

Dear ACPD Editor and Reviewers,

We thank the reviewers for their extremely valuable insights and salient key points, which attending to has improved and tightened the manuscript. We include the comments below (reviewer comments in bold), followed by our responses (regular black font) and text illustrating the changes in the manuscript (in italics and red). Line numbers in the reviewer comments are from the submitted manuscript, while the line numbers in the responses refer to the revised tracked changes document.

Reviewer 1:

- 1. This study evaluates climate-induced air-quality changes at regional scale across Europe and at urban scale for London in the 2090s under the RCP8.5 scenario, using a nested global–regional–urban modelling framework that links HadGEM2-ES, WRF–EMEP4UK, and ADMS-Urban. I think this will be a valuable read for anyone working on climate-AQ interactions in Europe, and I support publication in ACP.**

We thank the reviewer for their positive comments.

- 2. One suggestion, mainly to make the paper more interesting for readers outside Europe, is to more clearly state what the global-regional-urban setup adds compared to the more common global-regional approach. The authors mention this in a few places (e.g., around lines 675-679), but it is a bit spread out. My takeaway is that adding the urban model does not really change the main regional conclusions, but it does show much stronger spatial differences within the city. I think the paper would benefit from a more in-depth, focused paragraph that clearly says what changes (and what does not) when the urban model is included, and when this extra nesting is most worth doing in future studies.**

We agree and have added a new focussed paragraph to the conclusions text as recommended.

Line 848: *“The coupled modelling approach to enable nested regional and urban modelling of climate change is computationally intensive, but adds substantial value to conventional regional modelling approaches. For seasonal-average air pollutant concentrations over the UK, similar patterns and magnitudes of change are simulated using the 50 km× 50 km and 5 km × 5 km modelling domains, indicating that a regional modelling strategy may be sufficient for assessment of attainment and exceedances of long-term WHO air quality guidelines at the country or regional level. However, seasonal and diurnal cycle representation at the city scale is shown to be in closer agreement to observations when using the urban model compared to the 5 km × 5 km regional model for London, especially for the magnitude and spatial variation of surface NO₂ concentrations driven by sharp traffic-related gradients, as previously highlighted by Hood et al. (2018). Therefore, to evaluate long-term and short-term targets and guidelines at the city scale, and in particular for attainment of NO₂ AQGs, high resolution and explicit representation of its emission sources are crucial.”*

- 3. A related point, which the authors may choose to address or not, is how to place the results in the context of uncertainty in future climate projections (internal variability and differences across climate models). As the authors note, this kind of nested setup makes it hard to run large ensembles. In some cases, uncertainty from using one climate projection could be larger than the added benefit of going to finer scale, depending on the question and metric. I do not see this as a weakness of the paper. Urban-scale projections are rare, and even uncertain estimates are better than none, especially for policy use. Still, a brief**

note acknowledging this trade-off, and saying how the authors view the robustness of the main takeaways, could strengthen the paper.

We have performed additional calculations of uncertainty using data produced for the Zanis et al. (2022) study that used 3 (for surface PM_{2.5}) and 5 (for surface O₃) global earth system models as presented in The Royal Society (2021) report to quantify the uncertainty in surface O₃ and PM_{2.5} projections over the UK region (49-61°N, 10°W to 6°E) across the different models, which represents the structural uncertainty. We have discussed these results in Section 4.1 (UK distributions) and the Conclusions sections of the revised manuscript. However, as we acknowledge this is not a like-for-like comparison and hence this comparison can only be interpreted qualitatively.

