UK Air Quality Forecasting: Annual Report 2012

Report for Defra and the Devolved Administrations

Ricardo-AEA/R3384/ED48946
Issue Number 1
Date 30/08/2013

Customer:

Defra and the Devolved Administrations

Customer reference:

Contract RMP 5270

Confidentiality, copyright & reproduction:

Ricardo-AEA reference:

Ref: ED48946- Issue Number 1

Contact:

Paul Willis
Ricardo-AEA Ltd
Gemini Building, Harwell, Didcot, OX11 0QR
t: 01235 753639
e: paul.willis@ricardo-aea.com

Ricardo-AEA is certificated to ISO9001 and ISO14001

Author:

Agnieszka Griffin, Andrea Fraser

Approved By:

Paul Willis

Date:

30 August 2013

Signed:

Paul Willis

Executive summary

This report covers the operational activities carried out by Ricardo-AEA on the UK Air Quality Forecasting Contract for the year 2012. The work is funded by the Department for Environment Food and Rural Affairs, the Scottish Government, Welsh Government and the Department of the Environment in Northern Ireland.

During 2012, there was a total of forty days on which “Very high” and “High” air pollution was recorded across the UK. Thirty-six of these days were due to particulate PM_{10} and $PM_{2.5}$, three due to ozone, one due to SO_2 and none due to NO_2.

The forecasting success and accuracy for this year is summarised in Box 1, together with the results from the previous calendar year. The previous year comparison should be treated with caution because of the change in the Daily Air Quality Index (DAQI) in 2012. $PM_{2.5}$ is included in the analysis for the first time whilst the criteria for Moderate and High ozone and NO_2 have changed significantly.

The table shows that High forecasts in 2012 were improved in Agglomerations but deteriorated in Zones compared to 2011. This is due to improved urban forecasts for Particulates in 2012 but poorer regional forecasts.

The performance for the “Moderate” band forecasts was satisfactory, with an overall accuracy figure of 55%, slightly lower compared to previous years. However, this is perhaps the most meaningful and consistent figure from year-to-year. Please note that due to the current definition of +/- 1 index value in each band, success rates can be reported as greater than 100 %.

Box 1 – Forecast success/accuracy for incidents above “High” and above “Moderate” in 2012 *(and 2011)*

Region/Area	*“High” % success*	*% accuracy*	*“Moderate” % success*	*% accuracy*
Zones	39 *(57)*	16 *(39)*	117 *(119)*	63 *(89)*
Agglomerations	47 *(14)*	18 *(11)*	*118 (118)*	46 *(81)*

The % accuracy figure for Moderate dropped considerably in 2012 compared to 2011, indicating that these forecasts were probably over-cautious with too many false alarms of Moderate issued.

There were no reported breakdowns in the service over the year and all bulletins were delivered to the Air Quality Communications contractor on time.

We have continued to actively research ways of improving the air pollution forecasting system by:

1. Investigating the use of automatic software systems to streamline the activities within the forecasting process, thereby allowing forecasters to spend their time more efficiently in maximising forecast accuracy.
2. Improving the CMAQ model runs which can be used for daily and ad-hoc analysis.
3. Improving and updating the emissions inventories used in our models.

Table of contents

Appendices

1 Introduction

Ricardo-AEA is contracted by the Department for Environment, Food and Rural Affairs (Defra), Scottish Government, Welsh Assembly Government and the Department of the Environment in Northern Ireland (the DAs) to provide 24-hour air pollution forecasts which are widely disseminated through the media. The forecasts allow individuals who may be affected by episodes of high air pollutant concentrations to take appropriate actions. These can include increasing medication or taking steps to reduce exposure and dose.

A forecast of the following day's air pollution is prepared every day by the duty forecaster at Ricardo-AEA. The forecast is based on a combination of results from the CMAQ chemical transport model, measurement data from the UKs Automatic Urban and Rural network (AURN), and expert judgement. During 2012 these forecasts were reconfigured to be based on calendar days rather than a 24-hour forward look from their time of issue. The forecast includes simple text for the UK-AIR website together with numerical forecasts of the Daily Air Quality Index (DAQI).

The latest daily forecast summary is published on the UK-AIR website home page at http://uk-air.defra.gov.uk/. Full text forecasts for the current day, following day and a longer term outlook are published together with a mapped presentation of the numerical forecast on a new webpage introduced during 2012 at http://uk-air.defra.gov.uk/forecasting/.

The numerical forecast consists of a prediction of the daily air quality index for the worst-case situation in 16 zones and 16 agglomerations for the following day. Forecasts are disseminated in a number of ways to maximise public accessibility; these include the ”Pollution Forecast” UK-AIR web page and a Freephone telephone service. Work in 2013/4 will continue to look to improve access to the forecasting service through a number of means, including improving visibility/links to the webpage, increasing uptake/visibility of the UK-AIR twitter feed and publicising website plugins for incorporation forecasts into public facing websites.

Forecasts are routinely prepared at 1.00 p.m. each day (including weekends and bank holidays) for upload to the Internet by Defra’s Data Dissemination Unit contractor. They are then included in subsequent air quality bulletins, newspapers and many other interested organisations. Updated forecasts may be issued at other times of the day if the situation is changing rapidly.

This report covers and analyses the media forecasts issued during the 12 months from January 1st to December 31st 2012. Results from forecasting models are available each day and are used in constructing the forecast. The forecasters issue predictions for rural, urban background and roadside environments but, for the purposes of this report, these have been combined into a single “worst-case” category (i.e. the forecasts issued are not analysed by environment type within this report).

Twice per week, on Tuesdays and Fridays, Ricardo-AEA also provides a long-range pollution outlook. This takes the form of a short piece of text which is emailed to approximately 120 recipients in the Defra and other government Departments. The outlook is compiled by examining the outputs from our pollution models, which currently extend to 3 days ahead for Defra and the DAs, and by assessing the long-term weather situation.

Forecasts issued by Ricardo-AEA for UK regions and individual local authorities are checked for consistency with the overall UK forecasts issued on behalf of Defra and the DAs.

2 New Daily Air Quality Index

The Daily Air Quality Index (DAQI) and banding system first developed in the late 90-ties underwent the major changes after the recommendations were published in the *Review of the UK Air Quality Index* in June 2011 by the Committee on the Medical Effects of Air Pollutants (COMEAP). Defra and the Devolved Administrations implemented COMEAP recommendations from 1st January 2012.

The implemented recommendations include:

- A 10-point scale index with four bands of low, "Moderate", "High" and "Very high"
- Changes to the index bands for particulate matter (PM_{10}), nitrogen dioxide (NO_2) and ozone (O_3)
- Exclusion of carbon monoxide (CO) from the index
- Inclusion of fine particulate matter ($PM_{2.5}$) in the index
- Updated health advice
- Introduction of a concept of "trigger" values to allow for the prediction of episodes and elevated air pollution in real-time as they emerge
- Information on the long-term health effects of air pollution linked to the DAQI

Table 2-1: The UK's Daily Air Quality Index

Band	Index	Ozone	Nitrogen Dioxide	Sulphur Dioxide	$PM_{2.5}$ Particles (EU Reference Equivalent)	PM_{10} Particles (EU Reference Equivalent)
		Running 8 hourly mean	hourly mean	15 minute mean	24 hour mean	24 hour mean
		μgm^{-3}	μgm^{-3}	μgm^{-3}	μgm^{-3}	μgm^{-3}
Low	1	0-33	0-67	0-88	0-11	0-16
	2	34-66	68-134	89-177	12-23	17-33
	3	67-100	135-200	178-266	24-35	34-50
Moderate	4	101-120	201-267	267-354	36-41	51-58
	5	121-140	268-334	355-443	42-47	59-66
	6	141-160	335-400	444-532	48-53	67-75
High	7	161-187	401-467	533-710	54-64	84-91
	8	188-213	468-534	711-887	59-64	84-91
	9	214-240	535-600	888-1064	65-70	92-100
Very High	10	241 or more	601 or more	1065 or more	71 or more	101 or more

The new DAQI not only comprises of the COMEAP's recommendations but was also aligned with the EU Limit Values.

The implementation of the new DAQI resulted in the following changes to the index and banding concentrations:

- The minimum concentrations for the "Moderate" band have increased by one µg/m^3 to ensure that the "Low"/"Moderate" division is compatible with the EU Limit Values for PM_{10}, SO_2 and NO_2. Also to ensure that the ratio of $PM_{2.5}$ and PM_{10} is maintained throughout the index.
- The minimum concentration for the "High" and "Very high" bands were increased by one µg/m^3 for the consistency with the changes made to the "Moderate" band
- The divisions between index values have been adjusted so that they are more evenly distributed within each band

Defra adopted a "Low"/"Moderate" breakpoint of 100 µg/m^3 for ozone. WHO recommend a guideline of 100 µg/m^3 but noted that some effects may be expected in susceptible individuals within the population at this level. COMEAP took into account WHO recommendations and set the lower breakpoint of 80 µg/m^3. However, the value of 80 µg/m^3 is close to the hemispheric background level of ozone. It would be likely that this level would be exceeded on the majority of days in the UK (up to 80% of days in a year) and elsewhere, and air pollution would be described as "Moderate" on most days. If "Moderate" air pollution would be reported on the majority of days then the importance of messages warning of elevated air pollution would be lost.

Trigger values

The main purpose of DAQI is to provide health information to the public on short-term changes in air quality. The DAQI index is based on 24-hour mean concentrations for particulate matter and 8-hour mean concentrations for ozone. The long averaging times for these pollutants mean that sudden increases in measured levels of pollution may not affect the DAQI reported until the pollution episode is well established. The use of trigger values were recommended by COMEAP to complement the DAQI and should be used to provide early warning to the public about possible exposure to "Moderate", "High" and "Very high" levels of pollution.

During 2012 Defra implemented the recommended trigger values as part of its daily air quality forecasting service, to complement the other tools available to the duty forecaster.

Trigger values are defined as hourly measured concentrations which provide an early warning of a period of "Moderate", "High" or "Very high" air pollution later that same day. The triggers are based on two consecutive hourly mean concentrations. The first hourly mean has to be greater than or equal to a threshold and the second hourly mean concentration has to be greater or equal to the first.

3 Development of the WRF-CMAQ Air Quality Forecasting Model

This section provides a summary of 2012 development work on the WRF-CMAQ air quality forecasting model. During this year the main developments included extending the forecast from 48hrs to 72hrs and starting a 2km resolution model over London and the south-east in time for the London Olympics.

WRF (Weather Research and Forecasting) is a numerical weather model developed in the USA as a collaborative partnership between several agencies including: National Center for Atmospheric Research (NCAR), the National Centers for Environmental Prediction (NCEP), the Air Force Weather Agency and the Naval Research Laboratory. The WRF code and documentation are available at www.wrf-model.org.

The CMAQ (Community Multiscalar Air Quality) model was first developed under the US EPA Models-3 project (Byun and Ching, 1999). It is a comprehensive AQ model, incorporating meteorology, emissions, land use, chemistry and aerosol processes. For the UK AQ forecasting project it is driven by weather from WRF, and the emissions are generated using the NAEI (National Atmospheric Emissions Inventory) and EMEP (European Monitoring and Evaluation Program) Inventory, supplemented by natural emissions calculated using the Biogenic Potential Inventory. In 2012 the emissions were updated to run from 2009 base year emissions until after the Olympics then updated to 2010. CMAQ model code and documentation are available at www.cmaq-model.org.

Following the Defra-EA-USEPA meeting in June 2012 the data assimilation method used by the USEPA was tested on the UK forecast. The first indications are that this will improve the operational ozone, PM_{10} and $PM_{2.5}$ forecasts.

The maps and model evaluation tools were updated to reflect the new Daily Air Quality Index (DAQI) bands in January 2012. This included updating the colour schemes and indices for the daily air quality maps, creating animations for 8hr and 24hr averages, and adding $PM_{2.5}$.

3.1 Extension to the air quality forecast model for the London Olympics.

During the London Olympics a number of developments were introduced in order to enhance the UK forecasting service:

The scheduling and efficiency of model runs was improved so that the air quality forecast could be delivered earlier in the day. This was required in order to fit in with the timing of HPA and Defra situation reports. Forecasts were brought forwards to be issued between 1pm and 2pm daily.

The daily model run was extended from 48hrs to 72hrs in order to provide further advanced warning of changing conditions. This development was intended to be of particular benefit to Defra and the duty forecaster in cases where increased pollution events were expected over the weekend.

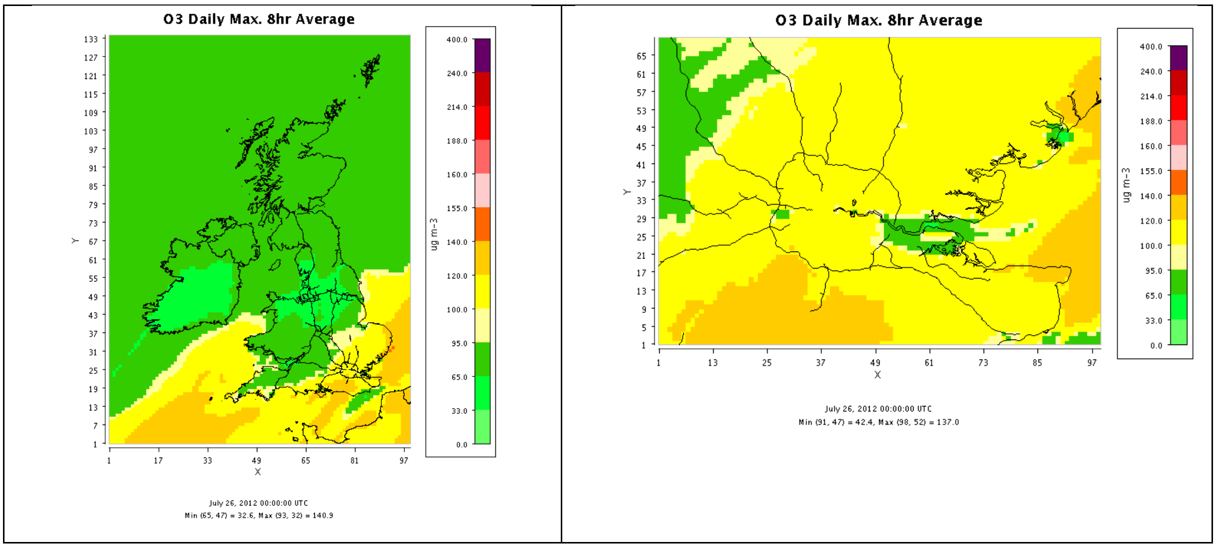
A finer resolution 2km forecast was introduced for London and the south-east in order to give improved visibility of local variation in pollutant concentrations around the Olympics venues. Increasing the model resolution does not necessarily improve the model performance when compared to monitoring sites. However the animations give a better indication of the anticipated extent and progress of a pollution event.

Key routes such as the torch relay, marathon, cycling and triathlon were overlaid onto the forecast maps so that the predicted concentrations in these areas could be analysed more carefully.

These developments were all successfully introduced well before the commencement of the torch relay and arrival of the athletes. However wet weather meant that the forecasts were straightforward for low pollution and not significantly tested during May, June and much of July 2012.

However, on the days before the Olympics started, amid much public and media interest, there was a spell of hot weather and evidence that ozone would increase and reach the “Moderate” air quality index. Figure 3-1 shows the improved model resolution over London and the south-east, giving a better indication of the spatial extent of the “Moderate” ozone. Having the 3 day forecast gave advanced warning of this event and also an indication that it would end before the Olympics started.

Figure 3-1: Ozone forecast for 25th July 2012.



3.2 Evaluating Data Assimilation

The air quality forecast prepared by Ricardo-AEA uses CMAQ to predict pollutant concentrations in the UK each day and for the following two days. Many researchers have proposed that data assimilation of observations can be used to improve the forecast accuracy.

Different sophisticated approaches exist, such as four-dimensional variational data assimilation (4DVAR) and Ensemble Kalman Filtering (EnKF). Whilst these techniques are potentially very powerful, they are also highly computation-intensive, requiring either the implementation of a model adjoint, or the simultaneous integration of several tens of model ensemble members.

In recent years, rather more simple bias adjustment techniques have emerged, in which the bias correction factors are estimated by means of the Kalman filter (KF) approach. These techniques are applied in post-processing (i.e., off-line) mode rather than as a part of the initialization of the deterministic forecast, and they are characterized by a very low computational cost.

The US EPA currently implements bias adjustment of its ozone and particulate matter forecasts using a Kalman filter approach. They kindly provided Ricardo-AEA with computer codes used to implement the Kalman filter algorithm. The EPA uses forecast and observed hourly concentrations for the previous two days at monitoring sites. It uses the Kalman filter approach to estimate a bias adjustment at each of the monitoring sites for each hour of the current day based on the bias adjustment calculated for the same hour of previous days.

As part of the UK CMAQ development project these computer codes were rewritten and adapted for use in UK forecasting, taking into account of the different formatting of the UK data.

The aims of the study carried out in 2012/2013 were:

- to investigate how previous forecasts and observations can be used to provide improved estimates of future concentrations, and
- to demonstrate the effectiveness of proposed methods

Below are examples of the results showing the effect of applying the filter to daily average $PM_{2.5}$ (Figure 3-2) and daily maximum ozone (Figure 3-3).

***Figure 3-2:* Scatter plots showing correlation between observations and predictions of daily average $PM_{2.5}$ at urban background sites between 01/01/13 and 15/03/13. The left hand plots shows the predicted data before (black circles) and after (red dots) adjustment by Kalman filter. The right hand density scatter plot shows the same adjusted data, but shows the density of sites, with high densities indicated by red and low densities by yellow.**

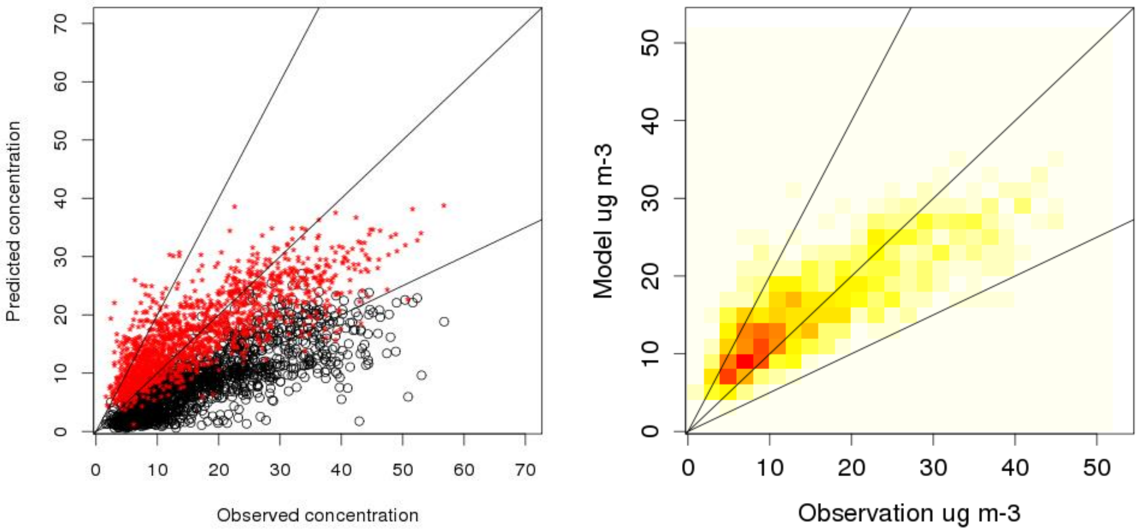
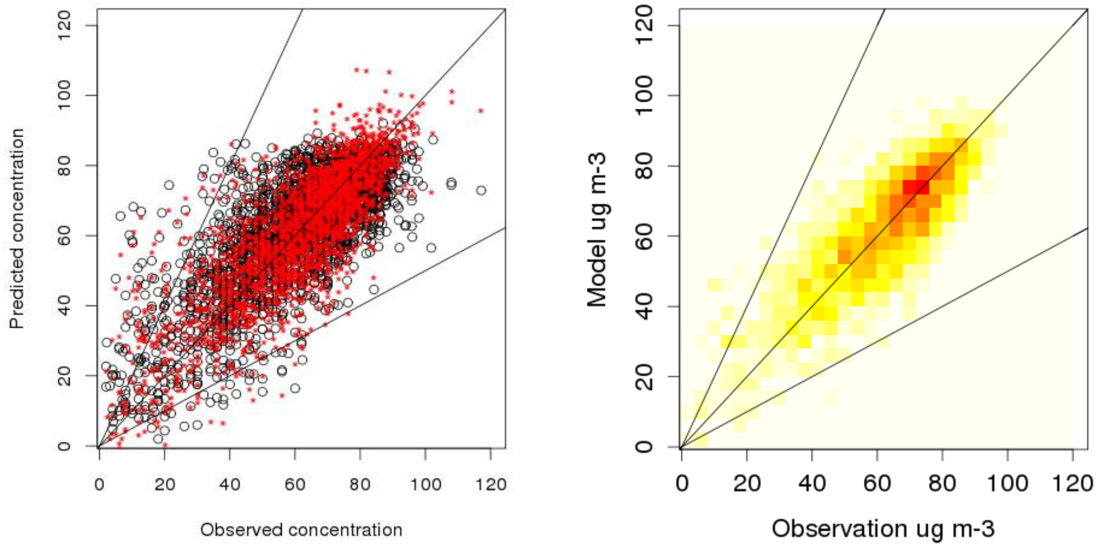


***Figure 3-3:* Scatter plots showing correlation between observations and predictions of daily maximum ozone at urban background sites between 01/01/13 and 15/03/13. The left hand plots shows the correlations to predicted data before (black circles) and after (red dots) adjustment**

by Kalman filter. The right hand density scatter plot shows the same adjusted data, but indicates the density of sites, with high densities indicated by red and low densities by yellow.



