Response to Referee #1:

General comments

This manuscript uses the HTAP2 models to quantify source-receptor relationships for surface ozone and fine particulate matter for emission reductions occurring in six world regions and globally, as well as within three emission sectors. These source-receptor relationships are then combined with concentration-response functions to estimate premature mortality due to intercontinental, within-region, and global emissions (which includes for three separate sectors). This manuscript builds on an existing body of literature coming out of HTAP1, and so, while not particularly novel in terms of methodology, it provides an important benchmark for comparison with earlier and future work.

Thank you for your careful review of our paper and constructive comments.

A serious weakness in the paper is the absence of model comparison to observations. At the very least the paper should include a summary of any evaluation of the HTAP2 models that may be appearing in other articles in this special issue, preferably ones that are already published. A stronger paper would evaluate the specific exposure metrics used to calculate health impacts. For example, observational estimates could be added to Table 1 for regions with ground-level networks. This seems especially relevant in light of the large discrepancies across the HTAP2 models. If some models could be discarded as unrealistic, it is possible that the uncertainty in the estimated numbers of premature mortalities due to the inter-model range may decrease.

Response:

Thank you for this comment. We had previously anticipated that other HTAP2 studies would include this comparison with observations. But we now see that while two papers do include comparisons in some regions, a full global comparison with observations for all of the models used in this study is desirable here. We have now included this model evaluation with ground level observations as described in the new section 2.2 (Lines 241-294):

"Measurements from multiple observation networks are employed in this study to evaluate the model performance around the world. We evaluate model performance for the 2010 baseline simulation for 11 TF-HTAP2 models for O_3 and 8 for $PM_{2.5}$ (Table S1). For O_3, we use ground level measurements from 2010 at 4,655 sites globally, collected by the Tropospheric Ozone Assessment Report (TOAR) (Schultz et al., 2017; Young et al., 2018). The TOAR dataset identifies stations as urban, rural and unclassified sites (Schultz et al., 2017). Model performance is evaluated for the average of daily 1-h maximum O_3 concentrations for the 3

consecutive months (3m1hmaxO_3) with the highest concentrations in each grid cell, including models that only report daily or monthly O_3 as described above. This metric for O_3 differs slightly from the 6-month average of daily 1-h maximum metric used for health impact assessment, and is chosen because TOAR reports the 3-month metric but not the 6-month metric. For $PM_{2.5}$, we compare the annual average $PM_{2.5}$, using $PM_{2.5}$ observations from 2010 at 3,157 sites globally selected for analysis by the Global Burden of Disease 2013 (GBD2013) (Forouzanfar et al., 2016). Statistical parameters including the normalized mean bias (NMB), normalized mean error (NME), and correlation coefficient (R) are selected to evaluate model performance.

Table S2 and S3 present statistical parameters of model evaluation for O_3 and $PM_{2.5}$, and Figures S3-S10 show the spatial O_3 and $PM_{2.5}$ evaluation as NMB around the world, and in North America, Europe and East Asia. For 3m1hmaxO_3, the model ensemble mean shows good agreement with measurements globally with NMB of 7.3% and NME of 13.2%, but moderate correlation with R of 0.53 (Table S2). For individual models, 8 models (CAM-chem, CHASER_T42, CHASER_T106, EMEPrv48, GEOSCHEMADJOINT, GEOS-Chem, GFDL_AM3 and HadGEM2-ES) overestimate 3m1hmaxO_3 with NMB of 9.2% to 23% while 3 models (C-IFS, OsloCTM3.v2 and RAQMS) underestimate by -10.8% to -19.4% globally (Figure S3). In the 6 perturbation regions, the model ensemble mean is also in good agreement with the measurements, with -11.2% to 25.3% for NMB, 9.8% to 25.3% for NME, and -0.09 to 0.98 for R. The ranges of NMB for individual models are -18.1% to 32.3%, -24.1% to 21.3%, -24.5% to 45.0%, -26.4% to 24.5%, -30.5% to 20.3%, -35.3% to 5.4%, in NAM, EUR, SAS, EAS, MDE, and RBU, respectively (Figure S4-S6). Note that some regions (SAS, MDE, and RBU) have very few observations for model evaluation, making the comparison less robust. The underestimated O_3 in the western US and overestimated O_3 in eastern US in most models is very close to the model performance result of Huang et al. (2017) who compare 8 TF-HTAP2 models with CASTNET observations (Figure S4) , as well as earlier studies under HTAP1 (Fiore et al. 2009). Similarly, Dong et al. (2018) find that O_3 is overestimated in EUR and EAS by 6 TF-HTAP2 models, consistent with our ensemble mean result in these two regions (Figure S5-S6).

For $PM_{2.5}$, the model ensemble mean agrees well with measurements globally, with NMB of -23.1%, NME of 35.4%, and R of 0.77 (Table S3). For individual models, only 1 model (GEOSCHEMADJOINT) overpredicts $PM_{2.5}$ by 20.3%, while the other 7 models underpredict $PM_{2.5}$ by -60.9% to -7.4% around the world (Figure S7). In 6 perturbation regions, the model ensemble mean is also in good agreement with measurements, with ranges of NMB of -49.7% to 19.4%, 21.2% to 49.7% for NME, and 0.50 to 1.00 for R. The range of NMB for individual models are -46.6% to 13.9%, -76.0% to 31.9%, -35.0% to 49.7%, -50.4% to 29.5%, -52.6% to 31.5%, and -74.1% to -

19.8%, in NAM, EUR, SAS, EAS, MDE, and RBU, respectively (Figure S8-S10). Dong et al. (2018) shows that $PM_{2.5}$ is underestimated in EUR and EAS by 6 TF-HTAP2 models, consistent with our ensemble mean result in these two regions (Figure S9-S10). Note that many observations used are located in urban areas, and models with coarse resolution may not be expected to have good model performance. Also several models neglect some $PM_{2.5}$ species, which may explain the tendency of models to underestimate."

In the abstract, some context could be provided as to whether the numbers here are in line with earlier work.

Response:

For impacts of intercontinental transport, we compare results from TF-HTAP2 with the previous TF-HTAP (Anenberg et al., 2009; 2014) and comparable studies (West et al., 2009; Duncan et al., 2008), and for sectoral reductions, we compare with previous studies by Crippa et al (2017), Lelieveld et al. (2015) and Silva et al. (2016a). We have modified the abstract to compare regional results with previous studies (Lines 63-66):

"Our findings that most avoided O_3-related deaths from emission reductions in NAM and EUR occur outside of those regions contrast with those of previous studies, while estimates of $PM_{2.5}$-related deaths from NAM, EUR, SAS and EAS emission reductions agree well."

And we have also modified the abstract to compare sectoral impacts (Lines 75-81):

"In sectoral emission reductions, TRN emissions account for the greatest fraction (26-53% of global emission reduction) of O_3-related premature deaths in most regions, in agreement with previous studies, except for EAS (58%) and RBU (38%) where PIN emissions dominate. In contrast, PIN emission reductions have the greatest fraction (38-78% of global emission reduction) of $PM_{2.5}$-related deaths in most regions, except for SAS (45%) where RES emission dominates, which differs with previous studies in which RES emissions dominate global health impacts."

Specific comments

1.Lines 63-68. Does this mean outside of any of the six regions?

Response:

This metric was not sufficiently clear in the previous draft. We now present two estimates of the impact of intercontinental transport on mortality, from the source and receptor perspectives, which are added to Tables 5 and 6. The estimate of the impact of intercontinental transport on mortality from the receptor perspective uses

the RERER metric that was introduced in previous HTAP studies. We express extra-regional deaths, as presented in the abstract, as the total avoided deaths outside of each source regions from six source emission reductions. We modified how this is presented in the abstract (Lines 67-70):
"For six regional emission reductions, the total avoided extra-regional mortality is estimated as 6,000 (-3,400, 15,500) deaths/year and 25,100 (8,200, 35,800) deaths/year through changes in O_3 and $PM_{2.5}$, respectively."

We added text to clarify how the RERER metric is defined (Lines 384-396):

"We estimate the impacts of extra-regional emission reductions on mortality by using the Response to Extra-Regional Emission Reduction (RERER) metric defined by TF-HTAP (Galmarini et al., 2017):

$$RERER_i = \frac{R_{global} - R_{region,i}}{R_{global}} \quad (4)$$

where for a given region *i*, R_{global} is the change in mortality in the global 20% reduction simulation (GLO) relative to the base simulation, and $R_{region,i}$ is the change in mortality in response to the 20% emission reduction from that same region *i*. A RERER value near 1 indicates a strong relative influence of foreign emissions on mortality within a region, while a value near 0 indicates a weak foreign influence. We also estimate the total avoided extra-regional mortality from a source perspective as the sum of avoided deaths outside of each of the 6 source regions, and from a receptor perspective by summing $R_{global} - R_{region,i}$ for all 6 regions."

We modified how the results are presented on these issues (Lines 558-561):
"Overall, adding results from all 6 regional reductions, interregional transport of air pollution from extra-regional contributions is estimated to lead to more avoided deaths through changes in $PM_{2.5}$ (25,100 (8,200, 35,800) deaths/year) than in O_3 (6,000 (-3,400, 15,500) deaths/year), consistent with Anenberg et al. (2009; 2014)."

We modified this in the discussion (Lines 649-653):
"Also, total avoided deaths through interregional air pollution transport are estimated as 6,000 (-3,400, 15,500) deaths/year for O_3 and 25,100 (8,200, 35,800) deaths/year for $PM_{2.5}$ in this study, in contrast with 7,300 (3,600, 11,200) deaths/year for O_3 and 11,500 (8,800, 14,200) deaths/year for $PM_{2.5}$ in Anenberg et al. (2009; 2014)."

And we modified this in the conclusions (Lines 732-735):
"Reductions from all six regions in the transport of air pollution between regions are

estimated to lead to more avoided deaths through changes in $PM_{2.5}$ (25,100 (8,200, 35,800) deaths/year) than for O_3 (6,000 (-3,400, 15,500) deaths/year)."

2.Many prior studies are mentioned in the introduction. Are there any robust findings across this prior body of work?

Response:

We have modified the introduction to point out the robust findings by prior studies (Lines 140-143):

"These prior studies have consistently concluded that most avoided O_3-related deaths from emission reductions in NAM and EUR occur outside of those regions, while most avoided $PM_{2.5}$-related deaths occur within the regions."

3.Lines 246-248. Is the actual value of β given somewhere?

Response:

We have added text to give the value of RR from the Jerrett study, from which Beta is calculated from equation 1 (Lines 316-318):

"For O_3, RR = 1.040 (95% Confidence Interval, CI: 1.013-1.067) for a 10 ppb increase in O_3 concentrations (Jerrett et al., 2009), which from eq. 1 gives values for β of 0.00392 (0.00129-0.00649)."

4.Line 261. Make sure all terms in equation 3 are defined.

Response:

Burnett et al. (2014) defines the function given and specifies three parameters (α, γ, δ) which they use to allow more flexibility in fitting the cause-specific RR. These parameters therefore do not have specific physical meaning, and are used in the functional fitting, so we refer the reader to Burnett's paper to understand these parameters more fully (Lines 329-334):

"RR is calculated as:

For $z<z_{cf}$, $RR_{IER}(z) = 1$ (2)

For $z \geqq z_{cf}$, $RR_{IER}(z) = 1 + \alpha\{1 - exp[-\gamma(z - z_{cf})^{\delta}]\}$ (3)

where z is the $PM_{2.5}$ concentration in μg/m^3 and z_{cf} is the counterfactual concentration below which no additional risk is assumed, and the parameters α, γ, and δ are used to fit the function for cause-specific RR (Burnett et al., 2014)."

5.Line 267-268. Elaborate on Zcf: does it vary from 5.8 to 8.8 g/m^3 in space and time?

Response:

We have revised to clarify (Lines 338-341):

"A uniform distribution from 5.8 μg/m^3 to 8.8 μg/m^3 is used for z_{cf} as suggested by

Burnett et al. (2014), which does not vary in space nor time. For uncertainty analysis, we use results from 1,000 Monte Carlo simulations of Burnett et al. (2014) to calculate RR in each grid cell by eq.2 or eq. 3."

6.Figures S8 and S9 are referred to several times in the text but are impossible to read. I suggest splitting them each into 4 figures, with half the models on each, one for the regional perturbations and one for the sectoral perturbations. The full range of the colorbar isn't used, so consider using a different color bar that allows for one to read the values off the figure more easily.

Response:

We split Figures S8-S9 into 6 pages, and we use a different color bar to show full range of data. See the updated Figs. S14-S17.

7.Lines 318-320. Is this intended to be a quantitative comparison? If so, are the metrics reported here and in the Lin et al. studies the same?

Response:

No, we don't intend to show a quantitative comparison with Lin et al. (2012 and 2017) due to the different ozone metrics evaluated. Instead, we suggest that the ozone responses in the western US to emission reductions from EAS are similar to those of Lin et al. (2012 and 2017) who show that a model can capture the measured western US ozone increases due to rising Asian emissions. We add this text (Lines 423-426):

"Our ensemble shows similar ozone responses in the western US to emission reductions from EAS (Figs. 1c) as those modeled by Lin et al. (2012 and 2017), who show that a model can capture the measured western US ozone increases due to rising Asian emissions."

8.Lines 449-459. This seems like methodology and could be included in the earlier section.

Response:

We have moved these lines into the method section (Lines 376-382):

"We also quantify the uncertainties in mortality due to the spread of air pollutant concentrations across models, RRs, and baseline mortality rates, as contributors to the overall uncertainty, expressed as a coefficient, of variation and compare the result with the Monte-Carlo analysis estimate. To do so, we hold two variables at their mean values and change the variable of interest within its uncertainty range; for example, using mean RRs and baseline mortality rates, we analyze the spread of the model ensemble to calculate the coefficient of variation caused by model

uncertainty."

9.Lines 545-547. Could the use of a different year make a difference here?

Response:

We agree with the reviewer that the different year could be responsible for part of the differences between studies. We have revised the text (Lines 653-657):

"These differences likely result from different concentration-response functions and the use of 6 regions here vs. 4 by Anenberg et al. (2009; 2014). In addition, updated atmospheric models and emissions inputs, as well as different atmospheric dynamics in the single years chosen in TF-HTAP1 vs. TF-HTAP2 may contribute to the differences."

10.Lines 559-560. This seems like an important point and suggest including in abstract and conclusions.

Response:

We have revised the abstract to add this comment (Lines 72-75):

"For NAM and EUR, our estimates of avoided mortality from regional and extra-regional emission reductions are comparable to those estimated by regional models for these same experiments."

And we have added it to the conclusions (Lines 735-738):

"For NAM and EUR, our estimates of avoided mortality from regional and extra-regional emission reductions are comparable to those estimated by regional models in AQMEII3 (Im et al., 2018) for these same emission reduction experiments."

11.Lines 609-610. Given the large ranges, is it really meaningful to report averages?

Response:

The overall percentage is derived from all 6 regional emission reductions altogether, not the average of percentages for each region. We've revised to clarify (Lines 722-727):

"For regional scenarios, 6 source emission reductions altogether can cause 84% of the global avoided O_3-related premature deaths within the source region, ranging from 21 to 95% among 6 regions, and 16% (5 to 79%) outside of the source region. For $PM_{2.5}$, 89% of global avoided $PM_{2.5}$-related premature deaths are within the source region, ranging from 32 to 94% among 6 regions, and 11% (6 to 68%) outside of the source region."

12.Table 4. What is an "empirical mean"?

Response:

Since we conduct 1,000 Monte Carlo simulations to propagate uncertainty from baseline mortality rates, modeled air pollutant concentrations, and the RRs in the health impact functions, the mean of the result is called the "empirical mean" as the mean of 1,000 simulations. We added this explanation into Table 4:

"Empirical mean is the mean of 1,000 Monte Carlo simulations."

We also revised the method section to explain where this result is used (Lines 373-375):

"The mean of the 1,000 Monte Carlo simulations (the "empirical mean") may differ from the mean when using the mean RR."