Section 4.1 Added text at line 535: *“To provide context to our results, the impact of model structural uncertainty on O₃ and PM_{2.5} climate change projections over the UK region is also assessed, using results from three to five global-scale models from Zanis et al. (2022) as presented in RS (2021). In this study, the impact of climate change was evaluated for a baseline SSP3-7.0 anthropogenic emissions scenario for simulations with SSTs for present-day compared to SSTs for SSP3-7.0 for the 2090s. The summer mean surface O₃ response to climate change over the UK is -2.4 ppbv ±1.1 ppbv across the five models (Table A2). These multiple global-scale results suggest structural uncertainty in projected surface O₃ mixing ratios is greater than the uncertainty associated with spatial resolution identified using EMEP4UK e.g., that shows a summer O₃ response of -2.8 ppbv at 50 km and -3.2 ppbv at 5 km resolution (Table A2). Annual-mean surface PM_{2.5} concentrations across the three models reduces slightly but the standard deviation is substantial (-0.14 ± 0.18 µg m⁻³, Table A2). The structural uncertainty in annual-mean PM_{2.5} projections across the three global models used in Zanis et al. (2022) appears considerably larger than the uncertainty associated the spatial resolution of the EMEP4UK model over the UK (which exhibits an annual-mean PM_{2.5} response to climate change of 1.4 µg m⁻³ at 50 km resolution and 1.1 µg m⁻³ at 5 km resolution). However, the underlying emissions used in Zanis et al. (2002) are very different to this study, and the UK regional extent is somewhat larger (and includes the surrounding ocean) than used in this study, hence only a qualitative comparison is possible.*

Conclusions. Added text at line 861: *“Dynamical downscaling studies to achieve finer spatial representation of atmospheric composition change, as presented here, limit the use of multiple models with ensemble members that would be required for a comprehensive quantification of uncertainties (scenario, structural, internal climate variability) related to climate change. Indeed, although not a like-for-like comparison, uncertainties associated with model spatial resolution (for EMEP4UK) are somewhat smaller than model structural uncertainty derived from (three to five) different global models.”*

4. Finally, one small item: in line 330, the reference to “Fig. 1c” seems like it should be “Fig. S1c” (please check and fix if needed).

Line 394: Corrected; now “Fig. A2c”

Reviewer comment 2:

1. The manuscript “From Continental to Street Scales: Climate Change Impacts on Atmospheric Composition over Europe and London” by Doherty et al. presents an analysis of the effect of climate change on air pollution over Europe, UK and London under present-day emissions and the RCP8.5 climate scenario. In general, the manuscript fits very well into the scope of ACP. Before publication, however, minor revisions are necessary. These revisions should address the following concerns:

We thank the reviewer for their positive comments here and address specific concerns below.

- 2. The modelling approach has various simplifications, e.g., prescribing lightning NO_x, using the same ozone boundary conditions, and employing a dust climatology, etc. This information, however, is spread among various parts of the model description. Therefore, a table that clearly lists the processes altered in the future simulation and those that are not would be very helpful.**

This is an excellent point. In Section 2.4. Present-day and future model experiments. We have added *“Table 1: Model experiment set-up and simulations”* and accompanying text.

Line 261: “The model *experiment set-up, shown in Table 1,...*”.

- 3. The authors should also add a short note on how these simplifications will affect the results.**

We have added text to discuss the impacts of the use of prescribed present-day (PD) emissions where there is evidence from the literature. For a number of these climate-sensitive processes the impacts of climate change are uncertain. Our original manuscript included a note on the minor impact of using prescribed lightning NO_x emissions over Europe which we have now clarified.

The lack of biomass burning emissions is the greatest simplification in this study, and means that we underestimate baseline emissions. Using the Emissions of atmospheric Compounds and Compilation of Ancillary Data (ECCAD) database for the “TNO Europe” region for 2012, CAMS anthropogenic CO emissions are 53.6 Tg(CO)/yr and GFASv1.3 biomass burning emissions are 4.6 Tg(CO)/yr which suggests that we underestimate CO emissions by about ~8%. The corresponding values for NO_x emissions are 13.7 and 0.137 Tg (NO)/yr, indicating an underestimate of 1%. For primary PM_{2.5} emissions the regional CAMS-REG-ANT dataset for 2016 (not available for 2012) suggests 4.0 Tg over Europe compared to GFASv1.3 biomass burning emissions of 0.38, so a similar underestimate as for CO. Since Pan et al. (2020) suggest Europe to be a region with “least agreement” between biomass burning emissions datasets we have added a simplified statement of the above to the revised manuscript.