The study concluded that use of the basic Kalman filter algorithm substantially improves forecasts of daily mean and daily maximum 1-hour concentrations of particulate matter $PM_{2.5}$ and ozone. It effectively eliminates the mean bias between the predictions and the measured concentrations and in most cases reduces the standard deviation of the bias

The study considered the bias at each monitoring site individually. The filter performance may potentially improve if the bias adjustment took account of the calculated bias at neighbouring monitoring sites. It would then be less influenced by "events" (e.g. short term pollutant releases close to the monitor). The Kalman filter algorithms can be adapted to take account of the bias at more than one monitoring station on the basis of existing mathematical formulations.

A further advantage of this development would be that it would be possible to map the filtered concentrations over the whole of the UK. Discussion, of how and when to introduce data assimilation into the UK AQ Forecast model, is on-going.

More details of the method are provided in Appendix 3.

3.3 Model Evaluation

Air quality forecast model evaluation is an on-going work-in-progress. Model values corresponding to monitoring sites are automatically extracted from the daily forecast output files and stored in a mySQL database along with the provisional and ratified monitoring data from the AURN monitoring network, and weather measurement data from European airports. R scripts are used to produce the daily and monthly evaluations for WRF and CMAQ. Air quality evaluations are produced separately for each forecast species and class of monitoring site (rural, urban background and urban centre). WRF is evaluated where AURN sites and weather monitoring data fall within the same model grid cell. Data are only available for part of 2012 for the London 2km model.

There are currently two distinct levels of model evaluation:

Daily rolling evaluation to provide guidance to the air quality forecasters of how well WRF and CMAQ represent the current conditions. This evaluation is available on the forecasting dashboard alongside the daily maps, giving an up-to-date indication of model performance under the current meteorological conditions. The WRF and CMAQ skill plots are updated every morning and cover the previous 14 day period. For CMAQ the line plot is extended to the previous 7 days and the next 3 days forecast.

The same evaluation is used for monthly, quarterly and yearly analysis using ratified (or provisional) monitoring data. This more extended analysis is used to evaluate overall model performance and to guide model development.

3.3.1 WRF Forecast data evaluation 2012

The performance of the WRF forecast for wind speed, wind direction and temperature has been evaluated by comparing the model outputs against surface observations collected from automated meteorological stations located at several different airports across the UK.

The results shown here are for a comparison of modelled WRF 10km gridded data with observed meteorological values. Table 3-1 and Table 3-2 show that WRF performs well for 2m temperature and wind speed. Wind direction shows a mean positive bias of 16 degrees, this is similar to previous years.

Table 3-1: Comparison of annual observations and WRF model forecast temperature and wind speed

	2m Temperature Observations	2m Temperature Model	10m Wind Speed (ms^{-1}) Obs.	10m Wind Speed (ms^{-1}) Model
Max	30	28	20	21
Min	-23	-10	0	0
Mean	10.22	9.64	4.15	5.04
Median	10	10	4	5

Table 3-2: Statistics for annual observations and WRF model forecast temperature, wind speed and wind direction

	2m Temperature	10m Wind Speed (ms^{-1})	10m wind direction (degrees)
Standard Deviation	2.14	2.00	42.96
Mean Absolute Error	1.56	1.63	32.09
Mean Bias	-0.40	0.77	16.09
Root Mean Sq. Error	2.18	2.14	45.87

3.3.2 CMAQ Forecast data evaluation 2012

Table 3-3 shows a summary of the annual CMAQ performance for 2012. It summarises the metrics recommended in the model evaluation protocol developed by Derwent et al. 2009. The results show a tendency to overestimate ozone and underestimate PM_{10}, $PM_{2.5}$ and NO_2 Performance is better at the rural sites than the urban background or urban centre sites. This would be expected as the rural sites are selected to reflect the air quality over a similar spatial scale as the 10km model. Annual results are not available for the 2km model.

For O_3, NO_2 and PM_{10} more than 50% of the paired values fall within a factor of 2 for each of the rural and urban background sites. Performance for $PM_{2.5}$ was below 50% of the paired values falling within a factor of 2 with on underestimate of the PM. The Odds Ratio Skill Score shows that the model performs well at predicting the low to "Moderate" threshold, performing best for ozone.

For SO_2 the relatively low number of modelled values falling within a factor of two of the observation (Table 3-3) reflects the high number of observations that are low and close to the lower limit of the measurement calibration. The actual bias is less than 7μgm^{-3}.

Table 3-3: Annual evaluation of forecast species for rural, urban background and urban centre AURN monitoring sites 10km model UK simulation and ratified observations

Pollutant	Area type	Normal Mean Bias %	Normal Mean Error %	Factor of 2 %	ORSS
O_3	Rural	22	30	82	0.96
	Urban BG	32	40	71	0.97
	Urban C	45	54	64	0.97
NO_2	Rural	11	66	54	*
	Urban BG	-19	49	55	0.98
	Urban C	-36	52	48	0.93
PM_{10}	Rural	-27	47	59	0.98
	Urban BG	-37	50	53	0.93
	Urban C	-44	52	46	0.93
$PM_{2.5}$	Rural	-48	55	40	0.98
	Urban BG	-50	58	40	0.94
	Urban C	-56	59	33	0.96
SO_2	Rural	3	6	30	*
	Urban BG	3	6	39	*
	Urban C	4	7	38	*

ORSS – Odds Ratio Skill Score – based on the low/"Moderate" AQI interface
* Only calculated where the threshold is exceeded

The standard evaluations are produced for all species and area type, a rolling 14 day analysis is updated daily. The same analyses are produced monthly and annually. In Appendix 3 there are examples of the annual skill evaluations for rural ozone and urban background ozone are shown in Table A3.4 and Table A3.5 respectively.

3.4 Summary of Model Development and Performance

WRF and CMAQ continue to be used successfully each day to produce daily predicted maps and animations to support the UK duty air quality forecasters. The WRF weather maps are published for public access on the UK-AIR website at http://uk-air.defra.gov.uk/forecasting/wrf-summary.

The currently operational version of WRF can accurately reproduce observed wind speed and temperature values but has a noticeable positive bias in the predicted wind direction. This bias has been observed in 2011 and by other WRF users, it does not appear to adversely affect the ability of CMAQ to forecast pollutant concentrations across the UK and Europe.

CMAQ forecasts are shown to accurately represent seasonal and daily variations in ozone and PM_{10} concentrations. There is evidence of a slight positive bias in ozone predictions, and an underestimation in PM_{10} concentrations. Both of these are likely to be due to inaccuracies in the underlying pollutant emissions inventories - which are updated with improved versions annually as they become available.

Further development and application of the basic Kalman filter algorithm will improve the forecasts for PM_{10} and $PM_{2.5}$.

4 Analysis of forecasting success rate

4.1 Introduction

Analysis of forecasting performance is carried out for each of the 16 zones and 16 agglomerations used in the daily forecasting service. Further details of these zones and agglomerations are presented in Appendix 2. Forecasting performance is analysed for a single, general pollutant category rather than for each individual pollutant and has been aligned to the forecasting day (a forecasting day runs from the issue time, generally 2 pm). This analysis of forecasting performance is mainly based on the provisional data available and used in the daily forecasting process. However, provisional data are only subject to initial screening checks and it is possible for erroneous results to be reported. To provide a fairer measure of true forecast accuracy any faulty data clearly identified during later stages of the ratification process have therefore been removed from the analysis presented here.

The analysis treats situations where the forecast index was within ±1 of the measured index as a successful prediction since this is the target accuracy we aim to obtain in the forecast. . Using this method it is possible to record rates in excess of 100% rather than 'true' percentages. Further details of the text descriptions and index code used for the forecasting are given in Appendix 1. The forecasting success rates for each zone and agglomeration for January - December 2012 are presented in Table 4-1 (forecasting performance in zones) and Table 4-2 (forecasting performance in agglomerations) for "High" days. Table 4-3 and

Table 4-4 show the same statistics for the "Moderate" band. Table 4-5 provides a summary for each pollutant of the number of days on which "High" and above pollution was measured, the maximum exceedance concentration and the day and site at which it was recorded. The forecasting performance Table 4-1and Table 4-2 give:

- The number of "High" days measured in the PROVISIONAL data
- The number of "High" days forecast
- The number of days with a correct forecast of "High" air pollution, within an agreement of ±1 index value. A "High" forecast is recorded as correct if air pollution is measured "High" and the forecast is within ±1 index value, or it is forecast "High" and the measurement is within ±1 index value. For example measured index 7 with forecast index 6 counts as correct, as does measured index 6 with forecast index 7.
- The number of days when "High" air pollution was forecast ('f' in the tables) but not measured ('m') on the following day to within an agreement of 1 index value.
- The number of days when "High" air pollution was measured ('m') but had not been forecast ('f') to within an agreement of 1 index value.

The two measures of forecasting performance used in this report are the 'success rate' and the 'forecasting accuracy'.

The forecast success rate (%) is calculated as:
- (Number of episodes successfully forecast/total number of episodes measured) x 100

The forecast accuracy (%) is calculated as:
- (Number of episodes successfully forecast/[Number of successful forecasts + number of wrong forecasts]) x 100

4.2 Forecast analysis for 2012

*Table 4-1: Forecast Analysis for UK Zones "High" band and above **

Zone	Central Scotland	East Midlands	Eastern	Greater London	Highland	North East	North East Scotland	North Wales	North West & Merseyside	Northern Ireland	Scottish Borders	South East	South Wales	South West	West Midlands	Yorkshire & Humberside	Overall
Measured days	*0*	*6*	*3(1 ozone)*	*6(1 ozone)*	*0*	*1*	*0*	*0*	*7*	*1*	*0*	*7(1 ozone, 1 SO2)*	*0*	*1 (ozone)*	*3*	*1*	*36*
Forecasted days	*2*	*2*	*7*	*11*	*0*	*3*	*0*	*0*	*8*	*4*	*0*	*14*	*1*	*3*	*4*	*4*	*63*
Ok (f and m)	*0*	*1*	*2*	*3*	*0*	*0*	*0*	*0*	*2*	*0*	*0*	*3*	*0*	*0*	*3*	*0*	*14*
Wrong (f not m)	*2*	*2*	*5*	*8*	*0*	*3*	*0*	*0*	*6*	*4*	*0*	*11*	*1*	*3*	*2*	*4*	*51*
Wrong (m not f)	*0*	*5*	*2*	*3*	*0*	*1*	*0*	*0*	*5*	*1*	*0*	*5*	*0*	*1*	*0*	*1*	*24*
Success %	*100*	*17*	*67*	*50*	*100*	*0*	*100*	*100*	*29*	*0*	*100*	*43*	*100*	*0*	*100*	*0*	*39*
Accuracy %	*0*	*13*	*22*	*21*	*0*	*0*	*0*	*0*	*15*	*0*	*0*	*16*	*0*	*0*	*60*	*0*	*16*

*Table 4-2:Forecast Analysis for UK Agglomerations "High" band and above **

Agglomerations	Belfast Metropolitan Urban Area	Brighton / Worthing / Littlehampton	Bristol Urban Area	Cardiff Urban Area	Edinburgh Urban Area	Glasgow Urban Area	Greater Manchester Urban Area	Leicester Urban Area	Liverpool Urban Area
Measured days	*3*	*0*	*4*	*0*	*1*	*2*	*5*	*1*	*2*
Forecasted days	*4*	*0*	*3*	*3*	*2*	*3*	*7*	*5*	*4*
Ok (f and m)	*4*	*0*	*1*	*0*	*0*	*0*	*1*	*2*	*2*
Wrong (f not m)	*1*	*0*	*2*	*3*	*2*	*3*	*6*	*3*	*2*
Wrong (m not f)	*0*	*0*	*3*	*0*	*1*	*2*	*4*	*0*	*0*
Success %	*133*	*100*	*25*	*100*	*0*	*0*	*20*	*200*	*100*
Accuracy %	*80*	*0*	*17*	*0*	*0*	*0*	*9*	*40*	*50*

*Table 4.2 (cont'd) - Forecast Analysis for UK Agglomerations "High" band and above **

Agglomerations	Nottingham Urban Area	Portsmouth Urban Area	Sheffield Urban Area	Swansea Urban Area	Tyneside	West Midlands Urban Area	West Yorkshire Urban Area	Overall
Measured days	*0*	*4*	*3*	*2*	*1*	*2*	*4*	*34*
Forecasted days	*1*	*8*	*5*	*3*	*1*	*5*	*9*	*63*
Ok (f and m)	*0*	*3*	*1*	*0*	*0*	*1*	*1*	*16*
Wrong (f not m)	*1*	*5*	*5*	*3*	*1*	*4*	*9*	*50*
Wrong (m not f)	*0*	*2*	*2*	*2*	*1*	*1*	*3*	*21*
Success %	*100*	*75*	*33*	*0*	*0*	*50*	*25*	*47*
Accuracy %	*0*	*30*	*13*	*0*	*0*	*17*	*8*	*18*

*Table 4-3:Forecast Analysis for UK Zones "Moderate" band and above **

Zone	Central Scotland	East Midlands	Eastern	Greater London	Highland	North East	North East Scotland	North Wales	North West & Merseyside	Northern Ireland	Scottish Borders	South East	South Wales	South West	West Midlands	Yorkshire & Humberside	Overall
Measured days	*29*	*36*	*51*	*109*	*19*	*19*	*5*	*10*	*24*	*19*	*11*	*52*	*12*	*38*	*32*	*33*	*499*
Forecasted days	*20*	*47*	*64*	*88*	*13*	*29*	*13*	*11*	*42*	*32*	*18*	*82*	*32*	*49*	*46*	*45*	*631*
Ok (f and m)	*24*	*39*	*55*	*119*	*23*	*17*	*13*	*16*	*30*	*27*	*16*	*63*	*21*	*42*	*42*	*37*	*584*
Wrong (f not m)	*7*	*20*	*28*	*31*	*1*	*18*	*4*	*1*	*21*	*12*	*6*	*31*	*13*	*18*	*19*	*22*	*252*
Wrong (m not f)	*13*	*8*	*10*	*9*	*1*	*5*	*1*	*0*	*6*	*7*	*1*	*10*	*3*	*5*	*2*	*6*	*87*
Success %	*83*	*108*	*108*	*109*	*121*	*89*	*260*	*160*	*125*	*142*	*145*	*121*	*175*	*111*	*131*	*112*	*117*
Accuracy %	*55*	*58*	*59*	*75*	*92*	*43*	*72*	*94*	*53*	*59*	*70*	*61*	*57*	*65*	*67*	*57*	*63*

Table 4-4: Forecast Analysis for UK Agglomerations "Moderate" band and above *

Agglomerations	Belfast Metropolitan Urban Area	Brighton/Worthing/ Littlehampton	Bristol Urban Area	Cardiff Urban Area	Edinburgh Urban Area	Glasgow Urban Area	Greater Manchester Urban Area	Leicester Urban Area	Liverpool Urban Area
Measured days	*8*	*19*	*17*	*15*	*7*	*23*	*26*	*14*	*9*
Forecasted days	*18*	*26*	*46*	*27*	*19*	*39*	*41*	*37*	*26*
Ok (f and m)	*10*	*26*	*21*	*14*	*12*	*27*	*24*	*16*	*11*
Wrong (f not m)	*8*	*9*	*28*	*16*	*10*	*21*	*21*	*24*	*15*
Wrong (m not f)	*2*	*2*	*5*	*5*	*2*	*5*	*10*	*3*	*4*
Success %	*125*	*137*	*124*	*93*	*171*	*117*	*92*	*114*	*122*
Accuracy %	*50*	*70*	*39*	*40*	*50*	*51*	*44*	*37*	*37*

***Table 4.4 (cont'd)* - Forecast Analysis for UK Agglomerations "Moderate" band and above ***

Agglomerations	Nottingham Urban Area	Portsmouth Urban Area	Sheffield Urban Area	Swansea Urban Area	Tyneside	West Midlands Urban Area	West Yorkshire Urban Area	Overall
Measured days	*14*	*18*	*18*	*16*	*8*	*24*	*26*	*262*
Forecasted days	*34*	*39*	*33*	*31*	*15*	*49*	*46*	*526*
Ok (f and m)	*19*	*25*	*21*	*18*	*12*	*28*	*25*	*309*
Wrong (f not m)	*19*	*21*	*19*	*15*	*7*	*28*	*23*	*284*
Wrong (m not f)	*2*	*3*	*4*	*8*	*2*	*6*	*9*	*72*
Success %	*136*	*139*	*117*	*113*	*150*	*117*	*96*	*118*
Accuracy %	*48*	*51*	*48*	*44*	*57*	*45*	*44*	*46*

*Please refer to the start of section 3 for an explanation of the derivation of the various statistics, success >100 % may occur.

Table 4-5 – Summary of "High" pollution episodes in 2012

Pollutant	*No. of HIGH or VERY HIGH days*	*Maximum concentration* (Index)*	*Site with max concentration*	*Zone or Agglomeration*	*Date of max conc.*
Ozone	3	169 (Index 7)	Yarner Wood	South-West	26/05
PM_{10}	11	109 (Index 10)	Leeds Headingley Kerbside	West Yorkshire Urban Area	24/03
NO_2	0	n/a	Belfast Centre	Belfast Metropolitan Urban Area	11/12
SO_2	1	571 (Index 7)	Harwell	South-East	13/03
$PM_{2.5}$	24	102 (Index 10)	Wigan Centre	North West & Merseyside	06/11

** Maximum concentrations relate to 8 hourly running mean, daily mean for PM_{10} and $PM_{2.5}$, hourly mean for NO_2 and 15 minute mean for SO_2 Units ug/m^3 throughout.*

4.2.1 General trends in monitoring site data

Three HIGH days were recorded for ozone during 2012, occurring on 26th May, 25th and 26th July. There were sixty seven days throughout the year when "Moderate" level of ozone were recorded, as shown in Figure 4-3. The highest measured 8-hourly concentration of 169 µg/m^3 was recorded at the rural Yarner Wood monitoring site.

There was one "Very high" PM_{10} day recorded on 24th March. There were eleven "High" band PM_{10} episodes experienced in 2012, and forty one "Moderate" band PM_{10} episodes, as shown in Figure 4-4. The highest daily measured concentration was 109 µg/m^3 recorded at the Leeds Headingley Kerbside monitoring site on 24th March.

Figure 4-5 shows that there were seven "Very high" $PM_{2.5}$ days, seventeen "High" $PM_{2.5}$ days and thirty two "Moderate" $PM_{2.5}$ days recorded in 2012. The highest daily $PM_{2.5}$ concentration of 102 µg/m^3 was measured at the Wigan Centre monitoring site on 6th November 2012.

There was one "High" day and nineteen "Moderate" days for SO_2 measured during the 2012 calendar year. The highest measured 15-minute concentration of 571 µg/m^3 was recorded at the rural Harwell monitoring station. There were forty "Moderate" days recorded at the Grangemouth industrial monitoring site. Figure 4-6 shows the frequency of the SO_2 exceedances for 2012.

Ninety five "Moderate" NO_2 days were measured throughout the year, as shown in Figure 4-7. The highest measured 1-hour concentration of 338 µg/m^3 was recorded at Belfast Centre on 11th December 2012. There were 58 "Moderate" days measured at London Marylebone Road, 17 "Moderate" days at Camden Kerbside, 16 "Moderate" days at Oxford Centre and 9 "Moderate" days at Glasgow Kerbside. There were no "High" NO_2 days recorded during 2012.

Possible causes of the 2012 air pollution episodes are detailed in the sections which follow the charts below.

Figure 4-1: Number of stations with air pollution levels of "High" and above for days throughout 2012.

