Geoscientific
Model Development

Improving the spatial resolution of air-quality modelling at a European scale – development and evaluation of the Air Quality Re-gridder Model (AQR v1.1)

Mark R. Theobald[1,2], David Simpson[3,4], and Massimo Vieno[5]

[1]Atmospheric Pollution Division, Research Centre for Energy, Environment and Technology (CIEMAT),
Madrid, 28040, Spain
[2]Dept. Agricultural Chemistry and Analysis, Higher Technical School of Agricultural Engineering,
Technical University of Madrid, 28040, Spain
[3]EMEP MSC-W, Norwegian Meteorological Institute, Oslo, 0313, Norway
[4]Dept. Earth & Space Sciences, Chalmers University of Technology, Gothenburg, 412 96, Sweden
[5]Centre for Ecology & Hydrology, Edinburgh Research Station, Penicuik, EH26 0QB, UK

Correspondence to: Mark R. Theobald (mark.theobald@ciemat.es)

Abstract. Currently, atmospheric chemistry and transport models (ACTMs) used to assess impacts of air quality, applied at a European scale, lack the spatial resolution necessary to simulate fine-scale spatial variability. This spatial variability is especially important for assessing the impacts to human health or ecosystems of short-lived pollutants, such as nitrogen dioxide (NO_2) or ammonia (NH_3). In order to simulate this spatial variability, the Air Quality Re-gridder (AQR) model has been developed to estimate the spatial distributions (at a spatial resolution of $1 \times 1\,km^2$) of annual mean atmospheric concentrations within the grid squares of an ACTM (in this case with a spatial resolution of $50 \times 50\,km^2$). This is done as a post-processing step by combining the coarse-resolution ACTM concentrations with high-spatial-resolution emission data and simple parameterisations of atmospheric dispersion. The AQR model was tested for two European sub-domains (the Netherlands and central Scotland) and evaluated using NO_2 and NH_3 concentration data from monitoring networks within each domain. A statistical comparison of the performance of the two models shows that AQR gives a substantial improvement on the predictions of the ACTM, reducing both mean model error (from 61 to 41 % for NO_2 and from 42 to 27 % for NH_3) and increasing the spatial correlation (r) with the measured concentrations (from 0.0 to 0.39 for NO_2 and from 0.74 to 0.84 for

NH_3). This improvement was greatest for monitoring locations close to pollutant sources. Although the model ideally requires high-spatial-resolution emission data, which are not available for the whole of Europe, the use of a Europe-wide emission dataset with a lower spatial resolution also gave an improvement on the ACTM predictions for the two test domains. The AQR model provides an easy-to-use and robust method to estimate sub-grid variability that can potentially be extended to different timescales and pollutants.

1 Introduction

The impacts of air pollution on human health and natural ecosystems are often evaluated using data from atmospheric dispersion models or atmospheric chemistry and transport models (ACTMs). The scale of these evaluations ranges from local assessments with domains of several kilometres (e.g. Dragosits et al., 2002; Aggarwal and Jain, 2015; Galvis et al., 2015) to global assessments using grid cells of 1–10° (see for example Dentener et al., 2006). The spatial resolution used in these assessments depends on many factors, including availability of input data, model assumptions, receptor type (e.g. people, forests) and computation time. Many of the impact assessments at a European scale are carried

out using atmospheric concentration or deposition predictions of the model developed by the Meteorological Synthesizing Centre-West (MSC-W) of the European Monitoring and Evaluation Programme (EMEP). The EMEP MSC-W model (Simpson et al., 2012), called the EMEP model hereafter, has commonly been applied for policy purposes at a spatial resolution of ca. $50 \times 50 \, km^2$ (e.g. Fagerli and Aas, 2008; Simpson et al., 2006). Although the model is increasingly used at even finer resolution (e.g. $0.1 \times 0.1°$) even for official MSC-W purposes (EMEP, 2015), such runs are extremely CPU-intensive for European-scale modelling, and cannot be used for the hundreds to thousands of simulations required by the source–receptor matrices, which are an important output of MSC-W (EMEP, 2015). EMEP model results also underpin the Greenhouse gas – Air pollution Interactions and Synergies (GAINS) model, which is a key tool in developing European policy within both the United Nations Economic Commission for Europe and the European Union (Amann et al., 2011). However, the resolution of the EMEP model (or any other European-scale ACTM), at least when run in typical policy mode, is not currently high enough to resolve the large horizontal concentration gradients found close to sources of relatively short-lived pollutants, such as ammonia (NH_3), nitrogen dioxide (NO_2) or sulfur dioxide (SO_2) (CLRTAP, 2014; Denby et al., 2011).

The EMEP model predicts the mean near-surface atmospheric concentrations within each grid square, assuming a constant deposition flux between the centre of the first vertical layer (ca. 45 m) and a height of 3 m (Simpson et al., 2012). However, within a grid square there may be concentrations an order of magnitude (or more) above and below this mean value, even if the mean prediction is correct. Neglecting this sub-grid variability (SGV) can strongly bias assessments of air pollution impacts. For example, Denby et al. (2011) estimated that urban background exposure to NO_2 is underestimated by an average of 44 % when the $50 \times 50 \, km^2$ grid concentrations of the EMEP model are used. This problem is not restricted to the low grid resolution used by the EMEP model, it also occurs in assessments with higher resolutions. For example, Hallsworth et al. (2010) used an ACTM to estimate NH_3 concentrations in the UK at spatial resolutions of $5 \times 5 \, km^2$ and $1 \times 1 \, km^2$. They found that the 5 km model estimated that the NH_3 critical level of $1 \, \mu g \, m^{-3}$ was exceeded for 40 % of the total area of UK Special Areas of Conservation (SAC), whereas the 1 km model estimated an exceedance for only 21 %. This reduction in the area of exceedance when the model resolution was increased was due to the ammonia sources (agricultural areas) and the SAC being separated spatially. Modelling at a higher resolution resolved the large horizontal concentration gradients better, thus predicting higher concentrations in the agricultural areas and lower concentrations within the SAC. By contrast, Oxley and ApSimon (2007) found that increasing model spatial resolution from 50 to 5 km and from 5 to 1 km increased the estimates of exposure to primary parti-

cles with a diameter of 10 μm or less (PM_{10}) in urban areas. This is because, in this case, the urban areas are also some of the largest sources of primary PM_{10}. A multi-model study involving five ACTMs to simulate pollutant concentrations across Europe found a large increase in annual mean concentration predictions of PM_{10} and NO_2 in urban locations when increasing the spatial resolution through the range 56, 28, 14 and 7 km (Cuvelier et al., 2013; Schaap et al., 2015). For most of the models, about 70 % of the model response to the change of resolution was due to the change in the spatial distribution of emissions. By comparing the concentration predictions in urban areas with measured values, model performance (slope, bias and correlation) was generally found to improve for all models as the resolution was increased. In order to resolve the large horizontal concentration gradients found in urban areas, Cuvelier et al. (2013) suggested that a resolution of a few kilometres down to 1 km would be needed, but added that this is not currently feasible for application across Europe. However, even this might not be sufficient for resolving the large horizontal concentration gradients of NO_2, for example.

Several potential methods could be used to estimate the SGV of the concentration predictions of short-lived air pollutants across Europe. Firstly, the EMEP model could be applied at a higher resolution. This has been done in the UK for a resolution of $5 \times 5 \, km^2$ (EMEP4UK) (Vieno et al., 2010, 2014), and for Europe at ca. $7 \times 7 \, km^2$ (Schaap et al., 2015; EMEP, 2015), but such runs are extremely CPU-demanding and are not suitable for routine use, especially where ACTMs need to be run tens to hundreds of times for emission control assessments, for example. A European application at $1 \times 1 \, km^2$ resolution or higher is currently not feasible, even for research purposes. As well as being too demanding on computation time, such runs would also require a consistent and accurate high-resolution emission dataset, which is not currently available. A second solution is the "stitching together" of national modelling simulations at a high resolution (see, for example, de Smet et al., 2013; Janssen and Thunis, 2016). This approach has the advantage of making use of national expertise and emission and meteorological datasets. However, the disadvantages are that it is likely to lead to "border effects" as a result of differing methodologies and/or input datasets used by neighbouring countries and results may not be available for all countries, making it difficult to carry out a consistent assessment for the whole of Europe. The third solution is to apply geo-statistical techniques to the low-resolution concentration data (e.g. from the EMEP model) that makes use of other relevant spatial datasets. These techniques can be used to either estimate the probability distribution of the concentration (or a related quantity) within each grid square or to explicitly estimate the spatial distribution of the concentration within the grid square. An example of the former approach is that of Denby et al. (2011), who estimated the population-weighted concentrations of NO_2, PM_{10} and O_3 within each EMEP $50 \times 50 \, km^2$ grid