We have already outlined in the conclusions section that wildfires are likely to become more important emissions source in the future.

Text added to Section 2.2 Regional-scale modelling:

Line 182: *“Biomass burning emissions are not included here. The standard EMEP4UK model uses prescribed daily biomass burning emissions derived from satellite-based inventories. To ensure that interannual variability in the model experiments is due to climate alone, biomass burning emissions are turned off altogether. Biomass burning makes a relatively small contribution to total regional emissions over Europe (~1% for NO, 8-10% for CO and PM_{2.5}), but estimates vary substantially (see e.g., Pan et al. 2020) and there is considerable uncertainty in how they will respond to climate change.”*

Line 192: *“Lightning NO_x emissions are prescribed here. Interactive schemes used in other studies have shown strong responses to climate change globally, but relatively small responses over Europe (Finney et al. 2018), and hence the impact of this simplification is likely small. The impacts of climate change on DMS are uncertain (Thornhill et al. 2021; Zhao et al. 2024; Joge et al. 2025. The import of*

Saharan dust is treated using a monthly climatology of fine and coarse dust concentrations. Since the response of dust to climate change is uncertain (Sect. 1), Saharan dust is also treated as invariant between present-day and future”.

Regarding boundary conditions, in the conclusions section of the original manuscript we highlighted the impact on background O₃ if present-day methane concentrations are used, noting that following future trajectories of methane concentration increases in RCP8.5/SSP3-7.0 scenarios Turnock et al. (2022) found ~50% higher background O₃ levels. We added further text to provide a more comprehensive discussion of the sensitivity of our results to assumptions made.

Text added at line 840: *“Whilst the focus of the model simulations is to isolate the climate change response, the assumption of present-day levels for anthropogenic emissions (and the lack of wildfire emissions), as well as atmospheric methane and CO₂ concentrations and boundary conditions for O₃ and other species, means the chemical environment and the resulting atmospheric chemistry kinetics and atmospheric composition changes would be different if emission changes under RCP8.5 (or another scenario/pathway) were also employed.”*

- 4. (Related to this) I was wondering how boundary conditions of CO and other longer-lived species or species important for long range transport (e.g. PAN) are considered. Is the boundary condition for Europe simply zero?**

We have clarified this in the table above and added the following text to regional model description in Section 2.2, line 203: *“Boundary conditions for gas-phase (CO, PAN, NO_x, SO_x, HNO₃, H₂O₂, VOCs) and inorganic aerosol species are also prescribed climatologies based on measurements (Simpson et al. 2012).”*

- 5. The changes in soil NO_x and isoprene emissions with climate change in EMEP4UK seem to be very strong. The increase of PM_{2.5} and ozone is, of course, strongly driven by these increases. It would therefore be helpful to place these changes in the context of comparable modelling systems (or even results from global models). This will help the reader judge whether your results lie on the extreme end.**

This is an extremely valuable point and we thank for reviewer for highlighting this. There are relatively few studies in the literature that provide estimates of isoprene and NO emission changes for comparison with our study especially at the regional scale, which seems an omission. We have now added the isoprene and soil emission changes as a new Fig. A1 and Table A1 in the Appendix and we have added detailed text to compare our results to the studies that do provide changes in isoprene and soil NO emissions alongside temperature changes. We find the areal extent the temperature change is calculated over is an important consideration and the literature often does not specify this clearly.

