Environment Agency

using science to create a better place

Integration of air quality modelling and monitoring methods: review and applications

Science Report – SC060037/SR1

The Environment Agency is the leading public body protecting and improving the environment in England and Wales.

It's our job to make sure that air, land and water are looked after by everyone in today's society, so that tomorrow's generations inherit a cleaner, healthier world.

Our work includes tackling flooding and pollution incidents, reducing industry's impacts on the environment, cleaning up rivers, coastal waters and contaminated land, and improving wildlife habitats.

This report is the result of research commissioned and funded by the Environment Agency's Science Programme.

Published by:
Environment Agency, Rio House, Waterside Drive, Aztec West, Almondsbury, Bristol, BS32 4UD
Tel: 01454 624400 Fax: 01454 624409
www.environment-agency.gov.uk

ISBN: 978-1-84432-896-3

Further copies of this report are available from:
The Environment Agency's National Customer Contact Centre by emailing:
enquiries@environment-agency.gov.uk
or by telephoning 08708 506506.

Author(s):
Ball, Angela, Hill, Richard & Jenkinson, Peter

Dissemination Status:
Released to all regions
Publicly available

Keywords:
Integrating, modelling, monitoring, network, design, air quality, assessment, data assimilation

Research Contractor:
Westlakes Scientific Consulting Ltd
Westlakes Science & Technology Park
Moor Row
Cumbria
CA24 3LN

Environment Agency's Project Manager:
Dr. Hope Brett,
Environment Agency,
Lower Bristol Road,
Bath
BA2 9ES

Collaborator(s): None

Science Project Number:
SC060037

Product Code:
SCHO0308BOAE-E-P

Science at the Environment Agency

Science underpins the work of the Environment Agency. It provides an up-to-date understanding of the world about us and helps us to develop monitoring tools and techniques to manage our environment as efficiently and effectively as possible.

The work of the Environment Agency's Science Department is a key ingredient in the partnership between research, policy and operations that enables the Environment Agency to protect and restore our environment.

The science programme focuses on five main areas of activity:

- **Setting the agenda**, by identifying where strategic science can inform our evidence-based policies, advisory and regulatory roles;
- **Funding science**, by supporting programmes, projects and people in response to long-term strategic needs, medium-term policy priorities and shorter-term operational requirements;
- **Managing science**, by ensuring that our programmes and projects are fit for purpose and executed according to international scientific standards;
- **Carrying out science**, by undertaking research – either by contracting it out to research organisations and consultancies or by doing it ourselves;
- **Delivering information, advice, tools and techniques**, by making appropriate products available to our policy and operations staff.

Steve Killeen

Head of Science

Executive summary

The Environment Agency is responsible for the regulation of Part A(1) industrial activities in England and Wales. Ambient monitoring and air quality modelling are the tools most frequently used to assess the impact of industrial emissions to air and ultimately provide evidence upon which regulatory decisions are based. However, the use of either modelling or monitoring alone may introduce a high degree of uncertainty into air quality impact assessment. In reality, monitoring and modelling are often used independently of one another. However, it is anticipated that by applying monitoring and modelling integration techniques, the strengths of both technologies can be complemented and the uncertainties associated with each technique reduced.

The purpose of this report is to:

i. Bring the Environment Agency up-to-date with recent developments in the field of monitoring and modelling integration methods;

ii. Inform the Environment Agency of methods that could be used to resolve challenges to regulatory decisions, through the application of monitoring and modelling integration techniques;

iii. Review the possibility of economic savings through more efficient use of the existing monitoring resources in England and Wales.

To achieve effective integration, it may be necessary to increase investment in monitoring resources and ensure the optimal placement of monitoring sites. Existing monitoring networks such as the AURN are of limited use as they are often located far from industrial sources and the potential to attribute pollutant concentrations to such sources may be difficult due to the interference of pollution from roads and urban areas. It is essential that any monitoring resource must be optimised in terms of the number and placement of monitors, both to ensure cost effectiveness and to optimise the potential for effective integration of monitoring and modelling data. This can be achieved with the adoption of decision-support tools, which are able to prioritise multiple objectives such as population protection, ecosystem protection or targeting specific air quality metrics and are invaluable in designing an effective monitoring network.

Improving the accuracy of modelled data is also crucial to the effective integration of modelling and monitoring data. This may be achieved through the use of more representative meteorological data and the application of more appropriate background correction factors.

Data assimilation (model calibration) methods applied to the Kincaid SO_2 validation data set indicated that the optimum number of monitoring sites necessary to maximise the accuracy of data assimilation methods is between 10 and 15. However, the success of the more complex integration techniques, e.g. kriging, is limited. Of the simple techniques, the simple ratio method provided the best calibration using both the Kincaid data set and the Aire Valley data set. The linear regression method performed on a similar level to the simple ratio method when using a higher number of data points, but the performance of this method decreased dramatically when less than three monitoring points were used to calibrate the modelled data. It should be noted that calibration is designed to reduce uncertainty in model outputs caused by systematic errors and is of limited use when uncertainty is due to random errors.

The application of modelling and monitoring integration methods to improve the extrapolation of short-term monitoring campaigns, currently calculated using the Environment Agency pro-rata extrapolation method, proved of limited value. The

Integration Scaling Method showed an improvement in the consistency of measured and predicted exceedences compared to the pro-rata method, although overall uncertainty was still high.

A full set of recommendations for the Environment Agency stemming from this review is listed in Section 8 of this report.

Contents

List of tables and figures

1 Introduction

The Environment Agency is responsible for the regulation of Part A(1) industrial activities in England and Wales, potentially the most polluting of all industrial sources. Part A(1) processes include activities within the energy industry, metal production and processing, the mineral industry, waste management, the chemical industry and intensive agriculture (PPC Regulations, 2000).

Ambient monitoring and air quality modelling are the tools most frequently used to assess the impact of industrial emissions to air and ultimately provide evidence on which regulatory decisions are based. However, the use of either modelling or monitoring alone may introduce a high degree of uncertainty into air quality impact assessment. In reality, monitoring and modelling are often used independently of one another. However, it is anticipated that by applying monitoring and modelling integration techniques, the strengths of both technologies can be complemented and the uncertainties associated with each technique reduced.

This report outlines the existing approaches adopted in England and Wales for the assessment of environmental impacts associated with aerial emissions from Part A(1) industrial sources. The uncertainties associated with these existing approaches (i.e. the independent use of either monitoring or modelling) are also discussed. Modelling and monitoring integration techniques for the improvement of air quality impact assessment and the design and optimisation of monitoring networks are reviewed. Several of these techniques are analysed using data from the Aire Valley monitoring network and the Kincaid SO_2 validation data set, to determine their potential for improving the regulation of local air quality surrounding Part A(1) industrial sources.

The purpose of this report is to:

i. Bring the Environment Agency up-to-date with recent developments in the field of monitoring and modelling integration methods;

ii. Inform the Environment Agency of methods that could be used to resolve challenges to regulatory decisions, through the application of monitoring and modelling integration techniques; and

iii. Review the possibility of economic savings through more efficient use of the existing monitoring resources in England and Wales.

2 Monitoring ambient air quality

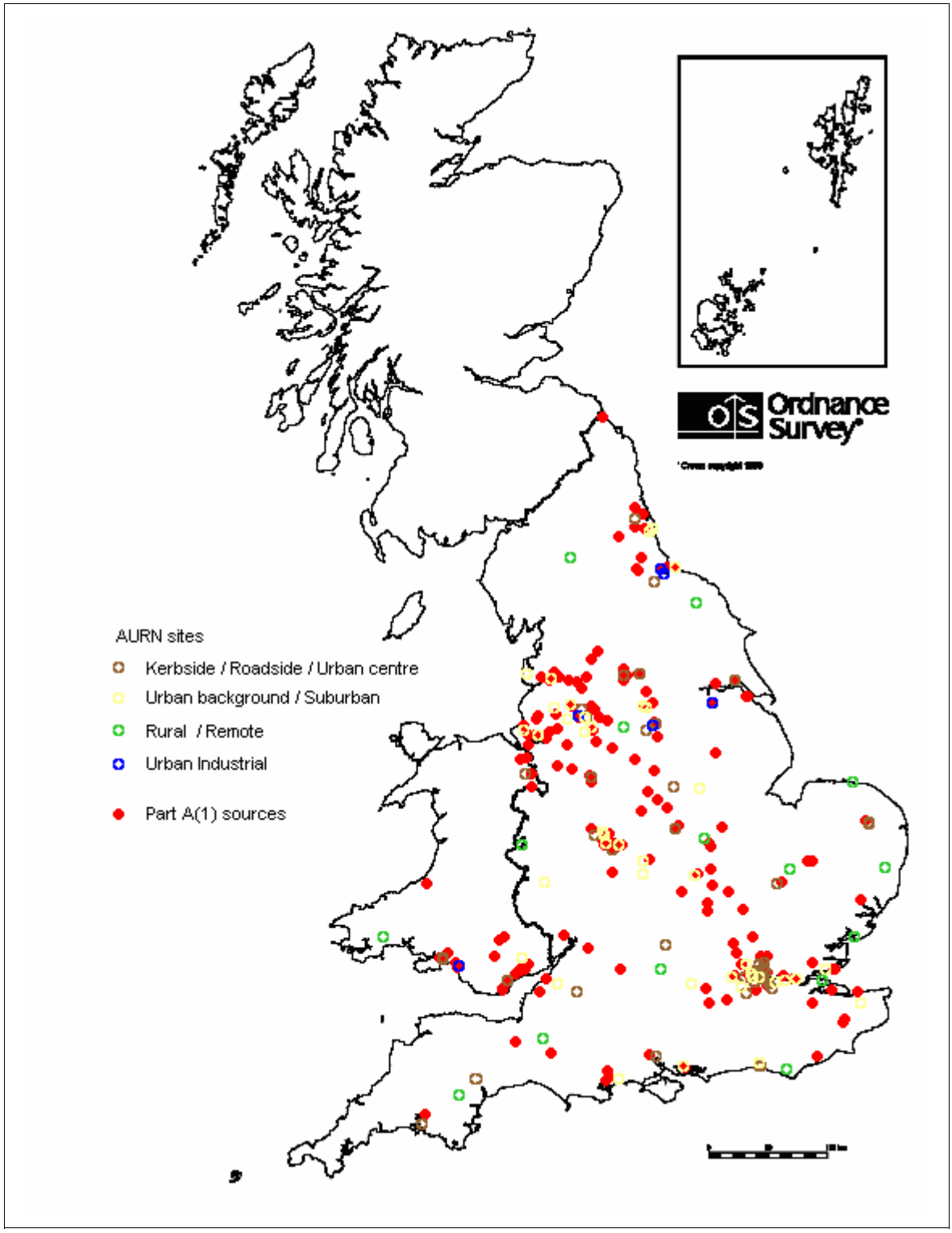
2.1 The history of air quality monitoring in the UK

Domestic emissions of air pollutants were first addressed by regulations set out in the 1926 Smoke Abatement Act and the Clean Air Acts of 1956 and 1968. However it was not until 1987 that the first air quality measurement network (the Statutory Urban Network or SUN) was established in the UK. After the Environmental Protection Act in 1990, the SUN network was expanded to monitor urban background air quality and was renamed the Enhanced Urban Network (EUN). The London Air Quality Network (LAQN) was subsequently formed in 1993 to co-ordinate and improve air pollution monitoring in London. In addition, under Section IV of the Environment Act 1995, each Local Authority was responsible for managing air quality within their area, which required both continuous and non-continuous monitoring of air pollutants. This expansion in monitoring led to the amalgamation of automatic monitoring sites in 1998 to form the Automatic Urban and Rural Network (AURN), largely managed by the Department of Environment, Food and Rural Affairs (Defra). However, the AURN also includes automatic monitoring sites owned by Local Authorities and operated to standards equivalent to Defra-managed AURN sites. The AURN monitors a range of pollutants including sulphur dioxide (SO_2), oxides of nitrogen (NO_x), carbon monoxide (CO), ozone (O_3) and fine particulate matter (PM_{10} and $PM_{2.5}$). Several sites within the AURN are also part of a smaller Hydrocarbon Network (HCN), which in addition to the aforementioned pollutants, monitors 25 Volatile Organic Compounds including benzene, 1,3-butadiene and O_3 precursors. To date (end of 2007) the AURN consists of 126 sites of which 107 are located in England and Wales (Figure 2.1).

In addition to the automatic sampling networks, there are several non-automatic networks currently monitoring air quality around the UK:

- Nitrogen Dioxide Diffusion Tube Network
- Smoke and Sulphur Dioxide Monitoring Network
- Rural Sulphur Dioxide Monitoring Network
- Non-automatic Hydrocarbon Network (monitoring benzene at urban background and roadside locations)
- Multi Element and Lead Monitoring Network
- Toxic Organic Micropollutants (TOMPS) Network
- Acid Deposition Network

The AURN exists primarily to satisfy the statutory requirements of EC Directives through the implementation of Local Air Quality Management and therefore does not focus on assessing the impact of emissions from industrial sources. AURN sites are classified into eight groups, listed in Table 2.1. AURN sites classified as ‘kerbside’, ‘roadside', ‘urban centre’ and ‘airport’ are considered inappropriate for measuring air concentrations arising from industrial sources because distant industrial air quality signals may be saturated by relatively close urban or traffic pollution sources. These classified sites account for 35% of all AURN sites. However, seven AURN sites are classified as ‘Urban Industrial’ (Table 2.2). These sites have been deliberately located In the vicinity of industrial sources and therefore may be of potential use in compliance assessment for Part A(1) processes.



Source: AURN site data supplied by AEA Energy and Environment and Part A(1) site data supplied by the Environment Agency.

Figure 2.1 Part A(1) industrial sources and AURN network monitoring sites in operation during 2006 in England and Wales.

Table 2.1 Classification of AURN sites in England, Wales, Scotland and Northern Ireland.

AURN Site Classification	Abbreviation	Description	Number of sites
Kerbside	K	Within 1 m of edge of busy road	3
Roadside	Ro	Within 5 m of kerbside	21
Urban centre	UC	Non-kerbside sites within towns or cities	25
Suburban	S	Sites typical of residential areas	11
Urban background	UB	Urban locations distanced from sources	37
Urban industrial	UI	Sites where industrial emissions make a significant contribution to measured pollution levels	7
Rural/Remote	Ru	Open country/isolated locations	21
Airport	A	Within the grounds of an airport	1

Source: http://www.bv-aurnsiteinfo.co.uk/

An analysis of the location of AURN sites relative to Part A(1) industrial sources in England and Wales reveals that only 38 from a total of 180 have a single AURN site within 5 km, only twelve have two AURN sites within 5 km and only one source has three AURN sites within 5 km. Similar statistics for distances ranging through 1, 2, 5, 7 and 10 km are shown in Table 2.3. Therefore AURN sites are of limited use in the regulation of industrial sources either because they are too far away from the source or there is interference from the contribution of pollutants from other sources e.g. roads or urban centres. However, AURN sites may be useful in supplying important information about background concentrations.

Table 2.2 AURN sites classified as Urban Industrial.

Site Location	Pollutants	Description
Billingham, Teesside Urban Area	NO_2	Situated in a council depot within a residential area, with a large complex of chemical/manufacturing plants 1-3 km to the south.
Middlesborough, Teeside Urban Area	CO, NO_2, O_3, PM_{10}, SO_2	Situated in a residential area near Longlands College of Further Education.
Scunthorpe Town, Yorkshire & Humberside	PM_{10}, SO_2	Situated in a mixed industrial/commercial/ residential area.
Sheffield Tinsley, Sheffield Urban Area	CO, NO_2	Situated in a mixed residential/industrial area, 200 metres from the M1.
Salford Eccles, Greater Manchester Urban Area	CO, NO_2, O_3, PM_{10}, SO_2	Situated in a residential area adjacent to Eccles town centre with parkland immediately to the south and a suburban road 7 metres to the north.
Grangemouth, Central Scotland	CO, NO_2, PM_{10}, SO_2	Situated in a residential area with BP refineries 300 metres north with the nearest main road 250 metres away.
Port Talbot, Swansea Urban Area	NO_2, O_3, PM_{10}, SO_2	Situated in the grounds of a small hospital, 700m from Port Talbot steel works and 75 metres from the M4 motorway.

Table 2.3 Number of Part A(1) sites with *x* AURN sites within 1-10 km in England and Wales.

Distance (km)	1 AURN	2 AURNs	3 AURNs	4 AURNs	5+ AURNs
0-1	7	0	0	0	0
0-2	21	2	0	0	0
0-5	38	12	1	0	0
0-7	45	15	7	0	0
0-10	56	24	11	3	6

2.2 Legislation and air quality monitoring

The recent growth in the UK air quality monitoring infrastructure, as described in the previous section, has mainly been driven by European legislation. The European Directives associated with improving air quality and the transposition of those Directives into UK law is briefly outlined in the following sections.

2.2.1 EU Air Quality Framework Directive

European Union (EU) law, in particular the Air Quality Framework Directive (1996/62/EC), requires member states to ensure that air pollution does not exceed certain legal limits. The Air Quality Framework Directive (AQFD) defines the policy framework for 13 air pollutants known to have a harmful effect on human health and the environment. The limit values (or in some cases target values) for the 13 pollutants are set through a series of Daughter Directives:

- First Daughter Directive (99/30/EC) sets limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air.
- Second Daughter Directive (2000/69/EC) sets limit values for benzene and carbon monoxide.
- Third Daughter Directive (2002/3/EC) sets target values relating to ozone.
- Fourth Daughter Directive (2004/107/EC) sets target values relating to arsenic, cadmium, mercury, nickel and Polycyclic Aromatic Hydrocarbons (PAHs) in ambient air.

In addition, the National Emissions Ceilings Directive (2001/81/EC) seeks to reduce emissions of those pollutants that cause acidification, eutrophication and ground-level ozone by setting annual emission limits for SO_2, NO_x, ammonia (NH_3) and Volatile Organic Compounds (VOCs) for each member state.

2.2.2 UK Air Quality Strategy

The UK Government's Air Quality Strategy (latest revision: Defra, 2007) makes it the responsibility of Local Authorities to work towards compliance with national objectives. In many cases these objectives replicate the legal limit values set in the EU AQ Daughter Directives. The Environment Agency is required to contribute to air quality action plans developed by Local Authorities and to have regard for the Air Quality Strategy.

The Air Quality Strategy (AQS) sets air quality standards and objectives for ten air pollutants (listed in AQ Daughter Directives) to be achieved between 2003 and 2020 in order to improve ambient air quality in the UK. Two key concepts provide the central structure to the AQS: standards and objectives. Standards are the concentrations of pollutants in the atmosphere, which can broadly be taken to achieve a certain level of environmental quality and are based on assessment of the effects of each pollutant on human health. Objectives are policy targets generally expressed as a maximum ambient concentration, to be achieved either without exception, or with a permitted number of exceedences within a specified time scale. Most of these objectives are derived from the limit and target values set in the AQ Daughter Directives, however, objectives are also generally based on the recommendations of the Expert Panel on Air Quality Standards (EPAQS). As a consequence, some objectives have to be achieved

sooner than EU-specified limit or target values. In addition, some AQS objectives are even more stringent than those specified in EU legislation, for example, the 15-minute average objective for SO_2 does not appear in EU legislation. The EU Daughter Directives also include limit values for oxides of nitrogen and sulphur dioxide for the protection of vegetation and ecosystems. However, the Government and the devolved administrations intend that these limit values should be treated as national objectives, against which compliance will be monitored at a national level.

Objectives for seven of the ten pollutants addressed in the AQS are shown in Table 2.4. These objectives are prescribed in regulations for the purposes of Local Air Quality Management (LAQM). The Environment Agency also has to have regard to the objectives set out in the AQS in its regulatory activities but the Pollution Prevention and Control Regulations for England and Wales (PPC, 2000) places a much stronger, direct duty on the Environment Agency for the achievement of EU limit values.

National AQS objectives are listed in Table 2.5 and include the national objectives to protect vegetation and ecosystems. Because of its trans-boundary nature, ozone continues to be a national objective, and is therefore outside the scope of LAQM. The release and dispersion of PAHs continues to have significant uncertainties, so it too is outside the scope of LAQM and is managed at a national level.

The Fourth Air Quality Daughter Directive (2004/107/EC) has only recently been transposed into UK law by the Air Quality Standards Regulations (2007) and therefore is not included in the current AQS. However, target values for pollutants listed in Daughter Directive 4 (known as Group B pollutants) are shown in Table 2.6.

The 2007 AQS introduced a new concept into the control of air quality, namely exposure reduction for particulates. This will be in terms of a set reduction in concentrations of $PM_{2.5}$ over a matter of years. Again this will be delivered at a national level and will be monitored at background sites in large conurbations (see Table 2.5).

For many substances released to air, national objectives have not been defined. Therefore, the Environment Agency has adopted interim values known as Environmental Assessment Levels (EALs) for these substances. These values are listed in Appendix D of the Agency's IPPC H1 Methodology, Environmental Assessment and Appraisal of BAT (Environment Agency, 2002a).

Table 2.4 Air Quality Strategy objectives included in regulations for the purposes of Local Air Quality Management (Group A pollutants).

Pollutant	Objective			To be achieved by and maintained thereafter
	Standard	Averaging period	Metric	
Benzene	16.25 µg/m^3	Running annual mean		31st December 2003
	*5.00 µg/m^3	Annual average		31st December 2010
1,3-Butadiene	2.25 µg/m^3	Running annual mean		31st December 2003
Carbon monoxide	10 mg/m^3	Maximum daily running 8-hour mean		31st December 2003
Lead	0.5 µg/m^3	Annual mean		31st December 2004
	0.25 µg/m^3	Annual mean		31st December 2008
Nitrogen dioxide	200 µg/m^3	1-hour mean	Not to be exceeded more than 18 times a year	31 December 2005
	40.0 µg/m^3	Annual mean		31 December 2005
Particulate Matter (PM_{10})	50.0 µg/m^3	24-hour mean	Not to be exceeded more than 35 times a year	31st December 2004
	40.0 µg/m^3	Annual mean		31st December 2004
Sulphur dioxide	266 µg/m^3	15-minute mean	Not to be exceeded more than 35 times a year	31st December 2005
	350 µg/m^3	1-hour mean	Not to be exceeded more than 24 times a year	31st December 2004
	125 µg/m^3	24-hour mean	Not to be exceeded more than 3 times a year	31st December 2004

* Only England & Wales objective shown

Table 2.5 Air Quality Strategy objectives included in regulations for delivery at a national level (Group A pollutants).

Pollutant	Objective		To be achieved by
	Standard	Averaging period	
Objectives for the protection of human health:			
Ozone	100 µg/m^3 (Not to be exceeded more than 10 times/ year)	8-hour mean	31 December 2005
Polycyclic Aromatic Hydrocarbons	0.25 ng/m^3 B[a]P	Annual mean	31 December 2010
Particulate Matter ($PM_{2.5}$)	*25 µg/m^3 Target of 15% reduction in concentrations at urban background	Annual mean Annual mean	2020 Between 2010 and 2020
Objectives for the protection of vegetation and ecosystems:			
Nitrogen oxides	30 µg/m^3	Annual mean	31 December 2000
Sulphur dioxide	20 µg/m^3 20 µg/m^3	Annual mean Winter average (1st October – 31st March)	31 December 2000 31 December 2000
Ozone	Target value of 18,000 µg/m^3 based on AOT40 to be calculated from 1-hour values from May to July, and to be achieved, so far as possible, by 2010	Average over 5 years	1 January 2010

* Only England & Wales objectives shown

Table 2.6 Air Quality Standards Regulations (2007) EU Target Values for delivery at a national level (Group B pollutants).

Pollutant	Objective		To be achieved by
	Concentration*	Averaging period	
Arsenic	6 ng/m^3	Annual mean	31st December 2012
Benzo(a)pyrene**	1 ng/m^3	Annual mean	31st December 2012
Cadmium	5 ng/m^3	Annual mean	31st December 2012
Nickel	20 ng/m^3	Annual mean	31st December 2012

*Total content of the relevant pollutant in the PM10 fraction averaged over one calendar year

**Polycyclic Aromatic Hydrocarbon (PAH)

2.3 Industrial processes and air quality monitoring

Air pollution from industry is regulated either by Local Authorities or by the Environment Agency, depending on the category of the industrial activity (Part A(1), Part A(2) or Part B processes). The Environment Agency regulates Part A(1) processes, potentially the most polluting of all industrial sources.

In order to reduce the impact of aerial emissions from industrial processes on human health and the environment, industrial activities are subject to Pollution Prevention and Control (PPC) Regulations. The system of Pollution, Prevention and Control has replaced the Integrated Pollution Control (IPC) regime set up under Part I of the Environmental Protection Act (EPA, 1990) and implements the European Directive (EC/96/61) on Integrated Pollution Prevention and Control (IPPC).

These regulations aim to ensure that industry adopts an integrated approach to pollution control in order to achieve a high level of protection for the environment and human health. Under these regulations, operators of Part A (1 and 2) and Part B industrial processes are required to obtain a permit to operate. In all cases, industrial operators must assess the impact of their emissions on human health and the environment and adopt Best Available Techniques (BAT) to minimise impacts. If a permit is issued, it will include conditions aimed at preventing or reducing pollution to acceptable levels. Depending on the size and scale of the process, stack emissions, as well as on-site and off-site ambient monitoring, may be required from the operator. In addition to these regulations, a number of complementary EU Directives specific to industrial emissions have been created. These include:

- The Large Combustion Plant Directive (2001/80/EC), which aims to reduce emissions from large combustion plants with a thermal output of greater than 50 MW (e.g. power stations, refinery boilers and large industrial boilers).
- The Sulphur Content of Liquid Fuels Directive (1999/32/EC), which aims to reduce emissions of SO_2 resulting from the combustion of heavy fuel oil and gas oil.
- The Solvents Directive (1999/13/EC), which aims to limit emissions of VOCs from the use of organic solvents in industrial processes.

2.3.1 Emissions monitoring

Under the requirements of IPPC regulations and the Large Combustion Plant Directive (LCPD), continuous emissions monitoring (CEM) of SO_2, NO_x and PM_{10} must be carried out by the operator for boilers with a thermal input greater than 100 MW. In addition to ensuring that the emission limits stated in the LCPD are adhered to, this can improve modelling of emissions to air from these sources. However, it should be recognised that there are still large uncertainties in continuously monitored concentrations and flow measurements.

For the remaining Part A(1) activities, continuous or non-continuous emissions monitoring of air pollutants may be required by the Environment Agency depending on the process and the scale of the activity. Table 2.7 provides examples from a range of IPPC operating permit submissions to the Environment Agency, listing source-specific requirements for stack emissions monitoring to be undertaken by Part A(1) operators.

2.3.2 Ambient monitoring

Ambient monitoring is also desirable to monitor the off-site impacts of emissions to air from industrial processes. However, while all the power generation operators are required to continuously monitor SO_2, NO_2 and PM_{10} concentrations off-site, requirements for off-site monitoring by other Part A(1) processes (e.g. refineries, steel and cement works) are examined on a case-by-case basis during the IPPC permit authorisation process. Compulsory on-site monitoring is limited to landfill sites, which are required to conduct ambient air monitoring along the entire site boundary and the landfill gas collection infrastructure on a monthly basis. Off-site monitoring by the landfill operator need only occur following a substantial odour complaint. Table 2.8 summarises the ambient monitoring of air quality undertaken by operators of power stations in the Aire Valley and Table 2.9 summarises both on- and off-site ambient monitoring required from operators of landfill sites.

Table 2.7 Examples from IPPC operating permit submissions to the Environment Agency, listing source-specific requirements for stack emissions monitoring for a range of Part A(1) processes.

Pollutant	Part A Process			
	Power	Refineries	Steel	Cement
Sulphur dioxide (SO_2)	CEM	CEM or 2/year[a]	CEM or 1/year[b]	CEM or 2/year
Nitrogen dioxide (NO_2)	CEM	CEM or 2/year[a]	CEM or 1/year[b]	CEM or 2/year
Particulate Matter (PM_{10})	CEM	CEM or 2/year[a]	CEM or 1/year[b]	CEM or 2/year
Polychlorinated dibenzodioxin (PCDD)			2/year	2/year
Polychlorinated dibenzofuran (PCDF)			2/year	2/year
Polychlorinated biphenyl (PCB)			2/year	2/year
Polycyclic aromatic hydrocarbons (PAHs)			2/year	2/year
Volatile organic compounds (VOCs)		2/year	2/year	2/year
Hydrochloric Acid (HCl)			2/year	2/year
Hydrogen fluoride (HF)			2/year	2/year
Carbon monoxide (CO)				2/year

CEM = Continuous Emissions Monitoring

[a]Dependent on process e.g. iron ore sintering, steel making, iron making, concast, rolling mills, re-heat furnaces, power plant boilers or coke ovens

[b]Dependent on boiler size e.g. < 100MW thermal input requires a monitoring frequency of 2/year, >100MW requires CEM

Table 2.8 Continuous ambient air quality monitoring undertaken by operators of power stations (Eggborough, Ferrybridge and Drax) in the Aire Valley as part of an IPPC operating permit.

Monitor Location	Pollutant monitored	Concentration averaging time
Carr Lane	Sulphur dioxide (SO_2)	15-min and 1-hour
Hemingbrough Landing	Sulphur dioxide (SO_2) Nitrogen dioxide (NO_2)	15-min and 1-hour
Downes Ground	Sulphur dioxide (SO_2)	15-min and 1-hour
West Bank	Sulphur dioxide (SO_2)	15-min and 1-hour
Smeathalls Farm	Sulphur dioxide (SO_2)	15-min and 1-hour
North Featherstone	Sulphur dioxide (SO_2)	15-min and 1-hour

Table 2.9 On- and off-site ambient monitoring requirements for landfill sites.

On-site monitoring			
Location	**Pollutant**	**Monitoring frequency**	**Trigger Level**
Site boundary (ambient air)	Methane (CH_4)	12/year or following a substantiated odour complaint	10 ppm
	Hydrogen Sulphide (H_2S)	12/year or following a substantiated odour complaint	5 ppm (average of 10 consecutive readings)
	Speciated VOCs	12/year or following a substantiated odour complaint	EAL or species odour threshold[b]
Gas infrastructure integrity survey	Methane (CH_4)	12/year	10 ppm or 100 ppm[c]
Permanently capped areas	Methane (CH_4)	1/year[a]	$1x10^{-3}$ mg m^{-2} s^{-1}
Temporarily capped areas	Methane (CH_4)	1/year[a]	$1x10^{-1}$ mg m^{-2} s^{-1}
Off-site monitoring			
Location	**Pollutant**	**Monitoring frequency**	**Trigger Level**
Maximum ambient methane exceedence location	Hydrogen Sulphide (H_2S)	When CH_4 trigger level exceeded or following a substantiated odour complaint	5 ppm (average of 10 consecutive readings)
Maximum ambient methane exceedence location	Speciated VOCs	When CH_4 trigger level exceeded or following a substantiated odour complaint	EAL or species odour threshold[b]

[a]Monitoring should only take place under calm conditions (< Beaufort Force 2) unless an odour complaint is substantiated.

[b] Environment Agency H1 Methodology (Environment Agency, 2002a) Environmental Assessment Level (EAL) or VOC species odour threshold, whichever is the greatest.

[c]10 ppm for leachate well heads, all pre-combustion pipework, manifolds, knockout pots and booster fans. 100 ppm for leachate well annuli.

2.3.3 Air quality monitoring: Environment Agency

In order to assist Part A(1) process operators to assess the impact of their aerial emissions on ambient air quality, or where the public have expressed concern over poor air quality arising from industrial emissions, the Environment Agency is able to conduct short-term monitoring campaigns. Sampling is undertaken using one of four Mobile Monitoring Facilities (MMFs), capable of continuously monitoring ambient concentrations of PM_{10}, $PM_{2.5}$, SO_2, NO_2, CO, hydrogen sulphide (H_2S), methane (CH_4) and a range of VOC species. Meteorological conditions (atmospheric pressure, temperature, relative humidity, wind direction and wind speed) are also measured by the MMFs. Table 2.10 summarises a number of short-term MMF campaigns undertaken by the Environment Agency in recent years.

Monitoring campaigns are typically four to six months in duration. Annual air quality statistics are determined on a pro-rata basis, making the assumption that the meteorological conditions that prevailed during the monitoring period are representative of a typical year. In order to take into account seasonal variation in meteorology, monitoring campaigns should ideally include an equal number of winter and summer months as detailed in the M8 Technical Guidance Note (Environment Agency, 2000). However, as can be seen from Table 2.10, this is often not the case.

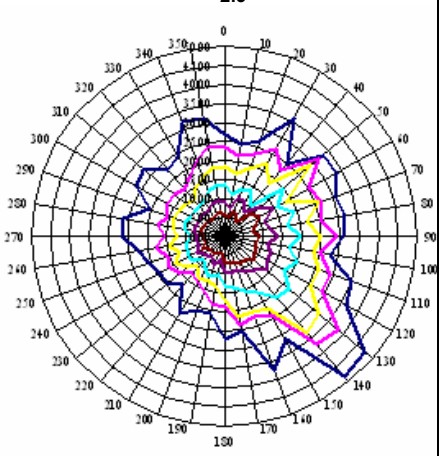
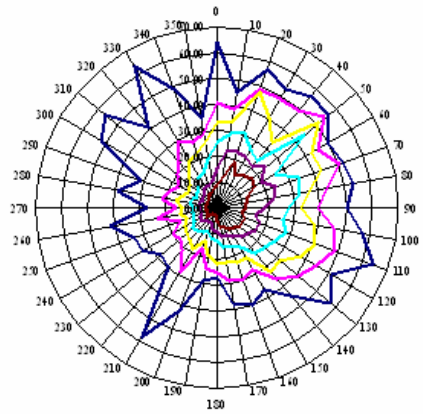
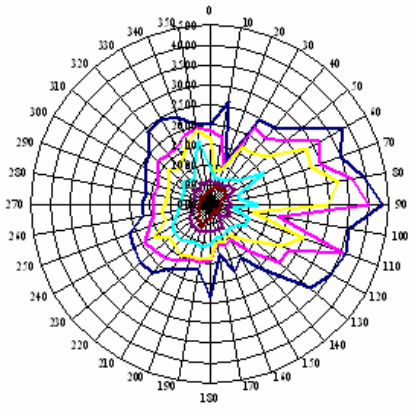
The analysis of data from MMF campaigns includes a comparison with AQS objectives and a detailed consideration of significant pollution events that occurred during the monitoring campaign. The pollution events do not necessarily constitute exceedences; rather they are described as events during which pollutant concentrations increased significantly above the average level. Each event is considered separately and the prevailing wind speed and wind direction examined throughout the duration of the event. The maximum pollutant concentrations during the events are also summarised, along with the corresponding wind speed, wind direction and time of the monitored maximum concentration.