13.Table S1. Why not calculate $PM_{2.5}$ consistently across models from the individual components?

Response:

Different models use different functions to represent $PM_{2.5}$, that are appropriate for each model based on how different species are defined in the models. We choose to use the reported $PM_{2.5}$ from each model, rather than to recalculate $PM_{2.5}$ based on their reported species concentrations. We include the functions used by each model in Table S1 to communicate the species that each model simulated, and other modeling differences, where for example some models may be missing important species, but we do not apply these functions ourselves in this study.

Response to Referee #2

General comments

This manuscript uses the multi-model results from HTAP2 project to estimate mortality for the baseline year 2010, and health benefits from reduced emissions in source regions. In general, it is well organized and written, and the multi-model results can provide more reasonable range than single model results in previous studies. However, some details are not well documented and explanations are too general, but important for readers.

Thank you for your careful review of our paper and constructive comments.

Specific comments

1.Page 3, line 100: It is better to provide some brief explanation of reasons for large differneces in estimates (4.2 and 2.1 million premature deaths).

Response:

We have added short discussion on this point (Lines 107-109):

"These differences in GBD estimates result mainly from differences in concentration response functions and estimates of pollutant concentrations."

2.Page 5, line 159: Please specify if the perturbation is increasing or decreasing.

Response:

We reduced the anthropogenic emissions by 20% (Line 170):

"Anthropogenic emissions were reduced by 20% in six source regions: …"

3.Page 6, line 190-203: how do these models perform in simulating ozone and $PM_{2.5}$

Response:

Thank you for this comment. We had previously anticipated that other HTAP2 studies would include this comparison with observations. But we now see that while two papers do include comparisons in some regions, a full global comparison with observations for all of the models used in this study is desirable here. We have now included this model evaluation with ground level observations as described in the new section 2.2 (Lines 241-294):

"Measurements from multiple observation networks are employed in this study to evaluate the model performance around the world. We evaluate model performance for the 2010 baseline simulation for 11 TF-HTAP2 models for O_3 and 8 for $PM_{2.5}$ (Table S1). For O_3, we use ground level measurements from 2010 at 4,655 sites globally, collected by the Tropospheric Ozone Assessment Report (TOAR)

(Schultz et al., 2017; Young et al., 2018). The TOAR dataset identifies stations as urban, rural and unclassified sites (Schultz et al., 2017). Model performance is evaluated for the average of daily 1-h maximum O_3 concentrations for the 3 consecutive months ($3m1hmaxO_3$) with the highest concentrations in each grid cell, including models that only report daily or monthly O_3 as described above. This metric for O_3 differs slightly from the 6-month average of daily 1-h maximum metric used for health impact assessment, and is chosen because TOAR reports the 3-month metric but not the 6-month metric. For $PM_{2.5}$, we compare the annual average $PM_{2.5}$, using $PM_{2.5}$ observations from 2010 at 3,157 sites globally selected for analysis by the Global Burden of Disease 2013 (GBD2013) (Forouzanfar et al., 2016). Statistical parameters including the normalized mean bias (NMB), normalized mean error (NME), and correlation coefficient (R) are selected to evaluate model performance.

Table S2 and S3 present statistical parameters of model evaluation for O_3 and $PM_{2.5}$, and Figures S3-S10 show the spatial O_3 and $PM_{2.5}$ evaluation as NMB around the world, and in North America, Europe and East Asia. For $3m1hmaxO_3$, the model ensemble mean shows good agreement with measurements globally with NMB of 7.3% and NME of 13.2%, but moderate correlation with R of 0.53 (Table S2). For individual models, 8 models (CAM-chem, CHASER_T42, CHASER_T106, EMEPrv48, GEOSCHEMADJOINT, GEOS-Chem, GFDL_AM3 and HadGEM2-ES) overestimate $3m1hmaxO_3$ with NMB of 9.2% to 23% while 3 models (C-IFS, OsloCTM3.v2 and RAQMS) underestimate by -10.8% to -19.4% globally (Figure S3). In the 6 perturbation regions, the model ensemble mean is also in good agreement with the measurements, with -11.2% to 25.3% for NMB, 9.8% to 25.3% for NME, and -0.09 to 0.98 for R. The ranges of NMB for individual models are -18.1% to 32.3%, -24.1% to 21.3%, -24.5% to 45.0%, -26.4% to 24.5%, -30.5% to 20.3%, -35.3% to 5.4%, in NAM, EUR, SAS, EAS, MDE, and RBU, respectively (Figure S4-S6). Note that some regions (SAS, MDE, and RBU) have very few observations for model evaluation, making the comparison less robust. The underestimated O_3 in the western US and overestimated O_3 in eastern US in most models is very close to the model performance result of Huang et al. (2017) who compare 8 TF-HTAP2 models with CASTNET observations (Figure S4) , as well as earlier studies under HTAP1 (Fiore et al. 2009). Similarly, Dong et al. (2018) find that O_3 is overestimated in EUR and EAS by 6 TF-HTAP2 models, consistent with our ensemble mean result in these two regions (Figure S5-S6).

For $PM_{2.5}$, the model ensemble mean agrees well with measurements globally, with NMB of -23.1%, NME of 35.4%, and R of 0.77 (Table S3). For individual models, only 1 model (GEOSCHEMADJOINT) overpredicts $PM_{2.5}$ by 20.3%, while the other 7 models underpredict $PM_{2.5}$ by -60.9% to -7.4% around the world (Figure S7). In 6 perturbation regions, the model ensemble mean is also in good agreement with

measurements, with ranges of NMB of -49.7% to 19.4%, 21.2% to 49.7% for NME, and 0.50 to 1.00 for R. The range of NMB for individual models are -46.6% to 13.9%, -76.0% to 31.9%, -35.0% to 49.7%, -50.4% to 29.5%, -52.6% to 31.5%, and -74.1% to -19.8%, in NAM, EUR, SAS, EAS, MDE, and RBU, respectively (Figure S8-S10). Dong et al. (2018) shows that $PM_{2.5}$ is underestimated in EUR and EAS by 6 TF-HTAP2 models, consistent with our ensemble mean result in these two regions (Figure S9-S10). Note that many observations used are located in urban areas, and models with coarse resolution may not be expected to have good model performance. Also several models neglect some $PM_{2.5}$ species, which may explain the tendency of models to underestimate."

4.Page 7, line 246-257: what beta value is used in this study? any source for the used RR=1.040? Please clarify..

Response:

We added this (Lines 316-318):

"For O_3, RR = 1.040 (95% Confidence Interval, CI: 1.013-1.067) for a 10 ppb increase in O_3 concentrations (Jerrett et al., 2009), which from eq. 1 gives values for β of 0.00392 (0.00129-0.00649)."

5.Page 8, line 264-271: The used RR framework here is not actually the latest. Please refer to Cohen et al. (2017).

Response:

Our work was nearly completed before Cohen et al. (2017) was published, and so we chose the most recent available RR from Burnett et al. (2014) for $PM_{2.5}$. This function was widely used in many studies, including by Silva et al (2016a), Lelieveld et al (2015), and GBD 2010 (Lim et al., 2012). However, we have added short discussion on this difference (Lines 473-481):

"Cohen et al. (2017) use RRs for particulate matter for IHD and stroke mortality that are modified from those used by Burnett et al (2014) and applied age modification to the RRs, fitting the IER model for each age group separately. The updated IER with estimated higher relative risks, together with greater global pollution and baseline mortality rates in the low-income and middle-income countries in east and south Asia leads to the higher absolute numbers of attributable deaths and disability-adjusted life-years in GBD 2015 than estimated in GBD 2013 (Forouzanfar et al., 2016). Also, GBD 2015 includes child lower respiratory infections estimate whereas we do not".

6.Page 8, line 276-277: Please clarify how you treat age distribution in the 2011

populaiton dataset.

Response:

We add text (Lines 355-358):

"For the population of adults aged 25 and older, we use ArcGIS 10.2 geoprocessing tools to estimate the population per 5-year age group in each cell by multiplying the country level percentage in each age group by the population in each cell."

7.Page 8, is sex difference considered in the estimation?

Response:

No, we only consider age-specific RR, as given by the health impact functions we use and the underlying epidemiological studies.

8.Page 8, line 282: Monte Carlo simulation is powerful to address uncertainty issues. However, the way of including model air pollutant concentrations is a bit misleading. The procedure in this study is actually the range of multi-model results. However, it is possible that this range deviate from the observations. Without showing model evaluation, we don't have confidence how reliable is the range from multi-models.

Response:

We've added the model evaluation in section 2.2 (Lines 241-294).

We also added an acknowledgement that the range of models in an ensemble is not a true reflection of the uncertainty in emissions to the method section (Lines 371-373):

"One should acknowledge that the range of models in an ensemble is not a true reflection of the uncertainty in emissions to concentration relationships."

10.Page 9, line 306: The texts refer to supplemental plots many times. I would suggest move some important figures from supplemental materials.

Response:

We've moved figures S6-S7 to figures 1-2 in the main paper. The order of figures has been updated to reflect this change in main paper as well as the supporting document.

11.Page 10, line 368-369: Please provide more details here: the updated baseline mortality rate in 2017, and how population is different. This comparison is too general here. In my understanding, the biggest change from GBD framework from old to latest (Cohen et al., 2017) is not just baseline mortality. In Cohen et al. (2017), the RR for stoke is totally different from previous version GBD, and LRI disease is added in addition to IHD, LC, COPD and stroke.

Response:

As stated before, we now provide details on how RRs were updated for use by Cohen et al. (2017) (Lines 473-481):

"Cohen et al. (2017) use RRs for particulate matter for IHD and stroke mortality that are modified from those used by Burnett et al (2014) and applied this age modification to the RRs, fitting the IER model for each age group separately. The updated IER with estimated higher relative risks, together with greater global pollution and baseline mortality rates in the low-income and middle-income countries in east and south Asia leads to the higher absolute numbers of attributable deaths and disability-adjusted life-years in GBD 2015 than estimated in GBD 2013 (Forouzanfar et al., 2016). Also, GBD 2015 includes child lower respiratory infections estimate whereas we do not."

12.Page 11 line 382-383: Please clarify how the avoid deaths is calculated. the IER model is not linear: at the high end large changes in pollutant will not result in large changes in death, some studies used average changes, some used marginal. How is this addressed here?

Response:

The percentage of the global change in O_3-related deaths within the source region is computed by the number of avoided deaths within source region divided by the number of avoided deaths globally from 20% source emission reduction. We've revised to clarify this calculation (Lines 495-496):

"The number of avoided deaths within source region is divided by the number of avoided deaths globally"

We added text to discuss the issue about IER model (Lines 343-352):

"However, in the IER model, the concentration–response function flattens off at higher $PM_{2.5}$ concentrations, yielding different estimates of avoided premature mortality for identical changes in air pollutant concentrations from less-polluted vs. highly-polluted regions. That is, one unit reduction of air pollution may have a stronger effect on avoided mortality in regions where pollution levels are lower (e.g., Europe, North America) compared with highly polluted regions (e.g., East Asia, India), which would not be the case for a log-linear function (Jerrett et al., 2009; Krewski et al., 2009). Therefore, using the IER model in this study may result in smaller changes in avoided mortality in highly polluted areas than using the linear model."

13.Page 11 line 406-408: The explanation here is not convincing.

Response:

We've revised this explanation to (Lines 520-522):
"In addition, updated atmospheric models and emissions inputs, as well as different atmospheric dynamics in the single years chosen in TF-HTAP1 vs. TF-HTAP2 may contribute to the differences."

14.It would be great to make a table to inter-compare the response of sector reductions, which is highly uncertain from different models, and please discuss it too.
Response:
We've listed this inter-comparison between models for sector reductions in TableS9-S10 and discussed these differences in Lines 616-625:
"Considering results from individual models, we found that O_3-and $PM_{2.5}$-related mortality from TRN emission reductions show greater relative uncertainty than from PIN or RES (Table 5-6 and Table S9-S10), reflecting a greater spread of results across models. Regional impacts from individual models also differ from the ensemble mean result - e.g., for O_3, GEOSCHEMADJOINT and OsloCTM3.v2 show that reducing PIN emissions causes the greatest fraction of avoided O_3-related deaths in EUR, while GEOSCHEMADJOINT, HadGM2-ES and OsloCTM3.v2 show that TRN emissions have the greatest fraction of avoided O_3-related deaths in RBU (Figs. S20). For $PM_{2.5}$, CHASER_t42 and GEOSCHEMADJOINT show that reducing PIN emissions causes the greatest fraction of avoided $PM_{2.5}$-related deaths in SAS (Figs. S21)"

In addition, we also compare our O_3 and $PM_{2.5}$-related premature deaths attributable to PIN, TRN and RES emissions with previous studies conducted by Silva et al. (2016) and Lelieveld et al. (2015) in table 7 and discuss the differences from our estimates in Lines 601-615 :
"In comparison with other studies (Table 7), our conclusion that PIN emissions cause the most O_3-related deaths and TRN emissions cause the greatest fraction of avoided deaths in most regions agrees well with Silva et al (2016a). For $PM_{2.5}$, reducing PIN emissions avoids the most $PM_{2.5}$-related premature deaths globally (128,000 (41,600, 179,000) deaths/year) and in most regions (38-78% of the global emission reduction), except for SAS (45%) where the RES emission dominates. Although these findings differ from those of Lelieveld et al (2015) and Silva et al (2016), who find that Residential emissions have the greatest of impact on $PM_{2.5}$ mortality globally and in most regions, all studies agree that PIN emissions have the greatest impact in NAM. Our result is also comparable with Crippa et al (2017) who find that PIN emissions have the greatest health impact in most countries. Although comparable emission inventories are used (i.e. Lelieveld et al (2015) use EDGAR emissions while Silva et al (2016) use RCP8.5. emissions), our lower mortality estimate for RES emissions may

be explained by our 20% reductions relative to the zero-out method, and the different years simulated."

and Lines 674-686 :
"Differences in our estimates of premature mortality attributable to air pollution from three emission sectors (multiplied by 5) may be explained by methodological differences relative to previous studies (Silva et al., 2016; Lelieveld et al., 2015), including our use of 20% emission reductions versus the zero-out method in those studies, different emission inventories, a multi-model ensemble versus single models, and differences in baseline mortality rates, population, and concentration response functions. Our finding that TRN emissions contribute the most avoided deaths for O_3 in most regions agrees well with the result by Silva et al (2016a), but differs for $PM_{2.5}$ mortality for which we find that PIN emissions cause the most deaths, while both Silva et al (2016a) and Lelieveld et al (2015) find that RES emissions are responsible for the most deaths. This discrepancy may be explained by different $PM_{2.5}$ species included in individual models, as we showed that changes in $PM_{2.5}$ concentration to TRN emission differ across models."

HTAP2 multi-model estimates of premature human mortality due to intercontinental transport of air pollution

Ciao-Kai Liang[1], J. Jason West[1], Raquel A. Silva[2], Huisheng Bian[3], Mian Chin[4], ~~Frank J. Dentener[5],~~ Yanko Davila[5,6], Frank J. Dentener[6], Louisa Emmons[7], Johannes Flemming[8], Gerd Folberth[8,9], ~~Johannes Flemming[9],~~ Daven Henze[6], Ulas Im[10], Jan Eiof Jonson[11], Terry J. Keating[12], Tom Kucsera[12,13], ~~Terry J. Keating[13], Marianne Tronstad Lund[14],~~ Allen Lenzen[14,15], Meiyun Lin[16,5], Marianne Tronstad Lund[16], Xiaohua Pan[17], Rokjin J. Park[18], R. Bradley Pierce[17,9], ~~Rokjin J. Park[18], Xiaohua Pan[19],~~ Takashi Sekiya[20], Kengo Sudo[20], Toshihiko Takemura[21]

[1]Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[2]Oak Ridge Institute for Science and Education at US Environmental Protection Agency, Research Triangle Park, NC, USA

[3]Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore, MD, USA

[4]Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, MD, USA

[5,6]Department of Mechanical Engineering, University of Colorado, Boulder, CO, USA

[6,5]European Commission, Joint Research Center, Ispra, Italy

~~[6]Department of Mechanical Engineering, University of Colorado, Boulder, CO, USA~~

[7]Atmospheric Chemistry Observations and Modeling Laboratory, National Center for Atmospheric Research (NCAR), Boulder, CO, USA

[8,9]European Center for Medium-R~~a~~range Weather Forecast~~s~~~~ing~~, Reading, UK

[8,9]UK Met Office Hadley Centre, Exeter, UK

~~[9]European Center for Medium-range Weather Forecasting, Reading, UK~~

[10]Aarhus University, Department of Environmental Science, Frederiksborgvej, DK-4000, Roskilde, Denmark.