Figure A1: *“Figure A1: Annual natural emissions for present-day and future of (a, b) isoprene and (c,d) soil NO and (e, f) differences between future and present-day due to climate change.”*

Table A1: *“European domain average temperature, isoprene (C₅H₈) and Soil NO emissions for present day (1996-2005) and future (2090-99) along with the % change and % change per 1°C.”*

Section 3.1 Annual and Seasonal mean changes, text on isoprene emissions added at line 324: *“Whilst anthropogenic emissions remain unaltered, biogenic isoprene emissions respond to elevated temperatures increasing from 10.4 to 22.5 mg m⁻² (116 %) associated with a 4.8°C increase over the European domain (35-70°N and 20°W-40°E), corresponding to 24 % increase per 1°C (Table A1). This doubling of annual-mean isoprene emissions (Fig. A1c), results in isoprene mixing ratios significantly*

elevated by up to 4 ppb in summer in parts of southern Europe, with much smaller changes in winter (Fig. 2c, h). These elevated isoprene levels are the main driver of higher O_3 and $PM_{2.5}$ in summer across continental Europe. Andersson and Engardt (2010) found increases in isoprene emissions of 83% over Europe in the 21st Century under the SRES A2 scenario, smaller than the changes reported here, but for a smaller temperature increase. Langner et al. (2012) showed that under the SRESA1B scenario isoprene emissions over Europe increased by 21-26 % in four out of five regional model simulations over the first four decades of the 21st Century associated with a 1.27°C temperature increase. The sensitivity of the isoprene response to climate change over Europe was comprehensively examined using the MEGAN-MOHYCAN model by Bauwens et al. (2018). Under RCP8.5 they suggest isoprene emission increases of 83% over the 21st Century for an average increase in temperature of 4°C (21% per 1°C). Larger increases were found when CO_2 fertilisation effects were included, but there were substantial decreases when considering CO_2 inhibition. Overall, the changes in isoprene emissions reported in our study are large but consistent with other studies that do not include CO_2 inhibition or fertilisation effects. Large uncertainties remain in the interplay between these complex effects on isoprene emissions (Do et al. 2025)."

We have also strengthened the conclusions text.

Line 835: "This study finds that robust projections of the magnitude of the impact of climate change on surface O_3 and $PM_{2.5}$ for Europe crucially rely on accurate representation of climate-sensitive biogenic emissions, which remain highly uncertain between studies."

Line 838: "Studies that assess climate change impacts on natural emissions focus on overall changes in air pollutant concentrations; few provide quantitative estimates of natural emission changes to compare with this study."

Section 3.1 Annual and Seasonal mean changes, text on soil NO emissions added at line 362:

"Annual-mean soil NO emissions increase under this large climate signal (Fig. A1f) across almost all of Europe yielding a 64 % increase in the future compared to present-day or a 13 % increase per 1°C (Table A1), potentially influencing land summer NO_x concentrations (Fig. 2i). Previous experimental studies have suggested a 100% (or doubling) of NO emissions for each 10°C rise, although sub-ranges of temperatures showed differing levels of linearity (Laville et al. 2009); with the parametrisation used in this study based on such results (Simpson et al. 2012). Few studies have reported soil NO emissions changes over Europe. A 9% increase in NO emissions averaged over Europe was simulated under a regional warming of 1.8°C on average across Europe in the 2030s compared to the 1990s (a 5.2 % increase per 1°C) by Kesik et al. (2006). The lower sensitivity to climate change compared to this study, is likely due to the inclusion of soil moisture effects on NO emissions. A recent experimental field study found that dryer soils in a warmer climate could reduce NO emissions (Huang et al. 2025), suggesting there is also uncertainty in the impact of this climate-sensitive emission process. Unlike isoprene, this natural soil emission source is minor compared to current anthropogenic NO_x emissions as noted by Simpson et al. (2012)".

6. While I like the idea of section 5, my major concern is that the authors present no evaluation with actual station measurements. The results of the threshold analyses strongly depend on whether the model has any positive or negative biases compared to observations in PD conditions. Please add such evaluation.