Various directional analysis methods are applied to the monitored data to determine source apportionment. Routine analysis includes the creation of directional concentration plots, either of annual mean concentration or percentile breakdown. The directional concentration plots indicate whether sources impact consistently or only occasionally and whether higher or lower percentiles (or both) are affected. An example of the interpretation of directional percentile concentration plots is given in Table 2.11, with examples taken from a report on ambient air quality at Pen-y-ffordd, Flintshire (Shutt *et al*., 2005). Further directional analysis explores the variation in pollutant concentrations for various wind sectors, comparing pollutant concentrations with wind speed and/or time of day. An example of the evaluation of pollutant concentration by wind speed and wind sector is given in Figure 2.2, taken from a report on ambient air quality at Aberthaw (Sheppard *et al*., 2001).

Table 2.10 Selected short-term monitoring campaigns undertaken by the Environment Agency during the period 2001-2005 using Mobile Monitoring Facilities (MMFs).

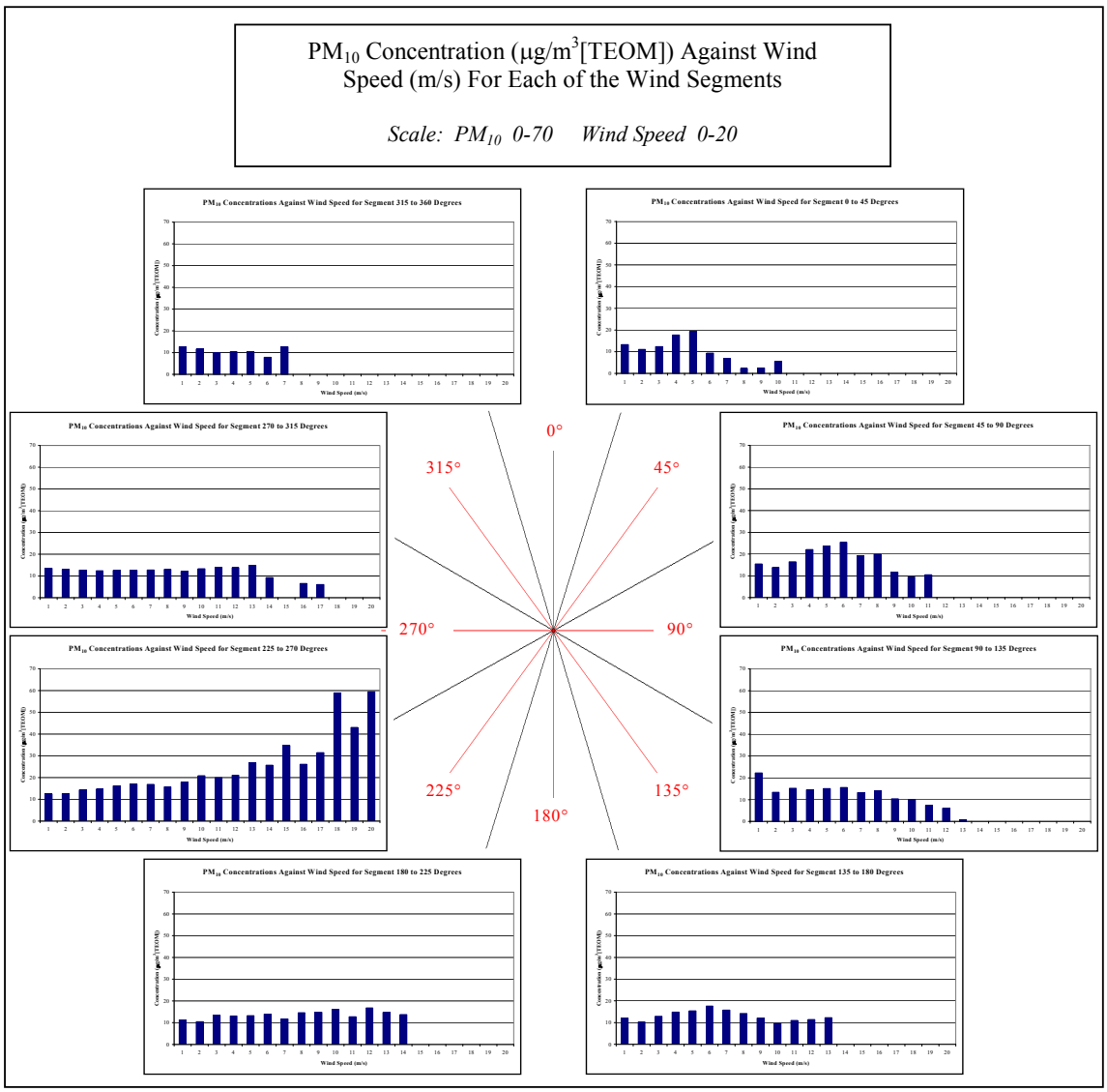
Monitoring site	Duration of monitoring	Purpose of campaign	Pollutants monitored
Port Talbot	5 months (21 June 2002 - 22 Oct 2002)	Assess impact of Corus steel works on Neath AQMA.	PM_{10}, SO_2, NO_2, CO, H_2S, selected VOCs (e.g. benzene)
Dowlais, Methyr Tydfil	5 months (17 Dec 2002 - 14 Apr 2003)	Assess impact of Trescatti landfill on local air quality	PM_{10}, SO_2, NO_2, CO, H_2S, selected VOCs (e.g. benzene), CH_4
Llanidloes, Powys	5 months (30 June 2003 - 28 Oct 2004)	Assess impact of Bryn Posteg landfill on local air quality	PM_{10}, SO_2, NO_2, CO, H_2S, selected VOCs (e.g. benzene), CH_4
Clydach Vale, Rhondda	4 months (29 Oct 2003 - 3 Feb 2004)	Assess impact of closed Nant-y-Gwyddon landfill on local air quality following public complaints	PM_{10}, H_2S, selected VOCs (e.g. benzene), CH_4
Port Talbot	7 months (20 Jul 2004 - 18 Jan 2005)	Assess impact of local pollution sources on air quality within an AQMA.	PM_{10}, SO_2, NO_2, CO, selected VOCs (including benzene and 1,3 butadiene), CH_4
East Aberthaw	6 months (21 Jan 2005 - 21 June 2005)	Assess impact of Aberthaw power station and cement works on local air quality	PM_{10}, SO_2, NO_2, CO
Scunthorpe	5 months (25 May 2005 - 15 Sept 2005)	Assess impact of Corus steel works on local air quality	PM_{10}, SO_2, NO_2, CO
Pen-y-Ffordd/ Dyke Farm, Flintshire	4 months (30 June 2005 - 26 Sept 2005)	Assess impact of Padeswood Cement works on local air quality.	PM_{10}, SO_2, NO_2, CO
	10 months (17 Nov 2004 - 16 Sept 2005)	Additional SO_2 analyser for triangulation of data from two sites.	SO_2 only

Table 2.11 Examples of emission source characteristics derived from the interpretation of pollution rose percentile breakdown.

Scenario	Interpretation	Example of percentile breakdown pollution rose
Contribution from a source from 50°-150° affecting lower percentiles. Contribution from a source from 130°-140° affecting high percentiles.	Sources in wind segment 50°-150° emit levels slightly higher than average but never enough to activate the higher percentiles. Source between 130°-140° emits relatively continuously and affects all percentiles.	$PM_{2.5}$ 
Contribution from sources between 0°-170° and 210°-330°. Concentration peaks affect the lower percentiles between 0°- 170° and the higher percentiles between 210°-330°.	Sources are assumed to emit continuously from between 0°- 170° but are never high enough to elevate the higher percentiles. Sources from 210°-330° only affect higher percentiles and are therefore intermittent.	NO_2 
Contribution from a source(s) from 50°-60°, 80°-90° and 110°. Concentration peaks noted to affect higher percentiles but not as evident in lower percentiles.	Sources only affect higher percentiles and are therefore intermittent and do not always affect SO_2 concentrations every time the wind is originating from these directions.	SO_2 

Source: Shutt *et al*. (2005).

Note: Percentiles are indicated by coloured lines on the pollution rose, from inner to outer, lines represent 25th 50th, 75th, 90th, 95th and 99th percentiles).



Source: Sheppard *et al*. (2001)

Figure 2.2 Directional analysis of PM_{10} concentrations against wind speed, measured during an Environment Agency MMF campaign at Aberthaw in 2000/2001 (from Sheppard *et al*., 2001).

2.4 Uncertainties in air quality monitoring

Although ambient monitoring can provide relatively accurate measurements of air quality at a single point in space and time, the number of monitors employed around industrial point sources in the UK and the frequency at which monitoring is undertaken is often limited. As a consequence, accurate temporal and spatial variations in air quality (environmental uncertainty) may not be captured. In addition, the error associated with a particular sampling instrument or analytical method may also introduce an element of measurement uncertainty.

2.4.1 Measurement uncertainty

Measurement uncertainty incorporates both sampling and analytical variability. Sampling variability is a measure of the inconsistencies in sampling methods and can be either random or systematic. For example, random variability may be introduced due to fluctuations in temperature or pressure that may affect the measurement instrument, whereas systematic variation may result from a built-in bias in the measurement equipment or procedure. Random sampling variability can be measured by the difference in concentrations between 'like samples', for example by collecting duplicate samples from the same location. In contrast, analytical variability is the difference in concentration measurements between 'like laboratory analysis', conducted on a common sample.

AQF Daughter Directives 1 and 3 (1999/30/EC, 2002/3/EC) recommend measurement accuracy objectives for the continuous measurement of SO_2, NO_2 and NO_x as 15% and for PM_{10} and Pb, 25%. In addition, the Environment Agency recommends quality control procedures for ambient air quality monitoring in Technical Guidance Note M8 (Environment Agency, 2000) which aims to limit sampling and analytical variability.

2.4.2 Environmental uncertainty

Environmental uncertainty describes the spatial and temporal variability in monitored concentrations. Spatial variability is the difference in concentration measurements over an area, and gives a measure of the degree to which measurements made at one point are representative of other locations in space. Temporal variability is the difference or change in concentrations over time (daily, seasonally or inter-annually). Capturing sufficient temporal variability depends on the capability of the sampling rate, whereas the number of monitoring sites needed to adequately capture the spatial variability depends on the properties of the pollutant, the nature of the source and the area over which the impacts are being measured.

2.4.2.1 Extrapolation from short-term monitoring campaigns

Short-term monitoring campaigns undertaken by the Environment Agency aim to monitor air quality around Part A(1) sources to assess the impact of installations in relation to AQS objectives. The outcome of these assessments is then fed back into permit conditions which set controls on the plant. As stated in Section 2.3.3, MMF monitoring campaigns are usually limited to six-months or less, with annual air quality statistics derived on a pro-rata basis, assuming that meteorological conditions which occurred during the monitoring period are representative of a typical year. Uncertainties associated with the derivation of annual air quality statistics from short-term monitoring campaigns are likely to be high. This can have implications for permit holders,

especially when the total number of calculated exceedences is close to or exceeds an AQS objective. For example, the Environment Agency conducted a short-term monitoring campaign to study ambient air quality in the vicinity of Aberthaw power station between August 2000 and January 2001 (Sheppard *et al*., 2001). The five-month campaign recorded 16 occasions when the 15-minute mean SO_2 concentration exceeded the objective (266 $\mu g\ m^{-3}$) set for this pollutant. As a result, it was projected that 266 $\mu g\ m^{-3}$ would be exceeded 38 times during the course of a single year, compared to the permitted 35 occasions allowed under the AQS.

In addition, MMF's are often sited where air quality complaints have been made by the local community (Sheppard *et al.*, 2001) or to complement existing monitoring locations (Shutt *et al.*, 2005). On many occasions, MMF's are not located in the predicted area of peak concentration or in the prevailing wind direction because of practical constraints e.g. the availability of power supply or site accessibility. It is necessary to consider the implications of these factors if short-term monitoring campaigns are used to try to assess likely compliance with air quality objectives.

2.4.3 Quantifying monitoring uncertainty

To decide how resources may best be allocated to reduce uncertainty in monitoring data, it is useful to determine the relative contributions of the different sources of variability. Bortnick and Stetzer (2002) applied a statistical model for partitioning and quantifying different sources of variability using data from the United States Environmental Protection Agency (USEPA) Urban Air Toxics Monitoring Program (UATMP). An analysis of variance (ANOVA) model with random effects was applied to a data set consisting of duplicate (co-located samplers) and replicate (repeat laboratory analysis) samples of VOC and carbonyl air pollutants sampled between 1996 and 1999. The method is similar to traditional ANOVA models in that each effect (spatial, temporal, sampling and analytical) is treated as a categorical factor (site, day, sampler and analysis) with multiple levels ("low", "medium" and "high" categories). However, unlike traditional ANOVA, the effects of the levels of each factor are treated as random deviations from a null effect. The random effects ANOVA model separates out each contributing component as variability corresponding to each random factor included in the model and may be written as follows:

$$\ln(y_{ijkl}) = \mu + \alpha_i + \beta_{j(i)} + \gamma_{k(ij)} + \varepsilon_{/(ijk)} \qquad \textbf{Equation 1}$$

where y_{ijkl} is the concentration for the *l*th replicate analysis (*l* = 1, 2) of the *k*th duplicate sample (*k* = 1, 2) collected on the *j*th day (*j* = 1, 2…, <7) at the *i*th monitoring location (*i* = 1, 2…, <10), μ the overall mean across all the data, α_i the random deviation from the overall mean of the data that is due to the spatial variability, $\beta_{j(i)}$ the random deviation from a site-specific mean due to the temporal variability, $\gamma_{k(ij)}$ the random deviation from a site $\times$ day-specific mean due to the sampling error and $\varepsilon_{l(ijk)}$ the residual error of the model due to the effect of the analytical error. Bortnick and Stetzer (2002) refer the reader to Searle et al. (1992) and Searle (1971) for a more in depth discussion on classifying effects in the model as fixed versus random and on estimating variance components.

The output from the model is a partitioning of the total data variability into each component (α_i, $\beta_{j(i)}$, $\gamma_{k(ij)}$ and $\varepsilon_{l(ijk)}$). Each individual variance component was transformed into an estimate of percent relative error (known as the % coefficient of

variation, %CV) using a formula which expresses the ratio of mean to standard deviation for log-normal data (Equation 2):

$$CV = 100 \times \left(\exp\left(\sigma^2\right) - 1\right)^{0.5} \qquad \textbf{Equation 2}$$

Environmental variability (especially temporal variability) was shown by Bortnick and Stetzer (2002) to be much more significant than measurement variability, as is shown in the results for benzene in Table 2.12. Note that Equation 2 (above) is a non-linear function of the individual variances and therefore the individual relative error components do not sum to the overall relative error. Temporal variability was not always found to be the most significant source of error; for some pollutants sampling variability became more important at lower concentrations. In addition, the proportion of sampling versus analytical error in measurement variability was found to vary with compound. Although this study is applied to a national network with 17 monitoring sites spread over multiple cities, the random effects ANOVA model methodology could also be applied to determine the partitioning of relative errors in monitoring data around single point sources. This would determine how best to allocate monitoring resources to limit the uncertainty in regulatory decisions based on monitoring data. However, in order to apply this model it is necessary to obtain data with repeated measures of each factor, at each level. Although unlikely with the current provision of monitoring sites around industrial point sources, the development of low cost, high accuracy sensor technologies (See Section 5.4) may enable such analysis to be carried out before permanent monitoring networks are installed.

Table 2.12 Sources of variability when monitoring benzene (from Bortnick and Stetzer, 2002).

Conc. level	No. of observations	No. of Sites	Mean No. of days	Total relative error (%CV)	Relative error due to environmental variability*		Relative error due to measurement variability*	
					Spatial	Temporal	Sampling	Analytical
Low	89	5	3	50.53	11.09	47.77	<1.00	9.83
Medium	81	5	4	39.06	15.67	27.55	<1.00	21.34
High	125	5	4	46.93	12.24	44.38	<1.00	6.63

*Results are expressed as a percentage coefficient of variation (%CV), calculated by the ratio of standard deviation to the mean.

3 Modelling ambient air quality

3.1 Modelling ambient air quality in the UK

To comply with PPC and IPPC regulations, as part of a permit application all industrial process operators are required to undertake an assessment of the impacts arising from aerial emissions with reference to national air quality objectives. The impact assessment is carried out using the H1 methodology (Environment Agency, 2002a) which requires the calculation of process contributions to air and the subsequent screening out of insignificant emissions. Advanced modelling studies are required if a H1 assessment reveals that the predicted emissions are likely to be significant. Compliance is monitored under the requirements of the specific PPC application, which may involve air quality monitoring by the operator. However, power station operators are required to carry out advanced dispersion modelling on an annual basis and continuous monitoring throughout the year to demonstrate compliance with AQS objectives. This is carried out by means of a Risk Management Framework (Hunter, 2004). Although the majority of modelling undertaken for PPC permit applications is undertaken by the process operator, the Environment Agency Air Quality Modelling and Assessment Unit (AQMAU) routinely audits air quality modelling assessments undertaken as part of PPC permit applications, compliance, enforcement and incident investigations.

A range of numerical dispersion modelling software packages is available for advanced modelling studies, which allows the prediction of ambient ground level pollutant concentrations on a far greater spatial and temporal resolution that can be achieved by monitoring alone. For point source emissions, simple Gaussian dispersion models such as R91 (Clark, 1979) use the Pasquill Gifford (PG) stability category scheme to determine the dispersion characteristics of the atmosphere. However, more advanced '*new generation*' Gaussian dispersion models, e.g. ADMS and AERMOD, describe the atmospheric boundary layer using continuous parameters (e.g. boundary layer height, h, Monin-Obukhov length, L_{MO}, sensible heat flux, $F_{\theta 0}$, and friction velocity, u_*), to allow continuous variation of boundary layer properties both spatially and temporally. Both ADMS (Carruthers *et al*., 1994) and AERMOD (Cimorelli *et al*., 1998) are used for regulatory purposes in the UK and are capable of simulating a wide range of buoyant and passive pollutant releases to the atmosphere. However, any model may be used, provided it is proven as 'fit-for-purpose'. Since the release of ADMS and AERMOD, both models have undergone continuing development. At the time of publication of this report, the latest available versions are ADMS 4 and AERMOD (07026).

3.2 Uncertainties in air quality modelling

Although the use of dispersion models can increase the spatial and temporal coverage of ambient pollutant concentrations, the uncertainties associated with the model predictions are much more complex than those associated with air quality monitoring. Model uncertainty defines the ability of a model to accurately simulate atmospheric dispersion. This depends not only on the input of accurate data into the model, e.g. meteorological data and emissions data, but also on the choice of model and the model configuration options selected by the user.

3.2.1 Choice of model and user configuration

Both ADMS and AERMOD use meteorological pre-processors to determine meteorological parameters necessary for estimating profiles of wind, turbulence and temperature from standard meteorological measurements. However, subtle differences in the way in which the ADMS and AERMOD meteorological pre-processors calculate surface parameters and the subsequent calculation of the boundary layer structure, can lead to differences in model output. A fundamental difference between ADMS and AERMET (the meteorological pre-processor used in the AERMOD system) is the parameterisation of the Atmospheric Boundary Layer (ABL). AERMET estimates of boundary layer height can be up to two or three times those of ADMS and furthermore, AERMET often predicts stronger unstable conditions (Brooke *et al*., 2003; Auld *et al*., 2003). Thé *et al*. (2001) suggest this may be due to the estimation tool used to allow AERMET to process meteorological data without upper air data. For the modelling of emissions from tall stacks (> 60 m) and highly buoyant plumes, the discrepancy in estimation of *h* can have a large effect on modelled concentration predictions. This is of particular importance during low boundary layer conditions, where small differences in boundary layer height can determine whether a plume rises above the boundary layer or is trapped beneath it. The reader is referred to Auld *et a*l. (2003) for a detailed description of the similarities and differences in the meteorological modelling approaches using ADMS 3 and AERMOD and the potential impacts of these approaches on dispersion calculations.

The difference in the treatment of terrain and buildings in ADMS and AERMOD is also a source of inconsistency in model output. The Environment Agency carried out an extensive model comparison study, focusing on an assessment of AERMOD (99351), AERMOD-PRIME (04300)[1] and ADMS 3.1 (Environment Agency, 2002b). The study tested the effects of plume rise, building entrainment and terrain on annual and hourly mean concentrations predicted by each of the three models using single meteorological conditions extracted from Lyneham meteorological data (1995). The results for the building entrainment analysis show that the difference between AERMOD (99351) and AERMOD-PRIME (04300) was limited to the wake region of the building, reflecting the inability of the earlier version of AERMOD to model the effects of building entrainment. Stack location and wind direction were both observed to have a significant impact on dispersion predictions from ADMS 3 and AERMOD-PRIME (04300). It is noteworthy that the treatment of buildings is approximate and neither model can be expected to provide an exact representation of flow and dispersion around such objects.

In areas of complex terrain, the accurate calculation of the wind-field is of primary importance to be able to predict dispersion. Both ADMS and AERMOD have algorithms for determining the influence of terrain features on dispersion, however the wind field is calculated differently in both models. ADMS uses a linear analytical solution to determine the wind field, whereas AERMOD does not model the wind field explicitly but instead models the interaction of the dispersing plume with the underlying terrain. Several authors have found substantial differences between ADMS 3 and AERMOD when modelling the effects of complex terrain on atmospheric dispersion (Environment Agency, 2002b; Brooke and Stiff, 2006; Hill *et al*., 2005). Much of the meteorological data applied in the UK for regulatory modelling is derived from distant meteorological sites. As such, consideration should be given to how representative meteorological data may be, when applied to dispersion modelling in areas of complex terrain. Additionally, if a greater number of monitoring sites were placed around sources in areas of complex terrain, more rigorous model validation studies could be undertaken. The reader is referred to Hill *et al*. (2005) for a comprehensive review of dispersion

[1] AERMOD-PRIME refers to the version of AERMOD (identified by a Julian date) which includes the PRIME plume rise and building downwash algorithms. Previous version of AERMOD also contained downwash and plume rise algorithms, although these algorithms were somewhat simpler than PRIME.

modelling in areas of complex terrain and Finardi *et al*. (1997) for guidance on model suitability for areas of complex terrain.

Continuous improvements in model development also mean that different versions of the same model can provide different results. Several model comparison studies compare earlier and later versions of ADMS and AERMOD (Brooke *et al*., 2003; Brooke and Stiff, 2006; Environment Agency, 2002b; Futter, 2000; Sidle *et al.,* 2004). Comparisons of the latest version of ADMS (Version 4) with previous versions of ADMS and AERMOD can be found in ADMS 4 model validation documentation (See: http://www.cerc.co.uk/software/publications.htm).

In addition to the intrinsic differences between ADMS and AERMOD, model user configuration can give rise to discrepancies in dispersion parameterisation, which can significantly influence model predictions. Table 3.1 defines the major differences in configuration options in ADMS 3 and AERMOD. Some of the most important aspects of model user configuration to be considered are detailed below.

3.2.1.1 Length of averaging period

Length of averaging period is a minor difference between ADMS 3 and AERMOD but one that is worth considering. AERMOD does not have the capacity to calculate concentrations for averaging times less than 1 hour, which is of particular importance for estimating concentrations for the 15-minute mean SO_2 objective. However, it may be argued that the resolution of meteorological data (normally one hour) may limit the capability of either model to produce accurate concentration predictions at a temporal frequency greater than once per hour. Under strongly convective conditions with high boundary layers, hourly average concentrations may vary significantly due to the time scales of convective motions, therefore, a period of several hours may be necessary to obtain a representative average (Auld *et al*., 2003).

3.2.1.2 Treatment of low wind speeds

The way models are configured to treat calm conditions is of particular importance. Calm conditions are synonymous with low wind speeds in both stable and unstable conditions and are often associated with peak concentration events, particularly for low-level sources. It is therefore necessary that these conditions can be accurately modelled. However, low wind speeds tend to give rise to uncertainty due to their effect on the dispersion of plumes and also from their treatment within models.

The minimum wind speed required to detect air flow varies between anemometers. For Numerical Weather Prediction (NWP) data, this speed is assumed to be 1 m s^{-1}, however site-specific measurements frequently use instruments that can detect speeds below this value. ADMS 3 and AERMET impose a minimum allowable wind speed for use in estimating the boundary layer parameters, below which dispersion calculations are not completed. For ADMS 3, the minimum allowable wind speed is 0.75 m s^{-1} (CERC, 2007) and for AERMOD this value is 0.28 m s^{-1} (Environmental Protection Agency, 2004). However minimum wind speeds may be set above these values if the minimum detectable (threshold) wind speed of the site-specific anemometer is higher than 0.75 m s^{-1} in the case of ADMS or 0.28 m s^{-1} in the case of AERMOD. Auld *et al*. (2003) compared peak (100^{th} percentile) short-term modelled concentrations in AERMOD using the default minimum wind speed of 0.28 m s^{-1} and the ADMS 3 minimum wind threshold value of 0.75 m s^{-1} and found peak concentrations were reduced by almost 50% when using the ADMS threshold value. When reporting the percentiles of a concentration distribution, it would therefore be advisable for modellers to report the minimum threshold wind speed used and reference the frequency of low

wind speed periods, i.e. the number of model time steps skipped due to below-threshold wind speeds.

If measured concentration exceedences commonly occur during low wind speed conditions, then consideration should be given to an alternative model which has the ability to treat calm conditions and low wind speeds more realistically, e.g. CALPUFF (Scire *et al*., 2000). ADMS 4 offers a new approach to treat calm conditions whereby at speeds of less than or equal to a user defined minimum wind speed (minimum allowed is 0.3 m s^{-1}), a radially symmetric solution is applied, assuming equal probability of all wind directions. At wind speeds above a critical value (U_{crit}), standard ADMS calculations for a Gaussian plume are applied and for meteorological lines with U between 0.5 m s^{-1} and U_{crit}, a weighted average of both solutions is applied (CERC, 2007).

3.2.1.3 Surface roughness length

The configuration of the roughness length (z_o) parameter, used by meteorological pre-processors to interpret the vertical profile of wind speed and the estimation of friction velocities, can also significantly affect model output. Historically, single values of z_o were used to represent surface characteristics at the emission source. More recently, it has been possible to include additional values for the meteorological mast site. Further developments in ADMS 3 now allow the input of a z_o grid over the model domain, whilst the AERMET preprocessor in AERMOD allows roughness length to vary on a seasonal and sectoral basis. Following an inter-comparison of the meteorological pre-processors in ADMS 3 and AERMOD, Auld *et al*. (2003) concluded that the parameterisation of surface roughness length was the largest contributor to differences in predicted ambient concentrations between the two models.

Table 3.1 ADMS (Version 3) and AERMOD configuration options.

Configuration option	ADMS	AERMOD
Meteorological data source	Observational or NWP data Meteorological Office data sets may be supplied in ADMS format	Observational data (including upper air measurements) or NWP data SCRAM, HUSWO, CD-144, NCDC and SAMSON formats may be used as input to AERMET
Meteorological data format	Hourly input (0-23) where the time refers to the mid-point of each hour	Hourly input (1-24) where the time refers to the end of each hour
Minimum averaging time	Sub-1 hour	1 hour
Threshold wind speed below which dispersion cannot be calculated	0.75 m s^{-1}	0.28 m s^{-1}
Surface energy balance parameters	Albedo (r) default value 0.23 (possible to input hourly) Priestly-Taylor Parameter (α) default value 1 (equivalent to β = 0.6), possible to input hourly)	Seasonal values of Albedo (*r*) and Bowen ratio (β) input on sector-based level
Surface roughness	Values of z_o can be assigned for the emission site and the meteorological site, in addition to a roughness length grid covering the model domain	Values of z_o can be assigned for the emission site and the meteorological site and also on a seasonal and sectoral basis
Boundary Layer height limits	50 – 4000 m	4 – 4000 m
Boundary layer parameterisation	Estimates *h* by incorporating mechanical and convective turbulence into boundary layer growth formulae	Estimates *h* due to convective and mechanical turbulence individually
Treatment of stacks	Capacity to model hourly varying flue diameter	Must represent multi-flue stacks as separate sources
Treatment of buildings	A maximum of 10 buildings, requires choice of main building to be made	Unlimited number of buildings
Treatment of complex terrain	Terrain data may be imported using a terrain file	Gridded terrain data may be imported using the AERMAP terrain preprocessor

3.2.2 Uncertainty in meteorological data

A major source of uncertainty in air quality modelling is due to the measurement and treatment of meteorological data. In Gaussian dispersion models, airflow measurements are typically parameterised by wind speed, wind direction, and atmospheric stability, which may be derived from temperature and cloud cover or solar radiation measurements. Uncertainty can therefore be introduced through inaccuracies in meteorological measurement instruments. However, uncertainty is also introduced through the use of meteorological data that are not representative of the conditions at the dispersion site.

3.2.2.1 Accuracy of measurement methods

The accuracy of measurement methods can be divided in terms of *measurement precision* and *methodological bias*.

Measurement precision is considered to be the amount by which a recorded value deviates from an accepted standard value. This standard can be quoted for individual parameter sensors or for the measurement system as a whole, allowing for data logging and signal processing limitations. The US EPA recommends system accuracy for *in situ* measurements of a number of variables for use in air quality models, based on standard averaging times of 1 hour (Table 3.2). The UK Met Office and other suppliers of UK meteorological data (e.g. ADM LTD) also follow these guidelines.

Methodological bias is more difficult to quantify than measurement precision and its importance varies with each meteorological parameter. Although not mentioned in Table 3.1, cloud cover is an important input in the calculation of further boundary layer parameters in dispersion models. Cloud cover is recorded in oktas (eighths of sky covered) as part of the UK Met Office standard observing routine. However, weather stations are becoming increasingly automated, with measurements of cloud density and height often being taken from laser cloud base recorders (LCBRs). A typical error of ± 1 okta may occur in the daytime if these data are manually recorded which may increase to ± 2 or 3 oktas at night due to the limitations of darkness (Auld *et al*., 2003). This uncertainty may be reduced with the use of LCBRs although rather than take in the whole arc of the sky, a LCBR will only measure cloud cover at the highest point above the horizon and hence may not detect the presence of an approaching bank of cloud. In addition, the models themselves cannot account for the differences between high and low level cloud, which may exert a greater or lesser influence on surface heat fluxes.

3.2.2.2 Representivity of meteorological parameters

Figure 3.1 shows the location of Met Office Automatic Weather Stations (AWS) in the UK, in relation to Part A(1) industrial sources. Representivity of meteorological data is an important aspect of model uncertainty. Data are considered representative if obtained from a site similar in land use, geographic location and terrain to the emission site. However, representivity can also include issues surrounding data quality, encompassing aspects such as the size of the meteorological data set and the presence of gaps in the data.

Table 3.2 System resolution and accuracy for *in-situ* digital measurements recommended by the US-EPA (from Auld *et al*., 2003)

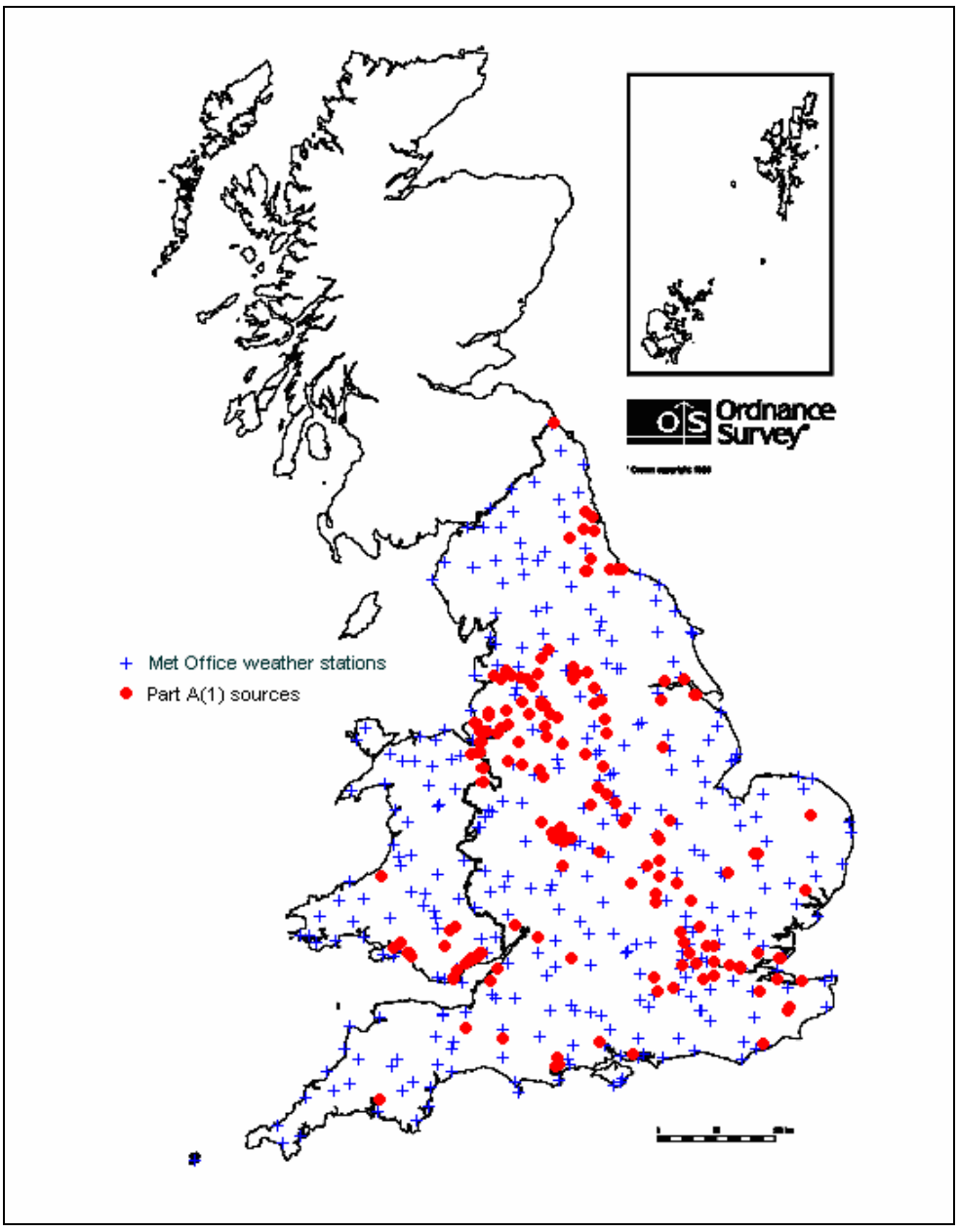
Variable	System Accuracy	Measurement resolution	Sensor specification
Wind speed	± 0.2 m s^{-1} (+ 5% of observed)	0.1 m s^{-1}	Threshold ≤ 0.5 m s^{-1}
Wind direction	± 5 degrees	1 degree	Threshold ≤ 0.5 m s^{-1} at 10 degrees
Ambient temperature	± 0.5 °C	0.1 °C	Time constant ≤ 1 min
Vertical temperature difference	± 0.1 °C	0.02 °C	Time constant ≤ 1 min
Precipitation	± 10% of observed or ± 5 mm	0.3 mm	
Solar radiation	± 5% of observed	10 W m^{-2}	2nd class standard pyranometer Spectral response 285 - 2800 nm
Pressure	± 3 mb (0.3 kPa)	0.5 mb	
Dew point temperature	± 1.5 °C	0.1 °C	Time constant ≤ 30 min

Bethan and Teasdale (2005) investigated the representivity of meteorological data by modelling dispersion of SO_2 from a generic power station emission source using three, 30-year meteorological data sets. The three data sets, Lyneham, Waddington and Aberporth, were contrasting in terms of both site characteristics and distance from the emission site. The meteorological years that produced the highest and lowest values of 99.9th percentile concentrations and exceedences above 266 μg m^{-3} (X266) were ranked accordingly (Table 3.3). On only one occasion did the year match for ranked statistics between meteorological data sets, highlighting the influence of meteorological data on modelled concentrations.