[11]Norwegian Meteorological Institute, Oslo, Norway

[13,2]US Environmental Protection Agency, Research Triangle Park, NC, USA

[12,13]Universities Space Research Association, Greenbelt, MD, USA

~~[13]US Environmental Protection Agency, Research Triangle Park, NC, USA~~

~~[14]CICERO Center for International Climate Research, Oslo, Norway~~

[15,4]Space Science & Engineering Center, University of Wisconsin -Madison, WI, USA

[16,5]Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA

[16,4]CICERO Center for International Climate Research, Oslo, Norway

[19,7]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA

[17]NOAA National Environmental Satellite, Data, and Information Service, Madison, WI, USA
[18]Seoul National University, Seoul, Korea
[19]NOAA National Environmental Satellite, Data, and Information Service, Madison, WI, USA
[19]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
[20]Nagoya University, Furocho, Chigusa-ku, Nagoya, Japan
[21]Research Institute for Applied Mechanics, Kyushu University, Fukuoka, Japan

Correspondence to: J. Jason West (jjwest@email.unc.edu)

Abstract

Ambient air pollution from ozone and fine particulate matter is associated with premature mortality. As emissions from one continent influence air quality over others, changes in emissions can also influence human health on other continents. We estimate global air pollution-related premature mortality from exposure to $PM_{2.5}$ and ozone , and the avoided deaths from 20% anthropogenic emission reductions from six source regions, North America (NAM), Europe (EUR), South Asia (SAS), East Asia (EAS), Russia/Belarus/Ukraine (RBU) and the Middle East (MDE), three global emission sectors, Power and Industry (PIN), Ground Transportation (TRN) and Residential (RES) and one global domain (GLO), using an ensemble of global chemical transport model simulations coordinated by the second phase of the Task Force on Hemispheric Transport of Air Pollution (TF-HTAP2), and epidemiologically-derived concentration-response functions. We build on results from previous studies of the TF-HTAP by using improved atmospheric models driven by new estimates of 2010 anthropogenic emissions (excluding methane), with more source and receptor regions, new consideration of source sector impacts, and new epidemiological mortality functions. We estimate 290,000 (95% CI: 30,000, 600,000) premature O_3-related deaths and 2.8 million (0.5 million, 4.6 million) $PM_{2.5}$-related premature deaths globally for the baseline year 2010. While 20% emission reductions from one region generally lead to more avoided deaths within the source region than outside, reducing emissions from MDE and RBU can avoid more O_3-related deaths outside of these regions than within, and reducing MDE emissions also avoids more $PM_{2.5}$-related deaths outside of MDE than within. Our findings that most avoided O_3-related deaths from emission reductions in NAM and EUR occur outside of those regions contrast with those of previous studies, while estimates of $PM_{2.5}$-related deaths from NAM, EUR, SAS and EAS emission

reductions agree well. In addition, EUR, MDE and RBU have more avoided O_3-related deaths from reducing foreign emissions than from domestic reductions. For six regional emission reductions, the total avoided extra-regional mortality is estimated as 6,000 (-3,400, 15,500)~~10,300 (6,700, 13,400)~~ deaths/year and 25,100 (8,200, 35,800)~~42,000 (12,400, 60,100)~~ deaths/year through changes in O_3 and $PM_{2.5}$, respectively. Interregional transport of air pollutants leads to more deaths through changes in $PM_{2.5}$ than in O_3, even though O_3 is transported more on interregional scales, since $PM_{2.5}$ has a stronger influence on mortality. For NAM and EUR, our estimates of avoided mortality from regional and extra-regional emission reductions are comparable to those estimated by regional models for these same experiments. In sectoral emission reductions, TRN emissions account for the greatest fraction (26-53% of global emission reduction) of O_3-related premature deaths in most regions, in agreement with previous studies, except for EAS (58%) and RBU (38%) where PIN emissions dominate. In contrast, PIN emission reductions have the greatest fraction (38-78% of global emission reduction) of $PM_{2.5}$-related deaths in most regions, except for SAS (45%) where RES emission dominates~~.~~, which differs with previous studies in which RES emissions dominate global health impacts. The spread of air pollutant concentration changes across models contributes most to the overall uncertainty in estimated avoided deaths, highlighting the uncertainty in results based on a single model. Despite uncertainties, the health benefits of reduced intercontinental air pollution transport suggest that international cooperation may be desirable to mitigate pollution transported over long distances.

1 Introduction

Ozone (O_3) and fine particulate matter with aerodynamic diameter less than 2.5 μm ($PM_{2.5}$) are two common air pollutants with known adverse health effects. Epidemiological studies have shown that both short-term and long-term exposures to O_3 and $PM_{2.5}$ are associated with elevated rates of premature mortality. Short-term exposure to O_3 is associated with respiratory morbidity and mortality (Bell et al., 2005; Bell et al., 2014; Gryparis et al., 2004; Ito et al., 2005; Levy et al., 2005; Stieb et al., 2009) while long-term exposure to O_3 has been associated with premature respiratory mortality (Jerrett et al., 2009, Turner et al., 2016). Short-term exposure to $PM_{2.5}$ has been associated with increases in daily mortality rates from all natural causes, and specifically from respiratory and cardiovascular causes (Bell et al., 2014; Du et al., 2016; Powell et al., 2015; Pope et al., 2011) while long-term exposure to $PM_{2.5}$ can have detrimental chronic health effects, including premature mortality due to cardiopulmonary diseases and lung cancer (Brook et al., 2010; Burnett et al., 2014;

Hamra et al., 2014; Krewski et al., 2009; Lepeule et al., 2012; Lim et al., 2012). The Global Burden of Disease Study 2015 (GBD 2015) estimated 254,000 deaths/year associated with ambient O_3 and 4.2 million associated with ambient $PM_{2.5}$ (Cohen et al. 2017). A comparable study using output from an ensemble of global chemistry–climate models estimated 470,000 deaths/year associated with O_3 and 2.1 million premature deaths/year associated with anthropogenic $PM_{2.5}$ (Silva et al. 2013). These differences in GBD estimates result mainly from differences in concentration response functions and estimates of pollutant concentrations.

Numerous observational and modeling studies have shown that anthropogenic emissions can affect O_3 and $PM_{2.5}$ concentrations across continents (Dentener et al., 2010; Heald et al., 2006; ~~TF-HTAP, 2010;~~ Leibensperger et al., 2011; Lin et al., 2012; Lin et al., 2017; Liu et al., 2009a; West et al., 2009a; Wild and Akimoto, 2001; Yu et al., 2008). As changes in emissions from one continent influence air quality over others, several studies have estimated the premature mortality from intercontinental transport (Anenberg et al., 2009; Anenberg et al., 2014; Bhalla et al., 2014; Duncan et al., 2008; Im et al., 2018~~7~~; Liu et al., 2009b; West et al., 2009b; Zhang et al., 2017). In 2005, the Task Force on Hemispheric Transport of Air Pollution (TF-HTAP) was launched under the United Nations Economic Commission for Europe (UNECE) Convention on Long-Range Transboundary Air Pollution (LRTAP). One of its tasks is to investigate the impacts of emission reductions on the intercontinental transport of air pollution, air quality, health, ecosystem and climate effects, using a multi-model ensemble to quantify uncertainties due to differences between models (Anenberg et al., 2009; Anenberg et al., 2014; Fiore et al., 2009; Fry et al., 2012; Huang et al., 2017; Stjern et al., 2016; Yu et al., 2013).

In the TF-HTAP Phase 1 (TF-HTAP1), human premature mortality due to 20% anthropogenic emission reductions in four large source regions was investigated by Anenberg et al. (2009 and 2014). They found that 20% foreign O_3 precursor emission reductions contribute approximately 30% to >50% of the deaths avoided by reducing precursor emissions in all four regions together (Anenberg et al., 2009). Similarly, reducing emissions in NA and EU was found to avoid more O_3-related premature deaths outside the source region than within (Anenberg et al., 2009), which agrees with other studies that together show for the first time that emission reductions in NA and EU have greater impacts on mortality outside the source region than within (Duncan et al., 2008; West et al., 2009). In contrast, Anenberg et al. (2014) estimate that 93–97 % of $PM_{2.5}$-related avoided deaths from reducing emissions in all four regions occurs within the source region while 3–7 % occur outside the source region from transport between continents. Despite the longer atmospheric lifetime of O_3 and its relatively larger scale of influence, $PM_{2.5}$ was found to cause more deaths from intercontinental transport

(Anenberg et al., 2009; 2014). These prior studies have consistently concluded that most avoided O_3-related deaths from emission reductions in NAM and EUR occur outside of those regions, while most avoided $PM_{2.5}$-related deaths occur within the regions. Similarly, an ensemble of regional models in the third phase of the Air Quality Modelling Evaluation International Initiative (AQMEII3) found that a 20% decrease of emissions within the source region avoids 54,000 and 27,500 premature deaths in Europe and the U.S. (from both O_3 and $PM_{2.5}$), while the reduction of foreign emissions alone avoids ~1,000 and 2,000 premature deaths in Europe and the U.S. (Im et al., ~~2017~~2018). Crippa et al (2017) used the TM5-FASST reduced-form model with HTAP2 emissions to estimate a global sensitivity to 20 % emission reductions of $PM_{2.5}$-related premature deaths of 401,000 globally, and 42,000 and 20,000 for Europe and the US respectively.

In addition, several studies have evaluated the relative importance of individual emissions sectors (Barrett et al., 2010; Bhalla et al., 2014; Chafe et al., 2014; Chambliss et al., 2014; Corbett et al., 2007) or multiple sectors (Lelieveld et al., 2015; Silva et al., 2016a) to ambient air pollution–related premature mortality. Lelieveld et al. (2015) estimated that residential energy use such as for heating and cooking has the largest mortality impact globally (for $PM_{2.5}$ and O_3 mortality combined), particularly in South and East Asia. Silva et al (2016) likewise found that residential & commercial emissions are most important for ambient $PM_{2.5}$-related mortality, but also found that land transportation had the greatest impact on O_3-related mortality, particularly in North America, South America, Europe, FSU and the Middle East. Understanding the impact of different sectors on the global burden and the relative importance of each sector among regions can help stimulate international efforts and region-specific air pollution control strategies. Nevertheless, those studies were limited by using a single atmospheric model, reflecting a need to understand whether results differ among models and apportionment approaches.

In this study, we estimate the impacts of interregional transport and of source sector emissions on human premature mortality from O_3 and $PM_{2.5}$, using an ensemble of global chemical transport models coordinated by the Task Force on Hemispheric Transport of Air Pollution Phase 2 (TF-HTAP2) (Galmarini et al., ~~2016~~2017; Huang et al., ~~2016~~2017; Janssens-Maenhout et al., 2015; Stjern et al., 2016). Anthropogenic emissions were ~~perturbed~~ reduced by 20% in six source regions: North America (NAM), Europe (EUR), South Asia (SAS), East Asia (EAS), Russia/Belarus/Ukraine (RBU) and the Middle East (MDE), three emission sectors: Power and Industry (PIN), Ground Transportation (TRN) and Residential (RES), and one worldwide region (GLO). Human premature mortality due to these reductions is calculated using a health impact function based on a log-linear model for O_3 (Jerrett et al. 2009) and an integrated

exposure-response model for $PM_{2.5}$ (Burnett et al. 2014), within the six source regions and elsewhere in the world. We conduct a Monte Carlo simulation to estimate the overall uncertainty due to uncertainties in relative risk, air pollutant concentrations (given by the spread of results among different models), and baseline mortality rates.

2 Method

2.1 Modeled O_3 and $PM_{2.5}$ surface concentration

Global numerical modelling experiments initiated by TF-HTAP2, the regional experiments by the Air Quality Model Evaluation International Initiative (AQMEII) over Europe and North America, and the Modelling Intercomparison Study-Asia (MICS-Asia) were coordinated to perform consistent emission perturbation modelling experiments across the global, hemispheric and continental/regional scales (Galmarini et al., ~~2016~~2017). Simulation periods, meteorology, emission inventories, boundary conditions, and model output are also consistent. The Joint Research Centre's (JRC) EDGAR (Emission Data Base for Global Research) team in collaboration with regional emission experts from the U.S. Environmental Protection Agency (US-EPA), European Monitoring and Evaluation Programme (EMEP), Centre on Emission Inventories and Projections (CEIP), Netherlands Organization for Applied Research (TNO), and the MICS-Asia Scientific Community and Regional Emission Activity Asia (REAS) provide a global emission inventory at 0.1^0x0.1^0 resolution for TF-HTAP2 modeling experiments (Janssens-Maenhout et al., 2015). The emissions dataset was constructed for SO_2, NO_X, CO, NMVOC, NH_3, PM_{10}, $PM_{2.5}$, BC and OC and seven emission sectors (shipping, aircraft, land transportation, agriculture, residential, industry and energy) for the year 2010 (Fig. S1).

This study uses outputs from 14 global models / model versions (Table S1) participating in TF-HTAP2. Overall, TF-HTAP2 model resolutions are finer than in TF-HTAP1. In TF-HTAP2, each model performed a baseline simulation and sensitivity simulations where the anthropogenic emissions in a defined source region or sector were perturbed (reduced by 20% in most cases). Based on the number of models that simulated different experiments, we choose to focus on emission reductions from six source regions, three emission sectors, and one global domain. More specifically, all anthropogenic emissions are reduced by 20% in the North America (NAM), Europe (EUR), South Asia (SAS), East Asia (EAS), Russia/Belarus/Ukraine (RBU) and the Middle East (MDE) continental regions, in the Power and Industry (PIN), Ground Transportation (TRN) and Residential (RES) emission sectors globally, and in one global domain (GLO) (Fig. S2). Unlike TF-HTAP1 (Dentener et al., 2010~~TF-HTAP, 2010~~) which defined rectangular regions that included ocean or some sparsely inhabited

regions, TF-HTAP2 regions are defined by geopolitical boundaries.

We selected output from the models that provided temporally resolved volume mixing ratios of O_3 and mass mixing ratios of $PM_{2.5}$ ("mmrpm2p5") for the baseline and at least one regional or sectoral emission reduction scenario. Among the 14 models, 11 models reported O_3 and 8 reported $PM_{2.5}$ for regional emission perturbation scenarios, 4 models reported O_3 and 4 reported $PM_{2.5}$ for sectoral emission perturbation scenarios, and 10 models reported O_3 and 8 reported $PM_{2.5}$ for the global emission perturbation. All models used prescribed meteorology for the year 2010, although this meteorology was derived from different (re-)analysis products and ~~was~~ not uniform across models. Modeled concentrations are processed by calculating metrics consistent with the underlying epidemiological studies to estimate premature mortality. For O_3, we calculate the average of daily 1-h maximum O_3 concentration for the 6 consecutive months with the highest concentrations in each grid cell (Jerrett et al., 2009), for the baseline and each 20% emission reduction scenario. While some models reported hourly O_3 metrics, others only reported daily or monthly O_3. We include these models by first calculating the ratio of the 6-month average of daily 1-h maximum O_3 to the annual average of O_3 in individual grid cells, for models reporting hourly O_3, and then applying that ratio to the annual average of ozone for those models that only report daily or monthly O_3, following Silva et al. (2013; 2016b). For $PM_{2.5}$, we calculate the annual average $PM_{2.5}$ concentration in each cell using the monthly total $PM_{2.5}$ concentrations reported by each model ("mmrpm2p5"). Model results for these two metrics are then regridded from each model's native grid resolution (varying from $0.5^o \times 0.5^o$ to $2.8^o \times 2.8^o$) to a consistent $0.5^o \times 0.5^o$ resolution used in mortality estimation. We estimate regional and sectoral multi-model averages for each 20% emission reduction scenario in the year 2010, but for each perturbation case, we only include models that report both the baseline and perturbation cases.