square using information on measured concentrations and their covariance with population density, which was then parameterised using emission and altitude data. Another example is the SGV parameterisation of Ching et al. (2006) for the CMAQ model based on sub-grid concentration distributions of benzene and formaldehyde, calculated using the ISCST3 short-range dispersion model. The same CMAQ simulations were used by Isakov et al. (2007) to develop a method to explicitly model the sub-grid spatial distributions of concentrations at a resolution of $200 \times 200 \, \text{m}^2$. Their method used relationships between the sub-grid concentrations and sub-grid emission strengths derived from short-range dispersion modelling results, although it was only applied to a small area (Philadelphia County). A different geo-statistical approach was used by Janssen et al. (2012), in which they estimated sub-grid concentrations for Belgium by using empirical relationships between long-term atmospheric concentrations and land-use characteristics. A Europe-wide approach was developed for NO_2 and particulate matter by Kiesewetter et al. (2013, 2014), although only at a resolution of $7 \times 7 \, \text{km}^2$. In their work, concentrations simulated by the EMEP model at a resolution of $28 \times 28 \, \text{km}^2$ were disaggregated using an "urban increment". This increment was calculated from the concentration predictions of the CHIMERE model (Bessagnet et al., 2004) at a resolution of $7 \times 7 \, \text{km}^2$. The relationship between the differences in the concentration predictions of the two models and the emission rate (from near-ground-level sources only) used for each 7 km grid square was used to calculate the urban increment. Model evaluation using annual mean concentrations from more than 1500 urban background monitoring stations showed that the model can predict concentrations within a factor of 2 of the measured value for most locations. The authors also developed a parameterisation to estimate the additional concentration increment at the locations of roadside air-quality stations, although this approach relies heavily on measurement data.

In this paper we present the development, testing and evaluation of a simple geo-statistical post-processing methodology (the Air Quality Re-gridder (AQR) model) that combines high-spatial-resolution emission data and a simple parameterisation of short-range dispersion to estimate the spatial distribution of concentrations of short-lived pollutants within the EMEP model grid squares. This sub-grid model is used to calculate the annual mean concentrations of NO_2 and NH_3 for 2008 at a resolution of $1 \times 1 \, \text{km}^2$ for two test domains (central Scotland and the Netherlands) and evaluated using monitoring network data from within the two domains. Section 2 provides information on the methods and datasets used and Sect. 3 describes the model development process. Section 4 presents the results of the sub-grid modelling, a model evaluation and an analysis of the sensitivity of the model to some of the parameters and datasets used, whilst Sect. 5 discusses model performance and its applicability, uncertainties and potential improvements and extensions.

2 Materials and methods

The two domains used in this study are central Scotland and the Netherlands (Fig. 1). These domains were chosen because they provide a contrast between a built-up, industrialised and agricultural region (the Netherlands) and a region with both large cities and intensive industrial and agricultural areas, as well as more extensively used or semi-natural areas (central Scotland). Both domains also have NH_3 and NO_x emission inventory data at a ca. $1 \times 1 \, \text{km}^2$ resolution. Spatially distributed annual NH_3 and NO_x emission data for the study year (2008) were obtained from the National Atmospheric Emissions Inventory (http://naei.defra.gov.uk/) for the Scottish domain and from the National Institute for Public Health and the Environment (RIVM), for the Netherlands (Fig. 1). In order to evaluate AQR for an emission dataset with a lower spatial resolution that could be used for a Europe-wide application of the model, the 2008 "EC4MACS" emissions with a spatial resolution of ca. $7 \times 7 \, \text{km}^2$ (EC4MACS, 2012, also used in Schaap et al., 2015) were also used for the two domains.

In order to parameterise the pollutant dispersion from source areas, three different atmospheric dispersion models were used. These were ADMS (v4.1) (Carruthers et al., 1994), AERMOD (v12345) (Cimorelli et al., 2002) and LADD (Dragosits et al., 2002). These three models were chosen because they have been extensively evaluated for the atmospheric dispersion of NO_2 and NH_3, with the exception of LADD, which has only been evaluated for NH_3 (Theobald et al., 2012). The meteorological data used for the atmospheric dispersion simulations were derived from the meteorological data used in the EMEP model simulation (generated by the Weather Research Forecast (WRF) model version 3.6.1; http://www.wrf-model.org). Surface and vertical profile data at the centre of each EMEP model grid square were extracted in AERMOD format using the Mesoscale Model Interface Program (MMIF; https://www3.epa.gov/ttn/scram/dispersion_related.htm#mmif) and subsequently converted into the input formats for ADMS and LADD. In order to test the sensitivity of AQR to the meteorological data used, additional simulations were carried out using two domain-specific real meteorological datasets and a synthetic meteorological dataset derived from data from an arbitrary location. The two domain-specific datasets used were from the Easter Bush experimental site, for Scotland (von Bobrutzki et al., 2010), and Cabauw, for the Netherlands (obtained from the Cesar Database: http://www.cesar-database.nl/). The synthetic dataset was derived from data from the Lyneham meteorological station in the UK for 1995 (LYNE95) (Spanton et al., 2004), which was a fairly typical year with regards to mean air temperature and wind speed. This dataset was chosen because it has been used in various model evaluation studies and has been made freely available to the dispersion modelling community (e.g. Hall et al., 2000; Theobald et al., 2012). This dataset was modified (LYNE95mod) by ran-
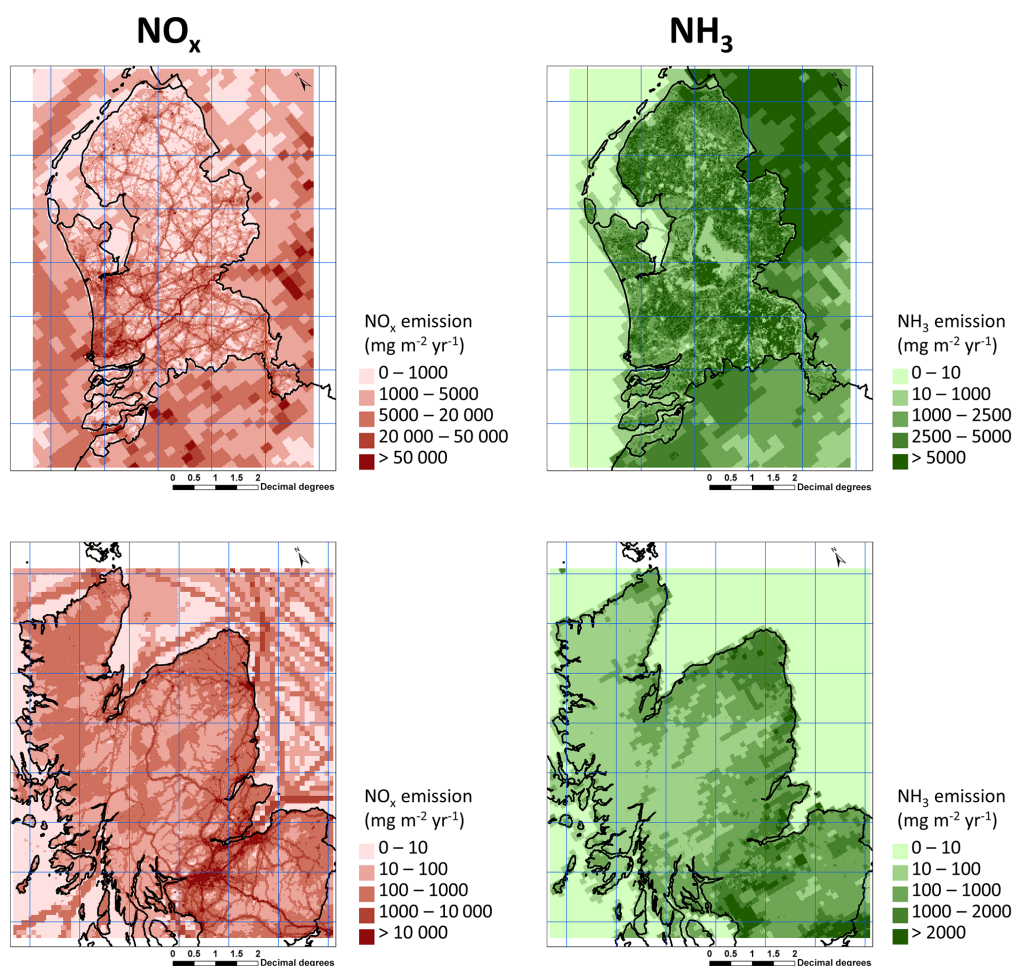
Figure 1. Spatial distributions of annual emissions of NO_x (left) and NH_3 (right), for the Dutch (top) and Scottish (bottom) domains. The EMEP $50 \times 50\,km^2$ grid is also shown (in blue).

domising the wind direction data and scaling the wind speed so that the annual mean value was equal to the annual domain mean value used in the EMEP model for the 2008 study year $(5.1\,m\,s^{-1})$. The wind directions were randomised for two reasons: (1) to make the meteorological data less location-specific so that they can be used within different modelling domains and (2) to provide a generic dispersion dataset that could be of use to the air-quality modelling community.