Bethan and Teasdale (2005) also demonstrated that variations in meteorology can cause the location of predicted maximum impact to vary both between years and within a single year. Figure 3.2 shows that between 1977 and 1980, the location of predicted maximum concentrations varies significantly, with the maximum location falling anywhere within an area as large as 3 km^2 for 100% load and 4 km^2 for 60% load. It is worth noting that the distance to maximum impact from the source is relatively consistent in each case. The authors conclude that for a power station with a single monitoring site, the site would lie outside the representative area[2] for one in every six years, and for a power station with two monitoring sites, the sites would lie outside the representative area for one in every ten years. The observations of Bethan and Teasdale (2005) have implications for monitoring stations that are sited according to

[2]According to AQMP compliance methodology, a representative monitoring site must be located within a contiguous area, bounded by a modelled 50% concentration isopleth corresponding to 50% of the maximum modelled 99.9th percentile 15-minute SO_2 concentration. Note: this is not analogous to the definition of meteorological data representivity.

the location of maximum concentration predicted by models configured with five years of meteorology. To minimise the uncertainty associated with non site-specific meteorological data, the Environment Agency's Air Quality Modelling and Assessment Unit (AQMAU) recommend using five years of representative hourly sequential meteorological data to account for inter-annual variability when performing dispersion modelling assessments for annual licensing purposes (Shi and Ng, 2002).



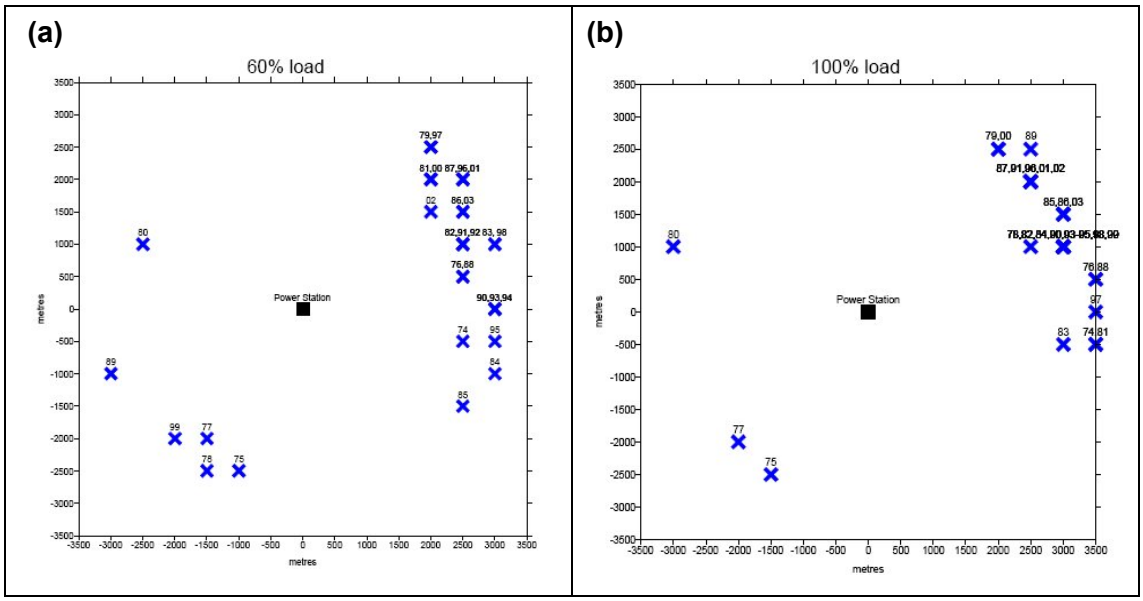
Source: Met Office weather station data supplied by the Met Office and Part A(1) site data supplied by the Environment Agency.

Figure 3.1 Part A(1) industrial sources and Met Office Automatic Weather Station sites open on 4th June 2006, measuring a minimum of wind speed, wind direction and dry bulb temperature.

Table 3.3 Ranking of years producing the highest and lowest modelled values of the 99.9th percentile 15-minute mean SO_2 concentration (99.9th) and the number of exceedences of 266 μg m^{-3} (X266) using three meteorological data sets (from Bethan and Teasdale, 2005).

Rank	Lyneham		Waddington		Aberporth	
	99.9th	X266	99.9th	X266	99.9th	X266
1st	1977	1977	1977	1977	2002	1990
2nd	1993	1993	1975	1975	1994	2002
3rd	2002	2002	1990	1986	1990	1994
4th	1984	1994	1986	1980	2001	2001
5th	1994	1976	1982	1982	1991	1996
29th	1989	1987	2003	1996	1982	1993
30th	1987	1995	2001	1987	1974	1989

Note: Modelling undertaken using a generic emission source for a full load scenario.



Source: Bethan and Teasdale (2005)

Figure 3.2 Location of maximum 99.9th percentile 15-minute mean SO_2 concentrations predicted using ADMS 3 configured for a generic power station operating at (a) 60% and (b) 100% load and using 30 years of meteorology between 1974 and 2003.

An alternative to using data provided by the Met Office is the use of site-specific meteorological data provided the fetch requirements for obtaining representative meteorological measurements are met by minimising the influence of local structures. However, site-specific data are used infrequently by the UK modelling community due to the lack of sufficient instrumentation to provide inputs containing all the necessary parameters for dispersion modelling (Auld *et al*., 2003).

3.2.2.3 Numerical Weather Prediction (NWP) data

Auld *et al*. (2003) suggest that the use of Numerical Weather Prediction (NWP) data should be considered when the nearest surface observation station is unrepresentative of the source location. NWP is the method behind modern weather forecasting and its primary purpose is to predict the state of the atmosphere into the future. In the current UK Met Office system, a mesoscale configuration of the Unified Model (UM) is used to generate NWP data in order to provide detailed weather forecasts for the UK. To generate these forecasts, the model requires meteorological observation data to be assimilated into the model from sources such as weather stations, radiosondes and satellites. The mesoscale model has a horizontal grid resolution of 12 km and a vertical resolution of 38 levels and provides advanced boundary layer parameters, which in atmospheric dispersion models are normally calculated using similarity profile relationships.

Nelson *et al*. (2002) reviewed the potential of NWP data for input into dispersion models and considered this a promising alternative in the search for representative meteorological data. Kidd (2002) compared meteorological measurements obtained at Sellafield, West Cumbria, with NWP data from the nearest grid location to assess the influence of these data sets on ADMS model predictions. The author acknowledged that NWP data and site measurements showed good agreement, both spatially and temporally, although the observed data tended to show more extreme conditions.

Lucas and Bethan (2004) compared concentrations modelled using NWP data, with those modelled using standard meteorological data supplied by the Met Office, to determine air quality around four UK power stations. In this study the NWP data were derived from the global configuration of the UM with a horizontal grid resolution of 60 km, rather than the higher resolution mesoscale configuration. The report concluded that the differences between modelled and measured concentrations when using NWP data are of the same order as the differences between modelled and measured concentrations when using standard meteorological data. On this basis, Lucas and Bethan conclude that NWP would be a valid substitute if representative meteorological data were not available for all, or part, of the time period required.

An advantage of NWP data over meteorological measurement data is the absence of gaps in the data series. If gaps in meteorological data are not accounted for, errors can occur in the reporting of percentile values. Data purchased from the Met Office is normally provided with greater than 99% coverage, with typically 2-3% calm conditions, however there is usually no information provided with the data on procedures that have been undertaken for gap filling (Auld *et al*. 2003). The US EPA recommends filling single hour gaps with interpolation techniques to maintain a low level of error. Data substitution from a site with similar surface conditions is recommended for longer gaps, although the error increases with the size of the gap.

3.2.3 Uncertainty in emissions data

Model uncertainty can also be introduced by the input of inaccurate or unrepresentative emissions data. Considering power stations, forecast emissions are often derived from anticipated hourly average power generation levels. However, emissions can also occur when a power station is in stand-by or start-up mode, which may be more difficult to forecast. Emissions during the start-up/stand-by process can be assigned to one of three categories: cold-start, hot-start and stand-by, as described in Table 3.4. Accurate modelling of emissions occurring during start-up and stand-by is particularly important for low load stations, which may be on stand-by for a large proportion of the time. An investigation by Hunter (2006) on the significance of stand-by emissions modelled using ADMS 3.1 reported that for very low load stations, impacts during stand-by

conditions are more significant than impacts which occur during main fuel burning, although these impacts are still small relative to air quality strategy thresholds. In addition, during start-up or stand-by, maximum ground level impacts were found to occur much closer to the station.

Table 3.4 Characteristics of start-up and warming of a coal- or oil-fired power station (from Hunter, 2006).

Start-up/stand-by conditions	Description	Time-frame
Cold start	Prolonged period of oil burning to raise temperature sufficiently	12 hours
	Followed by main fuel ignition and a steady increase to full load	3-4 hours
Hot start	Short period of oil burning to raise temperature sufficiently	2 hours
	Followed by main fuel ignition and a steady increase to full load	3-4 hours
Warming	Slow burning of oil to maintain temperature	Unlimited

The effects of increased emission rates during start-up conditions have also been found to occur in waste incinerators, in particular with respect to the release of polychlorinated dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs) (Environment Agency, 2006; Gass *et al*., 2002; Wang *et al*., 2007). A study by Wang *et al*. (2007) examines the emissions of PCDD/Fs from one continually operating and five intermittent waste incinerators in Taiwan. To test the influence of start-up conditions upon PCDD/F emissions, three consecutive stack flue gas samples were taken when incinerators reached stable combustion conditions after start-up. For all intermittent incinerators, the first sample was found to be two to three times higher than the mean of the remaining two samples and emissions during start-up were found to reach 96.9 ng I-TEQ N m^{-3}.

Dioxin releases in EU countries must meet the standards of the Waste Incineration Directive which has set the emission limit to 0.1 ng I-TEQ N m^{-3} (Environment Agency, 2006). An investigation by the Environment Agency (2006) into the performance of dioxin sampling systems, measured emissions following start-up of 0.057 and 0.068 ng I-TEQ N m^{-3} using a 6-hour Manual Sampling Train (MST). These emissions are much lower in comparison to those recorded by Wang *et al*. (2007). However, this report indicates further work measuring dioxin levels during start-up is required.

In addition to start-up and warming procedures, shut-down procedures have also been shown to increase emissions e.g. Lothgren and van Bravel (2005).

3.2.4 Model uncertainty limits

Data-quality objectives for the required accuracy of modelled concentrations are quoted in AQF Daughter Directives 1 and 3 (1999/30/EC, 2002/3/EC). The accuracy for modelling is defined as the maximum deviation of the measured and calculated concentration levels, over the period considered by the limit value, without taking into account the timing of the events. The uncertainty limit is 30% for annual averages, 50% for daily averages and 50-60% for hourly averages for SO_2, NO_2 and NO_x, whereas for PM_{10} the only defined uncertainty level is 50% for annual averages.

4 Summary: uncertainty in monitoring and modelling

It is evident that air quality impact assessment for the majority of Part A(1) industrial sources in England and Wales relies primarily on ambient air quality modelling. The use of numerical dispersion models such as ADMS and AERMOD provides a cheap and effective means of determining temporal and spatial variation in air quality over a large domain. However, uncertainties in model predictions associated with model choice, user configuration and the use of appropriate meteorological and emissions data means that effective integration of modelling and measurement data and improvements in the accuracy of model input and user configuration are necessary in order to reduce these uncertainties.

Although air quality monitoring provides a relatively accurate indication of the temporal variation in pollutant concentrations at the site of measurement, this technique provides a poor indication of the spatial variation in concentrations. In addition, due to financial and practical restrictions, the number of continuous monitors employed around industrial point sources is limited. Furthermore, short-term monitoring campaigns carried out by the Environment Agency are subject to additional uncertainty due to extrapolation for comparison with annual statistics. The usefulness of national monitoring networks e.g. the AURN network, in providing an additional resource for air quality monitoring around industrial point sources may be limited due to the distance of monitoring sites from industrial sites and interference from the contribution of pollutants from other sources. The use of passive monitors such as diffusion tubes is a cheaper alternative, and thus can be employed in a dense monitoring network. However, while this may provide a better indication of the spatial variability in air quality, the fact that measurements are passively recorded means that temporal variation in air quality is poorly defined. Therefore, this method may only be appropriate where the source is low level and where air quality strategy objectives are long term averages e.g. annual averages.

5 Integration of modelling and monitoring

The integration of modelling and monitoring could exploit the strengths of both techniques and provide better indications of the uncertainties associated with both methods. Integration currently has several applications within air quality impact assessment in England and Wales:

- **Model validation** - The independent testing of dispersion models using monitored data.
- **Data assimilation** – A procedure that combines model and observation data to improve a prediction (model calibration).
- **Data interpolation** - The interpolation of monitoring data to derive spatial maps of ambient concentrations.
- **Inverse modelling -** Ambient concentrations predicted by dispersion models using varying emission rates can be compared to monitored concentrations and thus assist in source strength estimation.
- **Placement of monitors** - Spatial patterns of ambient concentrations predicted by dispersion models can be used to identify the location of peak long-term and short-term concentrations and can therefore be used in the optimum placement of monitors for short and long-term monitoring campaigns. Modelled concentrations may also be used to locate monitors in order to discriminate between the impacts of two or more sources.
- **Data extrapolation** - Exceedences measured during short-term monitoring campaigns may be extrapolated to compare with annual standards.

It should be noted that data assimilation is designed to reduce uncertainty in model outputs caused by systematic errors due to model formulation and is of limited use when uncertainty is due to random errors e.g. the stochastic nature of atmospheric turbulence.

The following sections review existing examples of integration in modelling and monitoring for air quality assessments. The application of integration methods for network design optimisation and for the more representative extrapolation of short-term monitoring campaigns to long-term statistics will also be explored. The use of advanced methods of data assimilation and data driven statistical modelling are also reviewed to determine their suitability for use in air quality impact assessment.

5.1 Model validation

5.1.1 JEP Risk Management Framework

An example of integration for model validation is the Risk Management Framework adopted by the Joint Environment Programme (JEP), which comprises of all major

electricity generators in England, Scotland and Wales[3] (Hunter, 2004). The Risk Management Framework employed by the JEP encompasses an Air Quality Management Plan (AQMP) (Webb, 2005) which must be undertaken by all major coal- and oil-fired power stations to ensure compliance with AQS objectives, as agreed with the Environment Agency. The AQS objectives relevant to emissions from power stations are those for SO_2 (15-minute, hourly and annual means), NO_2 (hourly and annual means) and PM_{10} (daily and annual means).

The Framework uses a combination of dispersion modelling and ambient air quality monitoring in a three-stage process consisting of the following elements:

i. Pre-year projection

ii. Exceedence evaluation system

iii. Post-facto appraisal

5.1.1.1 Pre-year projection

Dispersion modelling is carried out for a pre-year projection using the anticipated generation scenario (variation in load factor typically at 4-hour resolution), the anticipated fuel sulphur content and five years of representative meteorological data, to investigate the potential impacts on air quality. A mean background correction is derived using Equation 3 and applied to all modelled annual mean and most short-term percentile values to account for the contribution from background sources that are not included in the model. In some locations, a higher background concentration is used for the percentile values.

The mean background concentration, $\overline{C_b}$, is determined from the monitored annual mean, $\overline{C_{Mon}}$ minus the modelled annual mean, $\overline{C_{Mod}}$ (JEP, 2002):

$$\overline{C_b} = \overline{C_{Mon}} - \overline{C_{Mod}} \qquad \textbf{Equation 3}$$

This background correction method is used only to calculate the contribution from background SO_2. The formation of NO_2 from NO_x emissions is limited by the presence of O_3, therefore, accounting for the prevailing background NO_x concentration usually makes no difference to NO_2 concentrations associated with the grounding of a plume (JEP, 2002). Estimates of background concentrations differ for each power station (or group of stations) and are reviewed annually on the basis of both recent and long-term trends in monitoring data. In reality, background concentrations may fluctuate over time, however the application of a constant correction for background is a reasonable approach for power stations as their buoyant plumes only mix to ground level where there is vigorous mixing and therefore generally low background concentrations.

Following the pre-year modelling projection, the risk of non-compliance is categorised as low, medium or high and adjustments may be made to the planned generation/fuel scenario to reduce the risk if necessary.

5.1.1.2 Exceedence evaluation system

In the second stage of the Risk Management Framework, an exceedence evaluation system is operated whereby measured exceedences are compared to modelled

[3] AEP, Drax Power Ltd, Eggborough Power Ltd, EDF Energy, International Power, E-on, RWE npower and Scottish Power

exceedences for an ongoing assessment of air quality impacts throughout the year. Given that seasonal variation in exceedences may occur, quarterly exceedence allowances are derived for each monitoring site by conducting modelling on a quarterly basis. Monitoring results are then evaluated by comparing the cumulative number of measured exceedences with the sum of exceedence allowances derived on a quarterly basis. As well as using modelled exceedences, the evaluation system uses percentage thresholds of AQS objectives to give broad indications of the risk of compromising an objective. For example, a threshold of 85% (equivalent to 30 exceedences of the 15-minute mean SO_2 concentration objective) is considered the threshold at which action is required to *ensure* future operations will not compromise AQS objectives. The status conditions are based on a 'red-amber-green' approach and the appropriate course of action taken when a particular threshold is crossed (either modelled exceedence allowance or AQS threshold) depends at which point in the year the threshold is reached. Table 5.1 details the thresholds and associated status conditions for the assessment of measured exceedences during each quarter. A detailed description of the exceedence evaluation process may be found in Hunter and Bethan (2005).

5.1.1.3 Post facto appraisal

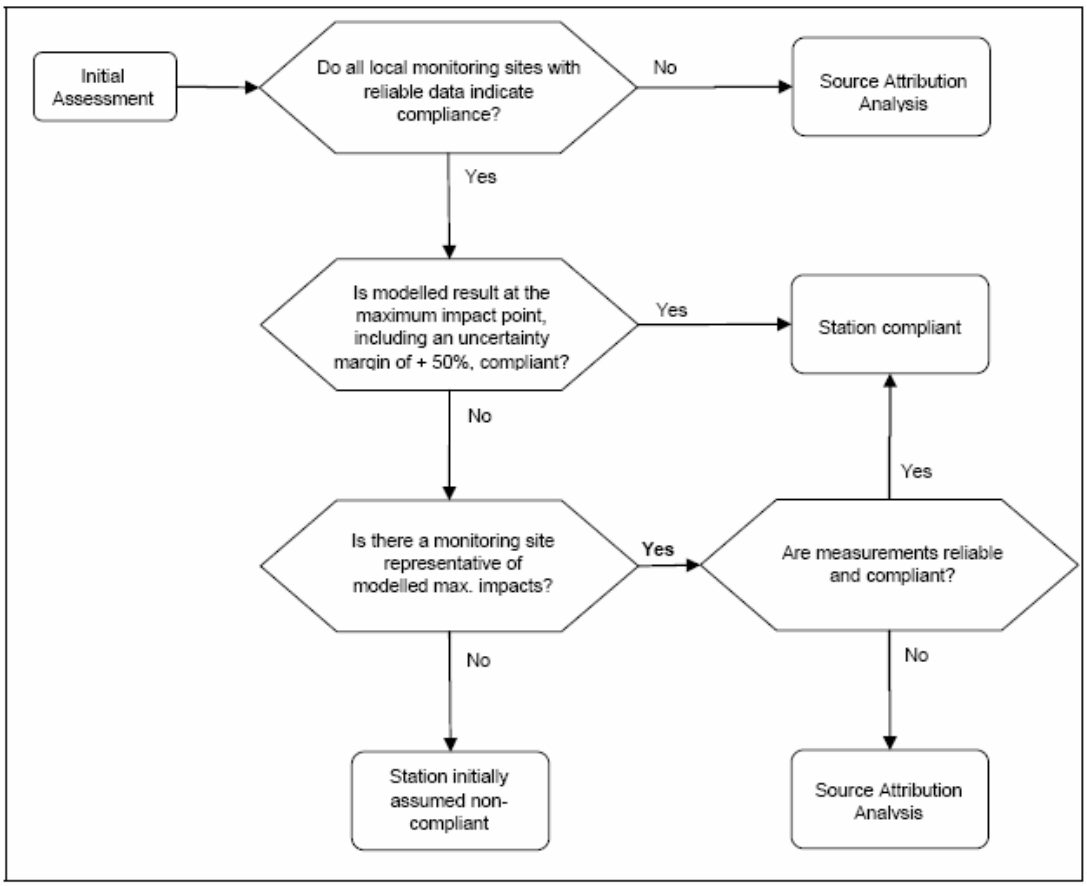
In the third stage of the Risk Management Framework, a post-facto appraisal is undertaken to compare monitored concentrations with modelled or predicted concentrations as a form of validation for the pre-year modelling forecast. Results are compiled in an Annual Review document for each power station. Model performance is considered "reasonable" if predicted values lie within a factor of two of measured values, corresponding to a modelled concentration falling between 50% and 200% of the measured concentration.

The primary method of assessing compliance is through the use of monitored data, although modelled data may also be used in the decision-making process, which is presented as a flow chart in Figure 5.1. However, if the monitoring site falls outside a defined area surrounding the pre-year modelled maximum concentration (Figure 5.2), then the monitoring data are considered 'unrepresentative'. In this case modelled concentrations are used as the primary method of assessing compliance.

An uncertainty margin of ± 50% is applied to all modelled concentrations used to assess compliance. For example, for the 15-minute mean metric of 266 μg m^{-3}, concentrations above 133 μg m^{-3} could be considered to be "potential exceedences" whilst concentrations above the upper limit of the uncertainty (400 μg m^{-3}) are considered as "highly likely" exceedences. Modelled concentrations are only considered exceedences if the upper limit of uncertainty is breached. This approach is necessary because there must be a high degree of confidence before the Environment Agency can consider further action on the basis of exceedences determined solely from modelling data.

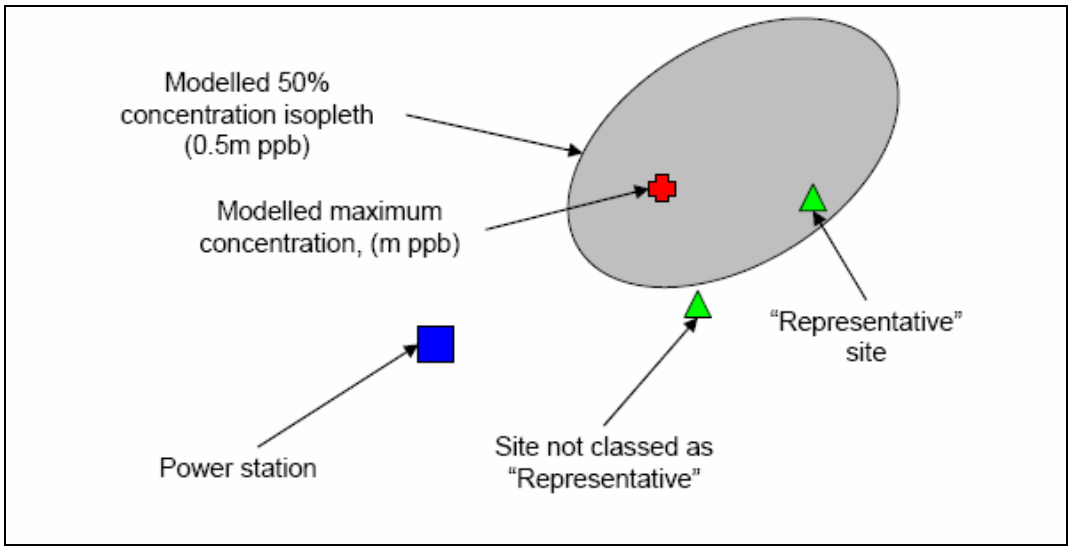
Table 5.1 Thresholds and associated status conditions for the assessment of measured exceedences during each quarter (from Hunter and Bethan, 2005).

Quarter	Criteria	Status	Action required
1	Less than 60% Q1	Green	None
	Greater than 60% Q1 but less than Q1 + Max (Q2/2, 15%*AQS) And less than 85% AQS	Green/Amber	Establish cause; review whether there are any compliance implications for remainder of year.
	Greater than Q1 + Max (Q2/2, 15% AQS) And less than 85% AQS	Amber/Red	Establish cause; review future operations and compliance implications. Need high degree of confidence that future operations will not jeopardise AQS objective; otherwise operational modifications required.
	Greater than 85% AQS	Red	Action required unless future operations will clearly not jeopardise AQS objective.
2	Less than 60% (Q1 + Q2)	Green	None
	Greater than 60% (Q1 + Q2) but less than (Q1 + Q2) + Max (Q3/2, 15% AQS) And less than 85% AQS	Green/Amber	Establish cause; review whether there are any compliance implications for remainder of year.
	Greater than (Q1 + Q2) + Max (Q3/2, 15% AQS) And less than 85% AQS	Amber/Red	Establish cause; review future operations and compliance implications. Need high degree of confidence that future operations will not jeopardise AQS objective; otherwise operational modifications required.
	Greater than 85% AQS	Red	Action required unless future operations will clearly not jeopardise AQS objective.
3	Less than 60% (Q1 + Q2 + Q3)	Green	None
	Greater than 60% (Q1 + Q2 + Q3) but less than 85% AQS	Amber	Establish cause; review whether there are any compliance implications for remainder of year.
	Greater than 85% AQS	Red	Action required unless future operations will clearly not jeopardise AQS objective.
4	Less than 60% AQS (20 exceedences)	Green	None
	Greater than 60% AQS but less than 85% AQS (30 exceedences)	Green/Amber	Establish cause; review future operations and compliance implications. Consider need for operational modifications.
	Greater than 85% AQS but less than AQS	Amber/Red	Establish cause; review planned operations. As headroom is very low need very high degree of confidence that operations will not jeopardise objective; otherwise operational modifications required.
	Greater than AQS	Red	Ensure station(s) can cause no further exceedences.



Source: Webb (2004)

Figure 5.1 Schematic of the compliance assessment procedure for UK coal- and oil-fired power stations.



Source: Webb (2004)

Figure 5.2 Schematic illustrating the JEP definition of a 'representative' monitoring site.

5.2 Data assimilation

Data assimilation may be defined as the combination of observational data with modelled data to provide a better estimation of the quantity of interest.

Data assimilation in the context of air quality assessment has several purposes:

- To improve the estimation of air quality and the associated uncertainties for compliance assessment.
- To provide improved estimates of source attribution and estimates of source strength.
- To optimise the design of monitoring networks in order to achieve maximum benefit from integration methods.

Data assimilation methods may be described as simple or complex. Simple methods basically involve a global adjustment of the modelled concentration field, with no spatial or temporal dimension. Complex methods may incorporate either a spatially or temporally varying adjustment of the modelled concentration field, or for the most complex methods, both adjustments combined.

Complex data assimilation methods for have been used frequently in the meteorological community over recent decades for Numerical Weather Prediction (NWP) e.g. Ghil *et al.* (1979), Rabier (2005) and Swinbank and O'Neill (1994). However, the use of these methods for the calibration of air quality assessments has been limited due to the complexity of the data assimilation methods involved. In addition, the effectiveness of these methods depends on an adequate spatial and temporal coverage of monitoring data.

The following sections describe several data assimilation methods currently in use within the air quality community and also methods that may be of particular value to air quality assessments but which are not commonly used at present. The assimilation methods include simple linear regression and ratio calibration of modelled concentration fields, simple and complex geostatistical methods including inverse distance interpolation, kriging and simulation, and finally the use of Bayesian assimilation methods will be discussed.

5.2.1 LAQM calibration methodology

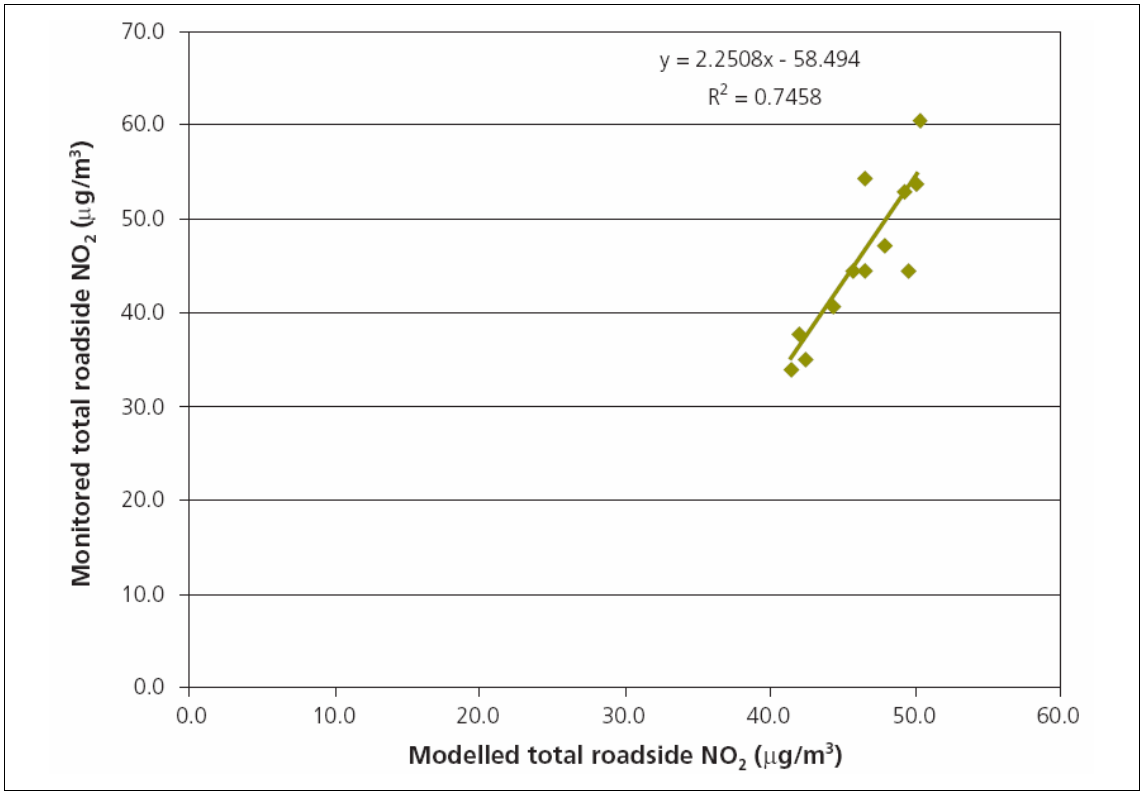
An example of a simple data assimilation (calibration) method is that recommended for use in Local Air Quality Management (LAQM) prescribed by LAQM Technical Guidance TG(03) (Defra, 2003). LAQM modelling assessments include all relevant point, area, grid and road sources that emit a pollutant of interest. Since results of the modelling studies are used to assist Local Authorities in making regulatory decisions, such as designating Air Quality Management Areas (AQMAs) within their district, the model needs to be calibrated and subject to uncertainty analysis. In order to achieve this, linear regression analysis is often used to determine the significance and strength of the relationship between predicted and measured concentrations. Linear regression is expressed using Equation 4, where M is the model field, M_{LR} is the resulting linear regression field and a and b are coefficients representing the y-intercept and the slope of the fitted line respectively.

$$M_{LR} = a + bM \qquad \text{Equation 4}$$

Linear regression is used to provide appropriate correction factors (slope and intercept) by which modelled concentrations are adjusted. The intercept parameter (a) also informs the modeller about the relevance of background concentrations. Figure 5.3 illustrates the application of the linear regression method for validation of urban air quality models from Defra Technical Guidance (Defra, 2003).

Analysis of the Root Mean Square Error (RMSE) allows the magnitude of error associated with the corrected model predictions to be derived. This is equivalent to one standard deviation from the mean and thus expresses uncertainty at the 68% confidence level. RMSE is calculated by taking the square root of the mean squared difference between monitored, O_i, and the adjusted modelled, M_i, concentrations (Equation 5).

$$RMSE = \sqrt{\frac{\sum (O_i - M_i)^2}{n}} \qquad \textbf{Equation 5}$$



Source: Defra (2003)

Figure 5.3 Relationship between ambient NO_2 predicted using an urban dispersion model and monitored ambient NO_2.

Linear regression is an example of a simple data assimilation method whereby air pollutant measurements are used to make controlled adjustments to model-predicted concentrations. However, linear regression models are global in nature and are therefore not appropriate when variations in modelled and measured concentrations are a result of local effects in wind, dispersion or emissions near monitoring sites (Denby *et al*., 2007). In addition, poor spatial representation of measured concentrations may mean that linear regression is not always the most appropriate

method to use. Ideally when applying this method, a large number of observations covering a range of concentrations (including regional background measurements) should be included.

