21st International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
27-30 September 2022, Aveiro, Portugal

DEFRA 2021 AIR QUALITY MODEL INTER-COMPARISON EXERCISE

David Carruthers[1], Jenny Stocker[1],Amy Stidworthy[1],Christina Hood[1], Rose Jackson[1], George McCosh[1], Stephen Smith[1], Daniel Connolly[1], Shyn Tong[2], Alison Davies[2], Sean Beevers[3], Nutthida Kitwiroon [3], Nosha Assareh [3], Vivien Bright[4], Lucy Neal[4], Paul Agnew[4], John Stedman[5], Beth Conlan[5], Massimo Vieno[6], Eiko Nemitz[6], Stefan Reis[6]

[1]CERC, 3 King's Parade, Cambridge, CB2 1SJ, UK
[2]Department for Environment, Food & Rural Affairs, Seacole Building, 2 Marsham Street, London, SW1P 4DF, UK
[3]Environmental Research Group, School of Public Health, Michael Uren Biomedical Engineering Hub, White City Campus, Wood Lane, London, W12 0BZ, UK
[4]Met Office, FitzRoy Road, Exeter, EX1 3PB, UK
[5]Ricardo, The Gemini Building, Fermi Avenue, Harwell, Didcot OX11 0QR, UK
[6]Centre for Ecology and Hydrology, Edinburgh, EH26 0QB, UK

Abstract: The UK takes a combined measurement and modelling approach to reporting associated with the Air Quality Standards Regulations (AQSR, previously the EU Air Quality Directive) pollutant metrics, with modelling currently being performed on behalf of the UK Department for Environment, Food & Rural Affairs (Defra) by Ricardo using the Pollution Climate Mapping (PCM) system. The primary purpose of the Defra 2021 Air Quality Model Inter-Comparison Exercise was to assess the capabilities of four air quality modelling systems in terms of their suitability for AQSR reporting, specifically: PCM; the CMAQ-Urban model driven by WRF meteorology (Environmental Research Group at Imperial College, London); the Air Quality model within the UK Met Office's Unified Model (AQUM-SPPO); and a WRF – EMEP application for the UK (UK Centre for Ecology and Hydrology). This paper provides a project overview and presents key conclusions. All models were configured to calculate pollutant concentrations for 2018 at over 400 monitor locations, gridded concentrations at the models' highest resolution over all of the UK, and, for three of the four models, near-road concentrations associated with the major road network. A wide range of metrics were calculated to assess model performance using NO_x, NO_2, O_3, $PM_{2.5}$ and PM_{10} measurement datasets. In addition to visual comparison of air quality maps, derived statistics such as areas in exceedance were calculated separately for 28 agglomeration and 15 non-agglomeration zones. A documented assessment of the models' formulations, configurations and inputs led to an informed model inter-comparison. Meteorological model performance has been evaluated at seven sites over the UK (wind speed, direction and temperature), and the relationship between modelled wind and pollutant concentrations has been investigated. Technical diagnostics have been used to assess how well the models account for NO_x chemistry, in addition to the models' ability to represent coarse and fine particulate concentrations. Conclusions of the study include: a quantitative inter-comparison of zonal exceedances, which are very similar for O_3 metrics but differ between models for NO_2 and particulates; and a qualitative discussion of the models' strengths and weaknesses in relation to AQSR reporting.

Key words: *Air quality, dispersion, inter-comparison, model, air quality directive, Defra, UK, AQSR*

INTRODUCTION

Four groups which run air quality (AQ) modelling systems with the potential to provide outputs suitable for assessing compliance with the Air Quality Standards Regulations (AQSR, 2010, previously the EU Air Quality Directive, AQD, 2008) were invited to participate in the Defra 2021 Air Quality Model Inter-Comparison Exercise (MIE). The Environmental Research Group at Imperial College London (ERG-ICL) used their CMAQ-Urban model (Beevers *et al.*, 2012), with WRF (Skamarock *et al.*, 2019) meteorological data; the Met Office (MO) ran their Air Quality Unified Model (AQUM-SPPO, Neal *et al.*, 2014); Ricardo supplied the Pollution Climate Mapping system (PCM, Brookes *et al.*, 2020); and the UK Centre for Ecology and Hydrology (UKCEH) ran EMEP (Simpson *et al.* 2012, Vieno *et al.*, 2016), also using WRF. The MIE comprised four tasks. The first task involved a mainly qualitative review and assessment of the models' formulations, configurations and inputs. The second task was an inter-comparison of AQSR pollutant metrics and maps. The third task involved a comprehensive evaluation of model performance relative to measured air pollutant concentrations. Modelled urban air quality was assessed in the final task, with a focus on two conurbations, Greater London and Greater Manchester.