Evaluation of the AQR model was carried out using 2008 annual mean concentration data from local and national monitoring networks in the two study domains. For Scotland, NO_2 data were obtained from the Air Quality in Scotland website (http://www.scottishairquality.co.uk/) (48 stations: 37 traffic and 11 non-traffic sites) and from RIVM for the Netherlands (43 stations: 13 traffic and 30 non-traffic). The evaluation was done for all sites and for the traffic and non-traffic sites separately since the traffic sites are strongly influenced by the exact site location and are unlikely to be representative of a $1 \times 1\,km^2$ grid square. For NH_3 concentrations in the Scottish domain, monitoring data were ob-

tained from the UK National Ammonia Monitoring Network (NAMN) (Sutton et al., 2001) (http://uk-air.defra.gov.uk/networks/network-info?view=_nh3), which has 14 sites within the domain. In addition, NH_3 monitoring data from 21 sites in a local network covering $36\,km^2$ (Vogt et al., 2013) were also used. For the Netherlands, NH_3 concentration data from the Measuring Ammonia in Nature (MAN) network (Lolkema et al., 2015) were provided by RIVM (108 stations). Model performance was assessed using the evaluation statistics of the R package "Openair" (Carslaw and Ropkins, 2012). Four performance metrics were used to compare the modelled concentrations with the observed values: fraction of model predictions within a factor of 2 of the observations (FAC2), normalised mean bias (NMB), normalised mean gross error (NMGE) and the Pearson correlation coefficient (r) (see Appendix A for definitions).
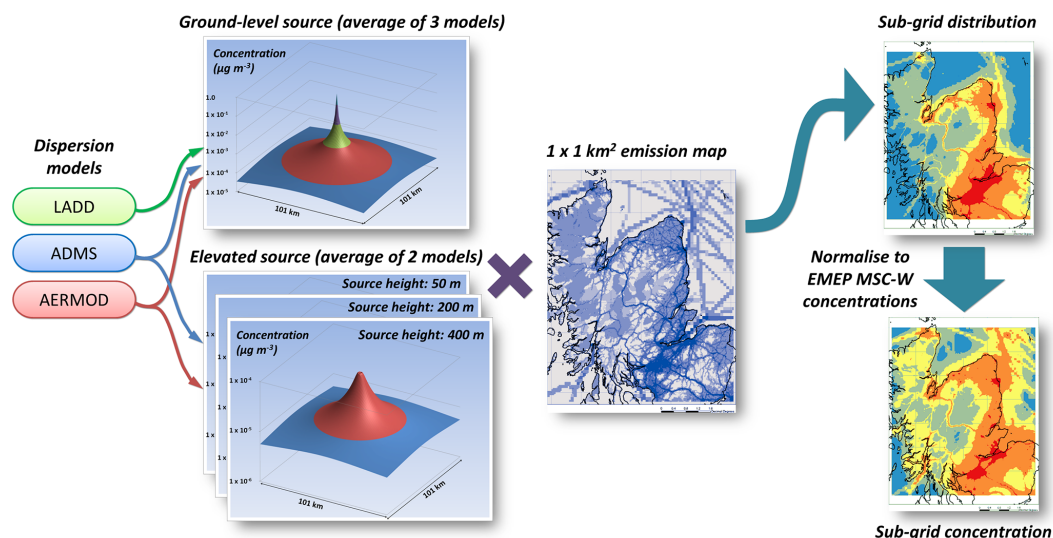
Figure 2. Schematic showing the process of producing the sub-grid concentration predictions from short-range dispersion model simulations and high-spatial-resolution emission data.

3 Model development

The sub-grid $1 \times 1 \, \text{km}^2$ concentration estimates were calculated from three components: the EMEP $50 \times 50 \, \text{km}^2$ concentration predictions, the $1 \times 1 \, \text{km}^2$ emission data and an estimate of short-range ($< 50 \, \text{km}$) pollutant dispersion. Figure 2 shows a schematic of the process. Short-range pollutant dispersion was parameterised using a simple scenario of a single $1 \times 1 \, \text{km}^2$ source with an emission rate of $1 \, \text{Mg} \, \text{km}^{-2} \, \text{yr}^{-1}$ in the centre of a square domain (of dimensions $101 \times 101 \, \text{km}^2$). Although individual sources are generally smaller than this, this value was used to match the spatial resolution of the emission data. For NO_2, the assumption was made that annual mean NO_2 concentrations are linearly correlated with those of NO_x. This allowed us to use the NO_x emissions for the calculation of NO_2 concentrations without considering photochemical reactions. An analysis of the 2008 mean annual concentrations for the 1478 sites in the Air Quality e-Reporting database (formerly AirBase) of the European Environment Agency shows that measured NO_2 and NO_x concentrations are approximately linearly correlated with a linear correlation coefficient, r^2, of 0.93. For the dispersion of NH_3, the source was assumed to be at ground level (a suitable approximation for most agricultural sources, which account for more than 90 % of emissions in Europe). Emissions of NO_x, on the other hand, can occur over a range of emission heights, depending on the source type. Since the emission height will affect the resulting NO_2 concentrations at ground level, it needs to be taken into account. This was done by assigning a representative emission height for each emission sector (Selected Nomenclature for Air Pollution (SNAP) code) that contributed more than 1 % of the total domain emissions (Table 1). These emis-

sion heights correspond approximately to the mean effective emission heights used in the EMEP model for the sector emissions. In order to test the sensitivity of the AQR model to the emission heights used, additional simulations were carried out using emission heights half and double these values. For the ground-level source, all three dispersion models (ADMS, AERMOD and LADD) were used to simulate the annual mean near-ground-level concentrations of NH_3 and NO_2 on a 1 km grid (for the $101 \times 101 \, \text{km}^2$ domain). For the elevated source scenarios, only ADMS and AERMOD were used to simulate the annual mean concentrations because the LADD model is not suitable for simulating dispersion from elevated sources (Theobald et al., 2012). A height of 1.5 m was used for the near-ground-level concentrations, because this height is commonly used for concentration monitoring and impact assessments (Cape et al., 2009). These short-range dispersion simulations were carried out using the meteorological data extracted from the WRF simulations at the centre of each EMEP model grid square. No removal processes (chemical reactions, dry or wet deposition, etc.) were simulated because these processes depend strongly on local conditions (concentrations of other chemical species, meteorological conditions, surface characteristics, etc.).

The result of these simulations was nine concentration fields (kernels), three for ground level sources (three models × one source height) and six for elevated sources (two models × three source heights) for each meteorological dataset (corresponding to each of the EMEP model grid squares). A model-average dispersion kernel (D) for each source height was obtained by taking the mean value of the dispersion model concentration estimates for each kernel grid cell.

Table 1. Emission heights used for each main emission sector.

SNAP code	Emission sector	Effective emission height used (m)
1	Combustion in energy and transformation industries	400
2	Non-industrial combustion plants	0
3	Combustion in manufacturing industry	400
4	Production processes	50
7	Road transport	0
8	Other mobile sources and machinery	0
9	Waste treatment and disposal	200

These model-average kernels were then combined with the emission data using a moving window approach to obtain the sub-grid concentration estimate (C):

$$C(i, j) = \sum_{s}^{\text{SNAP}} \sum_{i'}^{n} \sum_{j'}^{m} E(i', j') \, D'(i - i', j - j'),$$

where i and j are the sub-grid-cell coordinates, s is the emission sector, i' and j' are the emission grid cell coordinates, E is the emission rate of the emission grid cell ($\text{Mg km}^{-2} \text{yr}^{-1}$) and D' an interpolated dispersion kernel (inverse distance squared weighted interpolation of the kernels for the source EMEP grid square and the eight adjacent grid squares). Since the dispersion kernel has a size of 101×101 grid cells, the values of i' and j' range from $i-50$ and $j-50$ to $i+50$ and $j+50$, respectively, with the constraint that they lie within the modelling domain.

The resulting "sub-grid distributions" provide an estimate of the spatial variability of the concentrations at a $1 \times 1 \text{ km}^2$ resolution, which were then used to "redistribute" the EMEP predictions within each $50 \times 50 \text{ km}^2$ grid square. This step is necessary since AQR does not take into account large-scale processes such as long-range transport or chemical transformations of pollutants, processes that are included in the large-scale model (the EMEP model, in this case). The simplest way to do this redistribution would be to multiply the sub-grid distributions by the EMEP predictions and then divide by the mean value of the sub-grid distribution for each $50 \times 50 \text{ km}^2$ grid square. This approach conserves the sub-grid distribution for each $50 \times 50 \text{ km}^2$ square and also has the same mean concentration as the EMEP prediction. However, it also could lead to large discontinuities at the edges of the EMEP grid squares if the ratio between the mean of the sub-grid distribution and the EMEP prediction differ greatly from that of adjacent squares. To avoid this problem, the ratio of the EMEP predictions to the mean value of the sub-grid distribution for each $50 \times 50 \text{ km}^2$ square was interpolated on a $1 \times 1 \text{ km}^2$ grid (using a spline interpolation of the values at the centre of each grid square in ArcGIS 10.2 (Environmental Systems Research Institute, Redlands, CA, USA)). The interpolated field was then multiplied by the sub-grid distribution and then the process was repeated over 10 iterations.

