Dear Dr Duncan,

We thank both reviewers for their insightful comments that have aided the improvement of this manuscript considerably. Our responses to the reviewers' comments together with a description of the changes made to the manuscript can be found below. Changes not mentioned in this document are purely editorial. For clarity, the reviewers' comments are copied below in bold, followed by our responses, and in quotes and italics, modifications to the manuscript. In our revised manuscript, the modified text is shown using track changes.

Referee #1

This study compares health impacts estimated for ozone and $PM_{2.5}$ simulated at global versus regional chemical transport model resolutions, and analyzes the factors contributing to resulting differences in the health estimates. While several other studies have conducted similar analyses for air pollution health impacts in the US, at a range of spatial scales from 4km to _250 km, it hasn't been done for Europe. There is reason to believe that results from the US wouldn't directly apply to Europe due to differences in emissions magnitudes of pollution components and chemical processing in the atmosphere. Thus, while this paper is a relatively straightforward corollary to the US studies, it is interesting and a useful contribution to the literature. It also presents some interesting new results on seasonality of and factors contributing to the resolution effect.

We thank the reviewer for their positive comments.

Comments:

In reality, differences in disease rates in urban centers versus broader areas would also come into play in addition to model spatial resolution and population/pollution colocation. Please discuss this limitation in your approach and how its omission should affect interpretation of your results.

We agree with the reviewer that disease rates in urban centres versus broader areas could vary and thus may need to be taken into consideration when studying regional to urban-scale effects of air pollution on

morality. However we do not have this information. More importantly, in our study we did not examine changes in absolute mortality attributable to long-term exposure to O_3 and $PM_{2.5}$ but we examined changes in the Attributable Fraction (AF) of all-cause mortality to isolate the impact of changing the resolution on pollutant concentrations and the associated health impacts, from changes in baseline mortality rates across different countries. AF represents the fraction or percentage of the all-cause mortality which is attributable to the effects of O_3 and $PM_{2.5}$, and depends only on population weighted pollutant concentration and an appropriate concentration-response coefficient which is typically applied at a country or continental-scale level (e.g. see WHO 2013).

The paper made me wonder why so many papers have been written on spatial resolution issues, but not as much attention has been given to vertical resolution. Is the use of the first model layer (which is noted as 40m high) as ground concentrations adequate for capturing concentrations at the "nose level"? Should this be explored? Please provide guidance on this issue.

We thank the reviewer for this comment and acknowledge that using concentrations at lowest model level (with a height of 40 m) is a limitation of our study. The lowest model level is widely used as representative of surface concentrations in modelling studies and simulated concentrations are evaluated against measurements, but some studies e.g. Fiore et al. (2009) note uncertainties pertaining to vertical resolution in coarse global-scale models. Similarly, the lowest model layer is used when calculating health impact assessments (e.g. Punger and West, 2013). For pollutants with an extremely short lifetime such as NO_2 vertical resolution could be a very important issue but less so for longer lived O_3 and $PM_{2.5}$ investigated in this study. We have evaluated the model output using observations at ground level and found performance to be satisfactory (Sections 3.1 and 3.3) We have added the following text to the methods section 2.1 and have added a note to the conclusions in Section 5:

Page 5, line 148: "*All pollutant concentrations used in this study have been extracted at the lowest model level with a mid-point at 20 m. While this level is considered representative of surface or ground- level concentrations, local orographically driven flows or sharp gradients in mixing depths cannot be represented at this vertical resolution (Fiore et al. 2009)*."

Page 19, line 593: "*The pollutant concentrations used in this study have been extracted at the lowest model level with a mid-point at 20 m. The sensitivity of our simulated pollutant concentrations to vertical model resolution has not been examined*."

Line 160-167: The original PM2.5 epi study should be cited here, rather than simply referring to the HRAPIE recommendations. Please check whether all the health impact assessments referenced (Anenberg 2009, Punger and West 2013, and Thompson et al. 2014) actually used HRAPIE recommended effect estimates since some of these were published before HRAPIE.

We thank the reviewer for this comment. We have modified the text to clarify the sources of the concentrations response coefficients related to long-term exposure to O_3 and $PM_{2.5}$ used in our study and in the other studies cited (Section 2.3):

Page 6 line 173: "Although there is limited evidence available for the long-term health impacts of O_3 especially in Europe (The UK Committee on the Medical Effects of Air Pollution (COMEAP) 2015), *a number of studies have quantified the adverse health impacts associated with long-term exposure to* O_3. *In this study we apply the Health Risks of Air Pollution in Europe – HRAPIE project recommended coefficient for long-term exposure to* O_3 (WHO, 2013) *to investigate the sensitivity of health calculations to the model resolution used to simulate* O_3 *concentrations. This concentration–response coefficient is derived from the single-pollutant analysis of the American Cancer Society Cancer Prevention Study II (CPS II) cohort study data in 96 metropolitan areas of the US* (Jerrett et al., 2009) *which has been used by previous studies (e.g*. Anenberg et al., 2009; Punger and West, 2013; Thompson et al., 2014; Cohen et al., 2017); *but is re-scaled from 1-hour mean to 8-hour mean concentrations using the ratio 0.72, derived from the APHEA-2 project* (Gryparis et al., 2004). *The value recommended by HRAPIE for the concentration-response coefficient, or β value (Eq.1), for the effects of long-term* O_3 *exposure on respiratory mortality recommended is 1.014 (95% Confidence Interval (CI) = 1.005, 1.024) per 10 µg* m^{-3} *increase in MDA8* O_3 *during the warm season (April-September) with a threshold of 70 µg* m^{-3} (WHO, 2013). *For estimating the health impact of long-term exposure to* $PM_{2.5}$ *on all-cause (excluding external) mortality, HRAPIE (WHO 2013) recommends a relative risk coefficient of 1.062 (95% CI = 1.040, 1.083)*

per 10 µg m^{-3} increase in annual average concentrations (with no threshold) which is based on a meta-analysis of cohort studies by Hoek et al. (2013)."

Many health impact assessments are now employing non-linear concentration-response curves which flatten out considerably at higher concentrations, particularly for cardiovascular diseases. Please comment on how using such non-linear concentration-response functions would influence your results (e.g if the higher spatial resolution leads to higher PM concentrations, would those concentrations then fall on the flatter end of the CRF, leading to lower health impact estimates?)

To estimate the global burden of disease attributable to ambient fine particulate matter exposure, some recent studies have derived integrated concentration-response functions that come from integrating available relative risk information from various studies of ambient air pollution, second hand tobacco smoke, household solid cooking fuel and active smoking (e.g. Burnett et al., 2014). These functions are applied to cause-specific mortality associated with long-term exposure to $PM_{2.5}$, however in this study we have focused on all-cause $PM_{2.5}$-related mortality. We agree that by using the 'integrated' concentration-response function, the concentration response curves flatten out at high concentrations based on evidence from epidemiological studies. However for ambient air pollution the curve is log-linear (e.g. Fig. 1 and 2 Burnett et al., 2014). In addition the curve flattens out for annual $PM_{2.5}$ concentrations above approximately 100 µg m^{-3} (Burnett et al., 2014). Such high annual ambient $PM_{2.5}$ concentrations are common in cities across Asia and other developing regions (Brauer et al., 2012; Health Effects Institute, 2010). However, in our study across the whole European domain, the maximum annual mean $PM_{2.5}$ concentrations are 40 µg m^{-3} and 49 µg m^{-3} for the global and regional configuration, respectively. Given the magnitude of the concentrations in this Europe focused study we feel that applying a log-linear relationship is appropriate.

Same comment as above, but for low-concentration thresholds. We don't know whether PM health effects go down to zero, though some epi studies are showing relationships to very clean levels (2-5 ug/m3). It's useful to the reader to provide some guidance on how your results would be different

if you did apply a low-concentration threshold for PM2.5, perhaps set at the theoretical minimum risk level used in the GBD studies.

We thank the reviewer for this comment, and we have now investigated the impact of a low concentration threshold. We apply a threshold of 5.8 μg m^{-3} (following the minimum that is suggested by Burnett et al. (2014) which is derived from Lim et al. (2012)). We find differences in AF estimates associated with long-term exposure to population-weighted $PM_{2.5}$ range from -4.8% to +2.1% compared to -4.7% to +2.8% when no threshold is applied. The spatial distribution of these estimates remains unchanged with a large number of countries in Eastern Europe and the UK showing positive differences in AF between the global and regional resolutions and only slight changes in country rankings (see Fig. R1 below compared to Fig.8b in manuscript). Hence, in our study we find the effect of applying a low-concentration threshold for $PM_{2.5}$ to be small. We have added the following text to the manuscript to discuss these results in section 4.4 and added a sentence to the conclusions in Section 5.

Page 16, line 502: *"We also examine the impact of using a low-concentration threshold. We apply a threshold of 5.8 μg m-3 (suggested by Burnett et al. (2014) which is derived from Lim et al. (2012)) to annual mean $PM_{2.5}$ concentrations. Differences in AF estimates associated with long-term exposure to population-weighted $PM_{2.5}$ concentrations range from -4.8% to +2.1% (as compared to -4.7% to +2.8% above when no threshold is applied). The spatial distribution of these estimates remains unchanged and only slight changes in country rankings occur. Hence, the impact of applying a low concentrations threshold in this study for Europe is small."*

Page 19 line 583: *"In addition, these ranges in AF associated with long-term exposure to annual mean $PM_{2.5}$ were largely unaltered with the application of a low-concentration threshold for $PM_{2.5}$."*

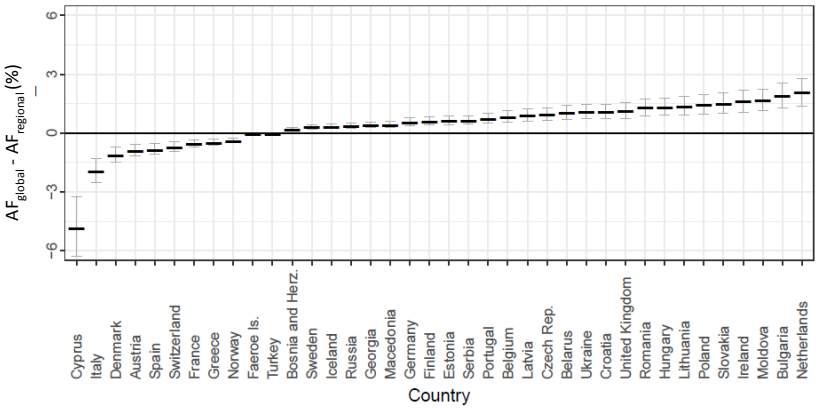


Figure R1: Differences in AF associated with long-term exposure to annual mean $PM_{2.5}$ between the two resolutions expressed as a percentage for each European country ($AF_{global} - AF_{regional}$) using a threshold of 5.8 µg m^{-3}. Grey lines show the 95 % C.I. which represents uncertainties associated only with the concentration-response coefficient used.

Also, the most recent American Cancer Society study update gives ozone-mortality relationships for annual average concentrations (Turner et al. Long-term ozone exposure and mortality in a large prospective study, American Journal of Respiratory and Critical Care Medicine, 193, 10, 1142, 2015). These relationships were used by Malley et al. to updated the ozone burden of disease values (Malley et al. Updated global estimates of respiratory mortality in adults >30 years of age attributable to long-term ozone exposure, Environmental Health Perspectives, 087021-1, 2017). Please comment on how your results would be different if you were to use these annual average ozone effect estimates, given the seasonality of the resolution effect on simulated concentrations.

We thank the reviewer for this interesting comment and we have investigated this effect. When using the concentration-response function (CRF) based on epidemiological studies, the time averaging period used for pollutant concentrations should match that used to quantify the CRF. For this reason we do not use

annual mean MDA8 O_3 concentrations in conjunction with the CRF used in our study (based on HRAPIE in turn based on Jerrett et al. 2009) as this was derived from warm season concentrations. Thus, in addition, we have estimated the differences in AF between the two resolutions following Turner et al. (2015) whereby we use: a) an annual-mean MDA8 O_3 concentration (instead of summer mean concentrations), b) a concentration response function of 1.06 (95% CI: 1.04-1.08) per 10 µg m^{-3} (instead of 1.014 (95% CI: 1.005-1.024) per 10 µg m^{-3} and c) a threshold of 53.4 µg m^{-3} (instead of 70 µg m^{-3}), with the values in parenthesis being those used to date in our study. Results are shown in Fig. R2 below (compare to Fig. 7 in manuscript).

Differences in AF between the two resolutions using annual-mean O_3 concentrations and CRF/threshold values from Turner et al. (2015) range from -2.3% to +12.0% across the countries compared to -0.9% to +2.6% when a summer mean MDA8 O_3 concentration with the CRF from the WHO HRAPIE project is used. The CRF quoted by Turner et al. (2015), applicable to annual-mean O_3, is approximately 4 times higher than the CRF used in our study which is derived from summer time MDA8 O_3 exposure. This is the main driver for a larger range in differences in AF between the resolutions when using the recommendations in Turner et al. (2015) for annual MDA8 O_3 concentrations.

In contrast, Turner et al. (2015) found similar results for O_3-mortality relationships for all-cause mortality, diseases of the circulatory system and cerebrovascular diseases when using summer and annual-mean O_3 concentrations. However results were attenuated when using a summer O_3 concentration for mortality due to dysrhythmias, heart failure, and cardiac arrest, diabetes and respiratory causes. Although as discussed previously we do not suggest applying the HRAPIE coefficient to annual-average concentrations, we have done this calculation as a sensitivity test (Fig. R3). When using the HRAPIE suggested coefficient derived from Jerrett et al. (2009) with annual MDA8 O_3 concentrations, differences in AF range from -0.5% to +3.7%. The range is slightly larger compared to the summer mean estimates as differences in annual mean MDA8 O_3 concentrations between the two resolutions are larger due to the seasonality noted in the Section 3 of the manuscript and as mentioned in the conclusion section.

We have added these results to our manuscript through the following text in section 4.3 and section 5 conclusions:

Page 15, line 474: *"Since, seasonal differences in simulated O_3 with resolution are considerable, the AF associated with long-term exposure to O_3 was also calculated based on annual-mean (as opposed to summer-mean) O_3 concentrations based on recommendations by Turner et al. (2015). Turner et al (2015) suggest a higher concentration response coefficient of 1.06 (95% CI: 1.04-1.08) per 10 µg m^{-3} and a slight lower MDA8 O_3 threshold of 53.4 µg m^{-3} compared to values used in our study for summer-mean MDA8 O_3. Using the values from Turner et al. (2015) the differences in AF are found to be of the same sign for the majority of the countries and the rankings across countries are largely similar. This similarity occurs because the difference in annual-mean MDA8 O_3 concentrations between the two resolutions shows generally similar spatial patterns to the differences in warm season MDA8 O_3 concentrations (not shown). However the ranges when using annual-mean O_3 concentrations and recommendations form Turner et al. (1015) are larger: -2.3% to +12.0%, compared to AF ranges given above for MDA8 O_3. From further sensitivity analyses it is found that these greater AF ranges can be attributed to the use of a higher concentration-response coefficient (by a factor of approximately 4) rather than differences in annual-mean compared to summer-mean concentrations.*

Page 19 line 586: *"When using annual-mean MDA8 O_3 concentrations alongside a recommended concentration-response coefficient and threshold suggested by Turner et al. (2015) the difference in AF between the two resolutions is considerably larger than our estimates using summer-mean MDA8 O_3 concentrations. This is driven by the higher concentration-response coefficient (by a factor of approximately 4) quoted in Turner et al. (2015) compared to that suggested by HRAPIE for summer mean MDA8 O_3 concentrations (WHO, 2013)."*

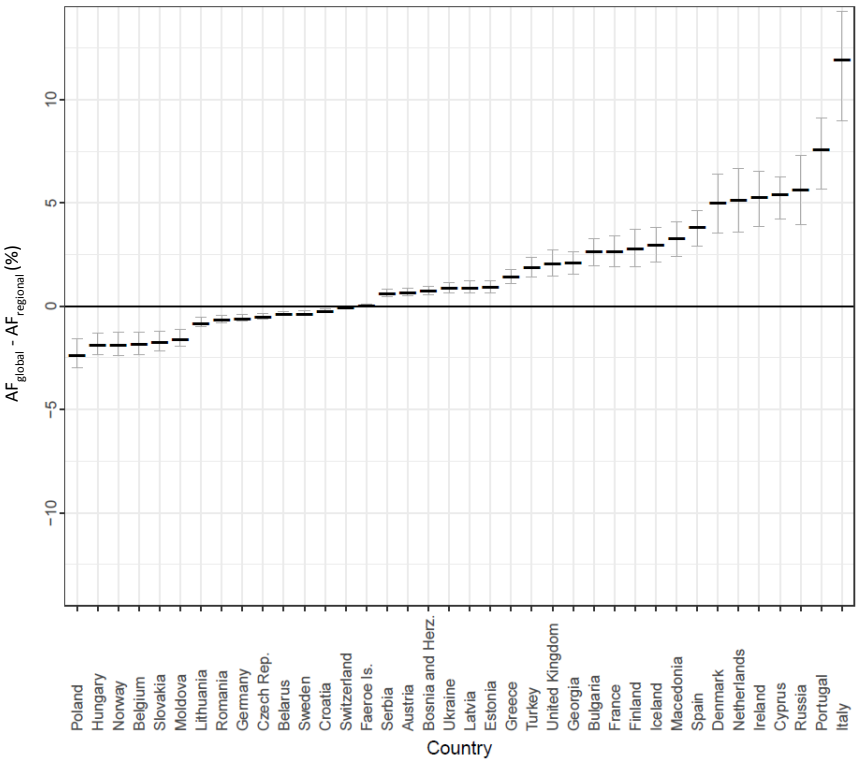


Figure R2: Differences in AF associated with long-term exposure to annual mean MDA8 O_3 between the two resolutions expressed as a percentage for each European country ($AF_{global} - AF_{regional}$) using a threshold of 53.4 µg m^{-3}. Grey lines show the 95 % C.I. which represents uncertainties associated only with the concentration-response coefficient quoted in Turner et al. (2015).