2.2 Model evaluation

Measurements from multiple observation networks are employed in this study to evaluate the model performance around the world. We evaluate model performance for the 2010 baseline simulation for 11 TF-HTAP2 models for O_3 and 8 for $PM_{2.5}$ (Table S1). For O_3, we use ground level measurements from 2010 at 4,655 sites globally, collected by the Tropospheric Ozone Assessment Report (TOAR) (Schultz et al., 2017; Young et al., 2018). The TOAR dataset identifies stations as urban, rural and unclassified sites (Schultz et al., 2017). Model performance is evaluated for the average of daily 1-h maximum O_3 concentrations for the 3 consecutive months ($3m1hmaxO_3$) with the highest concentrations in each grid cell, including models that only report daily or monthly O_3 as described above. This metric for O_3 differs slightly from the 6-month average of daily 1-h maximum metric used for health impact assessment, and is chosen

because TOAR reports the 3-month metric but not the 6-month metric. For $PM_{2.5}$, we compare the annual average $PM_{2.5}$, using $PM_{2.5}$ observations from 2010 at 3,157 sites globally selected for analysis by the Global Burden of Disease 2013 (GBD2013) (Forouzanfar et al., 2016). Statistical parameters including the normalized mean bias (NMB), normalized mean error (NME), and correlation coefficient (R) are selected to evaluate model performance.

Table S2 and S3 present statistical parameters of model evaluation for O_3 and $PM_{2.5}$, and Figures S3-S10 show the spatial O_3 and $PM_{2.5}$ evaluation as NMB around the world, and in North America, Europe and East Asia. For $3m1hmaxO_3$, the model ensemble mean shows good agreement with measurements globally with NMB of 7.3% and NME of 13.2%, but moderate correlation with R of 0.53 (Table S2). For individual models, 8 models (CAM-chem, CHASER_T42, CHASER_T106, EMEPrv48, GEOSCHEMADJOINT, GEOS-Chem, GFDL_AM3 and HadGEM2-ES) overestimate $3m1hmaxO_3$ with NMB of 9.2% to 23% while 3 models (C-IFS, OsloCTM3.v2 and RAQMS) underestimate by -10.8% to -19.4% globally (Figure S3). In the 6 perturbation regions, the model ensemble mean is also in good agreement with the measurements, with -11.2% to 25.3% for NMB, 9.8% to 25.3% for NME, and -0.09 to 0.98 for R. The ranges of NMB for individual models are -18.1% to 32.3%, -24.1% to 21.3%, -24.5% to 45.0%, -26.4% to 24.5%, -30.5% to 20.3%, -35.3% to 5.4%, in NAM, EUR, SAS, EAS, MDE, and RBU, respectively (Figure S4-S6). Note that some regions (SAS, MDE, and RBU) have very few observations for model evaluation, making the comparison less robust. The underestimated O_3 in the western US and overestimated O_3 in eastern US in most models is very close to the model performance result of Huang et al. (2017) who compare 8 TF-HTAP2 models with CASTNET observations (Figure S4) , as well as earlier studies under HTAP1 (Fiore et al. 2009). Similarly, Dong et al. (2018) find that O_3 is overestimated in EUR and EAS by 6 TF-HTAP2 models, consistent with our ensemble mean result in these two regions (Figure S5-S6).

For $PM_{2.5}$, the model ensemble mean agrees well with measurements globally, with NMB of -23.1%, NME of 35.4%, and R of 0.77 (Table S3). For individual models, only 1 model (GEOSCHEMADJOINT) overpredicts $PM_{2.5}$ by 20.3%, while the other 7 models underpredict $PM_{2.5}$ by -60.9% to -7.4% around the world (Figure S7). In 6 perturbation regions, the model ensemble mean is also in good agreement with measurements, with ranges of NMB of -49.7% to 19.4%, 21.2% to 49.7% for NME, and 0.50 to 1.00 for R. The range of NMB for individual models are -46.6% to 13.9%, -76.0% to 31.9%, -35.0% to 49.7%, -50.4% to 29.5%, -52.6% to 31.5%, and -74.1% to -19.8%, in NAM, EUR, SAS, EAS, MDE, and RBU, respectively (Figure S8-S10). Dong et al. (2018) shows that $PM_{2.5}$ is underestimated in EUR and EAS by 6 TF-HTAP2 models, consistent with our ensemble mean result in these two regions (Figure S9-S10).

Note that many observations used are located in urban areas, and models with coarse resolution may not be expected to have good model performance. Also several models neglect some $PM_{2.5}$ species, which may explain the tendency of models to underestimate.

2.3~~2~~ Health impact assessment

We use output from the TF-THAP2 model ensemble to estimate annual O_3- and $PM_{2.5}$-related global cause-specific premature mortality and avoided mortality from the 20% regional and sectoral emission reductions, following the same methods used by Silva et al. (2016a; 2016b). The annual O_3- and $PM_{2.5}$-related premature mortality is calculated using a health impact function based on epidemiological relationships between ambient air pollution concentration and mortality in each grid cell: $\Delta M = y_0 \times AF \times Pop$, where ΔM is premature mortality, y_0 is the baseline mortality rate (for the exposed population), $AF=1 - 1/RR$ is the attributable fraction, where RR is relative risk of death attributable to the change in air pollutant concentration ($RR=1$ when there is no increased risk of death associated with a change in pollutant concentration), and Pop is the exposed population (adults aged 25 and older).

For O_3 mortality, we use a log-linear model for chronic respiratory mortality (RESP) from the American Cancer Society (ACS) study (Jerrett et al 2009), following recent studies including the GBD (Cohen et al., 2017), but Turner et al. (2016) recently published new results for chronic ozone mortality, and adoption of these results would lead to more ozone-related deaths overall (Malley et al., 2017). RR is calculated as:

$$RR = e^{\beta \Delta x} \quad (1)$$

where β is the concentration-response factor, and Δx corresponds to the change in pollutant concentrations between simulations with perturbed emissions and the baseline simulation~~;~~. For O_3, RR = 1.040 (95% Confidence Interval, CI: 1.013-1.067) for a 10 ppb increase in O_3 concentrations (Jerrett et al., 2009), which from eq. 1 gives values for β of 0.00392 (0.00129-0.00649). ~~and $RR = 1.040$ (95% Confidence Interval, CI: 1.013-1.067) for a 10 ppb increase in O_3 concentrations.~~ We estimate O_3-related premature deaths due to respiratory disease (RESP) based on decreases or increases in O_3 concentration (i.e. Δx) due to 20% regional and sectoral emission reduction scenarios relative to the baseline. For regional and sectoral reductions, we do not assume a low-concentration threshold below which changes in O_3 have no mortality effects, as there is no clear evidence for such a threshold, following Anenberg et al (2009; 2010) and Silva et al. (2013; 2016a, b). However, we evaluate global O_3 premature mortality for the baseline 2010 simulation, relative to a counterfactual concentration of 37.6 ppb (Lim et al. 2012), for consistency with GBD estimates (Cohen et al., 2017).

For $PM_{2.5}$ mortality, we apply the Integrated Exposure–Response (IER) model, which is intended to better represent the risk of exposure to $PM_{2.5}$ at locations with high ambient concentrations (Burnett et al., 2014). RR is calculated as:

For $z<z_{cf}$, $RR_{IER}(z) = 1$ (2)

For $z \geqq z_{cf}$, $RR_{IER}(z) = 1 + \alpha\{1 - exp[-\gamma(z - z_{cf})^{\delta}]\}$ (3)

where z is the $PM_{2.5}$ concentration in $\mu g/m^3$ and z_{cf} is the counterfactual concentration below which no additional risk is assumed, and the parameters α, γ, and δ are used to fit the function for cause-specific RR (Burnett et al., 2014). The overall $PM_{2.5}$-related cause-specific premature deaths related to ischemic heart disease (IHD), cerebrovascular disease (STROKE), chronic obstructive pulmonary disease (COPD) and lung cancer (LC) are estimated using RRs per age group for IHD and STROKE and RRs for all ages for COPD and LC. A uniform distribution from 5.8 $\mu g/m^3$ to 8.8 $\mu g/m^3$ is used for z_{cf} as suggested by Burnett et al. (2014), which does not vary in space nor time. For uncertainty analysis, we use results from 1,000 Monte Carlo simulations of Burnett et al. (2014) to calculate RR in each grid cell by eq.2 or eq. 3. We estimate avoided premature mortality in 20% emission perturbation experiments by taking the difference in premature mortality estimates with the 2010 baseline. However, in the IER model, the concentration–response function flattens off at higher $PM_{2.5}$ concentrations, yielding different estimates of avoided premature mortality for identical changes in air pollutant concentrations from less-polluted vs. highly-polluted regions. That is, one unit reduction of air pollution may have a stronger effect on avoided mortality in regions where pollution levels are lower (e.g., Europe, North America) compared with highly polluted regions (e.g., East Asia, India), which would not be the case for a log-linear function (Jerrett et al., 2009; Krewski et al., 2009). Therefore, using the IER model in this study may result in smaller changes in avoided mortality in highly polluted areas than using the linear model.

For the exposed population, we use the Oak Ridge National Laboratory's Landscan 2011 Global Population Dataset at approximately 1 km resolution (30"x30") (Bright et al., 2012). For the population of adults aged 25 and older, we use ArcGIS 10.2 geoprocessing tools to estimate the population per 5-year age group in each cell by multiplying the country level percentage in each age group by the population in each cell. ~~We use ArcGIS 10.2 geoprocessing tools to estimate the population of adults aged 25 and older from Landscan.~~ We obtained cause-specific baseline mortality rates for 187 countries from the GBD 2010 mortality dataset (IHME, 2013). The population and baseline mortality per age group were regridded to the 0.5°×0.5° grid (Table ~~S2~~ S4 and Fig. S11~~3~~). Cause-specific baseline mortality rates vary geographically, e.g. RESP and COPD are relatively more dominant in South Asia, IHD in Europe, STROKE in Russia, and LC in North America.

Finally, we conduct 1,000 Monte Carlo simulations to propagate uncertainty from baseline mortality rates, modeled air pollutant concentrations, and the RRs in health impact functions. We use the reported 95% CIs for cause-specific baseline mortality rates, assuming lognormal distributions. For modeled O_3 and $PM_{2.5}$ concentrations we used the absolute value of the coefficient of variation among models in each grid cell, for each 20% emission perturbation case minus the baseline, assuming a normal distribution. For O_3 RRs, we use the reported 95% confidence intervals (CIs), assuming a normal distribution. For $PM_{2.5}$ RRs, we use the parameter values (i.e. α, γ, δ and z_{cf}) of Burnett et al. (2014) for 1,000 simulations. One should acknowledge that the range of modeled air pollution concentrations in an ensemble is not a true reflection of the uncertainty in emissions to concentration relationships. The mean health outcome of the 1,000 Monte Carlo simulations (the "empirical mean") may differ from the mean when using the mean RR.

We also quantify the uncertainties in mortality due to the spread of air pollutant concentrations across models, RRs, and baseline mortality rates, as contributors to the overall uncertainty, expressed as a coefficient, of variation and compare the result with the Monte-Carlo analysis estimate. To do so, we hold two variables at their mean values and change the variable of interest within its uncertainty range; for example, using mean RRs and baseline mortality rates, we analyze the spread of the model ensemble to calculate the coefficient of variation caused by model uncertainty. Given that our 0.5°×0.5° grid cell resolution can capture most of the population well in a given region, uncertainty associated with population was assumed to be negligible.

We estimate the impacts of extra-regional emission reductions on mortality by using the Response to Extra-Regional Emission Reduction (RERER) metric defined by TF-HTAP (Galmarini et al., 2017):

$$RERER_i = \frac{R_{global} - R_{region,i}}{R_{global}} \quad (4)$$

where for a given region i, R_{global} is the change in mortality in the global 20% reduction simulation (GLO) relative to the base simulation, and $R_{region,i}$ is the change in mortality in response to the 20% emission reduction from that same region i. A RERER value near 1 indicates a strong relative influence of foreign emissions on mortality within a region, while a value near 0 indicates a weak foreign influence. We also estimate the total avoided extra-regional mortality from a source perspective as the sum of avoided deaths outside of each of the 6 source regions, and from a receptor perspective by summing $R_{global} - R_{region,i}$ for all 6 regions.

3 Results

3.1 Response of O_3 and $PM_{2.5}$ concentrations to 20% regional and sectoral emission reductions

Previous TF-HTAP studies reported area-averaged concentrations to quantify source-receptor relationships averaging concentrations over a region (Doherty et al., 2013; Fiore et al., 2009; Fry et al., 2012; Huang et al., 2017; Stjern et al., 2016; ~~Sanderson et al., 2008;~~ Yu et al., 2013). Here, we present the population-weighted concentration over a region, which is more relevant for health. Among six receptor regions, the population-weighted multi-model mean O_3 concentrations range from 48.38±8.05 ppb in EUR to 65.72±10.08 ppb in SAS with a global average of 53.74± 8.03 ppb, while the annual population-weighted multi-model mean $PM_{2.5}$ concentrations range from 9.36±2.62 μg/m^3 in NAM to 39.27±13.50 μg/m^3 in EAS with a global average of 25.98±5.05 μg/m^3 (Table 1 and S5~~3~~-S6~~4~~ and Figs.S12~~4~~-S13~~5~~).

For 20% perturbation scenarios, in general the impact on the multi-model mean change in surface O_3 and $PM_{2.5}$ concentration is greater within the source region (i.e., domestic region) than outside of it (i.e., foreign region) (Figs. ~~S6~~1-~~S7~~2). This is also true for individual model results (Figs. S14~~8-S9 and -~~S16). Among six source regions, the emission reduction from SAS has the greatest impact on global population-weighted O_3 concentration (Tables 2 and ~~S3~~S5), while that from EAS has greatest impact on $PM_{2.5}$ (Tables 3 and ~~S4~~S6). The source-receptor pairs with the greatest changes in O_3 and $PM_{2.5}$ concentration reflect the geographical proximity between regions and the magnitude of emissions (Table 2-3) – e.g., EUR→MDE (0.34±0.08 ppb), EUR→RBU (0.34 ppb±0.09), EAS→NAM (0.29±0.14 ppb), EAS→RBU (0.27±0.12 ppb), and NAM→EUR (0.26±0.55 ppb) for O_3, and EUR→RBU (0.26±0.19 μg/m^3), EUR→MDE (0.18±0.08 μg/m^3), MDE→SAS (0.12±0.06 μg/m^3), SAS→EAS (0.08±0.08 μg/m^3), and EAS→SAS (0.08±0.07 μg/m^3) for $PM_{2.5}$. Our ensemble shows similar ozone responses in the western US to emission reductions from EAS (Figs. ~~S6c~~1c) as those modeled by Lin et al. (2012 and 2017), who show that a model can capture the measured western US ozone increases due to rising Asian emissions.

For each receptor region, reducing foreign anthropogenic emissions by 20% (estimated by global minus within-region reductions) can decrease population-weighted O_3 concentrations by 29-74% of the change in O_3 concentration and 8–41 % of the change in $PM_{2.5}$ concentration (Tables 2-3). In some cases, regional emission reductions cause small O_3 concentration increases within the source region or in foreign receptors, reflecting O_3 nonlinear responses (Figs. S~~8~~14 ~~and S10~~). For instance, C-IFS_v2 predicts O_3 concentration increases in EUR by 0.04 ppb from domestic

emission reductions, which is in agreement with results from TF-HTAP1 (Anenberg et al. 2009). Similarly, ~~CMAchem~~ CMAchem shows more local O_3 increases, particularly in SAS, than other models (Figs. S1~~8 and S104~~). The change in O_3 concentration in foreign receptors is broader than for $PM_{2.5}$, reflecting that O_3 has a longer atmospheric lifetime than $PM_{2.5}$.