In fact only four to five iterations were necessary to give concentration fields that differed by a maximum of 1 %. A more detailed description of the process is provided in the Supplement.

In order to test the sensitivity of the model to the meteorological data, the above process was repeated with the kernels obtained from the dispersion simulations, using the domain-specific meteorological data and with kernels derived from the dispersion simulations using the synthetic meteorological data (more details provided in the Supplement).

4 Results

4.1 Sub-grid concentration predictions and model evaluation

Figure 3 shows the sub-grid concentration predictions for NO_2 and NH_3 for the two domains (data for the individual domains are provided in Fig. S3.1 in the Supplement). The EMEP concentration fields are also shown for comparison. Table 2 shows the evaluation statistics of the EMEP and AQR models for annual mean NO_2 concentrations for the Dutch and Scottish monitoring data. In general, AQR is an improvement on the EMEP model alone because the latter generally underestimates concentrations (negative NMB). The mean error of the EMEP model is largest for the Scottish dataset with a NMGE of 82 and 70 % for the datasets with and without traffic stations, respectively. The model performs worst for the Scottish traffic stations with a mean underestimation of 84 %. The EMEP model performs considerably better for the Dutch dataset, with 91 % of predictions within a factor of 2 of the observed values, although this drops to 69 % when considering the traffic stations only. The AQR model (using $1 \times 1 \text{ km}^2$ emissions) also performed best for the Dutch dataset, with a smaller mean bias and error and a better correlation than the EMEP model alone. However, the EMEP model had a lower mean bias and error for the non-traffic stations. The AQR model is also an improvement on the EMEP model alone for the Scottish dataset (both with and without traffic stations), as well as for the combined dataset (Netherlands plus Scotland). Similarly to the EMEP
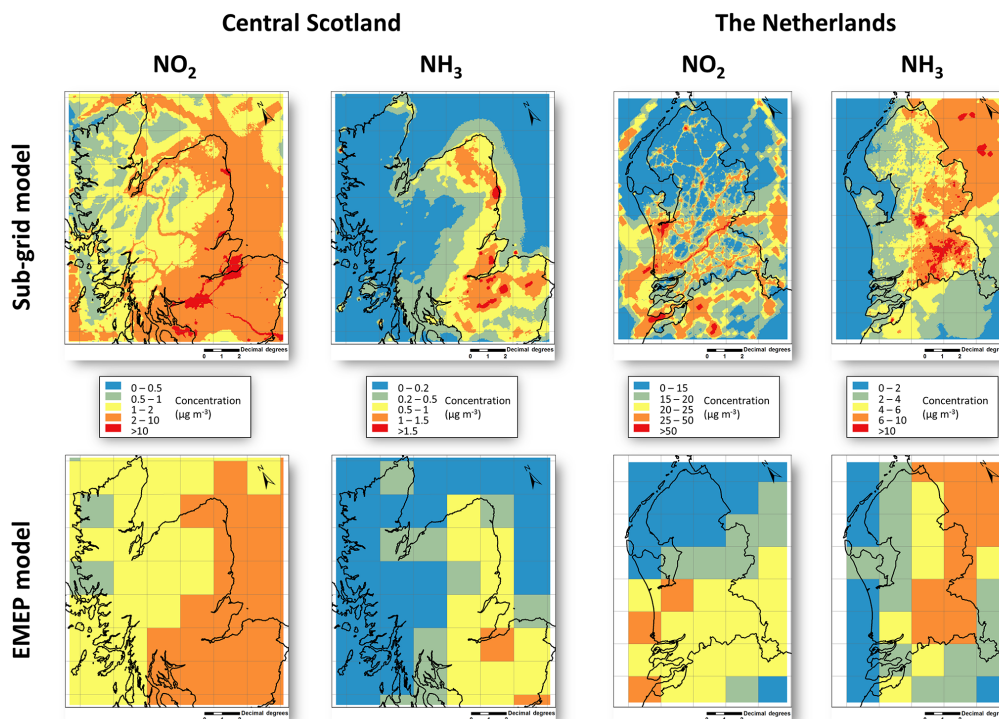
Figure 3. Sub-grid model predictions (top row) of annual mean concentrations of NO$_2$ and NH$_3$ for the two domains. EMEP model predictions at a resolution of 50×50 km^2 are shown for comparison (bottom row).

model, AQR performed worst for the Scottish traffic stations, although was a notable improvement over the EMEP model alone. The use of the lower-resolution emissions actually improved the performance of AQR for some of the statistics (most notably for the non-traffic stations in the Netherlands domain).

Table 3 shows the evaluation statistics of the EMEP and AQR models for annual mean NH$_3$ concentrations for the Dutch and Scottish monitoring data. In general AQR was an improvement on the EMEP model alone, which performed worse for the local monitoring network, as all monitoring locations were within a single EMEP 50×50 km^2 square. The AQR model (using 1×1 km^2 emissions) also performed worst for this dataset, although its performance was still an improvement on that of the EMEP model alone, as it was for all the datasets except for the National Ammonia Monitoring Network sites in Scotland. The use of the 7×7 km^2 emissions worsened the performance of AQR (with respect to the simulations using the 1×1 km^2 emissions) for all datasets except for the National Ammonia Monitoring Network sites, for which it had a similar performance to the model using the higher-resolution emissions. Figure 4 shows the scatterplots of NO$_2$ and NH$_3$ concentration predictions of the EMEP and AQR models vs. the observed values for all sites in both domains.

4.2 Sensitivity of sub-grid model predictions to model parameters

The use of alternative meteorological datasets only had a small effect on the concentration estimates of the AQR model (Fig. 5). The use of domain-specific data from a single location affected the concentration predictions by an average of 6 % for NO$_2$ and 5 % for NH$_3$ although differences of up to 23 % were found for individual measurement sites. Similarly, the use of the synthetic meteorological data affected concentrations, on average, by 6 and 5 % for NO$_2$ and NH$_3$, respectively, with a maximum difference of 28 %. Randomising the wind direction data of the domain-specific datasets gave very similar results to those using the synthetic meteorology dataset, with maximum differences of only 1 % (not shown). This suggests that the meteorological factor that most influences the estimates of the AQR model is the wind direction distribution.

The AQR model estimates are also not very sensitive to the NO$_x$ emission height. On average, the effect on the concentration predictions of halving or doubling the emission heights is less than 2 %, with a maximum difference of 6 % (not shown). This lack of sensitivity to the exact source height reflects the fact that ground-level sources contribute significantly more to near-source concentrations than elevated sources. Since the concentrations predicted by AQR were not greatly affected by the meteorological data or the

Table 2. Performance evaluation of the EMEP and sub-grid models for annual mean NO_2 concentrations. The best-performing model for each statistic is highlighted in bold. FAC2 is the fraction of model predictions within a factor of 2 of the observations, NMB is the normalised mean bias, NMGE is the normalised mean gross error and r is the Pearson correlation coefficient. Italic font highlights the model performance for the sub-grid model using the lower-resolution emission data.

Dataset	n	EMEP model				Sub-grid model				
		FAC2	NMB	NMGE	r	Emission data	FAC2	NMB	NMGE	r
Netherlands (All)	43	0.91	−0.24	0.31	0.54	$1 \times 1\,km^2$	0.98	**0.05**	0.24	**0.83**
						$7 \times 7\,km^2$	*1.0*	*−0.08*	*0.21*	*0.79*
Netherlands (No traffic stations)	30	**1.00**	−0.06	**0.18**	0.73	$1 \times 1\,km^2$	0.97	0.07	0.29	**0.86**
						$7 \times 7\,km^2$	*1.0*	*0.01*	*0.21*	*0.81*
Netherlands (Traffic stations only)	13	0.69	−0.45	0.45	0.17	$1 \times 1\,km^2$	**1.0**	**0.02**	**0.16**	0.48
						$7 \times 7\,km^2$	*1.0*	*−0.18*	*0.21*	*0.32*
Scotland (All)	48	0.06	−0.82	0.82	0.16	$1 \times 1\,km^2$	**0.46**	**−0.50**	**0.54**	0.43
						$7 \times 7\,km^2$	*0.23*	*−0.63*	*0.63*	*0.51*
Scotland (No traffic stations)	11	0.27	−0.70	0.70	0.40	$1 \times 1\,km^2$	**0.91**	**−0.08**	**0.30**	0.80
						$7 \times 7\,km^2$	*0.64*	*−0.38*	*0.39*	*0.85*
Scotland (Traffic stations only)	37	0.00	−0.84	0.84	0.05	$1 \times 1\,km^2$	**0.32**	**−0.56**	**0.57**	0.48
						$7 \times 7\,km^2$	*0.11*	*−0.67*	*0.67*	*0.51*
All	91	0.46	−0.58	0.61	–	$1 \times 1\,km^2$	**0.70**	**−0.27**	**0.41**	**0.39**
						$7 \times 7\,km^2$	*0.59*	*−0.40*	*0.46*	*0.27*

Table 3. Performance evaluation of the EMEP and sub-grid models for annual mean NH_3 concentrations. The best-performing model for each statistic is highlighted in bold. FAC2 is the fraction of model predictions within a factor of 2 of the observations, NMB is the normalised mean bias, NMGE is the normalised mean gross error and r is the Pearson correlation coefficient. Italic font highlights the model performance for the sub-grid model using the lower-resolution emission data.