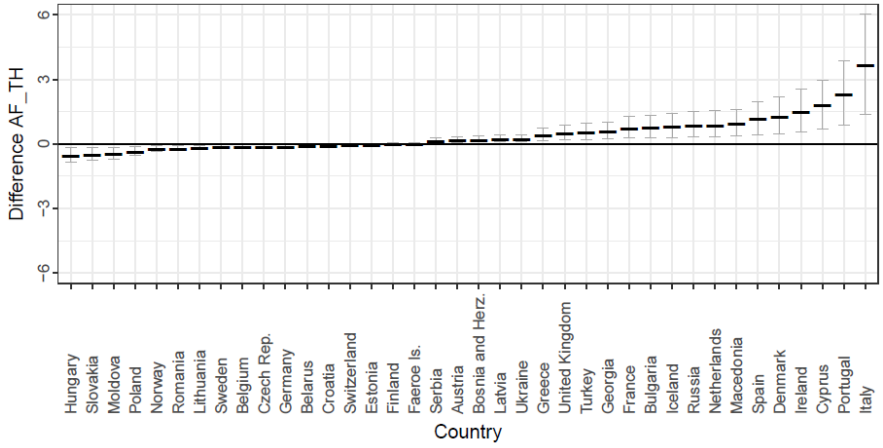


Figure R3: Differences in AF associated with long-term exposure to annual mean MDA8 O_3 between the two resolutions expressed as a percentage for each European country ($AF_{global} - AF_{regional}$) using a threshold of 70 µg m^{-3}. Grey lines show the 95 % C.I. which represents uncertainties associated only with the concentration-response coefficient quoted in HRAPIE which is based on Jerrett et al. (2009).

- Line 183: GPW data are at a much finer resolution. Were these regridded to 0.5x0.5 degrees?
We thank the review for pointing this out. The GPW were summed up to produce the total population that falls within each chemistry-climate model grid cell. To clarify this point we have edited the manuscript as follows.

Page 7 Line 215: "*Here, x_i represents the pollutant concentration within each model grid-cell i and p_i represents the total population (aged 30+ years) summed within each model grid-cell.*"

- Lines 426-433: should compare results also to Thompson et al. 2014, and also compare results for ozone from these studies.

We thank the reviewer for this point, although we do note the difficulty of definitive comparisons as all the studies we compare our results to are for the USA. To make clear that our study region is Europe we have added 'for Europe' to the title of the manuscript. We have also added the following text comparing our O_3 and $PM_{2.5}$ results to these U.S. based studies, respectively.

Page 15 line 465*: "For U.S. averaged mortality estimates, Punger and West (2013) show that mortality estimates related to long-term O_3 exposure, calculated using the O_3 concentrations at 36 km, were higher (by 12%) than estimates calculated at the 12 km resolution. Resolution was also found to play and important role in determining health benefits associated with differences in O_3 between 2005 and 2014 in the U.S. (Thompson et al. 2014). In particular, in urban areas, Thompson et al. (2014) estimate that the benefits calculated using coarse resolution results were on average two times greater than estimates calculated using the finer scale results. Both the studies mentioned are conducted in the U.S. and use a different concentration response coefficient and thus a definitive comparison between these studies and our estimates over Europe is not possible."*

Page 17 line 512: *"In contrast, Thompson et al. (2014) find that health benefits associated with changes in $PM_{2.5}$ concentrations between 2005 and 2014 in the U.S., were not sensitive to resolution."*

Referee #2

The manuscript by Fenech et al. considers the impact of model resolution (140 km vs 50 km) on the attributable fraction of premature mortality to O3 and PM2.5 in Europe. This question of model resolution influences on such health effects estimates has been raised previously in a few other targeted studies but has yet to be evaluated in Europe at these scales. The authors find that the impact of resolution is spatially variable, and significant. Hence, this study is of value of the community for better understanding health impact assessments in Europe, and contributes more broadly to a body of work that helps us understand the mechanisms governing scale dependencies. The manuscript is clearly organized and easy to read. There are through some areas where the analysis could be more focused, and I have some concerns related to model performance at the two different resolutions, and how that translates into a potential recommendation for future research into health impacts. These aspects and others are described in detail below; addressing them will constitute minor revisions, after which this paper will be suitable for publication in ACP.

We thank the reviewers for their positive comments and encouragement.

Major comments:

156: I understand the authors motivation here, to isolate the impact of model resolution from the impact of resolving differences in baseline mortalities. However, I disagree with their approach. But computing country-level AF and country-level baseline mortalities, the authors neglect any impact on mortality estimates that may come from sub-national variability in AF and baseline mortalities. It seems to me that a better (?) approach would be to map the O3 and PM2.5 concentrations from both the coarse and fine simulations to the same fine-scale resolution of the available population and baseline mortality rate information. This way they would have a consistent comparison that isolates the impact of the air quality model resolution, but their final estimates of mortality would be more accurate and more sensitive to differences in the air quality model resolutions. I'd suggest they at least consider this approach, which is just a postprocessing step and

doesn't involve any more model simulations, to see if it makes a significant difference, or explain why it isn't the recommended approach.

We thank the review for raising this point. Sub-national mortality rates that take into account variations in mortality within each country are not readily available across most European countries. However, we do account for sub-national variability in pollutant concentrations by applying population-weighting as described in Eq. (2) Section 2.3. If we were to calculate the differences in attributable fractions between the two resolutions at the model grid-level, their spatial distribution would be identical to that of the differences in warm season MDA8 O_3 and annual-mean $PM_{2.5}$ concentrations depicted in Fig. 5 as the AF is only dependent on the pollutant concentration and β (which is not available at the grid-cell nor country-level) following Eq. (1) Section 2.3. To illustrate this point we have calculated the difference in AF attributable to summer mean MDA8 O_3 at the grid-cell level (Fig. R4). Fig. R4 re-produces the spatial distributions of Fig. 5a in the manuscript with a scaling applied to the concentrations. For this reason we do not feel that calculating differences in AF at the grid-cell level between the two resolutions would add extra value to the manuscript. We have added the following text to section 4.3 to improve clarity and explain this point:

Page 14, line 442: *"If the AF was calculated for each model grid-cell rather than at the country level, the differences in AF for the two pollutants would have identical spatial distributions to the differences in warm season MDA8 O_3 and annual-mean $PM_{2.5}$ concentrations depicted in Fig. 5, as the AF is only dependent on the pollutant concentration and β (which is constant across all countries)."*

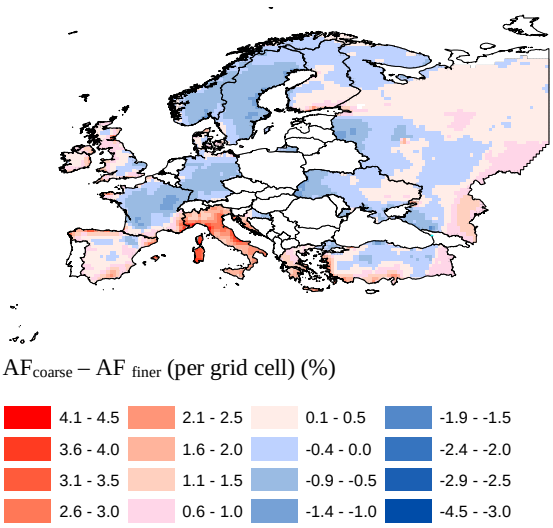


Figure R4: Differences in AF associated with long-term exposure to summer mean MDA8 O_3 between the two resolutions expressed as a percentage for each model grid-cell ($AF_{coarse} - AF_{finer}$) using a threshold of 70 µg m^{-3}.

When presenting the AF results, it would be interesting to know if the differences between the two scales of analysis are greater than the error bars in the AF estimates stemming from the uncertainty in the concentration response parameter (beta). In other words, when are the model-dependent differences significant, compared to the health-data uncertainties? See papers by Thompson et al. in this regard.

We thank the review for pointing out this omission in our results. We have added the following text to the manuscript on this point in Sections 4.3 and 4.4.

Page 15 Line 458: "*When considering the uncertainty associated with the concentration-response coefficient used, the sign of the difference of AF between the two model resolutions is unaltered (Fig. 7b). Over the majority of the countries, the AF attributable to long-term exposure to MDA8 O_3 by the coarse resolution fall within the range of uncertainty as calculated by the finer resolution (Fig. 7a). However, over Finland and Ireland, the coarse mean estimates fall outside the uncertainty range estimates using the finer resolution (Fig. 7a).*"

Page 16, line 499, *"For a number of countries, the mean AF attributable to long-term exposure to $PM_{2.5}$ using the coarse resolution falls outside the uncertainty range of the finer estimates in particular over Iceland and Ireland (Fig. 8a)"*

Page 17, line 523, "*For differences in AF attributable to long-term exposure to summer mean MDA8 O_3 and annual mean $PM_{2.5}$ concentrations, the uncertainty associated with the concentration-response coefficient used does not alter the sign of the difference of AF between the two model resolutions (Fig. 7b and 8b). The uncertainty ranges for the $PM_{2.5}$ –related estimates show a greater variability between the two resolutions for more countries compared to MDA8 O_3-related AF estimates. Using the concentration-response coefficient in Jerrett et al. (2009), Thompson et al. (2014) find that the avoided mortalities due to difference in ozone concentrations between 2005 and 2014 at a 36 km model resolution are within the 95% uncertainty range associated with the concentration-response coefficient used compared to estimates at a resolution of 12 km and 4 km. These authors also find avoided mortalities associated with long-term effects of $PM_{2.5}$ exposure at 36 km to fall within estimates at the 12 km and 4 km resolution for three different concentration-response coefficients. Thus our results are in agreement for summer mean O_3 but less for annual mean $PM_{2.5}$*"

It seems somewhat problematic, in terms of drawing conclusions, that the fine-scale simulated concentrations are, in many seasons, a poorer match to the observations than the coarse scale simulations, for both O3 and PM2.5. I strongly insist that the authors should present the statistical evaluation of biases in O3 during the warm sea-son and annual average PM2.5 in the main text, not the supplemental, as these are the scales most relevant to the focus of this work (health impacts). This is rather critical information that the reader shouldn't have to dig for.

We thank the reviewer for this valid point and we agree that the statistical evaluation of biases in O_3 during the warm season and annual average $PM_{2.5}$ currently in Table S1 would fit better in the main text. Hence Table S1 has now been moved to Table 3.

Overall, given these biases, would the authors recommend using the fine scale model over the coarse scale model for health impact analysis, especially for PM2.5 where the bias in the annual average concentration is higher at the finer scale? Or are there enough observations to say which is better at estimating exposure? This wasn't clear to me. I think this warrants some discussion, with conclusions in the abstract and conclusions.

This is an interesting point. However, we cannot and do not wish to state if one model resolution is 'better' than the other in terms of health impact analysis, because this depends on many factors and the specific comparison. The main message we wish to convey is that the differences in pollutant concentrations between the two model resolutions, which in turn drive differences in AF, vary spatially and that we can quantify the ranges of these differences and explain why these occur. As already mentioned in the abstract and conclusion, for $PM_{2.5}$ concentrations the coarse resolution results in a lower bias in spring and autumn, while the finer resolution results in a lower bias in winter and summer. For annual $PM_{2.5}$ concentrations, the absolute difference in mean concentrations between the two resolutions is small (1.1 μg m^{-3}) hence, it is difficult to derive robust conclusions about which model resolution produces better results. This is not to say that the model performance in world regions where the ranges in $PM_{2.5}$ are greater might differ substantially between the two resolutions.

Moreover, as mentioned in the manuscript (Page 11 Line 335), we note that the available sites measuring $PM_{2.5}$ during our study period are not representative of the whole domain as measurements are lacking in the eastern part of Europe where we find higher annual mean $PM_{2.5}$ concentrations simulated at the coarse compared to finer configuration, particularly in summer and autumn (Fig 1 for site locations and Figs. 4 and 5 for seasonal/annual-mean $PM_{2.5}$ differences between the two resolution in the manuscript). In addition, whilst the bias in seasonal mean O_3 concentrations is higher for the finer resolution compared

to the coarse resolution in most seasons, the concentrations at the finer resolution in some locations capture the high NO_2 and low O_3 concentrations associated with highly populated and thus polluted regions (Fig. 3). However, again we cannot definitively say which resolution more realistically estimates O_3 and NO_2 concentrations as available as site locations are not representative across the whole domain (Fig 1).

Also, model bias relative to observations should be considered when discussing regional differences in modeled spatial resolution of population-weighted concentrations (section 4.2) – in other words, is one model resolution notably better in heavily populated areas? This is a critical question which I couldn't find a direct evaluation of, although all the pieces are available to make the comparison. The same comment applies to comparison of AF (section 4.3).

We thank the reviewer for this comment and we agree this would be very insightful to add to the manuscript. However all the measurement locations in EMEP are urban background sites and not in densely populated areas, following the criteria for urban background site classification from Tørseth et al. (2012) and the EMEP manual (EMEP/CCC, 2001). For example, the minimum distance to emission and contamination sources from towns, power plants and major motorways is 50 km. Therefore these model to observation comparisons do not allow us to distinguish densely and non-densely populated locations. For clarification, we have added the following correction to the manuscript in Section 2.2 and Section 4.2.

Page 5 line 154: "*We note that all EMEP stations are classified based on a specific distance away from emission sources so as to be representative of larger areas. For example the minimum distance from large pollution sources such as towns and power plant is ~ 50 km (Torseth et al., 2012; EMEP/CCC, 2001).*"

Page 14 line 425: "*It would be insightful to examine these population-weighted results in relation to model-observation biases in densely populated areas. However, as outlined in Section 2.2, the available sites in the EMEP database are urban background stations which are required to be representative of a wide area and away from urban and industrial areas (EMEP/CCC,2001). Nonetheless we do note that in*

southern Europe, simulated summer mean MDA8 O_3 concentrations at the finer resolution are closer to observations than concentrations simulated at the coarse resolution. We find no consistent result for model biases in simulated annual mean $PM_{2.5}$ concentrations with respect to observations for the two model resolutions."

For AF estimates, comparison to observation is not possible as our estimates are calculated at the country level.

426 - 437: Regarding comparison to studies in the US, I think an interesting conclusion is that the differences owing to model resolution is not something that is consistent in sign, spatially (or that could thus be easily corrected for without knowing the spatial dependence). This is self consistent with their own evaluation of the variability of the difference across regions within Europe. Still, one might hypothesize about additional factors that control these differences. Did the authors consider the speciation of the PM2.5 and how this might affect the differences between coarse and fine scale simulations? For example, both Punger and West (2013) and Li et al. (2015) note that the differences are more significant for primary anthropogenic PM (e.g., BC) than secondary anthropogenic PM or primary natural PM. I contrast, Thompson 2014 noted the biggest impact of resolution going from 36 km to 4 km was for secondary PM. I didn't see PM2.5 speciation discussed anywhere in the present work.

We thank the reviewer for their insights. We had analysed differences in $PM_{2.5}$ components between the two resolutions. However we found no substantial differences. We have added the following text in Section 4.4 to highlight previous findings and our results.

Page 17 Line 512: "*In contrast, Thompson et al. (2014) find that health benefits associated with changes in $PM_{2.5}$ concentrations between 2005 and 2014 in the U.S., were not sensitive to resolution. Both Punger and West (2013) and Li et al. (2015) find that differences in $PM_{2.5}$ are mainly attributable to primary*

anthropogenic PM, while Thompson et al. (2014) attribute the greatest differences (between 36 km and 4 km resolutions) to secondary PM. However, in our study no substantial differences in $PM_{2.5}$ components between the two resolutions were found."

Minor comments:

1.13: Given that it is a regional modeling study, would make more sense to phrase as "at resolutions typical of global (_140 km) and regional (_50km) models." Throughout, it would make more sense to me if the results were referred to as "coarse" vs "fine" rather than "global" vs "regional", all results are regional and this could be misleading to someone just glancing at the figures. Further, most regional models these days are more like 12 km scale or finer.

Yes this is a good point and we agree. Therefore throughout the manuscript, the words "*global*" and "*regional*" have been changed to "*coarse*" and "*finer*". For further clarity as noted to our responses to Reviewer 1, we have added the words 'for Europe' to the title of the manuscript.

1.15: The differences seem a bit more modest, all less than 30% and most less than 10%. Not sure if "strong" is the right word.

We have now removed the word "*strong*" from lines 16, 293 and 539 to address this point.

28-33: Readers not familiar with AF may think these numbers are very small – the authors might wish instead (or additionally) to present the amount by which AF is changing owing to model resolution (i.e. a factor of two to three). So consider not the changes in total mortality (which is small, 5%) but instead the % changes in pollution attributed mortality. I think the authors should also state the differences in the total over the entire domain, rather than just the range across regions, even if the total benefits from some fortuitous cancelation of under and over estimates.

We agree that some readers may think these numbers are small. Therefore when stating the ranges of differences in AF between the two resolutions, we have added the following text to sections 4.3 and 4.4 to highlight these factors when discussing individual countries with the largest percentage changes. We

left the numbers in the abstract as is as the percentage values for differences in AF are of the same order of magnitude as the differences in concentrations between the two resolutions given above.

Page 15 Line 455*: "In Poland and Portugal the estimated AF at the finer resolution is 1.4 times and 0.7 times respectively that estimated at the coarse resolution."*

Page 16 Line 495*: "For Cyprus and Croatia, using the finer resolution results in an estimated AF that is 1.5 and 0.7 times that estimated using the coarse resolution."*

Differences in the AF total over the entire domain are outweighed by cancelation of under and over estimates leading to a very small average difference and since we wish to highlight regional differences we feel that this addition would not add value to the manuscript.

52-62: I'm not sure how "strong" the effects are that are being discussed in this paragraph – can the authors be more quantitative when describing previous works? The following paragraph on PM2.5 is much better in this regards.

We thank the reviewer for this suggestion. We have now updated these lines in Section 1 to include numerical values from the Stock et al. (2014) paper as indicated in the text below. However Valari and Menut (2008) state "the model is more sensitive to changes in the resolution of emissions than in meteorological input" with no quoted numerical values and only refer to Fig. 7. Hence we did not feel we could provide a further quantitative analysis since the authors did not explicitly state this.

Page 2 lines 60-65: "*Furthermore, Stock et al. (2014) found the impact of spatial resolution (150km vs. 40km) on simulated O_3 concentrations to vary with season across Europe. In winter, higher NO_x concentrations produced more pronounced titration effects on O_3 at 40 km resolution with a mean bias error (MBE) of 3.2%, leading to lower O_3 concentrations than at 150 km resolution (MBE = 14.4%). In summer, although similar results were found for O_3 concentrations simulated at the coarse (MBE =*

29.7%) and fine resolution (MBE = 32.8%) simulated boundary layer height was suggested to be largely responsible for the spatial differences in O_3 concentrations at the two resolutions."

68: The reference is Li et al. (2015). not 2016. The authors refer to the paper both ways in the text but only include an incorrect reference to 2016 in the bibliography.

We thank the review for pointing out this mistake. We have now corrected this reference to "*Li et al. (2015)"* both in the text and in the reference list.

78: The authors should also consider the results of Thomson and Selin (2012), which found there were some differences between O3-related premature deaths at the 36 km scale and finer (24, 12, and 4 km) scales, although these tended to lie within the range of uncertainty of the health impact estimates. Further, they should discuss the O3 health impact results from Thompson 2014, which were found to be more sensitive to resolution than PM2.5 health impacts.