5.2.2 FRAME calibration methodology

An example of a more complex data assimilation method is the standard calibration procedure routinely applied to output data from the FRAME (Fine Resolution Atmospheric Multi-pollutant Exchange) model (Dore *et al*., 2005). The FRAME model is used to forecast long-term annual mean deposition of nitrogen and sulphur at a 5km resolution over the United Kingdom, whereas mapped UK deposition estimates of nitrogen and sulphur for individual years are obtained from concentration-based measurements of wet deposition and gas concentrations (CBED). To compare CBED-derived concentrations for the present year directly with model-derived concentrations for future years may provide misleading results due to the different methods used. Therefore, a calibration procedure is applied to the modelled FRAME projections ($DEP_{(FRAMEproj)}$) to normalise the modelled projection by a factor determined by the relationship between the present CBED derived estimates ($DEP_{(CBEDpres)}$) and the modelled deposition estimated using FRAME for the same time period ($DEP_{(FRAMEpres)}$). This procedure is then used to calibrate the projected FRAME deposition estimate ($CALDEP_{(FRAMEproj)}$) for each 5 x 5 km grid square in the UK using Equation 6.

$$CALDEP_{(FRAME,proj)} = DEP_{(FRAMEproj)} * (DEP_{(CBEDpres)}/(DEP_{(FRAMEpres)}) \qquad \textbf{Equation 6}$$

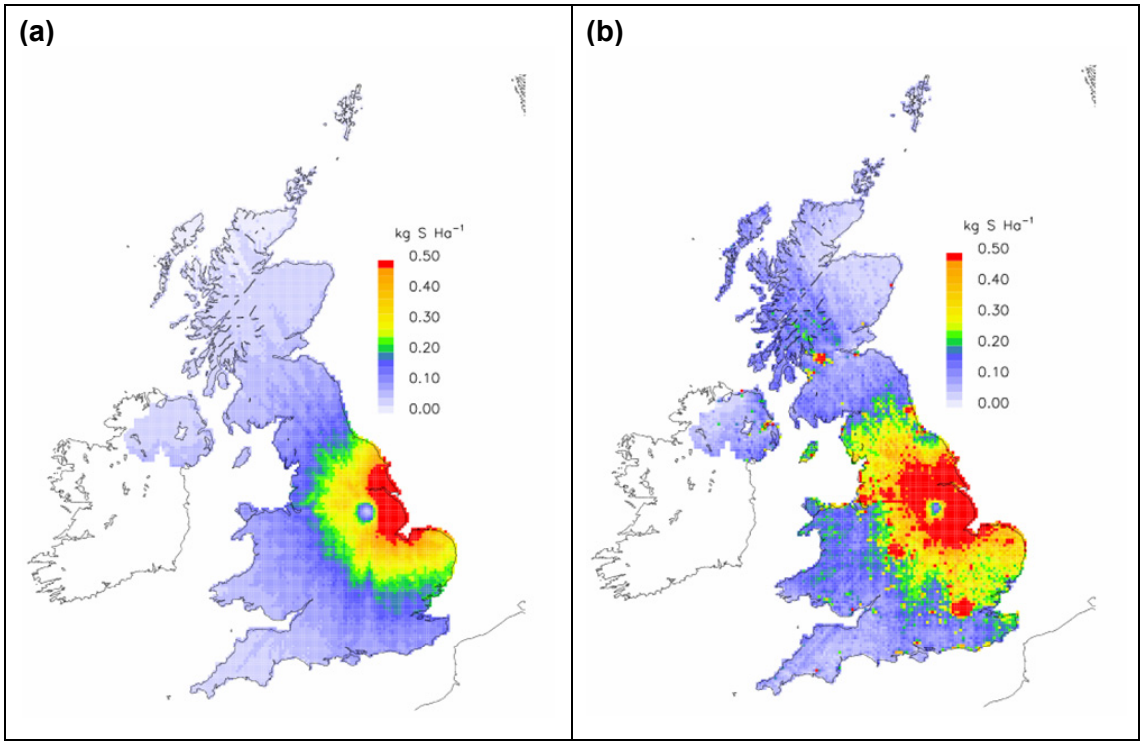
The calibration procedure is conducted for each 5 x 5 km grid square, thus producing a spatially varying model calibration, unlike the LAQM example which enables simply a global adjustment to the modelled concentration field. When applied to total deposition patterns from all sources in the UK, this method may be an appropriate way of calibrating measured and modelled data. However, for future projection, this method assumes that that ratio between measured and modelled estimates is constant year-upon-year.

Dore *et al*. (2005) applied this methodology to assess the magnitude and spatial distribution of individual deposition footprints of sulphur and nitrogen, associated with gaseous emissions from point sources in the UK. The application of this procedure was found to introduce unrealistic anomalies in the spatial distribution of deposition. Figure 5.4 shows a comparison of the calibrated and uncalibrated deposition estimates from a single emission source in 2005. Notable differences include evidence of anomalous areas of peak deposition in urban areas located far from the source and the increased spatial extent of the highest deposition contour. It is possible that these anomalies have arisen because CBED data derives deposition from all emission sources rather than from individual sources.

It is important to note that the FRAME calibration method is designed for application to relatively low resolution, multiple-source models, where widespread monitoring data are available to generate an interpolated concentration field from measurements. The FRAME calibration method is model-specific and therefore unsuitable for the calibration of high resolution, local area models e.g. ADMS or AERMOD. However the general principle of the ratio method for application to model calibration is explored in Section 7 of this report.

It is evident from the FRAME calibration example that it is necessary to account for the contribution of individual sources to measurement data prior to conducting model calibration. One way of addressing this problem would be to apply a correction to the modelled field to account for the influence of background sources. Another would be to

correct the measured data by isolating the contribution of individual sources. This method is discussed in Section 7 of this report. This problem does not occur in the LAQM example as all major sources of pollution are accounted for.



Source: Dore *et al*. (2005)

Figure 5.4 Uncalibrated (a) and calibrated (b) estimates of SO_2 deposition (keq Ha-1 yr-1) in 2005 arising from a single source using the FRAME model.

5.2.3 Geostatistics

Geostatistics is a powerful data assimilation tool, which may be used independently, to generate spatial information that would normally be modelled by an atmospheric dispersion model. The following are examples of where geostatistical techniques have been used in place of mathematical models, to devise an interpolated concentration field. This offers an alternative to deterministic model calibration using the integration of measurement data.

5.2.3.1 The Inverse Distance Weighting method

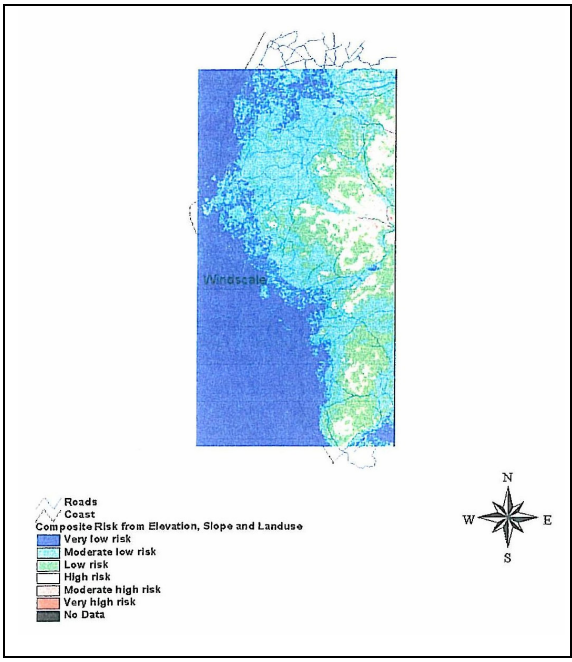
The Inverse Distance Weighting (IDW) method was used by Argyraki *et al*. (1999) to interpolate the results of an aerial radiometric survey carried out in West Cumbria in 1988. These survey data were used to validate a risk classification system, developed to optimise sampling following a release of radioactive material.

Argyraki *et al*. (1999) devised a system whereby a simple Gaussian plume model of airborne contamination could be integrated with information on meteorology, land cover and topography, to provide an objective and auditable means for assessing sampling strategies following the release of radioactive material. This method involved defining the area under consideration on the basis of deposition risk according to meteorological parameters, elevation, slope and land use. The composite map in Figure 5.5 shows the deposition risk categories based on elevation, slope and land use.

The calculation of the overall risk factor is achieved by multiplying modelled deposition by the deposition risk value for a particular area. The results of which are superimposed onto the plume footprint (Figure 5.6).

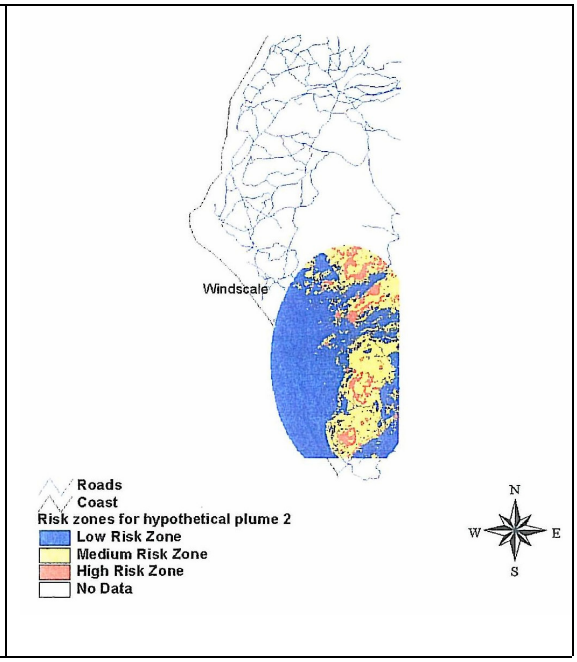
To verify the risk classification system, measured deposition data were interpolated using the IDW method. This method interpolates data at nodes on a pre-determined grid by applying a power weighting function to each measurement data point. The weight given to a particular data point when calculating a grid node is proportional to the inverse of the distance of the observation from the grid node. When calculating a grid node, the assigned weights are fractions, and the sum of all the weights is equal to 1. When an observation is coincident with a grid node, the observation is given a weight of 1, and all other observations are given a weight of 0.

Figure 5.7 illustrates the aerial radiometric survey deposition measurements and Figure 5.8 the interpolated spatial distribution of deposition. The interpolated surface enables the identification of areas where the highest concentrations occur for comparison with the deposition risk map (Figure 5.6). Areas of high, medium and low risk, defined by the risk classification system shown in Figure 5.6, can be seen to correspond well to the interpolated deposition surface in Figure 5.8.



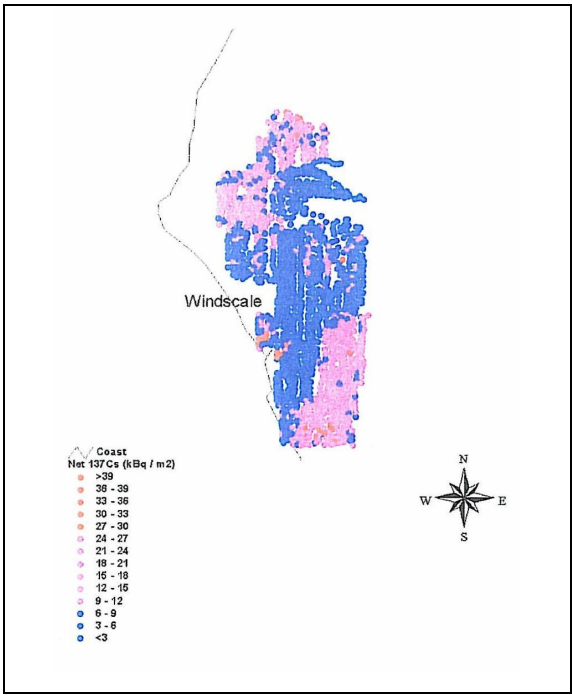
Source: Argyraki *et al*. (1999)

Figure 5.5 Composite map showing risk categories based on elevation, slope and land use.



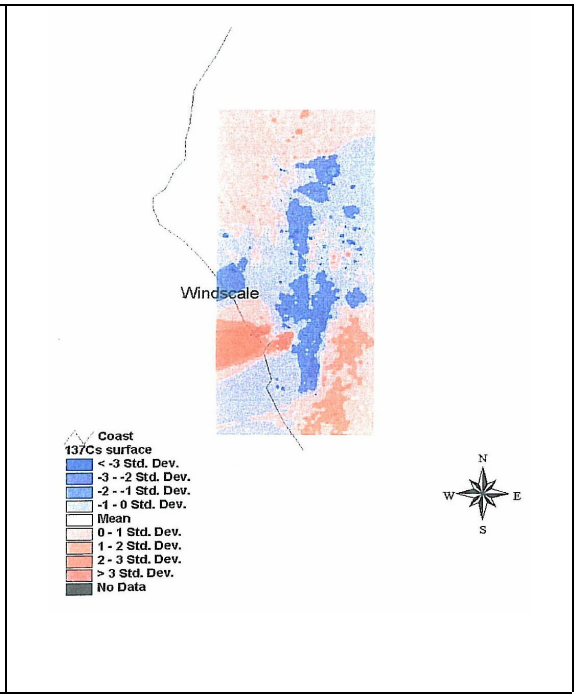
Source: Argyraki *et al*. (1999)

Figure 5.6 Combined plume deposition and risk map following the Windscale Fire in 1957.



Source: Argyraki *et al*. (1999)

Figure 5.7 Spatial distribution of ^{137}Cs deposits from the Windscale Fire in West Cumbria. Map based on data collected from an aerial radiometric survey in 1988.

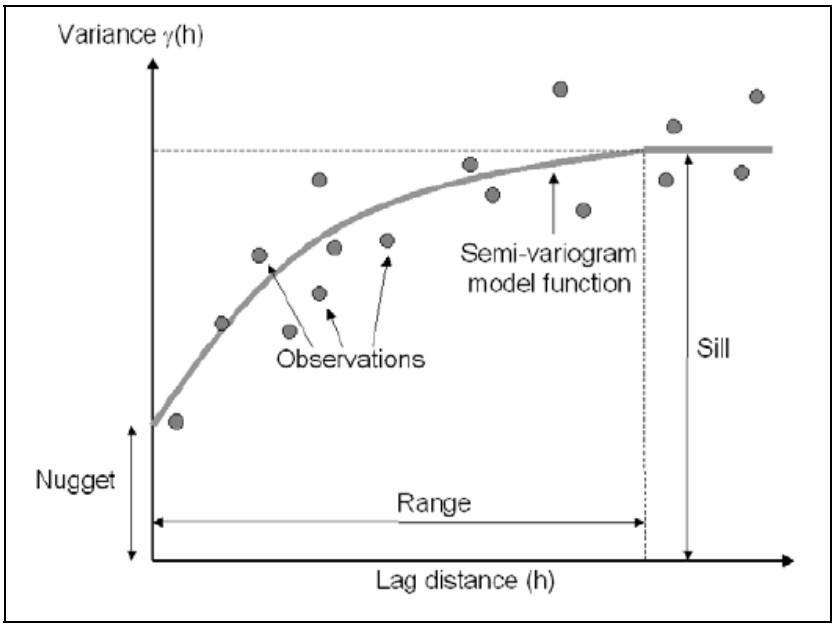


Source: Argyraki *et al*. (1999)

Figure 5.8 Interpolated surface for ^{137}Cs deposits with the classification based on the standard deviation of the element's concentration frequency distribution.

5.2.3.2 Kriging of the observed field

Kriging is a more sophisticated spatial interpolation method than the IDW method although both revolve around the assumption that there is a spatial correlation between points of known value. The spatial correlation is related to the distance between the points, i.e. the closer the points, the higher the correlation. In kriging the relationship is described by a spatial variance function called the semi-variogram (Figure 5.9). Attributes of the semi-variogram include the range and the nugget:sill ratio. The range indicates the distance at which the maximum variance is reached, i.e. the separation between observations at which no spatial correlation is found. The nugget:sill ratio is the ratio of the separation between observations at which variance is no longer constant (nugget) and the upper limit of variance (sill). The semi-variogram is then used to interpolate any point in space by weighting nearby measurement points so that the variance at the interpolation point is minimised.



Source: Denby *et al*. (2007)

Figure 5.9 Parameters and terms used to define the kriging semi-variogram function and how the semi-variogram function may be fitted to the observed variance.

Several types of kriging may be used, including the following three main types;

- Simple kriging
- Ordinary kriging
- Kriging-with-a-trend or Universal kriging

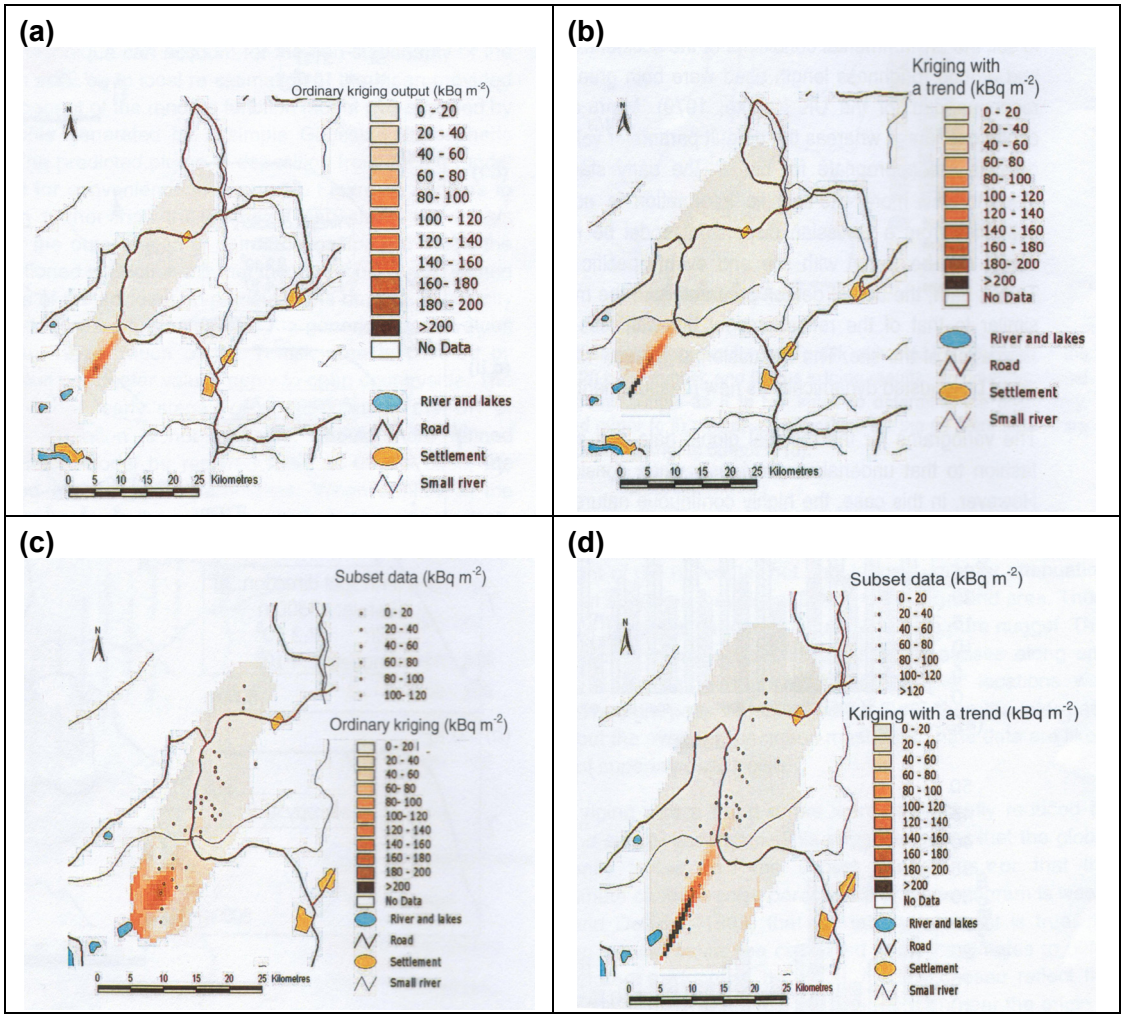
Simple kriging is the most basic form of kriging, which assumes that the measured values are realisations of a trend component with a constant known mean for the entire interpolated surface. In contrast, ordinary kriging uses a local average for each particular interpolation point based on a local 'search area' or search radius. As a result, ordinary kriging can be more accurate than simple kriging and therefore ordinary kriging is often used in preference. In both simple and ordinary kriging, the mean value

of the variable over the interpolated surface or search area is assumed constant. However, in some circumstances, a trend may be observed in the data such that the mean varies over the search area and is therefore no longer stationary. Kriging-with-a-trend is a variant of ordinary kriging that can incorporate the effect of a trend on the local mean. For a detailed description of these kriging methods and their derivations, e.g. block kriging and co-kriging, see Webster and Oliver (2001).

In terms of interpolating air quality monitoring data, ordinary kriging is more applicable to occasions where monitoring data are spread evenly over a whole area, e.g. urban scale air quality monitoring networks. Kriging-with-a-trend is more applicable to near-field contamination, i.e. the deposition pattern formed by an idealised Gaussian Plume. In this case, sampling is often more dense in the area of highest deposition closest to the release and decreases with distance from the area of plume impact.

Higgins *et al*. (2005) investigated the potential of kriging to provide guidance in establishing the likely extent of food restriction zones following the accidental release of radioactive substances to air. Kriging methods were applied to data from three nuclear accidents (Tomsk, Chernobyl and Windscale) to determine the usefulness of these techniques for a broad range of release scenarios.

Ordinary kriging and kriging-with-a-trend interpolation techniques were applied to aerial gamma survey data points obtained 5 months after the Tomsk accident in April 1993. Figure 5.10 compares the ordinary kriging and kriging-with-a-trend interpolations applied to the global data set (812 data points) with ordinary kriging and kriging-with-a-trend interpolations from a subset of 40 data points. Ordinary kriging was used for the initial interpolation (Figure 5.10 a.) because the density of data points was sufficient to enable adequate interpolation. Figures 5.10 c. and d. show ordinary kriging and kriging-with-a-trend applied to the random subset of 40 data points, each showing distinctly different interpolations of the deposition field. Figure 5.10 c. shows that although the general pattern of deposition is replicated, interpolation at the edges of the plume, especially nearer the source, is very noisy. This is because when estimating locations at the edge of the plume, the search radius is extended to include enough sample points to generate an estimate, which results in the kriging algorithm using a mean more like the local mean of the interior than of the edge. The application of kriging-with-a-trend to the subset (Figure 5.10 d.) gives a better estimate than ordinary kriging, with the surface appearing very similar to that derived when applying ordinary kriging to the global data set.



Source: Higgins *et al*. (2005)

Figure 5.10 A comparison of the Tomsk aerial gamma survey global data set interpolated using ordinary kriging (a), kriging-with-a-trend (b), with the interpolation of the deposition field from 40 data points using ordinary kriging (c), and kriging-with-a-trend (d).

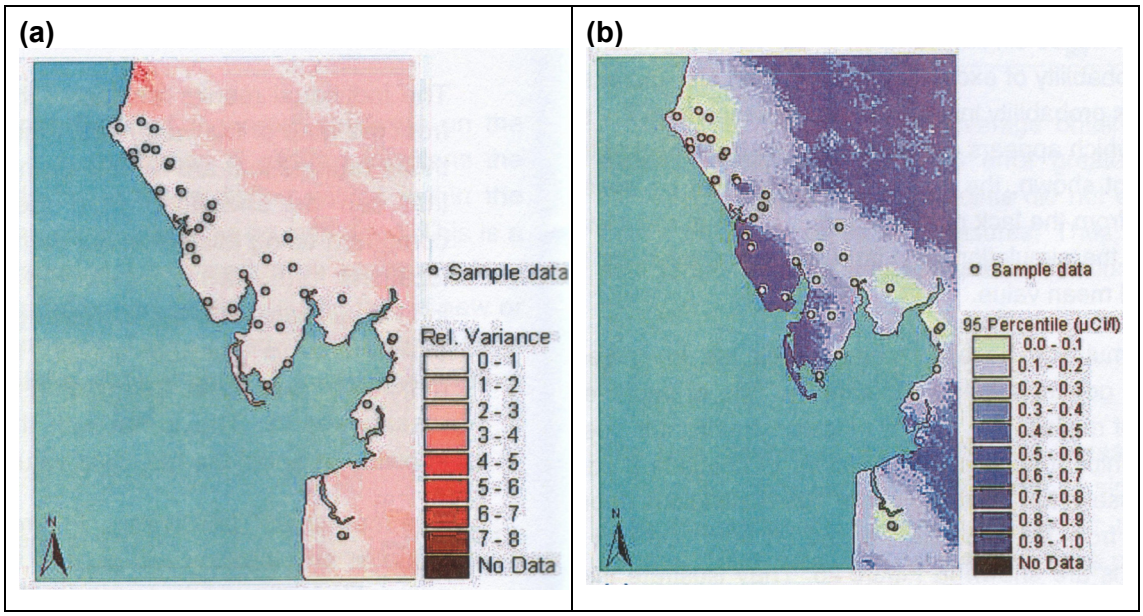
5.2.3.3 Simulation of the observed field

Simulation differs from kriging in that it attempts to model the variability of a surface rather than the best estimate value and hence can often produce a more realistic interpolated surface than that produced using kriging. Realisations are generated using a random number-based algorithm that obtains information about the spatial structure of the semi-variogram.

During simulation, large numbers of equally probable realisations are generated, therefore simulation may be used to calculate the probability distribution of a value occurring. Simulation can be used to indicate areas with the highest variance and hence where there is least knowledge about the true data values; the values of the different realisations will be more widely distributed the further they are from measurements. This can help identify areas where additional information could be most effectively gained by making further measurements.

A sequential simulation is one in which the conditioning data for any location includes not only the measured values but also previously simulated values within a specified region. Figure 5.11 illustrates how sequential Gaussian simulation can be used to generate probability contour maps. Such maps are useful to determine where further measurements should be made to reduce uncertainty in the interpretation of such data. Further information on the sequential Gaussian simulation approach is detailed in Deutsch and Journel (1998).

The simulations from Higgins *et al.* (2005) produce a clear visual representation of increased activity concentration in milk in the south-westerly direction of plume travel. Figure 5.11a. shows that variance is highest in the area remote from where measurements were taken. In these locations, sample data have less influence over the simulated concentration field thus resulting in higher variance. Figure 5.11 b. gives an indication of whether the measured value at any location will exceed the milk bank criterion. The green areas contain several sample locations, each of which is well below the milk ban criterion so it is possible to be confident that the measured value at a nearby location will be similarly low. However, as the value represented by the 95th percentile increases, the probability of exceeding the milk ban limit also increases. It may be noted from these maps that the probability increases in a north-easterly direction. Higgins *et al.* (2005) presume this to be an artefact stemming from the lack of measurement data at this distance. Since there are no real data available, the simulation algorithm returns a value close to the global mean for these locations.



Source: Higgins *et al*. (2005)

Figure 5.11 Maps generated from 100 realisations, showing (a) the relative variance and (b) the 95th percentile of ^{131}I activity concentration in milk (µ Ci l^{-1}) with reference to the milk ban criterion of 0.1 µ Ci l^{-1} (3700 Bq l^{-1}).

5.2.3.4 Geostatistical methods integrating model and monitoring data

At their simplest, geo-statistical techniques use only measured values, therefore the techniques can be applied in a very straightforward manner to interpolate a concentration field from measured data (as is described in the previous three sections). However, geostatistics may also be used to aid model calibration. The interpolation of a measured concentration field using kriging techniques, as conducted by Dore et al. (2005) using the FRAME model, can enable a modelled concentration field to be calibrated in a spatially varying manner. This is in contrast to simple data assimilation methods, which involve the global adjustment of a modelled concentration field, with no spatial or temporal dimension.

Denby *et al*. (2007) used ordinary kriging to calibrate annual mean modelled concentrations of SO_2, NO_2 and PM_{10} in the city of Prague. The ATEM model (a statistical Gaussian model) was used to generate the modelled data and provided annual mean pollutant concentrations at a resolution of 250 m. Twelve monitoring stations provided annual mean measured concentrations for the year 2003. This study was conducted as part of the Air4EU project, which was set up in 2004 to provide recommendations on air quality assessment by monitoring and modelling for regulated pollutants in Europe (http://www.air4eu.nl/index.html).

Denby et al. (2007) performed three kriging methods in order to interpolate measured annual mean concentration fields for SO_2, NO_2 and PM_{10}:

- Kriging of the observed field
- Kriging of the residual field (observed - modelled)
- Kriging of the normalised residual field (observed / modelled)

The cross-validation procedure, explained in Section 5.2.5 of this report, was used to assess the effectiveness of each method.

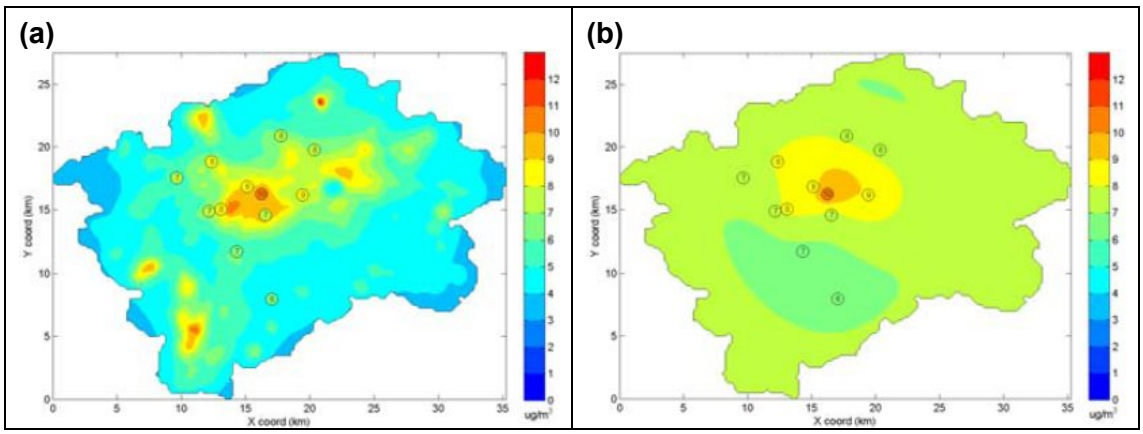
For kriging of the observed field to be effective, it is necessary to have enough observations so that the density of measurement points is equal to or greater than the spatial variation of the measurements. Figure 5.12 shows that the density of monitoring stations applied by Denby *et al*. (2007) was insufficient to generate a kriged concentration field purely from the observations as most of the detail shown in the modelled concentration field is lost in the kriged concentration field. In terms of the semi-variogram, this means that the empirically-derived variance does not give a clear definition of the actual spatial variance, therefore other methods must be used to determine the parameters of the semi-variogram.

In contrast to kriging of the observed field, kriging of the residual allows the modelled concentration field to retain its detailed spatial variation. The modelled concentration field is calibrated by simply adding (or subtracting) a kriged residual field. Denby *et al*. (2007) also apply the kriging technique to the normalised residual or the ratio of observed to modelled concentrations. Rather than adding (or subtracting) the kriged field, the modelled concentration field is multiplied by a kriged ratio field to scale the modelled field in a spatially varying manner. Kriging of the residuals has been shown by Blond *et al*. (2003) to be an effective method for improving mapped concentration fields on regional scales.

Table 5.2. compares the kriging methods used in the Denby et al. (2007) study using the cross-validation technique discussed in Section 5.2.5 of this report. The cross-validation RMSE has been calculated for the uncalibrated model using the mean of the observations. Improvements in model vs. measured concentrations are expressed by RMSE values lower than those achieved for the uncalibrated model. The data in Table

5.2 show that neither kriging of the observations nor kriging of the residuals (normalised and non-normalised) result in an improvement in the RMSE above that achieved using a simple linear regression method with varying slope and intercept. This is because the kriging semi-variogram is less well defined when kriging the residuals, as the spatial correlation between the observed and modelled data is often removed by subtracting the modelled from the observed concentrations. This has the effect of defining the semi-variogram using a constant variance (or nugget), which simply interpolates the regional mean, producing a result similar to that obtained by using a global concentration field calibration such as the linear regression method (Section 5.2.1). The extent to which the spatial correlation is removed through kriging of the residual is dependent on the size of the residual, the scale over which the residual variations occur, e.g. sub-grid or super-grid, and whether the residual is a consequence of real effects or artefacts introduced due to model inadequacy.

The ability of a modeller to interpret variogram characteristics is an important factor in determining the appropriate application of the above methods. Therefore, kriging would be difficult to apply in routine regulatory situations unless appropriate systems and methodologies were in place. This emphasises the need for substantial skill and expertise in the implementation and application of all complex data assimilation methods in connection with the integration of air quality models and observations.



Source: Denby *et al*. (2007)

Figure 5.12 Annual mean SO_2 concentrations for 2003 in the city of Prague, (a) modelled and (b) produced by kriging of observations.

Table 5.2 Cross-validation RMSE (μg m^{-3}) for a range of data assimilation techniques applied to annual mean SO_2, NO_2 and PM_{10} concentrations for 2003 from the city of Prague (Denby *et al*. 2007).

Data assimilation method	Annual mean SO_2	Annual mean NO_2	Annual mean PM_{10}
Uncalibrated model[1]	**1.13**	**7.80**	**5.90**
Linear regression[2]	0.95	6.19	5.81
Linear regression[3]	1.28	9.32	4.45
Linear regression[4]	1.25	6.24	4.51
Kriging of observations[5]	1.02	6.33	6.02
Kriging of residuals	1.19	6.84	4.49
Kriging of normalised residuals	1.29	8.81	4.79
Bayesian assimilation of modelled field	1.23	7.75	5.85

[1]Cross-validation using mean of observations

[2]Linear regression model with varying slope and intercept such that $M_{LR1} = a_1 + b_1 M$

[3]Linear regression model with intercept or background = zero (0_{BG}) and measured background concentrations subtracted from modelled concentrations such that $M_{LR2} = b_2 (M - M_{BG}) + 0_{BG}$

[4]Linear regression with slope = 1 but with varying intercept such that $M_{LR3} = a_3 + M$

[5]Nugget:sill ratio adjusted to represent observational uncertainty

5.2.4 Bayesian assimilation

There is a tendency with data assimilation methods, to treat observations as being error free. In reality there are uncertainties attached to observations, especially if observations are used to calibrate grid-based models where observations are assumed to be representative of the value averaged over a grid cell. Data assimilation methods based on Bayesian statistics can take into account these uncertainties when integrating modelled and measured data.

Denby *et al*. (2007) compared a Bayesian assimilation method with kriging and linear regression methods used on the 2003 Prague data set described previously. The observation uncertainty was determined by performing kriging of the observations and calculating the variance at every monitoring point in the kriged observation field. The model uncertainty was defined by calculating the RMSE of the model field using the observations. The assimilated field was subsequently generated by combining the observation field with the modelled field to produce a concentration field weighted towards modelled or measured concentrations depending on the relative uncertainties in both fields. In areas far from observations, the assimilated value is closer to the modelled value and in areas closer to observations, the assimilated value approaches the observational value but still contains an uncertainty component. A comparison of this method with the linear regression and kriging methods explored by Denby *et al*. (2007) can be found in Table 5.2. In this example, Bayesian assimilation only marginally improves the cross-validation RMSE for annual mean NO_2 and PM_{10} when compared to the uncalibrated model concentration field and provides no improvement upon the best performing non-Bayesian methods.

The example of Bayesian data assimilation discussed in Denby *et al*. (2007) may be described as ‘offline’ as this method does not require automated interaction of modelled and monitoring data. However, there are more complex data assimilation methods, which do require direct interaction with a model. Examples include 4D-Var, Kalman filter methods and newer methodologies such as Sequential Monte Carlo (SMC) methods, otherwise known as particle filter (PF) methods. The advantage of using more complex ‘online’ methods of data assimilation is that they can produce both spatial and temporal assimilation of data, whereas ‘offline’ methods can only be applied to concentration fields that do not have a temporal dimension. Walker *et al*. (2006) provides a description of these data assimilation methods, for which a summary is provided in Table 5.3.

Online Bayesian assimilation allows the data assimilation procedure to advance by analysis cycles whereby observations of the current state of a system are combined with a model forecast to produce the best estimate of the current state of the system. Described statistically, the prior expectations of critical model parameters are updated through the influence of measurements. This assimilation of data reduces uncertainty as predictions become increasingly based on measurements rather than on model results. In this way, Bayesian data assimilation may be described as a form of back-fitting model parameters. Offline methods are simply applied on a one-off basis and therefore require no analysis cycles.

To date, data assimilation methods applied to industrial point sources have been limited to emergency assessment methods for use after accidental radioactive release to the environment (e.g. Higgins *et al*., 2005; Politis and Robertson, 2003; Zheng *et al*. 2007). In the context of data assimilation after an accident, Bayesian analysis can provide a method of evolving the predictions of models to take account of measurement data as it arrives.

