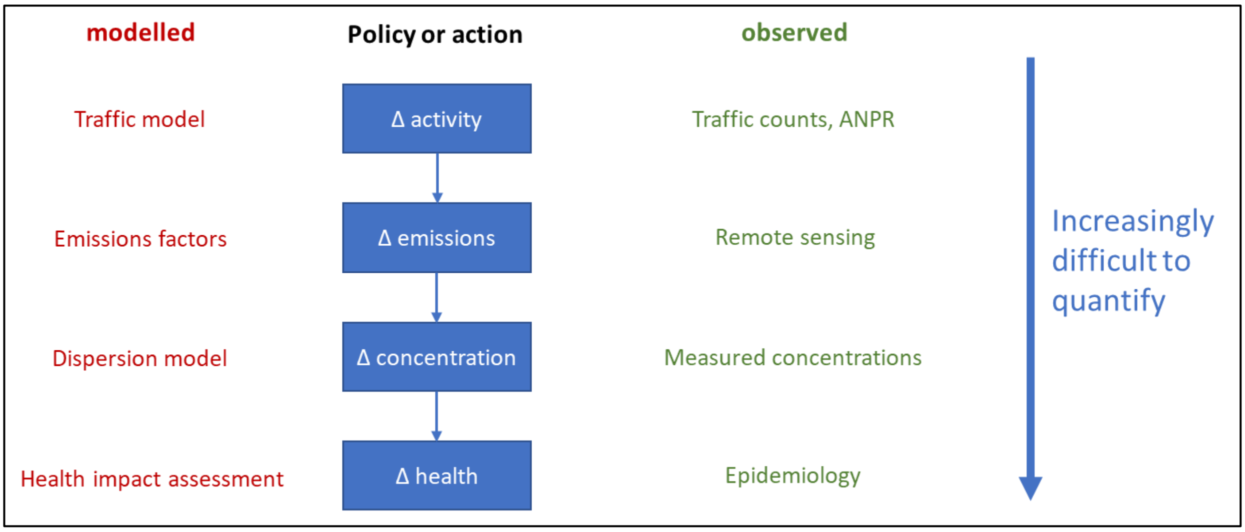


*Figure 1 Schematic of the stages involved in assessing interventions. In this example, the intervention is related to a road traffic intervention.*

Also shown in *Figure 1* are the types of data that can be used to assess each part of the accountability chain. On the left (in red) are the steps involved based on the use of modelled data. In green text is the parallel sequence of steps based on observations. In practice, assessments can involve a mix of measured and modelled quantities.

An important characteristic of the schematic shown in *Figure 1* is that as one moves down the chain, the quantification of the intervention becomes increasingly difficult. It is for example much easier to quantify a change in road vehicle flows than it is to determine the impacts of an intervention on human health. Indeed, as one moves down the chain, there are an increasing number of confounders that can make the analysis challenging. In the case of analysing measured concentrations for example, there will almost certainly be influences from other sources not associated with the intervention and the effect of meteorological confounders.

Greenbaum (2017) also highlights the importance of exploring the full “chain of accountability” by illustrating two studies of the traffic changes initiated to reduce congestion during the 1996 Olympics in Atlanta Georgia. The first study reported a 10–15% reduction in morning traffic, significant reductions in $O_3$, and reductions in asthma hospitalizations during the period of the Olympic Games (Friedman et al. 2001). However, a second much more detailed analysis found that while morning traffic did decrease, overall traffic volumes were unchanged (Peel et al. 2010). Peel et al. (2010) also found that the $O_3$ reductions were seen throughout the Southeast United States (in many areas unaffected by the Atlanta traffic patterns) and that hospitalizations, when adjusted for seasonal patterns for earlier and later years, showed no decline that could be tied to the reduced $O_3$. Such findings illustrate the need for a comprehensive approach to assessing interventions and the difficulties in establishing causal relationships.

The quantification of the effects of interventions is an active area of research and developments in methodology continue to be made.

# 2. Principles and approaches used for quantifying the effects of interventions

## 2.1. Challenges of quantifying interventions

There are many practical and methodological challenges related to the robust quantification of interventions. It is common for analyses to rely on existing data for intervention quantification rather than specifically collected data, which might not be optimal for quantification purposes. In this respect, it may prove challenging to quantify an intervention because the available data is not best suited to that purpose. Clearly, the suitability of the information available will depend on the situation being considered – but it would be expected that data collected specifically for quantifying interventions would be advantageous. There are however few examples of specifically designed experiments for that purpose. The London Low Emission Zone is one notable example where consideration was given to the optimum location of air quality measurement sites before the scheme started, as discussed in section 3.1.

There are also limitations in the use of proxy measurements and information to assess interventions. For example, a change in traffic flow is several steps away from a change in concentrations or health impacts. It is therefore important to understand some of the limitations of the information being used and the need perhaps to make additional assumptions about the change being considered.

A common issue when assessing interventions using ambient air quality data is that concentrations are usually affected by numerous other factors not associated with an intervention. The principal difficulty is establishing cause-effect against a backdrop that may be influenced by many other changes (e.g. general vehicle fleet turnover).

The processes involved in converting primary to secondary pollutants can be important and challenging to account for when assessing interventions. These issues will however depend on the pollutants being assessed and the spatial and temporal scale over which they are considered.

For many situations where the assessment of ambient concentrations is considered, the potentially large contribution from background air can be important. For example, a consideration of particulate matter concentrations in an urban area can be dominated by non-local sources and regional contributions. In these cases, the change in concentration due to an intervention may only contribute a small proportion of the absolute concentration, which could make the quantification of change difficult.

Arguably, the principal challenge related to the quantification of interventions using ambient air quality data is the influence of meteorology. Variations in meteorology can easily mask underlying changes to concentrations and can falsely mask or emphasise changes in concentration. Many of the approaches used to quantify interventions aim to remove or reduce the variations in concentration due to changes in meteorology. The types of analysis used to remove the effects of meteorology often share similarities with those used by epidemiologists, where confounding factors such as ambient temperature need to be ‘accounted for’ when attempting to quantify the health effects of pollutants.

## 2.2. Experimental design

As is shown in section 3, most practical examples of the analysis of interventions are not based on an initial experimental design, which would perhaps consider the most appropriate data to collect and how best to locate measurement sites to detect changes etc. It is much more common for the analysis to use data that exists for other purposes such as compliance monitoring.

The Health Effects Institute (HEI, 2016) has provided a study on how more systematic approaches to testing of causality (i.e., through use of causal inference frameworks and methods) could be adapted to the assessment of the effects of air pollution interventions on air quality and health. They successfully demonstrated the use of existing and newly developed methods in two case studies of regulatory actions: the designation of US counties to be in nonattainment with the National Ambient Air Quality Standards for $PM_{10}$ and the installation of $SO_2$ scrubbers on power plants.

The scrubber case study provided both newly developed methods and a rare comparison of two different but analogous statistical approaches — principal stratification and causal mediation analysis — applied to the same complex multipollutant problem. These methods seek to assess causality of an intervention by analysing statistical relationships between the intervention and the final outcome but also incorporating intermediate variables. For example, in their study one intervention was the installation of a scrubber on a power plant and the final outcome was impacts on health. The intermediate variables were emissions and ambient $PM_{10}$. Causal mediation analysis can incorporate several interventions, whereas principal stratification analysis is an analogous process but only using one intervention. Their work demonstrated the critical importance of involving multidisciplinary teams with detailed technical knowledge of the interventions to ensure appropriate study design and interpretation.  The HEI Review Committee in assessing the study concluded that these accountability methods are an important addition to the “toolkit” and should continue to be further explored, but cannot wholly substitute for accountability assessments that rely on evidence from other scientific methods, including more traditional epidemiology analyses.

Quantifying the impact an intervention has is also highly dependent on the pollutant considered. For example, it is more difficult to detect changes in $PM_{10}$ than $NO_x$ close to a road if the intervention affects road vehicles due to the large contributions from non-road sources to $PM_{10}$ concentrations. However, measuring certain PM components such as black carbon, could provide a much better indication of changes that relate to vehicle exhaust emissions.

## 2.3. Methods based on proxy data

Assessments of interventions can be made using proxies that are expected to be related to changes in ambient concentrations but are not in themselves a direct measure of concentration change. Examples include the measurement of changes in traffic flows (observed or from a traffic model) or some other measure of activity (see *Figure 1*). Such data can provide useful information that can help quantify the effect an intervention has but the use of such data will be very specific to the intervention being considered.

The measurement of some sort of activity such as road vehicle flows can be used as the basis of calculating changes in emissions. Quantifying potential changes in emissions can be highly useful in its own right and also as the basis of providing input to air quality models. The usefulness of such approaches will depend on how reasonable the emission estimates are, which will be dependent on the quality of the input data as well as the reliability of the emission factors themselves.

Building on the use of proxy input data and emission inventories is the use of air quality models to predict changes in concentrations. These methods are described in section 2.5.

## 2.4. Methods based on the analysis of ambient measurements

These methods are not dependent on assumptions about changes to emission sources or activity data but seek to detect and quantify the actual changes in measured concentrations. These methods have the benefit of aiming to determine the effect an intervention has on ambient concentrations directly. The approaches can include various data filtering methods (such as described in AQEG, 2015), through to sophisticated statistical analysis.

The efficacy of these approaches depends greatly on the pollutant being considered and the location of the measurement site(s). These considerations also include issues related to experimental design i.e. whether the measurement approach was chosen to specifically quantify the changes due to an intervention. Often, the analysis of concentration data to quantify the effects of interventions relies on existing measurement networks and species already measured, rather than being developed as part of an initial experimental design.

Statistical approaches can ‘account for’ changes to meteorology that can mask or emphasise changes in concentrations to better reveal changes to underlying source strength. The basis of these approaches is to use some sort of statistical model to explain concentrations of pollutants in terms of commonly available meteorological and other variables. If good models can be developed i.e. that can explain a large amount of the variation in concentrations in terms of meteorological variables, then new time series (trends) can be generated by choosing a fixed set of meteorological conditions and as input to the model and using the model to predict concentrations over a fixed set of conditions. These methods have been shown to be very effective, as summarised in section 3.

The tools available to carry out the analyses described above are increasingly becoming available, in part due to developments in statistical methodology but also the increased availability of the methods in open source software. A recent example is the *rmweather* R package developed at the University of York, which is freely available (see https://github.com/skgrange/rmweather) and Grange et al. (Grange and Carslaw, 2019).

### 2.4.1. Dealing with uncertainty in measurement

It is important to remember that many small-scale interventions will result in ambient air pollution changes that are too small to be detected. The reasons for the difficulty in detection include the changes in concentration being small compared with other wider scale changes such as the progressive fleet turnover, or smaller than the uncertainty in the ambient air pollution measurements. Although it might seem compelling to measure air pollution concentrations to demonstrate the effectiveness of an intervention, for minor schemes this may not be advisable.

Estimating the likely uncertainty in a measurement can be complicated. The Guide to the Assessment of Uncertainty in Measurement (GUM) (BIPM, 2008) provides a generic approach that requires knowledge of the measurement process and an estimate of the uncertainty in each component of the measurement (the gas cylinder concentrations, signal noise in the instrument etc.) which are then combined to provide an overall uncertainty dependent on the concentration and the time period of the measurement.

The uncertainty in longer term measurements such as annual mean concentrations will be less than that in short-term measurements. ISO 11222 (ISO, 2002) provides a method to estimate the uncertainty in longer-term measurements which starts by identifying the random and non-random components of the measurement uncertainty and reducing the impact of the random component over time.  In practice, this means that the uncertainty in a long-term measurement, such as an annual mean, can never be less than the uncertainty in the calibration standards used (around 4% or 5% being typical for gas standards such as NO in $N_2$, Sweeney et al 2015).

In practice, more pragmatic approaches can be used:

- When using EU reference methods or equivalent it is possible to assume the maximum specified in EU Directives e.g. 15% at the short-term limit value concentration for gases such as $NO_2$ and 25% for PM. This information can be used to estimate the uncertainty in short-term measurements. However, for longer-term measurements (of the order of months), the uncertainty would tend towards the underlying uncertainties in the calibration standards used.
- Assume the uncertainties from tests undertaken to show that an instrument is capable of making EU reference or reference equivalent measurements. Again, the ISO1122 approach is needed to consider uncertainties in longer term mean concentrations.
- To assume uncertainties given by a manufacturer or from other test procedures.

If measurements are undertaken to assess an intervention this will mostly likely require the comparison of at least two sets of measurements. These might be measurements made before and after the intervention or perhaps measurements made in the intervention location and in another an area that is used as a control.  The outcome of the intervention is then judged by looking at the difference between the two data sets.

An assessment of the uncertainty in the paired dataset comparison is required to determine if any difference is real or just an artefact of the measurement process.

Three approaches are possible here:

- Statistical tests of significant difference can be used. Air quality data is almost always non-parametric and contains outlier values so tests such as Mann Witney / Wilcoxon rank-sum or Kruskal–Wallis would be appropriate rather than the parametric Student’s t test.
- Estimates for uncertainty in the two datasets can be made using the approaches above.
- Estimates of uncertainty in the concentrations can be determined by field tests from a period of instrument colocation.

It should be noted that even if there are statistically significant differences between a set of measurements before and after the intervention, it does not necessarily prove that the change was caused by the intervention.

## 2.5. Methods based on Modelling

Modelling has a role to play both in the planning stage of interventions and in providing insight into the effectiveness of an intervention once implemented. To date, however, modelling approaches have not been extensively used in an air quality context to assist with intervention analysis.

### 2.5.1. Pre-intervention modelling

In the planning stage of an intervention modelling studies are used to inform decisions regarding whether a scheme is viable and to assist in its design. The studies can quantify the magnitude of changes in pollutant concentration that can be expected in different locations and therefore whether such changes are significant enough to be observed. In addition, they can inform measurement strategies designed to quantify the changes. It is important, however, to recognise that such studies are associated with significant uncertainty.

A typical example of a pre-intervention study is provided by Marner and Moorcroft (2016), who predicted the changes in the annual mean and 99.79th percentiles of 1-hour mean $NO_2$ concentrations associated with a scheme to introduce 2-way traffic, along with more pedestrian space, to Baker Street and Gloucester Place in Westminster. Changes in traffic movements were predicted using the VISSIM transport model. Traffic emissions were predicted using two average-speed emissions models: Defra's Emissions Factor Toolkit (Defra, 2016), and the CURED model (Marner, 2016). Concentrations were then predicted using the ADMS-Roads dispersion model, verified against annual mean measurements made at two monitoring sites. The results supported discussions on gaining permission for the scheme, and also informed the siting of pre- and post-construction air quality monitoring.

The impacts that are predicted in studies such as this are directly attributable to those model inputs that are deliberately altered. In this case, the variables were the lane positions, traffic volumes, and changes in average vehicle speeds. All other features, including meteorology, were unchanged in the different assessment scenarios. Modelling can be used to determine the cause of the predicted changes, however it is noted that, on account of modelling uncertainties, both the magnitude and the causes of the predicted changes are also subject to uncertainties; typically these have not been estimated for such studies.

An approach to dealing with uncertainties of predictive models has been presented by Tomlin et al. (2016). They demonstrated a global uncertainty and sensitivity methodology for a street scale model attempting to simulate the effects of changes to traffic demand on

$NO_x$ and $NO_2$ levels within a street canyon in York. Within the model uncertainties caused by possible errors in wind direction, chemical reaction rates and model parameterisations of surface roughness, it was possible to detect the influence of traffic demand and primary $NO_2$ emissions fractions. However, the method was only demonstrated for a single set of meteorological conditions and carried a significant computational cost since a large number of model runs was required for each case study to determine uncertainty.

A similar Monte Carlo sampling based study was also carried out for a Lagrangian photochemical pollution model by Bergin et al. (1999) for the Californian region. Their study evaluated the possible effects of reductions in $NO_x$ emissions on regional scale $O_3$ concentrations where model uncertainties were accounted for. A more recent study by Carnevale et al. (2016) carried out an assessment of proposed mitigation strategies for reducing $PM_{10}$ within the Lombardy region of Northern Italy using an Integrated Assessment Model. The work sets out a framework for assessing the possible effects of policy interventions using models, whilst taking into account model uncertainties.

### 2.5.2. Post intervention modelling

At the local and regional scale, atmospheric pollutant concentrations are affected by a wide range of emissions, chemical and physical processes, and as a result determining the causes of temporal and spatial changes solely from observations can be difficult, especially if the magnitude of the changes is small. Similar issues also exist for global pollution related problems such as climate change and stratospheric $O_3$ depletion. In these fields methodologies involving combined observation and modelling approaches have been applied in order to isolate the effects of changes in emissions on the target environmental impacts.

For example, a recent paper in Science (Solomon et al., 2016) compared different model predictions of changes to the size of the Antarctic $O_3$ hole. It was first established that the model simulated well, much of the observed year-to-year variability in September total $O_3$ for targets both at the South Pole and from satellite UV observations. The authors argue that this provides confidence in the attribution of changes to $O_3$ reductions to different causes which were isolated within model sensitivity studies. These included chemical contributions to the trends, as well as meteorology and aerosol loadings due to volcanic processes. Overall the work demonstrated reductions in $O_3$ depletion that could be attributed to reduced chlorine and bromine loading in the stratosphere.