This article provides a study overview: key aspects of the models' differing formulations and configurations are discussed; some examples from the extensive model evaluation exercise are presented followed by results relating to the application of each model to compliance reporting; and outcomes are summarised.

MODELS' FORMULATION AND CONFIGURATION

ERG-ICL, MO and UKCEH run regional meteorological and chemical transport models that predict pollutant concentrations at hourly resolution whereas Ricardo's modelling system calculates annual average concentration values. Three of the four models (CMAQ-Urban, AQUM-SPPO and PCM) calculate concentrations at roadside as well as regionally. CMAQ-Urban generates modelled concentrations that vary continuously between regional and local scales, up to 20 m grid resolution. MO and Ricardo generate separate gridded (resolutions ~12 km and 1 km respectively) and roadside datasets. EMEP's spatial resolution for this study was 1 km. MO and Ricardo apply post-processing calibration; the methods used differ, as do the measurement datasets used for the calibration. ERG-ICL and UKCEH do not apply any post-processing calibration, but use measurement datasets to refine model boundary conditions (O_3); refinement of model boundary conditions for O_3 using measurements is also carried out by the MO. Local modelling approaches differ significantly between groups. ERG-ICL use a near-road dispersion kernel based on ADMS-Roads version 5.0 (CERC, 2022), and a simple NO_X chemistry scheme; local modelling effects are included up to 225 m from each road source modelled, and account is taken of the influence of street canyons on dispersion. MO use a post-processing bias correction approach to estimate roadside concentrations; some regional variation of roadside increments is modelled, but no account is taken of specific road link features in terms of emissions or geometry such as carriageway widths or canyon properties. Ricardo uses near-source dispersion kernels derived from ADMS-Roads version 4.1 for roads, and from ADMS versions 3 and 5.2 for point sources, with an oxidant partitioning model for NO_X chemistry. In the form used by UKCEH in this MIE, EMEP does not conduct sub-grid scale modelling.

A range of land use and surface property values used as input to the hourly resolution meteorological models have been inter-compared at 15 sites throughout the UK, representing a variety of environments. The land use categories used by the different models are consistent for regional models using the same input data at relatively similar resolution (WRF at 2 and 1 km by ERG-ICL and UKCEH respectively). Input surface roughness lengths for meteorological modelling are broadly similar across all models although roughness lengths used by ERG-ICL are generally higher in urban areas than for the other models, and have more seasonal variation. WRF is also used in PCM but at much lower (50 km) resolution, and there are other configurations differences. Modelled meteorological parameters were evaluated using measurements from seven of the 15 sites used in the land use / surface roughness inter-comparison (Figure 1). FAIRMODE meteorological parameter benchmarks indicate that the MO meteorological model performs best, satisfying the benchmark criteria for all parameters evaluated (wind speed, wind direction and temperature). All other models demonstrate a slight negative bias for temperature, and ERG-ICL also underpredicts wind speed. There is broad consistency in terms of the anthropogenic emissions inputs used by the modelling teams, specifically data from the National Atmospheric Emissions Inventory (NAEI, 2020) for the UK and EMEP for Europe. ERG-ICL adjust their emissions from light duty vehicles using bottom-up calculations with emission factors derived from remote sensing data, resulting in total emission

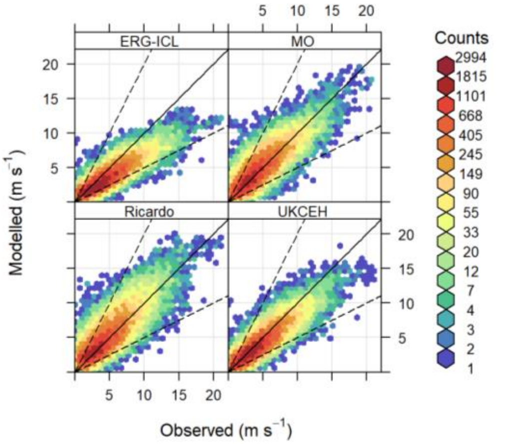

Figure 1. Frequency scatter plot of hourly modelled and observed wind speed data at 10 m above ground across seven meteorological evaluation sites; colours indicating the density of data points in each region of the graph.

increases of 5%, 30% and 150% for NO_X, $PM_{2.5}$ and PM_{10} respectively when compared to the base NAEI. The assumptions relating to the proportion of traffic NO_X emitted as primary NO_2 varies greatly between modelling groups. MO use the lowest value, assuming that all NO_X is emitted as NO, and ERG-ICL assume the highest proportion, with values ranging from 0.16 to 0.30. Non-road traffic primary NO_2 emissions assumptions also vary, with proportions ranging from 0 to 0.14.