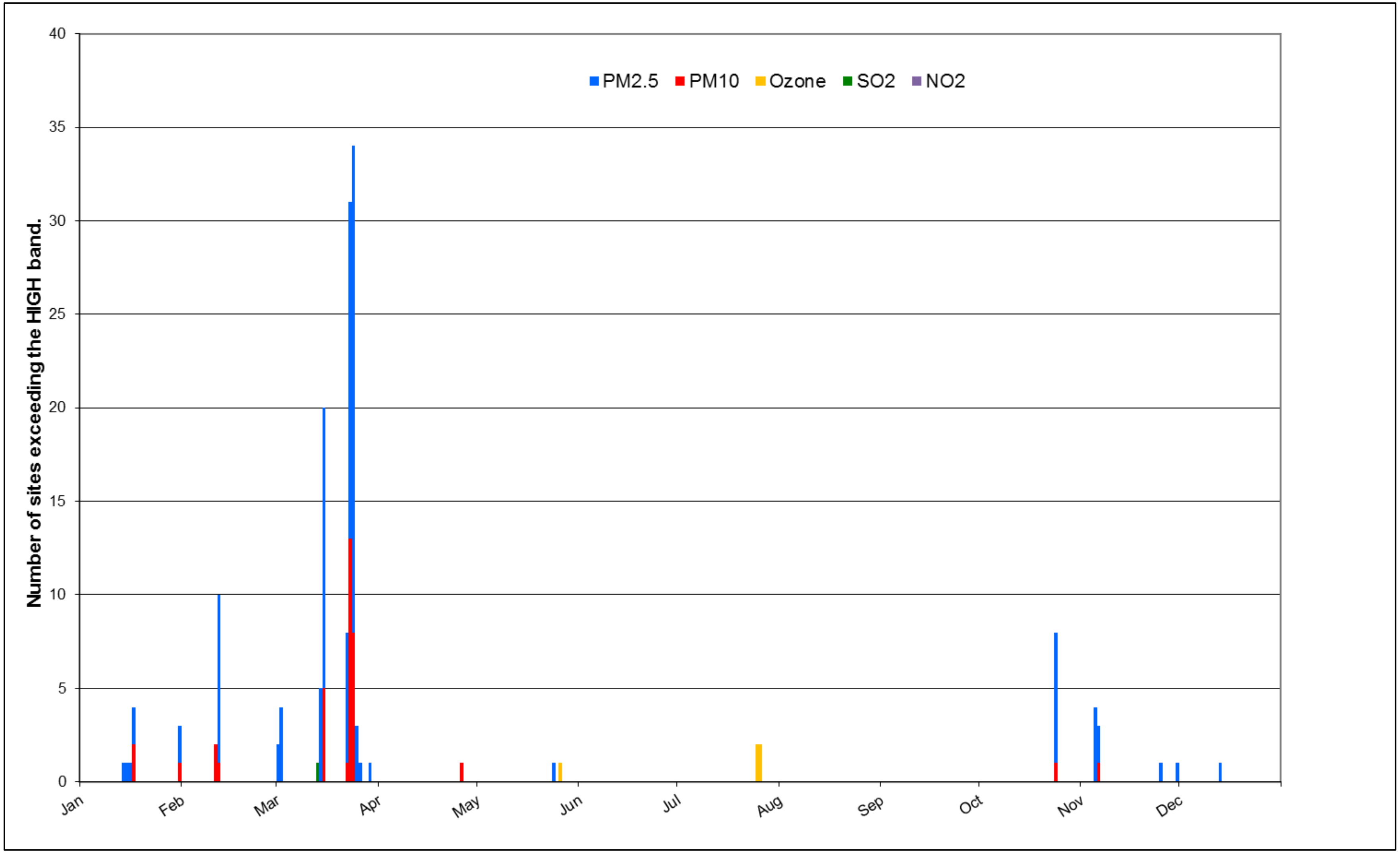


Figure 4-2: Maximum exceedance when air pollution levels were "High" and above for days throughout 2012.

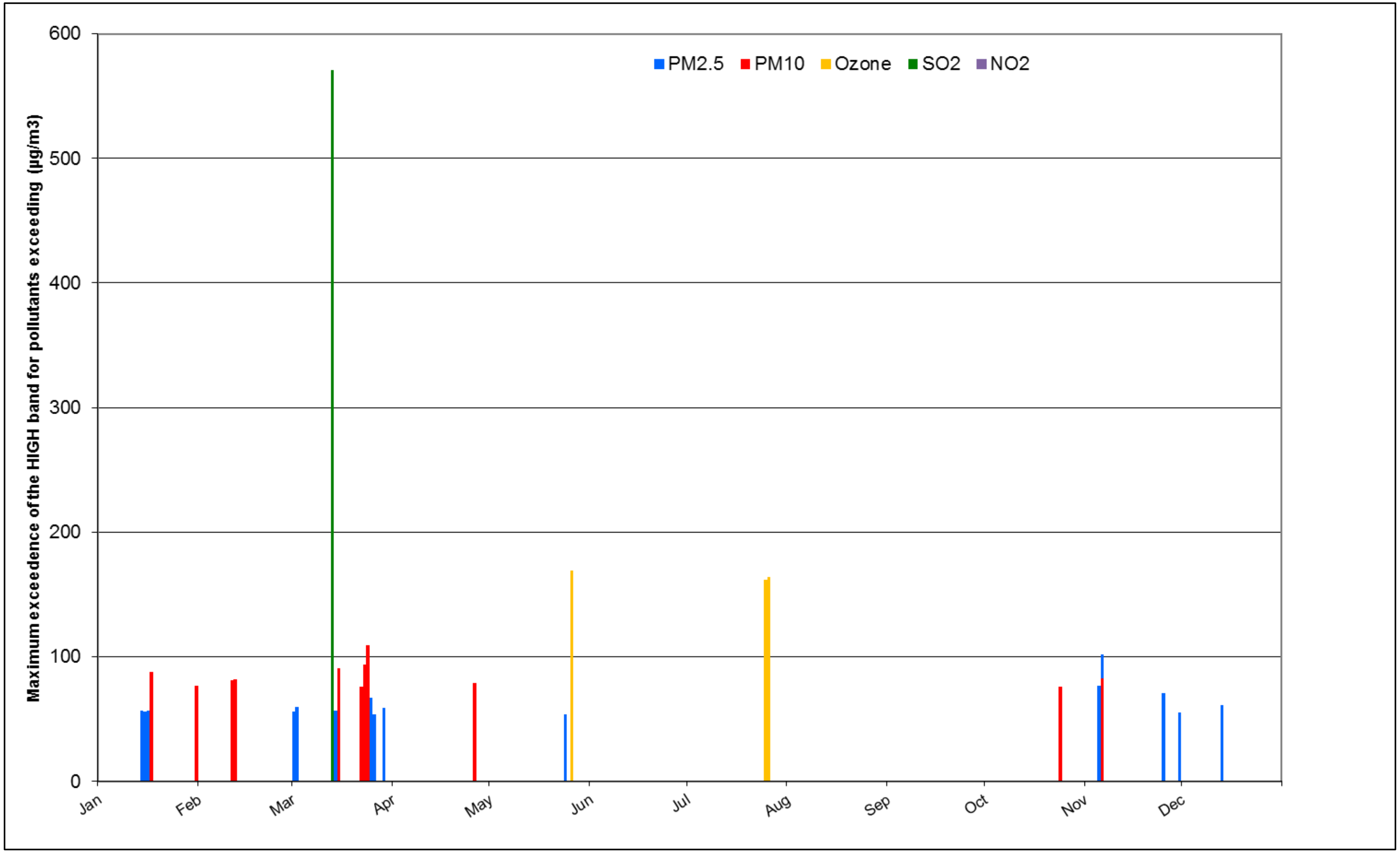


Figure 4-3: 8-hours running ozone concentration across AURN Network with total number of stations measuring "Moderate" or above levels of ozone during 2012.

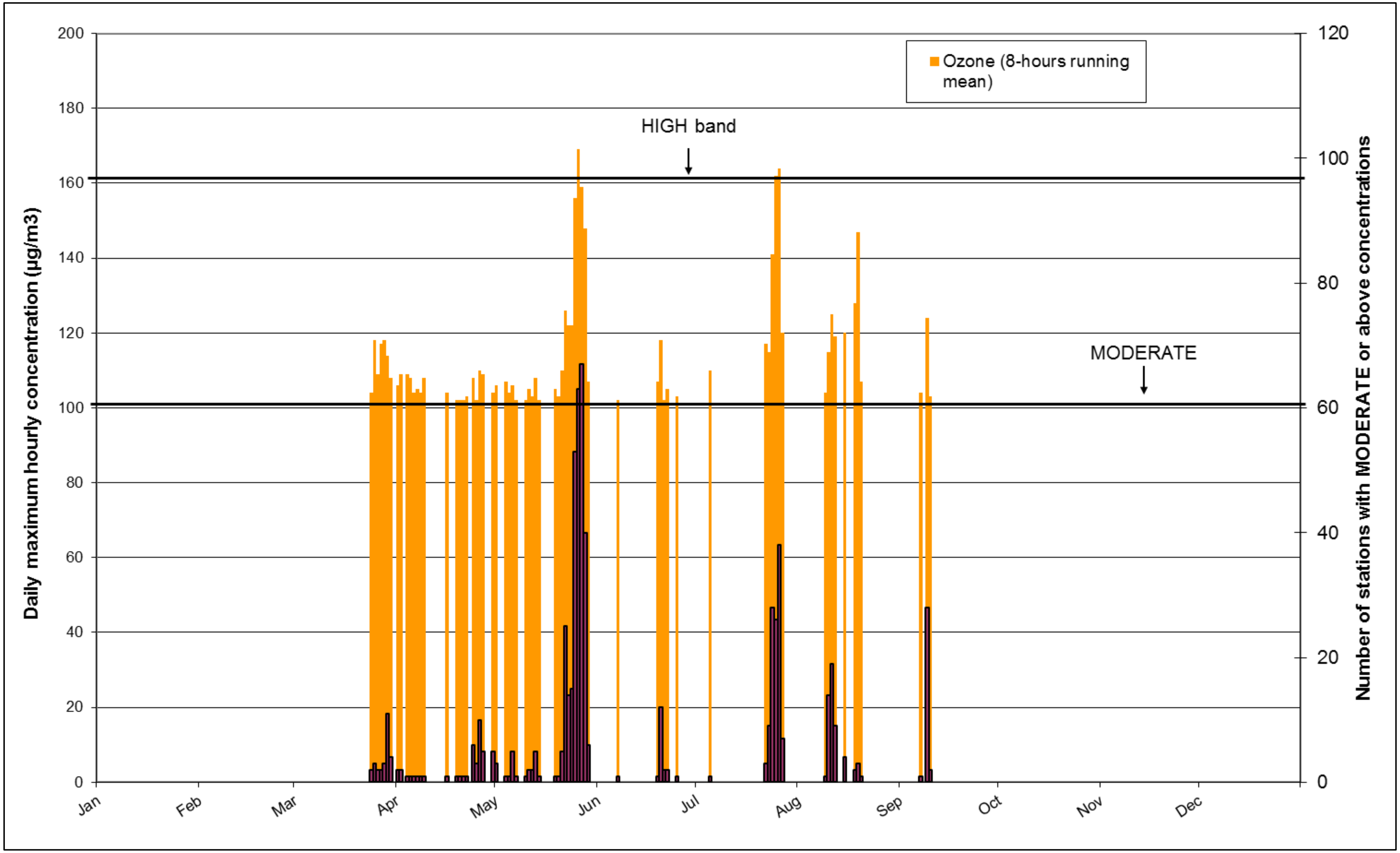


Figure 4-4: ***24-hour mean PM_{10} concentrations across AURN Network with total number of stations measuring "Moderate" or above levels during 2012.***

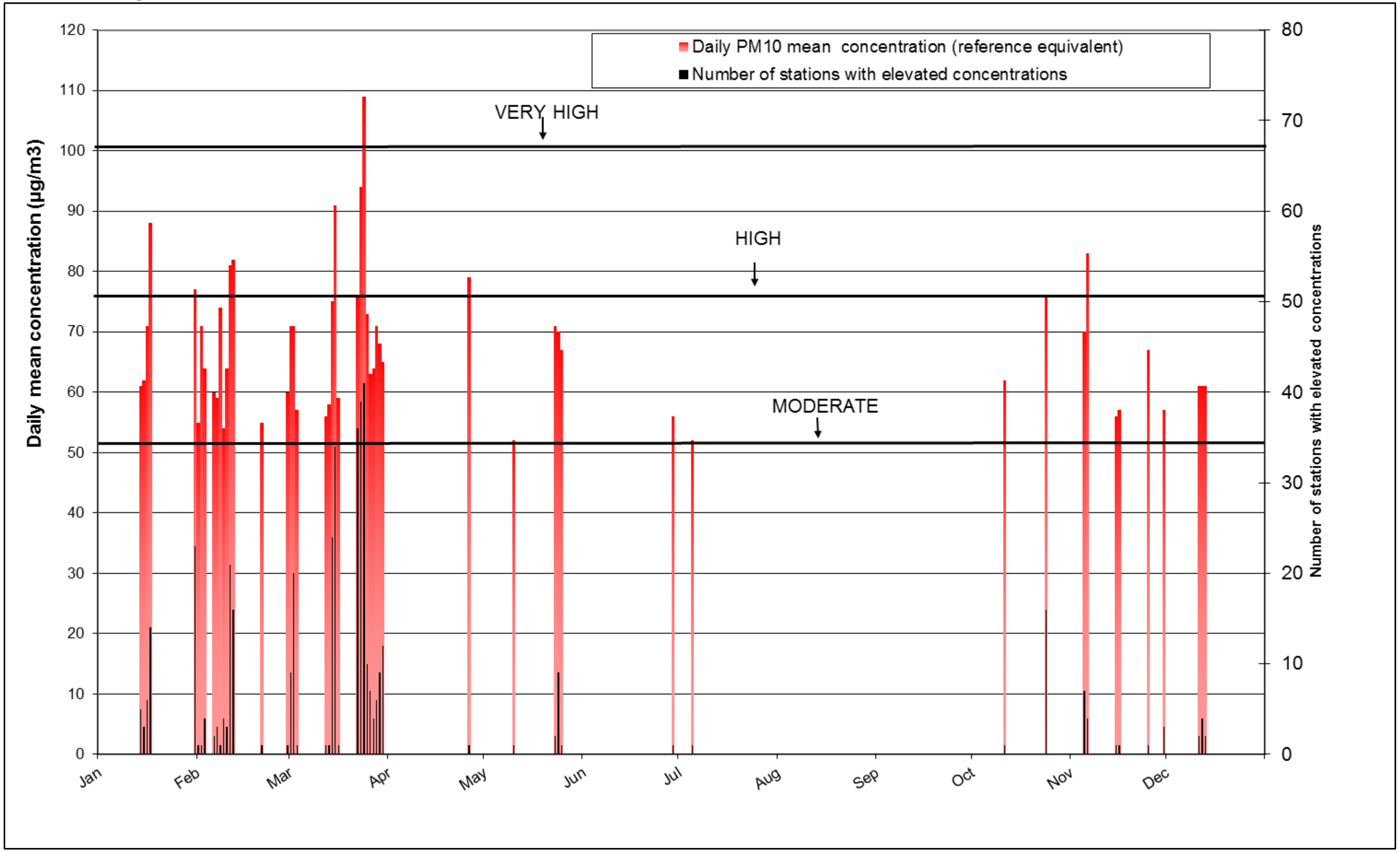


Figure 4-5: 24-hour mean $PM_{2.5}$ concentrations across AURN Network with total number of stations measuring "Moderate" or above levels during 2012.

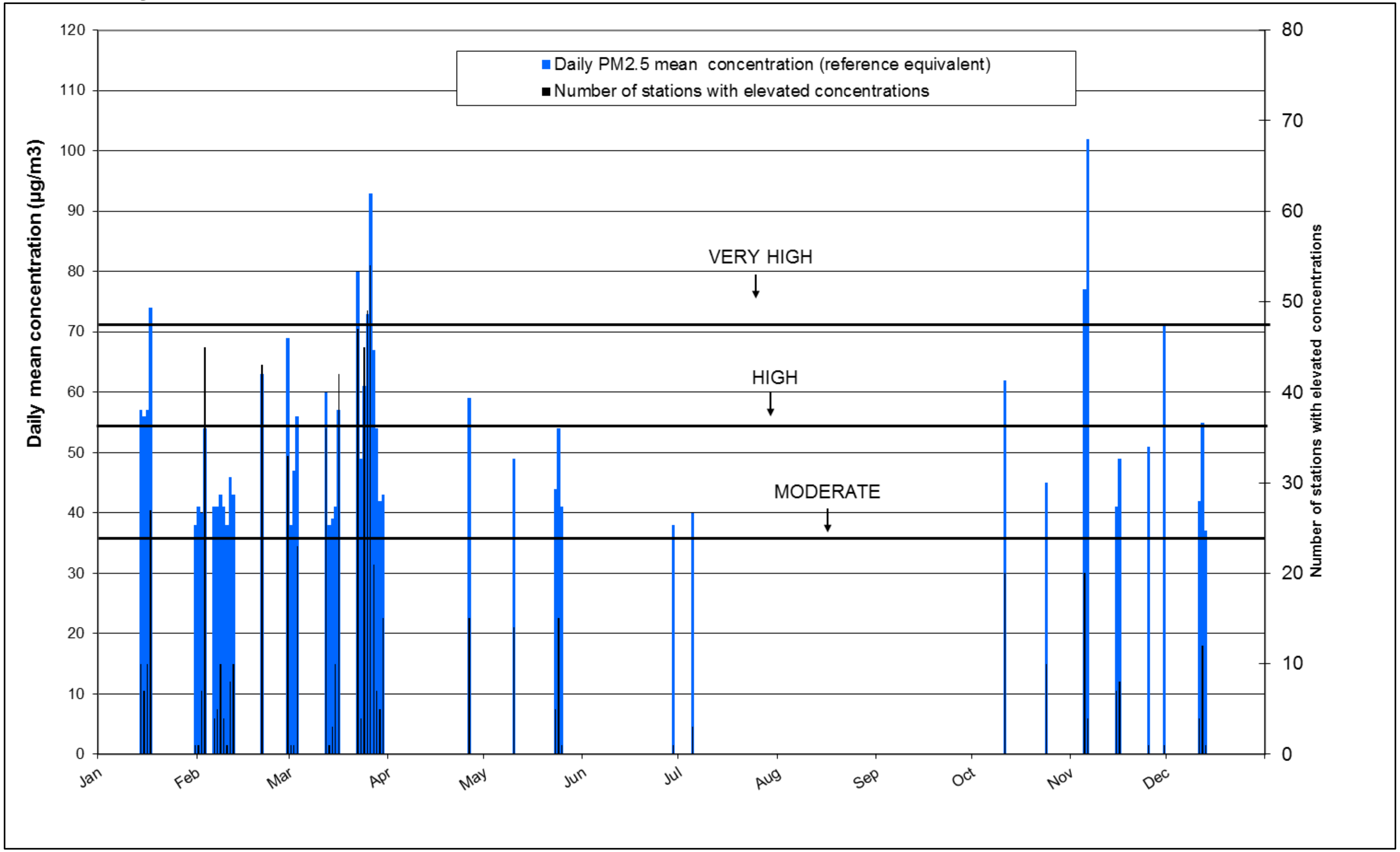


Figure 4-6: Maximum 15 minute average concentrations of SO_2 across AURN Network with total number of stations measuring "Moderate" or above levels during 2012.

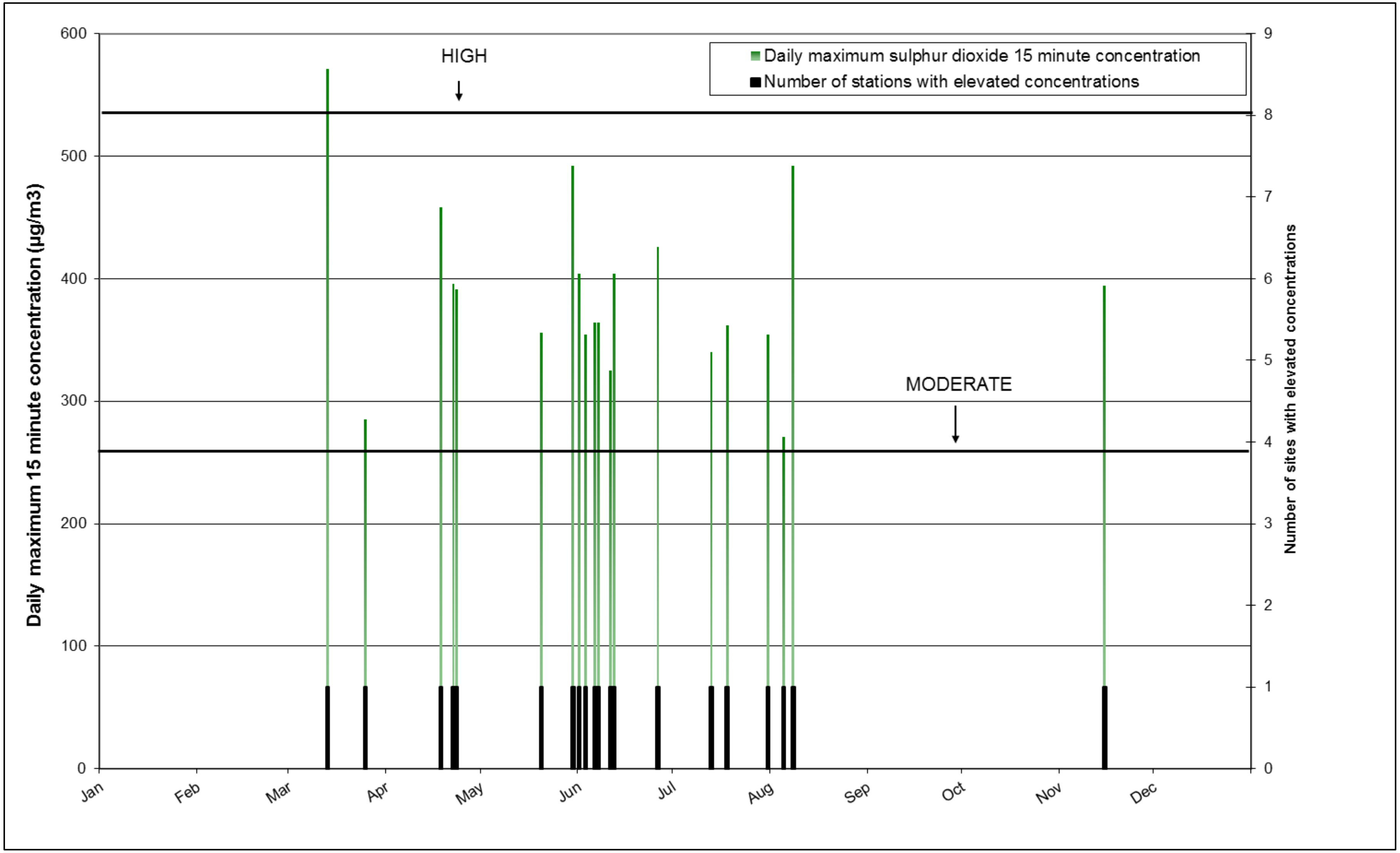


Figure 4-7: Daily Maximum hourly average of NO_2 across AURN Network with total number of stations measuring "Moderate" or above levels during 2012.