This is an excellent point. However, we were unable to perform a like-for-like model evaluation given the difference between the year of the emissions (2012) and the 10-year present-day meteorological period 1996-2005 based on HADGEM2-ES simulations using historical SSTs. Given the importance of this point, we have conducted a model evaluation for 2012, on the assumption that emissions are the primary driver of long-term concentrations responses. We perform this evaluation for model

configurations used to produce WHO guideline values in section 5: EMEP4UK at 50 km × 50 km over Europe and ADMS-Urban over London. We further note that observations are somewhat limited for 2012, and extremely sparse for representative background sites for the 1996-2005 period over Europe. Hence, we did not perform any further model evaluation over Europe for 1996-2005. In our evaluation for London, we have limited data to compare to observations over the period 1996-2005 and hence have also evaluated ADMS-Urban results for this period.

Our main finding is that for the most part (most locations and most target values) there are sufficient intervals between interim targets and air quality guideline values that our findings are robust, and where there is not the case, we describe the implications. We not only perform this evaluation for in section 5, but for consistency also for annual-average surface O₃ and PM_{2.5} concentrations over Europe (Fig. 1) in Section 3.1 and for seasonal and diurnal cycles of surface O₃, PM_{2.5} and NO₂ concentrations over London (Figs. 8 and 9) in sections 4.2 and 4.3, as well as throughout section 5. Specific additions to the text are given below including a description of the observational data and their model evaluation results.

Section 2.2 Regional-scale modelling, text added at line 207: *“For O₃ and NO₂ model evaluation (performed using the R openair package), hourly observational data over Europe for 2012 from “rural” sites were obtained from EMEP (<https://www.eea.europa.eu/data-and-maps/data/airbase-the-european-air-quality-database-8>), with 365 and 265 sites respectively meeting the 75% hourly data capture criteria. Slightly fewer sites met additional data capture criteria within the peak season required for calculating MDA8 O₃ (332). Observations from urban and suburban background sites were also included in the comparisons for PM_{2.5}, giving a total of 170 sites, due to the small number of rural PM_{2.5} monitoring sites available in 2012 (38). Model-observation comparisons are performed by extracting values from the closest EMEP4UK grid cell to each monitoring location.”*

Section 2.3 Coupled Regional and Urban-scale modelling, text added at line 242: *“Model evaluation (performed using the R openair package), used hourly observational data for London for the years 1996-2005 and 2012 from the LAQN for “background” and “near road” sites that met the requirement of at least 70 % data capture of hourly data during the relevant year. For 2012, 20, 11 and 42 sites for O₃, PM_{2.5} and NO₂ met this requirement (see Hood et al. 2018). For the 1996 to 2005 period, considerably fewer data were available: for O₃ (NO₂) the number of sites increased from 7 (9) in 1996 to 16 (40) in 2005. For PM_{2.5} for this period only “roadside” data were available (1 site from 1998 and 2 sites from 2004), hence the measurements do not reflect ambient conditions well; also, observations exhibit considerable year-to-year variation (between 28.9 and 56.0 µg m⁻³), which may reflect equipment error. Model-observation comparisons are made at directly corresponding locations due to the high spatial resolution of ADMS-Urban model output. The coupled system and the standard ADMS-Urban model configuration for 2012 is extensively evaluated in Hood et al. (2018).”*

Section 3.1 Annual and Seasonal mean changes. Additional text on model evaluation results:

Fig. 1 caption, *“Panels b) and c) include summary model-observation comparison statistics at rural observation sites (O₃) and background sites (PM_{2.5}) for year 2012. N: number of sites included in comparison; MB: mean bias; NMB: normalised mean bias; RMSE: root mean square error; r: correlation coefficient.”*

Line 292: *“These annual average modelled O₃ concentrations are compared to observations for rural sites for the year 2012, assuming that the dominant influence on long-term air pollutant concentrations arises from underlying anthropogenic emissions. The observed annual-average O₃ values are generally well captured by the model although slightly underestimated, with small*

magnitudes of mean bias (-1.8 ppbv), normalised mean bias (-6%) and RMSE=6 (Fig. 1). Some of the largest underestimates for individual sites occur when monitors are located at high elevation (> 1500m; Figure A1) e.g., in the Alps/Balkans. The spatial correlation coefficient is moderate $r=0.56$ (Fig. 1b); this may partly reflect differences in meteorological impacts on concentrations that arise through a comparison of observations from the year 2012 against the 1996-2005 modelled period average."