Using such approaches in global climate models, multi-model ensemble simulations are performed with and without anthropogenic emissions in order to show that, when neglecting anthropogenic emissions, the simulated trends do not overlap with observations even when taking into account model uncertainties (IPPC, 2013). This showing in turn that the *changes* to observations in the atmosphere, cryosphere and oceans cannot simply be due to changes in natural processes (Figure 2). Uncertainties are propagated in the model predictions based on ensembles of simulations using models from different research groups.

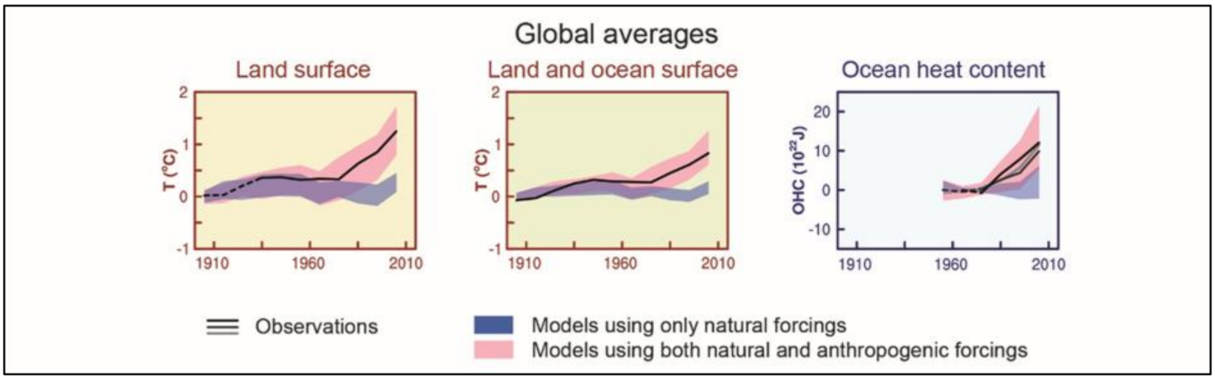


*Figure 2 Comparison of observed and simulated climate change based on three large-scale indicators in the atmosphere, the cryosphere and the ocean (IPPC, 2013). Uncertainty in the model simulations is shown by the shaded bands.*

To date there seems to have been a limited number of examples of such approaches being applied for urban air pollution problems, although a similar methodology was applied in Kelly et al. (2011) in a study of the impacts of the Congestion Charging Scheme (CCS) on air quality in London. Model simulations using the King's College London Air Pollution Toolkit, with a typical output grid resolution of 20 x 20 m, were applied to estimate the percentage changes in $NO_x$ and $PM_{10}$ concentrations across the Congestion Charging Zone (CCZ).

Using sensitivity analysis, the attributed changes over a North-South transect of the CCZ due to changes in traffic speed, traffic flow and speed, and vehicle-kilometres travelled (VKT), for cars, taxis and buses were estimated. The modelling and sensitivity studies attempted to separate effects due to the vehicle related interventions from those due to changes in meteorological conditions or influences from changes to regional background pollution levels. The detected changes in both the model and observations were fairly small, but some consistent trends were detected between the model and observations, the most notable being an increase in $NO_2$ concentrations at three background sites compared with levels at the control sites. The use of the model sensitivity analysis allowed these effects to be attributed to a separate intervention that introduced more diesel particulate filter equipped buses into the CCZ.

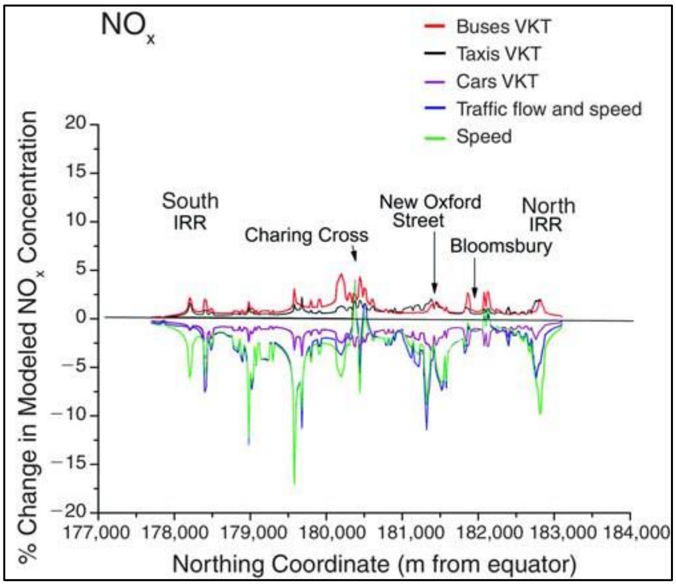


*Figure 3 The modelled impacts of the CCS on NOx concentrations across the CCZ. Graphs are percent changes in NOx concentrations due to changes in speed (green), traffic flow and speed (blue), VKT for cars (purple), VKT for taxis (black), and VKT for buses (red). IRR at each end ofthe graph is the boundary for the CCZ.*

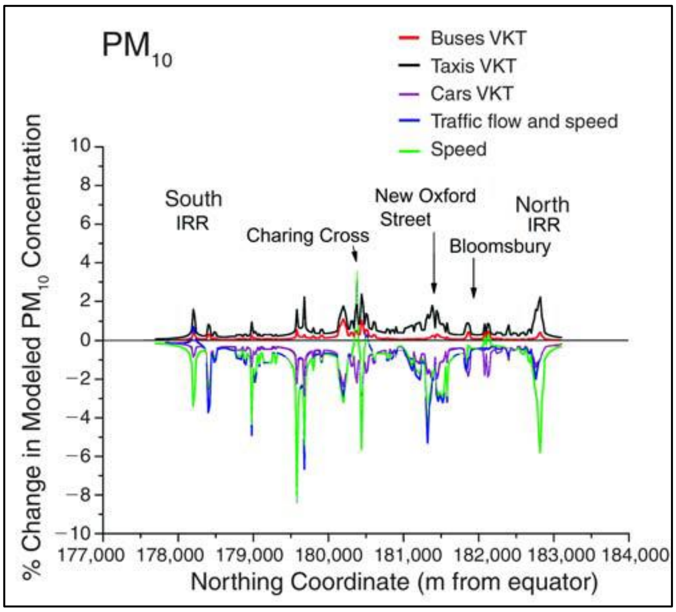


*Figure 4 The modelled impacts of the CCS on $PM_{10}$ concentrations across the CCZ. Graphs are percent changes in PM10 concentrations due to changes in speed (green), traffic flow and speed (blue), VKT for cars (purple), VKT for taxis (black), and VKT for buses (red). IRR at each end of the graph is the boundary for the CCZ.*

The main difference between the study of Kelly et al. and the methodologies applied within the climate change communities is the treatment of model uncertainties. Whilst the study of Kelly et al. included some sensitivity analysis with respect to emissions related factors, the model simulations did not carry error bounds and therefore it is difficult to assess the statistical significance of the modelled changes due to the CCZ and other policy related interventions.

# 3. Case Studies

This section considers some of the practical examples that aim to assess interventions. The list of interventions, while not definitive, cover a wide range of examples and techniques that have been used. There are many examples of how interventions have been considered, both in academic research and as part of various projects. These studies provide useful information on the types of approach used, their benefits and disadvantages and also can act as inspiration for conducting future analyses of interventions.

## 3.1. London Low Emission Zone

In 2008 London created the UK's first low emission zone (LEZ). The London zone is the world's largest and was the first built around a charging structure for non-compliant vehicles (thereby allowing occasional access) rather than a blanket ban. The scheme was targeted on $PM_{10}$ reduction but was also expected to have some impact on emissions of $NO_x$. Unlike London's congestion charging scheme, it operated at all times of day. The London LEZ is one of the few examples available where consideration was given to the experimental design before it was implemented and the location and types of measurements that would be made.

Unlike many schemes in Europe, the London LEZ scheme did not affect private cars. Instead, the focus was on heavy-duty vehicles reflecting their large emission per vehicle and pre-scheme evidence that London's heavy good vehicle fleet was amongst the oldest in the UK (TfL, 2008).

The zone was designed to broaden in scope over time. Phase one of the scheme in February 2008 applied to heavy goods vehicles > 12 t. Phase 2 from July 2008 extended the scope to coaches and goods vehicles greater than 3.5 t, covering around 5% of traffic. Phase 3 increased the scope still further to include light goods vehicles and mini-vans, a further 13% of traffic. This was postponed from 2010 to 2012 when it was implemented along with phase 4, which tightened the emissions standards from Euro III to Euro IV or retrofitted equivalent for PM.

The scheme was enforced using a network of automatic number plate recognition (ANPR) cameras.

### 3.1.1. Assessment strategy

An extensive assessment strategy was put in place (TfL, 2008). This assessment identified three levels of impacts from the scheme, broadly reflective of the HEI accountability chain shown in *Figure 1* (HEI, 2003).

Primary impacts were those directly attributable to the scheme; the improvement in the emissions characteristics of the vehicle fleet through the exclusion of the non-compliant Euro classes. Secondary impacts related mainly to air quality; both emissions and concentrations. Tertiary impacts were more societal and included those on health and businesses. A fourth set of impacts related to the area outside London through the possible accelerated fleet renewal or the displacement of non-compliant vehicles. Being mindful of the previous difficulties in detecting an air quality change from the earlier congestion charging scheme in 2003 (Kelly et al 2011a), it was stated at the outset that impacts from the scheme would be harder to detect at each stage in the hierarchy due to the influence of other confounding factors.

Primary impacts were assessed from a bespoke network of around 100 ANPR cameras supported by manual traffic surveys. The assessment of secondary impacts on air pollution is detailed in Barratt et al (2009), TfL (2008) and Kelly et al (2011b) are summarised below.

- Emissions estimates were based on measured traffic flows and vehicle Euro classes within the framework of an upgraded London Atmospheric Emissions Inventory.
- Seven roadside indicator sites were created to measure air pollution at locations identified as experiencing greatest changes from pre-scheme modelling. These included Marylebone Road and five other sites in the London Air Quality Network (LAQN), that were augmented to measure a common suite of pollutants: $PM_{10}$, $PM_{2.5}$, $NO_x$, $NO_2$ and $O_3$ (to assess possible changes in primary $NO_2$). A seventh monitoring site was installed alongside one of the main arterial roads with greatest predicted change. Meteorology was measured at five sites and three sites measured particle number and black carbon. ANPR cameras and traffic loop counters were installed alongside each site. The new equipment was installed in the year before the LEZ began.
- A further 13 long-term background measurement sites were selected from the LAQN to be used in the analysis to account for confounding changes in regional concentrations of these pollutants.
- PM receptor analysis was undertaken to identify concentrations from primary traffic emissions and those from regional and background sources. This analysis was undertaken at each indicator site and 16 other locations using the $NO_x$ as a tracer for primary PM as detailed in Fuller et al (2002, 2006). Detailed chemical analysis of $PM_{10}$ was undertaken at two locations in 2008 to support analysis of future LEZ phases.
- New data analysis methods were developed including the use of wind speed and direction data to focus on concentrations arising from roads sources. The assessment sought to control for the effects of meteorology by considering

background concentration and approaches that focused on the traffic derived pollution.

### 3.1.2. Key findings

Impacts of phases 1 and 2 of the London LEZ were published in TfL (2008, 2010), Barratt et al (2009) and Kelly et al (2011b). The main findings were:

- Substantial pre-scheme compliance was found as vehicle operators upgraded their fleets ahead of the phase 1 and 2. As shown in Figure 5 this was especially clear in the six months before phase 2. In 2008, HGV turn-over rates were 40% compared to an estimated rate of 5% in 2007. Given that the LEZ bought about a gradual change in the vehicle fleet no change in total measured concentrations of $PM_{10}$ were found at the times that phases 1 and 2 were implemented. Due to substantial pre-compliance analysis focused on changes in the thirteen months before phase 1, between phase 1 and phase 2 and in the 17 months following phase 2. A separate study by Ellison et al (2013) found substantial upgrades and pre-compliance in the LGVs in the year ahead of the implementation of phases 3 and 4.

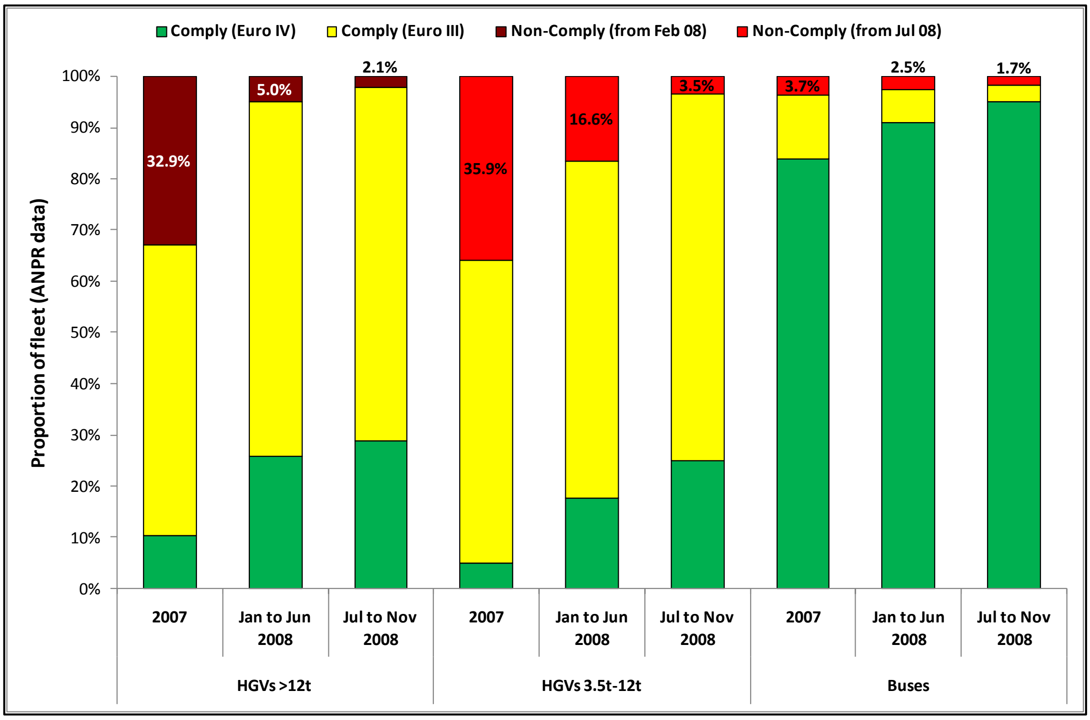


*Figure 5 Proportion of compliant and non-compliant vehicles at Hackney - Old Street in central London (Barratt et al 2009).*

- Techniques that focused on the concentrations associated with the road traffic using bivariate polar plots were used to detect changes across the LEZ implementation period. Decreases were found in black carbon and / or $PM_{2.5}$ from traffic alongside three roads in outer London (monitoring sites BT4, GR8 and TH4) which was not replicated in central London (MY1 and HK6). This was thought to be due to the different vehicle mixes in inner and outer London. A high proportion of buses and relatively low numbers of HGVs prevail in the central area. TfL buses were fitted with particle filters substantially ahead of the LEZ (2003 to 2006) and as shown in Figure 5, these were substantially compliant by 2007 before the LEZ. By contrast heavy diesel traffic in outer London was dominated by goods vehicles that were upgraded in the run-up to the LEZ.

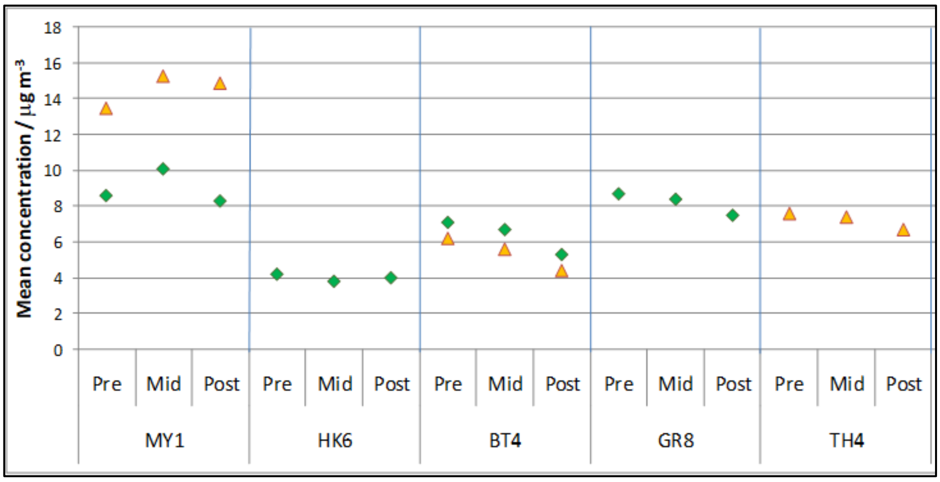


*Figure 6 Concentrations of black carbon (yellow triangles) and $PM_{2.5}$ (green diamonds) from local road traffic around the implementation of phase 1 and 2 of the LEZ (Barratt et al, 2009)*

- PM receptor analysis showed a steady increase in regional secondary and natural $PM_{10}$ between 2004 to mid-2007 was followed by decrease confounding any simple analysis of total measured concentrations. Very few of the 23 locations studied showed a change in primary $PM_{10}$ that was bigger than the confidence intervals for the model around the time of the LEZ implementation. Statistically significant increases in primary PM were seen at Marylebone Road.