As part of the SECTAR project[4], Higgins *et al*. (2005) investigated the potential application of Bayesian assimilation techniques using data from two nuclear accidents (Tomsk and Windscale). In this example, the Bayesian approach returned a measure of how well the predictions of the process model compared with the observed data, a term known as 'model inadequacy'. Model parameter uncertainties were adjusted before any observational data were incorporated so that model inadequacy could be examined only when the best parameter values had been determined. Model inadequacy was then examined, as with most calibration techniques, by comparing measurement data with model predictions. After determination of the initial probability distributions for the calibration parameters and the model inadequacy term, observations were used to evolve the model predictions. Such approaches are extremely useful in assessing the impact of pollution incidents associated with accidental releases of radioactivity. Similarly, they may also be applied to major air pollution incidents which require a fast response in terms of impact assessment and where data continues to be collected as the incident evolves, e.g. the 2005 Hertfordshire oil storage terminal fire.

5.2.5 Cross-validation

To provide an indication of the uncertainty in a calibration method, cross-validation of the assimilated data may be performed. Cross-validation involves carrying out the data assimilation process with all but one of the concentration measurements. The difference between the excluded measurement and the assimilated concentration is then calculated and the process is repeated for all measurements. Descriptive statistics, e.g. RMSE (Equation 5), NMSE (Equation 8) etc., may be used to give an indication of the error of the method using the cross-validation errors. This cross-validation method can be used to assess the suitability of integration methods using multiple monitoring observations.

[4] Statistical Estimation and Characterisation Techniques for use during Accident Response

Table 5.3 Description of data assimilation methods for use in air quality assessments (from Walker *et al*., 2006).

Method	Description
Variational Methods	Variational methods are derived from formulating the data assimilation problem as a variational problem. The search for an optimal assimilated model state is formulated as an optimisation problem involving the minimisation of a cost function. The cost functions are then constructed to describe the goodness of fit with both forecasted model state and with observations. Examples of variational methods include **3D-Var**, **4D-Var** and **Physical-space Statistical Analysis System (PSAS)**. The PSAS method is equivalent to the 3D-Var method but is more appropriate for situations with a smaller number of observations and when using grid cell concentrations only, e.g. Eulerian grid models without any sub-grid models. The 4D-Var approach is an extension of the 3D-Var approach whereby in addition to current observations at time step k, a set of future observations at time steps k+1,k+2…,k+L for some given time lag L > 0 is used to provide an estimate of the current model state.
Kalman Filter Methods	Kalman Filter methods automatically update the model error covariance matrix from one time step to the next using the dispersion model itself. These methods are ensemble based, i.e. the dispersion model is run N times to propagate an ensemble of model states to the next time step. The ensemble size depends on the problem to be solved but N typically ranges between 25 and 100. Examples include the **Ensemble Kalman Filter (EnKF)**, the **Reduced rank Kalman Filter (RKF)** and the **Ensemble Kalman Smoother (EnKS)** (similar to the 4D-Var method). An important feature of these methods is that model biases can be estimated and removed from the solution procedure.
Sequential Monte Carlo (SMC) or Particle Filter (PF) methods	SMC methods are based on statistical simulations of model error evolution using Monte Carlo random draw techniques and like Kalman Filter methods, are also ensemble-based. SMC methods are supported by a general Bayesian statistical framework, which makes no assumptions of linearity in model evolution or the Gaussian distribution of errors. Examples of this technique are **Sequential Importance Re-sampling (SIR)** and **Guided SIR**. Guided SIR is similar to 4D-Var and EnKS methods, whereby future observations are used to provide an estimate of the current model state.

5.3 Inverse modelling

Rates of emission (E, in g s^{-1}) can be predicted using inverse modelling methods, by dividing measured air concentrations (χ_a, in g m^{-3}) by modelled dispersion factors (D, in s m^{-3}). The dispersion factor can be calculated by conducting modelling assessments using a unit release rate providing model output as s m^{-3}. The calculated dispersion factor is a function of the prevailing meteorological conditions, the surface roughness, the release conditions (particularly the height and geometry of the source) and the potentially complex interactions of the dispersing material with buildings and terrain features.

Back-calculation techniques are particularly useful for application to diffuse emission sources and have been widely used in the determination of emission rates of atmospheric ammonia. Examples include determining emissions from single isolated sources, such as following the application of cattle slurry to grassland (McInnes *et al.*, 1985), from whole farms (Flesch *et al.*, 2005) and from farm waste stores (Hill *et al.*, 2007). The application of dispersion models by McInnes *et al*. (1985) and Flesch *et al*., (2005) benefited from significant reductions in modelling complexity due to well defined sources and on-site meteorological measurements. Furthermore, situations were considered where building effects were either not present (McInnes *et al*., 1985) or where dispersion estimates were made at sufficient downwind distances (>10 obstacle heights) that neglecting the complexity of flows close to the source would not contribute a significant error term (Flesch *et al*., 2005). This contrasts with the study of Hill *et al*. (2007) where the ADMS model was applied to disaggregate emissions from on-farm sources.

Other uses of back-calculation methods include determining release rates during and/or following accidents. For example, Davoine and Bocquet (2007) used inverse modelling to reconstruct the source term for the 1986 Chernobyl accident and Smith *et al*. (2007) applied the ADMS model to evaluate the source term arising from emissions of radioactive particles from the Windscale piles during the 1950s.

Further research on the application of models for evaluating release rates will be included in “Techniques to analyse monitoring data for source attribution and to determine exceedence conditions”, to be completed by Westlakes Scientific Consulting on behalf of the Environment Agency, scheduled for 2009/2010.

5.4 Optimisation of monitoring network design

The previous examples of integration have all highlighted the necessity for a sufficient number of monitoring locations to be able to apply integration techniques effectively. In reality, there is often an insufficient number of monitoring locations around Part A(1) sites to fully integrate observed and modelled data. There is very little guidance in Environment Agency literature about methods for determining the optimum number of sampling locations considered necessary for monitoring impacts from industrial sources. However, for the installation of multi-site networks, the Environment Agency recommends one of four options:

i. Location of sites on concentric circular lines around the area of interest

ii. Location of sites on typical trajectories of surface winds

iii. Location of a random heavy density of sites in the core of interest with random open spacing further out

iv. Location of sites on an equally spaced grid pattern

(Environment Agency, 2000)

When a large sampling network cannot be justified, the Environment Agency recommends sampling sites should be located to enable the maximum amount of information to be obtained from a minimum number of sites. For point sources, this is often achieved by the simple arrangement of upwind and downwind monitoring sites, located at a distance that depends on the height of the source. Area sources such as landfill sites require a different type of network design with monitors in the near-field and around the site boundary.

It would be impractical to suggest that every Part A(1) operator in England and Wales should have such detailed monitoring networks. However, there are several areas in the UK where a high density of Part A(1) industrial sites occur and could benefit from operating a collaborative network of greater density than could be achieved on an individual site. However, to ensure a sampling network is optimised, it is necessary to determine how many sampling sites are necessary and where they should be located in order to provide adequate measurement data for effective integration. This may be achieved using single or multi-objective approaches.

5.4.1 Single objective approaches

When systematic approaches to air quality monitoring were first introduced, site design studies were based on single objectives, predominantly locating sites in areas of maximum average concentration (Graves *et al*., 1981; Nakamori *et al*., 1979; Noll *et al*., 1977). The JEP Risk Management Framework adopts a single objective approach, which prioritises the most demanding AQS target (the 15-minute mean SO_2 objective) in order to locate sites where an individual source has a high frequency of impact. Monitoring sites are located as close as practicable to the modelled maximum 99.9^{th} percentile 15-minute SO_2 concentration. As the site of maximum impact changes on a yearly basis, this maximum is bounded by an isopleth corresponding to 50% of the maximum 99.9^{th} percentile 15-minute SO_2 concentration (Figure 5.2). For monitored concentrations to be considered representative of the maximum impact, the location of the monitoring site must fall within this isopleth boundary. The boundary is justified on the basis that the uncertainty applied to modelled concentrations (set to 50%) is such that any modelled concentration within the isopleth boundary is not significantly different from any other modelled concentration in that area at the 95% confidence level (Webb, 2004, Appendix A).

Along similar lines, Modak (1985) describes a method used to define the area around a monitoring site which can be considered representative of the measured concentration, otherwise known as the ‘detection area’. In this example, the size of a detection area can be determined by the spatial correlation of a monitoring site with adjacent sites. Elsom (1978) determined that a correlation coefficient of 0.8 was necessary to allow network rationalisation, whereby sites with a correlation of 0.8 or higher may be considered to provide the same information. Several authors have used spatial correlation to resolve detection areas (e.g. Liu *et al*., 1986; Langstaff *et al*., 1987; Arbeloa *et al*., 1993) with the approach also used to determine the potential for network rationalisation (Wu and Chan, 1997). It is important to note that an assumption of correlation analysis is that monitoring sites should be ideally located on a regular grid network on uniform terrain (Elsom, 1978; Handscombe *et al*., 1982). In using correlation analysis to determine the optimum number of monitoring locations for a point source, it may be appropriate to use modelled concentration data, where a

gridded network of point concentrations can be isolated and subsequent spatial correlation analysis performed.

The single objective approach to network rationalisation has been adopted in the industrial district of Marghera (Province of Venice). An industrial consortium network (Ente Zona Industriale) consisting of 13 stations was set up in the late 1970's to monitor the impact of emissions from process activities on air quality. The consortium network was established independently of another network consisting of 10 stations, located in the same district and operated by the Provincial Environmental Protection Department (Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto) (Figure 5.13). Given the density of monitors within a small area, both network operators are now in collaboration with a view to rationalising the number of monitors.

Basemap: Google Copyright

Figure 5.13 Map illustrating the location of monitors in the Marghera industrial district, Province of Venice.

5.4.2 Multi-objective approaches

The single-objective approach assumes that the optimal location for a monitoring site is determined by only one criterion which is most commonly the point of maximum impact or potential for violation of an air quality objective. However, in reality the optimal design of a monitoring network can be influenced by many criteria. These criteria characterise both the degree of exposure and the hazard or consequence of exposure to a pollutant. Hazard criteria may include short-term concentration, long-term concentration or the frequency of plume impact at a particular location. Exposure criteria may include population density, sensitivity of population receptors or sensitivity of ecological receptors. Aspects of data validity may also be included in multi-objective analysis, including criteria such as the spatial representivity of a particular monitoring location.

Multi-objective analysis enables the consideration of several criteria to be incorporated in a decision making process. Examples of research into multi-objective analysis include Truijillo-Ventura and Ellis (1991) who proposed a model which applied weighting to the criteria of spatial coverage, air quality objective violation potential and

data validity to find the most suitable design for a monitoring network in Tarragona, Spain.

Arbeola *et al.* (1993) used similar objectives (spatial coverage and violation potential) to find the optimal multi-pollutant network design for a hypothetical pot ash plant and two thermal power stations in open countryside, using a technique called *Pareto optimal design* (Modak, 1985). This technique calculates what is known as a 'utility function', taking account of the spatial coverage and violation potential of each monitoring site location for each pollutant. A global utility function is built as a weighted sum of the individual utility functions depending on the preference for an individual pollutant.

Kahn and Sadiq (2005) applied a fuzzy synthetic evaluation technique to determine a risk-based prioritisation of air pollution monitoring around a petroleum refinery. Potential monitoring locations were identified at 2 x 2 km grid node intersections of an 8 km x 12.5 km area grid. Fuzzy evaluation matrices were developed for all grid nodes, for hazard parameters calculated for a range of pollutants (CO, NO_x, PM_{10} and SO_x) and exposure parameters including population density, population location and population sensitivity. A modular hierarchical model was developed to provide a framework for the aggregation of hazard and exposure parameters in order to rank the potential monitoring sites.

Kao and Hsieh (2006) used a multi-objective mixed integer programming model to determine the optimum location for up to seven potential monitoring sites located around an industrial district in Taiwan. The following criteria were used to determine the optimal placement for monitors within the network:

i. The number of significant plumes[5] captured by a monitor (Maximum Detection Capability, DC)

ii. The maximum long-term exposure captured by a monitor (Maximum Dosage Detection Capability, DDC)

iii. The maximum detection area captured by a monitor (Maximum Detection Area, DA)

iv. The maximum population coverage captured by a monitor (Maximum Population Protection, PP)

The area affected by emissions from the industrial sources was divided into a series of grid squares to which the multi-objective model was applied (Figure 5.14b). In this study the Industrial Source Complex Short-Term (ISCST3) model was used to simulate pollutant distributions under hourly wind fields for an entire year, the results of which were then integrated with the multi-objective model. The multi-objective model allows the desired number of monitoring sites to be selected (in this example the desired number of monitoring stations ranges from 4 to 7) and also applies weighting to the criteria considered to be most important.

Figure 5.14c shows a scenario whereby optimal locations for five monitoring sites are modelled using single-objective analysis, from which the following conclusions were drawn. When 'Maximum Detection Capability' and 'Maximum Dosage Detection Capability' are the single objectives, the selected sites are distributed downwind of the major industrial sites. When 'Population Protection' is considered, sites are generally located close to a significant population concentration, shown in Figure 5.14b. When using the 'Maximum Detection Area' criterion, the established correlation coefficient

[5] A 'significant plume' is defined according to the 'potential zone' based on the concept of Noll et al. (1977), delineated by the daily average SO_2 concentration threshold (Kao and Hsieh, 2006).

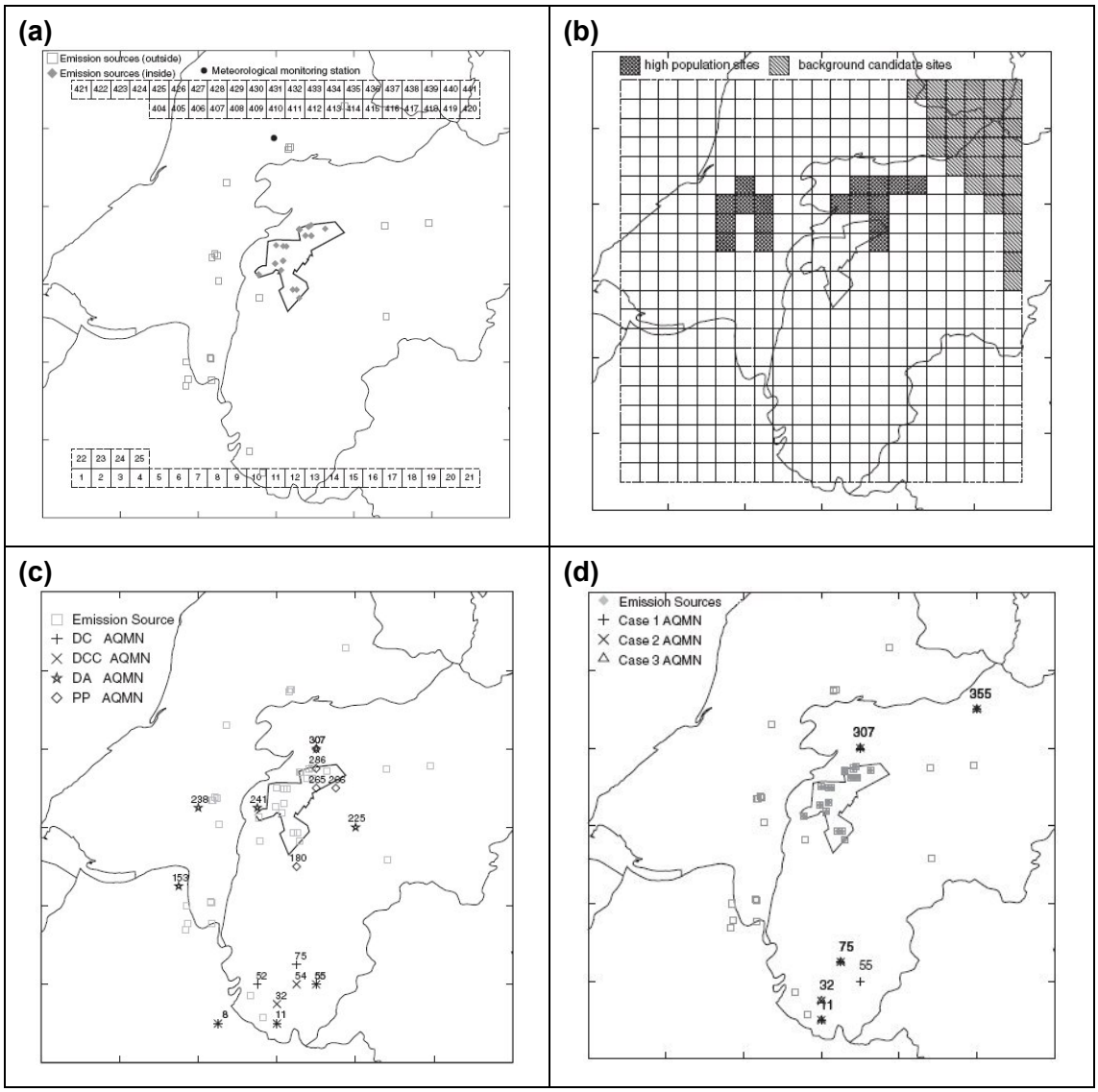
cut-off value is not related to concentration level and consequently sites based on the Maximum Detection Area objective are distributed throughout the model domain.

Figure 5.14d shows the results of multi-objective analysis for the same scenario, using three different criteria weighting scenarios (Case 1, 2 and 3). As the highest weighting has been given to the Maximum Detection Capability and Maximum Dosage Detection Capability criteria in all cases, the location of the monitoring sites has remained largely similar. However, the optimal monitoring locations are noticeably different to those derived using single-objective analysis.

There may be many more decision-making tools and techniques currently described in the literature, which could be applied to help locate monitors on the basis of multiple criteria. However, research into the application of such techniques to air pollution network design is limited. To determine the most applicable network design tool for industrial point sources, there is a need for a comprehensive review of the literature, both of existing techniques, e.g. *Pareto optimal design*, and techniques which are currently used as decision-making tools in other areas of scientific research.

5.4.3 Network design for source attribution

Where multiple sources are present, it is also desirable to be able to discriminate between the impacts of different sources. For example, Stewart et al. (2004) investigated a method to determine the optimum placement of monitoring sites within the Aire Valley, UK, for source attribution. Figure 5.15a depicts the “hour-glass formation”, relating to the area in which monitoring sites may be placed to determine the impacts of individual sources. The hourglass shape of Figure 5.15 is formed by assuming that it is possible to discriminate plume trajectories separated by angles greater than ± 20°. The current arrangement of monitoring sites was designed to measure maximum combined impacts from the Aire Valley stations and achieves either none or only partial separation of sources for source attribution. Figure 5.15b indicates the proportion of hours downwind of each power station for varying locations within the full separation zone. This identifies the optimum location for a monitoring site, which would allow the attribution of impacts to an individual power station source. The optimum monitoring site is the one that experiences the largest percentage of hours downwind of each power station in equal proportion.



Note: Abbreviations: DC, DCC, DA and PP denote the optimum location of monitoring sites based on maximum Detection Capability, maximum Dose Detection Capability, maximum Detection Area and maximum Population Protection, respectively.

Figure 5.14 Multi-objective monitoring network design in an industrial district in Taiwan. Emission sources, modelling grid and meteorological monitoring station in the study area (a), grids with high population density or low accumulated concentration (b), single-objective results (c) and multi-objective results (d) for 5 desired monitoring stations (from Kao and Hsieh, 2006).

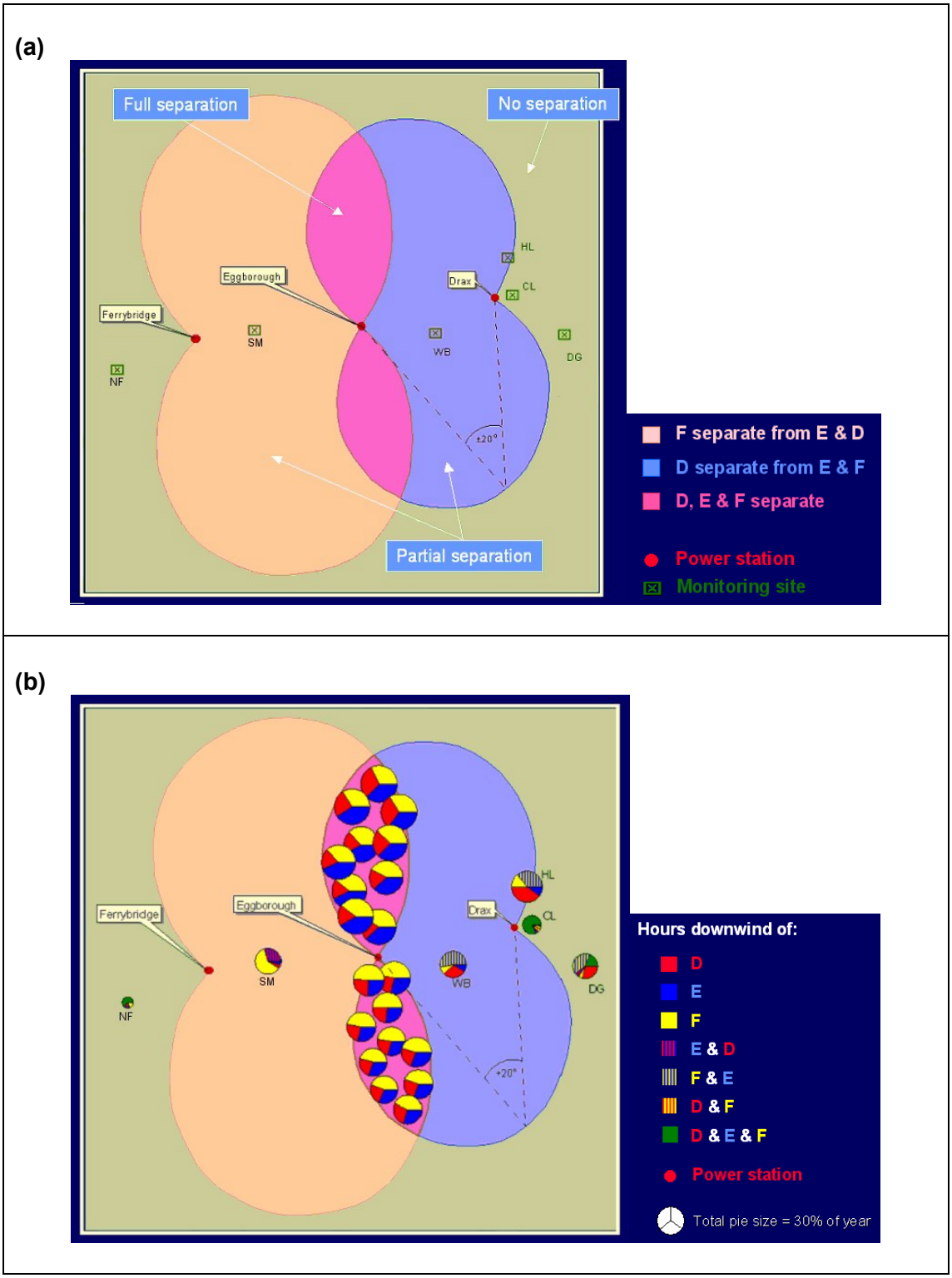


Figure 5.15 Network design for source attribution, showing full, partial and no separation zones (a) and optimal locations for monitoring sites in the full separation zone (b) (from Stewart *et al.*, 2004).

5.5 Emerging technologies

The number of monitoring stations required to generate sufficient amounts of data for effective integration with models far exceeds the infrastructure currently in place around industrial sites in the UK. However, improvements in accuracy and cost from new sensor technologies may mean that the use of integration techniques in air quality assessment may become both economically and practically justifiable. Examples of emerging technologies are discussed in this section.

New generation air pollution sensors, Generic Ultraviolet Sensors Technologies and Observations (GUSTO) are based on open-path DUVAS^{TM} (Differential Ultraviolet Absorption Spectroscopy) technology (Richards *et al*., 2006). These sensors are able to measure and transmit the volume-mixing ratio of key urban pollutants (SO_2, NO_x, NO_2, O_3, NH_3 and benzene) in real time. GUSTO sensors can record data on much shorter time scales than other analytical instruments (2 second scan rate) and have an open variable path of up to 30 metres, enabling spatially integrated measurements to be made. The sensors are also relatively cheap and robust and are therefore sufficient for large-scale deployment on moving platforms as well as stationary positions.

The downside of the development of large sensor networks is the need for data management, computation management, information management and knowledge discovery associated with the sensors and the generated data. Richards *et al*. (2006) developed the idea of a 'sensor grid' to standardise the way a sensor network is distributed and data collected, analysed and interpreted for specific end users. The author describes the Discovery Net system (Curcin *et al*., 2002) which is a form of high level informatics software, designed to support the analysis of large volumes of scientific data. Within this system, visual pollution trend analysis supported by Geographical Information Systems provides an effective way for an end user to monitor sensor data in real time and also to explore historical data sets. If the improvement of sensor technology enables the increased use of data-driven interpolation modelling, then the development of sensor grid assessment software will be necessary for the implementation of data analysis techniques.

6 Application of integration techniques: The extrapolation of short-term monitoring data

The Environment Agency conducts short-term monitoring campaigns using Mobile Monitoring Facilities (MMFs) in order to assess whether process operators are likely to meet air quality objectives. MMF campaigns are commonly limited to 3-6 months in duration and a pro-rata extrapolation method is used to convert the number of exceedences measured over the short-term to yearly exceedence figures. For example, if a monitoring campaign is 6 months in duration, then an estimate of the number of exceedences in a 12-month period is calculated by multiplying the short term total by 2. There is evidence that high ambient concentrations are associated with particular meteorological conditions (Bethan and Teasdale, 2005), therefore it is likely the Environment Agency extrapolation methodology may under- or over-estimate the true number of exceedences.

An analysis was conducted to test the efficiency of the Environment Agency's extrapolation method using the Aire Valley data set, comprising of six monitoring sites measuring the impact of emissions from three coal-fired power stations. The potential for the use of monitoring and modelling integration methods to improve the accuracy of short-term extrapolation was also examined.

6.1 Testing the Environment Agency pro-rata extrapolation method

6.1.1 Methodology

To examine the accuracy of the Environment Agency extrapolation method, hourly average SO_2 concentration data from 2003-2005 were analysed from all six monitoring sites. Sub-hourly average data (15-minute average) were also available but due to the complex analysis procedure it was only possible to examine hourly data.

A succession of rolling monitoring periods was simulated for each monitoring site from the 3-year dataset. Monitoring periods were defined as 3 and 6 months in length and were rolled on a monthly basis to produce 33, 3-month and 30, 6-month rolling monitoring periods for each monitoring site. The number of exceedences was then calculated for each rolling monitoring period at each monitoring site.

Table 6.1 shows the number of exceedences of the hourly SO_2 air quality objective (350 µg m^{-3}) for each monitoring site on a yearly basis. As it was necessary to compare predicted and measured exceedences to determine the accuracy of the method, the number of exceedences of the hourly objective were too few to use 350 μm^{-3} as the defined exceedence level. Therefore, rather than consider the formal SO_2 hourly average objective, it was considered necessary to use a lower concentration cut-off (200 μm^{-3}) to generate more data. Monitoring data for 2005 were excluded from the analysis as no exceedences of the 200 µg m^{-3} threshold occurred at any of the monitoring stations during this year. Figure 6.1 illustrates the general trend for decreasing emissions from all three power stations from 2003 to 2005. Most noticeable

is the decrease in emissions from Eggborough, which installed Flue Gas Desulphurisation (FGD) during 2004.

Table 6.1 Number of exceedences of hourly mean SO_2 concentrations: 350 μg m^{-3} and 200 μg m^{-3} at monitoring sites in the Aire Valley during 2003, 2004 and 2005.

SO_2 Cut-off	Year	Monitoring Site					
		CL	DG	HL	NF	SM	WB
350 μg m^{-3}	2003	1	0	0	0	4	0
	2004	0	0	0	0	0	1
	2005	0	0	0	0	0	0
	Total	**1**	**0**	**0**	**0**	**4**	**1**
200 μg m^{-3}	2003	6	5	4	3	6	11
	2004	1	2	2	0	3	5
	2005	0	0	0	0	0	0
	Total	**7**	**7**	**6**	**3**	**9**	**16**

CL = Carr Lane, DG = Downes Ground, HL = Hemingbrough Landing, NF = North Featherstone, SM = Smeathalls Farm, WB = West Bank

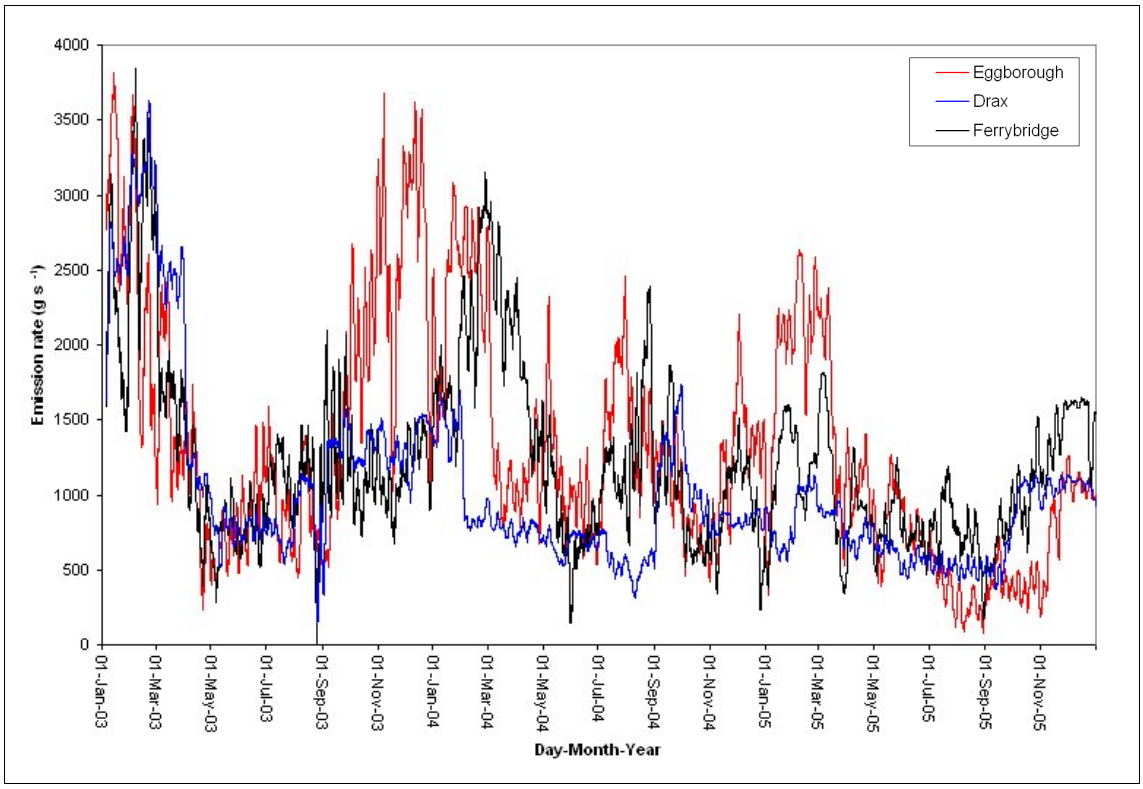


Figure 6.1 Emissions (g s^{-1}) of SO_2 from Aire Valley Power Stations (Eggborough, Drax, Ferrybridge) during 2003-2005 displayed as a rolling average of 100 hours.

The number of exceedences of 200 µg m^{-3} within each simulated monitoring period was calculated and extrapolated to 12 months using the pro-rata method. The extrapolated exceedence frequencies were then compared to the measured exceedence frequencies obtained from 12-month rolling monitoring data. However, where no exceedences were recorded in a 12-month monitoring period, the data were excluded from the analysis as any extrapolation based on this comparison would lead to an assumption of 100% accuracy of the method, which may be misleading.

6.1.2 Results and discussion

Table 6.2 shows the properties of linear regressions comparing measured annual exceedences with annual exceedences predicted using the Environment Agency pro-rate method. It is evident that extrapolating from a 6-month campaign provides a more accurate estimate of total exceedences than extrapolating from a 3-month campaign as 6-month gradient and r^2 values are generally much closer to 1. The exception is Smeathalls Farm (SM) where the gradient is equal to 1 for the 3-month monitoring period. However, the r^2 value for this relationship is very poor (0.10), and the RMSE very high (4.0). The RMSE statistic shows that uncertainty approximately doubles for a 3-month campaign when compared to a 6-month campaign.

Table 6.2 Properties of linear regressions comparing measured annual exceedences and annual exceedences predicted using the Environment Agency pro-rata method.

Monitoring Site	Length of simulated monitoring period	r^2	Gradient	RMSE	*N*
CL	3 months	0.66	2.08	3.0	16
	6 months	0.88	1.38	1.5	16
DG	3 months	0.18	2.09	3.7	20
	6 months	0.16	0.95	2.1	20
HL	3 months	0.49	2.88	2.6	20
	6 months	0.51	1.23	0.9	20
NF	3 months	0.01	-0.29	2.8	8
	6 months	0.16	0.71	1.8	8
SM	3 months	0.10	1.01	4.0	20
	6 months	0.52	1.36	2.4	20
WB	3 months	0.12	0.81	6.4	21
	6 months	0.45	0.98	3.2	21

CL = Carr Lane, DG = Downes Ground, HL = Hemingbrough Landing, NF = North Featherstone, SM = Smeathalls Farm, WB = West Bank

n = Number of rolling monitoring periods used in analysis

6.2 Integration Scaling Method (ISM)

There is potential to use a method that integrates monitoring and modelling to enable more accurate extrapolation from short-term monitoring campaigns. The Integration Scaling Method (ISM) uses modelled concentration data to produce a ratio of 6-month to 12-month exceedences, to provide scaling factors that can be used to extrapolate from short-term monitoring campaigns to annual exceedence estimates. The Aire Valley data set was used to test this method, in order to make a direct comparison with the previous analysis of the Environment Agency's extrapolation method.

6.2.1 Methodology

Within the Aire Valley, six monitoring sites measure the maximum combined impact of emissions from three power stations: Ferrybridge, Eggborough and Drax. The available data set contains measured 15-minute mean SO_2 concentrations for the period 2003 – 2005 for all six monitoring sites.

Modelled concentration data for each monitoring site within the Aire Valley were generated using ADMS 3.3. Hourly sequential meteorological data were supplied by the Met Office for Linton-on-Ouse (Grid reference 4492 4613), approximately 40 km north of the Aire Valley. A roughness length of 0.2 m was applied to the model domain, consistent with the roughness length used in the JEP modelling scenarios for the Aire Valley (Brooke *et al*., 2003). A separate roughness length was not assigned to the met data site as the roughness length at Linton was considered similar to that of the Aire Valley. The emissions data supplied by the Environment Agency were in the form of ADMS time-varying emission files, with emission rates measured in g s^{-1}. No correction was made for the influence of background concentrations. A higher than expected frequency of very calm conditions (18%) was recorded in 2003; for these hours concentrations were not modelled.

6.2.1.1 Frequency distribution analysis

To determine whether the generation of scaling factors from modelled data was appropriate, it was first necessary to examine the frequency distribution of measured and modelled concentrations, to determine the consistency of both data sets.