We thank the reviewer for this suggestion. We have now added the following text to the revised manuscript in Section 1 to address both points.

Page 3 line 79-84: **"***Thompson et al. (2014) also found that especially in urban areas, the human health impacts associated with differences in O_3 between 2005 and 2014 calculated using a coarse resolution model (36 km) were on average two times greater than those estimated using finer scale resolutions (12 km and 4 km). In addition,* Thompson and Selin (2012) *found that the estimated avoided O_3-related mortalities between a 2006 base case and a 2018 control policy scenario at a 36 km resolution were higher compared to estimates at finer resolutions (12 km , 4 km and 2 km). However, their health estimates at the 36 km resolution fall within the range of values obtained using concentrations simulated at the finer resolutions used ."*

Table 1: The placement of the "difference" row is confusing. It is the difference in the model simulated mean, and should be labeled as such and more clearly located directly below the row reporting the mean, not the rows reporting NMB and SD. The choice of significant figures for the

difference also seems odd. For example, why the DJF mean difference is reported as 10 rather than 9.6 isn't clear, when other numbers are resolved to the tenth of a ug/m3. Same comments apply to Table 2. Annual average PM2.5 and warm-season (April - Sept) O3 should be added to these tables, not put in the SI.

We agree it is more intuitive to have the difference row below the mean concentration. Tables 1 and 2 have been updated accordingly and percentages have been rounded to 1 d.p. to be consistent with the rest of the values quoted. We have double checked the figures and found that the difference in magnitude of the mean spring $PM_{2.5}$ concentrations between the two resolutions was incorrect (the sign remains unchanged). This has been updated from -27% to -5.5% in Table 2, Abstract, main text and Conclusions.

References

Anenberg, S. C., West, J. J., Fiore, A. M., Jaffe, D. A., Prather, M., Bergmann, D., Cuvelier, K. and Dentener, F. J.: Intercontinental Impacts of Ozone Pollution on Human Mortality, Environ. Sci. Technol., 43(17), 6482–6487, doi:10.1029/2008GM000741/summary, 2009.

Brauer, M., Amann, M., Burnett, R. T., Cohen, A., Dentener, F., Ezzati, M., Henderson, S. B., Krzyzanowski, M., Martin, R. V., Van Dingenen, R., Van Donkelaar, A. and Thurston, G. D.: Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution, Environ. Sci. Technol., 46(2), 652–660, doi:10.1021/es2025752, 2012.

Burnett, R. T., Arden Pope, C., Ezzati, M., Olives, C., Lim, S. S., Mehta, S., Shin, H. H., Singh, G., Hubbell, B., Brauer, M., Ross Anderson, H., Smith, K. R., Balmes, J. R., Bruce, N. G., Kan, H., Laden, F., Prüss-Ustün, A., Turner, M. C., Gapstur, S. M., Diver, W. R. and Cohen, A.: An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure, Environ. Health Perspect., 122(4), 397–403, doi:10.1289/ehp.1307049, 2014.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A.,

Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C. A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C. J. L. and Forouzanfar, M. H.: Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, Lancet, 389(10082), 1907–1918, doi:10.1016/S0140-6736(17)30505-6, 2017.

Colette, A., Bessagnet, B., Vautard, R., Szopa, S., Rao, S., Schucht, S., Klimont, Z., Menut, L., Clain, G., Meleux, F., Curci, G. and Rouïl, L.: European atmosphere in 2050, a regional air quality and climate perspective under CMIP5 scenarios, Atmos. Chem. Phys., 13(15), 7451–7471, doi:10.5194/acp-13-7451-2013, 2013.

COMEAP: Quantification of Mortality and Hospital Admissions Associated with Ground- level Ozone., 2015.

EMEP/CCC: EMEP Manual for Sampling and Analysis, Norwegion Inst. Air Res. [online] Available from: https://www.nilu.no/projects/ccc/manual/index.html (Accessed 14 March 2018), 2001.

Gryparis, A., Forsberg, B., Katsouyanni, K., Analitis, A., Touloumi, G., Schwartz, J., Samoli, E., Medina, S., Anderson, H. R., Niciu, E. M., Wichmann, H. E., Kriz, B., Kosnik, M., Skorkovsky, J., Vonk, J. M. and Dörtbudak, Z.: Acute effects of ozone on mortality from the "Air Pollution and Health: A European Approach" project, Am. J. Respir. Crit. Care Med., 170(10), 1080–1087, doi:10.1164/rccm.200403-333OC, 2004.

Health Effects Institute: Outdoor Air Pollution and Health in the Developing Countries of Asia: A Comprehensive Review., 2010.

Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B. and Kaufman, J. D.: Long-term air pollution exposure and cardio- respiratory mortality: a review, Environ. Heal., 12(1), 43, doi:10.1186/1476-069X-12-43, 2013.

Jerrett, M., Burnett, R. T., Pope, C. A., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E. and Thun, M.: Long-Term Ozone Exposure and Mortality, N. Engl. J. Med., 360(11), 1085–1095, doi:10.1056/NEJMoa0803894, 2009.

Li, Y., Henze, D. K., Jack, D. and Kinney, P. L.: The influence of air quality model resolution on health impact assessment for fine particulate matter and its components, Air Qual. Atmos. Heal., 9(1),

51–68, doi:10.1007/s11869-015-0321-z, 2015.

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., Blore, J. D., Blyth, F., Bonner, C., Borges, G., Bourne, R., Boussinesq, M., Brauer, M., Brooks, P., Bruce, N. G., Brunekreef, B., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Bull, F., Burnett, R. T., Byers, T. E., Calabria, B., Carapetis, J., Carnahan, E., Chafe, Z., Charlson, F., Chen, H., Chen, J. S., Cheng, A. T. A., Child, J. C., Cohen, A., Colson, K. E., Cowie, B. C., Darby, S., Darling, S., Davis, A., Degenhardt, L., Dentener, F., Des Jarlais, D. C., Devries, K., Dherani, M., Ding, E. L., Dorsey, E. R., Driscoll, T., Edmond, K., Ali, S. E., Engell, R. E., Erwin, P. J., Fahimi, S., Falder, G., Farzadfar, F., Ferrari, A., Finucane, M. M., Flaxman, S., Fowkes, F. G. R., Freedman, G., Freeman, M. K., Gakidou, E., Ghosh, S., Giovannucci, E., Gmel, G., Graham, K., Grainger, R., Grant, B., Gunnell, D., Gutierrez, H. R., Hall, W., Hoek, H. W., Hogan, A., Hosgood, H. D., Hoy, D., Hu, H., Hubbell, B. J., Hutchings, S. J., Ibeanusi, S. E., Jacklyn, G. L., Jasrasaria, R., Jonas, J. B., Kan, H., Kanis, J. A., Kassebaum, N., Kawakami, N., Khang, Y. H., Khatibzadeh, S., Khoo, J. P., Kok, C., et al.: A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010, Lancet, 380(9859), 2224–2260, doi:10.1016/S0140-6736(12)61766-8, 2012.

Malley, C. S., Henze, D. K., Kuylenstierna, J. C. I., Vallack, H. W., Davila, Y., Anenberg, S. C., Turner, M. C. and Ashmore, M. R.: Updated global estimates of respiratory mortality in adults ≥ 30 years of age attributable to long-term ozone exposure, Environ. Health Perspect., 125(8), 1–9, doi:10.1289/EHP1390, 2017.

Markakis, K., Valari, M., Perrussel, O., Sanchez, O. and Honore, C.: Climate-forced air-quality modeling at the urban scale: sensitivity to model resolution, emissions and meteorology, Atmos. Chem. Phys., 15(13), 7703–7723, doi:10.5194/acp-15-7703-2015, 2015.

Punger, E. M. and West, J. J.: The effect of grid resolution on estimates of the burden of ozone and fine particulate matter on premature mortality in the United States., Air Qual. Atmos. Health, 6(3),

563–573, doi:10.1007/s11869-013-0197-8, 2013.

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., Rouïl, 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, doi:10.1016/j.atmosenv.2015.04.003, 2015.

Stock, Z. S., Russo, M. R. and Pyle, J. A.: Representing ozone extremes in European megacities: the importance of resolution in a global chemistry climate model, Atmos. Chem. Phys., 14, 3899–3912, doi:10.5194/acp-14-3899-2014, 2014.

Thompson, T. M. and Selin, N. E.: Influence of air quality model resolution on uncertainty associated with health impacts, Atmos. Chem. Phys., 12(20), 9753–9762, doi:10.5194/acp-12-9753-2012, 2012.

Thompson, T. M., Saari, R. K. and Selin, N. E.: Air quality resolution for health impact assessment: influence of regional characteristics, Atmos. Chem. Phys., 14, 969–978, doi:10.5194/acp-14-969-2014, 2014.

Tie, X., Brasseur, G. and Ying, Z.: Impact of model resolution on chemical ozone formation in Mexico City: application of the WRF-Chem model, Atmos. Chem. Phys., 10(18), 8983–8995, doi:10.5194/acp-10-8983-2010, 2010.

Tørseth, K., Aas, W., Breivik, K., Fjeraa, A. M., Fiebig, M., Hjellbrekke, A. G., Lund Myhre, C., Solberg, S. and Yttri, K. E.: Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972-2009, Atmos. Chem. Phys., 12(12), 5447–5481, doi:10.5194/acp-12-5447-2012, 2012.

Turner, M. C., Jerrett, M., Pope, C. A., Krewski, D., Gapstur, S. M., Diver, W. R., Beckerman, B. S., Marshall, J. D., Su, J., Crouse, D. L. and Burnett, R. T.: Long-Term Ozone Exposure and Mortality in a Large Prospective Study, Am. J. Respir. Crit. Care Med., 193(10), 1134–1142, doi:10.1164/rccm.201508-1633OC, 2015.

Valari, M. and Menut, L.: Does an Increase in Air Quality Models' Resolution Bring Surface Ozone Concentrations Closer to Reality?, J. Atmos. Ocean. Technol., 25(11), 1955–1968,

doi:10.1175/2008JTECHA1123.1, 2008.

WHO: Health Risks of Air Pollution in Europe – HRAPIE project: Recommendations for concentration-response functions for cost-benefit analysis of particulate matter, ozone and nitrogen dioxide. [online] Available from: http://www.euro.who.int/__data/assets/pdf_file/0006/238956/Health-risks-of-air-pollution-in-Europe-HRAPIE-project,-Recommendations-for-concentrationresponse-functions-for-costbenefit-analysis-of-particulate-matter,-ozone-and-nitrogen-dioxide., 2013.

The influence of model spatial resolution on simulated ozone and fine particulate matter for Europe: implications for health impact assessments

Sara Fenech[1,2], Ruth M. Doherty[1], Clare Heaviside[2], Sotiris Vardoulakis[3], Helen L. Macintyre[2], Fiona M. O'Connor[4]

[1] School of GeoSciences, University of Edinburgh

[2] Centre for Radiation, Chemical and Environmental Hazards, Public Health England

[3] Institute of Occupational Medicine

[4] Met Office, Hadley Centre, Exeter, UK

Correspondence to: Sara Fenech (S.Fenech@sms.ed.ac.uk)

Abstract. We examine the impact of model horizontal resolution on simulated surface ozone (O_3) and particulate matter less than 2.5 μm ($PM_{2.5}$) concentrations, and the associated health impacts over Europe, using the HadGEM3-UKCA chemistry-climate model to simulate pollutant concentrations over Europe at a ~~global~~coarse (~ 140 km) and a ~~regional~~finer (~ 50 km) resolution. The attributable fraction (AF) of total mortality due to long-term exposure to warm season daily maximum 8-hour running mean (MDA8) O_3 and annual-average $PM_{2.5}$ concentrations is then calculated for each European country using pollutant concentrations simulated at each resolution. Our results highlight a ~~strong~~ seasonal variation in simulated O_3 and $PM_{2.5}$ differences between the two model resolutions in Europe. Compared to the ~~regional~~finer resolution results, simulated European O_3 concentrations at the ~~global~~coarse resolution are on average higher in winter and spring (10% and 6%, respectively). In contrast, simulated O_3 concentrations at the ~~global~~coarse resolution are lower in summer and autumn (-1% and -4%, respectively). These differences may partly be explained by differences in nitrogen dioxide (NO_2) concentrations simulated at the two resolutions. Compared to O_3, we find the opposite seasonality in simulated $PM_{2.5}$ differences between the two resolutions. In winter and spring, simulated $PM_{2.5}$ concentrations are lower at the ~~global~~coarse compared to the ~~regional~~finer resolution (-8% and -~~6~~27%, respectively) but higher in summer and autumn (29% and 8%, respectively) and are mostly related to differences in convective rainfall between the two resolutions for all seasons. These differences between the two resolutions exhibit clear spatial patterns for both pollutants that vary by season, and exert a strong influence on country to country variations in estimated AF for the two resolutions. Warm season MDA8 O_3 levels are higher in most of southern Europe, but lower in areas of northern and eastern Europe when simulated at the ~~global~~coarse resolution compared to the ~~regional~~finer resolution. Annual-average $PM_{2.5}$ concentrations are higher across most of northern and eastern Europe but lower over parts of southwest Europe at the ~~global~~coarse compared to the ~~regional~~finer resolution. Across Europe, differences in the AF associated with long-term exposure to population-weighted MDA8 O_3 range between -0.9 % and +2.6 % (largest positive

differences in southern Europe) while differences in the AF associated with long-term exposure to population-weighted annual mean $PM_{2.5}$ range from -4.7% to +2.8% (largest positive differences in eastern Europe) of the total mortality. Therefore this study, with its unique focus on Europe, demonstrates that health impact assessments calculated using modelled pollutant concentrations, are sensitive to a change in model resolution by up to ±5% of the total mortality across Europe.

1 Introduction

A substantial number of epidemiological studies have derived risk estimates for mortality associated with long-term exposure to ambient fine particulate matter with aerodynamic diameter less than 2.5 µm ($PM_{2.5}$) (Krewski, 2009;Brook et al., 2010; WHO, 2013) and also recently, to a lesser extent, for long-term exposure to ozone (O_3) (Jerrett et al., 2009; Forouzanfar et al., 2016; Turner et al., 2016). Differences in risk estimates produced from different epidemiological studies can be due to differences in methodologies, air pollution and health data used including the size and spatial extent of cohort populations. For O_3, these long-term risk estimates are derived from North American studies. In this region O_3 data is typically monitored only during the O_3 season (April-September), hence these derived O_3-risk estimates apply only to the ozone occurring in the warm season part of the year.

Air pollutant exposure estimated from concentrations measured at fixed monitoring stations, is often used to estimate health impacts at the cohort-scale (Cohen et al., 2004). However, quantifying the adverse health effects of air pollution at the continental-scale requires ~~global or regional~~ atmospheric models (with resolutions ranging from ~250 to 50 km) to simulate pollutant spatio-temporal distributions across these scales (e.g. West et al. 2009; Anenberg et al. 2010; Fang et al. 2013; Silva et al. 2013; Lelieveld et al. 2015, Malley et al., 2017). Amongst a number of factors, simulated air pollutant concentrations may vary depending on the three-dimensional chemistry model used, its set-up and the model resolution (e.g. Markakis et al. 2015; Schaap et al. 2015; Yu et al. 2016; Neal et al. 2017). Although the same model processes are represented at different model resolutions, simulated pollutant concentrations can vary due to differences in (i) the resolution of emissions, which may have a nonlinear effect on the chemical formation of pollutants, and (ii) the resolution of the driving meteorology (Valari and Menut 2008; Tie et al. 2010; Arunachalam et al. 2011; Colette et al. 2013; Markakis et al. 2015; Schaap et al. 2015)**.**

The impact of model horizontal resolution on simulated O_3 concentrations has been primarily linked to less dilution of emissions when using a finer resolution (Valari and Menut 2008; Tie et al. 2010; Colette et al. 2013; Stock et al. 2014; Schaap et al. 2015). Investigating the impact of increasing model horizontal resolution from 48 km to 6 km on O_3 concentrations in Paris, Valari and Menut (2008) found modelled surface O_3 to be more sensitive to the resolution of input emissions than to meteorology. A number of other studies note the sensitivity of simulated O_3 to simulated nitrogen oxide (NO_x) concentrations that determine the extent of titration of O_3 by nitrogen monoxide (NO) (Stock et al., 2014; Markakis et al., 2015; Schaap et al., 2015). Furthermore, Stock et al. (2014) ~~further~~ found the impact of spatial resolution (150km vs. 40km) on simulated O_3 concentrations to vary with season across Europe. In winter, higher NO_x concentrations produced more pronounced titration effects on O_3 at 40 km resolution with a mean bias error (MBE) of 3.2%, leading to lower O_3

concentrations than at 150 km resolution (MBE = 14.4%). In summer, although similar results were found for O_3 concentrations simulated at the coarse (MBE = 29.7%) and fine resolution (MBE = 32.8%) ~~the~~ simulated boundary layer height was suggested to be largely responsible for the spatial differences in O_3 concentrations at the two resolutions.

$PM_{2.5}$ concentrations have also been found to be sensitive to the model horizontal resolution (Arunachalam et al. 2011; Punger and West 2013; Markakis et al. 2015; Neal et al. 2017). In the U.S., Punger and West (2013) found population-weighted annual mean $PM_{2.5}$ concentrations to be 6% higher at 36 km compared to 12 km, but 27% lower when simulated at 408 km compared to 12 km. However in this study, statistical averaging was used to estimate pollutant concentrations at the coarsest resolutions, and therefore differences in emissions and meteorology and their atmospheric processing between the resolutions were not included. In contrast, Li et al. (2015) found annual mean $PM_{2.5}$ concentrations simulated at a resolution of ~ 2.5° in the U.S. to be similar to $PM_{2.5}$ concentrations simulated at a resolution of ~ 0.5° suggesting that the horizontal scales being compared and the methodology for comparison are important. However maximum $PM_{2.5}$ concentrations which occur in highly populated regions were found to be 21% lower at the coarse resolution (Li et al., 2015).

As outlined above, a number of studies have analysed the effect of model resolution on O_3 and $PM_{2.5}$ concentrations but few have looked at the sensitivity of the associated health impacts to model resolution (Punger and West 2013; Thompson et al. 2014; Li et al. 2015). Punger and West (2013) found mortality associated with long-term exposure to O_3 in the US to be 12% higher at a 36 km resolution compared to the mortality estimate at 12 km, as a result of higher O_3 simulated at the coarser-scale. Thompson et al. (2014) also found that especially in urban areas, the human health impacts associated with differences in O_3 between 2005 and 2014 calculated using a coarse resolution model (36 km) were on average two times greater than those estimated using finer scale resolutions (12 km and 4 km). In addition, ~~(~~Thompson and Selin~~,~~ (2012) found that the estimated avoided O_3-related mortalities between a 2006 base case and a 2018 control policy scenario at a 36 km resolution were higher compared to estimates at the finer resolutions (12 km , 4 km and 2 km) . However, their health estimates at the 36 km resolution fall within the range of values obtained using concentrations simulated at the finer resolutions used.