The assessment benefited from the pre-existing very large and high-quality measurement network in London along with the long-term involvement of a research university in the emissions and dispersion modelling, measurement and assessment of air pollution providing London with a substantial skills base research infrastructure.

Robust impacts assessment was given a high priority and additional international research funding was obtained. Resources were provided for measurement infrastructure and analysis. The hierarchical approach to impacts assessment led to monitoring of primary effect of traffic changes and not just the secondary impact on measured concentrations. The study identified substantial pre-scheme compliance with HGV operators changing their fleets more than six months ahead of the scheme start dates. An assessment approach that looked for a step change at the implementation date would have underestimated any impacts. Unexpected differences were found between impacts of the intervention in inner and outer London due to the confounding effect of another policy to retrofit buses with particle filters. This was not foreseen in pre-scheme modelling.

A number of the assessment methods were able to produce results with confidence intervals allowing genuine change to be separated from expected variation due to measurement or modelling uncertainty.

The study design was not able to consider a clear counter-factual i.e. what the changes would have been without a LEZ. The zone covered all of London so separation into intervention and control areas was not possible. Major differences between the vehicle fleets in London (with high numbers of buses and very old HGVs) and other UK urban areas meant that a comparable control city could not be found. Instead, the assessment relied on timing, the comparison of traffic and background monitoring sites and receptor analysis to focus on the changes due to the intervention. It was not able to clearly separate the effects of the LEZ from other policies that were affecting vehicle emissions at the same time, such as the take up of lower emission vehicles due to natural fleet turn-over.  No assessment was made of changes in traffic flow. As shown later by Font and Fuller (2016), medium-term changes in the flow of HGV and buses along roads in London can outweigh gains achieved by improved emissions controls.

## 3.2. Impact of multiple policy interventions in London

It is challenging to clearly identify the impacts of a single intervention against a backdrop of the diversity of policies that have been implemented to improve air quality and reduce population exposure across the UK and beyond. Rather than trying to focus on a specific intervention and develop strategies to control for confounding impacts from other policies Font and Fuller (2016) took an agnostic approach. Instead of looking for the impacts of specific inventions they sought identify the locations where the policy mixture was working best and other locations where the policy mixture was working less well.

Most routine air pollution analysis assesses policy success by attainment of legal limit values. Instead, Fuller and Font (2016) considered the rate of change as a metric of policy success. Fuller and Font (2016) considered changes in the traffic derived air pollution alongside 65 roads in London from 2005 to 2015.

### 3.2.1. Assessment strategy

Measured concentrations of $NO_x$, $NO_2$, $PM_{10}$, $PM_{2.5}$, $CO_2$ and black carbon were obtained from 66 LAQN and AURN monitoring sites in London. The study looked at traffic derived air pollution by considering the roadside increment in concentrations. Traffic counts were also obtained. The study period was divided into two five year periods approximating to the periods of sale of Euro IV and Euro V vehicles. Robust statistical approaches were used to calculate trends and techniques from meta-analysis were used to present overall trends. Clustering techniques were used to group locations with similar results.

The assessment accounted for changes in concentrations due to meteorology by considering five-year windows, statistical de-seasonalising of measurements and also by control for background concentrations.

### 3.2.2. Key findings

This view of city-wide air pollution revealed considerable variability across London with some roads showing significant decreases but others did not improve. Examples include the notable improvement in nitrogen dioxide from traffic alongside Putney High Street which was attributed to a substantial programme to retrofit buses with Selective Catalytic Reduction abatement (see section 3.5). Places showing deterioration were mainly slower, more congested roads where SCR is thought to be less effective. Improvements in both $PM_{10}$ and $PM_{2.5}$ were found in central London. However, outer London showed an increase in $PM_{10}$ despite decreases in $PM_{2.5}$ and black carbon. These findings were especially clear on roads with increased numbers of HGVs. Changes in bus flows were associated with changes in traffic-derived $NO_x$ and $CO_2$ and to a lesser extent $NO_2$ and PM.

The study concluded that current policy packages were not strong enough to counter changes in traffic flow nor bring about decreases in $NO_2$ and $PM_{10}$ from roads in all parts of the city.

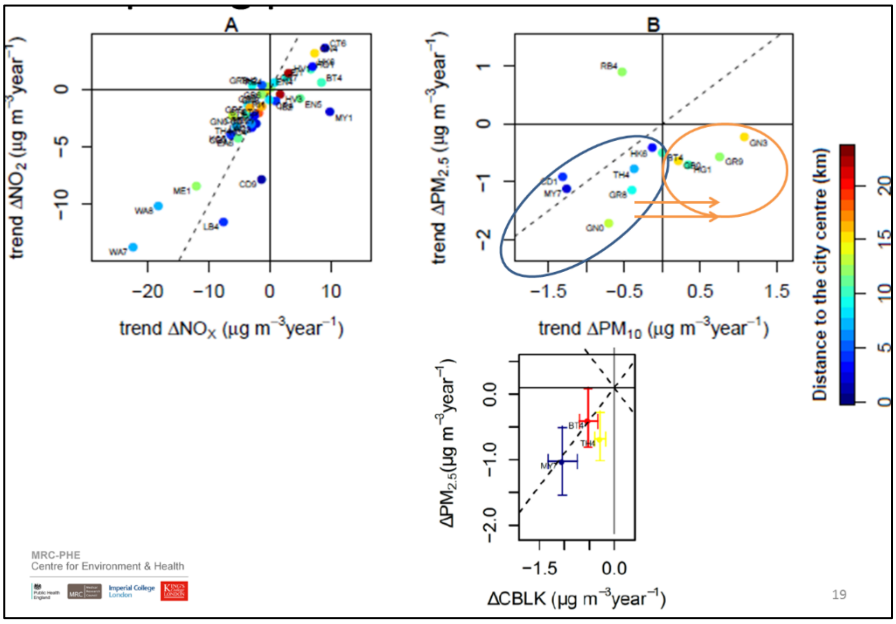


*Figure 7 Trend in road contribution between 2010 and the end of 2014. The top right-hand corner of panel A shows trends in $NO_x$ and $NO_2$ with those locations in the top right-hand corner experiencing increases in both pollutants. Panel B shows changes in $PM_{10}$ and $PM_{2.5}$ with the inset showing $PM_{2.5}$ and black carbon (CBLK). Colours denote distance from city centre.*

Overall, $NO_2$ increased between 2005 and 2010. Since 2010 most roads showed some improvement in $NO_2$ with an average decrease of five per cent per year but around three-quarters of roads still exceeded the $NO_2$ EU Limit Value in 2015. $PM_{2.5}$ decreased after 2010 due to the exhaust abatement technologies such as particle filters on newer diesel vehicles. There was little overall change in $PM_{10}$ from traffic.

The assessment focused on rate of change in both air pollution and the traffic activity. Many air pollution changes were associated with changes in traffic flows rather than exhaust abatement. The approach could simultaneously quantify changes at a city-wide level and on a road by road basis. All estimates of change were produced with confidence intervals enabling actual change to be separated from random variation. The approach could be adapted to look for the effects of a local intervention by comparing rates of change across a wider area.

The assessment method requires an extensive measurement network covering not just an area with a specific intervention and a control area but many more locations. It has applicability to large city networks such as those in Paris and London or at a national level. The study was the first stage in a larger proposed programme to initially identify and then explain concentrations changes. It therefore did not include differentials in fleet-turn over and Euro class along the different roads and the effects of different exhaust abatement upgrades to London’s bus fleet along each route.

## 3.3. German Low Emission Zones

There are many examples of interventions from German cities that include both low emission zones and traffic bans (Fensterer et al., 2014; Morfeld et al., 2015, 2014). These types of intervention have many similarities to the potential approaches that could be adopted in CAZ cities. Of interest and relevance are the city-scale changes to vehicle technologies and the influence they have on emissions and concentrations of pollutants.

Rather than focusing on the effectiveness of individual city LEZs, researchers pooled effects across many cities using statistical modelling tools that compared the changes in pollution concentrations in cities with a LEZ and those without. This approach takes account of changes such the larger scale changes to concentrations not associated with the LEZ, such as the ongoing turnover of vehicle stock.

Wolff (2013) tested two different strategies to match LEZ cities with appropriate control cities; a match based on similarities in $PM_{10}$ concentrations and another based on proximity. A model was created to include the effects of meteorology and to pool the results from multiple cities. They also considered cities with action plans and compared these to LEZ cities and those with no plan. Vehicle registration data was also considered to investigate if the LEZ accelerated vehicle turnover and a further investigation explored the hypothesis that non-compliant vehicles would simply divert around the zone. LEZs were found to be effective, decreasing $PM_{10}$ concentrations by 9% and there was no evidence of diversion or that non-LEZ plans were effective.

Malina and Schiffer (2015) also looked at the effectiveness of German LEZ for $PM_{10}$ using a statistical modelling approach that pooled and compared LEZ and non-LEZ cites and compared different types of LEZ. Meteorological factors were also considered along with the temporary vehicle scrappage scheme introduced during the financial crisis with results expressed in terms of avoided premature mortality and monetarised health benefit. More stringent stage 2 LEZs were found to be three times as effective as stage 1 LEZs in decreasing $PM_{10}$.

Morfeld et al. (2014) provide a comprehensive assessment of assessing LEZs in 17 German cities. This study is interesting because it encompasses many of the issues considered in this report including: the choice of instrument used to measure concentrations, the study design, accounting for variations not due to the LEZs (such as the influence of meteorology), careful statistical analysis and the evidence from published emission factors.

The study benefitted from considering the pooled effects across several cities rather than considering single cities in isolation, which gave the assessment more statistical power. By carefully considering ‘index sites’ and ‘reference sites’ (those within the LEZ and those outside the LEZ in each city, respectively), account was taken of the larger scale changes to concentrations not associated with the LEZ, such as the ongoing turnover of vehicle stock. The Morfeld et al. (2014) study is also notable for its comprehensive statistical approach to determining whether a concentration change had occurred and the likely magnitude of the change predicted. Even though the quantified changes to the concentrations, of NO, $NO_2$ and NOx were small (at most 2 µg $m^{-3}$, or about a 4% change at most), they were predicted to be statistically significant.

It is interesting that the Morfeld et al. (2014) study was undertaken at a time when changes to the diesel passenger car fleet would have had relatively small influences on emissions, and hence concentrations due to the lack of reduction in real-world emissions of $NO_x$. In that respect, the results from the Morfeld et al. (2014) study are consistent with the lack of progress in reducing emissions of $NO_x$ from Euro 1 to Euro 5 vehicles (Carslaw et al., 2011), as shown by the small estimated changes in concentration. These issues highlight how helpful it can be to have wider knowledge of the changes taking place and whether the changes are consistent with expectations given a knowledge of the other factors involved.

These statistical methods have the benefit of pooling results across many cities and controls to quantify the effectiveness of a type of intervention. This greatly increases the ability of the analysis to discern a change and also allows uncertainty to be estimated. As with many types of intervention analysis the matching of the intervention and a non-intervention control are key and this has produced debate among the German researchers working in this field as discussed in Morfeld (2015) and Cyrys et al (2017).

## 3.4. Short-term changes – Heathrow Airport

The quantification of interventions over short time scales is especially challenging because of the lack of data with which to assess the intervention.  Carslaw et al. (2012) used a statistical modelling technique based on boosted regression trees to consider the short-term impact of a flight ban at Heathrow airport. A flight ban was imposed at the airport for a period of 6 days due to the eruption of the Icelandic volcano Eyjafjallajökull. The flight ban provided an opportunity to consider whether changes in measured concentrations of $NO_x$ and $NO_2$ could be detected and quantified at measurement sites close to the airport.

The quantification of the changes in concentrations of $NO_x$ and $NO_2$ considered the counter-factual question as to what the concentrations would have been if there had been *no flight ban*. For this particular intervention, robust estimates of flight activity could be made because the number of flights and types of aircraft in operation is very similar each day. The statistical model was developed using aircraft movement data, meteorological observations and ambient concentrations for a period of two years before the flight ban operated. The model was then used to predict the expected concentration of $NO_x$ and $NO_2$ using typical hourly flight movement information. The approach yields a time series of $NO_x$

and $NO_2$ concentrations (with uncertainty) that can be compared to the measured concentrations, as shown in *Figure 8*. Overall it was shown that at the two nearest air pollution monitoring sites (LHR2 and Oaks Road), the concentration of $NO_x$ was predicted to be 29 µg $m^{-3}$ lower over the flight ban period compared with business as usual.

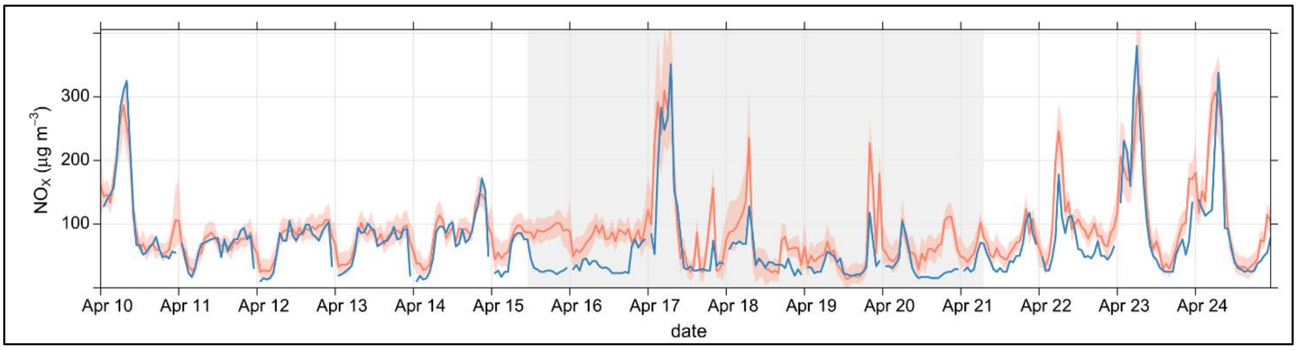


*Figure 8 Time series of $NO_x$ concentrations at the Oaks Road site. The blue line shows the observed concentration of $NO_x$, the red line is the predicted business as usual concentration with 95% confidence intervals shown. The grey shaded area shows the flight-ban duration.*

The modelling of the short-term ban also enabled an estimate to be made of the annual contribution made by aircraft to concentrations of $NO_x$ at the Oaks Road monitoring site to the south of the northern runway. Overall, it was predicted that the aircraft contribute 13.5 µg $m^{-3}$ $NO_x$, which is similar to the 12.0 µg $m^{-3}$ predicted through dispersion modelling. The 13.5 µg $m^{-3}$ corresponds to about 23% of the annual mean $NO_x$ concentration at the Oaks Road site. Furthermore, the pattern of change in $NO_x$ concentration (business as usual minus measured) was shown to be consistent with changes in aircraft sources, which helped to verify the validity of the analysis. Even though the duration of the flight ban was short (6 days), this work indicates that given good appropriate data, the potential impact of interventions can be quantified.

## 3.5. Changes in bus flows and technologies

Changing bus routes and the technologies used on buses are examples of interventions that are relevant to many local authorities. The influence that buses have on local air quality can be important because they can be an important contributor to emissions and traffic restrictions in urban areas can increase their importance (if for example traffic exclusions forbid other types of vehicle on certain roads).

There are several examples of where the specific influence of changes to the bus fleet have been quantified, including the London-wide analysis described in section 3.1, where locations with large changes in bus technologies show a different response to changes in $NO_x$, $NO_2$ and PM to other locations.

Earlier AQEG analyses of changes in $NO_2$ concentrations in Oxford using a cumulative sum (cusum) analysis revealed that changes in $NO_2$ concentrations were timed with known changes to either the flow of buses or changes in the emission control technologies used (AQEG, 2008). More recently, the detailed analysis of changes in concentrations of $NO_2$ at Putney High Street in London has shown that decreases in $NO_2$ concentrations at a roadside site were consistent with the TfL retrofit of their Euro III buses to use SCR (selective catalytic reduction) technology (Barratt and Carslaw, 2014; Carslaw and Priestman, 2015). The change in concentrations of $NO_2$ are shown in *Figure 9* together with information on when retrofit buses were introduced.