Dataset	n	EMEP				Sub-grid model				
		FAC2	NMB	NMGE	r	Emission data	FAC2	NMB	NMGE	r
Netherlands	108	0.85	0.23	0.39	0.69	$1 \times 1\,km^2$	**0.92**	**0.10**	**0.24**	**0.84**
						$7 \times 7\,km^2$	*0.90*	*0.28*	*0.40*	*0.71*
Scotland – Local network	21	0.52	−0.47	0.65	–	$1 \times 1\,km^2$	**0.62**	**−0.04**	**0.52**	**0.55**
						$7 \times 7\,km^2$	*0.52*	*−0.48*	*0.66*	*–*
Scotland (National Ammonia Monitoring Network)	14	**0.71**	**0.07**	0.46	0.73	$1 \times 1\,km^2$	0.57	0.26	**0.45**	0.77
						$7 \times 7\,km^2$	*0.57*	*0.07*	*0.43*	*0.81*
All	143	0.79	0.17	0.42	0.74	$1 \times 1\,km^2$	**0.84**	**0.09**	**0.27**	**0.84**
						$7 \times 7\,km^2$	*0.81*	*0.20*	*0.42*	*0.75*

emission heights, model performance was very similar (not shown).

5 Discussion

5.1 An improvement, but is it enough?

These results show that a simple and robust geostatistical approach can be used to improve the EMEP model predictions of NO_2 and NH_3 annual concentrations. This improvement is not surprising considering the large difference in spatial resolutions (50 km vs. 1 km) and the strong link between short-lived pollutants and the spatial distribution of emissions. In fact, it is worth looking at whether this improvement is mainly a result of the high-resolution emissions and

has very little to do with the use of short-range dispersion estimates. This can be done by repeating the analyses with the $1 \times 1\,km^2$ grid cell emissions as the initial sub-grid distribution. Figure 6 shows that doing this for NO_2 substantially overestimates concentrations for the mid-range of measured values, whereas for NH_3, concentrations are substantially underestimated at many sites. The model performance statistics for these simulations show that using just the emissions gives lower values of FAC2 (0.60 vs. 0.70 for NO_2 and 0.28 vs. 0.84 for NH_3) and larger bias and error (NMB: 0.36 vs. −0.27 for NO_2 and −0.36 vs. 0.09 for NH_3; NMGE: 0.72 vs. 0.41 for NO_2 and 0.79 vs. 0.27 for NH_3). Model error is even larger than that for the EMEP model alone (0.72 vs. 0.61 for NO_2 and 0.79 vs. 0.42 for NH_3), which demonstrates that
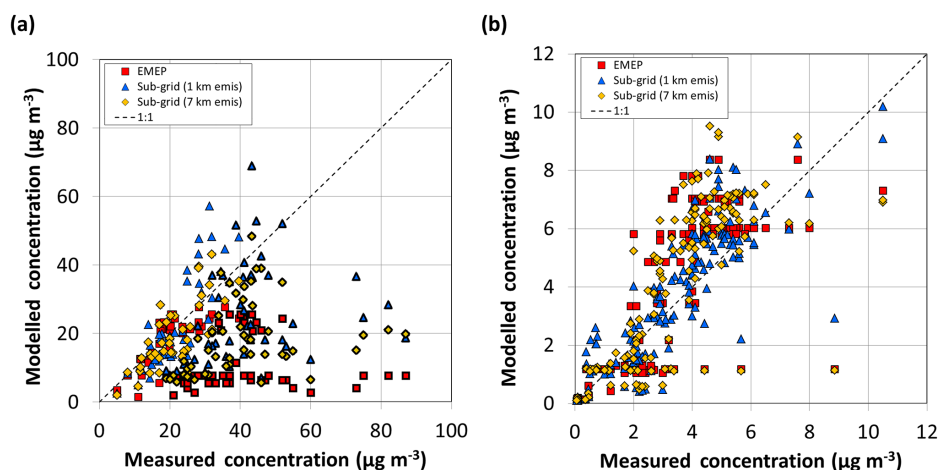
Figure 4. Modelled concentrations plotted against measured values for all sites for **(a)** NO$_2$ and **(b)** NH$_3$. NO$_2$ traffic stations are indicated by bold symbol outlines. Plot data provided in the Supplement.
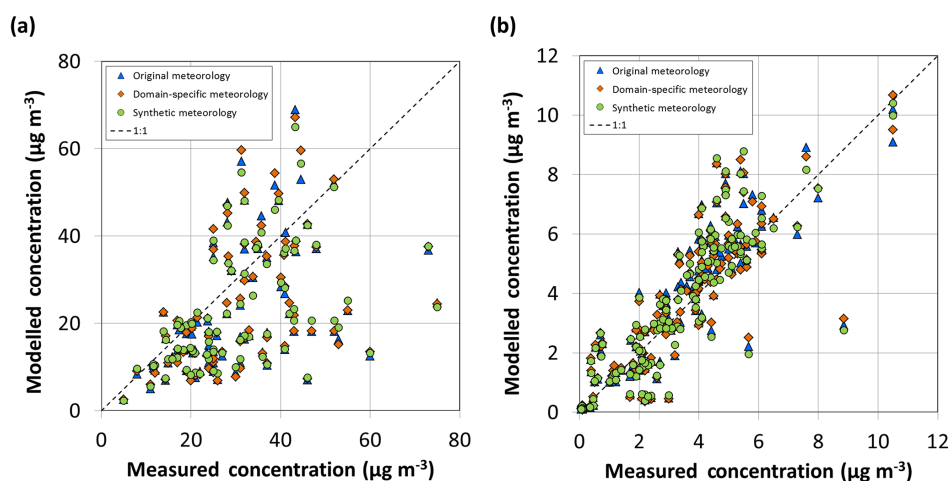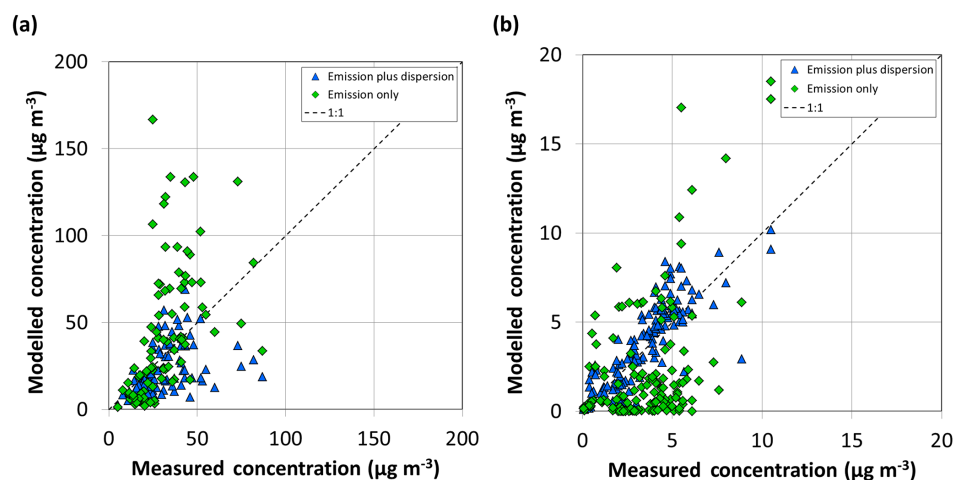

Figure 5. Modelled concentrations plotted against measured values for all sites for **(a)** NO$_2$ and **(b)** NH$_3$ using the original meteorology (as in Fig. 4) and using the domain-specific and synthetic meteorological datasets.

short-range dispersion estimates are necessary for improving on the EMEP model predictions.