MODEL EVALUATION AT AIR POLLUTION MONITORING SITES

Model predictions of core AQSR pollutants NO_X, NO_2, O_3, PM_{10} and $PM_{2.5}$ have been compared with hourly measurements from 415 UK monitoring sites for 2018. Models have been evaluated separately at background (including rural, suburban and urban background), roadside and industrial sites, with comparisons on three timescales: annual, hourly and, where relevant, the AQSR short-term averaging periods. PCM calculates only annual metrics, so hourly and AQSR short-term limit assessments exclude PCM. The CERC Model Evaluation Toolkit (2021), which uses tools from the openair package (Carslaw and Ropkins, 2012), was used to produce a comprehensive set of statistics and graphs to quantitatively assess each model's performance in relation to observations; the statistics include the mean, root mean square error (RMSE), normalized mean bias (NMB) and normalized mean square standard deviation (NMSD) for both the annual and hourly data. The number of short-term AQSR limit exceedances has been calculated. FAIRMODE metrics, which allow for measurement uncertainty, have also been calculated.

As an example, Table 1 presents a selection of statistics associated with evaluation of annual average NO_2. There is good overall agreement between modelled concentrations and observations for CMAQ-Urban and PCM; EMEP underestimates NO_2 at background sites. AQUM-SPPO has good agreement overall, but further categorisation of the statistics (not presented) indicates overestimation at rural sites. AQUM-SPPO and EMEP underestimate variability, whereas CMAQ-Urban overestimates variability. With regard to the FAIRMODE metric $MQI_{annual\ 90}$, both AQUM-SPPO and PCM achieve the annual threshold criteria (less than 1 for an acceptable model) at background sites; none of the models achieve this criteria at roadside. PCM gives the best prediction in terms of the number of sites exceeding the AQSR annual mean limit value (40 µg/m³), although CMAQ-Urban also demonstrates good performance for this metric at roadside sites.

Table 1. Model evaluation statistics for annual mean NO_2 (µg/m³) at background and roadside sites; best result per statistic per site type in bold; last column shows number of sites exceeding the annual limit value (40 µg/m³).

Site type	Modelling group	Model	Mean	RMSE	NMB	NMSD	$MQI_{annual\ 90}$	Sites exc. annual limit
Background		**Observed**	19.8					1
	ERG-ICL	CMAQ-Urban	21.6	5.8	0.09	0.32	1.06	11
	MO	AQUM-SPPO	18.3	6.0	-0.08	-0.46	0.96	0
	Ricardo	PCM	**18.5**	**4.7**	**-0.07**	**0.04**	**0.73**	**3**
	UKCEH	EMEP	15.0	6.9	-0.25	-0.18	1.10	0
Roadside		**Observed**	36.6					51
	ERG-ICL	CMAQ-Urban	38.8	13.1	**0.06**	0.36	1.43	56
	MO	AQUM-SPPO	33.0	12.4	-0.10	-0.67	1.45	22
	Ricardo	PCM	**34.6**	**9.6**	**-0.06**	**-0.20**	**1.27**	**47**
	UKCEH	EMEP	14.9	24.3	-0.59	-0.57	2.70	0

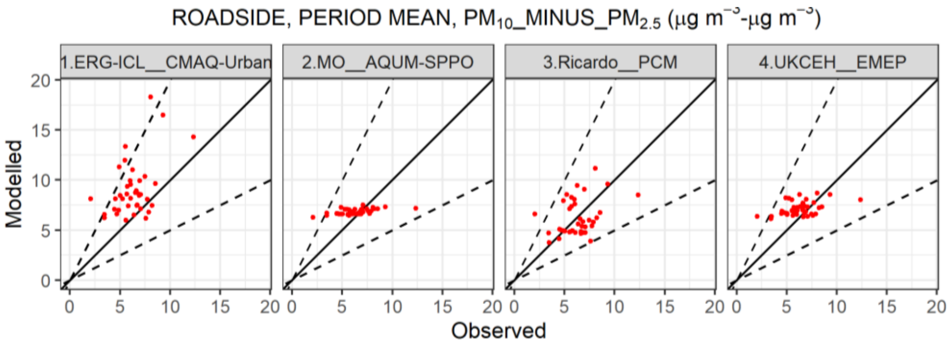

Figure 2. Modelled versus observed annual mean coarse particulate levels ($PM_{10} – PM_{2.5}$, µg/m³) at 40 roadside sites.