For $PM_{2.5}$-related health estimates, studies by Punger and West (2013) and Li et al. (2015) both found that attributable mortality associated with long-term exposure to $PM_{2.5}$ in the US was lower for their coarser resolution simulations (> 100 km) due to lower simulated $PM_{2.5}$ concentrations in densely populated regions. However, Thompson et al. (2014) found that using model horizontal resolutions of 36, 12 and 4 km had a negligible effect on changes in $PM_{2.5}$ concentrations and associated health impacts. This is likely due to the relatively small range of resolutions used by Thompson et al. (2014) compared to these other studies.

The majority of health effect studies relating to the impact of model resolution have been conducted in North America. Hence, similar studies are lacking over Europe. This study is therefore the first to examine the impact of two different model resolutions: a ~~global~~coarse (~ 140 km) and ~~regional~~ a finer resolution (~ 50 km) on O_3 and $PM_{2.5}$ concentrations, and their subsequent impacts on European-scale human health through long-term exposure to O_3 and $PM_{2.5}$. We define the sensitivity of health impacts to model resolution by calculating the attributable fraction (AF) of total mortality which is associated with

long-term exposure to O_3 and $PM_{2.5}$ for various European countries, based on simulated concentrations at both resolutions, and expressed as a percentage.

The remainder of the paper is organised as follows. Section 2 describes the modelling framework used for both the coarse and finer ~~global and regional~~ simulations and the methods used to calculate the AF of mortality associated with O_3 and $PM_{2.5}$ for various European countries. Section 3 presents differences in seasonal mean O_3 and $PM_{2.5}$ concentrations between the two resolutions. In section 4, we first analyse differences in warm season daily maximum 8-hour running mean (MDA8) O_3 concentrations and annual $PM_{2.5}$ concentrations between the two resolutions, then quantify differences in country-level population-weighted MDA8 O_3 and annual mean $PM_{2.5}$ concentrations. Secondly, the country-level AF associated with long-term exposure to MDA8 O_3 and annual mean $PM_{2.5}$ simulated at both resolutions are presented. The conclusions of this study are then presented in Section 5.

2 Methods

2.1 Model description and experimental setup

The two chemistry-climate configurations used in this study are based on the Global Atmosphere 3.0 (GA3.0) / Global Land (GL3.0) configuration of the Hadley Centre Global Environmental Model version 3 (HadGEM3, Walters et al., 2011), of the Met Office's Unified Model (MetUM, Brown et al., 2012). The coarse~~global~~ configuration has a horizontal resolution of 1.875° × 1.25° (~ 140 km, Walters et al., 2011) while the ~~regional~~finer configuration has a horizontal resolution of 0.44° × 0.44° (~50 km, Moufouma-Okia and Jones, 2014) with a domain covering most of Europe.

As this study focuses on health impacts, our analysis is restricted to European land regions. In both configurations, a 63 level hybrid height vertical co-ordinate system is used with 50 levels below 18 km and a surface level at 40 m. Gas phase chemistry is simulated within HadGEM3 by a tropospheric configuration of the United Kingdom Chemistry and Aerosol (UKCA) model (Morgenstern et al., 2009; O'Connor et al., 2014). The chemistry scheme used for both configurations is the UKCA Extended Tropospheric Chemistry (UKCA-ExtTC) scheme (Folberth et al., In prep.) which is an extension to the TropIsop standard chemistry scheme (O'Connor et al., 2014) and includes 89 chemical species. Boundary layer mixing for both configurations is based on Lock et al. (2000) and includes an explicit entrainment parametrisation and non-local mixing in unstable layers. The GA3.0/GL3.0 configuration of HadGEM3 (Walters et al., 2011) also includes an interactive aerosol scheme called CLASSIC (Coupled Large-scale Aerosol Simulator for Studies in Climate; Jones et al., 2001;Bellouin et al., 2011) from which $PM_{2.5}$ concentrations are estimated. CLASSIC simulates ammonium sulphate and nitrate, fossil-fuel organic carbon (FFOC), mineral dust, soot and biomass burning (BB) aerosol interactively. Biogenic secondary organic aerosols are prescribed from a climatology. Sea salt aerosol is diagnosed over ocean only and does not contribute to particulate matter over land.

The model simulations for both these configurations cover a period of one year and 9 months starting from April 2006, from which the first nine months were discarded as spin-up. The ~~global~~coarse configuration uses monthly mean

distributions of sea surface temperature (SST) and sea ice cover (SIC), derived for the present-day (1995-2005) from transient coupled atmosphere-ocean simulations (Jones et al., 2001) of the HadGEM2-ES model (Collins et al., 2011). Using a simple linear re-gridding algorithm, the SST and SIC climatologies developed for the ~~global~~coarse configuration were downscaled to the finer~~regional~~ configuration. The ~~global~~coarse configuration was set to produce lateral boundary conditions (LBCs) at six-hourly intervals which were then used to drive the ~~regional~~finer configuration.

A consistent set of baseline emissions have been used for both configurations by using the same source data and then re-gridding to the ~~global~~coarse and ~~regional~~finer resolutions of the chemistry-climate model. The surface emissions for chemical species were implemented from emission data at 0.5° by 0.5° resolution developed by Lamarque et al. (2010) for the Fifth Coupled Model Inter-comparison Project (CMIP5) report which include reactive gases and aerosols from anthropogenic and biomass burning sources. Both model configurations are driven by decadal mean present-day emissions from Lamarque et al. (2010), representative of the decade centred on 2000. Biogenic emission of isoprene and monoterpenes are calculated interactively following Pacifico et al. (2011) and the biogenic emissions of methanol and acetone are prescribed, taken from Guenther et al. (1995). A full description of other biogenic emissions and the ~~global~~coarse and ~~regional~~finer configurations can be found in Neal et al. (2017).

The two configurations are consistent in terms of driving meteorology and emissions as discussed above, however a change in model resolution also requires changes to model's dynamical time-step (from 20 min; ~~global~~coarse resolution to 12 min; ~~regional~~ finer resolution) as well as some of the parameters in the model parametrisations schemes that are resolution dependent. In this study we assume any such differences to be a model resolution effect. To compare pollutant concentrations simulated at the two resolutions, the ~~global~~ coarse model results were re-gridded to the finer~~regional~~ resolution via bi-linear interpolation and differences between the two configurations were then calculated at each grid box. For consistency, all figures, tables and values shown in the following sections show differences calculated as ~~global~~ coarse minus finer~~regional~~ results. All pollutant concentrations used in this study have been extracted at the lowest model level with a mid-point at 20 m. While this level is considered representative of surface or ground–level concentrations, local orographically driven flows or sharp gradients in mixing depths cannot be represented at this vertical resolution (Fiore et al. 2009).

2.2 Measurement data

Modelled seasonal mean O_3 and $PM_{2.5}$ concentrations for 2007 were evaluated using measurement data from the European Monitoring Evaluation Programme (EMEP) network (ebas.nilu.no). We note that all EMEP stations are classified based on a specific distance away from emission sources so as to be representative of larger areas. For example the minimum distance from large pollution sources such as towns and power plant is ~ 50 km (Tørseth et al., 2012; EMEP/CCC, 2001). We chose a sub-set of the available EMEP O_3 measurement sites with an altitude less than or equal to 200 m above sea level to focus on near-surface comparisons between measurements and simulated O_3 concentrations (52 sites – Fig. 1). As there are fewer measurements of $PM_{2.5}$ for 2007, all available EMEP measurement sites were used for $PM_{2.5}$ evaluation (25 sites – Fig. 1). All modelled O_3 and $PM_{2.5}$ concentrations shown in this study were taken from the lowest vertical model level which reaches a

height of 40 m. To perform an observation-model comparison, simulated pollutant concentrations were extracted at measurement site locations using bi-linear interpolation.

2.3 Health calculations

Annual total mortality estimates associated with long-term exposure to O_3 and $PM_{2.5}$ are frequently calculated by estimating the country-level Attributable Fraction (AF) of mortality, based on concentration-response relationships associated with each pollutant, and then multiplying the AF by the baseline mortality rate. Since we are interested in the effects of changing resolution on pollutant concentration, in our analysis, we focus on the absolute values and differences in the AF between the two resolutions, rather than calculating mortality associated with each pollutant, which also depends on underlying baseline mortality rates. This allows us to isolate the effect of model resolution on health impacts. We note that differences in AF will be the same as the differences in mortality between the two resolutions (expressed as a percentage of total mortality), if calculated as described in this section.

Although there is limited evidence available for the long-term health impacts of O_3 especially in Europe (The UK Committee on the Medical Effects of Air Pollution (COMEAP) 2015), a number of studies have quantified the adverse health impacts associated with long-term exposure to O_3. In this study we apply the Health Risks of Air Pollution in Europe – HRAPIE project recommended coefficient for long-term exposure to O_3 (WHO, 2013) to investigate the sensitivity of health calculations to the model resolution used to simulate O_3 concentrations. This concentration–response coefficient is derived from the single-pollutant analysis of the American Cancer Society Cancer Prevention Study II (CPS II) cohort study data in 96 metropolitan areas of the US (Jerrett et al., 2009) which has been used by previous studies (e.g. Anenberg et al., 2009; Punger and West, 2013; Thompson et al., 2014; Cohen et al., 2017); but is re-scaled from 1-hour means to 8-hour mean concentrations using the ratio 0.72, derived from the APHEA-2 project (Gryparis et al., 2004). The value recommended by HRAPIE for the concentration-response coefficient, or β value (Eq.1), for the effects of long-term O_3 exposure on respiratory mortality is 1.014 (95% Confidence Interval (CI) = 1.005, 1.024) per 10 μg m^{-3} increase in MDA8 O_3 during the warm season (April-September) with a threshold of 70 μg m^{-3} (WHO, 2013). For estimating the health impact of long-term exposure to $PM_{2.5}$ on all-cause (excluding external) mortality, HRAPIE (WHO 2013) recommends a relative risk coefficient of 1.062 (95% CI = 1.040, 1.083) per 10 μg m^{-3} increase in annual average concentrations (with no threshold) which is based on a meta-analysis of cohort studies by Hoek et al. (2013)~~Although there is limited evidence available for the long-term health impacts of O_3 especially in Europe (The UK Committee on the Medical Effects of Air Pollution (COMEAP) 2015), we apply the Health Risks of Air Pollution in Europe – HRAPIE project recommended coefficient for long-term exposure to O_3 (WHO, 2013) to investigate the sensitivity of health calculations to model resolutions as used in previous studies (Anenberg et al. 2009; Punger and West 2013; Thompson et al. 2014). The concentration-response coefficient, or β value (Eq.1), for the effects of long-term O_3 exposure on respiratory mortality recommended by HRAPIE is 1.014 (95% Confidence Interval (CI) = 1.005, 1.024) per 10 μg m^{-3} increase in MDA8 O_3 during the warm season (April-September) with a threshold of 70 μg m^{-3} (WHO, 2013). For estimating the health impact~~

~~of long-term exposure to $PM_{2.5}$ on all-cause (excluding external) mortality, HRAPIE (WHO 2013) recommends a relative risk coefficient of 1.062 (95% CI = 1.040, 1.083) per 10 μg m^{-3} increase in annual average concentrations (with no threshold).~~

For MDA8 O_3, the risk estimates above, suggested by HRAPIE, are based on data from the American Cancer Society (ACS) cohort (Jerrett et al., 2009) during the warm season re-scaled from 1-hour means to 8-hour means (WHO, 2013). Since MDA8 O_3 concentrations in the summer months exceed 70 μg m^{-3} for most areas included in the ACS study, little information exists on the shape of the concentration-response relationship at low levels. For this reason, following HRAPIE suggestions, only MDA8 O_3 concentrations exceeding 70 μg m^{-3} and averaged between April and September were used in the present study to calculate O_3-related health impacts. For $PM_{2.5}$-related health impacts we use annual averages with no threshold. As the β values used for O_3 and $PM_{2.5}$ are from the ACS cohort, the estimates in this study exclude people younger than 30 years.

For each model resolution, simulated air pollutant concentrations were used to calculate the country-average AF of respiratory or all-cause mortality associated with long-term exposure to O_3 and $PM_{2.5}$, respectively. Specifically, the country-average AF is derived from the country-averaged population-weighted pollutant concentration ($x_{country}$) and concentration-response coefficient (β) as shown in Eq. (1) (e.g. Anenberg et al. 2010; Gowers et al. 2014):

$$AF_{country} = 1 - e^{-\beta x_{country}} \quad (1)$$

The country-averaged population-weighted pollutant concentrations ($x_{country}$) were derived using gridded population data at a resolution of ~~0.05°~~5 km (GWPv3), obtained from the Socioeconomic Data and Applications Centre (SEDAC, http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-count-future-estimates/data-download), following Eq. (2).

$$x_{country} = \frac{\sum_{i \in country} (p_i \times x_i)}{\sum_{i \in country} p_i} \quad (2)$$

Here, x_i represents the pollutant concentration within each model grid-cell i and p_i represents the total population (aged 30+ years) summed within each model grid-cell. ~~Here, x_i represents the pollutant concentration within each model grid-cell i and p_i represents the number of people (aged 30+ years) exposed to the pollutant concentration also at the model grid-cell level.~~ For population-weighted $PM_{2.5}$ concentrations, the simulated $PM_{2.5}$ concentration for each model grid-cell was multiplied by the number of people within the same model grid-cell. This product was then summed for all grid-cells within the country and divided by the total population of the respective country. A similar procedure was used for MDA8 O_3 concentrations. However, for populated–weighted MDA8 O_3 concentrations, 70 μg m^{-3} was first subtracted from the simulated MDA8 O_3 concentration at each grid-cell before multiplying by the population (any resultant negative concentrations were set to zero).

3 The impact of model resolution on pollutant concentrations

3.1 The impact of model resolution on seasonal mean O_3: comparison with observations

Modelled and observed means and, standard deviations (SD), normalised mean bias (NMB) and percentage differences between the two resolutions for all four seasons at the 52 EMEP site locations are shown in Table 1. Similarly modelled means, SD and percentage differences between the two resolutions are also shown for all model cells within the European domain (discussed in Section 3.2). Compared to measurements, mean values simulated by the chemistry-climate model across the 52 station locations are lower in winter (DJF) and higher in summer (JJA) and autumn (SON) with NMB values up to -19%, 27% and 19%, respectively. In spring (MAM), simulated mean O_3 concentrations at the ~~regional~~finer resolution are closest to observations (NMB = ~ -4 %), whilst in all other three seasons the simulated values at the ~~global~~ coarse resolution are in closer agreement with observations (NMB = ~ -8%, ~24% and ~ 5%, respectively).

For all seasons, the SD of seasonal mean O_3 concentrations, simulated at the two resolutions are more similar to each other than to observations. However, the SD across all 52 sites, simulated at the ~~global~~ coarse resolution is higher than that simulated at the ~~regional~~finer resolution.

Modelled versus observed seasonal mean O_3 concentrations for each of the 52 EMEP station locations are shown in Fig. 2, with arrow lengths indicating the change in concentrations when simulated at ~~global~~coarse versus ~~regional~~finer resolutions. For both resolutions, higher O_3 concentrations are simulated during summer compared to observations as noted above (between 50 to 150 µg m^{-3}; Fig. 2). In winter, simulated O_3 concentrations are lower compared to measurements (< 30 µg m^{-3}), and are most similar to observations in spring and autumn in accordance with lower NMB (Table 1).

The magnitude of the differences in simulated O_3 concentrations between the two resolutions varies seasonally, with the smallest (coarse~~global~~-finer~~regional~~) differences in summer (green arrows – Fig. 3; -3 % ;Table 1) and the largest difference in spring, as noted above (16 % ;Table 1). Similar differences in July mean O_3 concentrations between a 150 km and a 40 km resolution were also found by Stock et al. (2014). Over the majority of the stations, during winter and spring, O_3 concentrations simulated at the finer~~regional~~ resolution are lower than concentrations simulated at the ~~global~~coarse resolution (downward arrows; Fig. 2, positive difference; Table 1). In contrast during summer and autumn, O_3 concentrations are higher when simulated at the finer~~regional~~ resolution (upward arrows; Fig. 2, negative difference; Table 1). These results are analysed further at the seasonal level in Fig. S1 of the Supplement to this article (Supplement S1~~2~~; Fig. S1).

3.2 The impact of model resolution on seasonal mean O_3: spatial differences

This section extends our investigation to examine the impact of model grid resolution on the spatial distribution of O_3 over the whole of Europe. The seasonal variation in O_3 concentrations simulated at the finer~~regional~~ resolution across Europe shows the same features as at the 52 site locations (section 3.1), with highest values in spring and summer (> 50 µg m^{-3} and up to 120 µg m^{-3}; Fig. 3b and 3c, respectively) and lowest values in autumn and winter (<55 µg m^{-3}; Fig. 3a and 3d). In all seasons,

except winter, there is a clear latitudinal gradient with higher O_3 concentrations in southern compared to northern Europe. In winter (Fig. 3a), very low O_3 concentrations are simulated across much of Europe (~30 µg m^{-3}).

For most of Europe, in winter and spring, mean O_3 concentrations are generally higher when simulated at the ~~global~~coarse compared to the ~~regional~~finer resolution (Fig. 3e and 3f, 10% and 6% respectively; Table 1), in agreement with the findings for the sub-set of 52 locations. However parts of northern Scandinavia and the UK, and parts of south-eastern Europe have lower O_3 concentrations simulated at the ~~global~~coarse resolution in these two seasons. In summer and autumn, O_3 concentrations are slightly lower when simulated at the ~~global~~ coarse compared to the ~~regional~~finer resolution (-1% and -4% respectively –Table 1) as found for the sub-set of locations, except in areas of easternmost Europe (especially in autumn) and parts of Spain and Italy (Fig. 3g and 3h). The greatest positive differences in simulated O_3 concentrations, i.e. higher values at the ~~global~~coarse resolution, are found in winter, especially in the far south of Europe in Spain (~ 20 µg m^{-3}; Fig. 3e). Some of these positive differences are clear around the coastal regions which is likely due to differences in the land/sea mask at the two resolutions, which leads to less deposition over oceanic grid-cells at the ~~global~~coarse resolution and higher simulated O_3 concentrations compared to the same locations that are designated as land at the ~~regional~~finer scale (Coleman et al., 2010). In addition, large positive differences in simulated O_3 concentrations between the two resolutions occur over the Alps, whereby simulated O_3 concentrations are higher at the ~~regional~~finer scale (Fig. 3e and 3h). This is most likely due to the differences in orography at the two resolutions with higher elevations at the ~~regional~~finer scale leading to higher O_3 concentrations.