The frequency distribution of modelled and measured concentrations for all monitoring sites collectively, was calculated for 6-monthly rolling periods, using the following concentration bins: <40, 40-80, 80-120, 120-160, 160-200 and >200 µg m^{-3} (see Appendix A1 and A2 for frequency distribution data). It was expected that the frequency distribution should exhibit a log-linear relationship, characteristic of point source monitoring data, whereby frequency decreases with increasing concentration. However, rolling monitoring periods including data from January to August 2003 showed an unusually high occurrence of concentrations in the >200 concentration bin relative to the 120-160 concentration bin (Figure 6.2).

Figure 6.3 compares the average frequency of measured and modelled concentrations for all Aire Valley monitoring sites, within five concentration bins for 31, 6-monthly rolling monitoring periods. The frequency of modelled concentrations exhibits the characteristic log linear relationship as outlined above, however the frequency of measured concentrations is higher than expected for the >200 concentration bin.

Table 6.3 shows that for 2003, modelled and measured frequency distributions were consistent in all concentration bins, apart from the >200 bin. However, for 2004 and 2005 the difference between modelled and measured frequency distributions in all

concentration bins increases, with modelled frequency consistently overestimating in all concentration bins >40 µg m^{-3} apart from the >200 bin. This suggests that the raw modelled data are not entirely representative of the raw measurement data. However, trend lines fitted to the frequency distribution of modelled and measured data (shown in Figure 6.3) are not significantly different and show high r^2 values (0.99 for measured data and 0.89 for modelled data). This indicates that the modelled trends may be sufficiently representative to determine integration scaling factors.

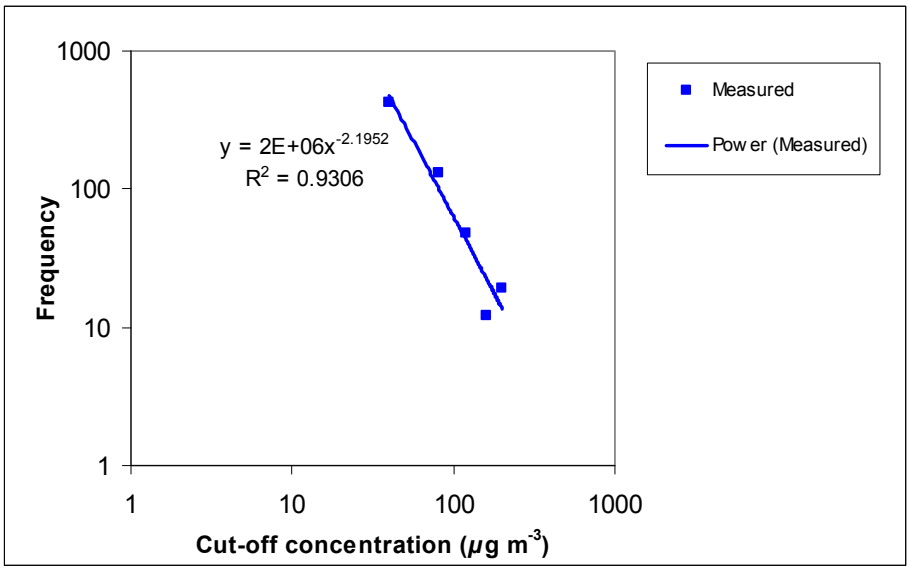


Figure 6.2 Comparison of the average frequency of measured concentrations for 8, 6-monthly rolling monitoring periods from January to August 2003, within five concentration bins; 40-80, 80-120, 120-160, 160-200, >200 µg m^{-3}.

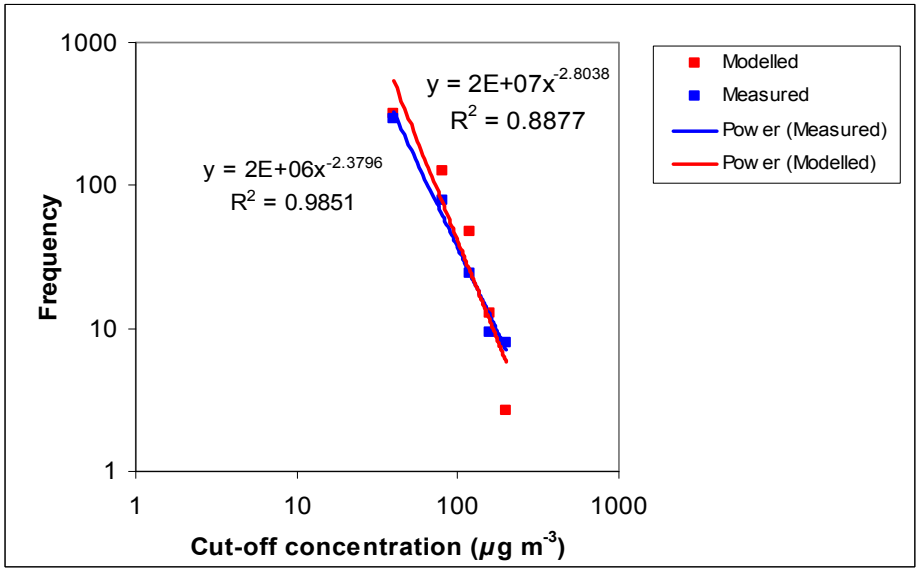


Figure 6.3 Comparison of the average frequency of measured and modelled concentrations for 31, 6-monthly rolling monitoring periods from 2003-2005 within five concentration bins, 40-80, 80-120, 120-160, 160-200, >200 µg m^{-3}.

Table 6.3 Comparison of average frequency distribution of measured concentrations for rolling 6-month periods for 2003, 2004, 2005 and all years.

Year	Measured/ Modelled	Frequency distribution of concentration bins (µg m^{-3})					
		<40	40-80	80-120	120-160	160-200	>200
2003	Measured	25675	439	135	50	13	19
	Modelled	25701	390	158	64	14	3
2004	Measured	25874	354	98	23	15	8
	Modelled	25633	413	202	92	28	5
2005	Measured	26112	160	28	9	1	0
	Modelled	25882	318	104	23	4	1
All years*	**Measured**	**25903**	**293**	**78**	**24**	**9**	**8**
	Modelled	**25813**	**318**	**127**	**47**	**13**	**3**

* Includes rolling 6-month periods overlapping years.

6.2.1.2 Generation of scaling factors

Modelled concentration data were used to generate scaling factors from which the number of exceedences measured during a 6-month campaign could be extrapolated to a 12-month period.

The 'best fit' power law relationship applied to the modelled and measured data (shown in Figure 6.3) was used to determine 'filtered' data sets to remove the 'noise' from the data. Filtered frequency distributions were derived for modelled concentration data for each of the monitoring sites in the Aire Valley. This process generated decimal numbers, which were subsequently rounded to integers. Scaling factors were subsequently calculated as the ratio of 6-month to 12-month modelled exceedences, using the filtered modelled data.

These scaling factors were then applied to the measured data to extrapolate the number of exceedences from 6 to 12 months. Scaling factors were applied to both filtered and un-filtered measured data for comparison. The number of exceedences predicted using the scaling factors was rounded to the nearest whole number before a comparison was made with measured 12-month exceedences. To maintain a similar approach to that used for testing the Environment Agency extrapolation method, the concentration defining an exceedence was classified as 200 µg m^{-3}.

To summarise, comparisons were made between measured 12-month exceedences (>200 µg m^{-3}) and;

- Predictions derived by applying scaling factors generated from filtered modelled data to unfiltered measured data (Integration Scaling Method 1)
- Predictions derived by applying scaling factors generated from filtered modelled data to filtered measured data (Integration Scaling Method 2)

6.2.2 Results and discussion

Table 6.5 compares the application of both scaling methods by examining properties of the linear regression (r^2 and *gradient*) fitted to measured vs predicted exceedences and

the RMSE statistic. The RMSE values indicate that neither method is consistently better than the other in accurately predicting 12-month exceedences.

When comparing these results to the results from the analysis of the pro-rata extrapolation method, the ISM methods provide a consistent improvement in RMSE, with the exception of West Bank (WB). However, the ISM method does not show any significant improvement over the Environment Agency data extrapolation method with respect to r^2 and *gradient* properties of the linear regression.

The integration scaling method aims to improve the accuracy of extrapolation of measured exceedences by utilising modelled concentration data. By comparing the cumulative frequency distributions of concentration bins for all monitoring sites, the frequency distributions for monitoring and modelling data showed some inconsistencies. The 'filtering' of modelled frequency distribution data aimed to reduce these inconsistencies, although it is possible that this had an adverse affect on the success of this method and that better agreement between the frequency distribution of modelled and measured data is necessary for this method to be successful.

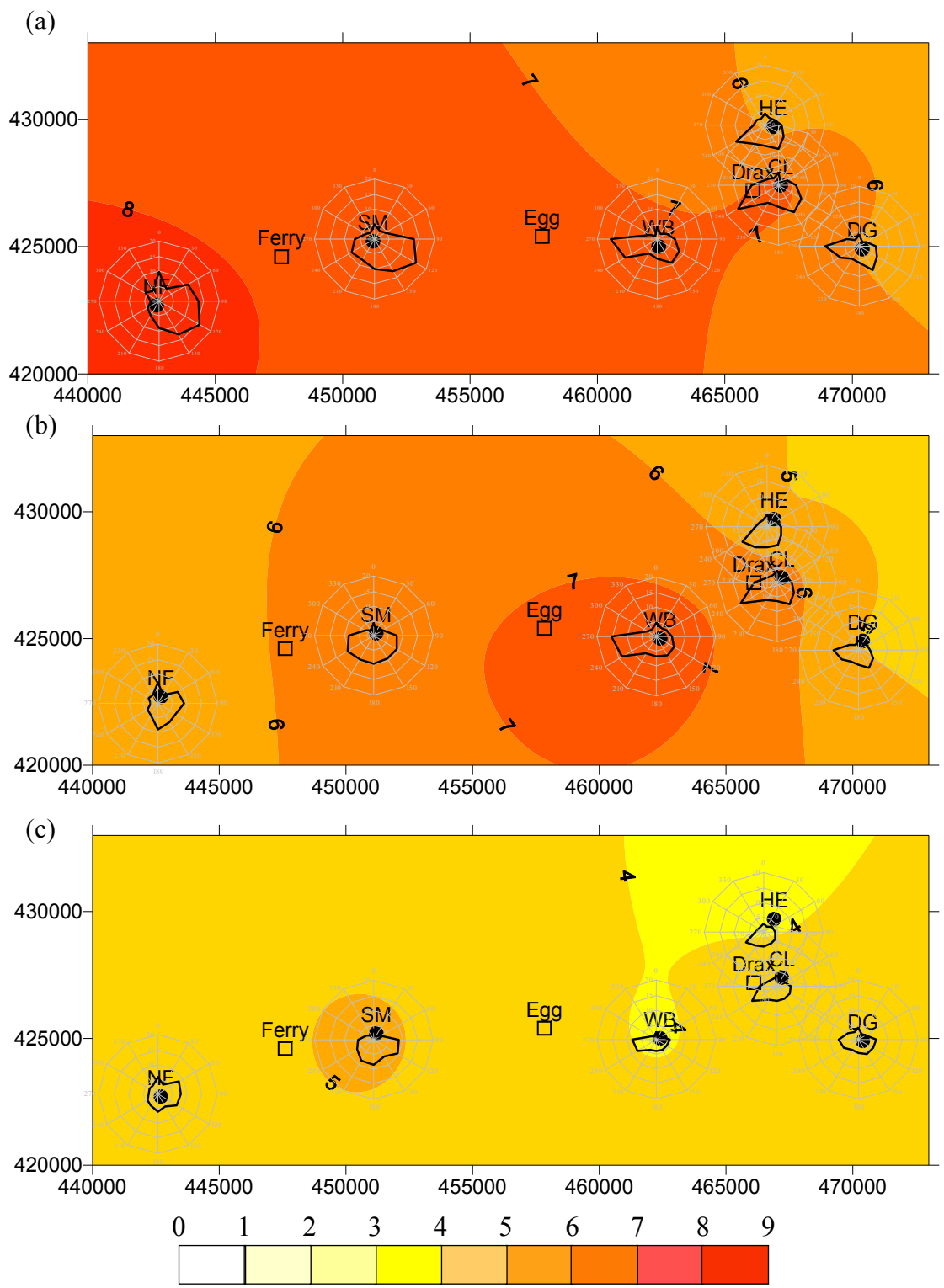
The modelled concentrations used in this analysis were not adjusted for background SO_2 concentrations, which may have influenced the scaling factors by affecting the frequency of modelled data in each concentration bin. However, upon closer examination of the SO_2 annual average and 99.9th percentile pollution rose plots (Figures 6.4 and 6.5), it is evident that a SO_2 source to the South East of the Aire Valley monitoring site was influencing concentrations measured at the monitoring sites to varying degrees. It is possible this source may have been the Trent/Soar Valley network of power stations (Environment Agency, 2003), although this is not conclusive and in any case, recent data suggest this source is now a historic issue. However, in the data shown here, a substantial proportion of the total number of exceedences at several Aire Valley monitoring sites may be directionally attributed to a source other than the three Aire Valley power stations. It is therefore not surprising that the scaling integration method fails to consistently improve the extrapolation of exceedences from short-term monitoring campaigns. It is necessary therefore to ensure that modelled concentration data are representative of measured concentrations before the integration scaling method is applied. Where appropriate, it may be possible to use wind direction filtering to exclude impacts from unmodelled sources to improve the performance of this technique.

Table 6.4 Results of the application of the scaling methodology measurement data.

Monitoring Site	Pro-rata Method			[a]Integration Scaling Method 1			[b]Integration Scaling Method 2		
	r^2	*Grad*	*RMSE*	r^2	*Grad*	*RMSE*	r^2	*Grad*	*RMSE*
CL	0.88	1.38	1.5	0.78	0.94	0.6	0.51	0.48	1.0
DG	0.16	0.95	2.1	0.19	1.41	1.9	0.12	0.79	1.7
HL	0.51	1.23	0.9	0.45	0.72	0.6	0.15	0.38	0.8
NF	0.16	0.71	1.8	0.00	0.07	1.0	0.03	0.29	1.1
SM	0.52	1.36	2.4	0.30	1.08	1.9	0.25	0.49	1.4
WB	0.45	0.98	3.2	0.28	1.05	3.9	0.18	0.57	3.5

CL = Carr Lane, DG = Downes Ground, HL = Hemingbrough Landing, NF = North Featherstone, SM = Smeathalls Farm, WB = West Bank

[a]Filtered ratio applied to unfiltered measured data; [b]Filtered ratio applied to filtered measured data

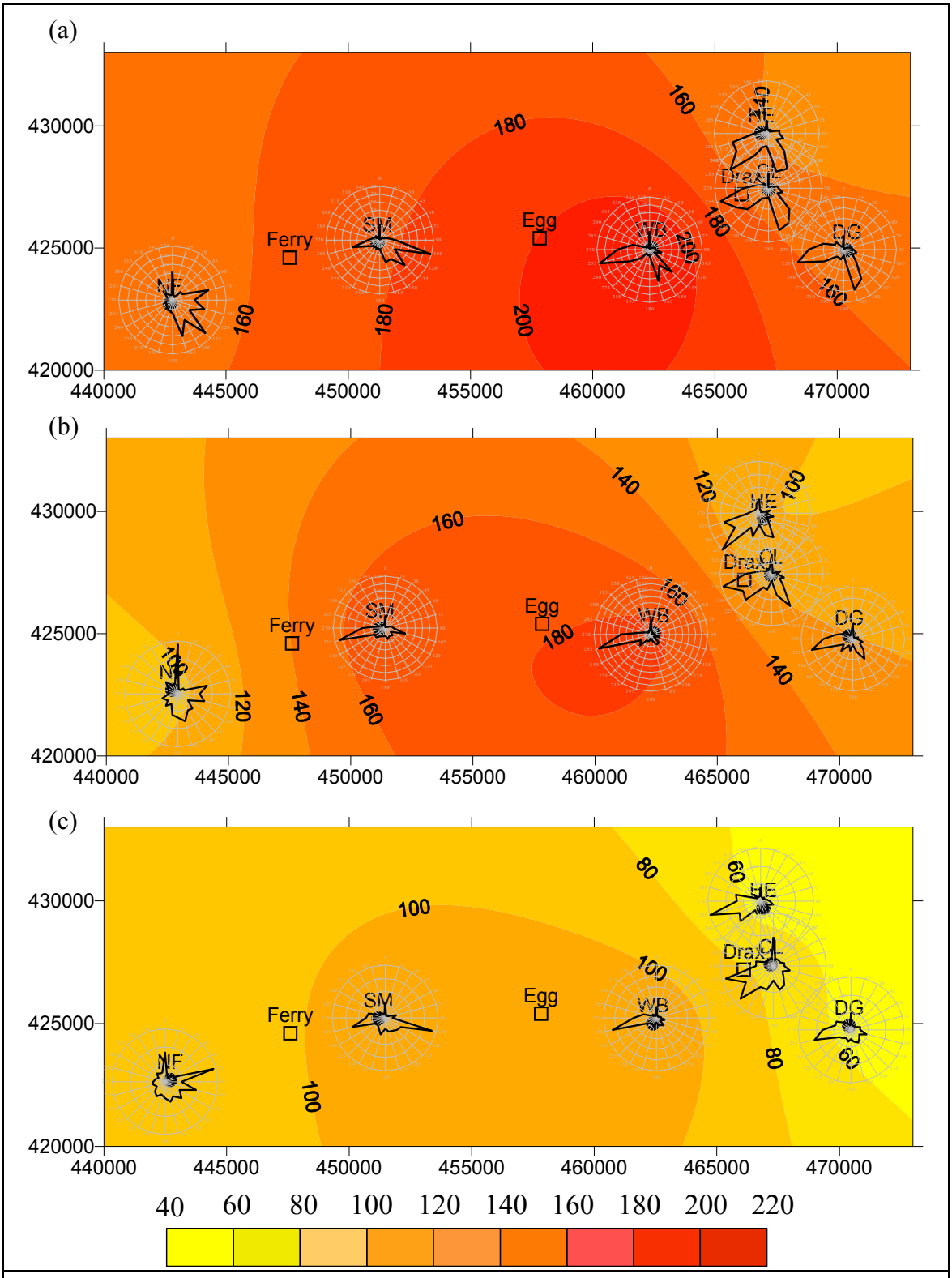


Note 1: Abbreviations; 'Ferry', 'Egg' and 'Drax' denote power stations; Ferrybridge 1 and 2, Eggborough and Drax, respectively.

Note 2: Abbreviations; NF, SM, WB, HE, CL and DG denote monitoring sites; North Featherstone, Smeathalls Farm, West Bank, Hemingbrough Landing, Carr Lane and Downes Ground, respectively.

Note 3: Pollution rose segments are split into 30˚ sectors with intervals of 5 μg m^{-3}.

Figure 6.4 Interpolated plots of annual average SO_2 concentrations (μg m^{-3}) for monitoring sites in the Aire Valley network, with superimposed sector-specific annual average pollution rose diagrams for (a) 2003, (b) 2004 and (c) 2005.



Note 1: Abbreviations; 'Ferry', 'Egg' and 'Drax' denote power stations; Ferrybridge 1 and 2, Eggborough and Drax, respectively.

Note 2: Abbreviations; NF, SM, WB, HE, CL and DG denote monitoring sites; North Featherstone, Smeathalls Farm, West Bank, Hemingbrough Landing, Carr Lane and Downes Ground, respectively.

Note 3: Pollution rose segments are split into 15˚ sectors with intervals of 20 μg m^{-3}.

Figure 6.5 Interpolated plots of 99.9th percentile SO_2 concentrations (μg m^{-3}) for monitoring sites in the Aire Valley network, with superimposed sector-specific 99.9th percentile pollution rose diagrams for (a) 2003, (b) 2004 and (c) 2005.

6.3 Meteorology-based extrapolation method

In contrast to the integration scaling method assessed above, there is evidence to suggest that exceedences may be commonly associated with specific meteorological conditions. The JEP have identified meteorological conditions associated with exceedences by examining modelled SO_2 concentrations for a receptor site using three meteorological data sets (Bethan and Teasdale, 2005). They identified that exceedences of the 15-minute SO_2 objective are associated with a weakly unstable boundary layer ($L_{MO} \approx -1000$ m), moderately high wind speeds (9-12 m s^{-1}) and a positive sensible heat flux (>70 W m^{-2}). In combination, these events are relatively rare occurrences and according to Bethan and Teasdale (2005), usually occur on a windy spring/summer day in the early afternoon. In contrast, they found conditions that give rise to the maximum concentration (100^{th} percentile) arise from very low heat flux values and wind speeds between 4-6 m s^{-1}. However, there are also other factors to consider when relating monitored concentrations to meteorological conditions. Both the strength of the emission source and the connectivity of the source to the receptor (determined by the wind direction) will have a crucial influence on whether exceedences occur at a receptor, regardless of atmospheric stability or wind speed.

6.3.1 Methodology

Multiple regression analysis was conducted to determine whether a simple model could be used to predict concentrations at receptor sites using local point source emission data and meteorological data. If this was shown to be true, then an extrapolation method similar to that used in the integrated scaling method could be applied. This would consist of estimating the ratio of exceedences for a period equivalent to a short-term monitoring campaign and a full year, according to the occurrence of specific meteorological conditions and source emissions. The ratio may then be applied to short-term measured exceedence figures to extrapolate the total number of exceedences in one year.

The response variable consisted of hourly average SO_2 concentrations from monitoring sites in the Aire Valley. The square root of the concentration data was used in an attempt to normalise the response variable to satisfy the assumptions of the analysis method. Explanatory variables included SO_2 emissions in g s^{-1} from three power stations (Eggborough, Drax and Ferrybridge) and meteorological variables (wind speed (m s^{-1}) and Monin-Obukhov Length). Wind direction was accounted for by including in the analysis only those hours in which wind direction originated from the direction of the source relative to the receptor. The multiple regression analysis for a specific receptor site only included hours when the wind direction occurred within a 25-degree wind arc from the receptor to the source. The analysis also considered overlapping wind arcs from multiple sources.

The statistical package Statistica® (Version 5) was used to perform multiple regression analysis in a forward stepwise process for each monitoring site. Forward stepwise regression starts with an empty model and subsequently adds the variable that has the smallest (most significant) P value. Each subsequent step adds the variable that has the smallest P value in the presence of the predictors already in the equation. Variables are added one-at-a-time as long as their P values are small enough, typically less than 0.05 or 0.10. Explanatory variables are removed from the model if they become non-significant as other predictors are added. This process continues until the model converges, i.e. no more explanatory variables are significant.

6.3.2 Results

Table 6.5 shows the significance of the explanatory variables included in the multiple regression analysis carried out for each of the monitoring sites. Generally, emissions emerge as the most significant explanatory variables in the regression and frequently the emission source emerging as most significant is the source closest to the monitoring site. The only exception to this is Carr Lane, for which the multiple regression includes all emission sources as their wind-arc boundaries overlap. In this case Drax is the closest power station to the receptor (Drax 1.1 km, Eggborough 9.6 km and Ferrybridge 19.9 km). However, emission from Ferrybridge is the only explanatory variable included when the model converges. In this instance, the short distance from Drax to Carr Lane monitoring site is unlikely to have been sufficient to allow Drax's plume to impact at ground level, meaning it simply over-flies the monitor.

Higher emissions are generally associated with higher concentrations, although there are exceptions to this rule associated with Drax emissions. Of the three occasions when Drax emissions are brought into the model before convergence, decreases in emissions are associated with increases in monitored concentrations on two occasions (Downes Ground and North Featherstone). However, in these instances the variable is not classed as significant at the 95% level.

For Downes Ground, Drax is the nearest emission source (4.9 km from the receptor), which appears to influence concentrations. However, Drax emissions are negatively correlated with concentration. In the second multiple regression including Eggborough and Ferrybridge (12.6 and 22.9 km from the receptor respectively), only Eggborough is included when the model converges.

For Hemingbrough Landing, Eggborough (10 km from the receptor) has a significant effect and although included in the same wind arc, Ferrybridge (20.1 km from the receptor) is not significant at the 95% level. In a separate analysis Drax (2.6 km from the receptor) is close to being significant at the 95% level and on this occasion has a positive relationship with concentration.

North Featherstone, is located to the west of all emission sources which in distance-order, occur as follows; Ferrybridge 5.3 km, Eggborough 15.4 km and Drax 23.8 km. Ferrybridge is included in the model and is significant but Eggborough is not included. Therefore the closest emission source is the most significant. Drax is also included but is again negatively correlated with concentration.

Smeathalls Farm is located between Ferrybridge to the west and Eggborough to the east, which are 3.7 and 6.6 km from the receptor respectively. Both emissions from Eggborough and from Ferrybridge are found to be significant in the separate analyses. Emissions from Drax are not included in the converged model.

West Bank is closest to Eggborough in the west and Drax to the east, with Ferrybridge also to the west of the receptor. For the wind-arc from the west, Eggborough is significant but Ferrybridge is not included in the model. From the East, no emission source is included in the converged model.

In terms of meteorological conditions (windspeed, U, and stability, $1/L_{MO}$), U is a highly significant variable on all but one occasion and frequently significant to the 99.9% level. L_{MO} is also included in the converged model for 6 out of 10 regression models, although is only significant on 4 occasions. However, the percentage of the variance explained by the models is very poor, indicated by the range of r^2 values (0.2-17%).

Table 6.5 Multiple regression results.

Monitoring site	Wind Arc Min	Wind Arc Max	P-values					Adjusted r^2
			Drax	Egg	Ferry	U	L_{MO}	
CL	245º	274º	-	-	0.00E+00	0.00E+00	-	1.32E-01
SM	70º	101º	-	1.61E-02		0.00E+00	3.98E-02	1.74E-01
	248º	273º			0.00E+00	0.00E+00	-	1.30E-01
DG	286º	311º	1.26E-01			-	0.00E+00	6.69E-02
	257º	285º		3.66E-04	-	1.20E-05	4.35E-02	7.84E-02
HL	232º	268º		3.09E-03	-	0.00E+00	-	1.58E-01
	185º	210º	5.63E-02			2.15E-02	3.21E-01	2.28E-01
NF	56º	92º	1.61E-01	-	2.00E-06	0.00E+00	-	1.37E-01
WB	256º	288º		1.67E-04	-	0.00E+00	6.09E-02	1.43E-01
	47º	72º	-			3.57E-02	1.89E-01	4.92E-02

Note: Emissions from sources outside the wind-arc boundary are denoted by grey cells in the table and are not included in the analysis.

Note: Variables that are included in the analysis but excluded from the converged model are denoted by the symbol, -.

6.3.3 Discussion

The results in Table 6.5 indicate that it is inappropriate to generate a simple model to explain the complex controls on monitored concentrations. However, an explanation for the poor results may be found in the choice of statistical method. Multiple regression is a parametric statistical technique with several assumptions and limitations:

- Assumption of linearity between response and explanatory variables
- Assumption of normality of residuals
- Potential for multi-colinearity between explanatory variables i.e. explanatory variables may be implicitly related
- Needs sufficient sample size (should be greater than ten times the number of explanatory variables)

It is likely that some or all of the basic assumptions of this statistical test are not met and therefore that the results are not representative of the true relationship between the response and explanatory variables. An investigation of the use of more complex non-parametric tests may be necessary for the further development of this method.

In addition to the problems outlined above, relating concentrations directly to emission rates implies that the impact of emissions is felt immediately at the receptor site. In reality, pollutants emitted from a source will impact on receptor sites at an unknown time after their release. This time lag will vary with wind speed and distance from source to the receptor. In addition, the use of time-series data may be inappropriate for multiple regression because of the potential for auto-correlation of the response variables, i.e. concentration measured during one hour may be affected by the concentration measured during adjacent hours at the same location.

7 Application of integration techniques: Exploring integration methods for model calibration

A major application for the integration of monitoring and modelling is model calibration. For retrospective assessment purposes, model calibration is useful in determining realistic impacts in areas for which monitoring data are not available and also to validate pre-year modelling assessments, e.g. the LAQM methodology (5.2.1). However, there is also the potential to apply model calibration to prospective modelling assessments, for example, the FRAME calibration methodology uses measurement data to normalise modelled data for projected estimates of deposition (5.2.2). This analysis will examine four calibration methods to determine their effectiveness in calibrating modelled predictions of air quality impacts from industrial point sources. Throughout the discussion, it should be remembered that although calibration techniques are able to address uncertainty in model outputs caused by systematic errors, they are unable to address the uncertainty due to random errors.

The four calibration methods examined are: 1) linear regression method; 2) simple ratio method; 3) kriging of the ratio; 4) kriging of the residual.

Both the linear regression and simple ratio methods are simple offline methods, which perform a global adjustment of the modelled concentration field. The linear regression method is similar to the LAQM methodology whereby a linear regression model is derived from a comparison between measured and modelled data and the modelled data are subsequently corrected according to the properties of the regression. The simple ratio method is similar to the FRAME integration methodology in that a ratio of measured to modelled concentrations is derived, however, in this application a global adjustment of the modelled concentration field is conducted using the average of all measured to modelled concentration ratios. The kriging methods are more complex in that they involve a spatially varying calibration of the modelled concentration field. It is possible to produce a kriged surface using raw measurement points. However, for the purposes of model calibration, ratios of modelled to measured concentrations and residuals (meaning the difference between modelled and measured concentrations) have been derived in order to generate a kriged surface with which to calibrate the modelled data.

The cross-validation technique referred to in Section 5.2.5 is used to determine the influence of monitoring site density on the performance of the calibration methods. However, in this case, rather than removing just one monitoring point from the data set, the systematic removal of several monitoring points will determine the number of monitoring points necessary to optimise the calibration. Specifically, optimisation is achieved when the rate of improvement in model calibration no longer increases with the use of additional monitoring data. The same analysis is also applied to the 2003 Aire Valley SO_2 monitoring data set, to determine whether results are consistent between data sets with multiple rather than singular sources. However, the Aire Valley is not a validation data set and therefore has a limited number of monitoring sites with which to conduct the cross-validation process.

7.1 Methodology

7.1.1 Kincaid data set

The Kincaid data set is a model evaluation data set containing 6024 hours of SO_2 data, which were reported at 28 sites around the Kincaid power plant in Illinois, USA, during 1980-1981. The Kincaid power plant is situated on flat terrain surrounded by farmland and lakes, consequently the roughness length used for the modelling lies between 5 and 15 cm and the height of the source stack is 187 metres. The reported data (monitoring and modelling predictions) are available from the US Environmental Protection Agency web site (http://www.epa.gov).

7.1.2 Aire Valley data set

Discussed previously in Section 6.2.1.

7.1.3 Sector-correcting measured concentrations

In an ideal scenario, a model will incorporate all major pollution sources and the calibration procedure then acts to compensate for model inadequacy and poorly defined background contributions. However, in both the Aire Valley and Kincaid data sets, the measured concentrations are affected by sources that are external to the model domain. Pollution rose plots clearly show that for both data sets, external sources influence the annual mean and 99.9^{th} percentile concentrations measured at monitoring sites. Figures 6.4 and 6.5 show that a strong influence from the south-east is apparent in the Aire Valley whereas Figures 7.1 and 7.2 show that for the Kincaid data set, the influence is greatest for monitoring sites in the north-west of the model domain. Therefore, in order to analyse the performance of the different calibration methods, it was necessary to remove the influence of external sources on the measured concentrations by 'sector-correcting' the measurement data. This was achieved by calculating mean concentrations using measured data only where the wind was blowing the plume from the source to the receptor site. The calculation of mean concentrations using wind arcs ranging from 5° to 100°, increasing in increments of 5° (Figure 7.3), was necessary to determine the optimum sector size. Note that in the Aire Valley, the JEP apply a constant background correction prior to comparison with modelled data, rather than using the sector-correction approach described here.

The Root Mean Square Error (Equation 5) and the regression r^2 of the modelled vs. sector-corrected monitored data were calculated for each sector and used to determine the sector size for which optimum data compatibility could be achieved.

The optimum sector size was determined where the RMSE is minimised, the r^2 is maximised, and where RMSE/R^2 is closest to zero. Determining the optimum sector size may be quite subjective, for example, for the Kincaid data set, the optimum sector size was determined to be 25° as even though the RMSE is not at its minimum, the r^2 rapidly decreases above this sector size (Figure 7.4). However, for the Aire Valley data set, the RMSE is fairly constant above a sector size of 25°, although the r^2 is maximised and RMSE minimised at 60° (Figure 7.5). Consequently, optimum sector sizes of 25° and 60° were used for the Kincaid and Aire Valley data sets respectively.

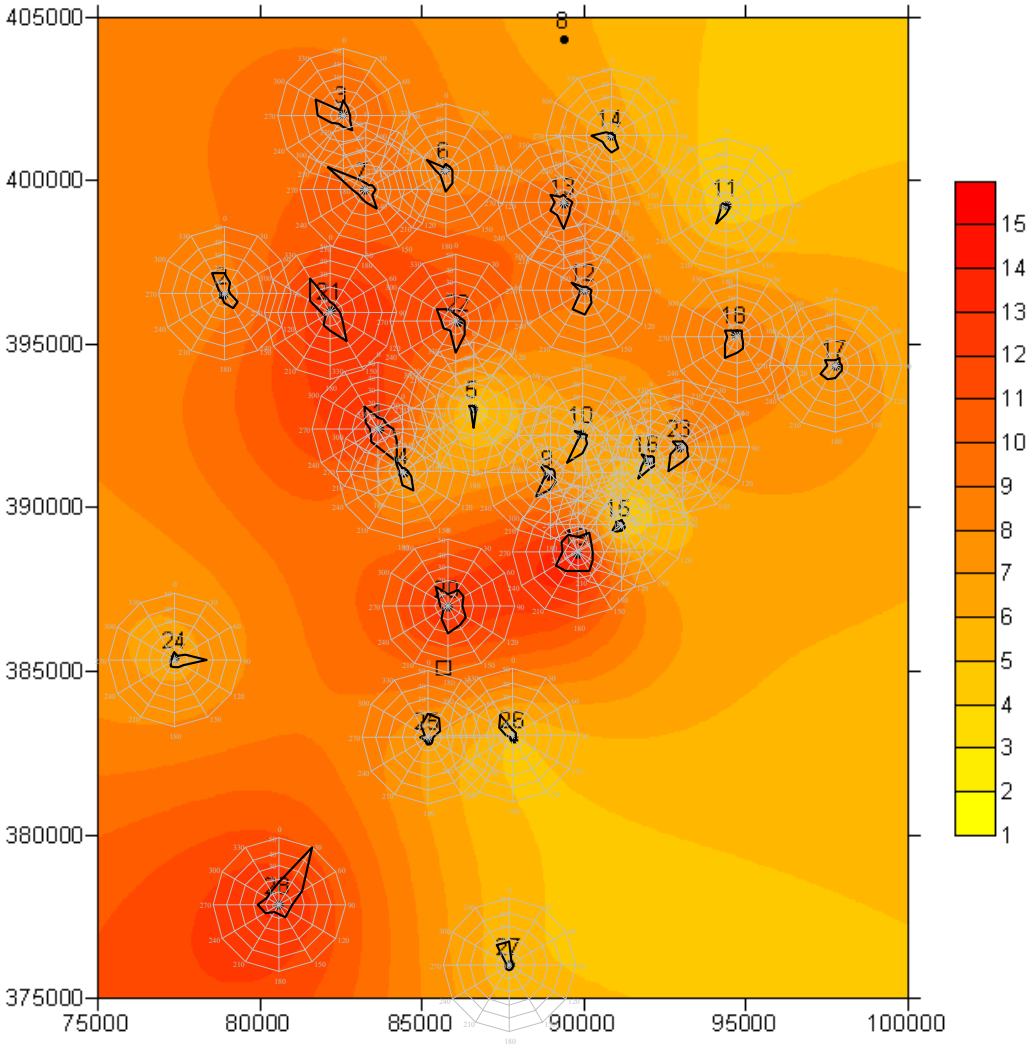


Figure 7.1 Interpolated plots of annual mean SO_2 concentrations (μg m^{-3}) for monitoring sites in the Kincaid network, with superimposed annual mean sector-specific pollution rose diagrams for data from 1980/81.