Line 304: "Annual average $PM_{2.5}$ concentrations are overestimated by the model in 2012, with a mean bias of $3 \mu g m^{-3}$ (NMB = 21%) and an RMSE =7, and a moderate spatial correlation ($r=0.63$)."

Section 4.2 Regional- and Urban-scale Seasonal cycles for London. Additional text on model evaluation results:

New Figure A5: "Seasonal cycles for the year 2012 across London from "background" and "near-road" site observations, EMEP4UK and ADMS-Urban for (a) surface O_3 , mixing ratios (ppbv) (b) $PM_{2.5}$ concentrations ($\mu g m^{-3}$) and (c) NO_2 mixing ratios (ppbv) respectively. For O_3 $n=20$ sites, for $PM_{2.5}$ $n=11$ sites, for NO_2 $n=42$ sites. Detailed site information can be found in Hood et al. (2018)."

New Table A4: "Median monthly values for "background" and "near-road" London sites for the year 2012 for surface O_3 mixing ratios (ppbv), $PM_{2.5}$ concentrations ($\mu g m^{-3}$) and NO_2 mixing ratios (ppbv) for the year 2012 from a) observations, b) EMEP4UK 5×5 km and c) from ADMS-Urban as used in Hood et al. (2018) that employed a similar coupling approach to this study. Month 1= January etc. For O_3 $n=20$ sites; for $PM_{2.5}$ $n=11$ sites; for NO_2 $n=42$ sites. Detailed site information can be found in Hood et al. (2018)."

Line 564: "To assess model performance of simulated seasonal cycles relative to observations; monthly O_3 , $PM_{2.5}$ and NO_2 distributions were calculated using data from the Hood et al. (2018) study, which employed the same set-up as used here, enabling a direct comparison for the year 2012. The amplitudes of the median values of surface O_3 seasonal cycles for both models are also ~ 20 ppb for 2012 which is an overestimate compared to observations of ~ 16 ppbv arising from a smaller spring/summer peak (Fig. A5, Table A4). Differences in magnitudes between the two models are also small (1-3 ppbv) for 2012 (Table A4)."

Line 578: "For the year 2012, observations show the wintertime peak in $PM_{2.5}$ concentrations occurs later in February/March and this timing is captured by the two models, but the maxima are substantially underestimated (by $13 \mu g m^{-3}$; Fig. A5; Table A4), which Hood et al. (2018) attribute to underestimated regional contributions. There are fewer sites available for evaluation of $PM_{2.5}$ concentrations and over 50% of these are near road sites which may be more challenging to simulate than background and rural locations."

Line 591: "For 2012, seasonal cycle peak and amplitude values for NO_2 are similar to those for the present-day period; median values are also higher for ADMS-Urban and in good agreement with observations (Fig. A4; Table A4)."

Line 598: "This is also the case for the year 2012, where the spatial variability simulated using the ADMS-Urban model is similar to (for NO_2) or smaller (for O_3 and $PM_{2.5}$) than for the observations."

Line 604: "This finding adds confidence to the conclusion that the street-scale simulation exhibits larger spatial variability in air pollutant concentrations, that agrees well with the observed spatial variability for 2012..."