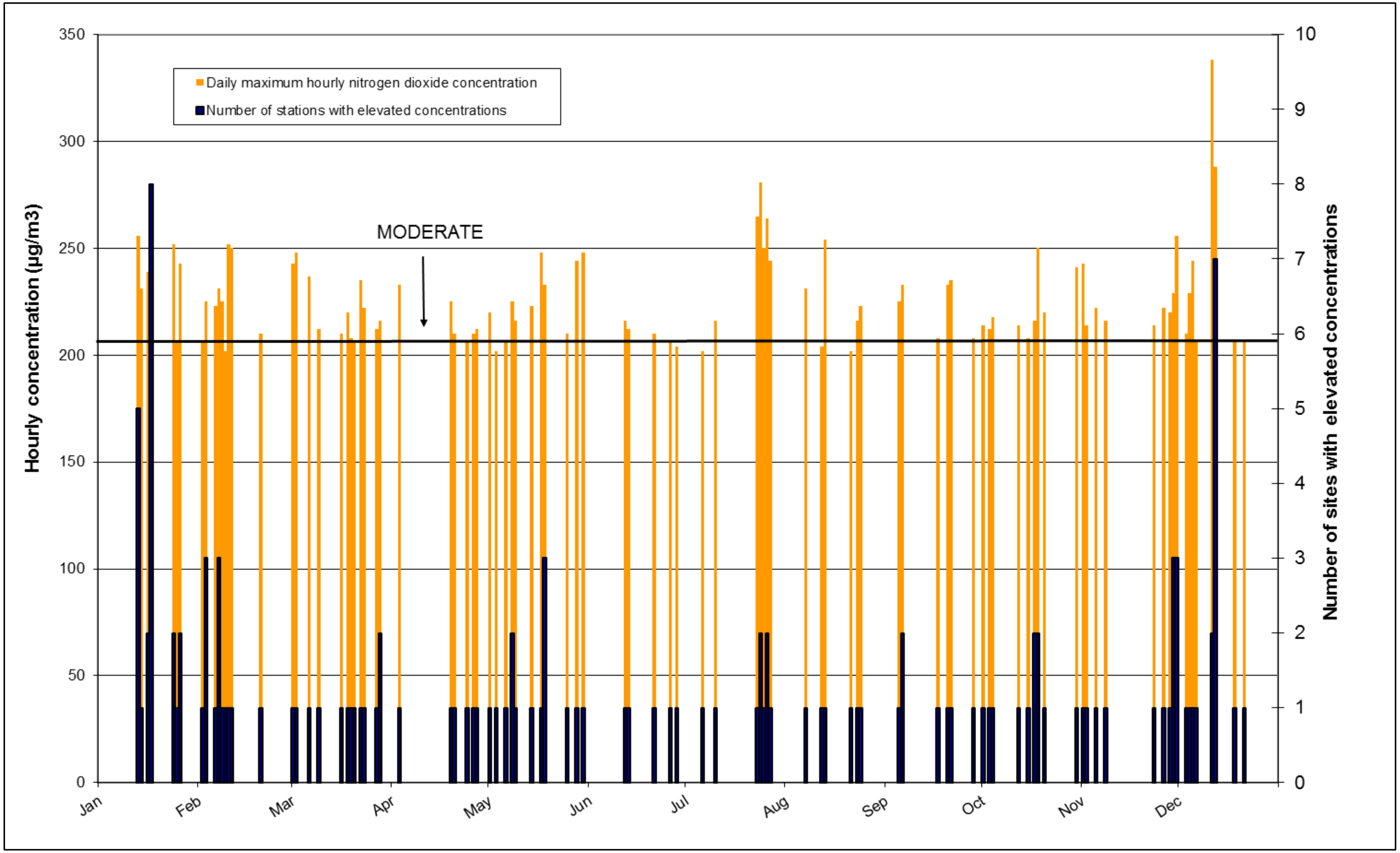


Figure 4-8: Number of pollutant days "Moderate" and above for each AURN Network station during 2012 (site names A-L)

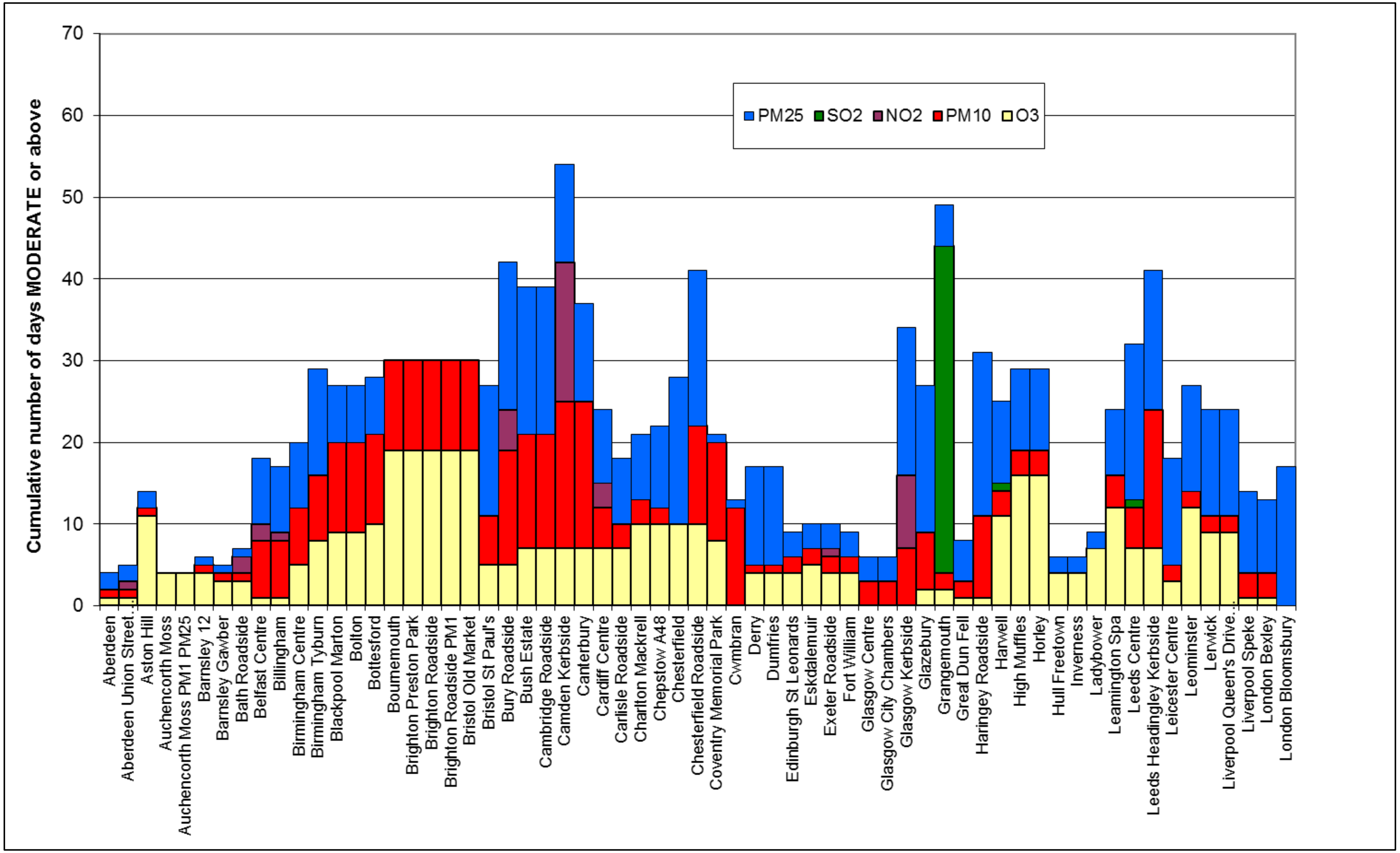
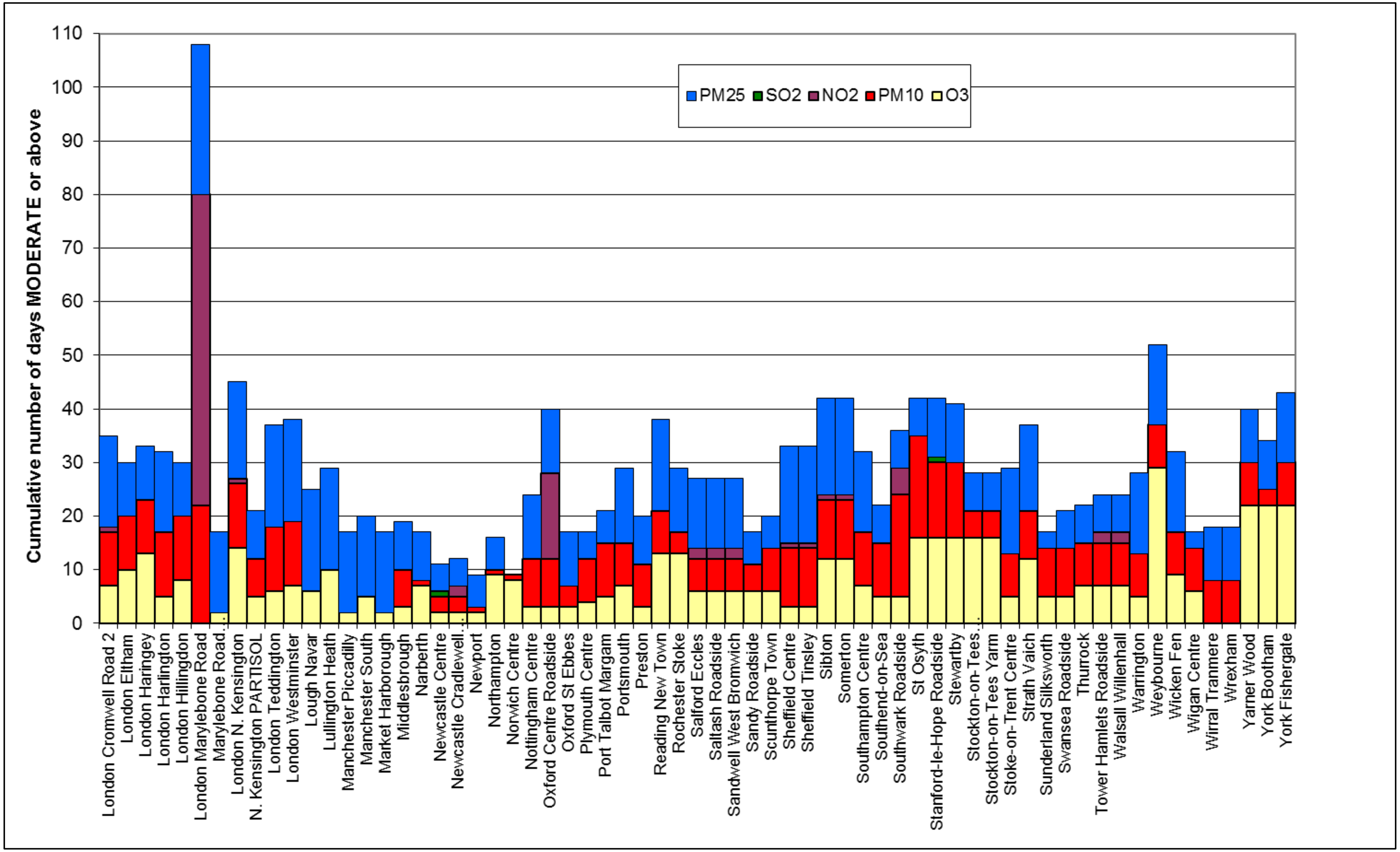


Figure 4-8 (continue): Number of pollutant days "Moderate" and above for each AURN Network station during 2012 (site names L-Y)



4.2.2 Ozone pollution episode summary

Air pollution episodes due to ozone usually occur in late spring and summer. In 2012 the first ozone episode was recorded between 22nd and 28th May. During the first four days of the episode many AURN monitoring sites measured "Moderate" ozone, reaching air pollution index 6 at Northampton.

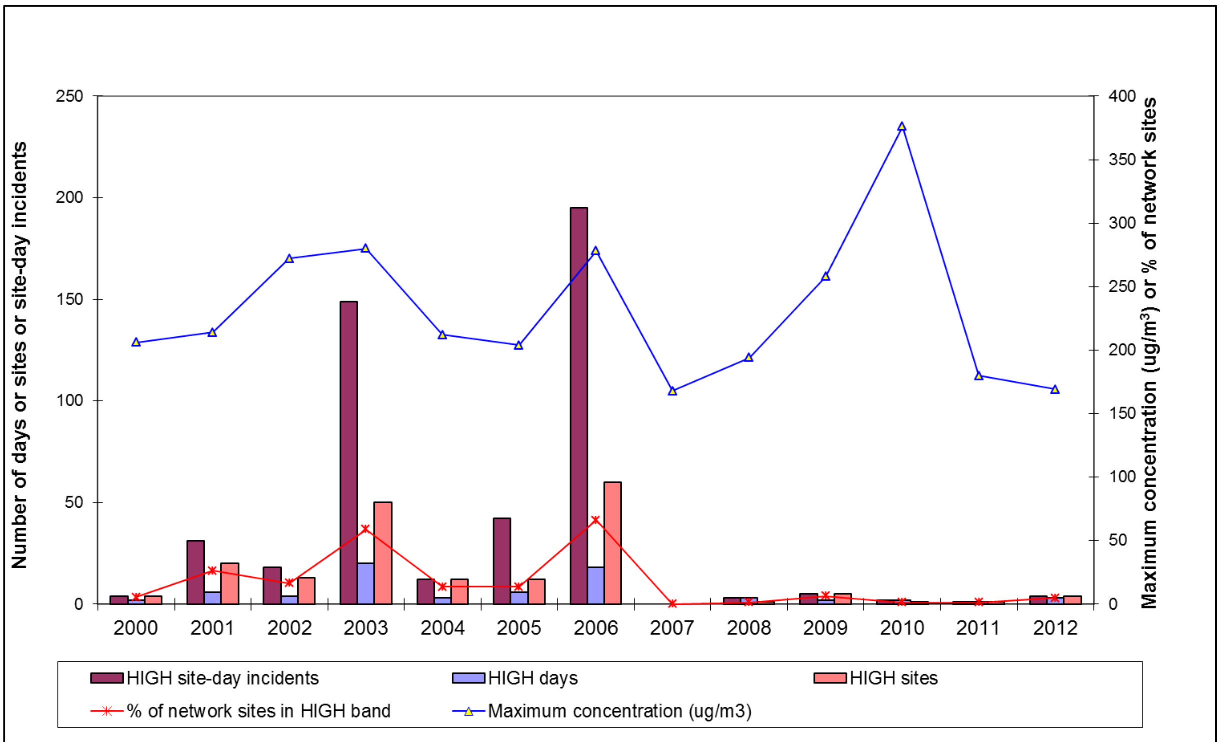
On 26th May 63 AURN sites recorded "Moderate"/"High" ozone across the UK reaching index 7, with the highest 8-hourly ozone concentration of 169 µg/m^3 measured at Yarner Wood. Widespread "Moderate" levels of ozone continued until 28th of May. At this time of the year ozone episodes are often due to the elevated hemispheric background level.

Elevated ozone levels were measured across London and southern England on Wednesday 25th and Thursday 26th July 2012. The Daily Air Quality Index reached band High (index 7) at the following sites: N. Kensington, Sibton, Brighton Preston Park, Northampton Kingsthorpe, Bournemouth and Charlton Mackrell.

There was an unusually late summer pollution episode at the mid-September caused by a period of stable high pressure and exceptionally high temperatures for this time of the year, as shown in Figure 4-3.

Figure 4-9 shows that 2012 was another low year for the number of ozone episodes.

Figure 4-9: UK ozone episodes summarized for years 2000 onwards.



4.2.3 Particulate matter (PM_{10} and $PM_{2.5}$) pollution episode summary

In 2012 there were several periods of elevated levels of particulates recorded by AURN monitoring sites across the UK. These were mainly between January and May and October to December 2012.

During the first episode of 14th to 17th January, 24-hour mean concentrations of PM_{10} and $PM_{2.5}$ reached the "Moderate" band at many AURN sites, with 14 AURN monitoring sites recording "High" on 17th January. This episode was mainly due to calm weather conditions, and a high pressure system centred over the British Isles. There were a number of further periods of

“Moderate” and “High” PM_{10} and $PM_{2.5}$ concentrations measured across the UK from Tuesday January 31st to Monday February 13th.

March began with a period of high pressure and low wind speeds bringing “Moderate” and “High” PM pollution to London, the south-east of England, the Midlands and Wales. Between March 5th and March 12th particulate concentrations remained mainly “Low”. From Tuesday March 13th until Saturday March 17th widespread increases in PM concentrations to “Moderate” and “High” levels were observed. After a brief lull PM concentrations rose again from Thursday March 22nd through to the end of the month, with instances of “Moderate”, “High” and “Very high” pollution observed at a number of different UK locations and varying from day-to-day.

“High” PM_{10} concentrations were measured at Port Talbot Morgam on Friday April 27th, this occurred during a period of rain with very strong south westerly winds from the direction of steelworks.

From Wednesday May 25th PM_{10} and $PM_{2.5}$ concentrations began rising across many areas with an influx of easterly air arriving over the UK. By Friday May 27th PM_{10} and $PM_{2.5}$ concentrations reached “Moderate” levels with “High” $PM_{2.5}$ levels at index 7 recorded at Chatham Roadside.

During autumn the first period of widespread “Moderate” and “High” PM_{10} and $PM_{2.5}$ in the south-east of England occurred between 23rd and 26th October, mainly due to the influx of air arriving from the continental Europe. A similar episode occurred between 15th and 17th November when a period of incoming continental air and low wind speeds brought PM_{10} and $PM_{2.5}$ levels to “Moderate” and “High” across many areas of England and South Wales.

Cold and frosty weather was well established over the UK from 25th November, bringing localised “Very High” $PM_{2.5}$ concentrations to Derry, and then more widespread “Moderate” and “High” levels of particulate matter across central and North-West England by the end of November. A similar pollution episode occurred between Tuesday 11th and 15th December as the cold weather and still weather continued bringing poor dispersion conditions with widespread “Moderate” PM_{10} concentrations reaching “High” across Northern Ireland and Glasgow including a “Very High” period measured at Armagh Roadside.
Figure 4-10 presents the number of sites annually exceeding “Moderate” and “High” PM_{10} bands during the “Bonfire Night” period between 1st and 10th November. Additionally

Figure 4-11 shows the overall number of “Moderate” or worse PM_{10} exceedances annually from all pollution sources from the year 2000 onwards.

Both graphs indicate that 2012 was a relatively low year for PM_{10}pollution episodes.

Figure 4-10: Number of sites exceeding the "Moderate" and "High" PM_{10} bands over 1st November to 10th November annually from the year 2000 onwards with additional descriptive statistics.