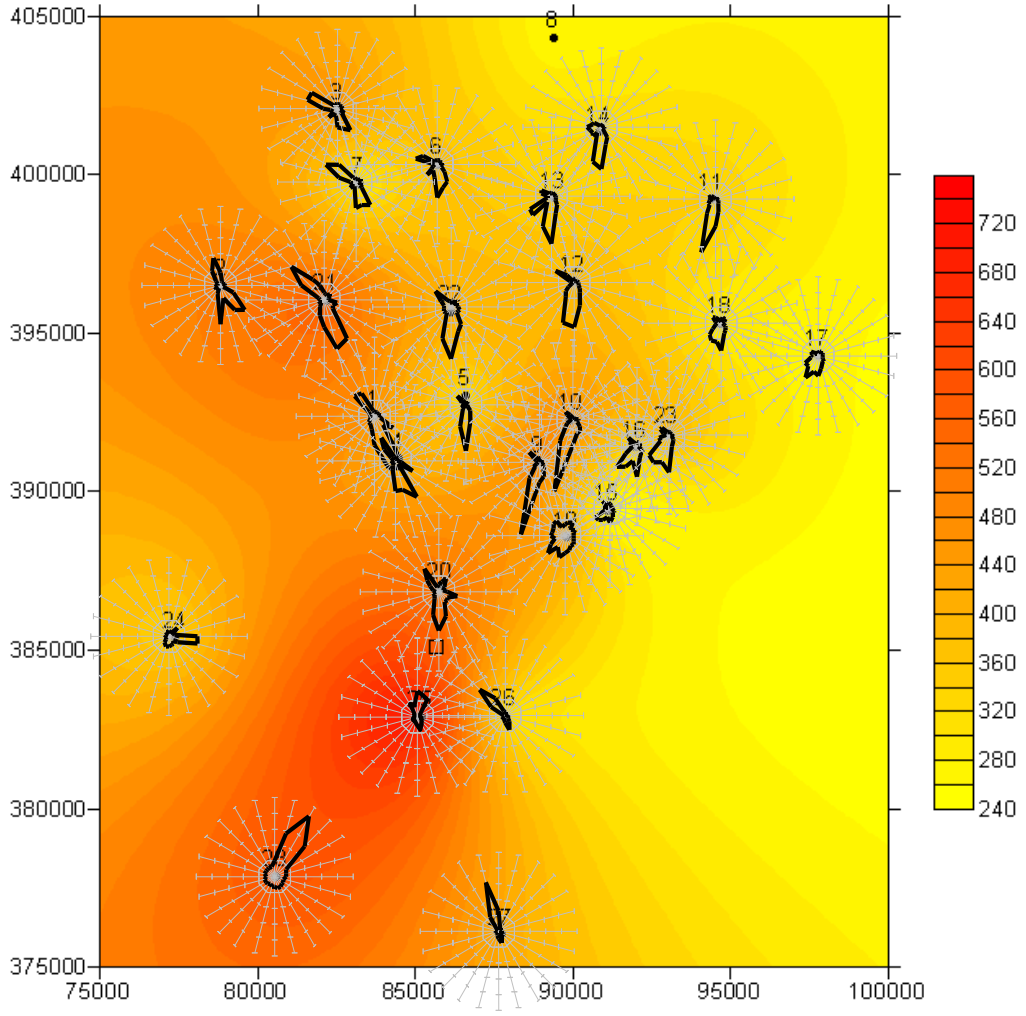


Figure 7.2 Interpolated plots of 99.9th percentile SO_2 concentrations (µg m^{-3}) for monitoring sites in the Kincaid network, with superimposed exceedence frequency pollution rose diagrams for data from 1980/81.

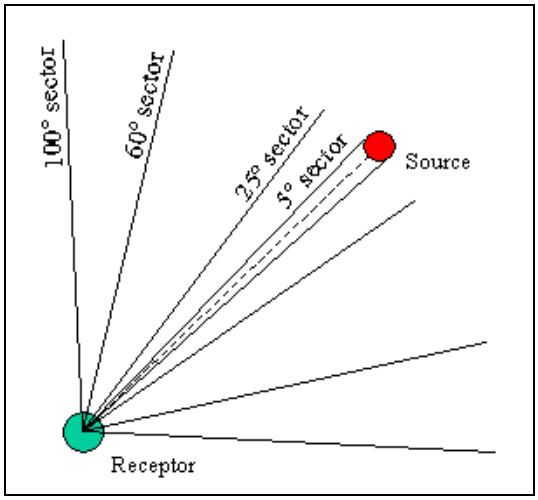


Figure 7.3 Sector correction analysis schematic.

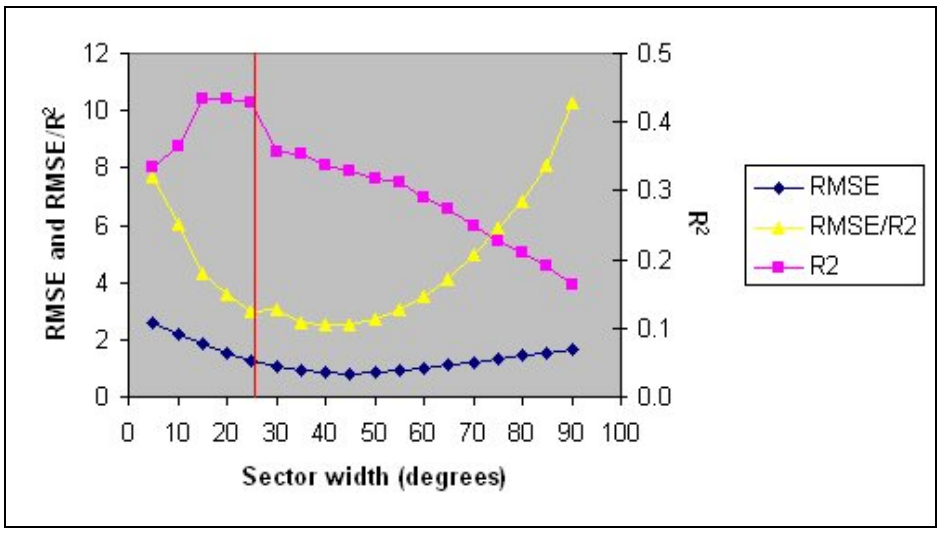


Figure 7.4 Sector correction analysis for Kincaid data set. Vertical red line shows optimal sector size.

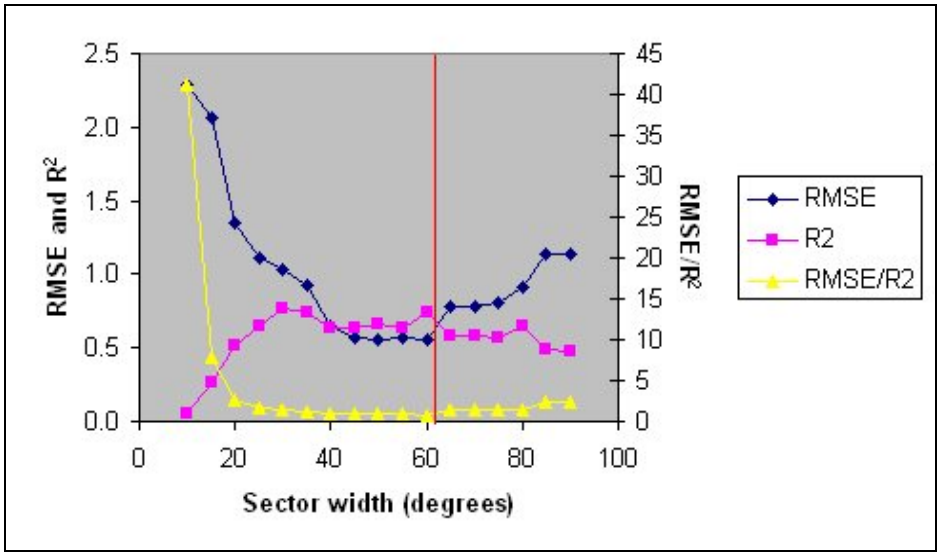


Figure 7.5 Sector correction analysis for Aire Valley data set. Vertical red line shows optimal sector size.

7.1.4 Analysis of the calibration methods

The calibration methods were analysed using a cross-validation technique, which involved the systematic removal of monitoring data. The retained monitoring data were used to conduct the calibration and the removed monitoring data were used to test the method by comparison with the calibrated model results. In this way, the performance of each calibration method could be assessed and the optimisation of each method on the basis of monitoring site density could be determined.

For the Kincaid data set, a maximum of twenty-four monitoring points was used to conduct the analysis of annual average concentrations and twenty-three for the analysis of 99.9th percentile concentrations. This difference was due to the presence of an anomalous modelled result at monitoring point 20 where the modelled 99.9th

percentile concentration was 817 µg m^{-3} compared to a corrected measured concentration of 271 µg m^{-3}.

For the Aire Valley data set, a maximum of five data points was used to conduct the analysis, as there are six monitoring sites in total with at least one site necessary to test the performance of the calibration. In terms of the minimum requirements for monitoring data, the simple ratio method requires a minimum of one data point, the linear regression method a minimum of two data points, and the kriging methods a minimum of three data points. Each calibration performed is termed a 'calibration scenario' which may be described by a specific calibration method and a specific number of data points used to conduct the calibration. In total, for the Kincaid data set, seven calibration scenarios were conducted for each calibration method for both mean and 99.9th percentile concentration metrics. For the Aire Valley data set, five scenario's were conducted using the simple ratio method, four using the linear regression method and three using the kriging methods, repeated for both mean and 99.9th percentile concentration metrics. This was the maximum number of scenarios possible with the available data. The kriging residual and kriging ratio interpolations were carried out using Surfer® (Golden Software) and the kriging method selected was ordinary kriging, ('Grid Method 1').

To test the performance of each calibration method, one hundred random combinations of monitoring sites (including repetitions where necessary) were generated for each calibration scenario and subsequently used to conduct one hundred separate calibrations. The random selection of monitoring site combinations means that any repetition will not skew the subsequent analysis. The remaining monitoring sites were then used to compare the calibrated modelled concentrations to the remaining measured concentrations. For example, there would be 400 randomly generated modelled vs. measured comparisons for a calibration scenario which used twenty-four monitoring points with which to conduct the calibration (i.e. one hundred times the four remaining monitoring points).

The randomly generated comparisons were then analysed using standard model evaluation statistics: Normalised Mean Square Error (Equation 7, 8), Mean Bias (Equation 9) and the percentage of modelled predictions within a factor of *x* (F*x*) of measured predictions (e.g. F1.2, F1.5, F2, F5 and F10). The F*x* statistic is expressed as a value from 0 to 1. The F2 statistic will be the only F*x* statistic discussed in this report but all summary statistics may be found in Appendix B.

$$MSE = \frac{\sum \left(conc_{mod} - conc_{meas} \right)^2}{n} \qquad \textbf{Equation 7}$$

$$NMSE = \frac{MSE}{\left(\overline{mod} \times \overline{meas} \right)} \qquad \textbf{Equation 8}$$

$$MB = \overline{mod} \div \overline{meas} \qquad \textbf{Equation 9}$$

7.2 Results

7.2.1 Kincaid data

Figures 7.6 (a, b and c) and 7.7 (a, b and c) show summary statistics (NMSE, F2 and MB) for the Kincaid data set, generated by comparing measured vs. calibrated model values for the mean and 99.9th percentile concentration metrics, respectively. Shown

below each diagram, in bold, is the average summary statistic (NMSE, F2 or MB) generated by comparing measured vs. uncalibrated model values.

For the mean concentrations, all summary statistics show that by using sector-corrected measured data, all four calibration methods have improved the consistency between modelled and measured concentrations when compared with uncalibrated modelled values. In addition, Figures 7.6 (a) and (b) indicate that as the number of data points used to perform the calibration increases, the F2 statistic increases, with the exception of the kriging ratio method in Figure 7.6 (b) which maintains a relatively constant F2 value of 0.93-0.96.

For the 99.9th percentile metric, shown in Figures 7.7 (a), (b) and (c), the summary statistics produced using the calibrated modelled data are very similar to those produced using the uncalibrated modelled data. However, it is clear that below a defined number of monitoring points, the NMSE rapidly increases and the F2 rapidly decreases as the calibration is based on too few data points, especially for the linear regression method. The number of data points that are required for optimum model calibration appears to fall between 10 and 15 for all calibration methods for both mean and 99.9th percentile metrics. This is based on the interpretation of Figures 7.6 and 7.7 (a) and (b), which show stabilising summary statistics around this point. However, the relationship with the number of data points used for the calibration and the mean bias (MB) statistic is complex and does not allow the same conclusions to be drawn. It is also important to remember that this analysis is based on an example with a single pollution source and for only one pollutant. For multiple-pollution sources in close proximity, the number of monitoring sites required for optimum calibration is unlikely to be linearly related to the number of pollution sources.

Overall, the simple ratio method appears to perform best as this method constantly produces the lowest NMSE and the highest F2 statistics, regardless of the number of data points used to conduct the calibration. This is surprising as the kriging method offers a spatially varying calibration. Comparing only the two kriging methods, Figure 7.8 shows that the kriging ratio method is superior to the kriging residual method, as the kriging ratio method minimises the NMSE and maximises the F2 statistic for both the mean and the percentile metrics.

7.2.2 Aire Valley data

Unlike the Kincaid data set, not all of the calibration methods improved the consistency between modelled and measured concentrations. Figures 7.6 (e) and (f) show that the linear regression method performs particularly poorly, especially when only 2 data points are used to perform the calibration.

Another inconsistency in the Aire Valley results is shown in Figure 7.6 (f) where for all calibration methods, the mean bias is lower using the uncalibrated rather than the calibrated modelled data.

The NMSE statistic in Figure 7.6 (d) indicates that the simple ratio method is the best performing calibration method when applied to the mean concentration metric using the Aire Valley data. This also appears to be the case for the 99.9th percentile metric, shown in Figure 7.7 (d), although for 4 and 5 data points, the linear regression method has an even lower NMSE.

It is difficult to tell how the kriging methods perform with the Aire Valley data set, as a minimum of three data points is needed to produce a kriged surface and again the maximum number of data points used for the calibration is limited to five. This is a consistent problem with this data set and limits any comparisons with the Kincaid data set.

For most calibration methods, NMSE and MB statistics show that as the number of data points used to perform the calibration increases, the calibration performance is improved. However, Figure 7.6 (e) illustrates that this is not the case for the linear regression method when used to calculate the annual average concentration. This may be because an outlier is affecting the linear regression properties and therefore producing a poor calibration. For example, the data in Table 7.1 shows that there is a significant difference between mean modelled and mean measured concentrations at West Bank and North Featherstone monitoring sites, possibly due to the influence of background sources.

The analysis of the Aire Valley data identified that calibration was optimised using five monitoring points (the maximum number available). Although Figures 7.6 (d) and (e) seem to show stabilising summary statistics for the simple ratio method, it is difficult to ascertain whether increasing the number of monitoring sites, beyond the six already established, would further improve the performance of the calibration methods.

Table 7.1 Aire Valley SO_2 concentrations (in µg m^{-3}) measured in 2003. Measured data sector corrected to 60-degree sectors.

Monitoring site	Mean Modelled	Mean Measured	99.9th%ile Modelled	99.9th%ile Measured
CL	3.8	3.1	189	108
DG	2.4	2.7	158	106
HL	3.1	3.2	111	112
NF	0.8	1.3	131	112
SM	3.2	3.5	140	142
WB	2.8	3.7	146	165

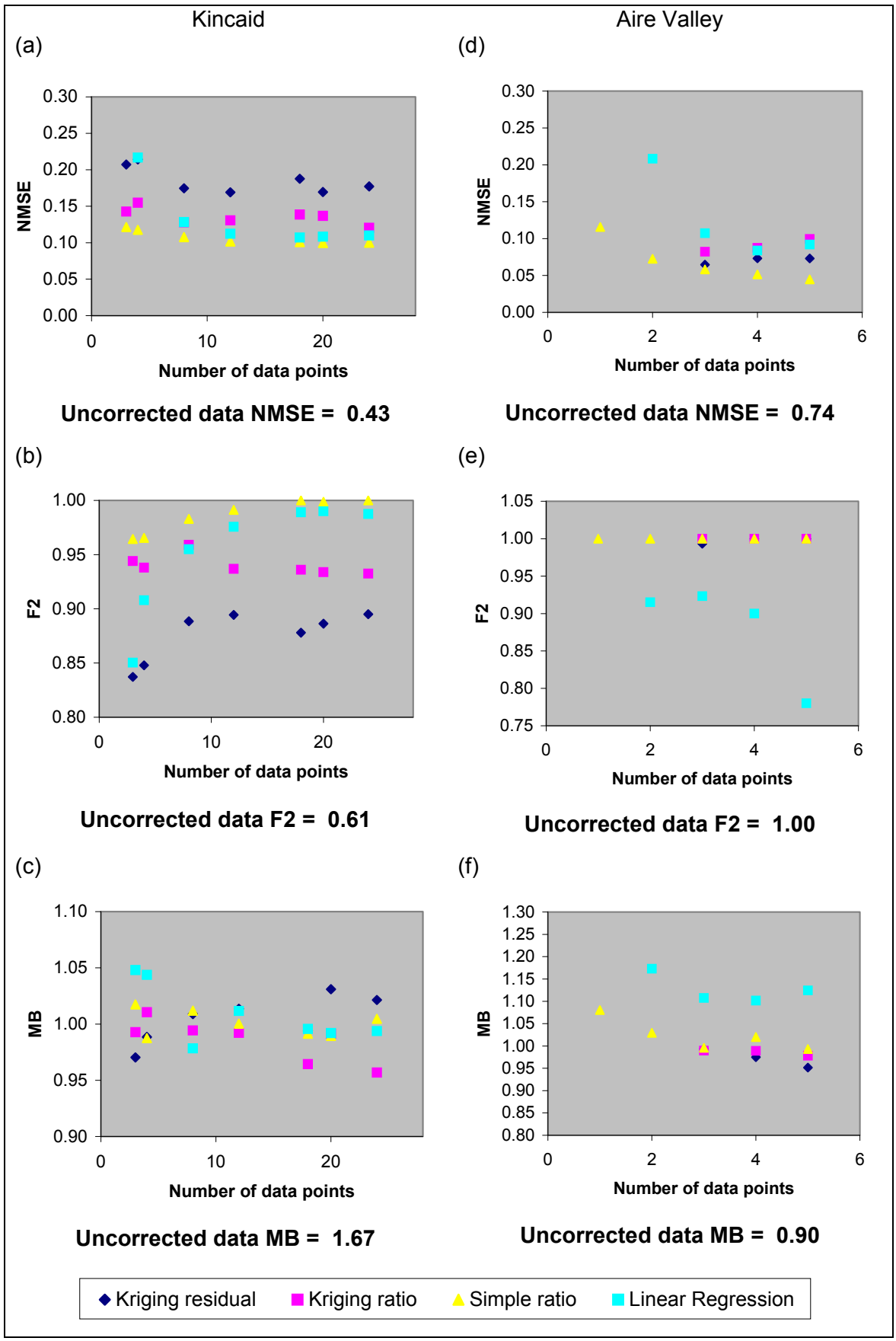


Figure 7.6 Statistical analysis of calibration method performance when applied to annual average concentrations for the Kincaid (1980/81) and Aire Valley (2003) data sets.

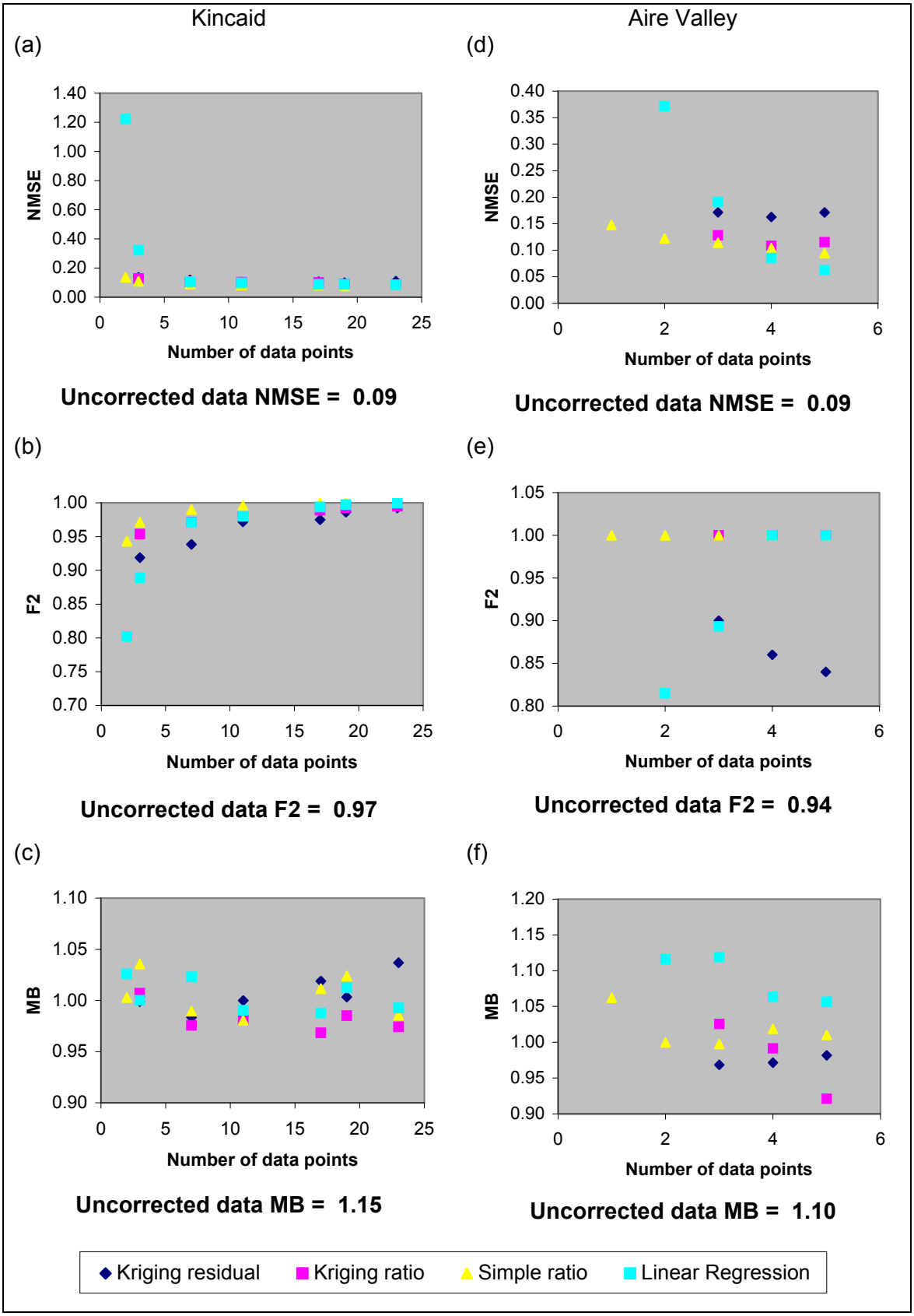


Figure 7.7 Statistical analysis of calibration method performance when applied to 99.9th percentile concentrations for the Kincaid (1980/81) and Aire Valley (2003) data sets.

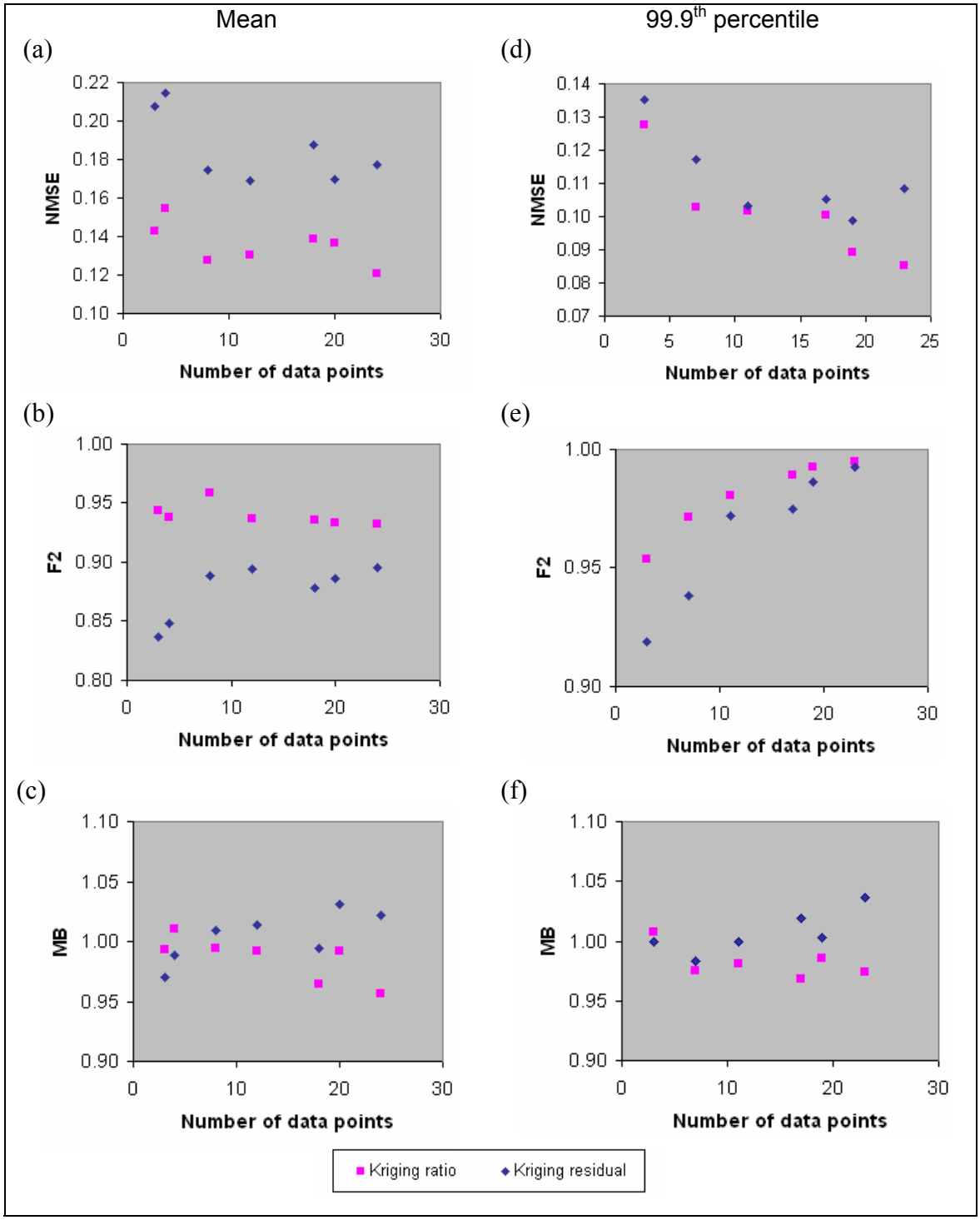


Figure 7.8 Statistical analysis of kriging calibration methods when applied to 99.9th percentile concentrations for the Kincaid (1980/81) data set.

7.3 Discussion

Four calibration methods have been analysed to determine their effectiveness in calibrating modelled predictions of air quality impacts from industrial point sources. The simple ratio method proved best using both the Kincaid data set and the Aire Valley data set. The linear regression method performed on a similar level to the simple ratio method when using a high number of data points but the performance of this method declined dramatically with the use of only two monitoring points. This is due to the poor definition of the linear regression when using a low number of data points.

The more complex kriging methods appear to be less effective than the simple methods, although this may be because the basic principle of kriging is undermined by the data used. Kriging revolves around the assumption of spatial correlation. If this spatial correlation is removed or is diluted, for example by kriging the residual or ratio rather than the original measured values, this may reduce the spatial correlation in the data such that the technique is unable to interpolate effectively between points of known value. The other potential problem with the kriging technique is the lack of monitoring points. It is likely that the spatial variation in the concentration field may be much larger than can be accounted for by the density of monitoring points and therefore this method may not be optimised by the number of available data points. When this occurs, ordinary kriging tends to return a value close to the global mean, which would therefore produce similar results to the simple ratio or linear regression methods.

The analysis of the Kincaid data set suggests that the number of monitoring points required to enable optimum calibration is between 10 and 15. The integration of data using more than 15 monitoring points was found not to provide further reductions in cross-validation error terms, possibly due to the effect of irreducible random uncertainties. Statistical optimisation of model calibration is only one consideration in determining the number of monitoring sites for regulatory purposes. Nonetheless, if the costs of monitoring decline through developments in low cost, high-resolution sensor technologies, these larger data fields may become achievable.

A major issue for both study sites seems to be how to correct modelled concentrations for influences external to the model domain. In the case of the Aire Valley, prospective modelling assessments conducted by the JEP are corrected for background concentrations by applying a correction equal to the monitored annual mean minus the contribution from modelled sources (Equation 3). However, ‘background’ concentrations in the 2003 Aire Valley data set may be partially attributed to the influence of point sources outside the model domain, e.g. the Trent/Soar Valley power stations to the South East of the Aire Valley. While the Trent/Soar stations are now unlikely to cause a significant problem regarding compliance with air quality standards in the Aire Valley, the point should be made that ‘external’ sources do not always contribute a relatively steady background to monitored concentrations. The temporal variation of ‘external’ point sources could have a significant impact on compliance with air quality objectives if the sources are of sufficient size and located upwind in the line of frequent wind directions at a distance where impacts could be noted.

If the influence of external point sources is found to be significant, it may be necessary to adopt the approach of sector-correcting measured concentrations rather than applying a constant background adjustment for modelled concentrations, especially if the contribution of local sources is to be isolated for regulatory purposes. However, to assess the total impact upon receptors, it would still be necessary for the model to take account of all sources, including those outside the model domain. In this case, the application of a time-varying background contribution may be a suitable solution and would be more realistic than a constant background correction, although this is rarely

feasible and would only be considered if air quality standards were significantly threatened.

8 Recommendations

The following are recommendations based on the review of modelling and monitoring for integration in air quality assessment and the findings from the investigative work in Sections 6 and 7 on the application of integration methods for the extrapolation of short-term monitoring data and model calibration.

8.1 Towards more effective integration of monitoring and modelling data

The JEP Risk Management Framework provides a working example of integration of modelling and monitoring procedures for industrial sources in the UK. Approaches using the integration of monitoring and modelling data could also be used for other Part A(1) sources, especially where there is evidence of frequent exceedences or where modelling studies are subject to a high degree of uncertainty, e.g. for sites located in coastal areas or areas of complex terrain. Though the exact specification of integration methods should be decided on a case-by-case basis due to the uncertainties highlighted herein.

8.2 Expansion of the monitoring infrastructure

The application of the integration approaches reviewed in this document, in particular the more complex techniques such as kriging and Bayesian assimilation, requires greater investment in monitoring infrastructure. Existing monitoring networks such as the AURN network are of limited use as they are often located far from industrial sources and the potential to attribute pollutant concentrations to such sources may be difficult due to the interference of pollutants from roads and urban areas. However, AURN sites classified as ‘suburban’, ‘urban background’ or ‘rural’ may be useful for defining background influences and should not be dismissed. While practical factors such as prohibitive costs limit the expansion of the monitoring infrastructure at present, developments in low cost, high-resolution sensor technologies may make this more achievable in the future.

8.3 Optimisation of the monitoring infrastructure

It is essential that any monitoring resource must be optimised in terms of the number and placement of monitors, both to ensure cost-effectiveness and to optimise the potential for effective integration of monitoring and modelling data. It may be necessary to first identify where uncertainty in monitoring data is most prevalent, for example: in sampling; analysis; temporal or spatial resolution. Techniques similar to those described by Bortnick and Stetzer (2002) could be applied. Once such uncertainties have been identified, the adoption of decision-making tools, which are able to prioritise multiple objectives such as population protection, ecosystem protection or targeting specific air quality metrics, for example the maximum 99.9th percentile concentration, are invaluable in designing an effective monitoring network.

If the number of monitoring sites around Part A(1) processes remains limited, the adoption of a flexible approach to off-site monitoring by both the Environment Agency (for MMF campaigns) and process operators is necessary as in these situations it will

be impossible to optimise a network for source attribution to distinguish between individual sources. For example, where several large industrial processes are located within close proximity to one another, operators could form a consortium to establish a jointly-funded ambient monitoring network. However, a network for an industrial zone cannot expect to have sites at all individual maxima; instead it would be more advantageous to have sites where there is good source attribution. These factors may also be incorporated into a multi-objective analysis procedure.

For Part A(1) sources located in areas of complex terrain, terrain grids should be included in dispersion model configurations when slope is greater than 1:10. Given the effect of complex terrain on dispersion, the provision of extra monitoring sites should be considered in these areas as this may allow model uncertainty to be better defined.

8.4 Determination of effective calibration methods

The exploration of data assimilation methods in this report followed on from the work of Denby *et al*. (2007) by applying several simple and complex offline methods to calibrate point source modelling data. Four offline data assimilation methods were applied to the Kincaid and Aire Valley data sets and cross-validation analysis was conducted in order to determine the number of monitoring points required for optimum model calibration.

The simple ratio method performed best using both the Kincaid and the Aire Valley data set. The linear regression method performed to a similar level when using a high number of data points although the performance of this method declined dramatically with the use of only two monitoring points. The more complex kriging methods appeared to be less effective than the simple methods, despite offering a spatially varying model calibration. However, ordinary kriging was the only kriging option used in the analysis of the data assimilation methods. Therefore, it may be useful to test the kriging of residual and kriging of ratio methods using other kriging options, e.g. kriging-with-a-trend, as this option is more suited to trended data, i.e. the deposition pattern formed by an idealised Gaussian Plume.

The analysis suggests that the number of monitoring points required for optimum model calibration using the techniques analysed is between 10 and 15. The integration of data using more than 15 monitoring points was found not to provide further reductions in cross-validation error terms, possibly due to the effect of irreducible random uncertainties. Statistical optimisation of model calibration is only one consideration in determining the number of monitoring sites for regulatory purposes. Nonetheless, if the costs of monitoring decline through developments in low cost, high-resolution sensor technologies, these larger data fields may become achievable.