Section 4.3 Urban-scale Diurnal cycles for London. Additional text on model evaluation results:

New Figure A6: *“Diurnal cycles for winter and summer for the year 2012 across London from “background” and “near-road” site observations, EMEP4UK and ADMS-Urban for (a, b) surface O₃, mixing ratios (ppbv) (c,d) PM_{2.5} concentrations (μg m⁻³) and (e,f) NO₂ mixing ratios (ppbv) respectively. For O₃ n=20 sites; for PM_{2.5} n=11 sites; for NO₂ n=42 sites. Detailed site information can be found in Hood et al. (2018).”*

Line 615: *“The close agreement between the simulated ADMS-Urban and observed diurnal cycles for 2012, that capture the key features described above, is evident in Fig. A6. The greater ability of the ADMS-Urban urban model to capture O₃ diurnal cycles for London as compared to the regional EMEP4UK model has been highlighted by Hood et al. (2018).”*

Line 632: *“The underestimate in simulated PM_{2.5} concentrations compared to observations, as noted in Sect. 4.2, is apparent in these diurnal cycles (Fig. A6).”*

Line 643: *“Observed and modelled NO₂ diurnal cycles for 2012 agree well (Fig. A6).”*

Section 5: Implications for achieving WHO guidelines. Additional text on model evaluation results:

Line 665: *“To assess the sensitivity of the results, long-term and short-term metrics are also calculated from available observations for 2012 (and for 1996-2005 for short-term metrics for London).”*

Figure 10 caption: *“Panels a), b) and c) include summary model-observation comparison statistics at rural observation site locations (O₃, NO₂) and background sites (PM_{2.5}) for year 2012. Note the PM_{2.5} statistics are the same as in Fig 1. N: number of sites included in comparison; MB: mean bias; NMB: normalised mean bias; RMSE: root mean square error; r: correlation coefficient.”*

Line 680: *“The model-observation comparison shows peak season MDA8 O₃ across 332 rural sites is slightly overestimated (MB=4.2 ppbv; NMB=5%; RMSE=17.8; Fig. 10a), although the MB is less than the difference between successive AQG targets. This MDA8 O₃ overestimate is most notable in the small area of southern Europe that has the highest simulated MDA8 O₃ values, suggesting that here the first interim target may have been met in the year 2012. However, MDA8 O₃ is underestimated over the UK and northern Scandinavia, suggesting that fewer areas meet the second interim target in the year 2012.”*

Line 696: *“The summary evaluation results for annual-average PM_{2.5} concentrations (Fig. 10b as in Sect. 3.1) outline a MB of 3 μg m⁻³ over 170 rural and background sites, which is also smaller than the difference between successive AQG targets. Overestimates in PM_{2.5} concentrations are most prominent for the Iberian Peninsula, suggesting greater attainment of the third and lower interim targets at these locations in 2012; whilst in parts of central Europe and the UK PM_{2.5} underestimates may reflect lesser regional attainment of interim targets/AQGs in these regions compared to observations for 2012.”*

Line 709: *“Annual average simulated NO₂ is overestimated compared to observations at 265 rural sites (MB=7.4 μg m⁻³; NMB=68%; RMSE=9.8; Fig. 10c). This may reflect influences of urban NO_x emissions being included in the same model grid cells as the rural observation sites. The spatial correlation coefficient (r=0.75) is higher than for the other air pollutants which may reflect a stronger influence of anthropogenic emissions on the spatial distribution of NO₂ given its shorter lifetime. However, the MB is less than the difference between successive AQG targets for NO₂. The*

observations for 2012 suggest that NO₂ concentrations at rural sites within these hotspot areas are overestimated. More locations in central Europe but fewer locations in western Europe meet the NO₂ AQG.”

Line 723: “To assess the robustness of these results for the present-day period, short-term metric values are also calculated from observations for 2012 and for the period 1996-2005 for the sites with available observations (which are fewer than the 56 sites used for comparison with the present-day ADMS-Urban simulations).”

Table A5: “Short-term 99th percentile values of MDA8 O₃, and 24-hour mean PM_{2.5} and NO₂ calculated annually from hourly data for present-day (1996-2005) as in Table 2, for 2012 from observations (Obs.) and ADMS-Urban results (Mod.), and for 1996-2005 from observations and ADMS-Urban results for sites/receptor locations where observations are available. Hence 1996-2005 model results will differ from the PD results because of differences in the number of sites are available for different years (for O₃ n=up to 16, for PM_{2.5} n=0-2, for NO₂ n= up to 40). Mean values across all locations are given in the 2nd column. The WHO short-term averaging period interim target and air quality guideline (AQG) values (WHO 2021) are shown in bold. Exceedance days per year are calculated for present-day and for 2012 for the WHO interim target values and the relevant AQG value (rightmost column). The spatial variation in exceedance days across the locations are represented by the standard deviation.