For sectors, TRN emission reductions cause the greatest decrease in global population-weighted O_3 by 1.13±0.19 ppb, while PIN emission reductions cause the greatest decrease in surface $PM_{2.5}$ by 1.46±0.56 μg/m^3 globally (Tables 2-3 and Figs. 1-2). The 20% emission reductions from individual sectors also have different effects in different regions. Of the three sectors, emission reductions from TRN have the greatest effect on population-weighted O_3 in NAM, EUR, SAS, MDE and MDE (40-50% of the global emission reduction) while PIN emission reductions dominate in EAS (57%). Emission reductions from PIN have the greatest effect on population-weighted $PM_{2.5}$ in NAM, EUR, EAS, MDE and MDE (41-84%) while RES emission reductions dominate in SAS (43%). The response of $PM_{2.5}$ concentration to sectoral emission reductions differs significantly across models, which reflects in part the $PM_{2.5}$ species simulated by each model (Table S1 and Figs. S15 ~~and~~ -S17). For instance, we found that models that simulate $PM_{2.5}$ nitrate (i.e. CHASER_t42 and GEOSCHEMADJOIN) predict a greater impact on $PM_{2.5}$ concentration from TRN emission reduction than those without nitrate (i.e. GOCARTv5 and SPRINTARS) (Fig. S17~~9~~).

3.2 Global mortality burden associated with anthropogenic air pollution

Table 4 shows the annual multi-model mean O_3- and $PM_{2.5}$-related premature deaths on 6 regions and globally for year 2010 baseline with 95% confidence intervals (CI) based on Monte Carlo sampling. Tables ~~S5~~S7-S8~~6~~ show estimates of premature deaths due to anthropogenic O_3 and $PM_{2.5}$ from individual models. For the ensemble model mean, we estimate 290,000 (30,000, 600,000) premature O_3-related deaths globally using a 37.6 ppb counterfactual concentration, and 2.8 million (0.5 million, 4.6 million) $PM_{2.5}$-related premature deaths using a uniform distribution of counterfactual concentration from 5.8 μg/m^3 to 8.8 μg/m^3. Highly populated areas of India and East Asia have the greatest O_3- and $PM_{2.5}$-related deaths, and those regions together account for 82% and 66% of the global total O_3- and $PM_{2.5}$-related deaths. Compared with the GBD 2015 (Cohen et al 2017), our global burden estimates are greater than the 254,000 (97,000, 422,000) premature deaths/year for O_3 from GBD, while less than 4.2 million (3.7 million, 4.8 million) premature deaths for $PM_{2.5}$. Lelieveld et al (2015) estimate 142,000 (CI: 90,000, 208,000) O_3-related deaths and 3.2 million (1.5 million, 4.6 million) $PM_{2.5}$-related premature deaths for 2015. These differences can be explained mainly by exposure estimates. Here we used a multi-model ensemble, whereas

Lelieveld et al. (2015) used a single model, and Cohen et al (2017) used a single model for O_3 and a single model combined with surface and satellite observations for $PM_{2.5}$. In addition, Cohen et al. (2017) use RRs for particulate matter for IHD and stroke mortality that are modified from those used by Burnett et al (2014) and applied age modification to the RRs, fitting the IER model for each age group separately. The updated IER with estimated higher relative risks, together with greater global pollution and baseline mortality rates in the low-income and middle-income countries in east and south Asia leads to the higher absolute numbers of attributable deaths and disability-adjusted life-years in GBD 2015 than estimated in GBD 2013 (Forouzanfar et al., 2016). Also, GBD 2015 includes child lower respiratory infections estimate whereas we do not. ~~Cohen et al (2017) use higher updated baseline mortality rate and population which leads to higher global premature deaths estimate.~~ Our wider range of uncertainty for the global mortality reflects the uncertainty in baseline rates, RRs and spread of air pollutant concentration across models whereas Cohen et al (2017) consider national-level population-weighted mean concentrations and uncertainty of IER function predictions at each concentration and Lelieveld et al. (2015) only account for the statistical uncertainty of the parameters used in the IER functions.

3.3 Effect of regional reductions on mortality

Reducing global anthropogenic emissions of air pollutant by 20% avoids 47,400 (11,300, 99,000) O_3-related deaths and 290,000 (67,100, 405,000) $PM_{2.5}$-related premature deaths (Tables 5-6 and ~~S7~~S9-S10~~8~~). Most avoided air pollution-related deaths were found within or close to the source region (Figs.3~~1~~-7~~4~~6). Reducing anthropogenic emissions by 20% from NAM, EUR, SAS, EAS, MDE and RBU can avoid 54%, 54%, 95%, 85%, 21%, and 22% of the global change in O_3-related deaths within the source region (The number of avoided deaths within source region is divided by the number of avoided deaths globally), and 93%, 81%, 93%, 94%, 32%, and 82% of the global change in $PM_{2.5}$-related deaths, respectively (Table 5-6). Whereas the most O_3-related premature deaths can be avoided by reducing SAS emissions (20,000 (3,600, 42,200) deaths/year), reducing EAS emissions avoids more O_3-related premature deaths (1,700 (-1,300, 5,400)) outside of the source region than for any other region (500 (180, 870) deaths/year to 1,300 (-1,200, 4,400) deaths/year (Table 5). Similarly, while reducing EAS emissions avoids the most $PM_{2.5}$-related premature deaths (96,600 (3,500, 136,000) deaths/year), reducing EUR emissions avoids more $PM_{2.5}$-related premature deaths (7,400 (930, 9,500) deaths/year) outside of the source region than for any other region (1,400 (-320, 2,300) deaths/year to 5,500 (3,000, 7,800) deaths/year) (Table 6). While emission reductions from one region generally lead to more avoided deaths within the source region than outside, 20% anthropogenic emission reductions

from MDE (i.e. 79% and 68% of global avoided deaths outside of source region for ~~both~~ O_3 and $PM_{2.5}$, respectively) and RBU (78% for O_3) can avoid more premature deaths outside of the source region than within (Table 5-6). This result for RBU is in agreement with West et al (2009). However, the results for NAM and EUR do not agree with previous studies that found that emission reductions in these regions cause more O_3-related avoided premature deaths outside of the source region than within (Anenberg et al., 2009; Duncan et al., 2008; West et al., 2009). For $PM_{2.5}$, our results are comparable with Anenberg et al. (2014) and Crippa et al. (2017) who found that for most regions, $PM_{2.5}$-related avoided premature deaths are higher within the source region than outside. The above difference in results with TF-HTAP1 may be in part because of the definition of regions. Whereas the TF-HTAP2 regions are defined by geopolitical boundaries, the TF-HTAP1 regions are defined by square domains which are larger and include more ocean areas (Anenberg et al., 2009). In addition, updated atmospheric models and emissions inputs, as well as different atmospheric dynamics in the single years chosen in TF-HTAP1 vs. TF-HTAP2 may contribute to the differences. ~~This could lead more emissions like aviation and shipping emission reduced by TF-HTAP1 experiment, reflecting the higher premature deaths can be avoided in downwind regions. TF-HTAP2 also adds new regions (RBU and MDE) that have strong influences on air quality in adjacent regions.~~

Using individual models, different conclusions may result for the relative importance of inter-regional transport. For example, for O_3, 8 models predict that NAM emission reductions cause more O_3-related premature deaths within NAM (i.e CAM-Chem, CHASER_T42, CHASER_T106, C-IFS, GEOSCHEMADJOINT, GEOS-Chem, GFDL_AM3 and HadGEM2-ES), whereas 2 models predict more deaths outside (i.e. EMEPrv48 and OsloCTM3.v2). 5 models suggest that EUR emission reductions cause more O_3-related premature deaths within EUR (i.e. CAM-chem, CHASER_T42, CHASER_T106, GFDL_AM3 and HadGEM2-ES), whereas 4 show more deaths outside (i.e. C-IFS, GEOSCHEMADJOINT, EMEPrv48 and OsloCTM3.v2). Each individual model shows that emission reductions from SAS and EAS avoid more O_3-related premature deaths within than outside, and that those from MDE and RBU avoid more O_3-related premature deaths outside than within (Fig. S~~8 and S10~~18). For $PM_{2.5}$, each individual model shows that emission reductions from NAM, EUR, SAS, EAS and RBU avoid more $PM_{2.5}$-related premature deaths within than outside, while for emission reductions from MDE, 3 models (EMEPrv48, GEOSCHEMADJOINT and SPRINARS) show more $PM_{2.5}$-related premature deaths within, while 3 (CHASER_T42 GEOS5 and GOCART) show more $PM_{2.5}$-related premature deaths outside (Fig. S~~9 and S11~~19). The variation of health effect reflects the differences in processing of natural emissions, atmospheric physical and chemical mechanisms,

numerics~~transport time step~~ etc across models.

For each receptor region, reducing domestic anthropogenic emissions by 20% contributes about 66%, 39%, 84%, 72%, 45% and 25% of the total O_3-related avoided premature mortality (from the global reduction), and 90%, 78%, 87%, 87%, 58% and 66% of the total $PM_{2.5}$-related avoided premature mortality (from the global reduction) in NAM, EUR, SAS, EAS, MDE and RBU, respectively (Table 5-6). Therefore, reducing emissions from foreign regions avoids more O_3 premature deaths in EUR (foreign emission account for 61% of total avoided deaths from the global reduction), MDE (55%) and RBU (75%) than reducing domestic emissions (Table 5-6), in agreement with the results for EUR from Anenberg et al (2009). Whereas EAS has the greatest number of avoided O_3-related premature deaths due to foreign emission reduction (3,800 (3,600, 3,900) deaths/year), RBU has the greatest fraction of O_3 mortality from foreign emission reductions (75%) (Table 5). Similarly, for $PM_{2.5}$, while EAS has greatest number of avoided $PM_{2.5}$-related premature deaths due to foreign emission reductions (13,600 (3,500, 18,800) deaths/year), MDE has the greatest fraction of $PM_{2.5}$ mortality from foreign emission reduction (42%) (Table 6).

Overall, adding results from all 6 regional reductions, interregional transport of air pollution from extra-regional contributions is estimated to lead to more avoided deaths through changes in $PM_{2.5}$ (25,100 (8,200, 35,800)~~42,000 (12,400, 60,100)~~ deaths/year) than in O_3 (6,000 (-3,400, 15,500)~~10,300 (6,700, 13,400)~~ deaths/year), consistent with Anenberg et al. (2009; 2014). This result is due to the greater influence of $PM_{2.5}$ on mortality, despite the shorter atmospheric lifetime of $PM_{2.5}$ relative to O_3.

The contributions of different factors to the overall uncertainties in mortality are shown in Tables S11-S12, considering uncertainties due to the spread of air pollutant concentrations across models, RRs, and baseline mortality rates, expressed as coefficients of variation. ~~We quantify the uncertainties in mortality due to the spread of air pollutant concentrations across models, RRs, and baseline mortality rates, as contributors to the overall uncertainty, expressed as a coefficient of variation and compare the result with the Monte Carlo analysis estimate (Tables S9-S10). For the spread of the model ensemble, we calculate the deterministic mean and standard deviation estimates for mortality with mean RRs and baseline mortality rates. For RRs, we use the 95%CI of RRs reported by Jerrett et al. (2009) for O_3 and upper and lower bound of RRs reported by Burnett et al. (2014) for $PM_{2.5}$ with mean baseline mortality rates to estimate the deterministic mean and standard deviation for mortality. For baseline mortality rates, we use upper and lower bound of baseline mortality rates with mean RRs to estimate the deterministic mean and standard deviation for mortality.~~ For both O_3 and $PM_{2.5}$ mortality, the spread of model results generally contributes most to the overall uncertainty, followed by uncertainty in RRs and in baseline mortality rates,

for most source-receptor pairs. The spread of model results is generally wider for $PM_{2.5}$ (14% to 3974% among source-receptor pairs) than for O_3 (13% to 1065%). The uncertainty in RRs for O_3 mortality has constant value (33% to 34%) due to the fixed uncertainty range of RRs from Jerrett et al. (2009), whereas $PM_{2.5}$ mortality leads to a wider range of uncertainty (1% to 247%) in RRs because the uncertainty differs at different $PM_{2.5}$ concentrations (Burnett et al., 2014). Low uncertainty in baseline mortality rate was found for most source-receptor pairs (<20%) except for the response of $PM_{2.5}$ mortality in SAS to 20% reduction from RBU (66%).

3.4 Effect of sectoral reductions on mortality

Reducing global anthropogenic emissions by 20% in 3 sectors (i.e. PIN, TRN and RES) together avoids 48,500 (7,100, 108,000) O_3-related premature deaths and 243,000 (66,800, 357,000) $PM_{2.5}$-related premature deaths globally (Tables 5-6), with the greatest avoided air pollution-related premature deaths located in highly populated areas (e.g., North America, Europe, India, China, etc.) (Figs.~~1~~3-~~6~~4). For instance, reducing anthropogenic emissions by 20% in 3 sectors together avoids the highest number of O_3-related deaths in SAS (24,000 (6,000, 49,600) deaths/year) and $PM_{2.5}$-related deaths in EAS (83,400 (29,400, 135,000) deaths/year). We compare our estimates of O_3 and $PM_{2.5}$-related premature deaths attributable to PIN, TRN and RES emissions with previous studies, by multiplying our results for 20% emission reductions by 5, and by combining their sectors to nearly match each of the three sectors in this study (Table 7). Compared with Silva et al (2016a), our estimate of O_3 and $PM_{2.5}$-related premature deaths attributable to PIN and TRN are very comparable, but that to RES is lower here. In comparison with Lelieveld et al (2015), we estimate greater O_3 and $PM_{2.5}$-related premature deaths attributable to PIN and TRN, but less for RES.

Like Silva et al. (2016a) and Lelieveld et al. (2015), different locations show relatively different mortality responses to changes in sectoral emissions. Whereas PIN emission reductions cause the greatest number of avoided O_3-related premature deaths globally (19,300 (1,400, 45,000) deaths/year), TRN emission reductions cause the greatest fraction of avoided deaths in most of the six regions (26-53% of the global emission reduction), except for EAS (58%) and RBU (38%) where the effect of reducing PIN emissions dominates. In comparison with other studies (Table 7), our conclusion that PIN emissions cause the most O_3-related deaths and TRN emissions cause the greatest fraction of avoided deaths in most regions agrees well with Silva et al (2016a). For $PM_{2.5}$, reducing PIN emissions avoids the most $PM_{2.5}$-related premature deaths globally (128,000 (41,600, 179,000) deaths/year) and in most regions (38-78% of the global emission reduction), except for SAS (45%) where the RES emission dominates. Although these findings differ from those of Lelieveld et al (2015) and Silva

et al (2016), who find that Residential emissions have the greatest of impact on $PM_{2.5}$ mortality globally and in most regions, all studies agree that PIN emissions have the greatest impact in NAM. Our result is also comparable with Crippa et al (2017) who find that PIN emissions have the greatest health impact in most countries. Although comparable emission inventories are used (i.e. Lelieveld et al (2015) and this study use EDGAR emissions while Silva et al (2016) use RCP8.5~~.~~ emissions), our lower mortality estimate for RES emissions may be explained by our 20% reductions relative to the zero-out method, and the different years simulated.

Considering results from individual models, we found that mortality from TRN emission reductions show greater relative uncertainty than from PIN or RES (Table 5-6 and Table S9-S10), reflecting a greater spread of results across models. Regional impacts from individual model also differ from the ensemble mean result - e.g., for O_3, GEOSCHEMADJOINT and OsloCTM3.v2 show that reducing PIN emissions causes the greatest fraction of avoided O_3-related deaths in EUR, while GEOSCHEMADJOINT, HadGM2-ES and OsloCTM3.v2 show that TRN emissions have the greatest fraction of avoided O_3-related deaths in RBU (Figs. S~~8 and S12~~20). For $PM_{2.5}$, CHASER_t42 and GEOSCHEMADJOINT show that reducing PIN emissions causes the greatest fraction of avoided $PM_{2.5}$-related deaths in SAS (Figs. S~~9 and S13~~21).