The analysis of the Kincaid data sets determined the number of monitoring sites necessary for optimum data assimilation. However, monitoring sites in the Kincaid data set are not organised according to optimal network design priorities. In order to determine both the correct number and arrangement of monitoring sites, it would be necessary to conduct a validation study that considered both aspects of network design. Initially, optimal placement of sites could be determined using multi-objective analysis applied to *modelled* concentration data, as this would enable the use of a large data set with minimal cost. This would determine the optimal locations for a varying number of monitoring sites. Each optimal monitoring arrangement could then be analysed to determine the number of monitoring sites necessary for optimal performance of a selected model calibration technique. However, this would require the use of real monitoring data. Emerging technologies in monitoring instrument design may enable investment into such validation studies. In addition, it is likely that the optimal number of monitoring sites will differ for different pollutants and different

sources, for example: point vs. area sources, such as landfill sites, or point sources with different stack heights. Further work is therefore necessary to determine the variation in the number of required monitoring sites according to these factors.

8.5 Application of suitable background correction factors

It is important to bear in mind that for the analysis of the integration methods the measurement data was sector-corrected to allow an unbiased analysis of the data assimilation methods by removing any influence from external sources that were not incorporated in the modelled scenario. The uncertainty introduced by the influence of sources external to the model domain is normally accounted for by using a background correction factor. If the external sources are relatively homogenous and widespread then a single correction can be applied to the model to correct for any background influence before calibration is carried out. However, in the case of the Kincaid and Aire Valley data sets, the background source can be identified as originating from a specific direction and seems to vary in magnitude temporally. In cases such as these, it may be necessary to consider the influence of background sources in more detail. This could include modelling the external source specifically or through conducting an analysis of background measurement data. However, an appropriate background correction may be difficult to apply if the calibration is intended for application to prospective modelling studies as opposed to retrospective model calibration for validation purposes. In this case, background correction factors from previous years could be examined and the most representative correction factor applied to the prospective model data prior to calibration. Further work is necessary to determine how effective retrospective calibration factors would be in adjusting prospective modelling assessments and what may be defined a ‘representative’ correction factor.

8.6 Extrapolating from short-term monitoring campaigns

From the analysis of the pro-rata extrapolation method, it is evident that extrapolating from a 6-month campaign provides a more accurate estimate of total exceedences than extrapolating from a 3-month campaign.

The Integration Scaling Method (ISM), an alternative to the pro-rata method, was analysed using the same routine used to test the pro-rata method. The ISM relies on the determination of an appropriate scaling factor, determined as the ratio of modelled exceedences for the duration of the monitoring period to modelled exceedences for the full 12-month period. The ratio of modelled 6-month to 12-month exceedences is then used to scale the short-term measurement data to obtain an estimate of the number of exceedences during a 12-month period. Although the ISM showed an improvement in the consistency of measured and predicted exceedences compared to the pro-rata method, the RMSE values were still relatively high. It would be beneficial to repeat this analysis with modelled concentration data that are more representative of the measured concentrations, i.e. modelled data which have been corrected for the influence of background concentrations and sources outside the model domain.

In contrast to the ISM, the multiple regression scaling method aimed to determine whether a multiple regression model could be used to relate SO_2 emissions and meteorological parameters in order to predict concentrations of SO_2 and therefore extrapolate the frequency of exceedences beyond those measured during short-term monitoring campaigns. However, the complexity of the controls on monitored

concentrations and the constraints of the assumptions of the multiple regression method resulted in very low percentage of explained variance (0.2-17%) so it was considered inappropriate to conduct further analysis of the application of this method.

References

ARBELOA, F.J.S., CASEIRAS, C.P. AND ANDRES, P.M.L., 1993. Air quality monitoring: optimisation of a network around a hypothetical potash plant in open countryside. Atmospheric Environment, 27A (5), 729-738.

AULD, V., HILL, R. AND TAYLOR, T.J., 2003. Uncertainty in deriving dispersion parameters from meteorological data. Westlakes Scientific Consulting report for the UK Atmospheric Dispersion Modelling Liaison Committee, ADMLC/2002/2.

ARGYRAKI, A. KINNERSLEY, R. AND RAMSEY, M.H., 1999. Optimising emergency sampling strategy following a release of radioactive material into the environment. MAFF Research Contract Final Report, Imperial College London.

BETHAN, S., 2006. Predicted and Measured SO_2 Concentrations Presented in AQMP Annual Reviews 2001-2004. Power Technology JEP Report, PT/05/BE1121/R.

BETHAN, S. AND TEASDALE, I., 2005. Variations in annual meteorology. Power Technology JEP Report, PT/04/BE1014/R.

BLOND, N., BEL, L. AND VAUTARD, R., 2003. Three-dimensional ozone data analysis with an air quality model over the Paris area. Journal of Geophysical Research, 108 (D23), 4744.

BORTNICK, S. M. AND STETZER, S. L., 2002. Sources of variability in ambient air toxics monitoring data. Atmospheric Environment, 36: 1783-1791.

BROOKE, D., LUCAS, V. AND BETHAN, S., 2003. A comparison of AERMOD with ADMS 3.1 and measured data from Yorkshire and the Lower Trent Valley in 1998 and 1999. RWE Innogy JEP Report, ENV/EEA/110/03.

BROOKE, D. AND STIFF, S., 2006. A comparison of results from ADMS 3.1, ADMS 3.3, and AERMOD-PRIME with measured data. RWE NPower JEP Report, PT/06/BE1149/R, ENV/136/2006.

CARRUTHERS, D.J., HOLROYD, R.J., HUNT, J.C.R., WENG, W.S., ROBINS, A.G. APSLEY, D.D., THOMPSON, D.J. AND SMITH, F.B., 1994. A new approach to modelling dispersion on the earth's surface. Journal of Wind Engineering Industry Aerodynamics, 52, 139-153.

CERC, 2007. ADMS 4 Users guide. Cambridge Environmental Research Consultants, Cambridge UK, (www.cerc.co.uk).

CIMORELLI A.J., PERRY S.G., VENKATRAM A, WEIL J.C., PAINE R.J., WILSON R.B., LEE R.F. AND PETERS W.D., 1998. AERMOD, description of model formulation. US Environmental Protection Agency, Research Triangle Park, NC.

CLARK, R.H., 1979. A model for short and medium range dispersion of radionuclides released into the atmosphere. Chilton, NRPB-R91.

CURCIN, V., GHANEM, M., GUO, Y., KŐHLER, M., ROWE, A., SYED, J. AND WENDEL, P., 2002. Discovery Net: towards a grid knowledge discovery. In: Proceedings of KDD-2002. The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

DAVOINE, X. AND BOCQUET, M., 2007. Inverse modelling-based reconstruction of the Chernobyl source term available for long-range transport. Atmospheric Chemistry and Physics, 7, 1549-1564.

DENBY, B., POCHMAN, M, MACOU, J, DUDEK, A. AND ELSHOUT, S.V.D., 2007. Individual Case Study Report 6, Basic data assimilation: Application to the urban scale. Air4EU report. Available from: http://www.air4eu.nl/reports_products.html [Accessed 13 August 2007].

DEFRA, 2003. Local Air Quality Management Technical Guidance LAQM TG(03). Department for Environment, Food and Rural Affairs Publication, PB7514.

DEFRA, 2007. The Air Quality Strategy for England, Scotland, Wales and Northern Ireland. Department for Environment, Food and Rural Affairs Publication, Cm 7169 NIA 61/06-07.

DEUTSCH, C.V. AND JOURNEL, A.G., 1998. GSLIB Geostatistical Software Library and User's Guide. Oxford University Press.

DORE, A.J., VIENO, M., BEALEY, W.J. AND SUTTON, M.A., 2005. Source Attribution, Report as part II of the project: Source Attribution and critical loads assessment for Special Areas of Conservation and Special Protection Areas in the UK. Centre for Ecology and Hydrology (Edinburgh Research Station), Bush Estate, Penicuik, Midlothian, EH26 0QB, UK.

ELSOM, D.M., 1978. Spatial correlation analysis of air pollution data in an urban area. Atmospheric Environment, 12, 1103-1107.

ENVIRONMENT AGENCY, 2000a. Environmental Monitoring Strategy – Ambient Air. Technical Guidance Note M8, ISBN 0113101759.

ENVIRONMENT AGENCY, 2002a. Integrated Pollution Prevention and Control (IPPC) Environmental Assessment and Appraisal of BAT. Horizontal Guidance Note IPPC H1, ISBN 0113101082.

ENVIRONMENT AGENCY, 2002b. An intercomparison of AERMOD, AERMOD-PRIME and ADMS 3.1. Prepared by Bethan Tuckett-Jones, Air Quality Modelling and Assessment Unit, Version 1 dated 8th May 2002.

ENVIRONMENT AGENCY, 2003. Integrated Pollution Prevention and Control (IPPC) Environmental Assessment and Appraisal of BAT, IPPC H1 Horizontal Guidance Note.

ENVIRONMENT AGENCY, 2006. Use of continuous isokinetic samplers for the measurement of dioxins and furans in emissions to the atmosphere. April 2006, ISBN 1844325504.

ENVIRONMENTAL PROTECTION AGENCY, 2004. Users Guide to the AERMOD meteorological preprocessor (AERMET). U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Emissions, Monitoring, and Analysis Division, Research Triangle Park, North Carolina 27711, EPA-454/B-03-002.

FINARDI, S., MORSELLI, .G. AND JEANNET, P., 1997. Wind flow models over complex terrain for dispersion calculations. Cost Action 710 Pre-processing of meteorological data for dispersion models. Report of Working Group 4.

FLESCH, T. K., WILSON J.D., HARPER L.A. AND CRENNA B.P., 2005. Estimating gas emissions from a farm with an inverse-dispersion technique, Atmospheric Environment 39, 4863-4974

FUTTER, D.N., 2000. A comparison between ADMS 3 and Breeze AERMOD Suite, Power Technology JEP Report, PT/00/EA386/R.

GASS, H.C., LUDER, K. AND WILKEN, M., 2002. PCDD/F emissions during cold start-up and shut-down of a municipal waste incinerator. Organohalogen Compounds. 56, 193-196.

GHIL, M., HALEM, M. AND ATLAS, R., 1979. Time continuous assimilation of remote sounding data and its effect on weather forecasting. Monthly Weather Review, 107(2): 140-171.

GRAVES, R.J., MCGINNIS Jr., L.F. AND LEE, T.D., 1981. Air monitoring network design: case study. Journal of ASCE Environmental Engineering Division, 107 (EE5), 941-955.

HANDSCOMBE, C.M. AND ELSOM, D.M., 1982. Rationalisation of the national survey of air pollution monitoring network of the United Kingdom using spatial correlation analysis: a case-study of the Greater London area. Atmospheric Environment, 16, 151-163.

HIGGINS, N.A., BURGE, F., CHARNOCK, T.W. AND TEALE, P., 2005. Statistical Estimation and Characterisation Techniques for use during Accident Response (SECTAR). Health Protection Agency (Radiation Protection Division) Report, HPA-RPD-009.

HILL, R.A., LUTMAN, E.R. AND ARNOTT, A.D., 2005. A review of atmospheric dispersion in complex terrain. Westlakes Scientific Consulting report for the UK Atmospheric Dispersion Modelling Liaison Committee, AMDLC/2005/1.

HILL R.A., SMITH K., RUSSELL K., MISSELBROOK T. AND BROOKMAN S., 2007. Emissions of ammonia from weeping wall stores and earth banked lagoons determined using passive sampling and atmospheric dispersion modelling. Journal of Atmospheric Chemistry. Available from: http://www.springerlink.com/content/c413t4m274j0331l/ [Accessed 15 October 2007].

HUNTER, G.C., 2004. A Risk Management Framework for the Air Quality Management Plan process. JEP Report, ENV/EEA/136/04/2004.

HUNTER, B. AND BETHAN, S., 2005. Trial of the exceedence evaluation system used in the Risk Management Framework for the Air Quality Management Plan Process. RWE npower JEP Report, ENV/EEA/167/2004.

HUNTER, G.C., 2006. Review of AQMP modelling procedures: Start-up and standby operation. RWE Power JEP Report, ENV/156/2006.

JEP, 2002. Appendix B: Methodology for the use and interpretation of monitoring and modelling for AQS management plans. Issue 5, January 2002.

KAHN, F.I. AND SADIQ R., 2005. Risk-based prioritisation of air pollution monitoring using a fuzzy-synthetic evaluation technique. Environmental Monitoring and Assessment, 105, (1-3), 261-283.

KAO, J. AND HSIEH, M., 2006. Utilizing multi-objective analysis to determine an air quality monitoring network in an industrial district. Atmospheric Environment, 40, 1092-1103.

KIDD, J., 2002. Comparison of NWP gridded weather data and site specific observed data for use in dispersion models. MSc Thesis for Applied Meteorology, University of Reading.

LANGSTAFF, J., SEIGNEUR, C. AND LUI, M.K., 1987. Design of an optimum air monitoring network for exposure assessments. Atmospheric Environment, 21 (6), 1393-1410.

LOTHGREN, C.J. AND VAN BAVEL, B., 2005. Dioxin emissions after installation of a polishing wet scrubber in a hazardous waste incineration facility. Chemosphere, 61, 405-412.

LUCAS, V. AND BETHAN, S., 2004. NWP parameters for use in ADMS 3.1 dispersion modelling. RWE Innogy JEP Report, ENV/EEA/147/04.

MCINNES, K.J., KISSEL D.E. AND KANEMASU, E.T., 1985. Estimating ammonia flux: a comparison between the integrated horizontal flux method and theoretical solutions of the diffusion profile. Agronomy Journal 77, 884-889.

MODAK, P.M., 1985. Optimization of ambient air quality monitoring networks (part I-III). Environmental Monitoring and Assessment, 5, 1-19.

NAKAMORI, Y., IKEDA, S. AND SAWARAGI, Y., 1979. Design of air pollutant monitoring system by spatial sample stratification. Atmospheric Environment, 13, 97-103.

NELSON, N., MIRZA, A.K. AND WEAVER, K.N., 2002. An assessment of alternative sources of met data for use in dispersion modelling. Met Office report for the UK Atmospheric Dispersion Modelling Liaison Committee, AMDLC/2002/1.

NOLL, K.E., MILLER, T.L., NORCO, J.E. AND RAUFER, R.K., 1977. An objective air monitoring site selection methodology for large point sources. Atmospheric Environment, 11, 1051-1059.

POLITIS, K. AND ROBERTSON, L., 2003. Bayesian updating of atmospheric dispersion after a nuclear accident. Applied Statistics, 53 (4), 583-600.

RABIER, F., 2005. Overview of data assimilation methods in numerical weather prediction centres. Quarterly Journal of the Royal Meteorological Society, 131 (613), 3215-3233, Part C.

RICHARDS, M., GHANEM, M., OSMOND, M., GUO, Y. AND HASSARD, J., 2006. Grid-based analysis of air pollution data. Ecological Modelling, 194, 274-286.

SCIRE, J.S., STRIMAITIS, D.G., YAMARTINO, R.J., 2000. A user's guide for the CALPUFF dispersion model (Version 5). Earth Tech Inc., 196 Baker Avenue, Concord, MA 01742. Available from: http://www.scr.com/calpuff/calpuff1.htm. [Accessed 19 November 2007].

SEARLE, S.R., 1971. Linear Models. Wiley, New York, pp. 376-493.

SEARLE, S.R., CASELLA, G. AND McCULLOCH, C.E., 1992. Variance components. Wiley, New York.

SHI, N.J. AND NG, B., 2002. Risk based pragmatic approach to address model uncertainty. Paper for the NSCA DMUG meeting, November 2002.

SHUTT, M. SHEPPARD, V. AND HICKEY, G., 2005. Study of Ambient Air Quality at Pen-y-fford, 30 June 2005 to 26 September 2005, Main Report. Environment Agency Report MAPG/TR/2005/006.

SHEPPARD, V., JONES, N. AND SHUTT, M., 2001. Study of Ambient Air Quality At Aberthaw, 9 August 2000 to 9 January 2001, Main Report. Environment Agency Report, NCAS/TR/2001/011.

SIDLE, C., TUCKETT-JONES, B., NG, B. AND SHI, J. P., 2004. Model intercomparison between ADMS 3.1, AERMOD and AERMOD PRIME. Proceedings of the 9th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Garmisch-Partenkirchen, Germany, pp 156-160.

SMITH, A. D., JONES S. R., GRAY J. AND MITCHELL K. A., 2007. A review of irradiated fuel particle releases from the Windscale Piles, 1950-1957. Journal of Radiological Protection, 27, 115-145.

STEWART, H., TIMMIS, R. AND WILSON, P., 2004 Using monitoring and meteorological analyses to assess the impact on air quality of SO_2 from power stations. Environment Agency Presentation, Dispersion Modelling User Group Meeting, 20th May 2004.

SWINBANK, R. AND O'NEILL, A., 1994. A stratosphere-troposphere data assimilation system. Monthly Weather Review, 122 (4), 686-702.

THE, J.L., LEE. R. AND BRODE, R.W., 2001. Worldwide Data Quality Effects on PBL Short-Range Regulatory Air Dispersion Models. Proceedings of the 7th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Belgirate, Italy, pp 202-206.

TRUJILLO-VENTURA, A. AND ELLIS, J.H., 1991. Multi-objective air pollution monitoring network design. Atmospheric Environment, 11, 1051-1059.

WALKER, S.E., SCHAAP, M. AND SLINI, L., 2006. Data Assimilation. WP 6 – Synthesis, Cross-cutting Issue Milestone Report 6.8 (M.5). Air4EU report. Available from: http://www.air4eu.nl/reports_products.html [Accessed 13 August 2007].

WANG L., HSING-CHENG, H., JUU-EN, C., XING-YI, Y., GUO-PING, C. AND WEI-SHAN, L., 2007. Influence of start-up on PCDD/F emissions of incinerators. Chemosphere 67, 1346-1353.

WEBB, A., 2005. Managing power station ambient air quality compliance – an overview of the Air Quality Management Plan process. RWE Power International JEP Report, ENV/EEA169/04.

WEBB, A., 2004. A methodology for the assessment of post-facto air quality compliance for use in AQMP Annual Reviews. RWE Power International JEP Report, ENV/EEA/154/2004.

WEBSTER, R. AND OLIVER, M. A., 2001. Geostatistics for Environmental Scientists. John Wiley and Sons, England.

WU, H.W.Y. AND CHAN, L.Y., 1997. Comparative study of air quality surveillance networks in Hong Kong. Atmospheric Environment, 31, 935-945.

ZHENG, D.Q., LEUNG, J.K.C, LEE, B.Y. AND LAM, H.Y., 2007. Data assimilation in the atmospheric dispersion model for nuclear accident assessments. Atmospheric Environment, 41, 2438-2446.

European Directives

Integrated Pollution Prevention and Control Directive (1996/61/EC)

National Emissions Ceiling Directive (2001/81/EC)

Large Combustion Plant Directive (2001/80/EC)

Sulphur Content of Liquid Fuel Directive (1999/32/EC)

Solvents Directive (1999/13/EC)

Large combustion plant directive (2001/80/EC)

Air Quality Framework Directive (1996/62/EC)

Air Quality Framework Daughter Directive 1 (1999/30/EC)

Air Quality Framework Daughter Directive 2 (2000/69/EC)

Air Quality Framework Daughter Directive 3 (2002/3/EC)

Air Quality Framework Daughter Directive 4 (2004/107/EC)

Air Quality Standards Regulations (2007) Available from: http://www.opsi.gov.uk [Accessed 04 June 2007]

The Pollution Prevention and Control (England and Wales) Regulations (2000) Available from: http://www.opsi.gov.uk [Accessed 04 June 2007]

Acts

Clean Air Act (1952)

Environmental Protection Act (1990)

Pollution Prevention Control Act (1999)

Appendices

Appendix A

1. Unfiltered frequency distribution for measured data – 6 month rolling, all sites

Start month	Concentration bin					
	<40	40-80	80-120	120-160	160-200	>200
2003_1	25349	489	134	58	14	14
2003_2	25366	495	123	53	13	14
2003_3	25779	490	137	58	14	18
2003_4	25698	430	136	51	15	22
2003_5	25884	399	135	45	12	21
2003_6	25704	414	149	51	11	23
2003_7	25943	357	129	37	9	21
2003_8	25987	331	115	34	9	20
2003_9	25792	294	80	25	9	8
2003_10	25956	293	72	18	8	5
2003_11	25793	316	69	14	9	7
2003_12	25963	303	63	8	10	5
2004_1	25822	297	66	8	10	5
2004_2	25740	354	85	12	12	5
2004_3	25942	393	109	25	15	12
2004_4	25770	401	123	30	18	10
2004_5	25979	354	109	29	17	8
2004_6	25866	336	98	27	17	8
2004_7	25997	344	98	30	19	8
2004_8	26108	252	82	28	18	8
2004_9	25802	184	52	14	11	1
2004_10	26045	132	21	6	4	0
2004_11	25878	151	23	8	4	0
2004_12	26032	144	21	8	3	0
2005_1	25889	141	25	7	2	0
2005_2	25851	170	31	11	1	0
2005_3	26265	187	32	11	1	0
2005_4	26116	190	34	11	1	0
2005_5	26297	160	29	9	1	0
2005_6	26159	156	27	9	1	0
2005_7	26208	118	19	7	0	0
Average	**25903**	**293**	**78**	**24**	**9**	**8**

2. Unfiltered frequency distribution for modelled data – 6 month rolling, all sites

Start month	Concentration bin					
	<40	40-80	80-120	120-160	160-200	>200
2003_1	29746	364	127	53	14	8
2003_2	29891	397	158	60	11	5
2003_3	30278	449	177	76	13	3
2003_4	30404	445	172	71	16	2
2003_5	30631	450	184	72	17	2
2003_6	30720	359	168	72	17	2
2003_7	31372	268	117	46	13	2
2003_8	31279	181	80	34	14	2
2003_9	30458	157	72	20	11	2
2003_10	30031	184	75	18	7	3
2003_11	30191	184	63	24	9	3
2003_12	30358	251	92	46	12	3
2004_1	29674	361	158	60	14	3
2004_2	30049	457	213	107	31	7
2004_3	30759	461	232	108	34	8
2004_4	30519	473	242	114	33	5
2004_5	30358	444	228	105	30	5
2004_6	30361	395	203	83	27	5
2004_7	30721	297	136	70	25	5
2004_8	30143	221	82	20	7	1
2004_9	29394	187	53	16	2	0
2004_10	29528	149	37	4	0	0
2004_11	29363	231	63	9	4	0
2004_12	29275	265	71	13	4	0
2005_1	29057	323	93	18	4	1
2005_2	29554	355	114	24	6	1
2005_3	30584	398	127	30	6	1
2005_4	30764	403	133	31	6	1
2005_5	31215	315	103	26	2	1
2005_6	31422	254	89	20	2	1
2005_7	31626	176	66	14	2	0
Average	**30315**	**318**	**127**	**47**	**13**	**3**

3. Filtered frequency distribution for measured data – 6 month rolling, all sites

Start month	Concentration bin				
	40-80	80-120	120-160	160-200	>200
2003_1	565	111	43	22	13
2003_2	550	106	41	21	12
2003_3	545	114	46	24	14
2003_4	464	109	47	26	16
2003_5	440	100	42	23	14
2003_6	470	106	45	24	15
2003_7	395	89	37	20	12
2003_8	357	82	35	19	12
2003_9	328	63	24	12	7
2003_10	333	54	19	9	5
2003_11	315	56	20	10	6
2003_12	293	48	16	8	4
2004_1	293	48	17	8	4
2004_2	381	60	20	9	5
2004_3	413	84	33	17	10
2004_4	458	90	35	18	10
2004_5	415	81	31	16	9
2004_6	381	76	30	15	9
2004_7	390	80	31	16	10
2004_8	287	67	29	16	10
2004_9	267	37	11	5	3
2004_10	127	21	7	3	2
2004_11	148	24	8	4	2
2004_12	145	22	7	3	2
2005_1	162	20	6	3	1
2005_2	233	22	5	2	1
2005_3	255	23	6	2	1
2005_4	264	23	6	2	1
2005_5	216	20	5	2	1
2005_6	207	20	5	2	1
2005_7	117	20	7	3	2
Average	**330**	**60**	**23**	**12**	**7**

4. Filtered frequency distribution for modelled data – 6 month rolling, all sites

Start month	Concentration bin				
	40-80	80-120	120-160	160-200	>200
2003_1	486	92	35	17	10
2003_2	630	95	31	14	8
2003_3	807	103	31	13	7
2003_4	840	100	29	12	6
2003_5	866	103	30	12	6
2003_6	705	93	29	12	6
2003_7	477	69	22	10	5
2003_8	296	52	19	9	5
2003_9	244	43	15	7	4
2003_10	262	44	15	7	4
2003_11	251	46	17	8	5
2003_12	388	65	23	11	6
2004_1	635	90	29	13	7
2004_2	741	140	53	27	16
2004_3	747	147	57	29	17
2004_4	851	145	51	25	14
2004_5	790	136	48	23	13
2004_6	678	120	43	21	12
2004_7	465	94	37	19	11
2004_8	387	44	12	5	2
2004_9	267	32	9	4	2
2004_10	185	21	6	2	1
2004_11	292	36	11	5	2
2004_12	344	43	13	5	3
2005_1	551	47	11	4	2
2005_2	645	57	14	5	2
2005_3	751	63	15	5	2
2005_4	774	65	15	5	2
2005_5	608	47	10	4	2
2005_6	474	40	9	3	2
2005_7	272	33	9	4	2
Average	**539**	**74**	**24**	**11**	**6**

Appendix B

1. Kincaid data, mean metric.

Linear regression method

Statistic	Number of data points used to perform calibration						
	3	4	8	12	18	20	24
NMSE	0.43	0.22	0.13	0.11	0.11	0.11	0.11
R2	0.07	0.14	0.27	0.31	0.31	0.34	0.36
F1.2	0.29	0.32	0.38	0.37	0.37	0.37	0.35
F1.5	0.61	0.66	0.74	0.78	0.82	0.82	0.83
F2	0.85	0.91	0.95	0.98	0.99	0.99	0.99
F5	0.99	0.99	1.00	1.00	1.00	1.00	1.00
F10	1.00	1.00	1.00	1.00	1.00	1.00	1.00
MB	1.05	1.04	0.98	1.01	1.00	0.99	0.99

Simple ratio method

Statistic	Number of data points used to perform calibration						
	3	4	8	12	18	20	24
NMSE	0.12	0.12	0.11	0.10	0.10	0.10	0.10
R2	0.29	0.30	0.33	0.37	0.38	0.38	0.41
F1.2	0.39	0.41	0.38	0.38	0.35	0.37	0.35
F1.5	0.74	0.76	0.80	0.83	0.85	0.85	0.83
F2	0.96	0.97	0.98	0.99	1.00	1.00	1.00
F5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
F10	1.00	1.00	1.00	1.00	1.00	1.00	1.00
MB	1.02	0.99	1.01	1.00	0.99	0.99	1.00

Kriging residual method

Statistic	Number of data points used to perform calibration						
	3	4	8	12	18	20	24
NMSE	0.21	0.21	0.17	0.17	0.19	0.17	0.18
R2	0.28	0.28	0.31	0.31	0.25	0.31	0.25
F1.2	0.30	0.30	0.35	0.39	0.38	0.43	0.43
F1.5	0.60	0.60	0.65	0.68	0.66	0.70	0.72
F2	0.84	0.85	0.89	0.89	0.88	0.89	0.89
F5	0.98	0.97	1.00	0.99	1.00	1.00	1.00
F10	0.99	0.99	1.00	0.99	1.00	1.00	1.00
MB	0.97	0.99	1.01	1.01	0.99	1.03	1.02

Kriging ratio method

Statistic	Number of data points used to perform calibration						
	3	4	8	12	18	20	24
NMSE	0.14	0.15	0.13	0.13	0.14	0.14	0.12
R2	0.24	0.22	0.31	0.30	0.26	0.29	0.35
F1.2	0.36	0.34	0.42	0.44	0.46	0.45	0.46
F1.5	0.71	0.70	0.75	0.74	0.74	0.75	0.76
F2	0.94	0.94	0.96	0.94	0.94	0.93	0.93
F5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
F10	1.00	1.00	1.00	1.00	1.00	1.00	1.00
MB	0.99	1.01	0.99	0.99	0.96	0.99	0.96

2. Kincaid data, 99.9th percentile metric

Linear regression method

Statistic	Number of data points used to perform calibration						
	2	3	7	11	17	19	23
NMSE	1.22	0.32	0.11	0.10	0.09	0.09	0.08
R2	0.02	0.05	0.20	0.20	0.24	0.24	0.25
F1.2	0.32	0.36	0.43	0.48	0.51	0.49	0.52
F1.5	0.59	0.64	0.75	0.75	0.76	0.74	0.77
F2	0.80	0.89	0.97	0.98	0.99	1.00	1.00
F5	0.97	0.99	1.00	1.00	1.00	1.00	1.00
F10	0.99	0.99	1.00	1.00	1.00	1.00	1.00
MB	1.03	1.00	1.02	0.99	0.99	1.01	0.99

Simple ratio method

Statistic	Number of data points used to perform calibration						
	2	3	7	11	17	19	23
NMSE	0.14	0.11	0.09	0.08	0.08	0.08	0.09
R2	0.16	0.22	0.28	0.30	0.32	0.33	0.28
F1.2	0.39	0.40	0.47	0.52	0.51	0.52	0.49
F1.5	0.73	0.76	0.80	0.81	0.82	0.84	0.82
F2	0.94	0.97	0.99	1.00	1.00	1.00	1.00
F5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
F10	1.00	1.00	1.00	1.00	1.00	1.00	1.00
MB	1.00	1.04	0.99	0.98	1.01	1.02	0.99

Kriging residual method

Statistic	Number of data points used to perform calibration					
	3	7	11	17	19	23
NMSE	0.14	0.12	0.10	0.11	0.10	0.11
R2	0.20	0.27	0.29	0.31	0.32	0.30
F1.2	0.36	0.41	0.44	0.45	0.47	0.46
F1.5	0.70	0.75	0.77	0.77	0.78	0.75
F2	0.92	0.94	0.97	0.98	0.99	0.99
F5	1.00	1.00	1.00	1.00	1.00	1.00
F10	1.00	1.00	1.00	1.00	1.00	1.00
MB	1.00	0.98	1.00	1.02	1.00	1.04

Kriging ratio method

Statistic	Number of data points used to perform calibration					
	3	7	11	17	19	23
NMSE	0.13	0.10	0.10	0.10	0.09	0.09
R2	0.20	0.30	0.30	0.32	0.35	0.37
F1.2	0.40	0.46	0.47	0.49	0.50	0.56
F1.5	0.74	0.80	0.78	0.77	0.82	0.81
F2	0.95	0.97	0.98	0.99	0.99	1.00
F5	1.00	1.00	1.00	1.00	1.00	1.00
F10	1.00	1.00	1.00	1.00	1.00	1.00
MB	1.01	0.98	0.98	0.97	0.99	0.97

3. Aire Valley data, mean metric

Linear regression method

Statistic	Number of data points used to perform calibration			
	2	3	4	5
NMSE	0.21	0.11	0.08	0.09
R2	0.04	0.01	0.10	0.09
F1.2	0.44	0.51	0.46	0.48
F1.5	0.82	0.91	0.86	0.78
F2	0.92	0.92	0.90	0.78
F5	0.98	1.00	1.00	1.00
F10	1.00	1.00	1.00	1.00
MB	1.17	1.11	1.10	1.12

Simple ratio method

Statistic	Number of data points used to perform calibration				
	1	2	3	4	5
NMSE	0.12	0.07	0.06	0.05	0.04
R2	0.49	0.58	0.60	0.57	0.67
F1.2	0.28	0.42	0.53	0.57	0.53
F1.5	0.74	0.87	0.83	0.89	0.83
F2	1.00	1.00	1.00	1.00	1.00
F5	1.00	1.00	1.00	1.00	1.00
F10	1.00	1.00	1.00	1.00	1.00
MB	1.08	1.03	1.00	1.02	0.99

Kriging ratio method

Statistic	Number of data points used to perform calibration		
	3	4	5
NMSE	0.08	0.09	0.10
R2	0.56	0.51	0.53
F1.2	0.23	0.13	0.00
F1.5	0.86	0.92	0.84
F2	1.00	1.00	1.00
F5	1.00	1.00	1.00
F10	1.00	1.00	1.00
MB	0.99	0.99	0.98

Kriging residual method

Statistic	Number of data points used to perform calibration		
	3	4	5
NMSE	0.06	0.07	0.07
R2	0.51	0.45	0.58
F1.2	0.25	0.25	0.40
F1.5	0.93	0.94	1.00
F2	0.99	1.00	1.00
F5	1.00	1.00	1.00
F10	1.00	1.00	1.00
MB	0.99	0.98	0.95

4. Aire Valley data, 99.9th percentile metric

Linear regression method

Statistic	Number of data points used to perform calibration			
	2	3	4	5
NMSE	0.37	0.19	0.09	0.06
R2	0.02	0.05	0.22	0.53
F1.2	0.35	0.36	0.46	0.30
F1.5	0.67	0.71	0.87	1.00
F2	0.81	0.89	1.00	1.00
F5	0.98	1.00	1.00	1.00
F10	1.00	1.00	1.00	1.00
MB	1.12	1.12	1.06	1.06

Simple ratio method

Statistic	Number of data points used to perform calibration				
	1	2	3	4	5
NMSE	0.15	0.12	0.11	0.10	0.09
R2	0.04	0.06	0.08	0.13	0.06
F1.2	0.41	0.31	0.32	0.29	0.12
F1.5	0.59	0.72	0.78	0.79	0.85
F2	1.00	1.00	1.00	1.00	1.00
F5	1.00	1.00	1.00	1.00	1.00
F10	1.00	1.00	1.00	1.00	1.00
MB	1.06	1.00	1.00	1.02	1.01

Kriging ratio method

Statistic	Number of data points used to perform calibration		
	3	4	5
NMSE	0.13	0.11	0.12
R2	0.02	0.00	0.00
F1.2	0.33	0.39	0.48
F1.5	0.66	0.64	0.48
F2	1.00	1.00	1.00
F5	1.00	1.00	1.00
F10	1.00	1.00	1.00
MB	1.03	0.99	0.92

Kriging residual method

Statistic	Number of data points used to perform calibration		
	3	4	5
NMSE	0.17	0.16	0.17
R2	0.02	0.00	0.00
F1.2	0.26	0.31	0.47
F1.5	0.57	0.55	0.47
F2	0.90	0.86	0.84
F5	1.00	1.00	1.00
F10	1.00	1.00	1.00
MB	0.97	0.97	0.98

We are The Environment Agency. It's our job to look after your environment and make it **a better place** – for you, and for future generations.

Your environment is the air you breathe, the water you drink and the ground you walk on. Working with business, Government and society as a whole, we are making your environment cleaner and healthier.

The Environment Agency. Out there, making your environment a better place.

Published by:

Environment Agency
Rio House
Waterside Drive, Aztec West
Almondsbury, Bristol BS32 4UD
Tel: 0870 8506506
Email: enquiries@environment-agency.gov.uk
www.environment-agency.gov.uk