Differences in simulated seasonal mean NO_2 concentrations at the two resolutions show similar, but less extensive differences and generally inverse patterns as for O_3 concentrations, with some negative differences, i.e. lower NO_2 values in winter and spring (Fig. 3i and 3j), when simulated at the ~~global~~coarse compared to the ~~regional~~finer resolution. In contrast, in summer and autumn, NO_2 concentrations are higher in some regions when simulated at the ~~global~~coarse compared to the ~~regional~~finer resolution (e.g. Italy; Fig. 3k and 3l). An inverse relationship i.e. a positive difference in O_3 concentrations and a negative difference in NO_2 concentrations is most prominent for locations in Spain (all year around) and Italy (winter and spring) and parts of the Benelux region (southern UK and Netherlands; all year around). This inverse relationship is driven by lower NO_x concentrations at the ~~global~~coarse resolution which lead to less O_3 titration by NO compared to the ~~regional~~finer resolution (Fig. 3i). This in turn results in higher simulated seasonal mean O_3 concentrations at the ~~global~~coarse resolution compared to the ~~regional~~finer resolution (Fig. 3e).

The planetary boundary layer (PBL) height is a key meteorological variable that affects the vertical transport of pollutants from the surface into the free troposphere from where they can then undergo strong horizontal transport. Thus we have also investigated the impact of changing model resolution on PBL height and how this impacts O_3 and NO_2 concentrations. Spatial differences in PBL height between the two resolutions are shown in Section S2~~3~~, Fig. S2 of the Supplement to this article. In all seasons, over most of western and central Europe and especially in summer, the PBL height is generally lower when simulated at the ~~global~~coarse resolution (negative differences up to 275m; Fig. S2c). In winter and spring (Fig. S2a and S2b), this lower height corresponds to generally higher O_3 concentrations but also lower NO_2 concentrations simulated at the ~~global~~coarse resolution, over the same region and vice versa in summer and autumn (Fig. S2c

and S2d). If a deeper PBL is the main driver of pollutant trapping producing higher O_3 levels, then we would also expect NO_2 concentrations to be higher with a lower PBL height at the ~~regional~~finer resolution, but their frequent inverse relationship suggest a stronger role for chemistry rather than PBL effects. However, these chemical and physical processes cannot be clearly separated.

In summary, we find a ~~strong~~ seasonal variation in simulated O_3 differences between the two resolutions. Simulated O_3 concentrations at the ~~global~~coarse resolution are higher in winter and spring and lower in summer and autumn compared to the ~~regional~~finer resolution. We also find that in a number of locations, NO_2 concentrations are lower at the ~~global~~coarse compared to the ~~regional~~finer resolution and correspond to higher O_3 concentrations at the ~~global~~coarse resolution as a result of reduced titration with lower NO_x levels. Orography also plays an important role in some coastal locations, leading to an overestimation of O_3 concentrations. The PBL height differs between the two resolutions especially during summer, with the finer~~regional~~ resolution resulting in a deeper boundary layer. However, it is not possible to separate chemistry and mixing effects on simulated O_3 concentrations.

3.3 The impact of model resolution on seasonal mean $PM_{2.5}$ – comparison with observations

Simulated seasonal mean $PM_{2.5}$ concentrations are compared to available EMEP observations at 25 sites (Table 2). Mean values for the observations are fairly similar across all seasons, with values in summer and autumn being slightly lower. $PM_{2.5}$ concentrations simulated at both the ~~global~~coarse and ~~regional~~finer resolutions are lower in winter and higher in summer compared to measurements. In addition, mean $PM_{2.5}$ concentrations simulated at the ~~regional~~finer resolution are higher than those simulated at the ~~global~~coarse resolution except in summer. The ~~global~~coarse resolution simulates $PM_{2.5}$ levels with the smallest bias during spring (NMB = -0.2%). In contrast, $PM_{2.5}$ concentrations simulated at the ~~regional~~finer resolution during spring have a large positive bias (NMB = 31%). Similarly in autumn NMB values are larger for $PM_{2.5}$ concentrations simulated at the ~~regional~~finer resolution. We find that the largest bias for both resolutions occurs in summer with the ~~global~~coarse resolution resulting in a NMB of 70%. Using, a similar ~~regional~~finer configuration, Neal et al. (2017) found a year-round small positive bias in simulated $PM_{2.5}$ concentrations averaged over a five year period (2001-2005) at two UK locations. The SD of $PM_{2.5}$ concentrations across the 25 sites is fairly similar between model results and measurements except in winter, when simulated SD values are lower at both resolutions compared to measurements and in autumn, when the SD at the ~~regional~~finer resolution is higher compared to measurements.

Modelled versus measured $PM_{2.5}$ concentrations across the 25 individual EMEP stations highlight the low simulated $PM_{2.5}$ concentrations in winter (Section S2~~3~~, Fig. S3 of the Supplement to this article). Large variations in $PM_{2.5}$ levels between the two resolutions are prominent in spring (-31%; Table 2). Smaller $PM_{2.5}$ concentrations simulated at the ~~global~~coarse resolution in winter, spring and autumn are apparent (upward arrows; Fig. S3, negative differences; Table 2).

3.4 Impact of model resolution on seasonal mean $PM_{2.5}$: spatial differences

Spatial distributions of $PM_{2.5}$ concentrations, simulated at the finer~~regional~~ resolution as well as differences between the two resolutions over the whole European domain are illustrated in Fig. 4. Over the whole domain, $PM_{2.5}$ concentrations simulated at the finer~~regional~~ resolution are lowest in winter (Fig. 4a) and highest in spring (Fig. 4b). As for O_3, there is clear latitudinal gradient with higher $PM_{2.5}$ levels in southern Europe in all seasons. Differences in seasonal mean $PM_{2.5}$ concentrations, between the coarse and fine resolutions, vary seasonally across the European domain with the smallest differences occurring during winter (± 3 µg m^{-3}; Fig. 4e, -8% ; Table 2) in agreement with the findings for the 25 EMEP stations described above (section 3.3). This suggests that at low $PM_{2.5}$ concentrations (~ 8 µg m^{-3}) in winter, model results do not differ greatly when increasing the model resolution from 150 km to 50 km. In spring, $PM_{2.5}$ concentrations simulated at the ~~global~~coarse are lower than at the ~~regional~~finer resolution over large parts of central and western Europe but are slightly higher in easternmost parts of Europe (negative differences ~-10 µg m^{-3} Fig. 4f; -6~~27~~% Table 2), as found at the 25 EMEP station locations. The opposite result occurs in summer with generally higher $PM_{2.5}$ concentrations simulated at the coarser resolution (positive differences ~ 10 µg m^{-3} Fig. 4g; 29% Table 2). In autumn, the differences in $PM_{2.5}$ concentrations at the two resolutions exhibit a marked east-west contrast, with lower values at the ~~global~~coarse resolution in western Europe (where the EMEP stations are generally located; Fig. 1) and higher values at the ~~global~~coarse resolution in eastern Europe (Fig. 4h). While $PM_{2.5}$ concentrations at the 25 EMEP site locations are on average lower when simulated at the ~~global~~coarse resolution (-23%), over all grid-cells, $PM_{2.5}$ concentrations are higher at the ~~global~~coarse resolution (8%). This highlights issues with representivity of the EMEP network across Europe, with much fewer EMEP measurement stations for $PM_{2.5}$ in eastern Europe.

The seasonality in $PM_{2.5}$ differences, brought about by a change in model horizontal resolution, can be partly explained by differences in PBL height between the two resolutions, as outlined in section 3.2. In particular, the deeper boundary layer in summer simulated at the ~~regional~~finer resolution may lead to greater vertical lofting from the surface, producing lower $PM_{2.5}$ levels compared to that simulated at the ~~global~~coarse resolution. In addition, differences in simulated precipitation (especially smaller-scale convective precipitation) between the two resolutions may be important, through its influence as the dominant mechanism in UKCA for removal of aerosols through wet deposition (O'Connor et al., 2014). Spatial patterns of convective precipitation differences between the two resolutions are shown in Section S2~~3~~, Fig. S4 of the Supplement to this article. In winter and spring, convective rainfall is higher at the ~~global~~coarse compared to the ~~regional~~fine resolution (Fig. S4a and S4b). Thus removal of $PM_{2.5}$ through wet deposition is greater, producing lower $PM_{2.5}$ concentrations at the coarser resolution (Fig. 4e and 4f). The opposite holds in summer and autumn as the convective rainfall is lower at the ~~global~~coarse compared to the ~~regional~~finer resolution (Fig. S4c and S4d) therefore resulting in higher $PM_{2.5}$ concentrations simulated at the ~~global~~coarse resolution (Fig. 4g and 4h).

Overall, we also find a large seasonal variation in simulated $PM_{2.5}$ concentrations between the two resolutions, with typically lower levels simulated in winter and spring at the ~~global~~coarse compared to the ~~regional~~finer resolution and the opposite result in summer and autumn. Hence, the seasonality of differences in simulated $PM_{2.5}$ concentrations between the

two model resolutions is generally the inverse of that found for O_3 in section 3.3. We find that these seasonal differences can be largely explained by meteorological effects: PBL height differences, especially in summer, and by differences in convective rainfall between the two resolutions.

4 Sensitivity of health impact estimates to model resolution

We now examine how the differences in O_3 and $PM_{2.5}$ concentrations simulated at the two resolutions, influence health impact estimations across Europe at the country level. For this analysis we use warm season daily maximum 8-hour running mean (MDA8) O_3 (above 70 µg m^{-3}) and annual-average $PM_{2.5}$ concentrations. To estimate health impacts, air pollution concentrations (with an averaging period consistent with that used in epidemiological studies) are combined with population estimates and concentration-response coefficients (Section 2.3).

4.1 Warm season MDA8 O_3 and annual-average $PM_{2.5}$ concentrations

Statistics for warm season MDA8 O_3 and annual $PM_{2.5}$ concentrations compared between EMEP measurements and model results at the two resolutions are provided in ~~Section S1,~~ Table 3~~S1 of the Supplement to this article~~. Mean simulated MDA8 O_3 levels in the warm season at the 52 EMEP locations for both resolutions, are higher compared to observations (NMB = 11% and 9 %; Table 3~~S1~~), in agreement with our findings for summer and autumn mean O_3 levels (c.f., Table 3~~S1~~, Table 1). The SD is also higher for both resolutions compared to observations. However, in contrast with summer mean O_3 levels, mean simulated MDA8 O_3 concentrations are 0.8~~1~~ % higher at the ~~global~~coarse compared to the ~~regional~~finer resolution at the 52 EMEP site locations (Table ~~S1~~3). Simulated annual mean $PM_{2.5}$ concentrations are also higher compared to observations at the 25 locations (NMB =~10-20%; Table ~~S1~~3) with concentrations being 8.7~~9~~% lower at the ~~global~~coarse compared to the ~~regional~~finer resolution. This represents the net effect of seasonality in NMB shown in Table 2.

Differences in warm season MDA8 O_3 and annual mean $PM_{2.5}$ concentrations, simulated at the ~~global~~coarse and ~~regional~~finer resolution, are shown in Fig. 5. The spatial distribution of differences in warm season MDA8 O_3 between the two resolutions (Fig. 5a) is most similar to the distribution of differences in summer mean O_3 concentrations (Fig. 3g). Differences in MDA8 O_3 concentrations range from ~ -7 µg m^{-3} in Northeast Europe to ~ +20 µg m^{-3} in Southern Europe, UK and Ireland (Fig. 5a). We note that if a different time-averaging period was chosen e.g., annual as opposed to warm season, the spatial patterns of MDA8 O_3 differences would alter considerably due to the seasonal variation displayed in Figure 3.

The spatial distribution of differences in annual mean $PM_{2.5}$ concentrations between the two resolutions (Fig. 5b) are most similar to the spatial distribution of differences in spring and especially autumn mean $PM_{2.5}$ concentrations notably with an east-west gradient (Fig. 5). Differences in $PM_{2.5}$ concentrations between the two resolutions range from ~ -8 µg m^{-3} in the southwestern part of Europe and Cyprus to ~ +4 µg m^{-3} in north and eastern Europe (Fig. 5b).

4.2 Effect of applying population-weighting to MDA8 O_3 and annual $PM_{2.5}$ concentrations

The warm season MDA8 O_3 concentrations and annual mean $PM_{2.5}$ concentrations, simulated at both resolutions, were weighted by population totals for each country to produce country average population-weighted concentrations (Section 2.3). Figure 6a shows the impact of the two resolutions on country-average warm season average MDA8 O_3 and the corresponding population-weighted MDA8 O_3 concentrations. Similarly differences in annual mean $PM_{2.5}$ concentrations between the two resolutions for non-population-weighted and population-weighted concentrations are shown in Fig. 6b.

Population-weighting of pollutant concentrations has different impacts across the European countries (Fig. 6a and 6b). In many countries, differences in population-weighted pollutant concentrations between the two resolutions are enhanced (i.e. larger positive or more negative differences) relative to non-population-weighted pollutant concentrations. However, in some countries population-weighting may reduce the positive or negative difference between the two resolutions. We examine several cases below.

For warm season MDA8 O_3 concentrations, the largest negative differences, implying lower MDA8 O_3 levels using ~~global~~coarse compared to the ~~regional~~finer resolution results, occur in eastern Europe (Fig. 5a). Hence, the largest negative differences in non-population-weighted and population-weighted MDA8 O_3 concentrations are found in eastern European countries (Fig. 6a). The difference between the two resolutions is greatest when population-weighting is applied. This is generally due to slightly lower population-weighted MDA8 O_3 concentrations compared to MDA8 O_3 concentrations derived from the ~~global~~coarse resolution results (Section S3, Fig. S5a of the Supplement to this article).

In the Netherlands warm season non-population-weighted MDA8 O_3 is also lower when derived from ~~global~~coarse compared to ~~regional~~finer resolution results (negative difference; Fig. 5a, 6a). However population-weighted MDA8 O_3 concentrations are higher when derived from the ~~global~~coarse resolution results (Fig. 6a). This is caused by lower MDA8 O_3 concentrations simulated at the ~~regional~~finer resolution when applying population-weighting (Fig. S5a). This suggests that in populated regions, MDA8 O_3 concentrations simulated at the ~~regional~~finer resolution are lower which might be linked to higher NO_2 concentrations.

Warm season MDA8 O_3 show the largest positive differences, with higher values simulated at the ~~global~~coarse resolution, for southern Europe and the UK/Ireland (Fig 5a). Thus, the largest positive differences for non-population-weighted and population-weighted MDA8 O_3 concentrations occurs in south European countries (Fig. 6a). Population–weighed MDA8 O_3 concentrations in Portugal are higher compared to MDA8 O_3 concentrations at the ~~global~~coarse but lower at the ~~regional~~finer resolution (Fig. S5a). This suggests that, at the ~~global~~coarse resolution, areas with high levels of O_3 are co-located with high population densities whilst at the ~~regional~~finer resolution areas with lower levels of O_3 are co-located with high population densities.

Annual-average $PM_{2.5}$ concentrations show the largest negative differences, with higher values simulated at the finer~~regional~~ resolution, in parts of western Europe (Fig. 5b). Hence, the largest negative non-population-weighted and population-weighted annual mean $PM_{2.5}$ concentrations are found for Cyprus, Italy and Spain (Fig. 6b). Conversely, higher annual-average $PM_{2.5}$

levels are simulated at the ~~global~~coarse resolution in eastern and northern Europe (Fig. 5b), hence larger positive non-population-weighted and population-weighted annual mean $PM_{2.5}$ concentrations occur for countries in eastern Europe and northern Europe (Fig. 6b).

In Cyprus, population-weighted annual mean $PM_{2.5}$ concentrations simulated at the ~~regional~~fine resolution are higher compared to concentrations with no population-weighting, due to denser populations being co-located with areas of higher $PM_{2.5}$ levels (Fig. S5b). In Croatia, population-weighted annual mean $PM_{2.5}$ concentrations simulated at the ~~global~~coarse resolution are greater than $PM_{2.5}$ concentrations with no population-weighted, again due to denser populations in regions of high concentrations but in this case when simulated at the ~~global~~coarse resolution (Fig. S5b). In a few countries (e.g. Switzerland), differences in population-weighted annual mean $PM_{2.5}$ concentrations between the two resolutions have an opposite sign to differences between concentration with no population-weighting (Fig. 6b). This indicates that annual mean $PM_{2.5}$ concentrations simulated at the ~~regional~~finer resolution are high in densely populated regions but are low in these same regions at the ~~global~~coarse resolution.

It would be insightful to examine these population-weighted results in relation to model-observation biases in densely populated areas. However, as outlined in Section 2.2, the available sites in the EMEP database are urban background stations which are required to be representative of a wide are and away from industrial areas (EMEP/CCC,2001). Nonetheless we do note that in southern Europe, simulated summer mean MDA8 O_3 concentrations at the finer resolution are closer to observations than concentrations simulated at the coarse resolution. We find no consistent result for model biases in simulated annual mean $PM_{2.5}$ concentrations with respect to observations for the two model resolutions.

4.3 Attributable fraction of mortality associated with long-term exposure to O_3

The Attributable Fraction (AF) associated with long-term exposure to MDA8 O_3, expressed as a percentage of total respiratory mortality and simulated at both resolutions, was calculated for each country (Fig. 7a), using the population-weighted warm season MDA8 O_3 concentrations (Fig. 6a) as discussed in Section 2.3. For both resolutions, the estimated AF is shown for each country, with the 95% confidence interval (95% C.I.) representing uncertainties associated only with the concentration-response coefficient (β) used (shown in grey). For all the countries considered, irrespective of the model resolution used, the AF of total respiratory mortality ranges from 1% (95% C.I. 0% - 2%) in Finland to 11 % (95% C.I. 4% - 18%) in Cyprus (Fig. 7a).

Differences in AF between the countries are solely attributed to differences in population-weighted MDA8 O_3 concentrations. Thus, countries with the highest population-weighted concentrations also have the highest AF. Similarly countries with the highest differences in population-weighted MDA8 O_3 concentrations between the two resolutions also have the largest differences in AF between the ~~global~~coarse and ~~regional~~finer resolution. If the AF was calculated for each model grid-cell rather than at the country level, the differences in AF for the two pollutants would have identical spatial distributions

to the differences in warm season MDA8 O_3 and annual-mean $PM_{2.5}$ concentrations depicted in Fig. 5, as the AF is only dependent on the pollutant concentration and β (which is constant across all countries).