However, is the improvement of AQR over the EMEP model alone large enough to warrant the inclusion of such a sub-grid model into the output processing options of a chemical transport model? In order to answer this question, we can use the concept of model acceptability suggested by Chang and Hanna (2004). This concept can be used to evaluate whether the EMEP model and/or the AQR model perform acceptably and, therefore, whether the AQR model represents an improvement on the EMEP model alone, in terms of model acceptability. Hanna and Chang (2012) suggested that an "acceptable" model is one that meets the criteria for more than half of a series of statistical tests. The performance metrics used are fractional bias, geometric mean bias, normalised mean square error, geometric variance and FAC2

(see Appendix A for definitions and acceptability criteria). In the current study, we define an acceptable model as one that meets at least three of these five criteria (for each dataset). Although the concept of model acceptability of Chang and Hanna (2004) was defined for research-grade experimental data, the fact that we are considering annual mean concentrations (instead of high-temporal-resolution measurements) should make the approach suitable for use with operational models and monitoring data, such as those used here. For the two combined datasets (NO$_2$-All and NH$_3$-All) shown in Fig. 4, the EMEP model meets none and five of the five criteria for NO$_2$ and NH$_3$, respectively, whereas AQR meets three and five criteria, respectively (Table 4). This suggests that AQR is a significant improvement (in terms of model acceptability) for NO$_2$ (even when the lower-resolution emission dataset is used), but not for NH$_3$, since the EMEP model

Figure 6. Modelled concentrations plotted against measured values for all sites for **(a)** NO_2 and **(b)** NH_3 using the original sub-grid parameterisation (Emission plus dispersion) and using just the spatial distribution of emissions as the sub-grid distribution (Emission only).

Table 4. Number of model acceptability criteria met for each model and evaluation dataset. Bold font represents acceptable model performance (≥ 3 criteria met).

Pollutant	Dataset		No. of criteria met		
			EMEP	Sub-grid ($1 \times 1\,km^2$ emissions)	Sub-grid ($7 \times 7\,km^2$ emissions)
	Netherlands	All	**5**	**5**	**5**
	Netherlands	No traffic stations	**5**	**5**	**5**
	Netherlands	Traffic stations only	**3**	**5**	**5**
NO_2	Scotland	All	0	2	1
	Scotland	No traffic stations	0	**5**	**3**
	Scotland	Traffic stations only	0	2	0
	All		0	**3**	**3**
	Netherlands		**5**	**5**	**5**
NH_3	Scotland	Local network	2	**5**	2
	Scotland	National Network	**5**	**5**	**5**
	All		**5**	**5**	**5**

alone already performs acceptably for this dataset. This can be explained by looking at the number of criteria met for the individual datasets (Table 4). For NO_2, The EMEP model performed acceptably for the Netherlands (All) but not for Scotland (All). This is partly due to the Dutch network having a larger proportion of non-traffic sites (70 % vs. 23 %), which would be more representative of the $50 \times 50\,km^2$ grid cells. However, the EMEP model also performed acceptably for the Dutch traffic stations but neither the EMEP model nor the AQR model performed acceptably for the Scottish traffic stations. Looking more carefully at the traffic stations used in the domains reveals that station siting may have an influence on model performance. According to the information available regarding the Scottish traffic sites, monitoring stations

are located between 0.5 and 16 m from the road edge. Although no information is available regarding the exact locations of the Dutch monitoring stations, Nguyen et al. (2013) point out that one station in the Amsterdam Municipal Health Service (GGD) network (not used in this study) "is very close to the road ($< 2.5\,m$)". This suggests that, in general, sites in the Dutch network are $> 2.5\,m$ from the road, whereas in the Scottish network 17 of the 37 traffic sites are closer than this. This difference in station siting could be the reason why neither the EMEP nor the AQR model performed acceptably for the Scottish dataset. For NH_3, the EMEP and AQR models performed acceptably for the two national networks but only AQR performed acceptably for the local network. This is probably because the national networks site

their monitoring stations far from the influence of individual emission sources in order to be representative of a large area, whereas the local network was located in an area with intensive poultry farming and was designed to assess the influence of individual sources. Since the majority (86 %) of the sites used in the analysis belonged to the national networks, overall model performance was similar to model performance for those networks. The sub-grid approach, therefore, is most useful where there are large horizontal concentration gradients, such as within large cities (for NO_2) or areas with intensive agriculture (for NH_3), which is where the largest impacts are most likely to occur.

It is also worth briefly comparing the improvements in model performance with those reported by other studies. Denby et al. (2011) showed that the population-weighted concentration for NO_2 was, on average, 44 % higher with their sub-grid parameterisation than that calculated using the original concentrations from the EMEP model. Although not directly comparable (since we do not calculate population-weighted concentrations), NO_2 concentrations estimated using the AQR model are, on average, 77 % higher than those of the EMEP model at the monitoring station locations. Despite this increase, the AQR estimates are still, on average, 27 % lower than the measured concentrations. Janssen et al. (2012) showed that their approach of downscaling modelled concentrations from $15 \times 15 \, km^2$ to $3 \times 3 \, km^2$ reduced model error by about 20 %. The AQR model for NO_2 reduced model error by 30–40 %, although for a larger change in resolution ($50 \times 50 \, km^2$ to $1 \times 1 \, km^2$). In the study by Schaap et al. (2015), increasing the spatial resolution from approx. $56 \times 56 \, km^2$ to $7 \times 7 \, km^2$ increased the correlation (r) between the models' predictions and hourly urban background NO_2 concentrations from approx. 0.1–0.4 to 0.6–0.7 and reduced model bias by approx. 60–90 % for most of the models. For a similar change in spatial resolution ($50 \times 50 \, km^2$ to $7 \times 7 \, km^2$), the AQR model for annual mean NO_2 concentrations using the low-resolution emissions increased r from 0.16–0.54 to 0.51–0.85 and reduced model bias by approx. 20–70 %.

5.2 How can the sub-grid approach be applied?

Two potential uses of the sub-grid approach can be envisaged: a Europe-wide application to provide a spatial assessment of exceedance of NO_2 and NH_3 annual limit values or critical levels and the assessment of individual emission hotspots in areas where detailed modelling assessments are not available but high-resolution emission data are. In the latter case, if the hotspot domain is located within a single EMEP $50 \times 50 \, km^2$ grid square, the smoothing step would not be necessary. The Europe-wide application would require high-spatial-resolution emission data for the whole domain. There is, as far as we are aware, currently no European emission inventory with a spatial resolution close to $1 \times 1 \, km^2$. The highest resolutions available are the $7 \times 7 \, km^2$ emission

inventories produced for various EU projects (Kuenen et al., 2014; EC4MACS, 2012). As shown above, the use of emission data at this resolution still gives an improvement on the concentration predictions and even performs better than the sub-grid model using the higher-resolution emissions, in some cases.

5.3 Advantage, disadvantages, uncertainties and potential improvements

The AQR model can provide more accurate concentration predictions than the EMEP model alone, especially close to emission sources. However, this approach has only been tested for annual mean NO_2 and NH_3 concentrations, although it could potentially be extended to other short-lived pollutants and shorter timescales (daily or hourly). This means that the model cannot currently be used to assess exceedance of short-term limit values (e.g. for Europe, an hourly mean concentration of $200 \, \mu g \, NO_2 \, m^{-3}$ more than 18 times in one year) although, as shown by Kiesewetter et al. (2013), the annual mean limit values for NO_2 are the more stringent target. Critical levels for ammonia are expressed as annual mean concentrations and so a sub-grid model with a higher temporal resolution is not necessary. The other limitation of the approach is the need for high-resolution emission data although, as shown above, the use of emission data with a resolution of $7 \times 7 \, km^2$ already produces improvements in model performance compared with the original ACTM concentration estimates.

The various assumptions and simplifications made in the development of AQR introduce uncertainty in the model estimates. The omission of NO_x photochemistry and the assumption that annual mean NO_2 concentrations are linearly correlated with those of NO_x was justified above by the fact that measured concentrations across Europe are approximately linearly correlated ($r^2 = 0.93$). However, a more in-depth analysis of the European measurements shows that if a constant factor is used to estimate NO_2 concentrations from the measurements of NO_x, the estimated NO_2 concentrations differ from the measure values by an average of 16 %, which is a small uncertainty compared with the uncertainty in emissions, meteorological conditions, etc. The uncertainty as a result of not modelling the chemical transformation of NH_3 (e.g. to particulate ammonium) is more difficult to quantify since the reactions depend on many factors such as the meteorological conditions and the concentrations of other pollutants. However, the fact that the errors (NMGE) in the AQR estimates of NH_3 concentrations are of a similar order of magnitude to the errors in the NO_2 estimates suggests that the benefits of AQR in handling sub-grid distributions outweigh any chemical impacts. In addition, such errors would be largest far from the sources, once NH_3 concentrations are diluted more to levels comparable to incoming sulphate or HNO_3 concentrations. Another source of uncertainty is the omission of deposition processes in the short-range disper-

sion parameterisations, but wet-deposition has been implicitly included in ACTM predictions, and timescales for dry deposition are usually far larger than those for sub-grid mixing. Again, given the AQR model has a mean error of 41 and 27 % for NO_2 and NH_3, respectively, the benefits of AQR seem greater than any uncertainty as a result of omitting these processes. Finally, another simplification is the use of a $1 \times 1\,km^2$ source for parameterising short-range dispersion. In reality sources are generally smaller than this and so this simplification may result in incorrect concentration gradients close to small or linear NO_x sources (e.g. chimney stacks or motorways). However, on average, transport emissions contribute more than 90 % of the estimated concentrations, most of which are in urban areas where a $1 \times 1\,km^2$ source is probably an adequate representation of a dense urban road network. In addition, we rarely know the location of stacks in emission inventories to better than 1 km resolution, and usually with no or very limited information on plume rise and height.