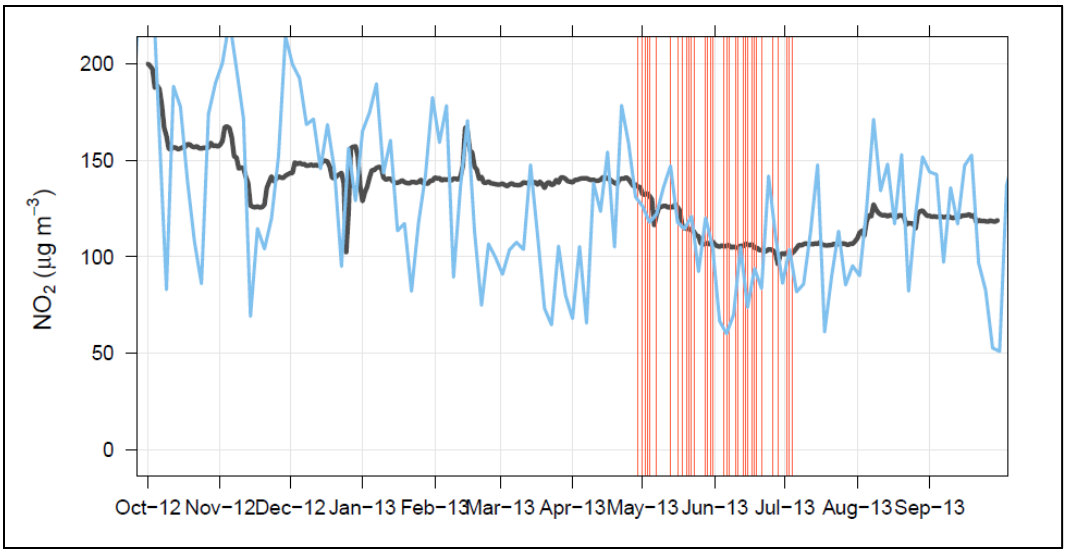


*Figure 9 Ambient concentrations of $NO_2$ at the Putney High Street LAQN site. The blue line shows the 3-day mean of the raw $NO_2$ data. The black line shows the meteorologically normalised $NO_2$ concentrations. The vertical lines show the dates when the SCRT system was fixed to individual buses.*

The analysis of the impacts at Putney High Street helps to illustrate several important aspects of quantifying the effects of a local intervention. First, it was very useful to have available information on when the buses were retrofitted (data available from TfL). Second, the ambient measurements were made at a kerbside location where the influence of the changes on ambient concentrations would be high. Additionally, the removal of the variations in concentrations due to changes in meteorology can greatly assist both the identification of when a change occurred as well as the magnitude of the change. The latter analysis does however require expertise on how to statistically analyse data to help remove the variation due to meteorology.

## 3.6. Fuel sulphur changes

Interventions related to fuel sulphur changes have been considered in two previous AQEG reports: Impacts of Shipping on UK Air Quality and the AQEG Report on Ultrafine Particles (UFP) (AQEG, 2017 and 2018).

In the first analysis, changes in the concentrations of $SO_2$ were considered at the Port of Dover. Changes in the concentration of $SO_2$ close to the Port of Dover might be expected because of legislation to control fuel sulphur content in ships. In this example, the analysis considered the change in concentrations of $SO_2$ at two monitoring sites close to the port. As discussed previously, it can be both difficult to detect changes and concentrations in the atmosphere and associate them with particular interventions. However, in this particular example, there are several aspects that make such an analysis more straightforward than many other case studies. First, concentrations of $SO_2$ at this location are dominated by a single source (shipping). Second, it would be expected that there would be a direct proportionate change in $SO_2$ concentrations associated with a fuel sulphur content change. Finally, there was a known date when the change might be expected to take place.

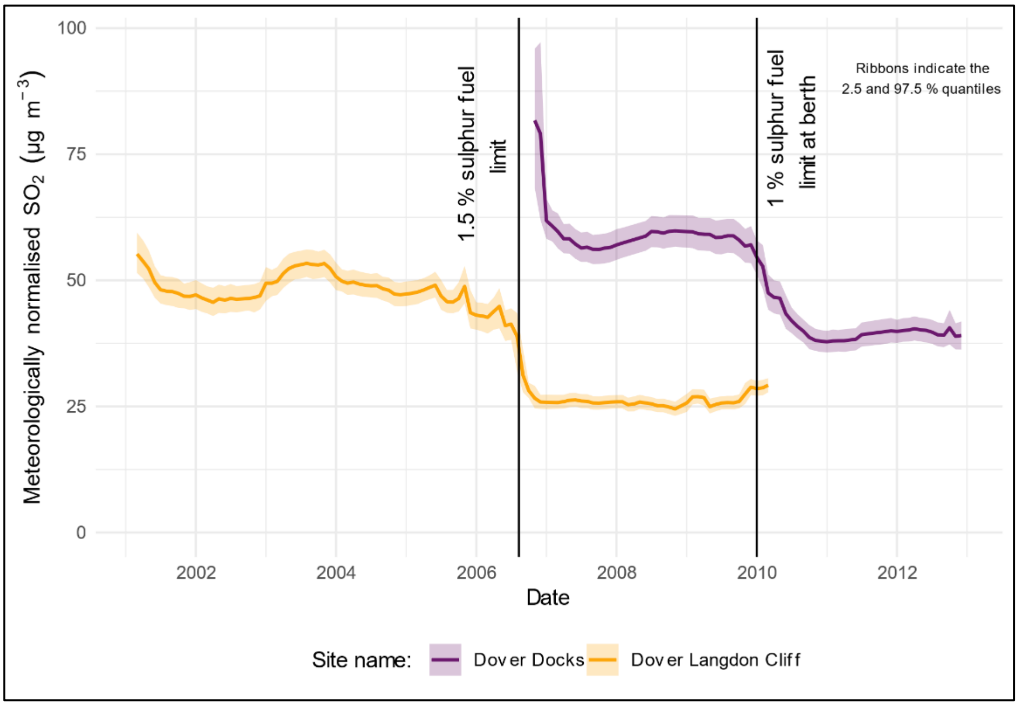


*Figure 10 The observed step-changes in $SO_2$ concentrations at Dover, coincident in timing with regulations reducing the maximum amount of sulphur in ship fuel used in the North Sea and English Channel in August 2006 and January 2010. The observations have been adjusted to remove the impact of variations in weather conditions on the trend in concentrations. The vertical dashed lines show the dates when new fuel sulphur regulations came into force* (Grange and Carslaw, 2019)*.*

Nevertheless, the analysis in the AQEG shipping report showed that accounting for changes in meteorology greatly increased the clarity of the nature, timing and magnitude of the change, as shown in *Figure 10*. In this case it was possible to show that the change in $SO_2$ concentration was very similar to the change in fuel sulphur limits and that the timing of the change matched with the date at which the legislation came into force. Therefore, in this example, there is very high confidence that the changes in concentration can be directly attributed to changes in fuel sulphur levels.

The second sulphur-related example is related to sulphur changes on vehicle fuels. In this case, the change in particle number (PN) concentrations were considered. PN concentrations are known to be strongly affected by fuel sulphur content and ambient concentrations of PN at roadside locations might be expected to reflect changes in concentration due to the change to sulphur-free (< 10 ppm S) road fuels. The analysis of ambient concentrations at Marylebone Road shown in *Figure 11* confirm there was a clear decrease in PN concentrations when sulphur-free fuels was first used in December 2007. Similar to the example for shipping, the timing and magnitude of the change are again much clearer once the effects of meteorology have been removed.

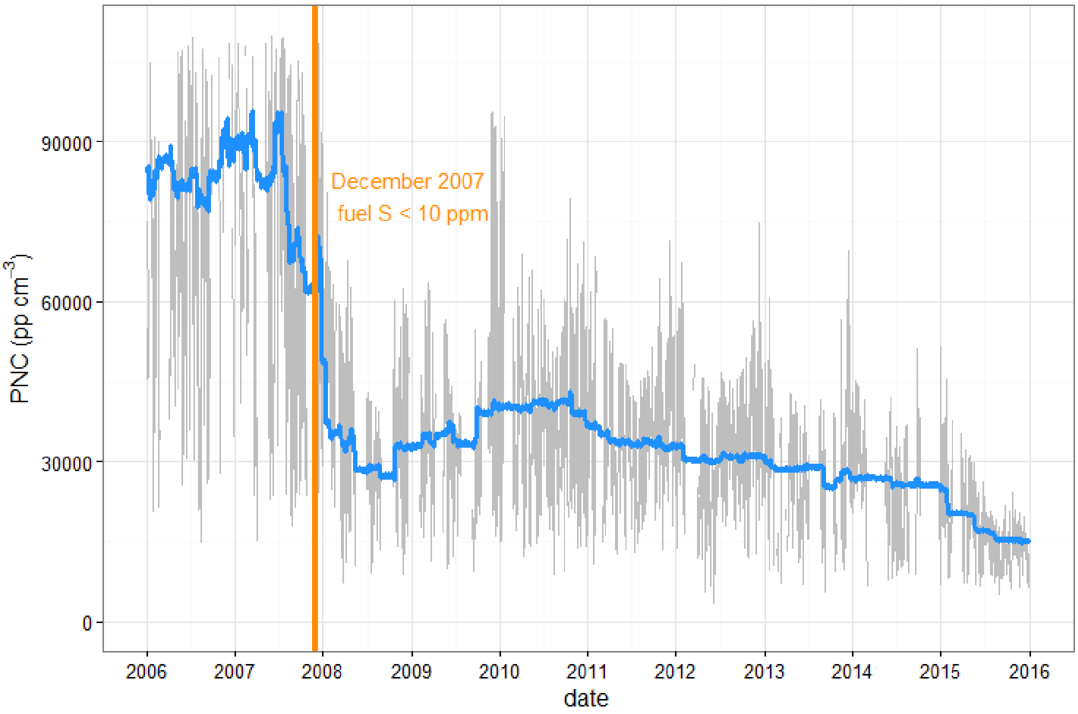


*Figure 11 Trend in PNC at Marylebone Road. The grey line shows the raw daily mean data and the blue line the trend with the meteorology removed. The vertical orange line shows the date when sulphur-free diesel and petrol were introduced.*

Both examples in this section will likely be easier to quantify and characterise than many of the other examples considered in this report for several reasons. These reasons include the influence of a single, dominant source, the magnitude of the change and the close proximity of where the measurements were made. Nevertheless, they do illustrate how concentration changes can be robustly linked to specific interventions.

## 3.7. The Dublin Coal ban

There are a number international examples of interventions where the impact has been followed though each stage of the accountability chain (Figure 1) from a change in polluting activity to a change in the health impact. One of the clearest is the Dublin coal ban of 1990. The Dublin coal ban was not motivated by the need to meet regulated air quality limits but by health evidence. In January 1982 an extra 54 inpatients died in St James' Hospital in the city when compared with previous years. Despite extensive investigations, no bacterial or viral cause could be found and investigation therefore turned to external factors. Home coal burning had increased a great deal in the city through the 1970s in response to increasing oil prices and encouragement from government grants. Concentrations of black smoke and $SO_2$ from the city council's monitoring network were found to be linked to increased in-hospital deaths (Kelly and Clancy, 1984). Rather than follow the UK route of declaring smoke control areas and upgrading home fires and boilers, the city council simply banned the sale, marketing and distribution of bituminous coal. This ban required people to burn smokeless or other fuels. Following the scheme in Dublin, the bituminous coal ban was extended to 11 more Irish cities.

Daily measurements of black smoke and $SO_2$ were obtained pre and post ban. Other data sets included temperature, deaths and hospital admissions along with fuel sales. Two assessments were made:

- The first investigation considered Dublin only. Data from five years before and after the ban were compared. No adjustment for meteorological changes were made to the air pollution data but this was accounted for in the epidemiological analysis along with changes in health metrics over the whole of Ireland (Clancy et al 2002).

- A second investigation of the coal ban was undertaken by the US Health Effects Institute. This study considered Dublin and the 11 other cities again using data from five years before and after each city / town ban. A comparison population from the peat burning areas of the Irish Midlands was also used as a control. Air pollution

> data was taken from multiple monitoring sites in each city. Six measurement locations were available for Dublin and Cork, four in Limerick and between zero and two in the other cities. Student *t* test statistics were used to compare the pre- and post-ban measurements and to determine the significance of any changes. (Dockery et al 2013).

Fuel sales were used to calculate the emission changes from the ban in Dublin. It was estimated that black smoke emissions reduced by 69% and $SO_2$ by 35%. Most of the black smoke changes (58%) arose from house holders switching from bituminous coal to smokeless and the remainder of the change was attributed to switches from solid fuel to natural gas heating (Clancy et al 2002).

The initial study in Dublin found that wintertime black smoke decreased by 70% and $SO_2$ decreased by 33% compared to the pre-ban period. This first analysis found that deaths from respiratory problems dropped by 16% and cardiovascular deaths fell by 10%. Across the city there were 116 fewer respiratory deaths and 243 fewer cardiovascular deaths per year (Clancy et al 2002).

The second wider study found decreases in black smoke between 45 to 70% but there were no clear changes in $SO_2$ across all 12 areas. The change in respiratory deaths found in the initial Dublin study were confirmed and similar changes were also found in Cork. The change in cardiovascular deaths found in Dublin were attributed to wider changes in health care and Ireland’s economy at that time (Dockery et al 2013).

The analysis is a great example of an intervention being assessed though the chain from a change in activity (in this case fuel use), emissions, and concentrations to health impact. The analysis strategy was based on a comparison data for five years of pre- and post-intervention. This approach constrained the analysis to the measurement infrastructure that was in place before the ban and induced a long time lag before the impact of the intervention could be reported. The inconclusive result for $SO_2$ might have been due to the lack of specificity for the $SO_2$ ‘bubbler’ measurement technique. The comparison between the health findings for the first and second analysis underlines the importance for controlling for wider societal changes in any assessment and highlights the comparison of before and after is not always sufficient to determine causality.

## 3.8. Case study: ‘APEC Blue’, Beijing

One of the most intensively-studied large-scale interventions in recent years were the controls enacted for the Asia-Pacific Economic Cooperation (APEC) Economic Leaders' Meeting in Beijing during 1-12 November 2014 (with the summit itself taking place on 10-

12 November) (Wang et al., 2016b). These followed the pollution-cutting measures enacted during the 2008 Olympics and included measures such as restricting vehicle use, limiting emissions from sources such as coal burning and temporarily shutting down certain industries. These measures applied to not just Beijing but the surrounding provinces as well, as much of the pollution in Beijing is recognised to be regional in nature. A dynamic approach was also adopted whereby additional measures were implemented if deemed necessary. For instance, weather forecasts indicated that for a period from the 4th of November, winds would be from the southwest which both air quality forecasts and empirical models predicted would be conducive to poor air quality in Beijing. In response, additional measures were enacted within 10 cities upwind of Beijing, such as shutting down certain industries and reducing power generation by 50%.

Pollutant concentrations were very intensively monitored during this period using a wide range of state of the art instruments in and around Beijing, ranging from conventional static monitoring sites, tower-based measurements, ground-based remote sensing and satellite observations (e.g. Wei et al., 2016; Tang et al., 2015; Tao et al., 2016; Huang et al., 2015; Chen et al., 2015; Wang et al., 2015). It is universally accepted that there was a very noticeable decrease in all key pollutant concentrations (except $O_3$, which increased slightly) for the duration of the summit and the unusually blue skies experienced gave rise to the nickname of ‘APEC Blue’. The scientific community within China has undertaken a very detailed interrogation of the data and there have been dozens of peer-reviewed papers published on the topic from a number of different institutes, often in conjunction with analysis of a subsequent intervention that was staged in September 2015 for the Victory Day Parade marking the 70th anniversary of the end of World War 2 (similarly nicknamed ‘Parade Blue’) (Li et al., 2016a).