The differences in country level AF associated with long-term exposure to warm season MDA8 O_3, simulated at the two resolutions, are shown in Fig. 7b. These values highlight the sensitivity of respiratory mortality attributable to long-term exposure to O_3 to a change in model resolution.

For most of northern and eastern Europe, the AF at the ~~global~~coarse resolution is lower than that at the ~~regional~~finer resolution (negative differences; Fig. 7b) as for differences in population-weighted warm season MDA8 O_3 concentrations in the same countries (Fig. 6a). In contrast, the AF at the ~~global~~coarse resolution is higher than that at the ~~regional~~finer resolution for countries in southern Europe (positive differences, Fig. 7b). Differences in AF range from -0.9% (95% C.I. -0.3% to -1.5%) in Poland to +2.6% (95% C.I. 1.0% to 4.1%) in Portugal (Fig. 7b) which directly correspond to the countries having the lowest and highest difference in population-weighted MDA8 O_3 concentration respectively (Fig. 6a; Note, although the differences in AF between the two resolution appear ~~to be~~ low, these are percentages of total respiratory mortality). In Poland and Portugal the estimated AF at the finer resolution is 1.4 times and 0.7 times that estimated at the coarse resolution. For approximately half of the European countries, the AF is higher for the ~~global~~coarse resolution compared to the ~~regional~~finer resolution and vice versa. When considering the uncertainty associated with the concentration-response coefficient used, the sign of the difference of AF between the two model resolutions is unaltered (Fig. 7b). Over the majority of the countries, the AF attributable to long-term exposure to MDA8 O_3 by the coarse resolution fall within the range of uncertainty as calculated by the finer resolution (Fig. 7a). However, over Finland and Ireland, the coarse mean estimates fall outside the uncertainty range estimates using the finer resolution (Fig. 7a).

~~The uncertainty associated with the concentration-response coefficient used does not alter the sign of the difference of AF between the two model resolutions.~~ For U.S. averaged mortality estimates, Punger and West (2013) show that mortality estimates related to warm season long-term O_3 exposure, calculated using the O_3 concentrations at 36 km, were higher (by 12%) than estimates calculated at the 12 km resolution. Resolution was also found to play and important role in determining health benefits associated with differences in O_3 between 2005 and 2014 in the U.S. (Thompson et al. 2014). In particular, in urban areas, Thompson et al. (2014) estimate that the benefits calculated using coarse resolution results were on average two times greater than estimates calculated using the finer scale results. Both the studies mentioned are conducted in the U.S. and use a different concentration response coefficient and thus a definitive comparison between these studies and our estimates over Europe is not possible.

Since, seasonal differences in simulated O_3 with resolution are considerable, the AF associated with long-term exposure to O_3 was also calculated based on annual-mean (as opposed to summer-mean) O_3 concentrations based on recommendations by Turner et al. (2015). Turner et al (2015) suggest a higher concentration response coefficient of -1.06 (95% CI: 1.04-1.08) per 10 μg m^{-3} -and a slight lower MDA8 O_3 threshold of 53.4 μg m^{-3} compared to values used in our study for

summer-mean MDA8 O_3. Using the values from Turner et al. (2015) the differences in AF are found to be of the same sign for the majority of the countries and the rankings across countries are largely similar. This similarity occurs because the difference in annual-mean MDA8 O_3 concentrations between the two resolutions shows generally similar spatial patterns to the differences –in warm season MDA8 O_3 concentrations (not shown). However the ranges when using annual-mean O_3 concentrations and recommendations from Turner et al. (2015) are larger: -2.3% to +12.0%, compared to AF ranges given above for MDA8 O_3. From further sensitivity analyses it is found that these greater AF ranges can be attributed to the use of a higher concentration-response coefficient (by a factor of approximately 4) rather than differences in annual-mean compared to summer-mean concentrations.

4.4 Attributable Fraction associated with long-term exposure to $PM_{2.5}$

The fraction of all-cause (excluding external) mortality attributable to long-term exposure to $PM_{2.5}$, is shown as percentages for each country in Fig. 8a. The AF for all countries, irrespective of the resolution used, ranges from 2% (95% C.I. 1% - 3%) in Iceland to 15% (95% C.I. 10% - 19%) in Cyprus (Fig. 8a). Differences in AF between the two resolutions are shown in Fig. 8b. Since the variability in AF differences across the countries is caused by variability in population-weighted annual mean PM2.5 differences, Cyprus and countries in parts of western Europe have the largest negative difference in percentage AF (Fig. 8b). In contrast, countries in eastern and northern Europe have the largest positive difference in percentage AF (Fig. 8b). These differences range from -4.7% (95% C.I. -6.1% to -3.2%) in Cyprus to 2.8% (95% C.I. 1.9% to 3.7%) in Croatia. For Cyprus and Croatia, using the finer resolution results in an estimated AF that is 1.5 and 0.7 times that estimated using the coarse resolution. Over most countries, annual mean population-weighted $PM_{2.5}$ concentrations are higher (positive difference; Fig. 6b) for the ~~global~~coarse compared to the ~~regional~~finer resolution, thus resulting in a higher AF when using the ~~global~~coarse resolution results. Note, similar to O_3, the uncertainty associated with the concentration-response coefficient for $PM_{2.5}$ does not alter the sign of the difference of AF between the two model resolutions (Fig. 8b). For a number of countries, the mean AF attributable to long-term exposure to $PM_{2.5}$ using the coarse resolution falls outside the uncertainty range of the finer estimates in particular over Iceland and Ireland (Fig. 8a).

We also examine the impact of using a low-concentration threshold. We apply a threshold of 5.8 μg m^{-3} (suggested by ~~(~~Burnett et al. (2014) which is derived from Lim et al. (2012)) to annual mean $PM_{2.5}$ concentrations. Differences in AF estimates associated with long-term exposure to population-weighted $PM_{2.5}$ concentrations range from -4.8% to +2.1% (as compared to -4.7% to +2.8% above when no threshold is applied). The spatial distribution of these estimates remains unchanged and only slight changes in country rankings occur. Hence, the impact of applying a low concentrations threshold in this study for Europe is small.

Our results are consistent with other studies, but not all, that examine the impact of model resolution on health estimates associated with long-term exposure to $PM_{2.5}$. Using concentrations simulated at the 36 km resolution, Punger and

West (2013) find that the U.S. national health estimate is higher (11%) than the estimate at 12 km resolution. Li et al. (2015) also show that averaged over the US, a coarse grid resolution (~ 200 km) results in a health estimate that is lower (8%) than the estimated based on the fine scale model results (~ 50 km), in contrast to our findings averaged across Europe. In contrast, Thompson et al. (2014) find that health benefits associated with changes in $PM_{2.5}$ concentrations between 2005 and 2014 in the U.S., were not sensitive to resolution. Both Punger and West (2013) and Li et al. (2015) find that differences in $PM_{2.5}$ are mainly attributable to primary anthropogenic PM, while Thompson et al. (2014) attribute the greatest differences (between 36 km and 4 km resolutions) to secondary PM. However, in our study no substantial differences in $PM_{2.5}$ components between the two resolutions were found. All the~~se~~ mentioned studies are conducted in the U.S. and hence definitive comparisons cannot be made with our results for Europe

In summary, our results suggest that differences in AF health estimates between ~~global~~coarse and finer~~regional~~ resolutions vary across the different European countries with clear differences between southern and eastern Europe for exposure to warm season MDA8 O_3 and west-east differences for exposure to annual-average $PM_{2.5}$ due to the dependence of AF on populated weighted MDA8 O_3 and annual $PM_{2.5}$ concentrations.

For differences in AF attributable to long-term exposure to summer mean MDA8 O_3 and annual mean $PM_{2.5}$ concentrations, the uncertainty associated with the concentration-response coefficient used does not alter the sign of the difference of AF between the two model resolutions (Fig. 7b and 8b). The uncertainty ranges for the $PM_{2.5}$ –related estimates show a greater variability between the two resolutions for more countries compared to MDA8 O_3-related AF estimates. Using the concentration-response coefficient in ~~(~~Jerrett et al.~~,~~ (2009), ~~(~~Thompson et al.~~,~~ (2014) find that the avoided mortalities due to difference in ozone concentrations between 2005 and 2014 at a 36 km model resolution are within the 95% uncertainty range associated with the concentration-response coefficient used compared to estimates at a resolution of 12 km and 4 km. These authors also find avoided mortalities associated with long-term effects of $PM_{2.5}$ exposure at 36 km to fall within estimates at the 12 km and 4 km resolution for three different concentration-response coefficients. Thus our results are in agreement for summer mean O_3 but less for annual mean $PM_{2.5}$.

5 Conclusions

Chemistry-climate model simulations were performed at two resolutions: a ~~global~~coarse resolution (~ 140 km) and a ~~regional~~finer resolution (~ 50 km) over Europe to quantify the impact of horizontal model resolution on simulated O_3 and $PM_{2.5}$ concentrations by season; and on the associated Attributable Fraction (AF) of mortality due to long-term exposure to these two pollutants. Simulated O_3 concentrations are lower in winter and higher in summer and autumn compared to measurements at both model resolutions. Results show a ~~strong~~ seasonal influence in the mean O_3 differences between the two resolutions. Simulated O_3 concentrations averaged across Europe at the ~~global~~coarse resolution are higher in winter and spring (10% and 6%, respectively), and lower in summer and autumn (-1% and -4%, respectively) compared to the ~~regional~~finer

resolution. In contrast during winter and spring, NO_2 concentrations are lower in some areas at the ~~global~~coarse compared to the ~~regional~~finer configuration, whilst in summer and autumn, there are more locations where NO_2 concentrations are higher at the ~~global~~coarse resolution. The lower O_3 concentrations simulated at the ~~regional~~finer compared to the ~~global~~coarse resolution can be partly explained by these higher NO_2 levels that enhance titration of O_3 at this finer resolution. The PBL height also differs between the two resolutions and may also account for differences in O_3 concentrations; however, it is not possible to clearly separate the effects of chemistry and mixing on simulated O_3.

Differences in $PM_{2.5}$ concentrations simulated at the two resolutions also vary seasonally. Modelled $PM_{2.5}$ concentrations are lower in winter and higher in summer compared to measurements at both resolutions. Simulated seasonal mean $PM_{2.5}$ concentrations averaged across Europe during winter and spring are lower at the ~~global~~coarse compared to the ~~regional~~finer resolution (-8% and -~~6~~27%, respectively) but higher in summer and autumn (29% and 8%, respectively). This seasonality in Europe-average differences in $PM_{2.5}$ concentrations is opposite to that found for differences in O_3 concentrations between the two resolutions. Differences in $PM_{2.5}$ concentrations simulated at the two resolutions are also influenced by PBL height, especially in summer when a deeper boundary layer at the ~~regional~~finer resolution leads to greater lofting and lower $PM_{2.5}$ concentrations. Furthermore, in all seasons, the differences in $PM_{2.5}$ levels between the two resolutions are closely related to differences in the convective rainfall rate. In winter and spring, the convective rainfall at the ~~global~~coarse resolution is higher than that at the ~~regional~~finer resolution thus resulting in lower $PM_{2.5}$ concentrations. The opposite result holds in summer and autumn.

Results show that differences in warm season mean MDA8 O_3 concentrations between the two resolutions are similar to summer mean differences in simulated O_3 concentrations, with spatial patterns of differences reveal clear and important contrasts. Warm season MDA8 O_3 levels are higher in most of southern Europe as well as the UK and Ireland, but lower in other areas of northern as well as eastern Europe when simulated at the ~~global~~coarse resolution compared to the ~~regional~~finer resolution. On the other hand, annual average $PM_{2.5}$ concentrations are higher across most of northern and eastern Europe but lower over parts of southwest Europe at the ~~global~~coarse compared to the ~~regional~~finer resolution.

Weighting the pollutant concentrations at both resolutions with the population within each country, results in some added differences between concentrations at the two resolutions which also vary across the countries. In many countries, weighting by population enhances either positive or negative differences in warm season MDA8 O_3 or annual mean $PM_{2.5}$ concentrations between the two resolution, which suggests that high levels of pollutant concentrations coincide with high population density at one resolution but low pollutant concentrations are co-located with high population density at the other resolution. Population-weighting pollutant concentrations also reduces differences between ~~global~~coarse and ~~regional~~finer resolution results in some countries.

The AF of respiratory mortality associated with long-term exposure to warm season MDA8 O_3 and annual mean $PM_{2.5}$ is also sensitive to resolution as is it is solely driven by the simulated population-weighted pollutant concentrations. For the AF associated with long-term exposure to O_3, countries in northern as well as eastern Europe have lower AF values at the ~~global~~coarse compared to the ~~regional~~finer resolution whilst the opposite result occurs for other countries in southern Europe

and Ireland. For the AF associated with long-term exposure to $PM_{2.5}$, a few countries in southwestern Europe and Cyprus have lower AF values for $PM_{2.5}$ concentrations simulated at the ~~global~~coarse resolution whilst more countries especially in eastern and northern Europe show a higher AF when using $PM_{2.5}$ concentrations simulated at the ~~global~~coarse resolution.

Overall, differences in the country-average AF associated with long term exposure to MDA8 O_3 range between -0.9 % and +2.6 % while differences in the AF associated with long-term exposure to annual mean $PM_{2.5}$ range from -4.7% to +2.8 % of the total baseline mortality. This result emphasizes the importance of model horizontal resolution when conducting country specific health impact studies. We also find~~note~~ that the impacts of a 95% C.I. in concentration-response coefficient is smaller than the impact of the model horizontal resolution. In addition, these ranges in AF associated with long-term exposure to annual mean $PM_{2.5}$ were largely unaltered with the application of a low-concentration threshold for $PM_{2.5}$.

Our calculation for O_3 health impacts only considers warm-season MDA8 O_3 impacts however these may differ to annual MDA8 O_3 impacts because of seasonal differences in simulated O_3 with resolution highlighted in this study. When using annual-mean MDA8 O_3 concentrations alongside a recommended concentration-response coefficient and threshold suggested by Turner et al. (2015) the difference in AF between the two resolutions is considerably larger than our estimates using summer-mean MDA8 O_3 concentrations. This is driven by the higher concentration-response coefficient (by a factor of approximately 4) quoted in Turner et al. (2015) compared to that suggested by HRAPIE for summer mean MDA8 O_3 concentrations (WHO, 2013). In addition, for our study we apply the same concentration-response coefficient to all populations and assumed that for $PM_{2.5}$-related health impacts, all $PM_{2.5}$ components have the same impact on mortality.

The pollutant concentrations used in this study have been extracted at the lowest model level with a mid-point at 20 m. The sensitivity of our simulated pollutant concentrations to vertical model resolution has not been examined. Future research focusing on the sensitivity of AF changes to different averaging periods or seasons would be beneficial. In addition, the use of concentration-response coefficients that are derived from European cohort data would be useful, although such data are limited. Nonetheless this study provides one of the first insights as to how air pollution related health impacts over Europe are influenced by the model resolution used to simulate pollutant concentrations.

Acknowledgements

Sara Fenech's PhD was funded by Public Health England. The development of the United Kingdom Chemistry and Aerosol (UKCA) model and Fiona M. O'Connor are supported by the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101).

References

Anenberg, S. C., West, J. J., Fiore, A. M., Jaffe, D. A., Prather, M., Bergmann, D., Cuvelier, K. and Dentener, F. J.: Intercontinental Impacts of Ozone Pollution on Human Mortality, Environ. Sci. Technol., 43(17), 6482–6487, doi:10.1029/2008GM000741/summary, 2009.

Formatted: Indent: Left: 0.3 cm, Hanging: 1.27 cm

Anenberg, S. C., Horowitz, L. W., Tong, D. Q. and West, J. J.: An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using atmospheric modeling, Environ. Health Perspect., 118(9), 1189–1195, doi:10.1289/ehp.0901220, 2010.

Arunachalam, S., Wang, B., Davis, N., Baek, B. H. and Levy, J. I.: Effect of chemistry-transport model scale and resolution on population exposure to PM2.5 from aircraft emissions during landing and takeoff, Atmos. Environ., 45(19), 3294–3300, doi:10.1016/j.atmosenv.2011.03.029, 2011.

Bellouin, N., Rae, J., Jones, A., Johnson, C., Haywood, J. and Boucher, O.: Aerosol forcing in the Climate Model Intercomparison Project (CMIP5) simulations by HadGEM2-ES and the role of ammonium nitrate, J. Geophys. Res. Atmos., 116(20), 1–25, doi:10.1029/2011JD016074, 2011.

Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. a., Peters, A., Siscovick, D., Smith, S. C., Whitsel, L. and Kaufman, J. D.: Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the american heart association, Circulation, 121(21), 2331–2378, doi:10.1161/CIR.0b013e3181dbece1, 2010.

Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J. and Shelly, A.: Unified modeling and prediction of weather and climate: A 25-year journey, Bull. Am. Meteorol. Soc., 93(12), 1865–1877, doi:10.1175/BAMS-D-12-00018.1, 2012.

Burnett, R. T., Arden Pope, C., Ezzati, M., Olives, C., Lim, S. S., Mehta, S., Shin, H. H., Singh, G., Hubbell, B., Brauer, M., Ross Anderson, H., Smith, K. R., Balmes, J. R., Bruce, N. G., Kan, H., Laden, F., Prüss-Ustün, A., Turner, M. C., Gapstur, S. M., Diver, W. R. and Cohen, A.: An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure, Environ. Health Perspect., 122(4), 397–403, doi:10.1289/ehp.1307049, 2014.

Cohen, A. J., Anderson, H. R., Ostro, B., Pandey, K. D., Krzyzanowski, M., Künzli, N., Gutschmidt, K., Iii, C. A. P., Romieu, I., Samet, J. M. and Smith, K. R.: Urban Air Pollution, in Urban air pollution. In: Comparative quantification of health risks: global and regional burden of disease due to selected major risk factors., edited by M. Ezzati, A. Lopez, A. Rodgers, and C. Murray, pp. 1353–1434, World Health Organization, Geneva., 2004.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C. A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C. J. L. and Forouzanfar, M. H.: Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, Lancet, 389(10082), 1907–1918, doi:10.1016/S0140-6736(17)30505-6, 2017.

Coleman, L., Varghese, S., Tripathi, O. P., Jennings, S. G. and O'Dowd, C. D.: Regional-Scale Ozone Deposition to North-East Atlantic Waters, Adv. Meteorol., 2010, 1–16, doi:10.1155/2010/243701, 2010.

Colette, A., Bessagnet, B., Vautard, R., Szopa, S., Rao, S., Schucht, S., Klimont, Z., Menut, L., Clain, G., Meleux, F., Curci, G. and Rouïl, L.: European atmosphere in 2050, a regional air quality and climate perspective under CMIP5 scenarios, Atmos. Chem. Phys., 13(15), 7451–7471, doi:10.5194/acp-13-7451-2013, 2013.