4 Discussion

We aggregate the avoided deaths attributable to 20% reductions from four corresponding source regions (i.e. NAM, EUR, SAS and EAS), and compare with the findings from TF-HTAP1. We estimate that these regional emission reductions are associated with 36,000 (-1,500, 90,300) avoided deaths globally through the change in O_3 and 207,000 (41,500, 304,000) avoided deaths through the change in $PM_{2.5}$, more than those estimated by Anenberg et al. (2009 and 2014) – 21,800 (10,600, 33,400) deaths for O_3 and 192,000 (146,000, 230,000) deaths for $PM_{2.5}$. This discrepancy might be attributed to different health impact function, emissions data sets, region definitions, updated population or baseline mortality rates. In particular, for O_3 respiratory mortality, we use a log-linear model for chronic mortality (Jerrett et al 2009), instead of the short-term O_3 mortality estimate based on a daily time-series study (Bell et al., 2004) used by Anenberg et al., (2009). For $PM_{2.5}$ mortality, Anenberg et al., (2014) only included the simulated changes in BC, particulate organic matter (POM=primary organic aerosol+secondary organic aerosol), and sulfate for $PM_{2.5}$ concentration, while we use the total model reported $PM_{2.5}$ concentration which includes more species for some models. We also apply the Integrated Exposure–Response (IER) model (Burnett et al.

2014) for $PM_{2.5}$, as opposed to the log-linear model of Krewski et al. (2009) used by Anenberg et al., (2014).

For regional reductions, our multi-model average results suggest that NAM and EUR emissions cause more deaths inside of those regions than outside, which disagrees with previous studies (Anenberg et al., 2009; Duncan et al., 2008; West et al., 2009) whereas similar regional impacts are found for EAS and SAS. Also, total avoided deaths through interregional air pollution transport are estimated as 6,000 (-3,400, 15,500)~~10,300 (6,700, 13,400)~~ deaths/year for O_3 and 25,100 (8,200, 35,800)~~42,000 (12,400, 60,100)~~ deaths/year for $PM_{2.5}$ in this study, in contrast with 7,300 (3,600, 11,200) deaths/year for O_3 and 11,500 (8,800, 14,200) deaths/year for $PM_{2.5}$ in Anenberg et al. (2009; 2014). These differences likely result from different concentration-response functions and the use of 6 regions here vs. 4 by Anenberg et al. (2009; 2014). In addition, updated atmospheric models and emissions inputs, as well as different atmospheric dynamics in the single years chosen in TF-HTAP1 vs. TF-HTAP2 may contribute to the differences. In addition, updated atmospheric models and emissions inputs, as well as different atmospheric dynamics in the single years chosen in HTAP vs. HTAP2 may contribute to the differences. ~~These differences likely result from different concentration-response functions and the use of 6 regions here vs. 4 by Anenberg et al. (2009; 2014).~~ Overall, whereas O_3 accounts for a higher percentage of the total deaths in foreign regions than $PM_{2.5}$, $PM_{2.5}$ leads to more deaths in general, which agrees well with the results of Anenberg et al. (2009; 2014).

Using regional models in AQMEII3, driven by a single global model (C-IFS_v2), Im et al. (~~2017~~2018) estimated that 20% domestic emission reductions would avoid 54,000 and 27,500 premature deaths (for O_3 and $PM_{2.5}$ combined) in Europe and the U.S., respectively, as opposed to ~1,000 and 2,000 premature deaths due to foreign emission reductions. These results are comparable to our estimates that 32,900 and 19,500 premature deaths result from 20% domestic emission reductions in Europe and the U.S., while 670 and 570 premature deaths result from foreign emission reductions. Although our defined U.S. region is slightly bigger than Im et al. (~~2017~~2018), the majority of U.S. emission sources and population are located within the region defined by Im et al. (~~2017~~2018). This comparison shows that regional and global models show similar impacts on mortality from air pollution transport.

Differences in our estimates of premature mortality attributable to air pollution from three emission sectors (multiplied by 5) may be explained by methodological differences relative to previous studies (Silva et al., 2016; Lelieveld et al., 2015), including our use of 20% emission reductions versus the zero-out method in those studies, different emission inventories, a multi-model ensemble versus single models, and differences in baseline mortality rates, population, and concentration response

functions. Our finding that TRN emissions contribute the most avoided deaths for O_3 in most regions agrees well with the result by Silva et al (2016a), but differs for $PM_{2.5}$ mortality for which we find that PIN emissions cause the most deaths, while both Silva et al (2016a) and Lelieveld et al (2015) find that RES emissions are responsible for the most deaths. This discrepancy may be explained by different $PM_{2.5}$ species included in individual models, as we showed that changes in $PM_{2.5}$ concentration to TRN emission differ across models.

By using an ensemble of multi-model results here, we highlight the relative importance of difference source-receptor pairs for mortality in a way that is more robust than using a single model, particularly since some individual models yielded different conclusions than the ensemble mean. The air pollutant concentration changes reported by the HTAP2 models may be different among models, it may result from variety of processes, e.g. atmospheric physical and chemical mechanisms, processing of natural emissions, and transport time step, etc. (Table S1), but not anthropogenic emissions since those were nearly identical among models. In addition, the coarse model resolution used by global models may underestimate health effects by misaligning peak concentration and population, particularly in urban areas and for $PM_{2.5}$ (Punger and West, 2013), but it is not known how model resolution would affect the relative contributions of extra-regional and intraregional health benefits. Future research should explore the possible bias from using coarse global models for extra-regional and intraregional mortality estimates in metropolitan regions by comparing with finer-resolution chemical transport models.

Another uncertainty in this paper (and other global studies) lies in applying the same RRs worldwide, because of lack of long-term records of the chronic influences of ambient air pollution on mortality outside of North America and Europe. We consider only the population of adults ≥25 years old, ignoring possible mortality effects on the younger population, and consequently we may underestimate premature mortality overall. Likewise, the effects of air pollution on several morbidity endpoints are omitted. We assume that all $PM_{2.5}$ is equally toxic, for lack of clear evidence for greater toxicity of some species. Inter-regional transport may also change the toxicity of $PM_{2.5}$ by changing the size distribution or chemical composition, where transport likely causes particles to become more oxidized (West et al., 2016). Future research on $PM_{2.5}$-related mortality should include estimating health effects for different $PM_{2.5}$ chemical components.

5 Conclusions

We estimate O_3- and $PM_{2.5}$-related premature mortality from simulations with 14

global CTMs participating in the TF-HTAP2 multi-model exercise for the year 2010. An estimate of 290,000 (30,000, 600,000) global premature O_3-related deaths and 2.8 million (0.5 million, 4.6 million) global $PM_{2.5}$-related premature deaths is obtained from the ensemble for the year 2010 in the baseline case. We focus on model experiments simulating 20% regional air pollutant emission reductions (excluding methane) in 6 regions, 3 sectors and 1 global domain. For regional scenarios, 6 source emission reductions altogether can cause 84% of the global avoided O_3-related premature deaths within the source region, ranging from 21 to 95% among 6 regions, and 16% (5 to 79%) outside of the source region. For $PM_{2.5}$, 89% of global avoided $PM_{2.5}$-related premature deaths are within the source region, ranging from 32 to 94% among 6 regions, and 11% (6 to 68%) outside of the source region. While most avoided mortality generally occurs within the source region, we find that emission reductions from RBU (only for O_3) and MDE (for both O_3 and $PM_{2.5}$) can avoid more premature deaths outside of these regions than within. Considering the effects of foreign emissions on receptor regions, 20% foreign emission reductions lead to more avoided O_3-related premature deaths in EUR, MDE and RBU than domestic reductions. Reductions from all six regions in the transport of air pollution between regions are estimated to lead to more avoided deaths through changes in $PM_{2.5}$ (25,100 (8,200, 35,800)~~42,000 (12,400, 60,100)~~ deaths/year) than for O_3 (6,000 (-3,400, 15,500)~~10,300 (6,700, 13,400)~~ deaths/year). For NAM and EUR, our estimates of avoided mortality from regional and extra-regional emission reductions are comparable to those estimated by regional models in AQMEII3 (Im et al., 2018) for these same emission reduction experiments. Overall, the spread of modeled air pollutant concentrations contributes most to the uncertainty in mortality estimates, highlighting that using a single model may lead to erroneous conclusions and may underestimate uncertainty in mortality estimates.

For sectoral emission reductions, reducing anthropogenic emissions by 20% in 3 sectors together avoids 48,500 (7,100, 108,000) O_3-related premature deaths and 243,000 (66,800, 357,000) $PM_{2.5}$-related premature deaths globally. Of the 3 sectors, TRN had the greatest fraction (26-53%) of O_3-related premature deaths globally and in most regions, except for EAS (58%) and RBU (38%) where PIN emissions dominate. For $PM_{2.5}$ mortality, PIN emissions cause the most deaths in most regions (38-78%), except for SAS (45%) where the TRN emissions dominate.

In this study, we have gone beyond previous TF-HTAP1 studies that quantified premature mortality from interregional air pollution transport, by using more source regions, analyzing source emission sectors, and using updated atmospheric models and health impact functions. The estimate of air transport premature mortality could vary due to differences in exposure estimate (single model vs ensemble model), health impact function, regional definitions, and grid resolutions. These discrepancies

highlight uncertainty estimated by different methods in previous studies. Despite uncertainties, our results suggest that reducing pollution transported over a long distance would be beneficial for health, with impacts from all foreign emission reductions combined that may be comparable to or even exceed the impacts of emission reductions within a region. Additionally, actions to reduce emissions should target specific sectors within world regions, as different sectors dominate the health effects in different regions. This work highlights the importance of long-range air pollution transport, and suggests that estimates of the health benefits of emission reductions on local, national, or continental scales may underestimate the overall health benefits globally, when interregional transport is accounted for. International cooperation to reduce air pollution transported over long distances may therefore be desirable.

Acknowledgments We sincerely acknowledge the contribution of modeling groups from the second phase of Task Force on Hemispheric Transport of Air Pollution (TF-HTAP2). This work was supported by a scholarship from the Taiwan Ministry of Education, grants from NIEHS (1 R21 ES022600-01), and NASA (NNX16AQ30G and NNX16AQ26G), funding from BEIS under the Hadley Centre Climate Programme contract (GA01101) and from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 641816 (CRESCENDO). The National Center for Atmospheric Research is sponsored by the National Science Foundation. We thank Dr. Owen Cooper who provided the TOAR ground level O_3 observation dataset, and Michael Brauer for the GBD2013 ground level $PM_{2.5}$ observation dataset.

Supporting information A detailed description of the models participating in the ensemble, a map of six priority regions used in this analysis, and additional results can be found in the Supporting Information.

References

Anenberg, S.C., West, J.J., Fiore, A.M., Jaffe, D.A., Prather, M.J., Bergmann, D., Cuvelier, K., Dentener, F.J., Duncan, B.N., Gauss, M., Hess, P., Jonson, J.E., Lupu, A., MacKenzie, I.A., Marmer, E., Park, R.J., Sanderson, M.G., Schultz, M., Shindell, D.T., Szopa, S., Vivanco, M.G., Wild, O., Zeng, G.: Intercontinental Impacts of Ozone Pollution on Human Mortality. Environmental Science & Technology 43, 6482-6487, doi: 10.1021/es900518z, 2009.

Anenberg, S.C., West, J.J., Yu, H., Chin, M., Schulz, M., Bergmann, D., Bey, I., Bian, H., Diehl, T., Fiore, A., Hess, P., Marmer, E., Montanaro, V., Park, R., Shindell, D., Takemura, T., Dentener, F.: Impacts of intercontinental transport of anthropogenic

826 fine particulate matter on human mortality. Air Quality, Atmosphere & Health 7, 369-
827 379, https://doi.org/10.1007/s11869-014-0248-9, 2014
828 Barrett, S.R., Britter, R.E., Waitz, I.A.: Global mortality attributable to aircraft
829 cruise emissions. Environ. Sci. Technol 44, 7736-7742, doi: 10.1021/es101325, 2010
830 Bell, M. L., Ebisu, K., Leaderer, B. P., Gent, J. F., Lee, H. J., Koutrakis, P. et al.:
831 Associations of PM2.5 constituents and sources with hospital admissions: analysis of
832 four counties in Connecticut and Massachusetts (USA) for persons≥ 65 years of age.
833 Environmental health perspectives, 122, 138. doi: 10.1289/ehp.1306656, 2014.
834 Bell, M.L., Dominici, F., Samet, J.M.: A meta-analysis of time-series studies of
835 ozone and mortality with comparison to the national morbidity, mortality, and air
836 pollution study. Epidemiology 16, 436-445, 2005.
837 Bell, M.L., Zanobetti, A.F., Dominici, F.: Who is more affected by ozone pollution?
838 A systematic review and meta-analysis. Am J Epidemiol 180, 15-28, doi:
839 10.1093/aje/kwu115, 2014
840 Bhalla, K., Shotten, M., Cohen, A., Brauer, M., Shahraz, S., Burnett, R. et al.:
841 Transport for health: the global burden of disease from motorized road transport., 2014
842 Bright, E. A., Coleman, P. R., Rose, A. N., and Urban, M. L.: Land-Scan 2011,
843 Oak Ridge National Laboratory SE, Oak Ridge, TN, 2012.
844 Brook, R.D., Rajagopalan, S., Pope, C.A., 3rd, Brook, J.R., Bhatnagar, A., Diez-
845 Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A.,
846 Siscovick, D., Smith, S.C., Jr., Whitsel, L., Kaufman, J.D., American Heart Association
847 Council on, E., Prevention, C.o.t.K.i.C.D., Council on Nutrition, P.A., Metabolism:
848 Particulate matter air pollution and cardiovascular disease: An update to the scientific
849 statement from the American Heart Association. Circulation 121, 2331-2378, doi:
850 10.1161/CIR.0b013e3181dbece1, 2010.
851 Burnett, R. T., Pope III, C. A., Ezzati, M., Olives, C., Lim, S. S., Mehta, S. et al.:
852 An integrated risk function for estimating the global burden of disease attributable to
853 ambient fine particulate matter exposure. Environmental health perspectives, 122, 397,
854 doi: 10.1289/ehp.1307049, 2014.
855 Cohen, Aaron J., et al.: "Estimates and 25-year trends of the global burden of
856 disease attributable to ambient air pollution: an analysis of data from the Global Burden
857 of Diseases Study 2015." The Lancet 389.10082: 1907-1918, DOI:
858 https://doi.org/10.1016/S0140-6736(17)30505-6, 2017.
859 Chafe, Z.A., Brauer, M., Klimont, Z., Van Dingenen, R., Mehta, S., Rao, S., Riahi,
860 K., Dentener, F., Smith, K.R.: Household cooking with solid fuels contributes to
861 ambient $PM_{2.5}$ air pollution and the burden of disease. Environmental health
862 perspectives 122, 1314, doi:10.1289/ehp.1206340, 2014.
863 Chambliss, S., Silva, R., West, J., Zeinali, M., Minjares, R.: Estimating source-

attributable health impacts of ambient fine particulate matter exposure: global premature mortality from surface transportation emissions in 2005. Environmental Research Letters 9, 104009, 2014.

Corbett, J.J., Winebrake, J.J., Green, E.H., Kasibhatla, P., Eyring, V., Lauer, A.: Mortality from ship emissions: a global assessment. Environ. Sci. Technol 41, 8512-8518, 2007.

Crippa, M., Janssens-Maenhout, G., Guizzardi, D., Van Dingenen, R., Dentener, F.: Sectorial and regional uncertainty analysis of the contribution of anthropogenic emissions to regional and global PM2.5 health impacts. Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-779, in review, 2017.