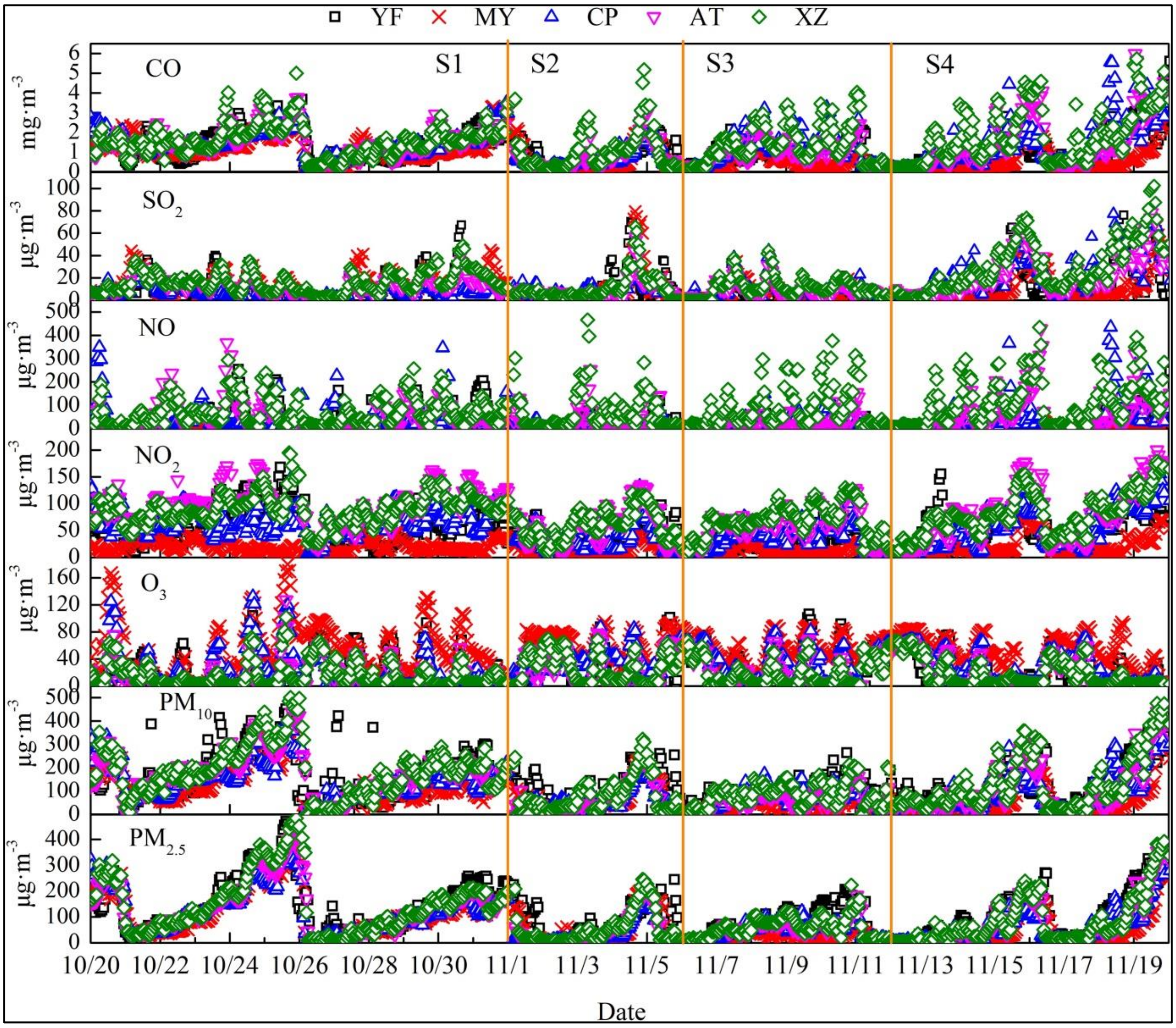


*Figure 12: Key pollutant concentrations from 5 monitoring sites in Beijing, with the summit occurring towards the end of phase S3 (Wang et al., 2016b).*

While the decrease in pollutants is not in doubt, the major overarching scientific goal of the analyses was to not just assess the effectiveness of the short-term and rather drastic interventions, but also to use these case studies to quantify the contributions of various sectors to pollution in Beijing, such that more long-term and sustained policies and interventions can be formulated. However, during analysis, it quickly became apparent that the reduction in pollutants could not solely be attributed to the control measures alone, as the meteorological conditions during the period had coincidentally been very favourable from an air quality perspective. Because of local geography, Beijing can experience some very extreme and abrupt transitions in air quality in response to changes in weather conditions and quantitatively unpicking the ‘signature’ of the temporary controls from this has proved extremely challenging for the scientific community. A wide variety of methods have been employed ranging from regional modelling of pollutants to empirical statistical analysis of the data and the quantitative attribution of the various factors is a topic of much debate within the scientific community (e.g. Guo et al., 2016; Zhai et al., 2016; Li et al., 2016b; Wang and Dai, 2016; Zhang et al., 2016; Wang et al., 2017; Wang et al., 2016a). The broad consensus is that once meteorology is taken account of, the emissions controls did still have a significant effect, however conclusions differ when it comes to attributing it to specific aspects of the controls.

Within the wider context of the assessment of interventions, this highlights the importance of taking a systematic and measured approach to data analysis rather than jumping to conclusions based on favourable data. While the intended result of the controls – a reduction of pollution during the APEC summit – did occur, it is fair to say that the outcomes of the scientific analyses that would usefully inform policymaking have not been as clear-cut as initially thought. It is worth highlighting that while very significant in nature, these interventions were relatively short-lived, so in scientific terms the case study may intrinsically lack the statistical power needed to satisfactorily overcome confounding issues such as variations in meteorology, in spite of the extremely rich observational dataset accompanying it. Hypothetically, a smaller-scale intervention that lasted for a longer period that could capture a wider variety of meteorological conditions could conceptually perform better. But whichever way, the collective effort that has been directed at this problem underscores the value in performing a comprehensive and multi-faceted approach to analysing the data available.

## 3.9. Newcastle Air Quality Management Area

The assessment work by Newcastle City Council is fairly typical of that carried out under the Local Air Quality Management (LAQM) regime. Monitoring and modelling carried out by the Council identified exceedances of the UK objective for annual mean $NO_2$ concentrations and an Air Quality Management Area (AQMA) was declared. An action plan of measures to reduce $NO_2$ concentrations was then formulated (Newcastle City Council, 2006 ). The action plan measures implemented prior to 2016 included: 20 mph speed restrictions, parking controls, bus priority enforcement, removal of specific bus types from key roads, a bus retrofit programme, increased electric vehicle infrastructure, use of electric vehicles for Council services, and promoting electric car clubs (Newcastle City Council, 2017).

The Council reported on progress with its action plan following the LAQM regime; providing implementation dates of all completed measures, as well as planned completion dates for pending measures. It also collated recent monitoring results from sites within the AQMA. This included 62 diffusion tube sites and four chemiluminescence samplers (including one AURN site), although the number of diffusion tube sites was reduced to 35 in 2015.

Annual mean concentrations were presented as time-series stretching from 2012 until 2016 (*Figure 13*). It was highlighted that there were continued exceedances of the objective at a large number of sites. From a visual interpretation of the data reproduced in *Figure 13*, it was concluded that there had been little, if any, reduction in $NO_2$ concentrations since 2012 (Newcastle City Council, 2017).

Ultimately, the effectiveness of the action plan is being assessed in terms of whether or not there are continued exceedances of the annual mean $NO_2$ objective. No attempt has been made to assign a pattern of causation to any observed changes or to quantify measure-specific effects. This fulfils the requirements of the LAQM regime, but provides no additional information as to whether individual measures may have had an effect, and thus might be useful elsewhere.

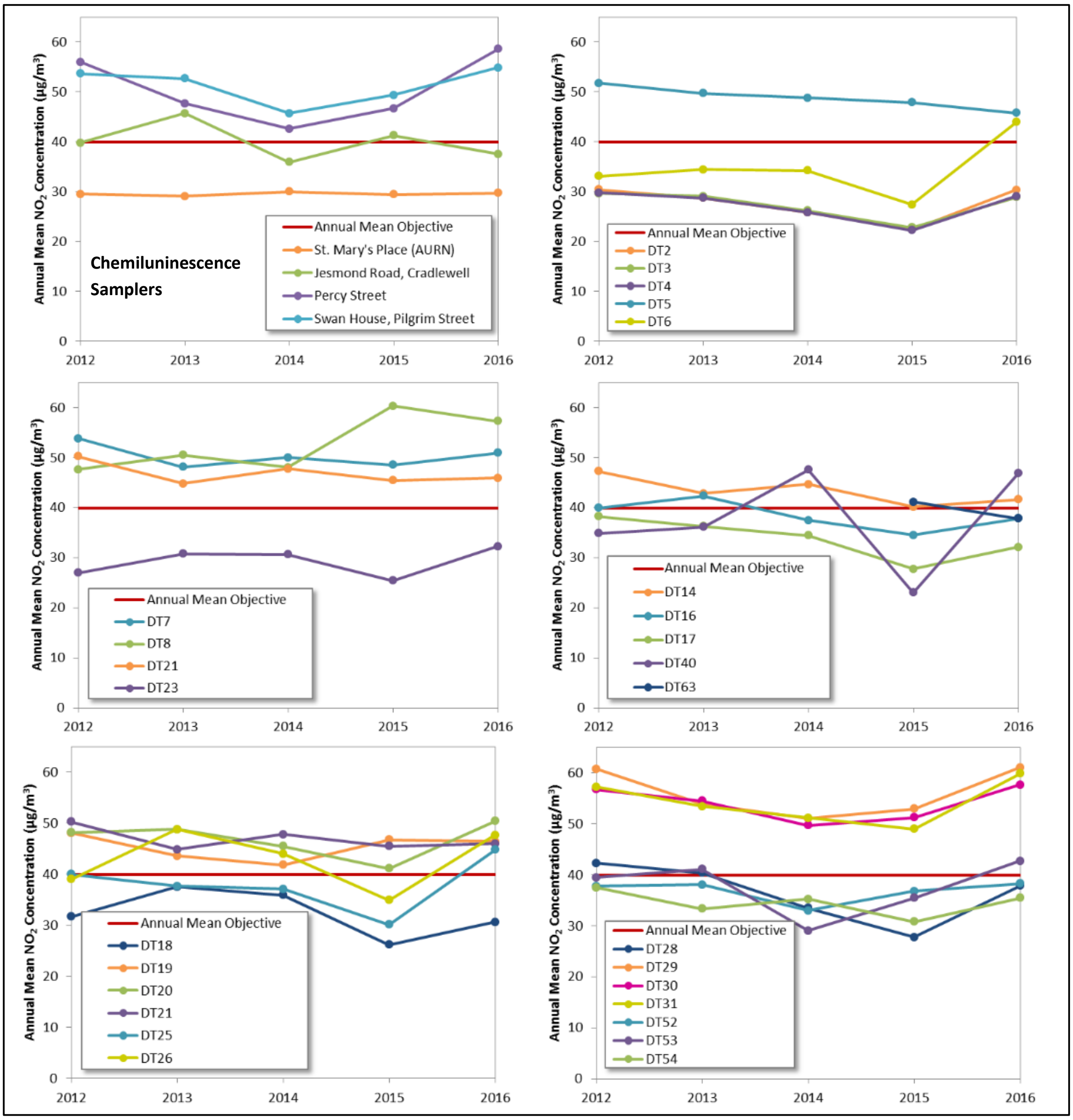


*Figure 13 Measured Annual Mean $NO_2$ in Newcastle City Centre AQMA over Five Years. All charts except the upper left show the results from diffusion tubes. Taken from Newcastle City Council, 2017.*

Most assessments carried out by local authorities under the LAQM regime have shown similar results to those of Newcastle City Council, and have not shown any clear effects of local-scale interventions. Reasons for this include:

1) there is no explicit LAQM requirement to quantify the effect of interventions;
2) robust analyses can be difficult; requiring expertise at both the experimental design and data analyses stages which is often unavailable to local authorities;
3) the measures themselves have often been ineffective; mainly because of reliance on vehicle type-approval emissions standards that have not delivered the improvements that Local Authorities were instructed, by Defra, to expect;
4) in order to expedite effects, multiple interventions have often begun concurrently. This makes it difficult to assign a cause to any changes in measured concentrations;
5) the spatial scale over which an intervention may have an effect is often poorly understood, making it difficult to define suitable counter-factual monitoring sites;
6) while long-term monitoring sites are valuable, reliance on pre-existing sites means that they were not located specifically for the purpose of assessing the intervention;
7) most local authority monitoring uses diffusion tubes, providing 1-month average concentrations with considerable month-by-month uncertainty. These data are not ideal for analysing what can be quite subtle changes in concentrations over time or between sites; and
8) notwithstanding Points 3-7, and for the reasons outlined in Section 2.4.1, the *predicted* effects of many interventions have been too small to realistically measure.

## 3.10. A419 Blunsdon Bypass

The assessment work carried out for the A419 Blunsdon bypass scheme is typical of that provided for most highways interventions. The A419/A417 connects the M4 near Swindon to the M5 near Gloucester. The section of the route around Blunsdon, immediately north of Swindon, historically passed close to a number of residential properties. The Blunsdon bypass scheme sought to re-route traffic along a new section of road, with the intention of reducing congestion and, amongst other perceived benefits, improving air quality close to the original route.

During the design phase, the Highways Agency measured $NO_2$ concentrations over a three-month period at four sites using passive diffusion tubes. These measurements were adjusted to represent annual mean concentrations, which ranged from 19 µg $m^3$ well away from any roads to 58 µg $m^3$ at the kerbside (Highways Agency, 2005).

A predictive air quality assessment was carried out at the planning stage (Highways Agency, 2005). Using a simple screening model, reductions of up to 32% and 18% in annual mean concentrations of $NO_2$ and $PM_{10}$ respectively were predicted. Increased concentrations of both pollutants were also predicted at properties close to the bypass itself, but these were smaller than the predicted improvements.

The scheme obtained the necessary permissions and opened in 2009. Post Opening Project Evaluation (POPE) studies were completed at one year and five years after opening of the scheme. Both studies collated traffic data from automatic traffic counters operated by both the Highways Agency and Swindon Borough Council. Temporary traffic surveys were also commissioned. Analyses of national, regional, and local background traffic growth were also carried out. This allowed the effects of the scheme on traffic flows to be isolated from these, wider-scale, trends. No post-opening air quality monitoring was carried out.

The one year after completion study (Highways Agency, 2012) noted that daily average traffic flows on the bypassed road section had reduced by 93%, but remained 37% higher than had been predicted at the planning stage. Total traffic flows on the bypass itself were similar to the earlier predictions, although the numbers of heavy duty vehicles were lower. From these observations, it was inferred that the air quality improvements, while present, were not as large as had been predicted, and that air quality was likely to be better than expected in the immediate vicinity of the bypass itself. The five years after completion study (Highways England, 2015) showed that observed traffic flows on both the bypass and the bypassed section of road, were lower than had been predicted. The changes in traffic flows were used to infer that the air quality benefits were as expected at the planning stage, while the disbenefits were smaller than expected.

Swindon Borough operated a long-term diffusion tube monitoring site at the A419 in Blunsdon up until 2007, but this site closed before the scheme opened in 2009.

By focusing on observed changes in traffic flows, the POPE studies provided a simple and straightforward assessment of the effect of the intervention. In particular, the analyses of background traffic growth are analogous to some of the principals covered elsewhere in this report, but with fewer potentially confounding factors. This level of assessment is very useful, but does not accurately show the effect that the intervention has had on air quality.

Ultimately, in order to definitively demonstrate an effect on concentrations, local air quality monitoring would be required. There are, however, other steps on the accountability chain (Figure 1) between measuring traffic volumes and measuring ambient concentrations, and some form of post-opening modelling could, thus, be useful. It should, though, be recognised that the planning-stage modelling carried out for interventions such as this typically relies on average-speed vehicle emissions factors and national-level fleet compositions. This means that, with the exception of being able to take account of updated information which is unrelated to the intervention itself, repeating the same modelling with observed traffic flows might add little when compared with the more qualitative discussion of traffic flows that is already included in most POPE studies. There may, however, be other aspects of traffic flows which could usefully be reported and discussed. Routine observations made by Highways England can often be used to indicate congestion-induced acceleration, which is likely to have a strong influence on emissions from the strategic road network. Furthermore, a close interrogation of observed and modelled traffic data might provide fresh insights into journey purpose, which can relate to vehicle type and age. Thus, even without additional air quality monitoring, and without re-running the planning-stage predictive models, it is likely that a more targeted examination of those data which are already being collected by Highways England might provide fresh insights regarding the effects that its interventions have had. Such analyses might not, however, be straightforward to carry out on a routine basis.

## 3.11. London City Airport

One example of post-intervention monitoring being required through the planning system is at London City Airport. Over a number of years, the airport has received various planning permissions, which allow an overall increase in activity at the airport. Conditions attached to recent permissions have required the airport to implement an air quality action plan, which contains measures intended to reduce emissions from on-airport operations as well as from surface access. There is also a requirement to measure concentrations of $NO_2$ and $PM_{10}$ in the vicinity of the airport, which is done using two chemiluminescence samplers, one FDMS-TEOM, and at 18 $NO_2$ diffusion tube sites. Progress with the action plan, as well as results from the monitoring survey, are reported annually to the local planning authority.

The airport's 2017 report to the planning authority (London City Airport, 2017) summarised progress with implementing the air quality action plan. It was accompanied by an air quality monitoring report (Nunn et al., 2017), which provided an update on measurements made over the previous year.

The monitoring report compared the measurements to the air quality objectives, as well as the Daily Air Quality Index (Defra, 2017). The timing of 1-hour peak concentrations was also compared against time-series from other sites in London. From a visual examination of the plots reproduced in *Figure 14*, it was concluded that the timing of high and low concentrations was consistent across all sites and that the airport was thus not the main cause of the highest measured concentrations.

Openair software (Carslaw and Ropkins, 2012) was used to generate bivariate pollution roses, which were used to identify the relative contribution that different sources might be making to measured NOx concentrations. While ground-level off-airport, sources were associated with most of the higher measured concentrations, the influences of ground-level on-airport, as well as elevated emission sources, were also detected.

Openair was also used to determine multi-year trends in NOx and $NO_2$ at the two airside chemiluminescence sites, as well as at seven off-airport comparator sites. *Figure 15* shows the trend for $NO_2$, as well as the results from Thiel-Sen analyses for both $NO_2$ and NOx. Over the period shown, there were several changes that could be expected to have affected emissions from the airport, including changes to passenger and aircraft numbers, decommissioning older mobile ground power units, an increase in the use of fixed electrical ground power units, and improvements to the on-airport ground-vehicle fleet. It was noted that the multi-year trend for $NO_2$ and NOx at both airport sites followed the same overall downward pattern as at the comparator sites. Furthermore, the rate of change at the airport was broadly within the range observed at other sites. It was thus concluded that airport-related activities were neither driving, nor significantly offsetting, these trends.

This assessment was carried out to demonstrate whether or not the airport was adversely affecting local air quality in order to fulfil the specific requirements of a planning condition. The analysis does, to an extent, identify some of the main airport and non-airport contributions to measured concentrations, and also seeks to determine the net effect of all interventions made by the airport (which in this context might include increases as well as reductions in emissions). It does not, however, attempt to disentangle the individual effects of any of these interventions.

As with the LAQM example, this case study is characterised by multiple different interventions, all operating over different spatial scales, and taking effect at different times. In such a case, it can be challenging even to record the precise rate at which each measure is implemented. This makes it extremely difficult to determine the effect of any individual intervention. The position taken in this study is that there is no real purpose in attempting to disentangle individual effects, since it is the net effect, including national-level trends, that is the principal concern.