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A. and Woodward, S.: Development and evaluation of an Earth-system model – HadGEM2, Geosci. Model Dev. Discuss., 4(2), 997–1062, doi:10.5194/gmdd-4-997-2011, 2011.

COMEAP: Quantification of Mortality and Hospital Admissions Associated with Ground- level Ozone., 2015.

EMEP/CCC: EMEP Manual for Sampling and Analysis, Norwegion Inst. Air Res. [online] Available from: https://www.nilu.no/projects/ccc/manual/index.html (Accessed 14 March 2018), 2001.

Fang, Y., Naik, V., Horowitz, L. W. and Mauzerall, D. L.: Air pollution and associated human mortality: The role of air pollutant emissions, climate change and methane concentration increases from the preindustrial period to present, Atmos. Chem. Phys., 13(3), 1377–1394, doi:10.5194/acp-13-1377-2013, 2013.

Folberth, G. A., Abraham, N. L., Johnson, C. E., Morgenstern, O., O'Connor, F. M., Pacifico, F., Young, P. A., Collins, W. J. and Pyle, J. a.: Evaluation of the new UKCA climate-composition model. Part III. Extension to Tropospheric Chemistry and Biogeochemical Coupling between Atmosphere and Biosphere, In prep.

Forouzanfar, M. H., Afshin, A., Alexander, L. T., Biryukov, S., Brauer, M., Cercy, K., Charlson, F. J., Cohen, A. J., Dandona, L., Estep, K., Ferrari, A. J., Frostad, J. J., Fullman, N., Godwin, W. W., Griswold, M., Hay, S. I., Kyu, H. H., Larson, H. J., Lim, S. S., Liu, P. Y., Lopez, A. D., Lozano, R., Marczak, L., Mokdad, A. H., Moradi-Lakeh, M., Naghavi, M., Reitsma, M. B., Roth, G. A., Sur, P. J., Vos, T., Wagner, J. A., Wang, H., Zhao, Y., Zhou, M., Barber, R. M., Bell, B., Blore, J. D., Casey, D. C., Coates, M. M., Cooperrider, K., Cornaby, L., Dicker, D., Erskine, H. E., Fleming, T., Foreman, K., Gakidou, E., Haagsma, J. A., Johnson, C. O., Kemmer, L., Ku, T., Leung, J., Masiye, F., Millear, A., Mirarefin, M., Misganaw, A., Mullany, E., Mumford, J. E., Ng, M., Olsen, H., Rao, P., Reinig, N., Roman, Y., Sandar, L., Santomauro, D. F., Slepak, E. L., Sorensen, R. J. D., Thomas, B. A., Vollset, S. E., Whiteford, H. A., Zipkin, B., Murray, C. J. L., Mock, C. N., Anderson, B. O., Futran, N. D., Anderson, H. R., Bhutta, Z. A., Nisar, M. I., Akseer, N., Krueger, H., Gotay, C. C., Kissoon, N., Kopec, J. A., Pourmalek, F., Burnett, R., Abajobir, A. A., Knibbs, L. D., Veerman, J. L., Lalloo, R., Scott, J. G., Alam, N. K. M., Gouda, H. N., Guo, Y., McGrath, J. J., Charlson, F. J., Erskine, H. E., Jeemon, P., Dandona, R., Goenka, S., Kumar, G. A., et al.: Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015, Lancet, 388(10053), 1659–1724, doi:10.1016/S0140-6736(16)31679-8, 2016.

Gowers, a M., Miller, B. G. and Stedman, J. R.: Estimating Local Mortality Burdens associated with Particulate Air Pollution. [online] Available from: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317141074607, 2014.

Gryparis, A., Forsberg, B., Katsouyanni, K., Analitis, A., Touloumi, G., Schwartz, J., Samoli, E., Medina, S., Anderson, H. R., Niciu, E. M., Wichmann, H. E., Kriz, B., Kosnik, M., Skorkovsky, J., Vonk, J. M. and Dörtbudak, Z.: Acute effects of ozone on mortality from the "Air Pollution and Health: A European Approach" project, Am. J. Respir. Crit. Care Med., 170(10), 1080–1087, doi:10.1164/rccm.200403-333OC, 2004.

Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger, L., Lerdau, M., McKay, W. A., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R., Taylor, J. and Zimmerman, P.: A global model of natural volatile organic compound emissions, J. Geophys. Res., 100(D5), 8873, doi:10.1029/94JD02950, 1995.

Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B. and Kaufman, J. D.: Long-term air pollution exposure and cardio- respiratory mortality: a review, Environ. Heal., 12(1), 43, doi:10.1186/1476-069X-12-43, 2013.

Jerrett, M., Burnett, R. T., Pope, C. A., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E. and Thun, M.: Long-Term Ozone Exposure and Mortality, N. Engl. J. Med., 360(11), 1085–1095, doi:10.1056/NEJMoa0803894, 2009.

Jones, A., Roberts, D. L., Woodage, M. J. and Johnson, C. E.: Indirect sulphate aerosol forcing in a climate model with an interactive sulphur cycle, J. Geophys. Res., 106, 20293–20310, 2001.

Krewski, D.: Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality., 2009.

Lamarque, J. F., Bond, T. C., Eyring, V., Granier, C., Heil, a., Klimont, Z., Lee, D., Liousse, C., Mieville, a., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K. and Van Vuuren, D. P.: Historical (1850-2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application, Atmos. Chem. Phys., 10(15), 7017–7039, doi:10.5194/acp-10-7017-2010, 2010.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D. and Pozzer, A.: The contribution of outdoor air pollution sources to premature mortality on a global scale., Nature, 525, 367–71, doi:10.1038/nature15371, 2015.

Li, Y., Henze, D. K., Jack, D. and Kinney, P. L.: The influence of air quality model resolution on health impact assessment for fine particulate matter and its components, Air Qual. Atmos. Heal., 9(1), 51–68, doi:10.1007/s11869-015-0321-z, 2015.

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., Blore, J. D., Blyth, F., Bonner, C., Borges, G., Bourne, R., Boussinesq, M., Brauer, M., Brooks, P., Bruce, N. G., Brunekreef, B., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Bull, F., Burnett, R. T., Byers, T. E., Calabria, B., Carapetis, J., Carnahan, E., Chafe, Z., Charlson, F., Chen, H., Chen, J. S., Cheng, A. T. A., Child, J. C., Cohen, A., Colson, K. E., Cowie, B. C., Darby, S., Darling, S., Davis, A., Degenhardt, L., Dentener, F., Des Jarlais, D. C., Devries, K., Dherani, M., Ding, E. L., Dorsey, E. R., Driscoll, T., Edmond, K., Ali, S. E., Engell, R. E., Erwin, P. J., Fahimi, S., Falder, G., Farzadfar, F., Ferrari, A., Finucane, M. M., Flaxman, S., Fowkes, F. G. R., Freedman, G., Freeman, M. K., Gakidou, E., Ghosh, S., Giovannucci, E., Gmel, G., Graham, K., Grainger, R., Grant, B., Gunnell, D., Gutierrez, H. R., Hall, W., Hoek, H. W., Hogan, A., Hosgood, H. D., Hoy, D., Hu, H., Hubbell, B. J., Hutchings, S. J., Ibeanusi, S. E., Jacklyn, G. L., Jasrasaria, R., Jonas, J. B., Kan, H., Kanis, J. A., Kassebaum, N., Kawakami, N., Khang, Y. H., Khatibzadeh, S., Khoo, J. P., Kok, C., et al.: A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010, Lancet, 380(9859), 2224–2260, doi:10.1016/S0140-6736(12)61766-8, 2012.

Lock, A. P., Brown, A. R., Bush, M. R., Martin, G. M. and Smith, R. N. B.: A New Boundary Layer Mixing Scheme. Part I: Scheme Description and Single-Column Model Tests, Mon. Weather Rev., 128(9), 3187–3199, doi:10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2, 2000.

Malley, C. S., Henze, D. K., Kuylenstierna, J. C. I., Vallack, H. W., Davila, Y., Anenberg, S. C., Turner, M. C. and Ashmore, M. R.: Updated global estimates of respiratory mortality in adults ≥ 30 years of age attributable to long-term ozone exposure, Environ. Health Perspect., 125(8), 1–9, doi:10.1289/EHP1390, 2017.

Markakis, K., Valari, M., Perrussel, O., Sanchez, O. and Honore, C.: Climate-forced air-quality modeling at the urban scale: sensitivity to model resolution, emissions and meteorology, Atmos. Chem. Phys., 15(13), 7703–7723, doi:10.5194/acp-15-7703-2015, 2015.

Morgenstern, O., Braesicke, P., O'Connor, F. M., Bushell, A. C., Johnson, C. E., Osprey, S. M. and Pyle, J. A.: Evaluation of the new UKCA climate-composition model – Part 1: The stratosphere, Geosci. Model Dev., 2(1), 43–57, doi:10.5194/gmd-2-43-2009, 2009.

Moufouma-Okia, W. and Jones, R.: Resolution dependence in simulating the African hydroclimate with the HadGEM3-RA regional climate model, Clim. Dyn., 44(3–4), 609–632, doi:10.1007/s00382-014-2322-2, 2014.

Neal, L. S., Dalvi, M., Folberth, G., McInnes, R. N., Agnew, P., O'Connor, F. M., Savage, N. H. and Tilbee, M.: A description and evaluation of an air quality model nested within global and regional composition-climate models using MetUM, Geosci. Model Dev., 10, 3941–3962, doi:10.5194/gmd-2017-73, 2017.

O'Connor, F. M., Johnson, C. E., Morgenstern, O., Abraham, N. L., Braesicke, P., Dalvi, M., Folberth, G. a., Sanderson, M. G., Telford, P. J., Voulgarakis, A., Young, P. J., Zeng, G., Collins, W. J. and Pyle, J. A.: Evaluation of the new UKCA climate-composition model-Part 2: The troposphere, Geosci. Model Dev., 7, 41–91, doi:10.5194/gmd-7-41-2014, 2014.

Pacifico, F., Harrison, S. P., Jones, C. D., Arneth, A., Sitch, S., Weedon, G. P., Barkley, M. P., Palmer, P. I., Serça, D., Potosnak, M., Fu, T. M., Goldstein, A., Bai, J. and Schurgers, G.: Evaluation of a photosynthesis-based biogenic isoprene emission scheme in JULES and simulation of isoprene emissions under present-day climate conditions,

Atmos. Chem. Phys., 11(9), 4371–4389, doi:10.5194/acp-11-4371-2011, 2011.

Punger, E. M. and West, J. J.: The effect of grid resolution on estimates of the burden of ozone and fine particulate matter on premature mortality in the United States., Air Qual. Atmos. Health, 6(3), 563–573, doi:10.1007/s11869-013-0197-8, 2013.

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., Rouïl, 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, doi:10.1016/j.atmosenv.2015.04.003, 2015.

Silva, R. A., West, J. J., Zhang, Y., Anenberg, S. C., Lamarque, J.-F., Shindell, D. T., Collins, W. J., Dalsoren, S., Faluvegi, G., Folberth, G., Horowitz, L. W., Nagashima, T., Naik, V., Rumbold, S., Skeie, R., Sudo, K., Takemura, T., Bergmann, D., Cameron-Smith, P., Cionni, I., Doherty, R. M., Eyring, V., Josse, B., MacKenzie, I. a, Plummer, D., Righi, M., Stevenson, D. S., Strode, S., Szopa, S. and Zeng, G.: Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change, Environ. Res. Lett., 8, 1748–9326, doi:10.1088/1748-9326/8/3/034005, 2013.

Stock, Z. S., Russo, M. R. and Pyle, J. A.: Representing ozone extremes in European megacities: the importance of resolution in a global chemistry climate model, Atmos. Chem. Phys., 14, 3899–3912, doi:10.5194/acp-14-3899-2014, 2014.

Thompson, T. M. and Selin, N. E.: Influence of air quality model resolution on uncertainty associated with health impacts, Atmos. Chem. Phys., 12(20), 9753–9762, doi:10.5194/acp-12-9753-2012, 2012.

Thompson, T. M., Saari, R. K. and Selin, N. E.: Air quality resolution for health impact assessment: influence of regional characteristics, Atmos. Chem. Phys., 14, 969–978, doi:10.5194/acp-14-969-2014, 2014.

Tie, X., Brasseur, G. and Ying, Z.: Impact of model resolution on chemical ozone formation in Mexico City: application of the WRF-Chem model, Atmos. Chem. Phys., 10(18), 8983–8995, doi:10.5194/acp-10-8983-2010, 2010.

Tørseth, K., Aas, W., Breivik, K., Fjeraa, A. M., Fiebig, M., Hjellbrekke, A. G., Lund Myhre, C., Solberg, S. and Yttri, K. E.: Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972-2009, Atmos. Chem. Phys., 12(12), 5447–5481, doi:10.5194/acp-12-5447-2012, 2012.

Turner, M. C., Jerrett, M., Pope, C. A., Krewski, D., Gapstur, S. M., Diver, W. R., Beckerman, B. S., Marshall, J. D., Su, J., Crouse, D. L. and Burnett, R. T.: Long-Term Ozone Exposure and Mortality in a Large Prospective Study, Am. J. Respir. Crit. Care Med., 193(10), 1134–1142, doi:10.1164/rccm.201508-1633OC, 2015.

Valari, M. and Menut, L.: Does an Increase in Air Quality Models' Resolution Bring Surface Ozone Concentrations Closer to Reality?, J. Atmos. Ocean. Technol., 25(11), 1955–1968, doi:10.1175/2008JTECHA1123.1, 2008.

Walters, D. N., Best, M. J., Bushell, A. C., Copsey, D., Edwards, J. M., Falloon, P. D., Harris, C. M., Lock, A. P., Manners, J. C., Morcrette, C. J., Roberts, M. J., Stratton, R. A., Webster, S., Wilkinson, J. M., Willett, M. R., Boutle, I. A., Earnshaw, P. D., Hill, P. G., MacLachlan, C., Martin, G. M., Moufouma-Okia, W., Palmer, M. D., Petch, J. C., Rooney, G. G., Scaife, A. A. and Williams, K. D.: The Met Office Unified Model Global Atmosphere 3.0/3.1 and JULES Global Land 3.0/3.1 configurations, Geosci. Model Dev., 4(4), 919–941, doi:10.5194/gmd-4-919-2011, 2011.

West, J. J., Naik, V., Horowitz, L. W. and Fiore, a. M.: Effect of regional precursor emission controls on long-range ozone transport – Part 2: steady-state changes in ozone air quality and impacts on human mortality, Atmos. Chem. Phys. Discuss., 9, 6095–6107, doi:10.5194/acpd-9-7079-2009, 2009.

WHO: Health Risks of Air Pollution in Europe – HRAPIE project: Recommendations for concentration-response functions for cost-benefit analysis of particulate matter, ozone and nitrogen dioxide. [online] Available from:

http://www.euro.who.int/__data/assets/pdf_file/0006/238956/Health-risks-of-air-pollution-in-Europe-HRAPIE-project,-Recommendations-for-concentrationresponse-functions-for-costbenefit-analysis-of-particulate-matter,-ozone-and-nitrogen-dioxide., 2013.

Yu, K., Jacob, D. J., Fisher, J. a., Kim, P. S., Marais, E. a., Miller, C. C., Travis, K. R., Zhu, L., Yantosca, R. M., Sulprizio, M. P., Cohen, R. C., Dibb, J. E., Fried, A., Mikoviny, T., Ryerson, T. B., Wennberg, P. O. and Wisthaler, A.: Sensitivity to grid resolution in the ability of a chemical transport model to simulate observed oxidant chemistry under high-isoprene conditions, Atmos. Chem. Phys., 16(7), 4369–4378, doi:10.5194/acp-16-4369-2016, 2016.

Table 1: Statistical results comparing seasonal mean O_3 concentrations simulated at the ~~global~~coarse and ~~regional~~finer resolutions to observations from 52 stations within the EMEP network in 2007. Statistical results for all model grid-cells of both resolutions are also shown. Percentage differences between the two model resolutions are calculated as (O_3 ~~global~~coarse resolution $-O_3$ ~~regional~~finer resolution)/(O_3 ~~global~~coarse resolution).

			52 sites		all grid-cells	
Season		**Obs.**	**Model**		**Model**	
			140 km	**50 km**	**140 km**	**50 km**
DJF	Mean (µg m^{-3})	52.8	48.5	42.6	35.1	31.7
	Difference in model mean (%)		12.2		9.7	
	NMB (%)		-8.1	-19.2		
	SD (µg m^{-3})	11.0	17.0	16.0	17.3	16.5
	~~Difference (%)~~		~~12~~		~~10~~	
MAM	Mean (µg m^{-3})	70.4	80.7	67.9	75.7	71.5
	Difference in model mean (%)		15.9		5.5	
	NMB (%)		14.6	-3.6		
	SD (µg m^{-3})	8.9	13.7	12.8	12.9	12.9
	~~Difference (%)~~		~~16~~		~~6~~	
JJA	Mean (µg m^{-3})	63.6	78.6	80.8	84.4	85.6
	Difference in model mean (%)		-2.8		-1.4	
	NMB (%)		23.7	27.1		
	SD (µg m^{-3})	10.2	16.3	15.1	20.6	20.5
	~~Difference (%)~~		~~3~~		~~1~~	
SON	Mean (µg m^{-3})	46.3	48.6	55.0	52.7	54.9
	Difference in model mean (%)		-13.2		-4.2	
	NMB (%)		4.9	18.8		
	SD (µg m^{-3})	10.2	15.0	14.2	15.2	14.1
	~~Difference (%)~~		~~13~~		~~4~~	

Table 2: Statistical results comparing seasonal mean $PM_{2.5}$ concentrations simulated at the ~~global~~coarse and ~~regional~~finer resolutions to observations from 25 stations within the EMEP network in 2007. Statistical results for all model grid-cells of both resolutions are also shown. Percentage differences between the two model resolutions are calculated as ($PM_{2.5}$ ~~global~~coarse resolution − $PM_{2.5}$ ~~regional~~finer resolution)/($PM_{2.5}$ ~~global~~coarse resolution).