With regards to potential improvements, in addition to the extension to shorter time periods, it also should be possible to incorporate stack parameters (effective emission heights and the contribution of stack emissions to the emissions of a particular grid square) from officially reported data and/or other data sources, if these become more readily available. This would potentially improve concentration estimates close to large stack sources. As shown above, model performance is poorer for sites very close to roads and so the inclusion of a roadside increment model could also improve the model estimates. However, by increasing the complexity of the model, we have to be careful not to lose sight of the objective of the AQR model, which is to provide a robust and simple method of post-processing concentrations estimated by an ACTM.

The sub-grid approach also has the potential to be applied to other pollutants for which there is a strong relationship between emissions and concentrations. Zhang and Wu (2013) analysed air-quality simulations of the CMAQ model to quantify the influence of a range of processes on the atmospheric concentrations of several pollutants. The species that were most strongly influenced by emission processes were NH_3, NO, NO_2, SO_2, $PM_{2.5}$, SO_4^{2-}, elemental carbon and primary organic aerosol and are, therefore, potential candidates for an extension of the model. The spatial distribution of ozone, a secondary pollutant, cannot be estimated based on emissions but its inverse relationship with NO_x could be exploited to model the sub-grid variability. Apart from concentrations, it may also be possible to develop a sub-grid model for processes such as wet deposition of nitrogen or sulphur, for which high-resolution rainfall maps could be used to estimate the sub-grid distributions. Dry deposition of reduced nitrogen could also be modelled using the NH_3 concentration distribution and land-cover parameters, assuming that most of the deposition is in the form of NH_3. Dry deposition of oxidised nitrogen would be more difficult since there is no one dominant species that contributes.

6 Conclusions

The sub-grid spatial variability of the annual mean NO_2 and NH_3 concentrations predicted by an atmospheric chemistry and transport model can be estimated by combining the predictions with high-spatial-resolution emission datasets and short-range dispersion fields. This paper describes the development of the Air Quality Re-gridder (AQR) model and its application to two test domains in Europe. Comparison of annual mean concentrations estimated by AQR with measured values within both domains shows that the AQR model represents an improvement on the predictions of the atmospheric chemistry and transport model, reducing both model error and bias and increasing the spatial correlation with the measured concentrations.

7 Code/data availability

The AQR model code (in the R programming language) plus example input and output files for the simulations using synthetic meteorological data are provided in the Supplement.

The data shown in Figs. 4, 5 and 6 are provided in the Supplement.

Appendix A: Descriptions of the performance metrics used

Table A1. The four metrics relating modelled concentrations (M_i) with the observed values (O_i), used for evaluating model performance.

Performance measure	Definition
Fraction of model predictions within a factor of 2 of the observations (FAC2):	$0.5 \leq \frac{M_i}{O_i} \leq 2.0$
Normalised mean bias:	$\text{NMB} = \dfrac{\sum\limits_{i=1}^{n} M_i - O_i}{\sum\limits_{i=1}^{n} O_i}$
Normalised mean gross error:	$\text{NMGE} = \dfrac{\sum\limits_{i=1}^{n} \lvert M_i - O_i \rvert}{\sum\limits_{i=1}^{n} O_i}$
Pearson correlation coefficient:	$r = \frac{1}{(n-1)} \sum\limits_{i=1}^{n} \left(\frac{M_i - \overline{M}}{\sigma_M} \right) \left(\frac{O_i - \overline{O}}{\sigma_O} \right)$

Table A2. The five performance measures relating modelled concentrations (M_i) with the observed values (O_i), used to assess model acceptability.

Performance measure	Definition	Optimum value	Acceptability criterion
Fractional bias (FB)	$\text{FB} = \frac{2\,(\overline{O} - \overline{M})}{(\overline{O} + \overline{M})}$	0	$\lvert \text{FB} \rvert < 0.3$
Geometric mean bias (MG)	$\text{MG} = \exp\left(\overline{\ln O} - \overline{\ln M} \right)$	1	$0.7 < \text{MG} < 1.3$
Normalised mean square error (NMSE)	$\text{NMSE} = \frac{\overline{(O-M)^2}}{\overline{O}\ \overline{M}}$	0	$\text{NMSE} < 1.5$
Geometric variance (VG)	$\text{VG} = \exp\left[\overline{(\ln O - \ln M)^2} \right]$	1	$\text{VG} < 4$
Fraction of model predictions within a factor of 2 of the observations (FAC2)	$0.5 \leq \frac{M_i}{O_i} \leq 2.0$	1	$\text{FAC2} > 0.5$

References

Aggarwal, P. and Jain, S.: Impact of air pollutants from surface transport sources on human health: A modeling and epidemiological approach, Environ. Int., 83, 146–157, 2015.

Amann, M., Bertok, I., Borken-Kleefeld, J., Cofala, J., Heyes, C., Höglund-Isaksson, L., Klimont, Z., Nguyen, B., Posch, M., and Rafaj, P.: Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications, Environ. Modell. Softw., 26, 1489–1501, 2011.

Bessagnet, B., Hodzic, A., Vautard, R., Beekmann, M., Cheinet, S., Honoré, C., Liousse, C., and Rouil, L.: Aerosol modeling with CHIMERE – preliminary evaluation at the continental scale, Atmos. Environ., 38, 2803–2817, 2004.

Cape, J. N., van der Eerden, L. J., Sheppard, L. J., Leith, I. D., and Sutton, M. A.: Evidence for changing the critical level for ammonia, Environ. Pollut., 157, 1033–1037, 2009.

Carruthers, D. J., Holroyd, R. J., Hunt, J. C. R., Weng, W. S., Robins, A. G., Apsley, D. D., Thompson, D. J., and Smith, F. B.: UK-ADMS: A new approach to modelling dispersion in the earth's atmospheric boundary layer, J. Wind. Eng. Ind. Aerod., 52, 139–153, 1994.

Carslaw, D. C. and Ropkins, K.: Openair – an R package for air quality data analysis, Environ. Modell. Softw., 27, 52–61, 2012.

Chang, J. C. and Hanna, S. R.: Air quality model performance evaluation, Meteorol. Atmos. Phys., 87, 167–196, 2004.

Ching, J., Isakov, V., Majeed, M., and Irwin, J.: An approach for incorporating sub-grid variability information into air quality modelling, Proceedings of the 14th Joint Conference on the Applications of Air Pollution Meteorology with the Air and Waste Management Association, Atlanta, USA, 27 January–3 February 2006.

Cimorelli, A. J., Perry, S. G., Venkatram, A., Weil, J. C., Paine, R. J., Wilson, R. B., Lee, R. F., Peters, W. D., Brode, R. W., and Pauimer, J. O.: AERMOD: Description of Model Formulation Version 02222. United States Environmental Protection Agency, Research Triangle Park, NC 27711, USA, 85 pp., 2002.

CLRTAP, Chapter 2 of the Draft Manual on Methodologies and Criteria for Modelling and Mapping Critical Loads & Levels and Air Pollution Effects, Risks and Trends, available at: http://www.rivm.nl/media/documenten/cce/manual/Ch2-MapMan-2014-08.pdf (last access: 12 December 2016), 2014.

Cuvelier, C., Thunis, P., Karam, D., Schaap, M., Hendriks, C., Kranenburg, R., Fagerli, H., Nyiri, A., Simpson, D., Wind, P., Schulz, M., Bessagnet, B., Colette, A., Terrenoire, E., Rouïl, L., Stern, R., Graff, A., Baldasano, J. M., and Pay, M. T.: ScaleDep: performance of European chemistry-transport models as function of horizontal special resolution, EMEP Technical report 1/2013, 63 pp., 2013.

Denby, B., Cassiani, M., de Smet, P., de Leeuw, F., and Horálek, J.: Sub-grid variability and its impact on European wide air quality exposure assessment, Atmos. Environ., 45, 4220–4229, 2011.

Dentener, F., Drevet, J., Lamarque, J. F., Bey, I., Eickhout, B., Fiore, A. M., Hauglustaine, D., Horowitz, L.W., Krol, M., Kulshrestha, U. C., Lawrence, M., Galy-Lacaux, C., Rast, S., Shindell, D., Stevenson, D., Van Noije, T., Atherton, C., Bell, N., Bergman, D., Butler, T., Cofala, J., Collins, B., Doherty, R., Ellingsen, K., Galloway, J., Gauss, M., Montanaro, V., Müller, J. F., Pitari, G., Rodriguez, J., Sanderson, M., Solmon, F., Strahan, S., Schultz, M., Sudo, K., Szopa, S., and Wild, O.: Nitrogen and sulfur deposition on regional and global scales: a multimodel evaluation, Global Biogeochem. Cy., 20, GB4003, doi:10.1029/2005GB002672, 2006.

de Smet, P., de Leeuw, F., Horálek, J., and Kurfürst, P.: A European compilation of national air quality maps based on modelling, ETC/ACM Technical Paper 2013/3, March 2013.