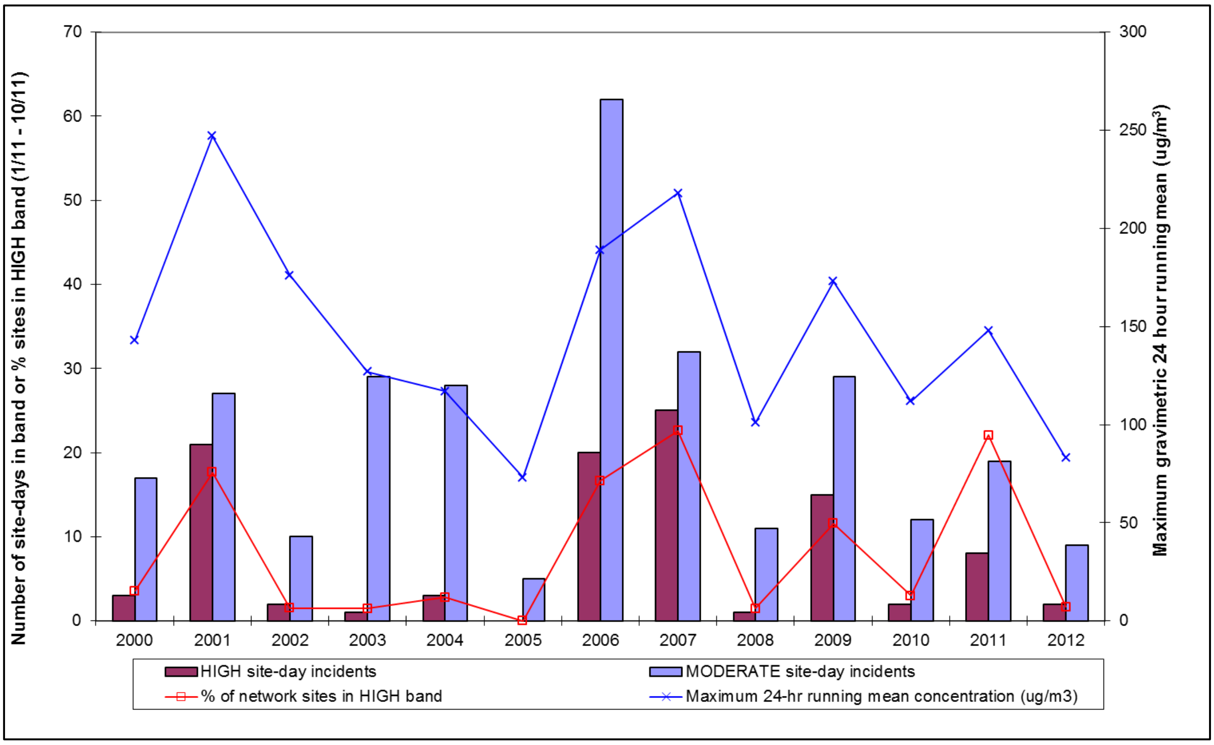
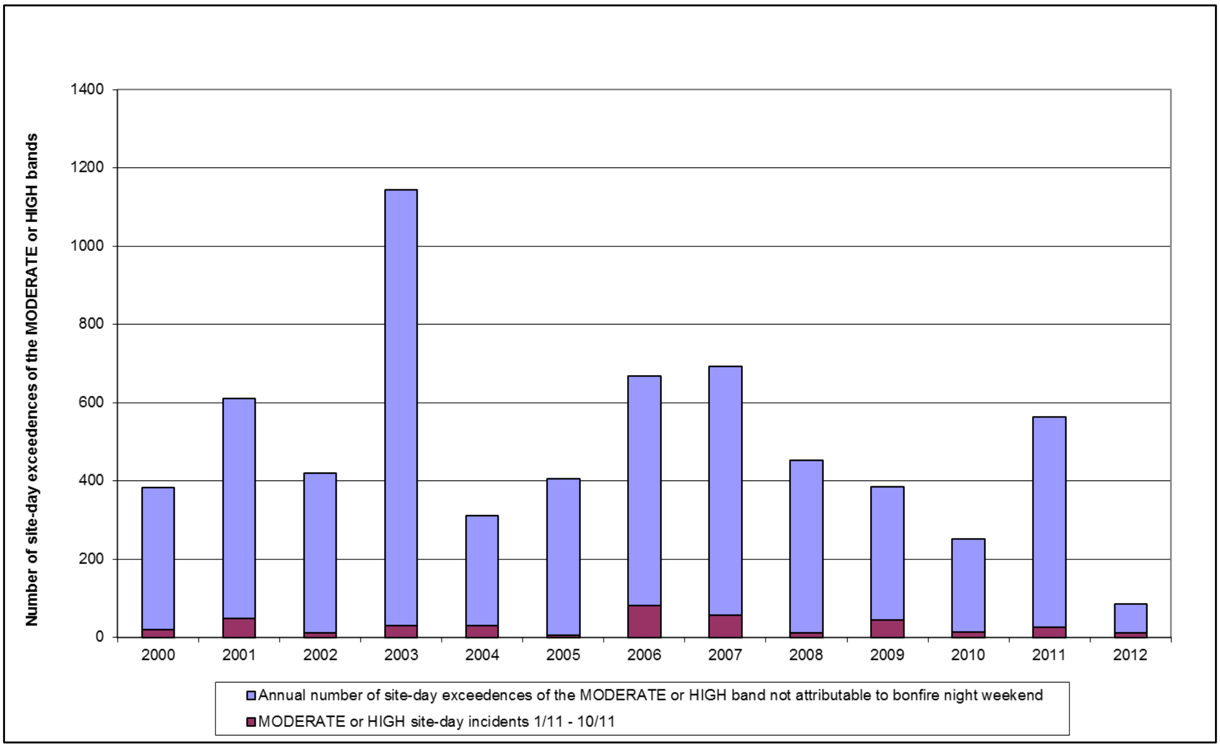


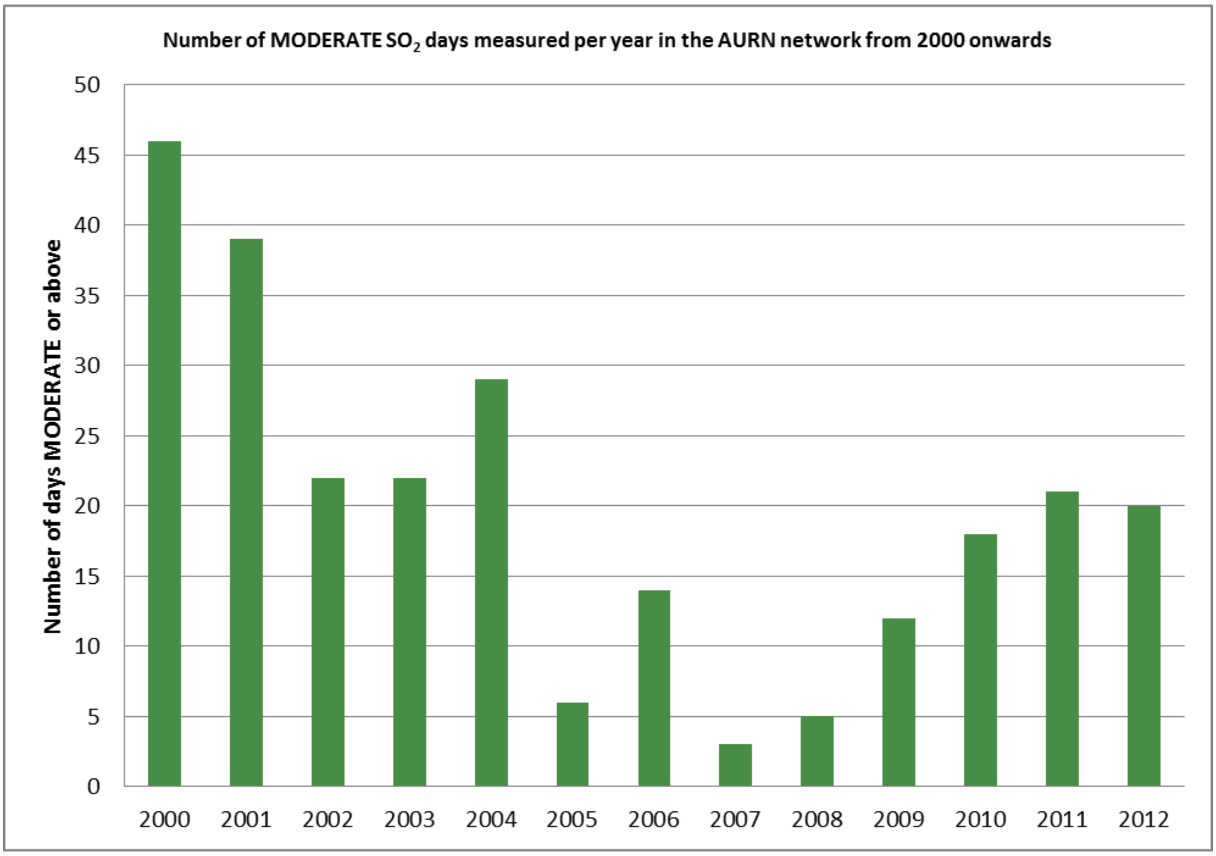
Figure 4-11: Annual number of site-day exceedances of the "Moderate" or "High" PM_{10} band for 2000 – 2012.



4.2.4 Sulphur Dioxide pollution episode summary

There were 46 sites in the AURN air quality monitoring network measuring SO_2 concentrations in 2012. The number of "Moderate" or above days per annum measured in the network is shown in Figure 4-12 from the year 2000 onwards. There were nineteen "Moderate" SO_2 days in 2012. The number of days of "Moderate" exceedances per year is low but has been rising since 2007. The exceedances continue to be at monitoring sites in mainly industrial locations; however in 2012 the highest concentration was recorded at the Harwell rural monitoring site.

Figure 4-12: Number of "Moderate" or worse SO_2 network days measured per annum.



A significant reduction in the number of exceedances over years is likely to be the result of an improvement in and proliferation of abatement technologies to control the release of sulphur dioxide and other pollutant species coupled with a downturn in the use of coal for domestic heating.

4.2.5 Nitrogen Dioxide pollution episode summary

Many "Moderate" days for nitrogen dioxide were measured during the year 2012. In January, February, March, November and December cold and still weather condition combined with the continental easterly airflow contributed to increased levels of NO_2 concentrations at many AURN monitoring sites.

During stable meteorology in April, May, June, July, August and September many kerbside and roadside sites also experienced increased level of NO_2 due to their proximity to road traffic and poor dispersion condition.

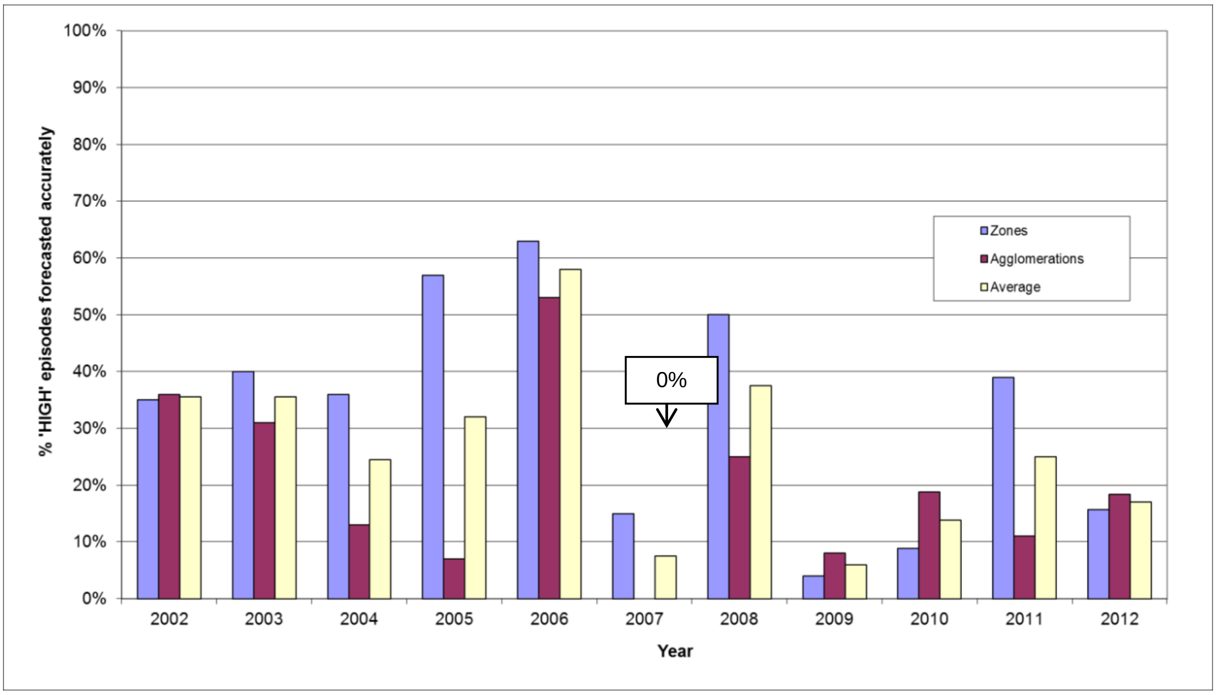
There was no "High" day measured at any AURN site throughout the year 2012.

4.3 Comparison with years 2002 onwards

4.3.1 Overall Forecasting Accuracy Rate

Figure 4-13 shows the forecasting accuracy rates for “High” pollution episodes for the whole of the UK for years 2002 to 2012. This is the percentage of “High” days that were accurately forecast according to the criteria agreed with Defra and specified at the beginning of section 4 of this report.

Figure 4-13: Forecasting Accuracy rate for “High” pollution episodes for the UK, 2002-2012



* 2002 was a partial year for forecasting analysis calculations.

The overall forecasting success rate for the “High” band in 2012 was better than in 2009 and 2010 but down by 8% compared to 2012 at only 17%.

The forecast accuracy rate changes from year to year depending on the type and frequency of pollution episodes. A step change was also expected in 2012 as the DAQI was revised and forecasts for $PM_{2.5}$ were issued for the first time. The forecasting model was less accurate for $PM_{2.5}$ than other pollutants but as described earlier its performance can be improved by data assimilation techniques.

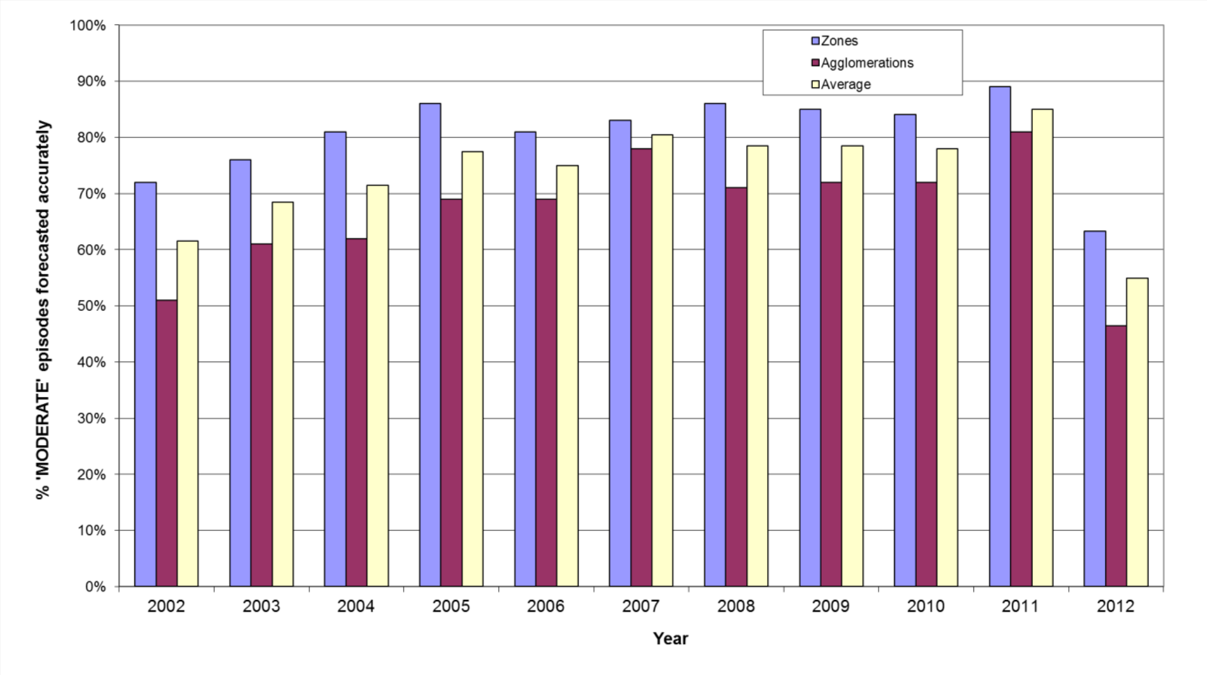
The use of the recommended DAQI trigger values only assists in prediction of pollution episodes later on the same day. The implementation of the triggers as part of forecasting service does not therefore affect the statistics presented here which are for the following days forecast.

In general due to the complex origins of PM pollution our capacity to successfully predict elevated PM levels remains less than that for ozone using the forecast models available.

Because of the infrequent nature of “High” UK pollution episodes in recent years the percentage of “Moderate” days that were accurately forecast is perhaps a better measure of forecast performance.

Figure 4-14 shows that this statistic remained stable or increased slightly over between 2002 and 2011 but again there is a noticeable drop caused by the implementation of the new DAQI in 2012. It is believed that this drop is due mainly to over prediction of moderate ozone pollution episodes which were much reduced in 2012 because of the change in averaging period from 1-hour to 8-hour in the new DAQI.

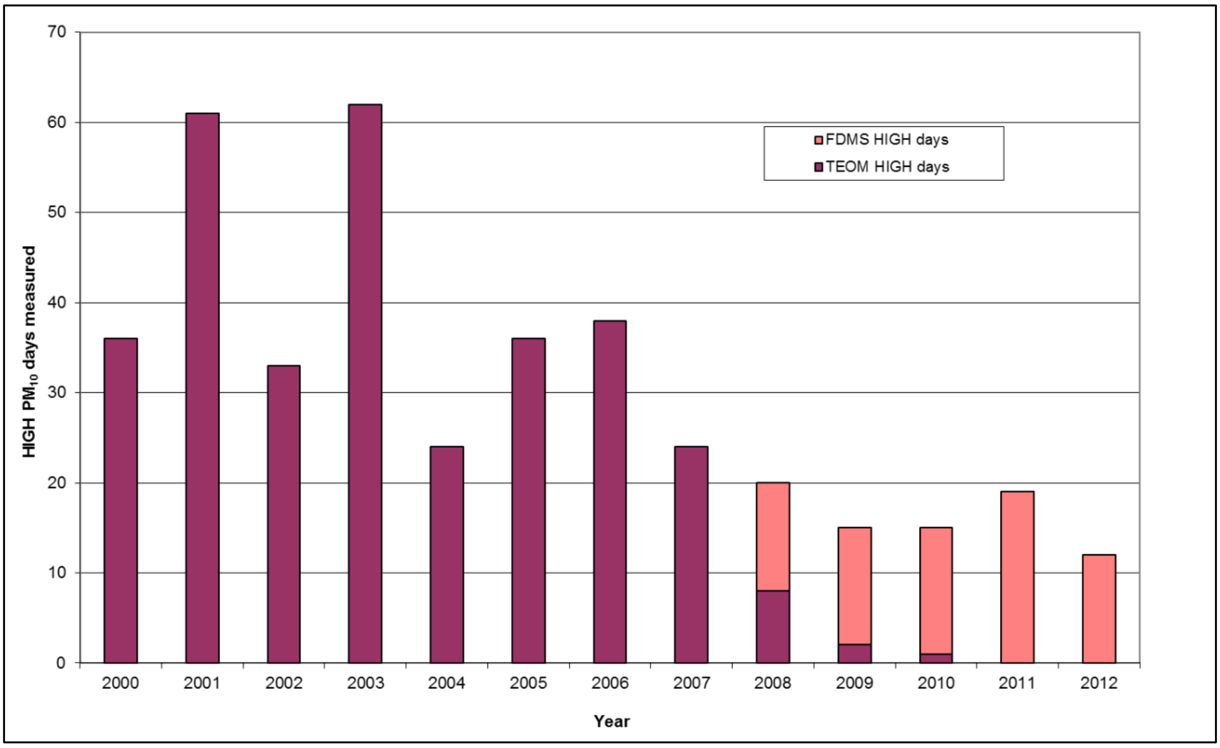
Figure 4-14: Forecasting Accuracy rate for "Moderate" pollution episodes for the UK, 2002-2012.



* 2002 was a partial year for forecasting analysis calculations.

Figure 4-15 below confirms that 2012 was typical of recent years in terms of the overall low number of "High" band PM_{10} measurements recorded.

Figure 4-15: Number of "High" band measurements for PM_{10} in the UK, 2000-2012.



4.3.2 LOCALISED INFLUENCES

In addition to the difficulties of forecasting long range transport of particulates, there are also problems in forecasting accurately in areas where local effects on pollution are significant and unpredictable. Port Talbot Margam is an example of such a site that reported “High” concentrations during 2012.

5 Breakdowns in the service

All bulletins were successfully delivered to Defra’s Data Dissemination Unit (DDU) contractor on time and there were no reported breakdowns in the service over the year.

6 Additional or enhanced forecasts

During the 2012 London Olympics formal enhanced forecasts were issued by the forecasting team. There were a few changes made to the routine circulation of air quality forecast emails. These included:

1) The email messages were circulated to an expanded circulation including those with responsibility for dealing with air quality issues during the Olympics period.

2) The format of the message in the email was changed to include a short introductory summary at the top with more detailed analysis following.

3) There were more details in the forecast for London and south-east England although the whole of the UK situation was also covered.

4) From July 2nd until September 21st the frequency of email updates was increased from twice-weekly to daily (including weekends and bank holidays).

5) From September 21st onwards the frequency of updates returned to bi-weekly.

There were numerous meetings between the Ricardo-AEA forecasters, the HPA and Defra to agree the scope of the forecasting service provided during this period.

The air pollution forecast is always re-issued to Web and Freephone services at 10.00 a.m. local time each day, but this is only updated when the pollution situation is changing.

The bi-weekly air pollution outlooks have continued to be delivered successfully to Defra and other government departments by email on Tuesdays and Fridays.

7 Ad-hoc Services

A paper on validation of WRF meteorological forecasts vs. measured data was prepared using ad-hoc funding http://uk-air.defra.gov.uk/reports/cat20/1310100848_Evaluation_of_meteorological_data_for_UK_forecasting.pdf.

8 On-going Research

Ricardo-AEA continues to develop the air quality forecasting systems by:

1. Investigating ways of using automatic software systems to streamline the activities within the forecasting process, thus allowing forecasters to spend their time more efficiently considering the most accurate forecasts.
2. Researching the chemistry used in our models, in particular the CMAQ chemical schemes for secondary PM_{10} and ozone.
3. Improving the automated validation analysis and plots.
4. Improving and updating the emissions inventories used in our models.
5. Considering how best to implement the findings of the CMAQ data assimilation study within the real-time forecasting service.

The UK Air Quality Forecasting project maintains close links with the Defra CMAQ model development project in order that real-time operational and off-line developments are closely aligned and optimised wherever possible.

9 Project and other related meetings

9.1 Project meetings

Regular six-monthly project meetings continued to be held at Harwell and London over the course of the year.

9.2 CMAS

The message from the Keynote address by Dr. Len Peters (Secretary of the Cabinet for Energy and Environment, Commonwealth of Kentucky) was:

Models should be used for guidance and support, and they can provide a means of understanding the science when experiments are not possible. However they are often too complicated for the non-expert, and require good data interpretation. Air Quality models cannot have all the answers particularly when some of the fundamental problems of meteorology models have not been resolved e.g. the physics of turbulence.

A model is a compass......not a GPS - it gives direction not an absolute.