In the case of $PM_{2.5}$ and PM_{10} (not shown), all models show good overall agreement between modelled concentrations and observations for annual means, and all pass the corresponding FAIRMODE threshold criteria at both background and roadside sites apart from CMAQ-Urban at roadside sites. This likely relates to CMAQ-Urban's tendency to overestimate coarse particulate traffic emissions, demonstrated by the annual mean evaluation, Figure 2**Error! Reference source not found.** which highlights the need for consideration of a range of metrics when undertaking an evaluation study; for instance, calibration ensures that AQUM-SPPO predicts the correct mean coarse component, but model variability is significantly lower than observed.

COMPLIANCE MAPPING AND STATISTICS

Only annual metrics have been considered in the compliance reporting metric assessment, to allow direct comparison with PCM outputs. The compliance reporting calculation methodology follows that used by Ricardo in their AQSR reporting work for Defra. Pollutant concentration metrics have been derived for 28 agglomeration (urban) zones and 15 non-agglomeration (rural) zones. Separate calculations have been performed using gridded and roadside datasets. Zonal exceedances are calculated as the maximum concentrations over gridded and road datasets. PCM and MO calculate only one 'roadside' concentration associated with each of the 8586 UK urban road 'sections' modelled in the study. ERG-ICL's roadside concentrations have been calculated as an average over pavements, defined as a 2.5 m wide buffer on either side of the full modelled road network. The ERG-ICL CMAQ-Urban grid is fine resolution (20 m) for all pollutants excluding O_3, which is 2km resolution.

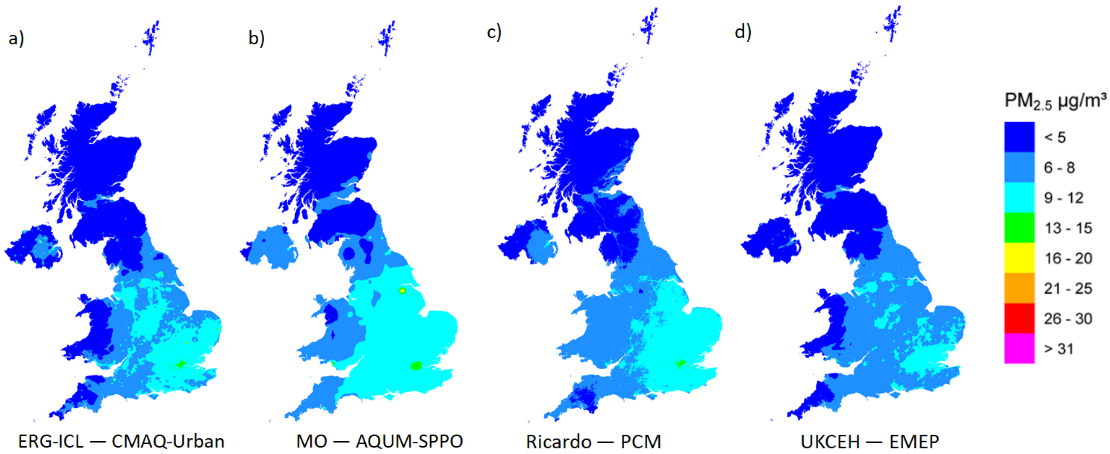

Figure 3. Maps of the United Kingdom showing gridded annual average $PM_{2.5}$ concentrations as modelled by a) ERG-ICL (CMAQ-Urban) b) MO (AQUM-SPPO) c) Ricardo (PCM) and d) UKCEH (EMEP). Note: grid resolutions differ for each modelling group.

Gridded concentrations of NO_2, $PM_{2.5}$, PM_{10} and O_3 have been mapped using colour scales corresponding to AQSR limit, target and long-term objective values. Example $PM_{2.5}$ air pollution maps are shown in Figure 3. All models show a similar spatial distribution of $PM_{2.5}$ for the lower concentration ranges (less than 10 µg/m³). However, map details differ between models in terms of peak concentrations, partly due to the differing approaches taken to modelling non-traffic sources such as industry and calcium rich dust. CMAQ-Urban predicts exceedances of the annual average limit value (25 µg/m³) in approximately 50% of zones, which is likely to be an overestimate due to the assumed release height of some non-road sources and road carriageways not being excluded from the exceedance calculations. None of the other models predict exceedances of this limit value. Table 2 summarises outcomes of the compliance reporting calculations using the four models. For NO_2, roadside exceedances are broadly consistent between CMAQ-Urban and PCM; gridded exceedances of NO_2 include road carriageways so are not entirely consistent with the AQSR for CMAQ-Urban. AQUM-SPPO and EMEP predict few or zero exceedances of the NO_2 limit value. Zonal exceedances relating to both O_3 metrics are consistent across all models, with models predicting few, or no, zonal exceedances of the target values, but with exceedances of the long-term objectives in most zones.