Dong, X., Fu, J. S., Zhu, Q., Sun, J., Tan, J., Keating, T., Sekiya, T., Sudo, K., Emmons, L., Tilmes, S., Jonson, J. E., Schulz, M., Bian, H., Chin, M., Davila, Y., Henze, D., Takemura, T., Benedictow, A. M. K., and Huang, K.: Long-range Transport Impacts on Surface Aerosol Concentrations and the Contributions to Haze Events in China: an HTAP2 Multi-Model Study, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-91, in review, 2018.

Dentener, F., Keating, T., and Akimoto, H. (Eds.): HTAP Hemispheric Transport of Air Pollution, Part A: Ozone and particulate matter, United Nations Publications, Geneva, Switzerland, 2010.

Du, Y., Xu, X., Chu, M., Guo, Y., Wang, J.: Air particulate matter and cardiovascular disease: the epidemiological, biomedical and clinical evidence. Journal of thoracic disease, 8, E8. doi: 10.3978/j.issn.2072-1439.2015.11.37, 2016.

Doherty, R.M., Wild, O., Shindell, D.T., Zeng, G., MacKenzie, I.A., Collins, W.J., Fiore, A.M., Stevenson, D.S., Dentener, F.J., Schultz, M.G., Hess, P., Derwent, R.G., Keating, T.J.: Impacts of climate change on surface ozone and intercontinental ozone pollution: A multi-model study. J. Geophys. Res. Atmos., 118, 3744–3763, doi:10.1002/jgrd.50266, 2013.

Duncan, B.N., West, J.J., Yoshida, Y., Fiore, A.M., Ziemke, J.R.: The influence of European pollution on ozone in the Near East and northern Africa. Atmos. Chem. Phys., 8, 2267-2283, https://doi.org/10.5194/acp-8-2267-2008, 2008.

Fiore, A., Dentener, F.J., Wild, O., Cuvelier, C., Schultz, M.G., Hess, P., Textor, C., Schulz, M., Doherty, R.M., Horowitz, L.W.: Multimodel estimates of intercontinental source-receptor relationships for ozone pollution. J. Geophys. Res., 114, D04301, doi:10.1029/2008JD010816m, 2009.

Fry, M.M., Naik, V., West, J.J., Schwarzkopf, M.D., Fiore, A.M., Collins, W.J., Dentener, F.J., Shindell, D.T., Atherton, C., Bergmann, D.: The influence of ozone precursor emissions from four world regions on tropospheric composition and radiative climate forcing. J. Geophys. Res., 117, D07306, doi:10.1029/2011JD017134, 2012.

Forouzanfar, Mohammad H 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. Lancet2016; 388: 1659–724, DOI: https://doi.org/10.1016/S0140-6736(16)31679-8, 2016.

Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., Schulz, M., Benedictow, A., Griesfeller, J. J., Janssens-Maenhout, G., Carmichael, G., Fu, J., and Dentener, F.: Technical note: Coordination and harmonization of the multi-scale, multi-model activities HTAP2, AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats, Atmos. Chem. Phys., 17, 1543-1555, https://doi.org/10.5194/acp-17-1543-2017, 2017.~~Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., Schulz, M., Benedictow, A., Griesfeller, J.J., Janssens-Maenhout, G., Carmichael, G., Fu, J., Dentener, F.: Technical note: Harmonization of the multi-scale multi-model activities HTAP, AQMEII and MICS-Asia: simulations, emission inventories, boundary conditions and output formats. Atmos. Chem. Phys., 1-20, 1543-1555, https://doi.org/10.5194/acp-17-1543-2017, 2016.~~

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., Dortbudak, Z.: Acute effects of ozone on mortality from the "air pollution and health: a European approach" project. Am J Respir Crit Care Med 170, 1080-1087, 2004.

Huang, M., Carmichael, G. R., Pierce, R. B., Jo, D. S., Park, R. J., Flemming, J., Emmons, L. K., Bowman, K. W., Henze, D. K., Davila, Y., Sudo, K., Jonson, J. E., Tronstad Lund, M., Janssens-Maenhout, G., Dentener, F. J., Keating, T. J., Oetjen, H., and Payne, V. H.: Impact of intercontinental pollution transport on North American ozone air pollution: an HTAP phase 2 multi-model study, Atmos. Chem. Phys., 17, 5721-5750, https://doi.org/10.5194/acp-17-5721-2017, 2017.

Hamra, G.B., Guha, N., Cohen, A., Laden, F., Raaschou-Nielsen, O., Samet, J.M., Vineis, P., Forastiere, F., Saldiva, P., Yorifuji, T., Loomis, D.: Outdoor particulate matter exposure and lung cancer: a systematic review and meta-analysis. Environ Health Perspect 122, 906-911, doi:10.1289/ehp.1408092, 2014.

Heald, C.L., Jacob, D.J., Park, R.J., Alexander, B., Fairlie, T.D., Yantosca, R.M., Chu, D.A.: Transpacific transport of Asian anthropogenic aerosols and its impact on surface air quality in the United States. J. Geophys. Res., 111, D14310, doi:10.1029/2005JD006847, 2006.

Institute for Health Metrics and Evaluation (IHME): Global Burden of Disease Study 2010 (GBD 2010) Results by Cause 1990-2010 - Country Level. Seattle, United

States, 2013.

~~Task Force Hemispheric Transport of Air pollution (TF-HTAP): Part A, Ozone and Particulate Matter, F. Dentener, T. Keating and H. Akimoto (eds) Economic Commission for Europe, Air Pollution Studies, 17, ISBN, 978-92-1-117043-6, UNECE, Geneva, 2010.~~Huang, M., Carmichael, G.R., Pierce, R.B., Jo, D.S., Park, R.J., Flemming, J., Emmons, L.K., Bowman, K.W., Henze, D.K., Davila, Y., Sudo, K., Jonson, J.E., Lund, M.T., Janssens-Maenhout, G., Dentener, F.J., Keating, T.J., Oetjen, H., Payne, V.H.: Impact of Intercontinental Pollution Transport on North American Ozone Air Pollution: An HTAP Phase II Multi-model Study. Atmos. Chem. Phys., 17, 5721-5750, https://doi.org/10.5194/acp-17-5721-2017, 2017.

Im, U., Brandt, J., Geels, C., Hansen, K. M., Christensen, J. H., Andersen, M. S., Solazzo, E., Kioutsioukis, I., Alyuz, U., Balzarini, A., Baro, R., Bellasio, R., Bianconi, R., Bieser, J., Colette, A., Curci, G., Farrow, A., Flemming, J., Fraser, A., Jimenez-Guerrero, P., Kitwiroon, N., Liang, C.-K., Nopmongcol, U., Pirovano, G., Pozzoli, L., Prank, M., Rose, R., Sokhi, R., Tuccella, P., Unal, A., Vivanco, M. G., West, J., Yarwood, G., Hogrefe, C., and Galmarini, S.: Assessment and economic valuation of air pollution impacts on human health over Europe and the United States as calculated by a multi-model ensemble in the framework of AQMEII3, Atmos. Chem. Phys., 18, 5967-5989, https://doi.org/10.5194/acp-18-5967-2018, 2018.~~Im, U., Brandt, J., Geels, C., Hansen, K. M., Christensen, J. H., Andersen, M. S., Solazzo, E., Kioutsioukis, I., Alyuz, U., Balzarini, A., Baro, R., Bellasio, R., Bianconi, R., Bieser, J., Colette, A., Curci, G., Farrow, A., Flemming, J., Fraser, A., Jimenez-Guerrero, P., Kitwiroon, N., Liang, C.-K., Pirovano, G., Pozzoli, L., Prank, M., Rose, R., Sokhi, R., Tuccella, P., Unal, A., Vivanco, M. G., West, J., Yarwood, G., Hogrefe, C., and Galmarini, S.: Assessment and economic valuation of air pollution impacts on human health over Europe and the United States as calculated by a multi-model ensemble in the frame work of AQMEII3, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-751, in review, 2017.

Ito, K., De Leon, S.F., Lippmann, M.: Associations between Ozone and Daily Mortality: Analysis and Meta-Analysis. Epidemiology 16, 446-457, 2005.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier van der Gon, H., Kuenen, J.J.P., Klimont, Z., Frost, G., Darras, S., Koffi, B., Li, M.: HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution. Atmos. Chem. Phys., 15, 11411-11432, https://doi.org/10.5194/acp-15-11411-2015, 2015.

Jerrett, M., Burnett, R.T., Pope, C.A., 3rd, Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E., Thun, M.: Long-term ozone exposure and mortality. N Engl J Med 360, 1085-1095, doi: 10.1056/NEJMoa0803894, 2009.

978 Krewski, D., Jerrett, M., Burnett, R.T., Ma, R., Hughes, E., Shi, Y., Turner, M.C.,
979 Pope, C.A., 3rd, Thurston, G., Calle, E.E., Thun, M.J., Beckerman, B., DeLuca, P.,
980 Finkelstein, N., Ito, K., Moore, D.K., Newbold, K.B., Ramsay, T., Ross, Z., Shin, H.,
981 Tempalski, B.: Extended follow-up and spatial analysis of the American Cancer Society
982 study linking particulate air pollution and mortality. Res Rep Health Eff Inst, 5-114;
983 discussion 115-136, 2009.
984 Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., Pozzer, A.: The contribution
985 of outdoor air pollution sources to premature mortality on a global scale. Nature, 525,
986 367-371, doi:10.1038/nature15371, 2015.
987 Leibensperger, E.M., Mickley, L.J., Jacob, D.J., Barrett, S.R.H.: Intercontinental
988 influence of NOx and CO emissions on particulate matter air quality. Atmospheric
989 Environment 45, 3318-3324, https://doi.org/10.1016/j.atmosenv.2011.02.023, 2011.
990 Lepeule, J., Laden, F., Dockery, D., Schwartz, J.: Chronic exposure to fine
991 particles and mortality: an extended follow-up of the Harvard Six Cities study from
992 1974 to 2009. Environ Health Perspect 120, 965-970, doi: 10.1289/ehp.1104660, 2012.
993 Levy, J.I., Chemerynski, S.M., Sarnat, J.A.: Ozone exposure and mortality: an
994 empiric bayes metaregression analysis. Epidemiology 16, 458-468, 2005.
995 Lim, S.S., Vos, T., Flaxman, A.D., Danaei, G., Shibuya, K., Adair-Rohani, H.,
996 Amann, M., Anderson, H.R. et al.: A comparative risk assessment of burden of disease
997 and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-
998 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380,
999 2224-2260. doi: http://dx.doi.org/10.1016/S0140-6736(12)61766-8, 2012.
1000 Lin, M., Fiore, A.M., Horowitz, L.W., Cooper, O.R., Naik, V., Holloway, J.,
1001 Johnson, B.J., Middlebrook, A.M., Oltmans, S.J., Pollack, I.B.: Transport of Asian
1002 ozone pollution into surface air over the western United States in spring. J. Geophys.
1003 Res., 117, D00V07, doi:10.1029/2011JD016961, 2012.
1004 Lin, M., Horowitz, L.W., Payton, R., Fiore, A.M., Tonnesen, G.: US surface ozone
1005 trends and extremes from 1980 to 2014: quantifying the roles of rising Asian emissions,
1006 domestic controls, wildfires, and climate. Atmos. Chem. Phys., 17, 2943-2970,
1007 https://doi.org/10.5194/acp-17-2943-2017, 2017.
1008 Liu, J., Mauzerall, D.L., Horowitz, L.W.: Evaluating inter-continental transport of
1009 fine aerosols:(2) Global health impact. Atmospheric Environment 43, 4339-4347, doi:
1010 10.1016/j.atmosenv.2009.05.032, 2009a.
1011 Liu, J., Mauzerall, D.L., Horowitz, L.W., Ginoux, P., Fiore, A.M.: Evaluating inter-
1012 continental transport of fine aerosols: (1) Methodology, global aerosol distribution and
1013 optical depth. Atmospheric Environment 43, 4327-4338,
1014 https://doi.org/10.1016/j.atmosenv.2009.03.054, 2009b.
1015 Malley, C., Vallack, H. W., Ashmore, M. R., Kuylenstierna, J. C. I., Henze, D.,

Davila, Y., et al: Updated Global Estimates of Respiratory Mortality in Adults ≥30 Years of Age Attributable to Long-Term Ozone Exposure. Environmental health perspectives, 125, 1390, doi: 10.1289/ehp.0901220, 2017.

Powell, H., Krall, J. R., Wang, Y., Bell, M. L., and Peng, R. D.: Ambient coarse particulate matter and hospital admissions in the Medicare Cohort Air Pollution Study, 1999–2010. Environmental health perspectives, 123, 1152. doi: 10.1289/ehp.1408720, 2015.

Pope III, C. A., Burnett, R. T., Turner, M. C., Cohen, A., Krewski, D., Jerrett, M. et al.: Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure–response relationships. Environmental health perspectives, 119, 1616. doi: 10.1289/ehp.1103639, 2011.

Punger, E.M., West, J.J.: The effect of grid resolution on estimates of the burden of ozone and fine particulate matter on premature mortality in the USA. Air Quality, Atmosphere & Health 6, 563-573, doi:10.1007/s11869-013-0197-8, 2013.

Sanderson, M.G., Dentener, F.J., Fiore, A.M., Cuvelier, C., Keating, T.J., Zuber, A., Atherton, C.S., Bergmann, D.J., Diehl, T., Doherty, R.M., Duncan, B.N., Hess, P., Horowitz, L.W., Jacob, D.J., Jonson, J.E., Kaminski, J.W., Lupu, A., MacKenzie, I.A., Mancini, E., Marmer, E., Park, R., Pitari, G., Prather, M.J., Pringle, K.J., Schroeder, S., Schultz, M.G., Shindell, D.T., Szopa, S., Wild, O., Wind, P.: A multi-model study of the hemispheric transport and deposition of oxidised nitrogen. Geophysical Research Letters 35, L17815, doi:10.1029/2008GL035389, 2008.

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., Zeng, G.: Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change. Environmental Research Letters 8, doi: 10.1088/1748-9326/8/3/034005, 2013.

Silva, R.A., Adelman, Z., Fry, M.M., West, J.J.: The Impact of Individual Anthropogenic Emissions Sectors on the Global Burden of Human Mortality due to Ambient Air Pollution. Environ. Health Perspect., Environ. Health Perspect. 124:1776–1784, 2016a.

Silva, R. A., West, J. J., Lamarque, J.-F., Shindell, D.T., et al.: The effect of future ambient air pollution on human premature mortality to 2100 using output from the ACCMIP model ensemble, Atmos. Chem. Phys., 16, 9847-9862, https://doi.org/10.5194/acp-16-9847-2016, 2016b.

Stieb, D.M., Szyszkowicz, M., Rowe, B.H., Leech, J.A.: Air pollution and emergency department visits for cardiac and respiratory conditions: a multi-city time-

series analysis. Environmental Health 8, 25, doi: 10.1186/1476-069X-8-25, 2009.

Stjern, C.W., Samset, B.H., Myhre, G., Bian, H., Chin, M., Davila, Y., Dentener, F., Emmons, L., Flemming, J., Haslerud, A.S., Henze, D., Jonson, J.E., Kucsera, T., Lund, M.T., Schulz, M., Sudo, K., Takemura, T., Tilmes, S.: Global and regional radiative forcing from 20 % reductions in BC, OC and SO4 – an HTAP2 multi-model study. Atmos. Chem. Phys., 16, 13579-13599, https://doi.org/10.5194/acp-16-13579-2016, 2016.

Schultz MG, Schröder S, Lyapina O, Cooper O, Galbally I, Petropavlovskikh I, et al.: Tropospheric Ozone Assessment Report: Database and Metrics Data of Global Surface Ozone Observations. Elem Sci Anth.,5:58, DOI:http://doi.org/10.1525/elementa.244, 2017

Turner, M. C., Jerrett, M., Pope, C. A., Krewski, D., Gapstur, S. M., Diver, W. R., et al. :~~(2016)~~ Long-Term Ozone Exposure and Mortality in a Large Prospective Study. American Journal of Respiratory and Critical Care Medicine, 193, 1134–1142. http://doi.org/10.1164/rccm.201508-1633OC, 2016.