Line 729: “There is close agreement in observational and model-derived results of the 99th percentile MDA8 O₃ values in 2012 that show the first and second interim targets are met, while the AQG is exceeded on 7-8 days (Table A5). For 1996-2005, the first and second interim targets are met when using the observations, but the second interim target is not met using ADMS-Urban model results. Although the magnitude of the short term MDA8 O₃ metric is highly sensitive to the underlying emissions and meteorology, the overall result that the AQG is not met for the present-day day period is consistent between datasets.”

Line 745: “The ADMS-Urban results underestimate 99th percentile 24-mean PM_{2.5} concentrations compared to observations in 2012 (as noted in Sects. 4.2 and 4.3), but overestimate this PM_{2.5} metric over 1996-2005. Observational-based estimates of the 99th percentile of 24-hour mean PM_{2.5} concentrations for 2012 suggest the first but not the second interim target is met; whilst for 1996-2005 the first and second interim targets are met (Table A5). When utilising ADMS-Urban results, the third interim target is narrowly met for 2012, and for 1996-2005 the first and second interim targets are met. The number of exceedance days of the PM_{2.5} short-term AQG in 2012 is 115 days based on observations and 59 days when using ADMS-Urban results. As noted in Sect. 4.2, PM_{2.5} observations are limited over the present-day time period considered, with only 11 sites available for 2012 and only 1-2 sites available for 1996-2005. The short-term PM_{2.5} AQG is not met whichever dataset is used.”

Line 757: “These results are also found using observations for 2012, but when using observations over the 1996-2005 the first interim target for the 99th percentile values NO₂ is met (Table A5). In all cases the number of exceedance days of the AQG is similar at between 300-310 days.”

Line 761: “Hence, for the three air pollutants, the short-term air quality guideline values are not met for present-day or in the future when utilising 2012 anthropogenic emissions in these simulations. Consistent AQG exceedance results are found when using observations for 2012.”

Conclusions:

Line 773: *“Simulated surface O₃ mixing ratios are slightly underestimated (MB=-1.8 ppbv) and surface PM_{2.5} concentrations are somewhat overestimated (MB=3 µg m⁻³).”*

Line 806: *“; with PM_{2.5} concentrations underestimated compared to observations for 2012....For all air pollutants, the urban model simulates substantially greater spatial variability..., in good agreement with observations for 2012.”*

Line 813: *“Monthly mean NO₂ magnitudes are ~10% higher at the street scale using ADMS-Urban compared to the regional EMEP4UK model, with the ADMS-Urban model results capturing key features of observed seasonal and diurnal NO₂ cycles.”*

Line 818: *“except in parts of northern Europe in the present day (with observations for 2012 showing less attainment for this region than EMEP4UK model simulations).”*

Line 824: *“Very limited areas of northern Europe meet the stringent WHO guidelines for annual mean PM_{2.5} of 5 µg m⁻³ for present-day (consistent with observations for 2012 in this region).”*

Line 828: *“...NO₂...(and for observations in 2012)”*

7. -p6l193 I don't fully understand the statement. Are biomass burning emissions included or not? And if not – why?

We have revised the text for clarity at line 182: *“Biomass burning emissions are not included. The standard EMEP4UK model uses prescribed daily biomass burning emissions derived from satellite-based inventories. To ensure that interannual variability in the model experiments is due to climate alone, biomass burning emissions are turned off altogether. Biomass burning makes a relatively small contribution to total regional emissions over Europe (~1% for NO, 8-10% for CO and PM_{2.5}), but estimates vary substantially (see e.g., Pan et al. 2020) and there is considerable uncertainty in how they will respond to climate change.”*