Table 2. Overall modelled exceedances (combined gridded and roadside) over the defined 43 zones. Note: there are instances of overlap between gridded and roadside zonal exceedances. Limit Value = LV, Target Value = TV, Long-Term Objective = LTO; *Grid exceedances include road carriageways; **CMAQ only for O_3.

Pollutant	**Time period**	**Threshold**	**Grid / road**	**ERG-ICL CMAQ-Urban***	**MO AQUM-SPPO**	**Ricardo PCM**	**UKCEH EMEP**
NO_2	Annual	40 µg/m³ (LV)	Grid	42*	0	1	0
			Road	32	2	34	0
			Total	**42***	**2**	**34**	**0**
O_3	No. of days 8-hour rolling mean > 120 µg/m³	25 days (TV)	Grid	1	1	0	0
		1 day (LTO)		42	43	43	43
	Seasonal (May to July) – AOT40	18000 µg/m³.h (TV)		0	1	0	4
		6000 µg/m³.h (LTO)		40	40	38	43
$PM_{2.5}$	Annual	25 µg/m³ (LV)	Grid	18	0	0	0
			Road	1	0	0	0
			Total	**18**	**0**	**0**	**0**
PM_{10}	Annual	40 µg/m³ (LV)	Grid	38	0	0	0
			Road	3	0	0	0
			Total	**38**	**0**	**0**	**0**

DISCUSSION

The MIE has identified the four models' strengths and weaknesses, and conclusions have been drawn in relation to the models' suitability for AQSR reporting purposes. This study was comprehensive, but only a small subset of results is presented in this article.

ACKNOWLEDGEMENTS

This work undertaken under a Memorandum of Agreement formed between Defra, CERC, Imperial College London, the UK Met Office, Ricardo and the UK Centre for Ecology and Hydrology.

REFERENCES

AQD, 2008: Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. Official Journal of the European Union, 152, 1–44.

AQSR, 2010: UK Statutory Instruments 2010 No. 1001 The Air Quality Standards Regulations 2010.

Beevers, S. D., Kitwiroon, N., Williams, M. L., & Carslaw, D. C., 2012: One way coupling of CMAQ and a road source dispersion model for fine scale air pollution predictions. *Atmos. Env.*, **59**, 47–58.

Skamarock, W. C. et al., 2019: A Description of the Advanced Research WRF Model Version 4 (No. NCAR/TN-556+STR).

Neal, L. S., Agnew, P., Mosely, S., Ordóñez, C., Savage, N. H., Tilbee, M., 2014: Application of a statistical post-processing technique to a gridded, operational air quality forecast. *Atmos. Env.*, **98**, 385-393.

Brookes D, Stedman J, Kent A, Whiting S, Rose R, Williams C., Pugsley, K., 2020: Technical report on UK supplementary assessment under The Air Quality Directive (2008/50/EC), The Air Quality Framework Directive (96/62/EC) and Fourth Daughter Directive (2004/107/EC) for 2018. Ricardo report reference ED12633 - Issue Number 1. Ricardo Energy & Environment/R/3470.

Simpson, D. *et al.*, 2012: The EMEP MSC-W chemical transport model – technical description. *Atmos. Chem. and Phys.*, **12**, 7825-7865.

Vieno, M., Heal, M.R., Williams, M.L., Carnell, E.J., Nemitz, E., Stedman, J.R. and Reis, S., 2016: The sensitivities of emissions reductions for the mitigation of UK PM 2.5. *Atmos. Chem. and Phys.*, **16(1)**, 265-276.

CERC, 2022: ADMS Technical Specifications, available online at https://www.cerc.co.uk/TechSpec (accessed April 2022)

NAEI, 2020: 13 Richmond B, Misra A, Broomfield M, Brown P, Karagianni E, Murrells T, Pang Y, Passant N, Pearson B, Stewart R, Thistlethwaite G, Wakeling D, Walker C, Wiltshire J, Hobson M, Gibbs M, Misselbrook T, Dragosits U, Tomlinson S (2020). UK Informative Inventory Report (1990 to 2018). National Atmospheric Emissions Inventory, Ricardo Energy & Environment.

Model Evaluation Toolkit (version 5.1), 2021: https://www.cerc.co.uk/ModelEvaluationToolkit

Carslaw, D. C. and Ropkins, K., 2012: openair --- an R package for air quality data analysis. *Env. Mod. & Soft*. **Volume 27-28**, 52-61.