In 2011 there was a special session on forecasting and introduced a new version of CMAQ, the 2012 meeting was more general. These are the key points to note:

- Air Quality Forecasting: Spain has upgrade from CMAQv4.6 to CMAQv5.01 improved the forecast particularly the diurnal effect for $PM_{2.5}$ and ozone.
- Emissions:
    - Aviation Environmental Design Tool (AEDT) – to simulate aircraft take-off, cruising and landing emissions within the model grid.
    - SMOKE-MOVES have developed surrogates for road types that can be used to create road traffic emissions. The general view is that using the MOVES factors improves the emissions but it is to slow as an explicate model.
    - Point sources, improvement to the processing time for plume dispersion and have added aerosol to the in-plume chemistry for planned release in 2013.
- Model resolution: there was a recurring theme that as spatial and temporal resolution increases so does the emission uncertainty. For modelling applications it is a balance for resolution / uncertainty e.g. for health studies where AQ required at high resolution but emissions are more uncertain.
- Wild fires: over the last few years there has been more interest in wildfires, this was particularly so this year with an increase in fires this summer
- Evaluation: measurements of ozone and $PM_{2.5}$ relevant to air quality policy were well predicted over multi-year analysis but there was a systematic bias.

10 Related projects

Ricardo-AEA ensures that any forecasts, issued under separate contracts, are consistent with the national forecasts issued for Defra and the Devolved Administrations.

UK regional forecasts are used in the uBreathe iPhone app which was launched in 2012 and is available at http://www.ricardo-aea.com/ubreathe/.

Forecasts for Scottish regions are used to populate the Scottish Government's Know & Respond SMS Alert service. Please see http://www.scottishairquality.co.uk/know-and-respond/ for further information.

The KentAir forecast has continued to be issued as a short piece of descriptive text detailing the pollution levels expected in the Kent area for the current and following day. In addition to the AURN network sites, air quality levels measured at sites in the Kent AQ network are also taken into account when making an assessment of the forecast for the region. The forecast issued is also sent to the KentAir website at http://www.kentair.org.uk.

11 Scientific Literature Review

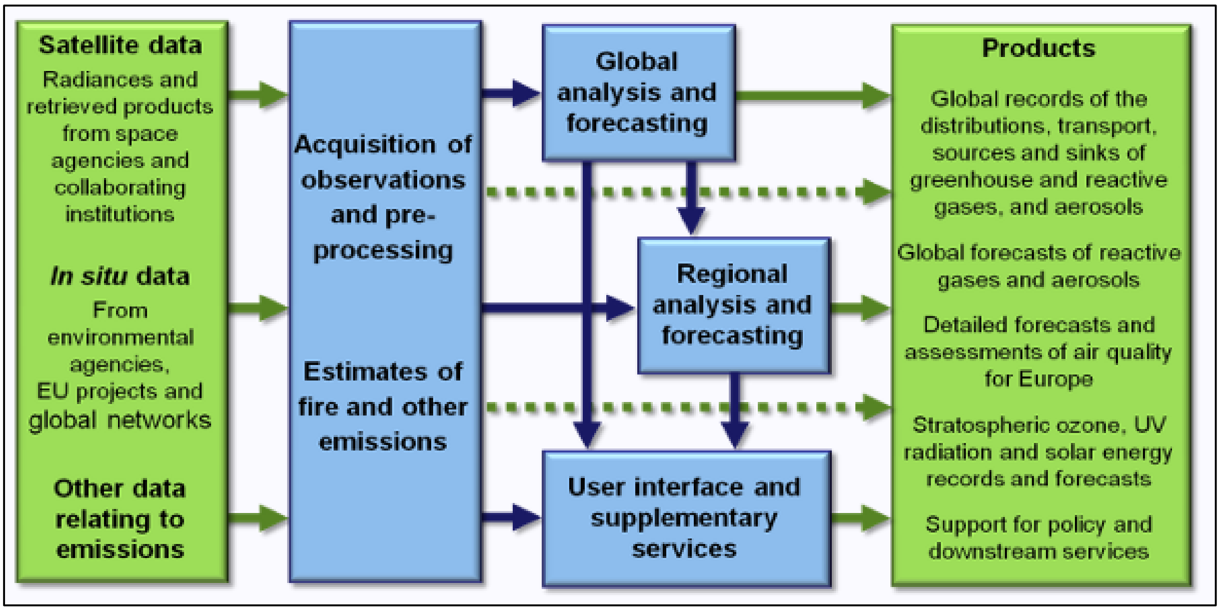
This section reviews a selection of the scientific literature available in the public domain that is relevant to air quality forecasting in 2012.

Recent developments concerned with air quality forecasting are summarised below, with relevant internet links provided at the end of each section.

11.1 MACC-II - Monitoring Atmospheric Composition and Climate - Interim Implementation

MACC-II - Monitoring Atmospheric Composition and Climate - Interim Implementation - is the project that is establishing the core global and regional atmospheric environmental services delivered as a component of Europe's GMES initiative. It is funded under the Seventh Framework Programme of the European Union and began on 1 November 2011. MACC is undertaken by a consortium drawn largely from the partners in the earlier MACC project, whose core systems and service lines provided the starting point for MACC-II.

Figure 11-1: Schematic illustration of data flow



MACC-II takes as its input comprehensive sets of satellite data from many tens of instruments supplying information on atmospheric dynamics, thermodynamics and composition, made available by space agencies and institutions with which the agencies collaborate to produce retrieved data products. The satellite data are supplemented by in-situ data from meteorological networks and a limited amount of data from networks providing in-situ measurements of atmospheric composition. Data are processed to provide a range of products related to climate forcing, air quality, stratospheric ozone, UV radiation at the earth's surface and resources for solar power generation. Additional in-situ data are used for validating the processing systems and the products they supply. MACC operates a value-adding chain which extracts information from as wide a range of observing systems as possible and combines the information in a set of data and graphical products that have more complete spatial and temporal coverage and are more readily applicable than the data provided directly by the observing systems.

http://www.gmes-atmosphere.eu/about/

11.2 Comprehensive Modelling of the Earth System for Better Climate Prediction and Projection (COMBINE)

The European integrating project COMBINE brings together research groups to advance Earth system models (ESMs) for more accurate climate projections and for reduced uncertainty in the prediction of climate and climate change in the next decades. COMBINE will contribute to better assessments of changes in the physical climate system and of their impacts in the societal and economic system. The proposed work will strengthen the scientific base for environmental policies of the EU for the climate negotiations, and will provide input to the IPCC/AR5 process.

http://www.combine-project.eu/

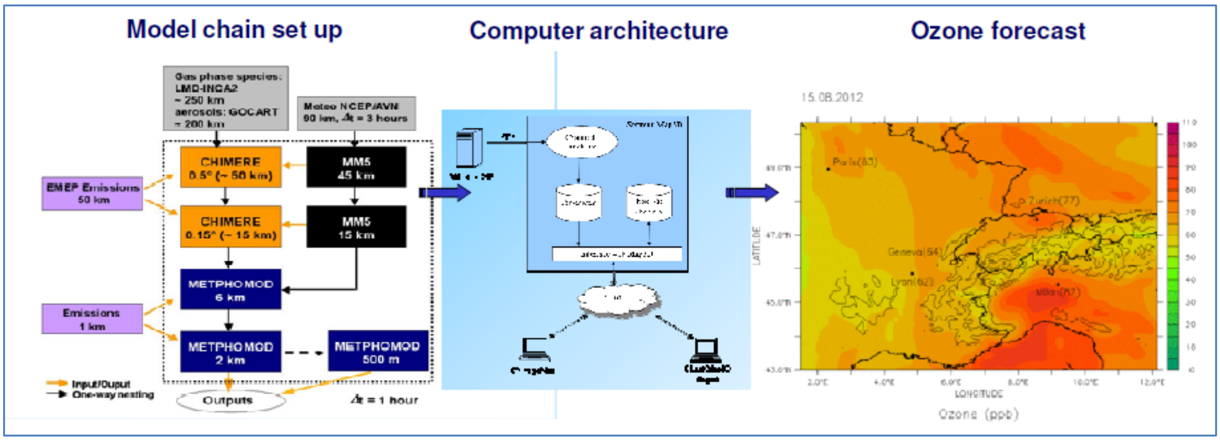
11.3 AIRNow-International

The goal of AIRNow-International is to strengthen relationships among governments and international organizations by sharing the technology to transform air quality data into vital information. AIRNow-International is poised to become the centerpiece of the United States Environmental Protection Agency's (EPA) real-time air quality reporting and forecasting program. The system is a redesign of the AIRNow information technology infrastructure that distributes current air quality information for the United States and Canada. The AIRNow-International software suite is being built to support and embrace the Global Earth Observation System of Systems (GEOSS) concept. The new U.S. EPA AIRNow system, which became operational in Spring 2009, is based on the AIRNow-International system software but with an added forecasting module to store the forecast information provided by U.S. air agencies.
http://www.earthzine.org/2010/01/25/airnow-international-the-future-of-the-united-states-real-time-air-quality-reporting-and-forecasting-program-with-geoss-participation

11.4 Gaiasens modeling tool MAP3D

GAIASENS combines measurements and modelling environmental phenomenons. One of the unique function of the gaiasens tool is the use of the meteorological real time measurements in extreme regions. The information from these measurements helps authorities to better understand and manage natural risks. GAIASENS provides air quality forecast and also allows investigating and analysing periods of “High” air pollution. It was developed at the Swiss Institute of Technology thanks to an INNOGRANT awards. Air pollution forecast over the Alps and Europe are calculated daily in partnership with IUCN and are freely available through GAIASENS Map3D interface.

Figure 11-2: Map3D structure



http://map3d.iucn.org/map3d/pmwiki/uploads/Main/poster_map3d_iwaqfr_v8.pdf

12 Forward work plan for 2013

The two tables below summarise both the weekly and annual planned activity for 2013 (Table 12-1 and 12-2 respectively).

Table 12-1: Weekly Activity Chart

Task 1	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Daily Forecast							
Forecast Outlook Summary							

Table 12-2 Annual Activity Chart

Task 2	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Monthly Reports												
Six-monthly Progress Meetings												
Annual reports												

13 Hardware and software inventory

Defra and the Devolved Administrations have funded the development of the WRF and CMAQ models for UK Air Quality Forecasting purposes. Defra and the Devolved Administrations also own the web pages used to display the forecasts.

No computer hardware being used on this project is currently owned by Defra and the Devolved Administrations.

14 References/Internet links

UK Air Quality Forecasting reports on the UK-AIR library:

http://uk-air.defra.gov.uk/library/

www.cmaq-model.org

www.wrf-model.org

http://www.rmets.org/

Atmospheric Environment Journal:

http://www.uea.ac.uk/~e044/ae_newpages/atmosenv.html

The KentAir website:

http://www.kentair.org.uk/pollutionlevels.php

Agnew et al. 2007 Evaluation of GEMS Regional Air Quality Forecasts http://gems.ecmwf.int/do/get/PublicDocuments/1533/1402

http://www.meas.ncsu.edu/aqforecasting/research.html

http://www.cerc.co.uk/air-quality-forecasting/austria.html

http://web.t-online.hu/dasy/forecast/Budapest.htm

http://www.earthzine.org/2010/01/25/airnow-international-the-future-of-the-united-states-real-time-air-quality-reporting-and-forecasting-program-with-geoss-participation

http://www.combine-project.eu/

http://www.gmes-atmosphere.eu/

Williams et al. 2011, Review of Air Quality Modelling in Defra

Lingard et al. 2013, Statistical evaluation of meteorological data used for UK air quality forecasting. AEAT/ENV/R/3273_ED48946_Issue Number 1http://uk-air.defra.gov.uk/reports/cat20/1310100848_Evaluation_of_meteorological_data_for_UK_forecasting.pdf

Derwent et al. 2009, Evaluating the Performance of Air Quality Models. AEAT/ENV/R/2873 - Issue 2

Appendices

Appendix 1: UK Air Pollution Index

Appendix 2: UK Forecasting Zones and Agglomerations

Appendix 3: Development of Data Assimilation in UK Air Quality Forecasting

Appendix 1 – UK Air Pollution Index

CONTENTS

The Committee on the Medical Effects of Air Pollutants (COMEAP) published its Review of the UK Air Quality Index in June 2011. Following these recommendations, Defra and the Devolved Administrations implemented a new index The Daily Air Quality Index (DAQI) from 1st January 2012.

Banding	Index	Ozone	Nitrogen Dioxide	Sulphur Dioxide	$PM_{2.5}$ Particulates (EU Reference Equivalent)	PM_{10} Particulates (EU Reference Equivalent)
		Running 8-hourly mean	Hourly mean	15-minute mean	24-hour mean	24-hour mean
		μgm^{-3}	μgm^{-3}	μgm^{-3}	μgm^{-3}	μgm^{-3}
Low	1	0-33	0-67	0-88	0-11	0-16
	2	34-66	68-134	89-177	12-23	17-33
	3	67-100	135-200	178-266	24-35	34-50
Moderate	4	101-120	201-267	267-354	36-41	51-58
	5	121-140	268–334	355-443	42-47	59-66
	6	141-160	335-400	444-532	48-53	67-75
High	7	161-187	401-467	533-710	54-58	76-83
	8	188-213	468-534	711-887	59-64	84-91
	9	214-240	535-600	888-1064	65-70	92-100
Very high	10	$\geq$ 241 μgm^{-3}	$\geq$ 601 μgm^{-3}	$\geq$1065 μgm^{-3}	$\geq$ 71 μgm^{-3}	$\geq$ 101 μgm^{-3}

Banding	***Index***	***Accompanying health messages for at-risk individuals****	***Accompanying health messages for the general population***
Low	1	**Enjoy** your usual outdoor activities.	**Enjoy** your usual outdoor activities.
	2		
	3		
Moderate	4	Adults and children with lung problems, and adults with heart problems, **who experience symptoms**, should **consider reducing** strenuous physical activity, particularly outdoors.	**Enjoy** your usual outdoor activities.
	5		
	6		
High	7	Adults and children with lung problems, and adults with heart problems, should **reduce** strenuous physical exertion, particularly outdoors, and particularly if they experience symptoms. People with asthma may find they need to use their reliever inhaler more often. Older people should also **reduce** physical exertion.	Anyone experiencing discomfort such as sore eyes, cough or sore throat should **consider reducing** activity, particularly outdoors.
	8		
	9		
Very high	10	Adults and children with lung problems, adults with heart problems, and older people, should **avoid** strenuous physical activity. People with asthma may find they need to use their reliever inhaler more often.	**Reduce** physical exertion, particularly outdoors, especially if you experience symptoms such as cough or sore throat.

Appendix 2 – UK Forecasting Zones and Agglomerations

CONTENTS

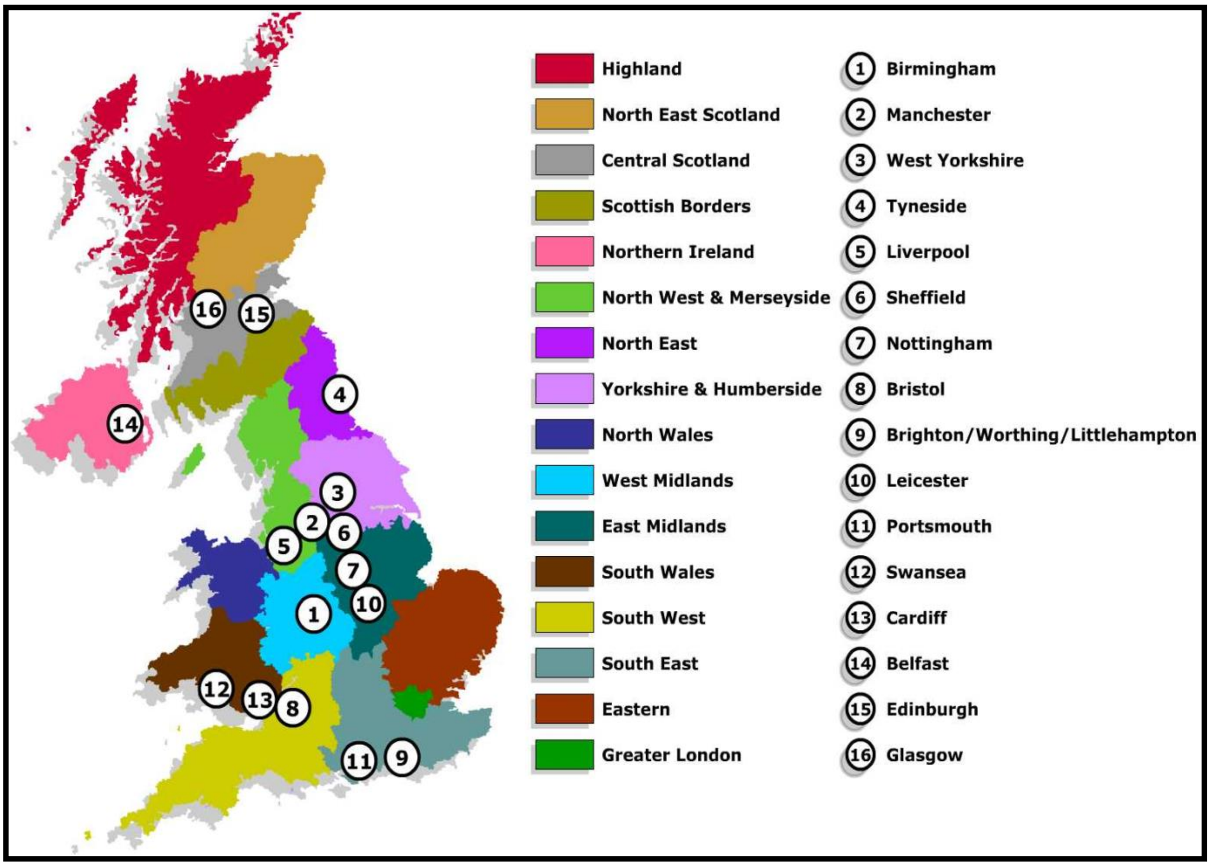
UK Forecasting Zones

Zone	*Population*
East Midlands	3519887
Eastern	5397646
Greater London	8864078
North East	1529031
North West and Merseyside	3560388
South East	6739450
South West	4403790
West Midlands	2738539
Yorkshire and Humberside	3169623
South Wales	1785578
North Wales	749613
Central Scotland	1929140
Highland	393901
North East Scotland	1120962
Scottish Borders	265338
Northern Ireland	1278732

UK Forecasting Agglomerations

Agglomeration	*Population*
Brighton/Worthing/Littlehampton	413892
Bristol Urban Area	533341
Greater Manchester Urban Area	2042681
Leicester	428132
Liverpool Urban Area	736751
Nottingham Urban Area	587683
Portsmouth	389220
Sheffield Urban Area	566262
Tyneside	758610
West Midlands Urban Area	2277340
West Yorkshire Urban Area	1276045
Cardiff	321769
Swansea/Neath/Port Talbot	201718
Edinburgh Urban Area	469052
Glasgow Urban Area	1117379
Belfast	532080

Map of UK forecasting zones and agglomerations



Appendix 3: Development of Data Assimilation in UK Air Quality Forecasting

Introduction

The air quality forecasts prepared by Ricardo-AEA use a variety of methods to predict pollutant concentrations in the UK each day and for the following two days. The methods used include the use of the CMAQ air quality model and the analysis of the trends in observed concentrations. The modelled concentrations provided by the CMAQ model differ from the observed concentrations, usually by a small amount. The aims of this project are

- to investigate how previous forecasts and observations can be used to provide improved estimates of future concentrations, and
- to demonstrate the effectiveness of proposed methods

Different sophisticated approaches exist, such as four-dimensional variational data assimilation (4DVAR) and Ensemble Kalman Filtering (EnKF). While these techniques are potentially very powerful, they are also highly computation-intensive, requiring either the implementation of a model adjoint, or the simultaneous integration of several tens of model ensemble members.

In recent years, rather simple bias adjustment techniques have emerged, in which the bias correction factors are estimated by means of the Kalman filter (KF) approach. These techniques are applied in post-processing (i.e., off-line) mode rather than as a part of the initialization of the deterministic forecast, and they are characterized by a very low computational cost.

The US EPA currently implements bias adjustment of its ozone and particulate matter forecasts using a Kalman filter approach. They have provided Ricardo-AEA with computer codes used to implement the Kalman filter algorithm. The EPA uses forecast and observed hourly concentrations for the previous two days at monitoring sites. It uses the Kalman filter approach to estimate a bias adjustment at each of the monitoring sites for each hour of the current day based on the bias adjustment calculated for the same hour of previous days. The computer codes have been rewritten and adapted for use in UK forecasting in this project, taking into account the different formatting of the UK data. The results of the application of the US EPA Kalman filter algorithm are described in this report.

This project also considered the application of a simpler, basic Kalman filter to provide bias adjustment for daily mean and daily maximum ozone and particulate matter, $PM_{2.5}$ concentrations. The results of this analysis are described first.

Basic Kalman filter bias adjustment

Basic Kalman filter bias adjustment was applied to measured and modelled daily mean ozone and $PM_{2.5}$ concentrations. It was applied separately and individually at monitoring sites throughout the UK. For each monitoring site, the daily mean concentrations were calculated. The observed bias for each day, y, was then calculated as the difference between the modelled and measured daily mean concentrations. The observed bias values can change substantially from day to day. The Kalman filter calculates an estimate of the underlying bias, x_k for day k from the measured and modelled concentrations:

$$x_k = x_{k-1} + k_{gk}(y_{k-1} - x_{k-1})$$

where k_g is a gain factor in the range 0-1.

If the gain factor is set to one, the calculated underlying bias is equal to the previous days observed bias. If the gain factor is set to zero, the underlying bias calculation ignores the observed bias values. The Kalman filter method calculates an "optimal" gain. The optimal gain takes account of the relative sizes of the variance in the process noise and the variance of the observation noise. In this context, the process noise can be thought of as the error associated with the uncertainty in the CMAQ model parameters and inputs. The observation noise includes measurement errors and errors arising from factors that are not taken into account in the CMAQ model. In practice, the relative size of the variances is only known approximately and so the choice of the gain requires some judgement. For this assessment, the gain was calculated assuming that the variance of the process noise was 0.06 times the variance of the observation noise, based on the US EPA algorithm.