Dragosits, U., Theobald, M. R., Place, C. J., Lord, E., Webb, J., Hill, J., ApSimon, H. M., and Sutton, M. A.: Ammonia emission, deposition and impact assessment at the field scale: a case study of sub-grid spatial variability, Environ. Pollut., 117, 147–158, 2002.

EC4MACS.: The GAINS Integrated Assessment Model, EC4MACS Report, available at: http://www.ec4macs.eu/content/report/EC4MACS_Publications/MR_Finalinpdf/GAINS_Methodologies_Final.pdf (last access: 12 December 2016), 2012.

EMEP: Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components, Status Report 1/2015, EMEP MSC-W, The Norwegian Meteorological Institute, Oslo, Norway, 2015.

Fagerli, H. and Aas, W.: Trends of nitrogen in air and precipitation: Model results and observations at EMEP sites in Europe, 1980–2003, Environ. Pollut., 154, 448–461, 2008.

Galvis, B., Bergin, M., Boylan, J., Huang, Y., Bergin, M., and Russell, A. G.: Air quality impacts and health-benefit valuation of a low-emission technology for rail yard locomotives in Atlanta Georgia, Sci. Total Environ., 533, 156–164, 2015.

Hall, D. J., Spanton, A. M., Dunkerley, F., Bennett, M., and Griffiths, R. F.: An Intercomparison of the AERMOD, ADMS and

ISC Dispersion Models for Regulatory Applications, Report No. P362. Environment Agency, Bristol, UK, 2000.

Hallsworth, S., Dore, A., Bealey, W., Dragosits, U., Vieno, M., Hellsten, S., Tang, Y., and Sutton, M.: The role of indicator choice in quantifying the threat of atmospheric ammonia to the "Natura 2000"network, Environ. Sci. Policy, 13, 671–687, 2010.

Hanna, S. R. and Chang, J.: Setting Acceptance Criteria for Air Quality Models, Air Pollution Modeling and its Application XXI, 479–484, 2012.

Isakov, V., Irwin, J. S. and Ching, J.: Using CMAQ for exposure modeling and characterizing the subgrid variability for exposure estimates, J. Appl. Meteorol. Clim., 46, 1354–1371, 2007.

Janssen, S. and Thunis, P.: Fairmode's composite mapping exercise, Proceedings of the 10th International Conference on Air Quality – Science and Application, Milan, Italy, 14–18 March 2016, available at: http://www.airqualityconference.org/, last access: 12 December 2016.

Janssen, S., Dumont, G., Fierens, F., Deutsch, F., Maiheu, B., Celis, D., Trimpeneers, E., and Mensink, C.: Land use to characterize spatial representativeness of air quality monitoring stations and its relevance for model validation, Atmos. Environ., 59, 492–500, 2012.

Kiesewetter, G., Borken-Kleefeld, J., Heyes, C., Bertok, I., Schoepp, W., Thunis, P., Bessagnet, B., Terrenoire, E., and Amann, M.: Modelling compliance with NO_2 and PM_{10} air quality limit values in the GAINS model. TSAP Report no. 9, IIASA., 2013.

Kiesewetter, G., Borken-Kleefeld, J., Schöpp, W., Heyes, C., Thunis, P., Bessagnet, B., Terrenoire, E., Gsella, A., and Amann, M.: Modelling NO_2 concentrations at the street level in the GAINS integrated assessment model: projections under current legislation, Atmos. Chem. Phys., 14, 813–829, doi:10.5194/acp-14-813-2014, 2014.

Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M., and Denier van der Gon, H. A. C.: TNO-MACC_II emission inventory; a multi-year (2003–2009) consistent high-resolution European emission inventory for air quality modelling, Atmos. Chem. Phys., 14, 10963–10976, doi:10.5194/acp-14-10963-2014, 2014.

Lolkema, D. E., Noordijk, H., Stolk, A. P., Hoogerbrugge, R., van Zanten, M. C., and van Pul, W. A. J.: The Measuring Ammonia in Nature (MAN) network in the Netherlands, Biogeosciences, 12, 5133–5142, doi:10.5194/bg-12-5133-2015, 2015.

Nguyen, P., Stefess, G., de Jonge, D., Snijder, A., Hermans, P., van Loon, S., and Hoogerbrugge, R.: Evaluation of the representativeness of the Dutch air quality monitoring stations: The national, Amsterdam, Noord-Holland, Rijnmond-area, Limburg and Noord-Brabant networks, RIVM report 680704021, 2013.

Oxley, T. and ApSimon, H.: Space, time and nesting integrated assessment models, Environ. Modell. Softw., 22, 1732–1749, 2007.

Schaap, M., Cuvelier, C., Hendriks, C., Bessagnet, B., Baldasano, J. M., Colette, A., Thunis, P., Karam, D., Fagerli, H., Graff, A., Kranenburg, R., Nyiri, A., Pay, M. T., Rouil, L., Schulz, M., Simpson, D., Stern, R., Terrenoire, E., and Wind, P.: Performance of European chemistry transport models as function of horizontal resolution, Atmos. Environ., 112, 90–105, 2015.

Simpson, D., Butterbach-Bahl, K., Fagerli, H., Kesik, M., Skiba, U., and Tang, S.: Deposition and emissions of reactive nitrogen over European forests: a modelling study, Atmos. Environ., 40, 5712–5726, 2006.

Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyíri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, Á., and Wind, P.: The EMEP MSC-W chemical transport model – technical description, Atmos. Chem. Phys., 12, 7825–7865, doi:10.5194/acp-12-7825-2012, 2012.

Spanton, A. M., Hall, D. J., Dunkerley, F., Griffiths, R. F., and Bennett, M.: A Dispersion Model Intercomparison Archive. Proceedings of 9th Int. Conf. on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Garmisch-Partenkirchen, Germany, 1–4 June 2004.

Sutton, M. A., Tang, Y. S., Dragosits, U., Fournier, N., Dore, A. J., Smith, R. I., Weston, K. J., and Fowler, D.: A spatial analysis of atmospheric ammonia and ammonium in the UK, The Scientific World Journal, 1 Suppl 2, 275–286, 2001.

Theobald, M. R., Løfstrøm, P., Walker, J., Andersen, H. V., Pedersen, P., Vallejo, A., and Sutton, M. A.: An intercomparison of models used to simulate the short-range atmospheric dispersion of agricultural ammonia emissions, Environ. Modell. Softw., 37, 90–102, 2012.

Vieno, M., Dore, A. J., Stevenson, D. S., Doherty, R., Heal, M. R., Reis, S., Hallsworth, S., Tarrason, L., Wind, P., Fowler, D., Simpson, D., and Sutton, M. A.: Modelling surface ozone during the 2003 heat-wave in the UK, Atmos. Chem. Phys., 10, 7963–7978, doi:10.5194/acp-10-7963-2010, 2010.

Vieno, M., Heal, M. R., Hallsworth, S., Famulari, D., Doherty, R. M., Dore, A. J., Tang, Y. S., Braban, C. F., Leaver, D., Sutton, M. A., and Reis, S.: The role of long-range transport and domestic emissions in determining atmospheric secondary inorganic particle concentrations across the UK, Atmos. Chem. Phys., 14, 8435–8447, doi:10.5194/acp-14-8435-2014, 2014.

Vogt, E., Dragosits, U., Braban, C. F., Theobald, M. R., Dore, A. J., van Dijk, N., Tang, Y. S., McDonald, C., Murray, S., Rees, R. M., and Sutton, M. A.: Heterogeneity of atmospheric ammonia at the landscape scale and consequences for environmental impact assessment, Environ. Pollut., 179, 120–131, 2013.

von Bobrutzki, K., Braban, C. F., Famulari, D., Jones, S. K., Blackall, T., Smith, T. E. L., Blom, M., Coe, H., Gallagher, M., Ghalaieny, M., McGillen, M. R., Percival, C. J., Whitehead, J. D., Ellis, R., Murphy, J., Mohacsi, A., Pogany, A., Junninen, H., Rantanen, S., Sutton, M. A., and Nemitz, E.: Field inter-comparison of eleven atmospheric ammonia measurement techniques, Atmos. Meas. Tech., 3, 91–112, doi:10.5194/amt-3-91-2010, 2010.

Zhang, Y. and Wu, S.: Understanding of the Fate of Atmospheric Pollutants Using a Process Analysis Tool in a 3-D Regional Air Quality Model at a Fine Grid Scale, Atmospheric and Climate Sciences, 3, 18-30, doi:10.4236/acs.2013.31004, 2013.