Season		25 sites			All grid-cells	
		Obs.	Model		Model	
			140 km	50 km	140 km	50 km
DJF	Mean (μg m^{-3})	12.1	8.3	9.5	5.1	5.5
	Difference in model mean (%)		-14.5		-7.8	
	NMB (%)		-31.0	-21.3		
	SD (μg m^{-3})	9.2	2.5	3.1	3.1	3.7
	~~Difference (%)~~		~~-14~~		~~-8~~	
MAM	Mean (μg m^{-3})	12.6	12.4	16.2	9.0	9.5
	Difference in model mean (%)		-30.6		-5.5	
	NMB (%)		-0.2	31.1		
	SD (μg m^{-3})	5.1	2.6	5.4	4.9	6.2
	~~Difference (%)~~		~~-31~~		~~-27~~	
JJA	Mean (μg m^{-3})	10.6	18.0	14.9	11.9	8.4
	Difference in model mean (%)		17.2		29.4	
	NMB (%)		70.0	40.1		
	SD (μg m^{-3})	4.0	5.4	6.4	7.0	6.2
	~~Difference (%)~~		~~17~~		~~29~~	
SON	Mean (μg m^{-3})	11.0	10.7	13.2	12.3	11.3
	Difference in model mean (%)		-23.4		8.1	
	NMB (%)		-2.4	22.0		
	SD (μg m^{-3})	4.8	4.1	10.3	7.0	6.7
	~~Difference (%)~~		~~-23~~		~~8~~	

Table 3: Warm season (April-September) mean of daily maximum 8-hour running mean O_3 concentrations (MDA8 O_3) and annual mean $PM_{2.5}$ concentrations at the coarse and finer resolutions compared to observations from 52 and 25 stations within the EMEP network, respectively.

Season		Obs.	140 km	50 km
MDA8 O_3 (Apr - Sept)	Mean (µg m^{-3})	86.3	95.6	94.8
	Difference in model mean (%)		0.8	
	NMB (%)		10.9	8.9
	SD (µg m^{-3})	9.2	14.7	14.2
$PM_{2.5}$ (Annual)	Mean (µg m^{-3})	11.4	12.6	13.7
	Difference in model mean (%)		-8.7	
	NMB (%)		10.5	20.2
	SD (µg m^{-3})	5.1	2.8	5.0

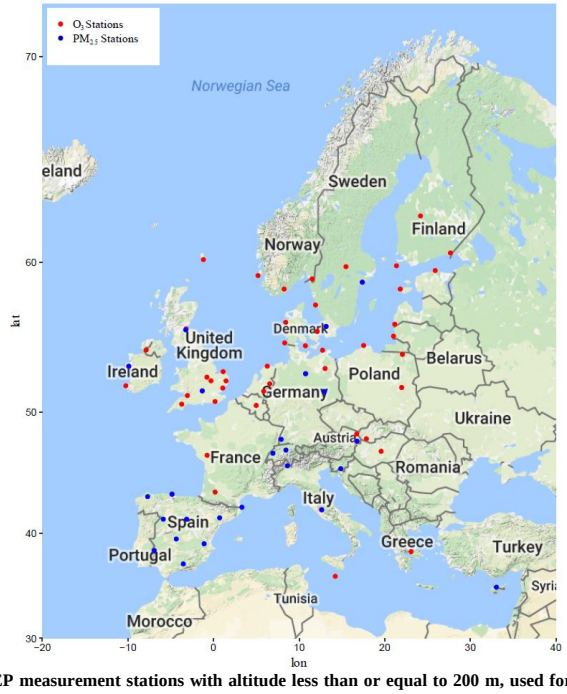


Figure 5: EMEP measurement stations with altitude less than or equal to 200 m, used for seasonal mean surface O_3 comparison to modelled concentrations (52 sites – red) and EMEP measurement stations used for seasonal mean $PM_{2.5}$ comparison to modelled concentrations (25 sites - blue)

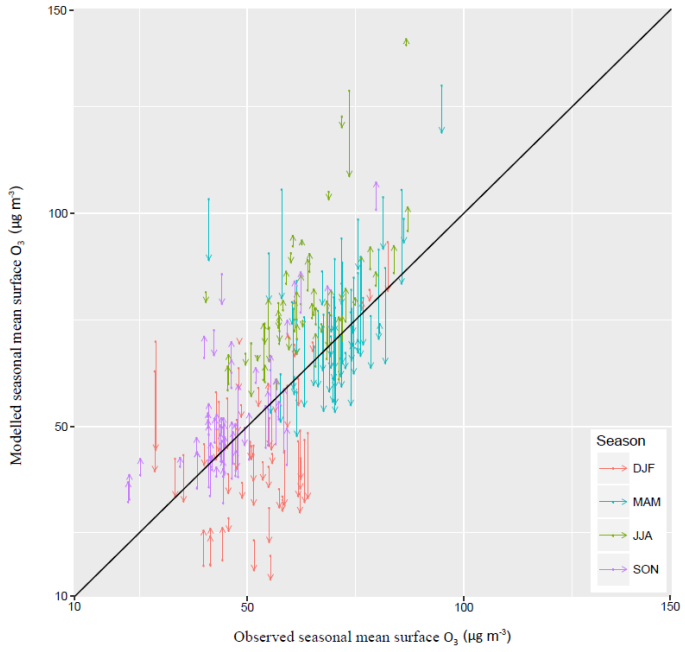


Figure 6: Seasonal mean modelled vs observed O_3 for 52 sites across the EMEP network for the year 2007. The arrow tails mark O_3 concentrations at the ~~global~~coarse resolution while the arrow heads represent the corresponding O_3 concentrations at the ~~regional~~finer resolution. The 1:1 line shows agreement between observed and simulated O_3.

860

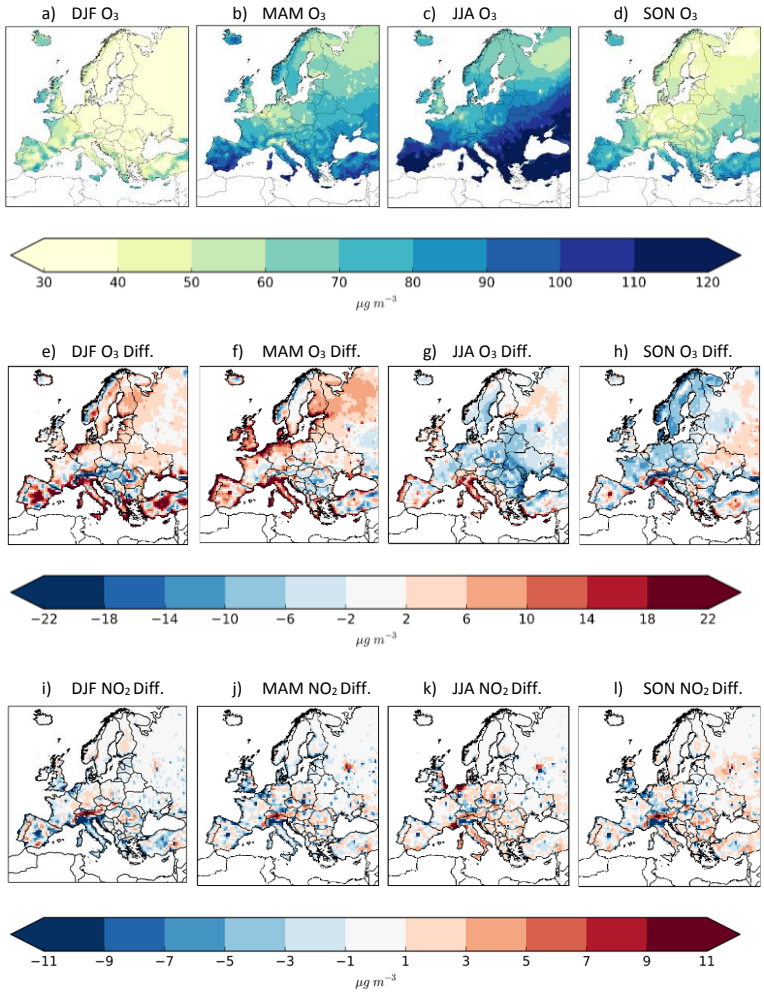


Figure 7: Seasonal mean O_3 simulated at the finer~~regional~~ resolution (top panel), differences in seasonal mean O_3 between the ~~global~~coarse and ~~regional~~finer resolutions (O_3 ~~global~~coarse resolution − O_3 ~~regional~~finer resolution) (middle panel) and NO_2 (NO_2 ~~global~~coarse resolution − NO_2 ~~regional~~finer resolution) (bottom panel). Blue regions in middle and bottom panels indicate that pollutant concentrations at the ~~global~~coarse resolution are lower (negative difference) while red regions indicate that concentrations are higher (positive difference) than those at the ~~regional~~finer resolution.

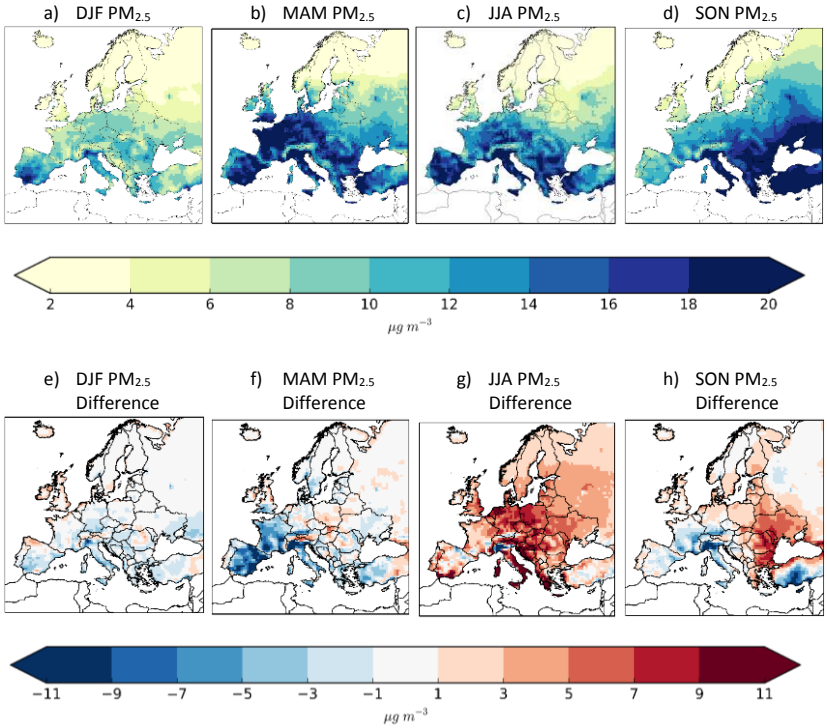
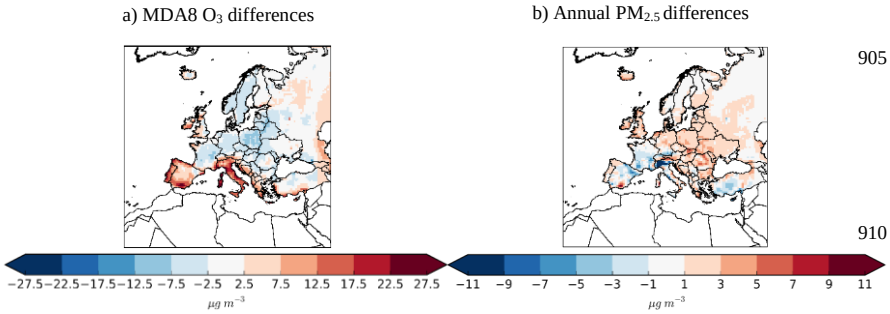


Figure 8: Seasonal mean $PM_{2.5}$ simulate at the ~~regional~~finer resolution (top panel) and differences between seasonal mean $PM_{2.5}$ at the ~~global~~coarse and ~~regional~~finer resolution in 2007 ($PM_{2.5}$ ~~global~~coarse resolution − $PM_{2.5}$ ~~regional~~finer resolution) (bottom panel).



905

910

Figure 9: Differences in a) warm season (April-September) mean of daily maximum 8-hour running mean O_3 (concentrations above 70 µg m^{-3}) and b) annual mean $PM_{2.5}$ between the ~~global~~coarse and ~~regional~~finer resolution (~~global~~coarse – finer~~regional~~).

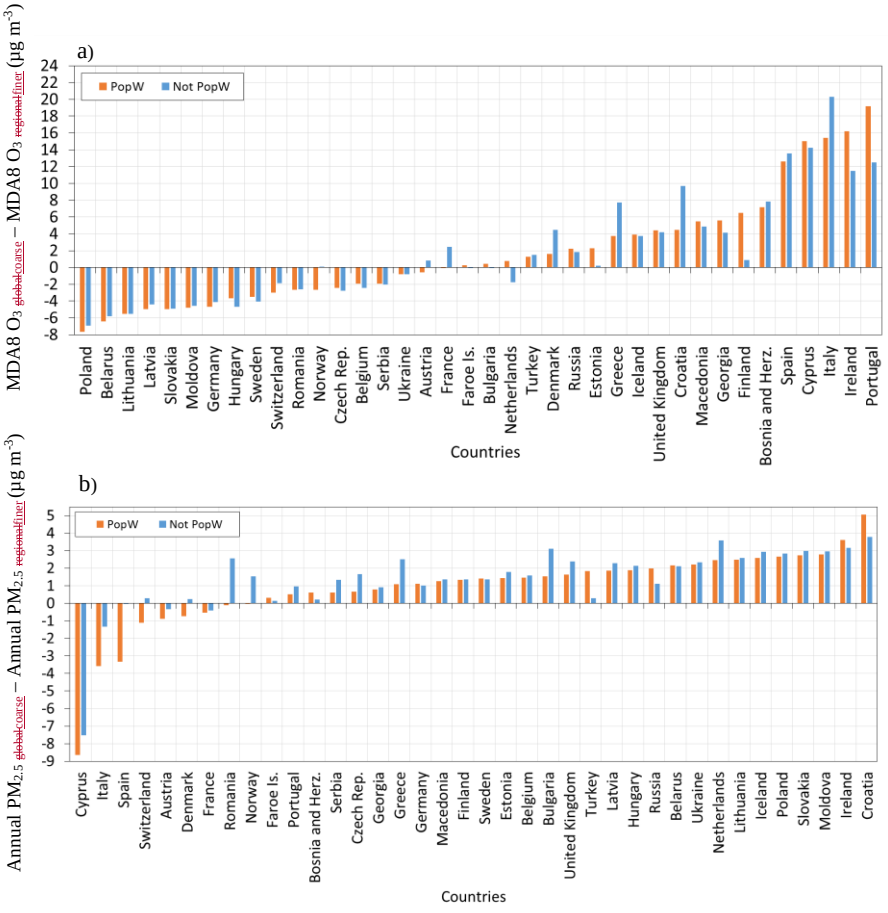


Figure 10: a) Differences between warm season mean daily maximum 8-hour running mean (MDA8) O_3 concentrations simulated at the two resolutions (~~global~~coarse – ~~regional~~finer) for population-weighted (PopW) concentrations (orange bars) and concentrations with no population-weighting (blue bars) b) same holds for annual mean $PM_{2.5}$ concentrations. Countries are ordered by differences in PopW pollutant concentrations between the two resolutions.

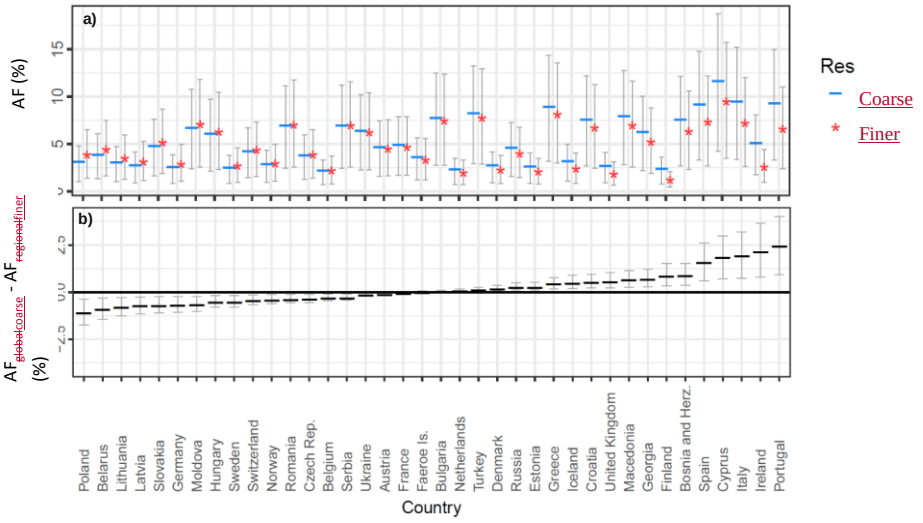


Figure 11: a) AF associated with long term exposure to daily maximum 8-hour running mean O_3 for each model resolution expressed as a percentage b) Differences in AF between the two resolutions expressed as a percentage for each European country ($AF_{global coarse} - AF_{regional finer}$). Grey lines show the 95% C.I. which represents uncertainties associated only with the concentration-response coefficient used.

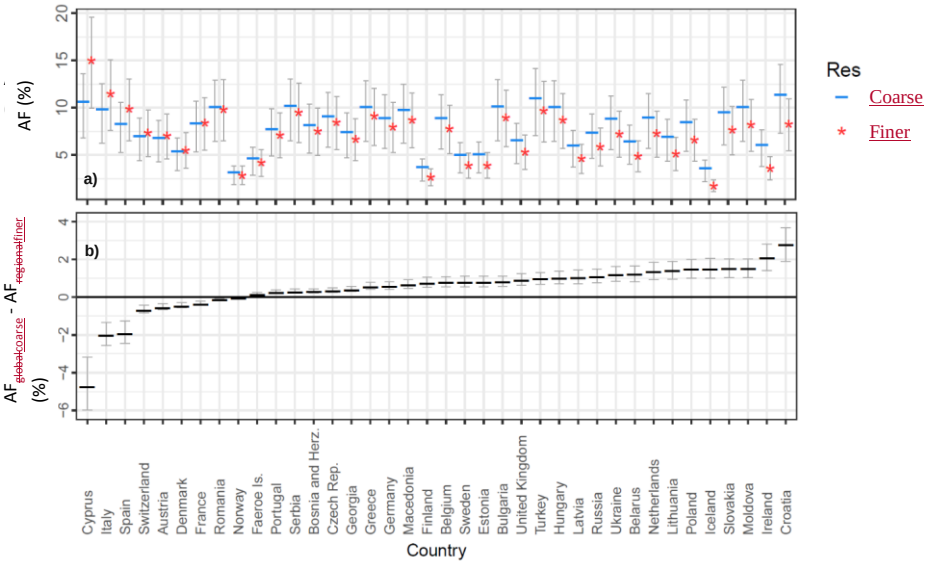


Figure 12: a) AF associated with long-term exposure to $PM_{2.5}$ for each model resolution expressed as a percentage b) Differences in AF between the two resolutions expressed as a percentage for each European country ($AF_{~~global~~coarse} - AF_{~~regional~~finer}$). Grey lines show the 95 % C.I. which represents uncertainties associated only with the concentration-response coefficient used.

940

945

950

955

960