$PM_{2.5}$ pollution events in February and March 2013

"Moderate" particulate pollution events in mid-February and a "High" $PM_{2.5}$ pollution event between 5th and 8th March 2013 provided a good test case for evaluating the basic Kalman filter methodology. Fig. A3.1 shows a scatter plot of the predicted daily mean concentration plotted against the observed values for all urban background sites throughout the UK for the period 1 January 2013 to 15 March 2013. The scatter plots show that after filtering, the predictions tend to overestimate the concentration for low concentrations of $PM_{2.5}$ and underestimate the concentrations when the observed concentration of $PM_{2.5}$ is "High". However, it is clear that applying the Kalman filter to the 24 hour average data improves the agreement between predicted and observed data. Table A3.1 provides a statistical summary of the data for the urban background sites and also for urban and rural sites. It shows the mean bias and the standard deviation of the bias. The mean bias (observed – modelled) is effectively eliminated when the Kalman filter is applied. The scatter is also reduced with the standard deviation of the difference between the observation and prediction falling by ~7-10%.

Figure A3.2 shows a similar scatter plot of daily maximum 1-hour average concentration data. Table A3.1 also includes summary statistics for the daily maximum 1-hour average concentrations. The mean bias is effectively eliminated following application of the Kalman filter. As with the daily average data the scatter is also reduced with the standard deviation of the difference between the observation and prediction falling by ~30%. However, examination of Fig. A3.2 indicates that the use of the filter increases the number of outliers where the predictions fail to match the observations.

Figure A3.1: Scatter plots showing correlation between observations and predictions of daily average $PM_{2.5}$ at urban background sites between 01/01/13 and 15/03/13. The left hand plots shows the predicted data before (black circles) and after (red dots) adjustment by Kalman filter. The right hand density scatter plot shows the same adjusted data, but shows the density of sites, with high densities indicated by red and low densities by yellow.

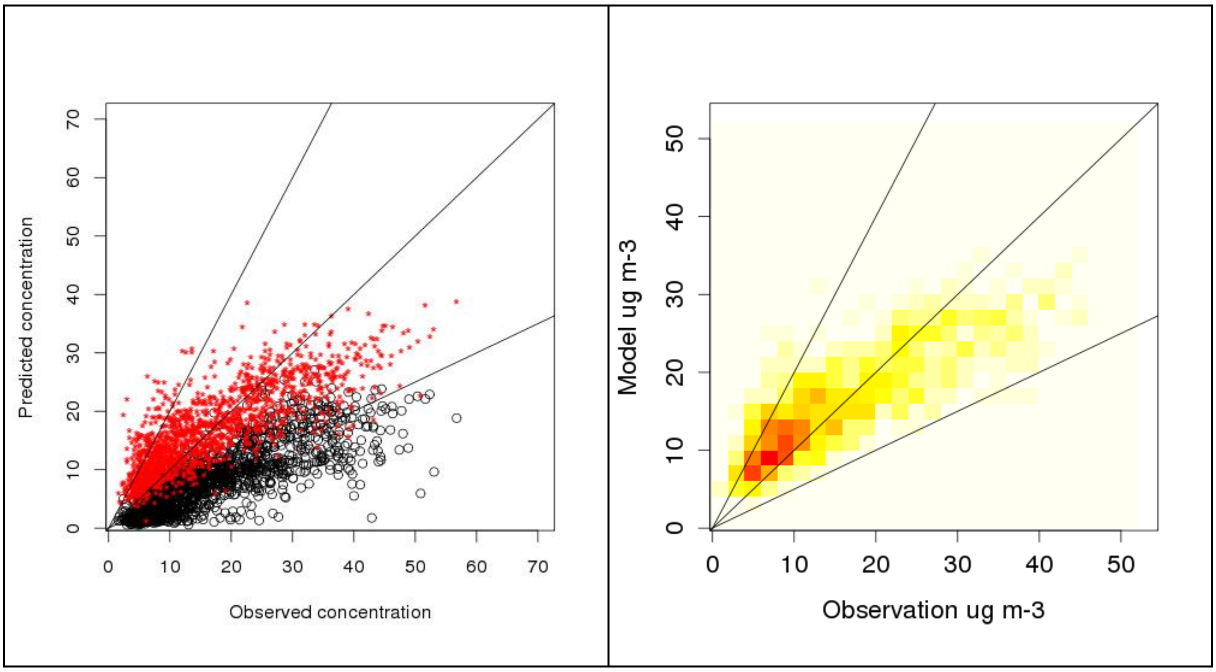


Table A3:1: Statistical assessment of the quality of Kalman filter adjustment of $PM_{2.5}$ data. The mean and standard deviation of the residual between the predicted and measured $PM_{2.5}$ concentrations are provided before and after adjustment by the Kalman filter. Results are shown for analysis of daily average and daily maximum concentration data at three different site types. A positive value of the mean basis indicates modelled concentrations are lower than observed values

Site type		Residual statistics	No Kalman filter, µg m^{-3}	With Kalman filter, µg m^{-3}
Urban background	Daily Average	Mean bias	8.49	0.01
		Standard deviation	6.91	6.35
	Daily maximum	Mean bias	16.04	0.15
		Standard deviation	13.94	8.28
Urban	Daily Average	Mean bias	10.88	0.06
		Standard deviation	8.16	7.64
	Daily maximum	Mean bias	18.37	0.13
		Standard deviation	13.12	8.03
Rural	Daily Average	Mean bias	9.27	0.04
		Standard deviation	8.19	7.32
	Daily maximum	Mean bias	16.72	0.05
		Standard deviation	14.60	8.73

To investigate the potential of the method for improving forecasts, Figure A3.3 plots daily mean observed and predicted $PM_{2.5}$ concentrations at six different sites in the UK between 1st February and 11th March 2013. The horizontal lines on the plots mark the boundaries of the daily air quality low, "Moderate", "High" and very high indexes. In general, it is seen that the unadjusted predictions underestimate the observed concentrations in $PM_{2.5}$, but do tend to rise and fall with the same time dependence. Adjustment of the predictions with the Kalman filter improves the agreement between predictions and observations, though typically maxima in the data are under-predicted, while minima in the vicinity of a maximum in the data tend to be over-predicted. It is interesting to note in these plots that the predictions (without filtering) never forecast outside of the low category during the pollution event while predictions which have been adjusted by the Kalman filter typically do not under- predict a maximum by more than one category. This is an improvement compared to the raw, unfiltered predictions and would likely be advantageous in advising the duty pollution forecaster.

Figure A3.2: Scatter plots showing correlation between observations and predictions of daily maximum $PM_{2.5}$ at urban background sites between 01/01/13 and 04/03/13. The left hand plots shows the correlations to predicted data before (black circles) and after (red

dots) adjustment by Kalman filter. The right hand density scatter plot shows the same adjusted data, but indicates the density of sites, with high densities indicated by red and low densities by yellow.

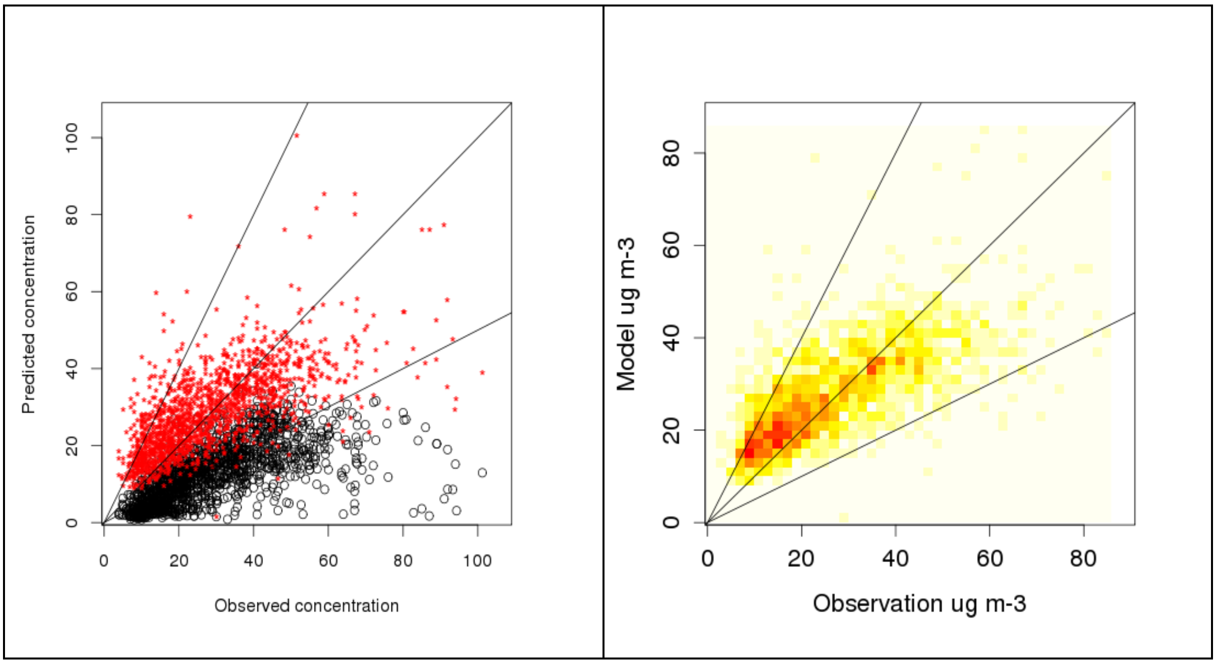
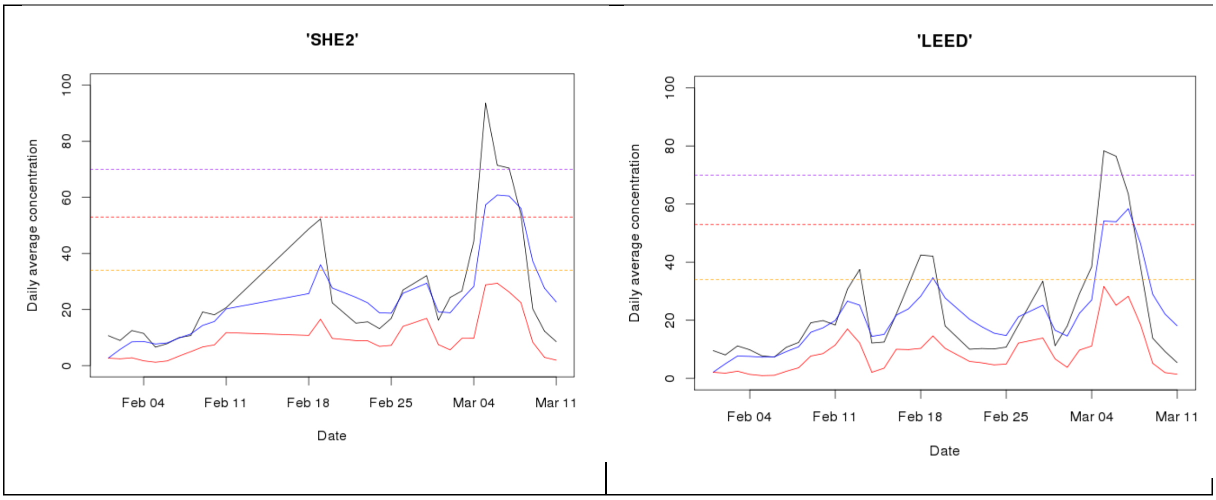


Figure A3.3: Example time series plots showing the influence of the Kalman filter on the predicted 24 hour averaged data for $PM_{2.5}$ concentrations at six locations between 1st February and 11th March 2013. The site observations (black lines), prediction without Kalman filter (red line) and prediction adjusted after application of the Kalman filter (blue line) are plotted. The dashed lines represent the lower bounds of the "Moderate" (orange), "High" (red) and "Very high" (purple) categories. Note that no site observations were recorded in Sheffield City Centre (SHE1) between 11th and 17th February.



Ozone daily average and daily maximum concentration data

Further analysis was carried out to test the potential benefits of applying the basic Kalman filter bias adjustment to daily average and daily maximum predicted ozone concentration data. Figures A3.4 and A3.5 show scatter plots of the predicted and observed concentrations of ozone before and after applying the Kalman filter adjustment to daily average and daily maximum data respectively. A statistical analysis of the results is presented in Table A3.2. The plots display concentration data for urban background sites between 1st January until 15th March 2013, but are representative of the trends observed across all site types over this time period. As was seen in the previous section for $PM_{2.5}$ data, the agreement between predicted and forecast ozone data improves after bias adjustment. The scatter between the daily average data is also reduced by ~15% after Kalman filter adjustment, but there is little change in the scatter of the daily maximum data. The improved agreement between predicted and measured concentrations after adjustment by the Kalman filter is further demonstrated in plots of daily mean and daily maximum observed and predicted ozone concentrations at six different locations in Figures A3.6 and A3.7 respectively.

Figure A3.4: Scatter plots showing correlation between observations and predictions of daily average ozone at urban background sites between 01/01/13 and 15/03/13. The left hand plots shows the correlations to predicted data before (black circles) and after (red dots) adjustment by Kalman filter. The right hand density scatter plot shows the same adjusted data, but indicates the density of sites, with high densities indicated by red and low densities by yellow.

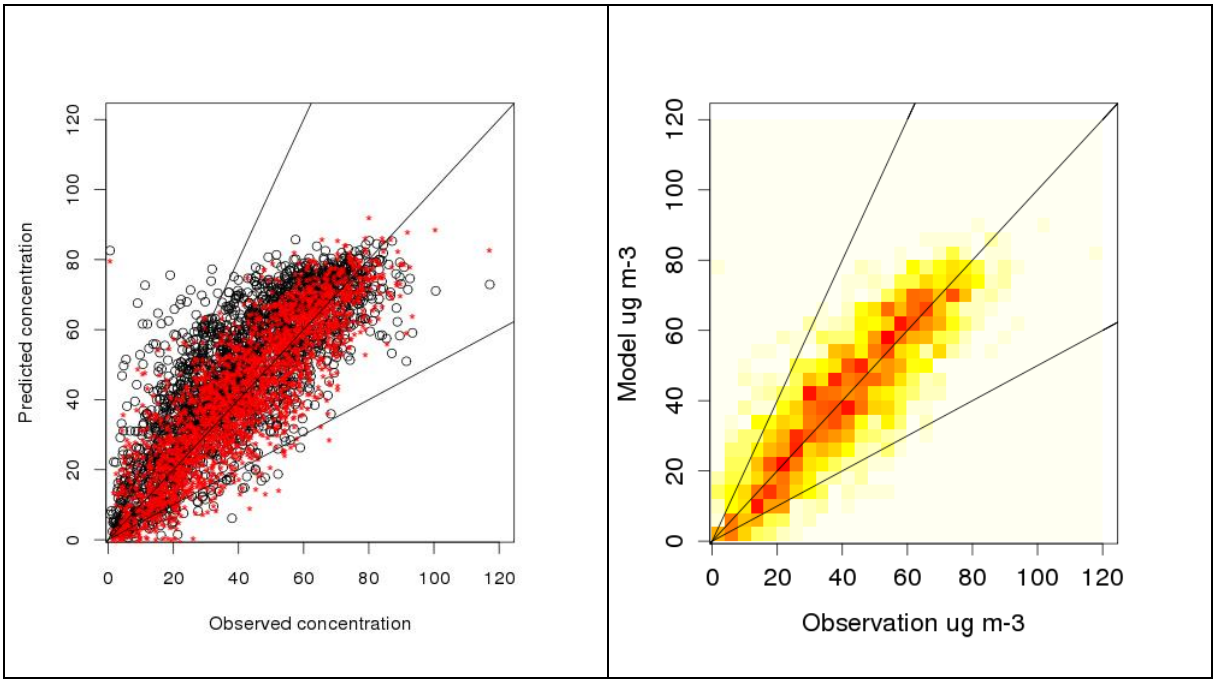


Figure A3.5: Scatter plots showing correlation between observations and predictions of daily maximum ozone at urban background sites between 01/01/13 and 15/03/13. The left hand plots shows the correlations to predicted data before (black circles) and after (red dots) adjustment by Kalman filter. The right hand density scatter plot shows the same adjusted data, but indicates the density of sites, with high densities indicated by red and low densities by yellow.

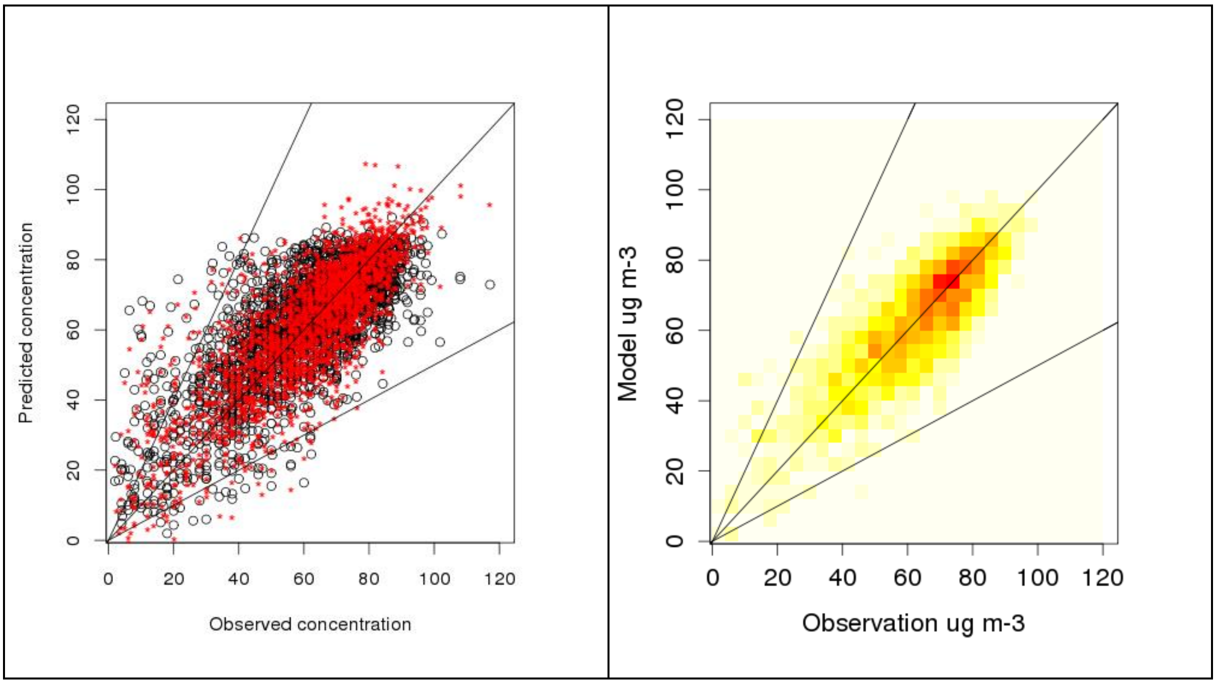
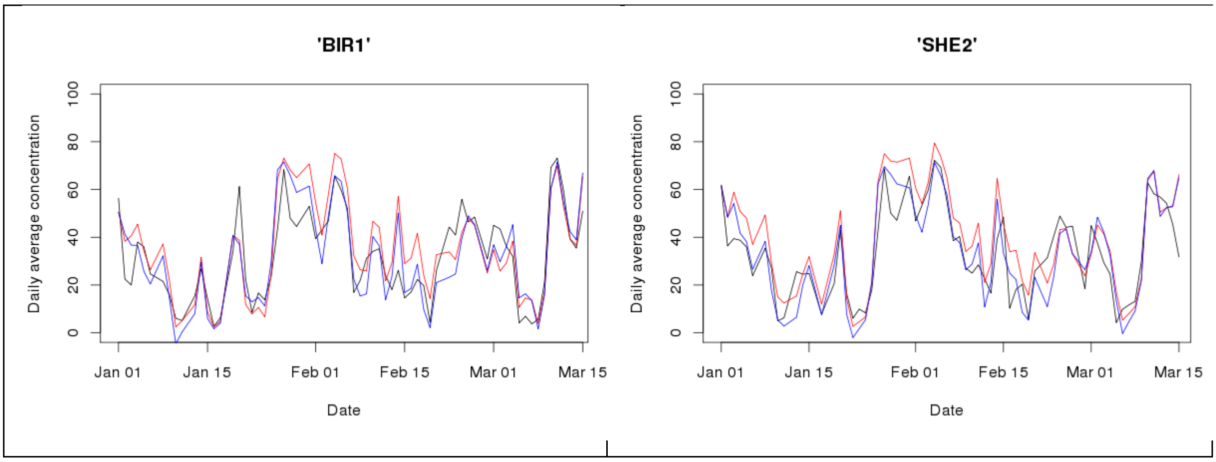


Table A3.2: Statistical assessment of the quality of Kalman filter adjustment of ozone data. The mean and standard deviation of the residual between the predicted and measured ozone concentrations are provided before and after adjustment by the Kalman filter. Results are shown for analysis of daily average and daily maximum concentration data at three different site types. A positive value of the mean basis indicates modelled concentrations are lower than observed values

Site type		Residual statistics	No Kalman filter, µg m^{-3}	With Kalman filter, µg m^{-3}
Urban background	Daily Average	Mean bias	-5.96	0.02
		Standard deviation	12.82	10.94
	Daily maximum	Mean bias	-1.03	0.02
		Standard deviation	13.17	13.19
Urban	Daily Average	Mean bias	-11.79	0.04
		Standard deviation	12.73	10.82
	Daily maximum	Mean bias	-7.23	0.00
		Standard deviation	13.23	12.95
Rural	Daily Average	Mean bias	-3.76	-0.00
		Standard deviation	12.36	10.26
	Daily maximum	Mean bias	-0.50	0.01
		Standard deviation	12.97	12.42

Figure A3.6: Example time series plots showing the effect of the Kalman filter on the predicted 24 hour averaged data for ozone concentrations at six locations between 1st January and 15th March 2013. The site observations (black lines), prediction without Kalman filter (red line) and prediction adjusted after application of the Kalman filter (blue line) are plotted.