More generally in relation to the planning system, it is worthwhile noting that while mitigation is frequently specified through planning conditions, there is typically no requirement for accompanying pollution monitoring. In those cases where post-opening monitoring is required, there is usually no formal requirement to assess the effectiveness of the mitigation. Ultimately for many new developments, there is also is no practical mechanism for a planning authority to impose additional mitigation requirements in the event that the agreed measures are ineffective. For these reasons, detailed studies into the effectiveness of interventions are not typically carried out under the planning regime.

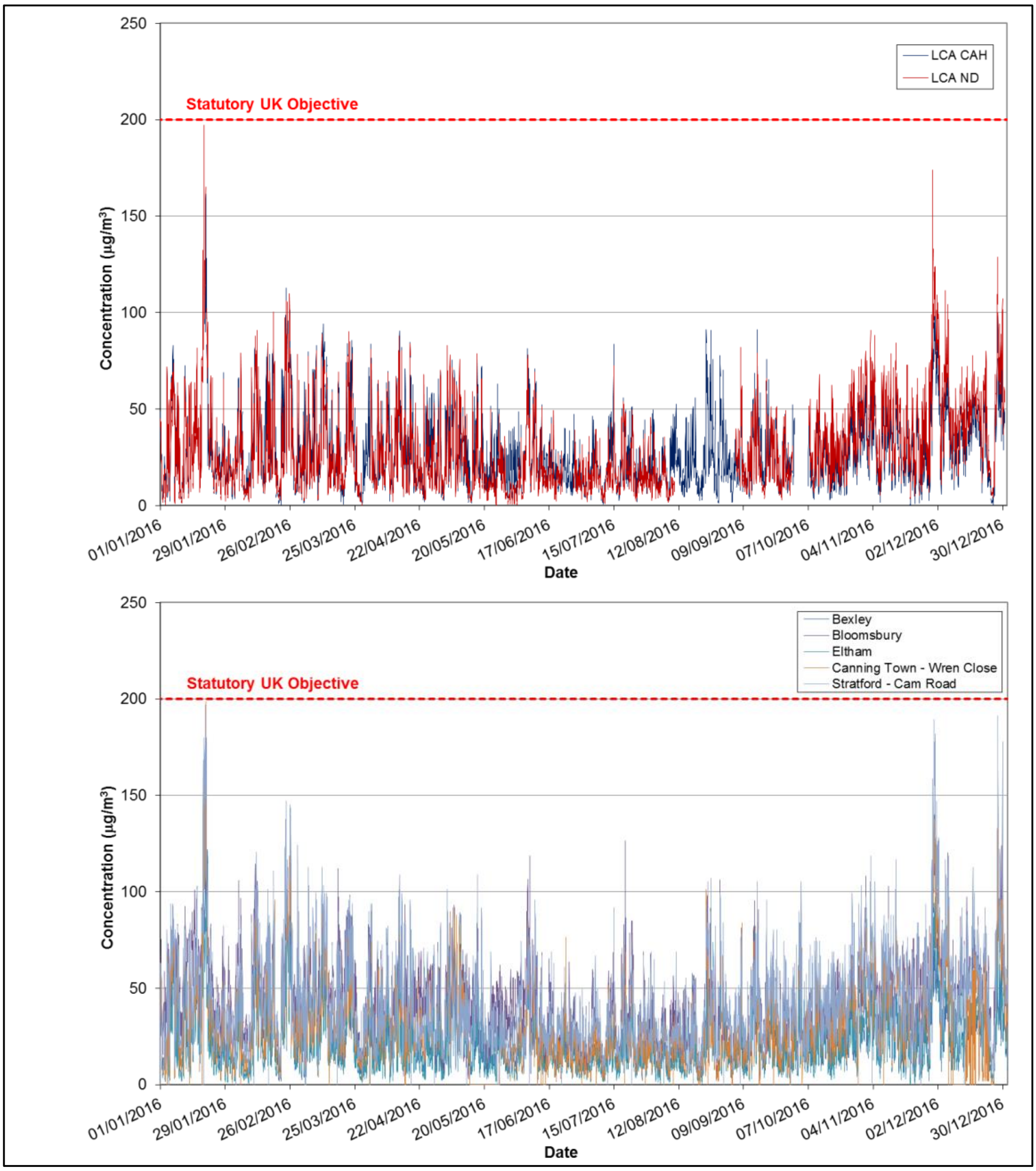
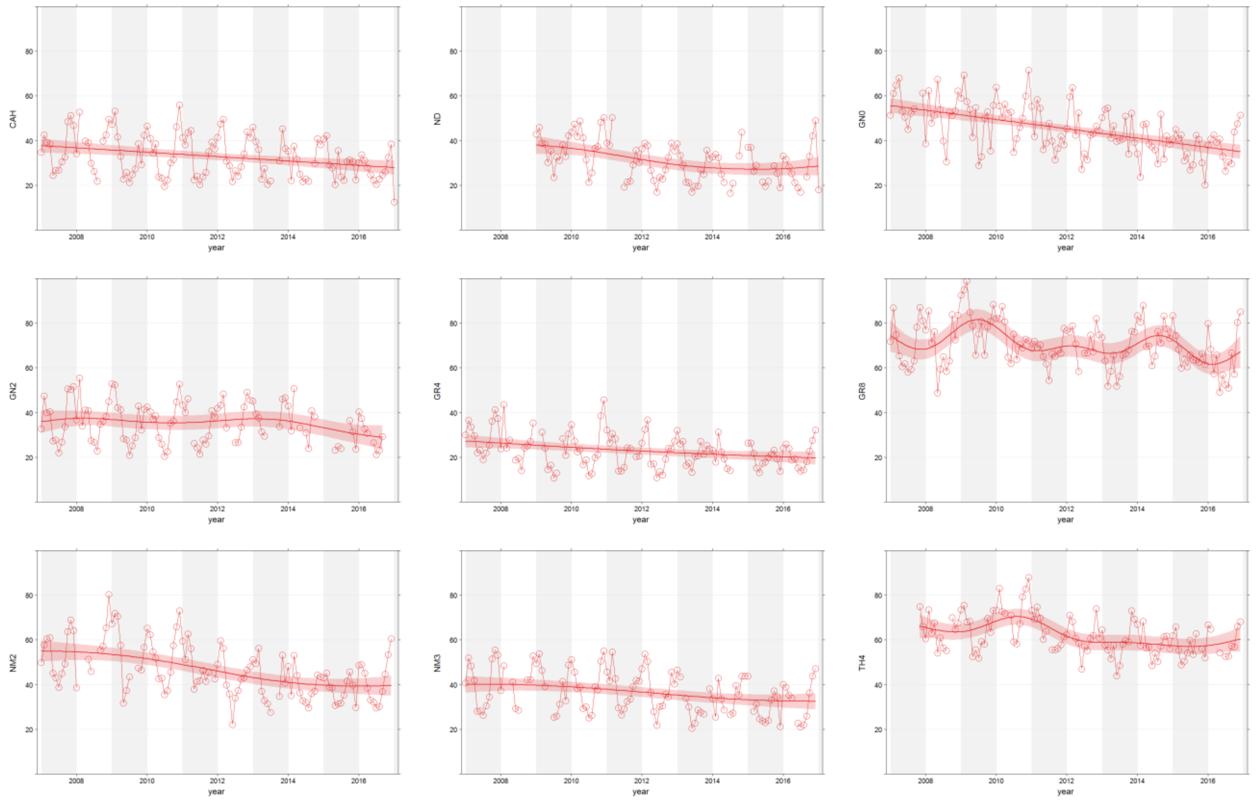


*Figure 14 1-hour mean $NO_2$ in 2016 at: Airport sites City Aviation House (CAH) and Newham Dockside (ND) (top chart), and Bexley, Bloomsbury, Eltham, Canning Town (Wren Close) and Stratford (Cam Road) (bottom chart)).*



| | $NO_2$ | NOx |
|---|---|---|
| City Aviation House (Airport) | -0.95 [-1.62, -0.43] | -1.79 [-3.17, -0.73] |
| Newham Dockside (Airport) | -1.68 [-2.45, -0.9] | -5.82 [-8.32, -3.58] |
| Greenwich Burrage Grove | -2.16 [-2.75, -1.54] | -5.25 [-6.91, -3.29] |
| Greenwich Millennium Village | -0.53 [-1.18, 0.13] | -1.72 [-3.9, 0.13] |
| Greenwich Eltham | -0.73 [-1.19, -0.28] | -0.99 [-2.04, -0.03] |
| Greenwich Woolwich Flyover | -1.02 [-1.78, -0.16] | -3.96 [-7.43, -0.5] |
| Newham Cam Road | -2.01 [-2.65, -1.4] | -4.74 [-6.49, -2.92] |
| Newham Wren Close | -1.15 [-1.78, -0.53] | -1.62 [-3.2, -0.31] |
| Tower Hamlets Blackwall | -1.2 [-1.88, -0.67] | -3.76 [-6.46, -0.61] |

*Figure 15 Smooth trend analysis, hourly $NO_2$ 2007 – 2016 at (Left to Right) City Aviation House (Airport), Newham Dockside (Airport), Greenwich Burrage Grove, Greenwich Millennium Village, Greenwich Eltham, Greenwich Woolwich Flyover, Newham Cam Road, Newham Wren Close, Tower Hamlets Blackwall. Circles show monthly means, the central red line is a smoothed trend calculated using generalised additive modelling, red shading shows the 95% confidence interval. Tabulated values show Theil-Sen results for both $NO_2$ and NOx (The first value is the slope, while the numbers in brackets are the upper and lower 95th percentile confidence intervals).*

# 4. Example of the role of air quality modelling

This section considers the use of the air quality model ADMS Urban to illustrate some of the important issues related to quantifying the effect of interventions. A simplified scenario is constructed where hourly concentrations of $NO_x$ at the London Kensington Knightsbridge site are considered. The concentrations of $NO_x$ at this site were modelled using ADMS Urban, with the road adjacent to the site subject to different emission reduction scenarios. Hourly concentration predictions are considered for 2012. In this example, the modelled road accounts for about 75% of the overall concentration of 192 µg m$^{-3}$. In practice, it would be likely there would be more than one year of data available to conduct such analyses.

Model predictions have been made for base case conditions and several scenarios that consider the effect of a reduction in the road $NO_x$ emission contribution from 1 July 2012 to 31 December 2012. Reductions in the road traffic contribution range from 5 to 50% in 5% intervals. These scenarios therefore reduce the emission of $NO_x$ by differing amounts in the second half of 2012.

*Figure 16* shows the daily mean $NO_x$ concentration at the Knightsbridge site for base case conditions ($NO_x$ reduction in road emissions) and the effect of a 20 and 50% reduction in road $NO_x$ emissions in the second half of the year. Daily mean concentrations are shown instead of hourly means to reduce the variability and highlight overall trends. It is very difficult from a consideration of *Figure 16* alone to discern any obvious reduction in $NO_x$ in the second half of the year, even with a 50% reduction on road traffic emissions.

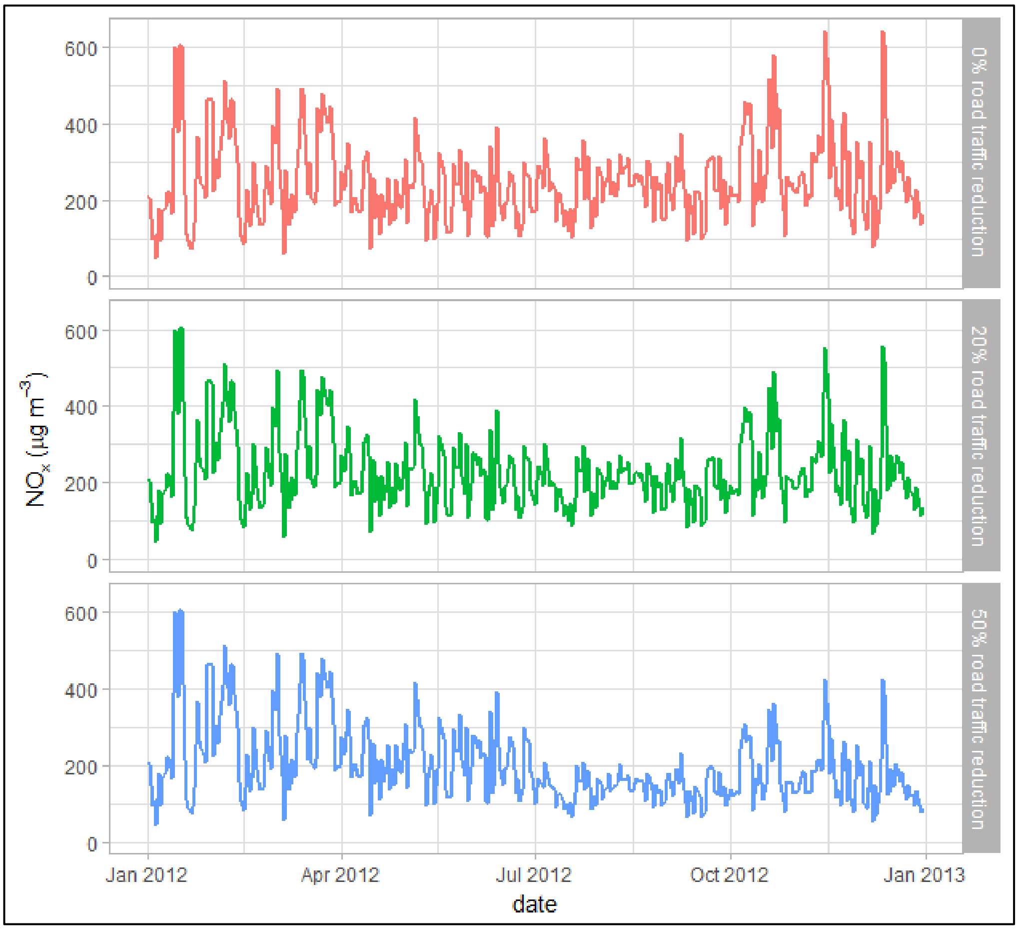


*Figure 16 Daily mean concentrations of NO$_x$ at the London Knightsbridge site for 2012 showing the base case concentrations (i.e. no road traffic reduction) and the effect of a 20 and 50% reduction in the road traffic NO$_x$ emissions in the second half of the year.*

It is useful to split the data into a ‘before’ and ‘after’ case where the before case is the concentration in the first half of the year and the after the concentration of NO$_x$ in the second half of the year. A comparison of the mean concentration and 95% confidence interval in the mean for the before/after pairs can be made as shown in *Figure 17*.

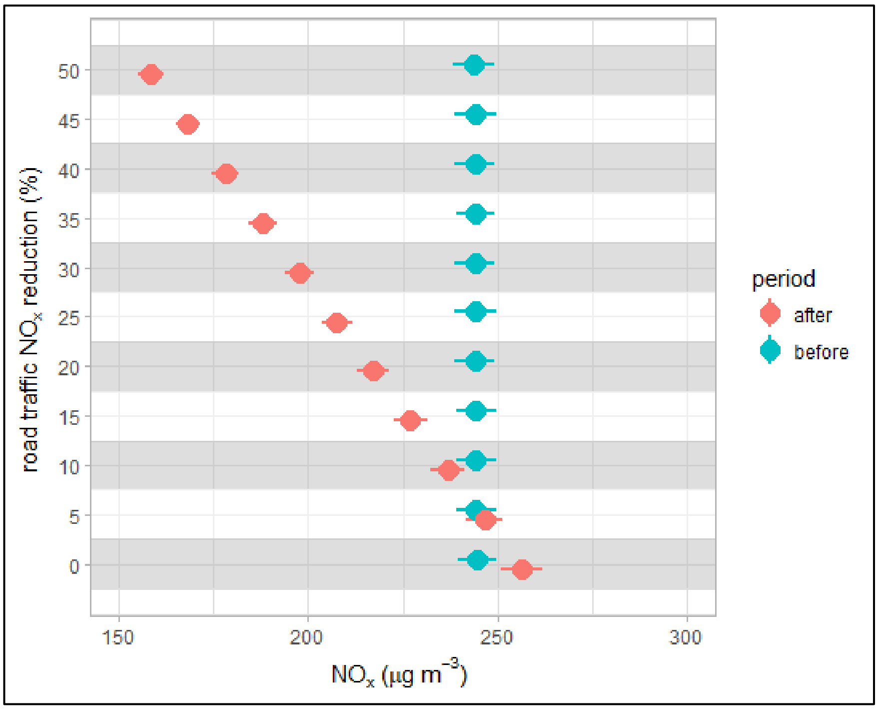


*Figure 17 Mean concentrations in the first half of the year (before) and second half of the year (after) for different road traffic $NO_x$ reduction scenarios. The error bars show the 95% confidence interval in the mean concentration. The 'before' concentrations remain the same for each scenario.*

The results shown in *Figure 17* illustrate some of the characteristics that are relevant to real interventions. The first is that considering the before/after concentrations of $NO_x$, the concentration of $NO_x$ for the after period is *higher* in the second half of the year than the first under base case conditions i.e. 0 % $NO_x$ reduction. These results show that due to meteorological variation, there are increased concentrations of $NO_x$ in the second half of the year that are not associated with changes in $NO_x$ emissions due to the intervention.

It is not until there is about a 15% reduction in $NO_x$ emissions that concentrations in the second half of the year are below those in the first half. From the raw data alone, it is not obvious that higher or lower concentrations of $NO_x$ in the second half of the year compared with the first are due to changes in emissions or changes in meteorology – or indeed both. Just because the concentration is lower in the 'after' period, it does not mean they are lower because of the intervention. This issue is central to the evaluation of any interventions based on the analysis of ambient measurements and is a recurring theme in many of the case studies considered in this report.

To illustrate some of the potential benefits of removing much of the variation due to meteorology, the analysis technique of Carslaw and Taylor (2009) has been used, as described earlier in the report. Briefly, statistical models were developed to explain the concentration of hourly $NO_x$ concentrations in terms of meteorological variables including wind speed, wind direction and ambient temperature based on hourly data available from London Heathrow. Once a model was developed, it was run several hundred times with randomly selected meteorology as input and the results averaged to provide a new predicted time series. The new time series effectively represents the trend in $NO_x$ under what might be thought of as ‘average’ meteorological conditions.

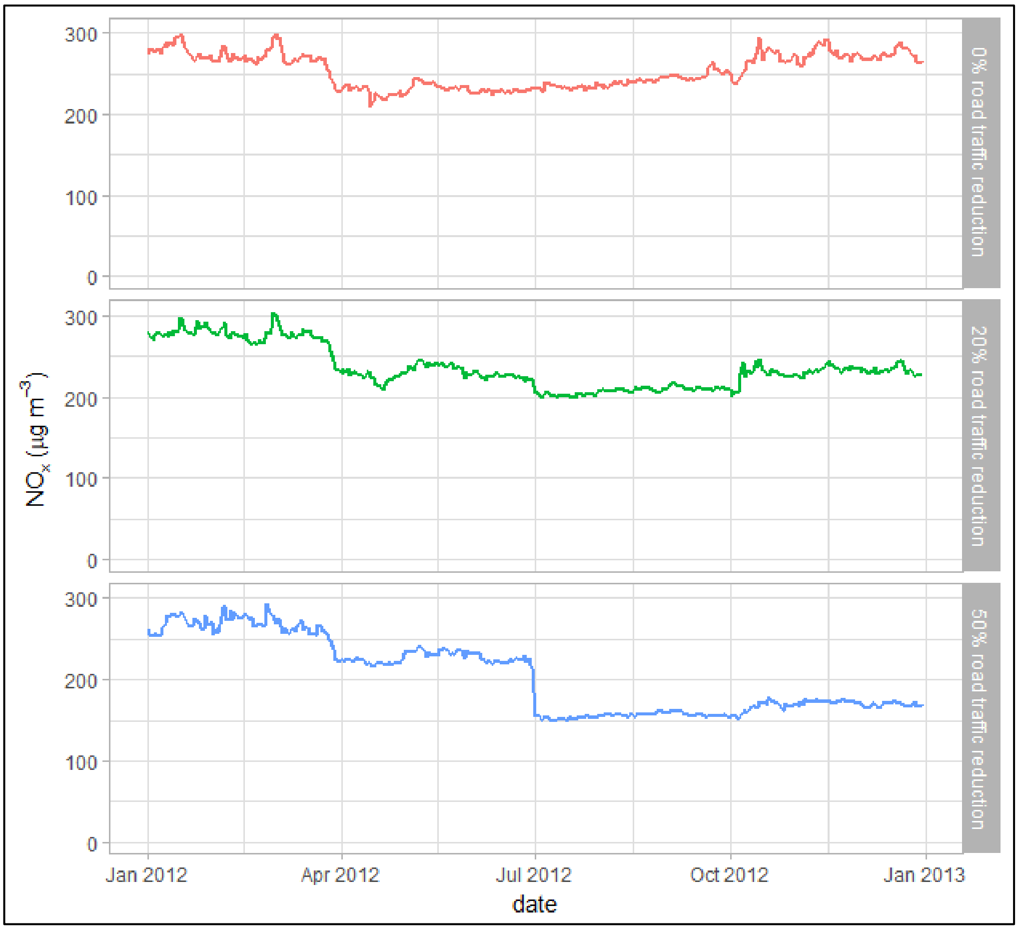


*Figure 18 Meteorologically-averaged $NO_x$ concentrations for Knightsbridge for base case conditions (no $NO_x$ reduction) and the 20 and 50% road traffic $NO_x$ reduction scenarios. The results have been daily-averaged.*

The results from these simulations are shown in *Figure 18*. All the results for *Figure 18* show considerably less variation in $NO_x$ concentrations due to the removal of meteorological variation (compare the time series with those in *Figure 16*). In the 50% $NO_x$ reduction case, it is now very clear that concentrations decrease sharply at the end of June 2012 i.e. consistent with the scenario being considered. It is also possible to see the change for the 20% reduction case. However, it is also apparent that there are many other changes (not associated with the intervention) that could be due to many (possibly unknown) factors. These results highlight the difficulty of establishing causation – while changes can be detected from time series using these types of methods, it can remain difficult to link them directly to underlying interventions.

# 5. Conclusions

Interventions that affect air quality take many different forms from policies that directly aim to improve air quality through to interventions that have an indirect effect on air quality. Interventions potentially cover an enormous range of spatial scales from very local changes e.g. affecting a few roads to continental-scale changes. This report focused on the types of intervention that would typically be relevant at a local scale and separate from national or international policies. The assessment of interventions is important for many reasons. One of the principal reasons is the potential to link a particular policy with a known, quantifiable outcome on air quality.

The quantification of interventions can be challenging and there is no single, accepted approach that has been widely adopted. However, frameworks for establishing and quantifying interventions have been developed – most notably the US Health Effects Institute ‘accountability chain’. The HEI approach provides a clear step-wise way of thinking about interventions from the original policy through changes in activity, atmospheric change and final impacts on health. The HEI framework is useful in the context of considering the effects of polices such as Clean Air Zones.

A common, ultimate goal of many intervention studies is to quantify a change in ambient pollutant concentrations or quantify some change in health impact. However, this is a rather narrow view of the whole ‘accountability chain’. In many cases it may only be possible to estimate the impact of an intervention based on changes in activity data such as changes in the vehicle fleet, followed by modelling studies to quantify the likely effect on concentrations and subsequent health impacts. In many practical situations, interventions can be analysed through a mix of direct measurements (such as traffic flow changes or ambient concentrations) and modelling (such as traffic model output and air quality modelling).

While it is often desirable to consider changes to atmospheric concentrations due to an intervention, quantifying such changes can be difficult. Without careful experimental design, the location of a measurement site (or sites) and the pollutants measured may not be ideal for detecting and quantifying changes. Indeed, many studies rely on existing measurement locations rather than optimally choosing new locations. Another important aspect of using ambient measurements to quantify changes is the often-dominant effect of meteorology, which can falsely mask or emphasise changes. For these reasons, it can be useful to apply statistical techniques to reduce or remove variations due to meteorology.

An understanding of the sources of uncertainty in evaluating interventions is important. There are many potential sources of uncertainty such as in the activity data, ambient measurements and dispersion modelling. An understanding of these uncertainties will help inform whether changes can reasonably be identified and quantified within the uncertainties of the data and techniques available to quantify interventions. For example, these considerations may lead to the conclusion that a monitoring program would not be appropriate because the likely change in emissions or concentrations is too small to practically detect. Under these situations, potentially expensive monitoring of changes may not be justified.

The wider context under which an intervention takes place is very important. It is very common for changes not associated with an intervention to occur at the same time, frustrating the task of quantification. For example, a local change that affects road vehicles against a backdrop of ongoing fleet turnover.

Air quality modelling has a potentially useful and important role in the assessment of interventions, even if there are relatively few examples where models have been used in such situations. Air quality models can provide information on the likely expected changes an intervention may have on ambient concentrations and support experimental design, which is something that can be undertaken before an intervention has started. Models are also useful for considering different scenarios and to understand the sensitivity of different input data and assumptions to concentration changes. Furthermore, models can be used to quantify changes that may be too small to detect through changes in measured concentrations. To date, there is limited information on the use of air quality models in these contexts but it is clear they could have an important role to play.

The air quality model simulations presented in Section 4 highlight many of the important issues in this report relating to the quantification of changes in ambient concentrations brought about by interventions. In particular, they show how difficult it can be to robustly identify a change in concentration when so much variation is driven by meteorological effects. However, the simulations also show the benefit of using statistical models to remove or reduce meteorological variation.

Even where changes in indicators (such as traffic flow or pollutant concentration) relating to interventions can be detected and quantified, the question of causality can remain. The strength of evidence relating to causality will vary depending on the nature of the intervention itself and to some extent will need to be determined on a case by case basis. Furthermore, it is unlikely that a single indicator will provide the necessary evidence that a change can be associated with an intervention and a ‘balance of evidence’ approach will be necessary that takes a wider view of all the evidence available.

## 5.1. Overview of case study findings

In general, the success of an intervention assessment is measured by whether it demonstrates the effectiveness of an intervention with statistical significance. In that respect, a common thread of all the intervention studies presented here is the success being dependent on the appropriate application of statistical tools, which sometimes requires a significant amount of effort and expertise.

The analysis is the most straightforward and the conclusions most clear-cut when a measurement is dominated by a single source type that is the subject of the intervention, through virtue of the siting of the measurement or the pollutant in question. This is certainly the case for the sulphur and bus route studies, where the measurement sites were well located to capture changes in the emissions. However, it is also possible to draw out conclusions within more ambiguous measurements, providing that the intervention is well defined and large enough in magnitude to demonstrate the quantitative influence, such as with the Heathrow study. Something that all these studies have in common is well-understood step changes in activity and/or emissions.

In complex environments with multiple pollution sources, such as interventions in large cities, the situation is much more challenging. Based on the case studies presented here, we can summarise the following to be of much benefit to the success of a study:

1. Extensive measurements: The increased amount of data will generally improve the overall statistical power of the analysis, but it is also important to cover a suitably long time period (such that meteorological factors can be accounted for) and a diverse number of sites, affected and control sites but also noting that different sites will show different pollutant responses to changes in activity, as was shown in the London multiple policy study. It is also beneficial to compare multiple interventions systematically, as was performed in the German LEZ study.

2. Monitoring of activity: Numerous studies presented here, such as the London LEZ study and Blunsdon Bypass assessment, highlight the value of monitoring changes in the targeted activity (e.g. road use) in addition to the pollution measures.

3. Predicting outcomes: It is also useful to make predictions of the interventions (or the counter-factual) for the purposes of contextualising results and hypothesis testing. Predictions of this type are often a requirement of the planning process (as with London City Airport and the Blunsdon Bypass) or AQMA action plans, such as Newcastle.

4. Taking account of meteorology: Because meteorology is a major controlling factor in air quality (Beijing being an extreme example of this), it is vital that the influence of this be accounted for when trying to quantify the effect of interventions. A number of the more successful case studies presented here had specific activities surrounding this theme.

The studies presented here are even further complicated when multiple interventions are superimposed; while statistically detecting a change in air quality in response to a single intervention can be difficult, quantitatively attributing change to individual overlapping interventions in a manner useful to policymakers is extremely challenging. The difficulties with attribution in the APEC Blue case study and unpicking the effect of the London LEZ against the other interventions are strong examples of this, but some coincident interventions are often less discrete, such as the problem being able to disentangle a discrete intervention against the backdrop of generally improving emissions (e.g. due to fleet turnover). The London multi-intervention analysis and the German multi-city LEZ study are good examples of overcoming this though a careful analysis of diverse datasets.

## 5.2. General advice on conducting analyses

How then should interventions be assessed in practice? There are two principal targets to assess, first the question of whether the desired outcome (e.g. air quality, health outcomes) has actually changed, and second the question of whether any observed change was causally related to the intervention. *Figure 19* shows the suggested steps involved in assessing an intervention.

Addressing the first point has been discussed above and this requires an assessment of the ‘baseline’ air quality and desired outcomes. To do this, ideally a sufficiently long period of advanced planning is required to set up the necessary monitoring and modelling systems. One may be fortunate in that existing monitoring could suffice (and this has been the case for some examples discussed above – fuel sulphur changes for example) but in the case of the local scale interventions that we are focussing on here, new monitoring is likely to be needed.

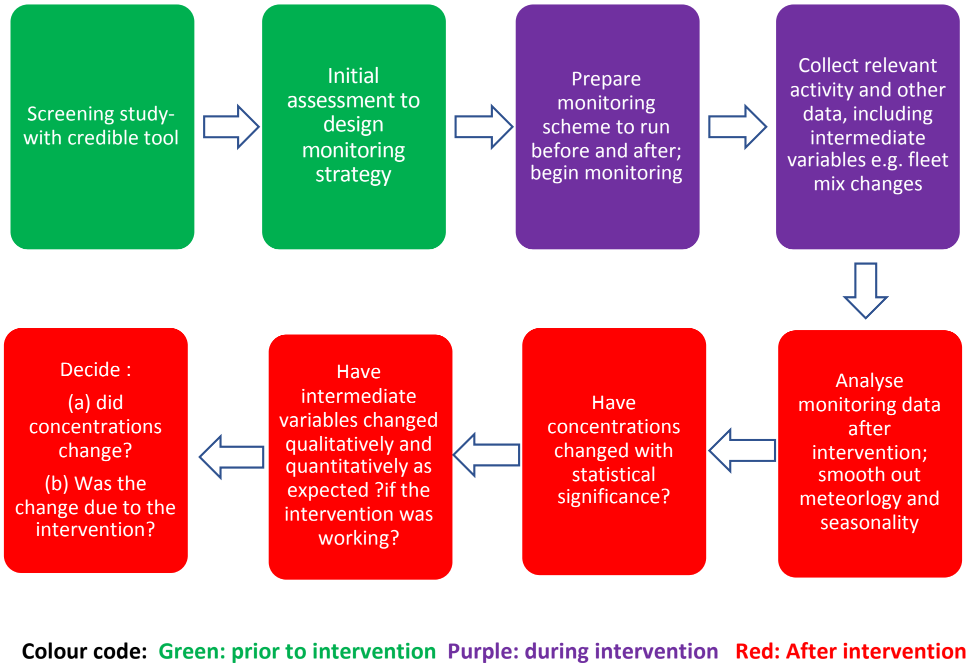


*Figure 19 Flow chart of the steps taken to conduct an analysis of an intervention.*

Before setting up new monitoring, ideally a study should be carried out to assess the likely scale of the changes that might occur. Judgements will need to be made about the sophistication of the model used at this stage e.g. whether a screening tool is suitable. Modelling can inform not just the magnitude of likely changes but also where the largest changes are likely to occur and can therefore be used to design an appropriate monitoring strategy. Establishing a statistically significant change in air quality or another outcome relies on this monitoring being of sufficient quality to be able to detect changes which might be small.

Consideration should be paid to monitoring the intervention through changes in activity in addition to air quality data. There may be occasions where it is not possible to obtain suitable air quality data or the available data is confounded by other issues. Activity data such as traffic or fuel usage are relatively easier to obtain and while these don’t give direct insights into changes in air quality that result, they can still be used as the basis of modelling studies which will allow the quantitative impacts on air quality to be estimated. Furthermore, these data will have additional uses, such as better understanding the underlying processes of the effects of an intervention. However, the desired end-point of quantifying changes in concentration shown in *Figure 19* is likely to provide a more robust basis for analysis than proxy indicators alone.

Changes in health outcomes can be even more difficult to demonstrate. Epidemiological studies would need to have sufficient power to show statistically significant changes. Well-characterised measurements of effects such as lung function in cohorts in the intervention area and in counter-factual areas could potentially be used, but again establishing statistically significant differences would need care.

Evaluating the effectiveness of the intervention is in general more difficult as it involves establishing causality against a background of other factors which may also change air pollution concentrations. The most obvious of these is meteorology and the first step in any assessment of concentration changes should account for meteorological changes. Models are now available to do this as discussed above.

**Advice for local authorities**

The methods, data requirements and skills needed to evaluate interventions in a comprehensive way can be complex and challenging. However, the very act of thinking about the steps involved, the types of data that are useful and some of the limitations is in itself highly valuable. Not all interventions result in a measurable changes in the concentrations of pollutants. For this, and other reasons, it is useful to start by considering the activity data related to an intervention. Furthermore, much of the effort involved in analysing interventions is related to controlling for confounders i.e. other factors that could result in change but which are not directly related to the intervention being considered.

Of specific current interest is the evaluation of concentrations of $NO_2$. In this case, almost all local authorities will have access to diffusion tube measurements rather than those from continuous chemiluminescent analysers. Because diffusion tubes provide longer-term averages and are less accurate than continuous analysers it may be difficult to detect changes in $NO_2$ concentrations robustly, even with a long time series from many sites. For that reason, it may be better to focus on the analysis of data from continuous analysers.

In their assessment of $NO_2$ concentrations, local authorities will also have conducted air quality modelling. This modelling is potentially highly useful from the perspective of intervention analysis, as discussed in this report. In carrying out air quality modelling, local authorities would have compiled a considerable amount of activity data and developed emission estimates for both base case conditions and different mitigation scenarios. These data sources are exactly of the type needed to consider the likely changes in total emissions of $NO_x$.