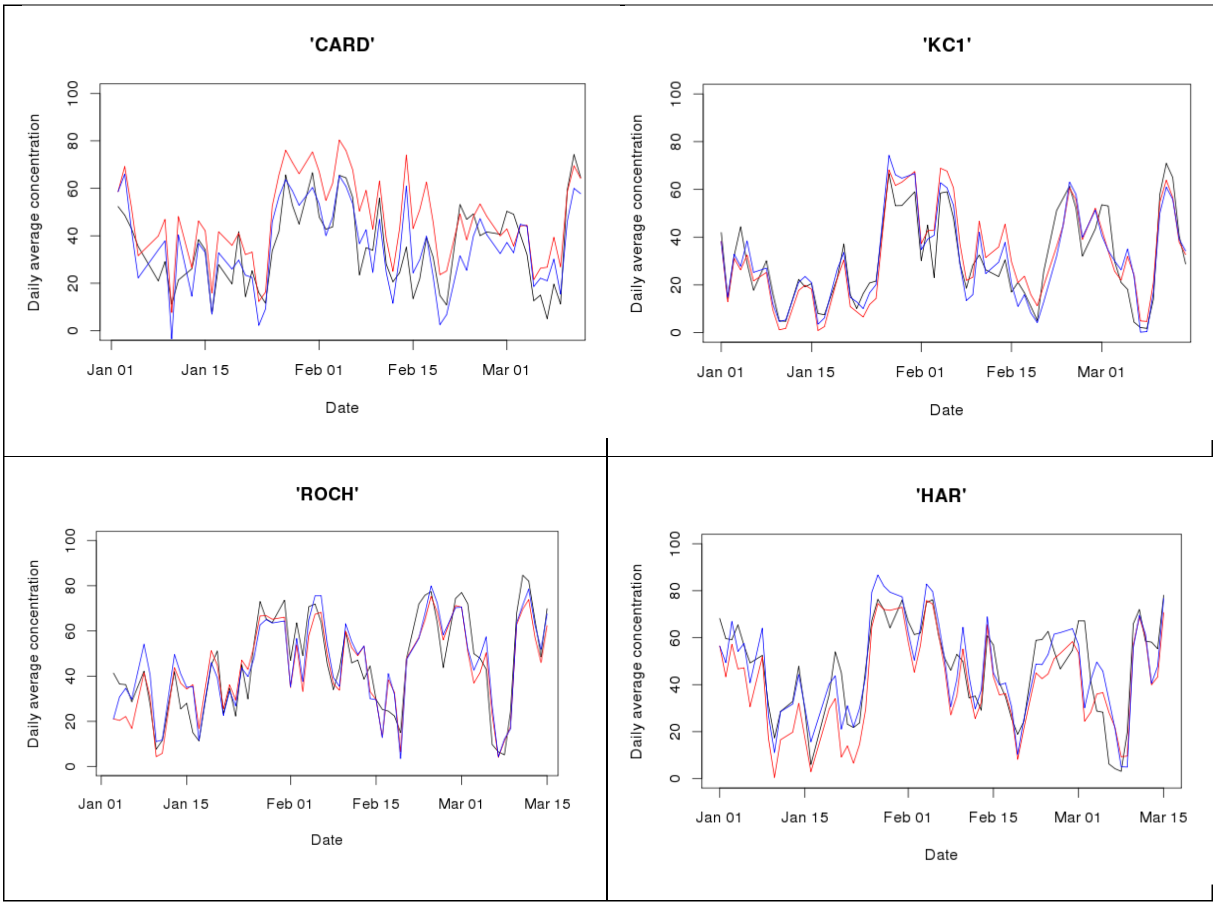
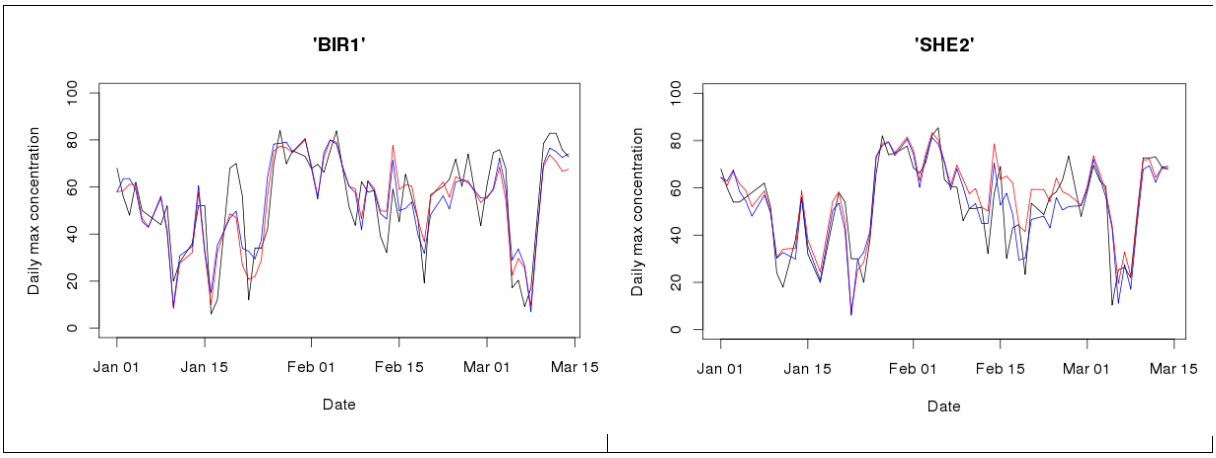
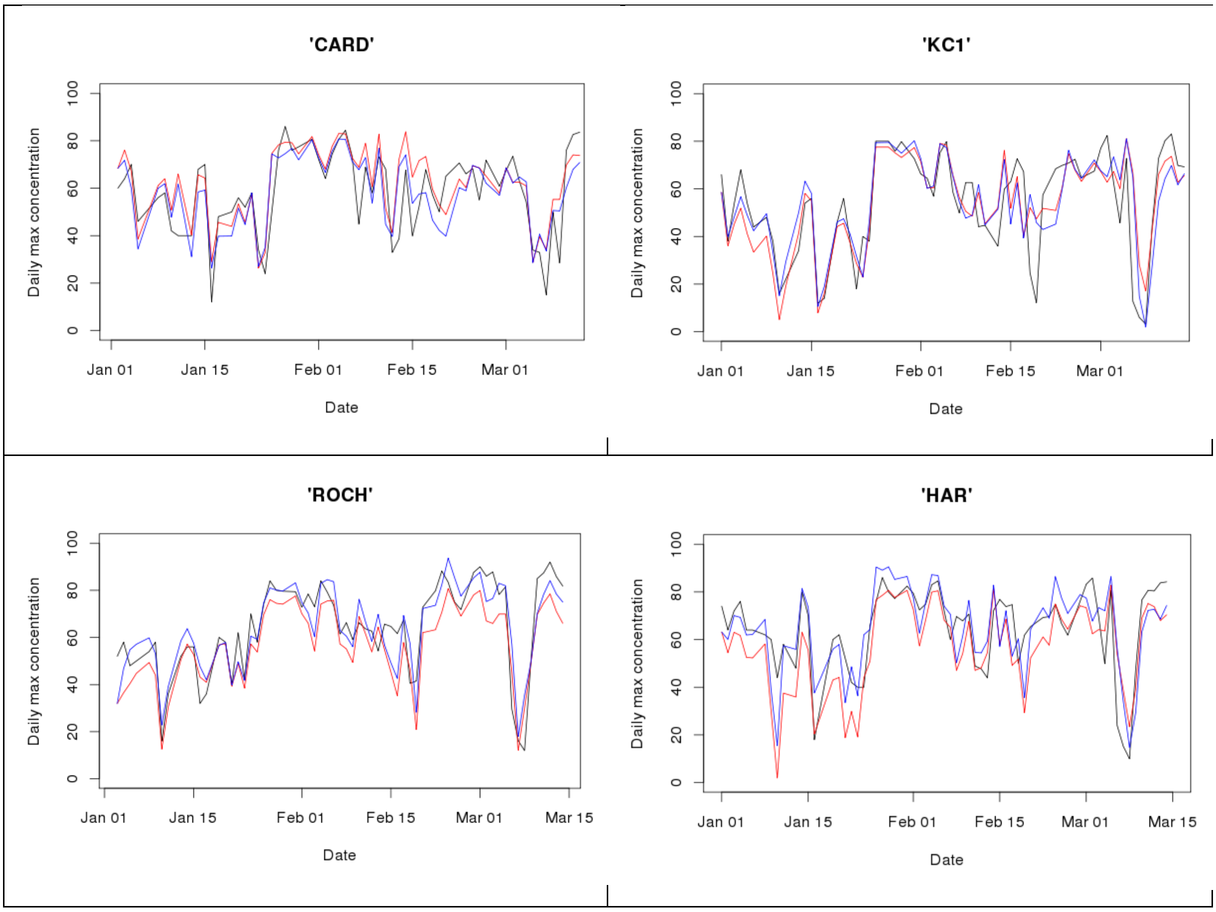


Figure A3.7: Example time series plots showing the effect of the Kalman filter on the daily maximum 1-hour averaged ozone concentrations at six locations between 1st January and 15th March 2013. The site observations (black lines), prediction without Kalman filter (red line) and prediction adjusted after application of the Kalman filter (blue line) are plotted.





US EPA Kalman filter algorithm

The EPA Kalman filter algorithm uses forecast and observed hourly concentrations for the previous two days at monitoring sites. It uses the Kalman filter approach to estimate a bias adjustment at each of the monitoring sites **for each hour** of the current day based on the bias adjustment calculated for the same hour of previous days. The computer codes have been rewritten and adapted for use in UK forecasting in this project, taking into account the different formatting of the UK data.

The US EPA Kalman filter approach was tested on data from several different UK monitoring sites across the whole of 2012. Three different pollutants were looked at depending on the availability of data available from each of the different sites: O_3, NO_2 and PM_{10}. The different sites were selected to provide an overview of the quality of the Kalman filter adjustment to the quality of the forecast for both rural and urban site and for sites which for which the forecast tends to show a high level of accuracy and sites for which the forecast tends to be less accurate or biased.

The Kalman filter was tested in two different modes, in the first the Kalman filter algorithm acts directly on the historical observed and forecast concentration data for the previous two days to provide the adjusted forecast for the third day. In the second the filter acts on the logarithm of the historical concentration data. The former can be understood to provide a shift bias correction to the correlation between observations and forecast data, while the latter provides a gradient bias correction to the correlation. Where data is missing due to missing forecast data the Kalman filter operates on the closest available days data from the time series. If observation data is missing that hours observation is set equal to the previous hours observation for the purpose of the bias correction.

The results of the US EPA Kalman filter bias correction tests based on the 2012 hourly data are presented below. Table A3.3 provides a summary of the mean and standard deviation of the residuals between the observed and forecast data before and after application of the correction method. Figures A3.8 and A3.9 present example plots for ozone and PM_{10} showing the correlation between the observed and forecast data before and after application of the bias correction for the North Kensington site, an urban background site which show trends typical of most sites. The data show that application of the US EPA Kalman filter typically reduces biases in the forecast data. However, for the majority of cases under investigation, particularly where the Kalman filter operates on the logarithmic bias, it does this at the expense of increased scatter between the bias corrected forecast data and the observed pollutant concentrations.

"Moderate" particulate pollution events in mid-February 2013 provided a good test case for evaluating the US EPA Kalman filter methodology. Analysis was performed as described above between 13/02/2013 and 24/02/2013 on the hourly concentration of PM_{10} and $PM_{2.5}$ for a number of sites which were grouped according to their type, rural, urban or urban background sites. Figure A3.10 shows the resulting correlations between the observations and the predicted $PM_{2.5}$ concentrations with and without correction by Kalman filter for urban background sites. Analysis of the mean and standard deviation statistics for all site types indicate that, as was seen in the analysis of data from 2012, that use of the US EPA Kalman filter typically reduces biases in the forecast data, but at the expense of increased scatter (i.e. increases likelihood of disagreement) between the bias corrected forecast data and the observed pollutant concentrations.

Table A3.3: Summary of the quality of Kalman filtering bias adjustment of forecasted pollution for several different monitoring sites in 2012. The mean and standard deviation of the residuals between the observations and forecast and the forecast after Kalman filter adjustment applied directly to the concentration data (linear KF) and applied to the logarithm of concentration. A positive value of the mean basis indicates modelled concentrations are lower than observed values

Site	species		Forecast, µg m^{-3}	Direct KF, µg m^{-3}	Log KF, µg m^{-3}
Harwell	O_3	mean	1.88	0.09	0.12
		stdv	14.73	15.13	15.51
	PM_{10}	mean	5.15	0.08	-2.79
		stdv	8.60	9.03	13.37
	NO_2	mean	-11.85	-4.30	1.33
		stdv	11.98	10.87	9.58
North Kensington	O_3	mean	-7.29	-0.94	-3.30
		stdv	16.81	16.55	31.02
	PM_{10}	mean	8.04	0.51	-0.94
		stdv	10.99	11.60	16.55
	NO_2	mean	2.91	0.08	-0.09
		stdv	15.94	15.98	17.19
	O_3	mean	-23.60	-1.81	-0.09

Market Harborough		stdv	15.81	15.40	14.63
	NO_2	mean	16.51	0.71	-3.91
		stdv	22.01	23.29	36.29
Aberdeen	O_3	mean	-14.08	-0.18	4.25
		stdv	17.37	18.04	19.34
	PM10	mean	4.87	0.16	0.18
		stdv	8.00	9.01	8.41
	NO_2	mean	6.40	0.17	1.39
		stdv	14.29	14.55	15.30
London Hillingdon	O_3	mean	-14.6	-3.56	-2.59
		stdv	20.85	19.65	29.57
	NO_2	mean	19.36	0.28	-0.96
		stdv	30.27	29.826	37.38
Eskdalemuir	O_3	mean	-17.82	-0.23	2.76
		stdv	14.63	15.23	16.35
	NO_2	mean	0.62	0.02	-0.66
		stdv	3.91	3.99	4.50
Lullington Heath	NO_2	mean	-2.35	-0.75	-0.79
		stdv	7.67	7.40	7.37
	O_3	mean	-13.67	-0.30	0.78
		stdv	14.67	15.18	15.42

Figure A3.8: Application of the Kalman filter bias correction method to the prediction of ozone concentrations at the North Kensington site in 2012. The left hand plot shows the correlation between observations and predictions before bias correction, the middle plot shows the correlation after bias correction of the prediction applied directly to the data, and the right hand middle plot shows the correlation after bias correction of the prediction applied to the log of the data. The colour of the scatter points indicates the density of events in a region with increasing density from yellow to red.

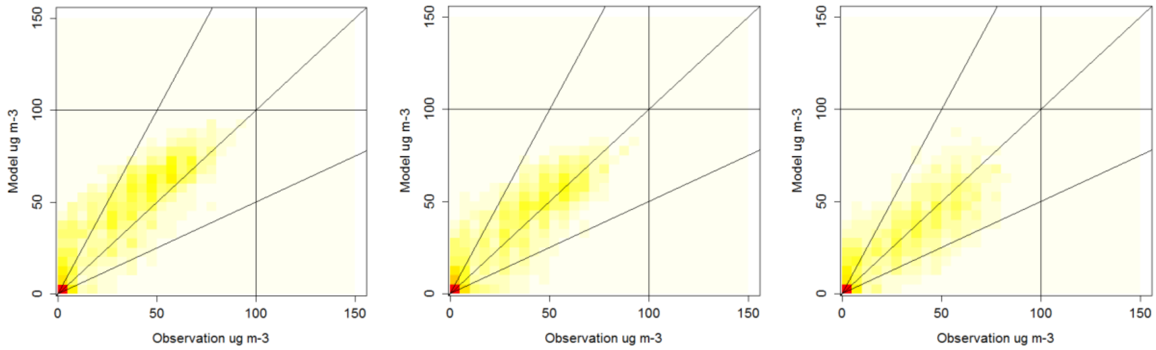


Figure A3.9: Application of the Kalman filter bias correction method to the prediction of PM_{10} concentrations at the North Kensington site in 2012. The left hand plot shows the correlation between observations and predictions before bias correction, the middle plot shows the correlation after bias correction of the prediction applied directly to the data, and the right hand middle plot shows the correlation after bias correction of the prediction applied to the log of the data. The colour of the scatter points indicates the density of events in a region with increasing density from yellow to red.

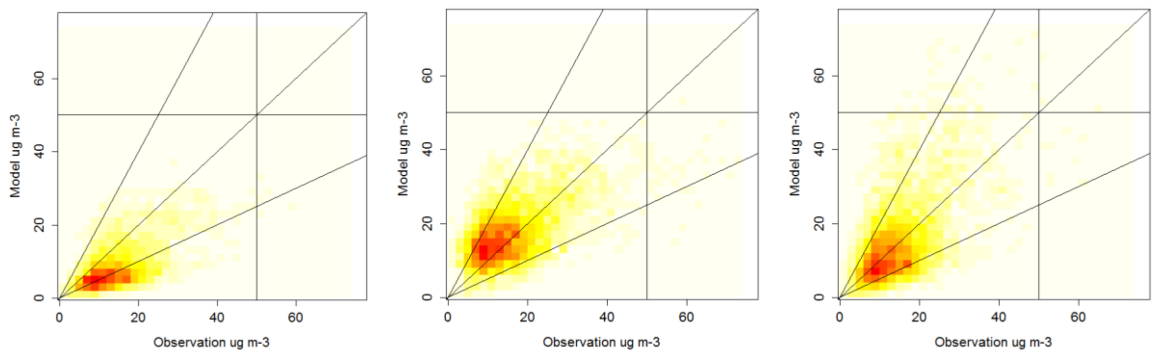
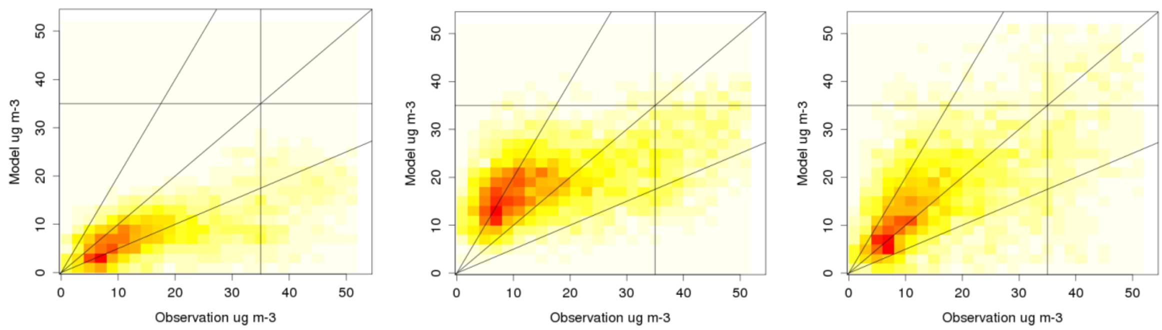


Figure A3.10: Scatter plots showing correlations between observations and predicted (model) $PM_{2.5}$ concentrations at urban background sites in February 2013. Left hand figure show the correlation to predicted data, middle shows the correlation to predicted data after adjustment with a linear Kalman filter and right hand figure shows the correlation to predicted data after adjustment by Kalman filter acting on logarithm of concentration data.



Recommendations

Use of the basic Kalman filter algorithm substantially improves forecasts of daily mean and daily maximum 1-hour concentrations of particulate matter $PM_{2.5}$ and ozone. It effectively eliminates the mean bias between the predictions and the measured concentrations and in most cases reduces the standard deviation of the bias. **It is therefore recommended that the approach is used routinely in the preparation of air pollutant forecasts**. The forecaster should be presented with both the filtered and unfiltered forecasts.

The US EPA Kalman filter algorithm attempts the more demanding task of predicting hourly average concentrations. Although the method reduces the mean bias between the predictions and the measurement, in many cases it can increase the scatter. This approach is therefore not recommended for application to UK forecasts.

This study has not tuned the Kalman gain. Further improvements may be obtained if the gain was tuned to reduce the scatter. It is recommended that further work is carried out to investigate the sensitivity of the performance of the filter to the gain.

This study has considered the prediction of mean concentration for the next calendar day from the previous day's concentration. In practice, model outputs are provided part way through the day. It is recommended that the algorithms are adapted to take account of the times of the model output.

This study has considered the bias at each monitoring site individually. The filter performance may potentially improve if the bias adjustment took account of the calculated bias at neighbouring monitoring sites. It would then be less influenced by "events" (e.g. short term pollutant releases close to the monitor). The Kalman filter algorithms can be adapted to take account of the bias at more than one monitoring station on the basis of existing mathematical formulations. A further advantage of this development would be that it would be possible to map the filtered concentrations over the whole of the UK.

RICARDO-AEA

The Gemini Building
Fermi Avenue
Harwell
Didcot
Oxfordshire
OX11 0QR

Tel: 0870 190 1900
Fax: 0870 190 6318

www.ricardo-aea.com