Having accounted for meteorology, in some specific cases which involve a technical measure causality may be easier to demonstrate – using low-emission buses along specific routes or the pedestrianisation of a street for example. However, where interventions involve behaviour change or in other ways involve more steps and ‘intermediate variables’ demonstrating causality may be less straightforward. This has been discussed to some extent above (see Figure 1 for example) and assessing causality in these cases would ideally involve gathering evidence at each stage of the accountability chain.

A traffic measure such as a low emission zone involves encouraging a faster turnover of the fleet towards a higher proportion of newer ‘cleaner’ vehicles entering the zone. One would therefore need to measure the fleet make up before and after the intervention (for example with ANPR) and also compare findings with a ‘counter-factual’ of fleet changes in an area without the intervention but which were otherwise similar. One might envisage an analogous process to assess the effectiveness of other ‘behavioural’ measures such as graduated parking charges, traffic light management (measure speeds and accelerations before and after).

# 6. Recommendations

The following recommendations have been identified.

1. The design of the assessment of an intervention should be considered at the planning stage, with the practitioners receiving pragmatic advice on the process. To date, the information relating to the analysis of interventions often resides in academic journals and is not easily accessible by most practitioners.

2. Many areas of policy should seek to incorporate the approaches and thinking of intervention analysis, e.g. in determining the efficacy or otherwise of technologies that offer potential reductions to emissions of different pollutants.

3. An analysis of the intermediate steps which the intervention is designed to influence should be carried out, to assess the causality of the relationship between the intervention and any measured concentration changes. For example, if the intervention is to change a pattern of fuel use, has fuel consumption changed as intended? Furthermore, there is also a need to take account of other changes that may also have affected the assessment of an intervention and to develop an understanding of whether any changes in concentrations specifically resulting from an intervention can reasonably be quantified.

4. Results from the implementation of local plans to mitigate air pollution should be pooled to derive a statistically more robust overall assessment of local measures.

5. Where new air quality monitoring sites are planned, their relevance for assessing interventions should be considered.

6. Intervention analysis is an active area of research that continues to evolve and be refined. Defra should retain a watching brief in this area to understand new developments and promote good practice.

# 7. References

Air Quality Expert Group, 2007. Trends in primary nitrogen dioxide in the UK. https://uk-air.defra.gov.uk/library/assets/documents/reports/aqeg/primary-no-trends.pdf

Air Quality Expert Group, 2017. Impacts of Shipping on UK Air Quality, https://uk-air.defra.gov.uk/library/reports.php?report_id=934

Air Quality Expert Group, 2018. Ultrafine Particles (UFP) in the UK, https://uk-air.defra.gov.uk/library/reports.php?report_id=968

Air Quality Expert Group, 2015. Linking Emission Inventories and Ambient Measurements, https://uk-air.defra.gov.uk/library/reports.php?report_id=828.

Barratt, B., Carslaw, D.C., 2014. Impacts of the SCRT bus retrofit programme on NO2 concentrations along Putney High Street. May 2014.

Barratt, B.M. and Fuller, G.W., 2014. Intervention assessments in the control of PM10 emissions from an urban waste transfer station. Environmental Science: Processes & Impacts, 16(6)1328-1337.

Bergin, M. S., G. S. Noblet, K. Petrini, J. R. Dhieux, J. B. Milford, AND R. A. Harley, 1999. Formal uncertainty analysis of a Lagrangian photochemical air pollution model. Environ. Sci. Technol., 33, 1116-1126.

Bureau international des poids et mesures (BIMP), 2008. Guide to the expression of uncertainty in measurement. Available at: https://www.bipm.org/en/publications/guides/gum.html. Accessed 25th November 2017.

Carnevale, C., J. Douros, G. Finzi, A. Graff, G. Guariso, Z. Nahorski, E. Pisoni, J-L. Ponche, E. Real, E. Turrini, Ch. Vlachokostas., 2016. Uncertainty evaluation in air quality planning decisions: a case study for Northern Italy, Environ. Sci. Policy, 65, 39-47.

Carslaw, D., Priestman, M., 2015. Analysis of the vehicle emission remote sensing campaigns data.

Carslaw, D.C., Beevers, S.D., Tate, J.E., Westmoreland, E.J., Williams, M.L., 2011. Recent evidence concerning higher NOx emissions from passenger cars and light duty vehicles. Atmospheric Environment 45.

Carslaw, D.C., Williams, M.L., Barratt, B., 2012. A short-term intervention study - Impact of airport closure due to the eruption of Eyjafjallajokull on near-field air quality. Atmospheric Environment 54.

Carslaw, D.C. and Ropkins, K., 2012. Openair—an R package for air quality data analysis. Environmental Modelling & Software, 27, 52-61.

Carslaw, D.C. and Taylor, P.J., 2009. Analysis of air pollution data at a mixed source location using boosted regression trees. Atmospheric Environment, 43(22-23), 3563-3570.

Chen, C., Sun, Y. L., Xu, W. Q., Du, W., Zhou, L. B., Han, T. T., Wang, Q. Q., Fu, P. Q., Wang, Z. F., Gao, Z. Q., Zhang, Q., and Worsnop, D. R., 2015. Characteristics and sources of submicron aerosols above the urban canopy (260 m) in Beijing, China, during the 2014 APEC summit, Atmos. Chem. Phys., 15, 12879-12895.

Clancy, L., Goodman, P., Sinclair, H. and Dockery, D.W. 2002. Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study 9341, 2002, The Lancet 360, 9341.

Cyrys, J., Gu, J., Gu, Soentgen, J. 2017. Analyse der Wirksamkeit von Umweltzonen in drei deutschen Städten: Berlin, München und Augsburg. Umweltbundesamt, Dessau-Roßlau, Germany.

Defra, 2016. Emissions Factors Toolkit V6.02. Available at: https://laqm.defra.gov.uk/review-and-assessment/tools/emissions-factors-toolkit.html.

Defra, 2017. Daily Air Quality Index. https://uk-air.defra.gov.uk/air-pollution/daqi Accessed 18/11/17.

Dockery, D.W., Rich, D.Q., Goodman, P.G., Clancy, L., Ohman-Strickland, P., George, P. and Kotlov, T. 2013. Effect of air pollution control on mortality and hospital admissions in Ireland, Health Effects Institute, 176, 3-109.

Ellison, R.B., Greaves, S.P. and Hensher, D.A., 2013. Five years of London's low emission zone: Effects on vehicle fleet composition and air quality. Transportation Research Part D: Transport and Environment, 23, 25-33.

Fensterer, V., Küchenhoff, H., Maier, V., Wichmann, H.E., Breitner, S., Peters, A., Gu, J. and Cyrys, J., 2014. Evaluation of the impact of low emission zone and heavy traffic ban in Munich (Germany) on the reduction of PM10 in ambient air. International journal of environmental research and public health, 11(5), 5094-5112.

Font, A. and Fuller, G.W., 2016. Did policies to abate atmospheric emissions from traffic have a positive effect in London?, Environmental pollution, 218, 463-474.

Friedman, M.S., Powell, K.E., Hutwagner, L., Graham, L.M. and Teague, W.G., 2001. Impact of changes in transportation and commuting behaviors during the 1996 Summer Olympic Games in Atlanta on air quality and childhood asthma. Jama, 285(7), 897-905.

Font, A. and Fuller, G.W., 2016. Did policies to abate atmospheric emissions from traffic have a positive effect in London? Environmental pollution, 218, 463-474.

Fuller, G.W., Carslaw, D.C. and Lodge, H.W., 2002. An empirical approach for the prediction of daily mean PM10 concentrations. Atmospheric Environment, 36(9), 1431-1441.

Fuller, G.W. and Green, D., 2006. Evidence for increasing concentrations of primary PM10 in London. Atmospheric Environment, 40(32), pp.6134-6145.

Grange, S.K., Carslaw, D.C., 2019. Using meteorological normalisation to detect interventions in air quality time series. Science of The Total Environment 653, 578–588.

Greenbaum, D.S., Org, D., 2017. Learning about "'cause'" and " 'effect'" through well-designed studies of air quality interventions. International Journal of Public Health 62, 719–720.

Grange, S (2017). normalweatherr: Package to conduct meteorological/weather normalisation on air quality data. R package version 0.0.16. https://github.com/skgrange/normalweatherr

Guo, J. P., He, J., Liu, H. L., Miao, Y. C., Liu, H., and Zhai, P. M.: Impact of various emission control schemes on air quality using WRF-Chem during APEC China 2014, Atmos. Environ., 140, 311-319, 10.1016/j.atmosenv.2016.05.046, 2016.

HEI Health Review Committee, 2016. Causal inference methods for estimating long-term health effects of air quality regulations. Research report (Health Effects Institute), (187), pp.5-49.

Highways Agency, 2005. A419 Blunsdon Bypass Environmental Statement. Volume 2: Technical Appendices.

Highways Agency, 2012. Post Opening Project Evaluation A419 Blunsdon Bypass One Year After Study. Document no. 5084038.753. http://assets.highways.gov.uk/our-road-network/pope/major-schemes/A419%20Blunsdon%20Bypass/A419%20Blundson%20BP%20POPE%20OYA%20%20-%20website%20version.pdf. Accessed 18/11/17

Highways England, 2015. Post Opening Project Evaluation A419 Blunsdon Bypass Five Years After Opening. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/636102/Full_report.pdf. Accessed 18/11/17

Huang, K., Zhang, X. Y., and Lin, Y. F., 2015. The "APEC Blue" phenomenon: Regional emission control effects observed from space, Atmos. Res., 164, 65-75.

International Standards Organisation (ISO), 2002. ISO 11222:2002 Air quality -- Determination of the uncertainty of the time average of air quality measurements. ISO Switzerland. Available at: https://www.iso.org/standard/32066.html. Accessed 25th November 2017

IPCC Working Group I, 2013, Summary for Policy Makers, Figure SPM.6 Available from http://www.climatechange2013.org/images/report/WG1AR5_SPM_FINAL.pdf

Kelly, F., H.R. Anderson, B. Armstrong, R. Atkinson, B. Barratt, S. Beevers, D. Derwent, D. Green, I. Mudway, and P. Wilkinson, 2011. The Impact of the Congestion Charging Scheme on Air Quality in London, Res. Rep. Health Eff. Inst., 155, 5-71.

Kelly, F., Armstrong, B., Atkinson, R., Anderson, H.R., Barratt, B., Beevers, S., Cook, D., Green, D., Derwent, D., Mudway, I. and Wilkinson, P., 2011. The London low emission zone baseline study. Health Effects Institute, Boston, USA.

Kelly, I. and Clancy, L.  1984 Mortality in a general hospital and urban air pollution. Irish Medical Journal, 77, 322-324.

Li, H. Y., Zhang, Q., Duan, F. K., Zheng, B., and He, K. B., 2016a. The "Parade Blue": effects of short-term emission control on aerosol chemistry, Faraday Discuss., 189, 317-335, 10.1039/c6fd00004e.

Li, R. P., Mao, H. J., Wu, L., He, J. J., Ren, P. P., and Li, X. Y., 2016b. The evaluation of emission control to PM concentration during Beijing APEC in 2014, Atmos Pollut Res, 7, 363-369.

London City Airport, 2017.  London City Airport 2016 Annual Performance Report (Compliance with Planning Permission 07/01510/VAR) 01 July 2017 https://assets.ctfassets.net/ggj4kbqgcch2/4EnG31r1tCgSw6oEIOQWGa/6aa0a4d1f3a0a32ae0227bbc69338e25/LCY_Annual_Performance_Report_2016_AW.pdf

Malina, C. and Scheffler, F., 2015. The impact of Low Emission Zones on particulate matter concentration and public health. Transportation Research Part A: Policy and Practice, 77, pp.372-385.

Marner, B. 2016.  CURED emissions model V1A.  Available at: http://www.aqconsultants.co.uk/Resources/Download-Reports.aspx

Marner, B., Moorcroft, J.S., 2016. Air Quality Assessment: Baker Street and Gloucester Place 2-way System.  http://www.bakerstreettwoway.co.uk/pdfs/Baker-Street-Air-Quality-Assessment. pdf  Accessed 18/11/17.

Morfeld, P., Groneberg, D.A. and Spallek, M.F., 2014. Effectiveness of low emission zones: large scale analysis of changes in environmental NO2, NO and NOx concentrations in 17 German cities. PloS one, 9(8).

Morfeld, P., Groneberg, D.A. and Spallek, M., 2015. The impact of Low Emission Zones on particulate matter concentration and public health: A Comment. Transportation Research Part A: Policy and Practice, 82, pp.255-256.

Newcastle City Council, 2006.  City of Newcastle upon Tyne Air Quality Action Plan Newcastle City Centre AQMA. https://www.newcastle.gov.uk/sites/default/files/wwwfileroot/environment-and-waste/pollution/city_centre_aqma_action_plan_0.pdf  Accessed 18/11/17.

Nunn, J., Hodgson, S., Moorcroft, S.M., Laxen, D., 2017. London City Airport Air Quality Measurement Programme:  Annual Report 2016.

Peel, J.L., Klein, M., Flanders, W.D., Mulholland, J.A. and Tolbert, P.E., 2010. HEI Health Review Committee Impact of improved air quality during the 1996 Summer Olympic Games in Atlanta on multiple cardiovascular and respiratory outcomes. Res Rep Health Eff Inst, 148, 3-23.

Solomon, S., D.J. Ivy, D. Kinnison, M.J. Mills, R.R. Neely III, A. Schmidt, 2016. Emergence of healing in the Antarctic ozone layer, Science, 353, 269-274.

Sweeney, B. P., Quincey, P. G., Green, D., & Fuller, G. W., 2015. Quantifying the impact of nitric oxide calibration gas mixture oxidation on reported nitrogen dioxide concentrations. Atmospheric Environment, 105, 169-172.

Tang, G., Zhu, X., Hu, B., Xin, J., Wang, L., Munkel, C., Mao, G., and Wang, Y., 2015. Impact of emission controls on air quality in Beijing during APEC 2014: lidar ceilometer observations, Atmos. Chem. Phys., 15, 12667-12680.

Tao, J., Gao, J., Zhang, L. M., Wang, H., Qiu, X. H., Zhang, Z. S., Wu, Y. F., Chai, F. H., and Wang, S. L., 2016. Chemical and optical characteristics of atmospheric aerosols in Beijing during the Asia-Pacific Economic Cooperation China 2014, Atmos. Environ., 144, 8-16.

Transport for London (TfL), 2008. London Low Emission Zone Impacts Monitoring Baseline Report, July 2008. Transport for London, London.

Transport for London (TfL), 2010. Travel in London Report 3. Transport for London, London.

Tomlin, A.S.,  T. Ziehn, P. Goodman, J.E. Tate, N.S. Dixon, 2016. The treatment of uncertainties in reactive pollution dispersion models at urban scales, Faraday Discussions 189, 567-587.

Young, M., 2005 Dublin ban on sales of bituminous coal. Department of Environment, Heritage & Local Government, Dublin.

Wang, G., Cheng, S. Y., Wei, W., Yang, X. W., Wang, X. Q., Jia, J., Lang, J. L., and Lv, Z., 2017. Characteristics and emission-reduction measures evaluation of PM2.5 during the two major events: APEC and Parade, Sci. Total Environ., 595, 81-92.

Wang, P., and Dai, X. G., 2016. "APEC Blue" association with emission control and meteorological conditions detected by multi-scale statistics, Atmos. Res., 178, 497-505.

Wang, Y. Q., Zhang, Y., Schauer, J. J., de Foy, B., Guo, B., and Zhang, Y. X., 2016a. Relative impact of emissions controls and meteorology on air pollution mitigation associated with the Asia-Pacific Economic Cooperation (APEC) conference in Beijing, China, Sci. Total Environ., 571, 1467-1476.

Wang, Z. S., Li, Y. T., Chen, T., Li, L. J., Liu, B. X., Zhang, D. W., Sun, F., Wei, Q., Jiang, L., and Pan, L. B., 2015. Changes in atmospheric composition during the 2014 APEC conference in Beijing, J. Geophys. Res.-Atmos., 120, 12695-12707.

Wang, Z. S., Li, Y. T., Chen, T., Zhang, D. W., Li, L. J., Liu, B. X., Li, J. X., Sun, F., and Pan, L. B., 2016b. SCIENCE-POLICY INTERPLAY Improvement of Air Quality from 2008 to 2014 in Beijing and the Scientific Approach to Achieve APEC Blue, B. Am. Meteorol. Soc., 97, 553-559.

Wei, X., Gu, X. F., Chen, H., Cheng, T. H., Wang, Y., Guo, H., Bao, F. W., and Xiang, K.S., 2016. Multi-Scale Observations of Atmosphere Environment and Aerosol Properties over North China during APEC Meeting Periods, Atmosphere-Basel, 7.

Wolff, H., 2014. Keep Your Clunker in the Suburb: Low-emission Zones and Adoption of Green Vehicles. The Economic Journal, 124(578), 481-512.

Zhai, S. X., An, X. Q., Liu, Z., Sun, Z. B., and Hou, Q., 2016. Model assessment of atmospheric pollution control schemes for critical emission regions, Atmos. Environ., 124, 367-377.

Zhang, L., Shao, J. Y., Lu, X., Zhao, Y. H., Hu, Y. Y., Henze, D. K., Liao, H., Gong, S. L., and Zhang, Q., 2016. Sources and Processes Affecting Fine Particulate Matter Pollution over North China: An Adjoint Analysis of the Beijing APEC Period, Environ. Sci. Technol., 50, 8731-8740.