West, J. J., Naik, V., Horowitz, L. W., and Fiore, A. M.: Effect of regional precursor emission controls on long-range ozone transport – Part 1: Short-term changes in ozone air quality, Atmos. Chem. Phys., 9, 6077-6093, https://doi.org/10.5194/acp-9-6077-2009, 2009a.

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., 9, 6095-6107, https://doi.org/10.5194/acp-9-6095-2009, 2009b.

West, J. J., A. Cohen, F. Dentener, B. Burnekreef, T. Zhu, B. Armstrong, M. L. et al.: What we breathe impacts our health: improving understanding of the link between air pollution and health, Environmental Science & Technology, 50: 4895-4904, doi: 10.1021/acs.est.5b03827, 2016.

Wild, O., Akimoto, H.: Intercontinental transport of ozone and its precursors in a three-dimensional global CTM. J. Geophys. Res., 106, 27729–27744, doi:10.1029/2000JD000123, 2001.

~~Yi, K., Liu, J., Ban-Weiss, G., Zhang, J., Tao, W., Cheng, Y., Tao, S.: Response of the global surface ozone distribution to Northern Hemisphere sea surface temperature changes: implications for long-range transport. Atmos. Chem. Phys., 17, 8771-8788, https://doi.org/10.5194/acp-17-8771-2017, 2017.~~Yu, H., Chin, M., West, J.J., Atherton, C.S., Bellouin, N., Bergmann, D., Bey, I., Bian, H., Diehl, T., Forberth, G.: A multimodel assessment of the influence of regional anthropogenic emission reductions on aerosol direct radiative forcing and the role of intercontinental transport. Journal of Geophysical Research: Atmospheres 118, 700-720, doi: 10.1029/2012JD018148, 2013.

Yu, H., L. A. Remer, M. Chin, H. Bian, R. G. et al.: A satellite-based assessment of transpacific transport of pollution aerosol, J. Geophys. Res., 113, D14S12, doi:10.1029/2007JD009349, 2008.

Young PJ, Naik V, Fiore AM, Gaudel A, Guo J, Lin MY, et al.: Tropospheric Ozone Assessment Report: Assessment of global-scale model performance for global and regional ozone distributions, variability, and trends. Elem Sci Anth. 6(1):10, DOI: http://doi.org/10.1525/elementa.265, 2018.Zhang, Q., Jiang, X., Tong, D., Davis, S.J., Zhao, H., Geng, G., Feng, T., Zheng, B., Lu, Z., Streets, D.G.: Transboundary health impacts of transported global air pollution and international trade. Nature 543, 705-709, doi:10.1038/nature21712, 2017.

Table 1. Population-weighted multi-model mean O_3 (ppb) and $PM_{2.5}$ concentration ($\mu g/m^3$) for the 2010 baseline, for the 6-month O_3 season average of 1-hr. daily maximum O_3 and annual average $PM_{2.5}$, shown with the standard deviation among models.

Scenarios	Receptor regions						
	NAM	EUR	SAS	EAS	MDE	RBU	World
O_3 (11 models)	56.51±9.40	48.38±8.05	65.72±10.08	59.10±10.46	61.11±9.79	46.79±7.53	53.74±8.03
$PM_{2.5}$ (8 models)	9.36±2.62	10.75±3.87	37.05±8.74	39.27±13.50	34.49±17.64	11.61±3.52	25.98±5.05

Table 2. Population-weighted multi-model mean change in O_3 (ppb) in receptor regions due to 20% regional (NAM, EUR, SAS, MDE and RBU), sectoral (PIN, TRN and RES) and global (GLO) anthropogenic emission reductions, for the 6-month O_3 season average of 1-hr. daily maximum. The diagonal, showing the effect of each region on itself, is underlined. All numbers are rounded to the nearest hundredth, and are shown with standard deviations among models.

Source	Receptor region						
regions/sectors	NAM	EUR	SAS	EAS	MDE	RBU	World
NAM	<u>-1.88±0.06</u>	-0.26±0.55	-0.04±0.14	-0.11±0.06	-0.23±0.12	-0.21±0.09	-0.19±0.07
EUR	-0.08±0.04	<u>-0.80±0.55</u>	0.01±0.14	-0.10±0.06	-0.34±0.08	-0.34±0.09	-0.14±0.07
SAS	-0.05±0.02	-0.04±0.02	<u>-3.65±0.94</u>	-0.08±0.04	-0.11±0.04	-0.04±0.03	-0.90±0.22
EAS	-0.29±0.14	-0.25±0.13	-0.09±0.22	<u>-1.96±1.10</u>	-0.23±0.12	-0.27±0.12	-0.58±0.25
MDE	-0.04±0.02	-0.05±0.01	-0.07±0.15	-0.03±0.01	<u>-1.23±0.66</u>	-0.11±0.01	-0.09±0.04
RBU	-0.05±0.04	-0.13±0.05	0.03±0.16	-0.08±0.06	-0.10±0.07	<u>-0.45±0.38</u>	-0.05±0.06
PIN	-1.13±0.28	-0.70±0.19	-1.43±0.18	-1.58±0.88	-1.09±0.45	-0.69±0.31	-1.11±0.25
TRN	-1.26±0.42	-0.81±0.34	-2.05±0.32	-0.73±0.32	-1.40±0.17	-0.71±0.19	-1.13±0.19
RES	-0.24±0.09	-0.21±0.04	-1.19±0.44	-0.62±0.10	-0.23±0.06	-0.18±0.03	-0.57±0.14
GLO	-2.86±0.77	-1.98±0.66	-4.40±1.04	-2.77±1.21	-2.84±0.70	-1.76±0.52	<u>-2.82±0.53</u>

Table 3. Population-weighted multi-model annual average change in $PM_{2.5}$ concentrations ($\mu g/m^3$) in receptor regions due to 20% regional (NAM, EUR, SAS, MDE and RBU), sectoral (PIN, TRN and RES) and global (GLO) anthropogenic emission reductions. The diagonal, showing the effect of each region on itself, is underlined. All numbers are rounded to the nearest hundredth, and are shown with standard deviations among models.

Source	Receptor region						
regions/sectors	NAM	EUR	SAS	EAS	MDE	RBU	World
NAM	<u>-1.33±0.66</u>	-0.03±0.02	0.00±0.01	-0.02±0.02	-0.01±0.01	-0.01±0.01	-0.08±0.04
EUR	-0.01±0.00	<u>-1.17±0.87</u>	-0.01±0.01	-0.02±0.01	-0.18±0.08	-0.26±0.19	-0.13±0.09
SAS	<-0.01	<-0.01	<u>-4.86±2.17</u>	-0.08±0.08	-0.03±0.02	<-0.01	-1.16±0.51
EAS	-0.03±0.01	-0.02±0.01	-0.08±0.07	<u>-6.19±3.08</u>	<-0.01	-0.04±0.02	-1.45±0.71
MDE	<-0.01	-0.03±0.01	-0.12±0.06	-0.01±0.02	<u>-0.91±0.38</u>	-0.05±0.03	-0.08±0.03
RBU	<-0.01	-0.07±0.05	-0.01±0.02	-0.04±0.02	-0.03±0.02	<u>-0.78±0.50</u>	-0.05±0.03
PIN	-0.61±0.18	-0.57±0.26	-1.73±0.71	-2.75±0.99	-0.92±0.14	-0.58±0.19	-1.46±0.56
TRN	-0.27±0.20	-0.38±0.41	-0.82±0.88	-0.54±0.43	-0.09±0.06	-0.15±0.16	-0.40±0.37
RES	-0.20±0.05	-0.27±0.12	-1.93±0.40	-1.70±0.28	-0.08±0.02	-0.20±0.05	-1.17±0.31
GLO	-1.47±0.72	-1.52±1.04	-5.40±2.31	-6.76±3.29	-1.55±0.75	-1.19±0.73	<u>-3.49±1.51</u>

Table 4. Annual multi-model empirical mean O_3- and $PM_{2.5}$-related premature deaths with 95% CI from Monte-Carlo simulations in parenthesis (including uncertainty in baseline mortality rates, RRs and air pollutant concentration across models) in year 2010 baseline. All numbers are rounded to three significant figures or the nearest 100 deaths. Empirical mean is the mean of 1,000 Monte Carlo simulations.

	Receptor region						
	NAM	EUR	SAS	EAS	MDE	RBU	World
O_3 (11 models)	15,000 (900－30,000)	13,000 (600－28,000)	136,000 (23,000－277,000)	100,000 (3,900－213,000)	3,200 (300－7,000)	2,900 (100－6,600)	291,000 (30,000－596,000)
$PM_{2.5}$ (8 models)	72,000 (1,500－158,000)	203,000 (2,700－463,000)	732,000 (328,000－1,110,000)	1,120,000 (159,000－1,720,000)	79,000 (600－133,000)	177,000 (2,700－358,000)	2,770,000 (514,000－4,640,000)

Table 5. Annual avoided multi-model empirical mean O_3-related premature respiratory deaths with 95% CI from Monte-Carlo simulations in parenthesis due to 20 % regional (NAM, EUR, SAS, MDE and RBU), sectoral (PIN, TRN and RES) and global (GLO) anthropogenic emission reductions in each region and worldwide. The diagonal, showing the effect of each region on itself, is underlined. For regional reductions, we also the RERER (eq. 4) as the percent of total avoided deaths in each receptor region that result from foreign emission reductions, as well as the percent of global avoided deaths from emission reductions in each source region. All numbers are rounded to three significant figures or the nearest 10 deaths.

Source regions/sectors	Receptor region NAM	EUR	SAS	EAS	MDE	RBU	World	Impact on foreign receptor regions
NAM	1,500 (-170－4,000)	330 (10－780)	170 (-250－690)	500 (-910－2,200)	30 (0－80)	70 (0－170)	2,800 (-1,300－8,400)	46%
EUR	60 (-80－240)	930 (-70－2,400)	-80 (-880－670)	490 (-1,100－2,300)	50 (10－110)	110 (10－250)	1,700 (-490－4,900)	45%
SAS	40 (-40－130)	50 (-30－160)	19,000 (4,000－42,000)	420 (-340－1,400)	20 (0－40)	10 (-10－40)	20,000 (3,600－42,200)	5%
EAS	230 (-50－630)	310 (-50－850)	450 (-1,300－2,400)	9,700 (-2,000－26,400)	30 (0－100)	80 (-10－230)	11,400 (-3,300－31,800)	15%
MDE	30 (-30－120)	60 (-50－190)	310 (-90－910)	160 (-120－520)	180 (-10－480)	30 (0－70)	870 (-330－2,600)	79%
RBU	40 (-60－170)	150 (-50－440)	-200 (-1,700－1,200)	420 (-620－1,700)	20 (-10－60)	140 (-60－420)	640 (120－1,300)	78%
PIN	900 (100－2,100)	850 (40－2,100)	7,400 (1,800－15,400)	7,800 (3,100－20,900)	140 (30－330)	210 (-100－650)	19,300 (1,400－45,000)	-
TRN	1,000 (-20－2,600)	970 (-270－2,800)	10,600 (2,600－22,000)	3,500 (-420－9,300)	210 (50－440)	200 (20－490)	18,800 (3,000－41,600)	-
RES	200 (-20－510)	250 (40－550)	6,000 (1,600－12,200)	3,000 (670－6,300)	30 (0－80)	60 (10－120)	10,400 (2,700－21,100)	-
GLO	2,300 (80－5,600)	2,400 (250－5,400)	22,600 (6,200－46,000)	13,500 (1,500－30,300)	400 (80－940)	550 (80－1,210)	47,400 (11,300－99,000)	-
RERER	34%	61%	16%	28%	55%	75%	-	

Table 6. Annual avoided multi-model empirical mean $PM_{2.5}$-related premature deaths (IHD+STROKE+COPD+LC) with 95% CI from Monte-Carlo simulations in parenthesis due to 20 % regional (NAM, EUR, SAS, MDE and RBU), sectoral (PIN, TRN and RES) and global (GLO) anthropogenic emission reductions in each region and worldwide. The diagonal, showing the effect of each region on itself, is underlined. For regional reductions, we also the RERER (eq. 4) as the percent of total avoided deaths in each receptor region that result from foreign emission reductions, as well as the percent of global avoided deaths from emission reductions in each source region. All numbers are rounded to three significant figures or the nearest 10 deaths.

Source regions/sectors	Receptor region NAM	EUR	SAS	EAS	MDE	RBU	World	Impact on foreign receptor regions
NAM	18,000 (630−28,300)	640 (80−1,100)	10 (-210−80)	200 (-300−370)	10 (0−30)	250 (90−420)	19,400 (310−30,600)	7%
EUR	60 (20−110)	31,900 (4,500−53,900)	120 (-60−190)	390 (-20−550)	400 (30−1,400)	2,700 (680−8,000)	39,400 (5,500−63,400)	19%
SAS	50 (-10−90)	110 (0−200)	47,900 (30,000−68,500)	1,400 (-70−2,100)	40 (0−150)	40 (10−110)	51,300 (32,300−73,300)	7%
EAS	340 (40−510)	400 (20−690)	900 (590−1,400)	91,100 (440−128,700)	10 (0−30)	800 (0−1,300)	96,600 (3,500−136,000)	6%
MDE	30 (0−60)	420 (90−850)	1,400 (740−2,400)	180 (-610−460)	1,600 (240−4,500)	640 (30−1,600)	5,000 (1,900−11,100)	68%
RBU	40 (10−60)	2,200 (300−3,700)	90 (-220−190)	810 (330−1,100)	80 (10−220)	17,600 (390−25,700)	21,500 (900−31,000)	18%
PIN	9,300 (940−13,000)	15,700 (1,900−24,700)	21,000 (8,400−30,700)	47,310 (22,600−69,700)	2,200 (200−6,100)	14,300 (0−24,100)	128,000 (41,600−179,000)	-
TRN	3,600 (-320−7,000)	8,900 (130−17,400)	6,200 (-12,800−14,400)	6,800 (-6,400−12,200)	230 (10−770)	3,100 (0−5,400)	31,900 (-16,500−58,300)	-
RES	2,900 (110−4,400)	6,900 (210−11,300)	25,000 (15,100−40,700)	29,300 (13,200−52,900)	200 (10−520)	4,600 (0−8,100)	83,400 (41,700−120,000)	-
GLO	19,900 (710−31,300)	40,900 (4,900−68,100)	55,300 (36,500−78,300)	105,000 (4,000−147,000)	2,800 (330−8,400)	26,700 (2,300−36,000)	290,000 (67,100−405,000)	-
RERER	10%	22%	13%	13%	42%	34%	-	

Table 7. Comparison of O_3 and $PM_{2.5}$-related premature deaths attributable to PIN, TRN and RES emissions with previous studies. Results from this study (for 20% reductions) are multiplied by 5. For Silva et al. (2016), we combine results for "Energy" and "Industry" to represent PIN, and use "Land transportation" to represent TRN and "Residential & Commercial" to represent RES. For Lelieveld et al. (2015), we combine the "Power generation" and "Industry" sectors to represent PIN, and use "Land Traffic" to represent TRN, and "Residential Energy" to represent RES.

Emission source sector	This study	Silva et al. (2016)	Lelieveld et al. (2015)
PIN	O_3: 96,500 (7,000, 225,000) $PM_{2.5}$: 640,000 (208,000, 895,000)	O_3 : 111,000 (23,200, 240,000) $PM_{2.5}$:613,000 (422,000, 816,000)	$O_3 + PM_{2.5}$ (692,000)
TRN	O_3: 94,000 (15,000, 208,000) $PM_{2.5}$: 160,000 (-82,500, 292,000)	O_3: 80,900 (17,400, 180,000) $PM_{2.5}$: 212,000 (114,000, 292,000)	$O_3 + PM_{2.5}$ (165,000)
RES	O_3: 52,000 (13,500, 106,000) $PM_{2.5}$:417,000 (209,000, 600,000)	O_3: 53,700(12,300, 116,000) $PM_{2.5}$:675,000 (428,000, 899,000)	$O_3 + PM_{2.5}$ (1,020,000)

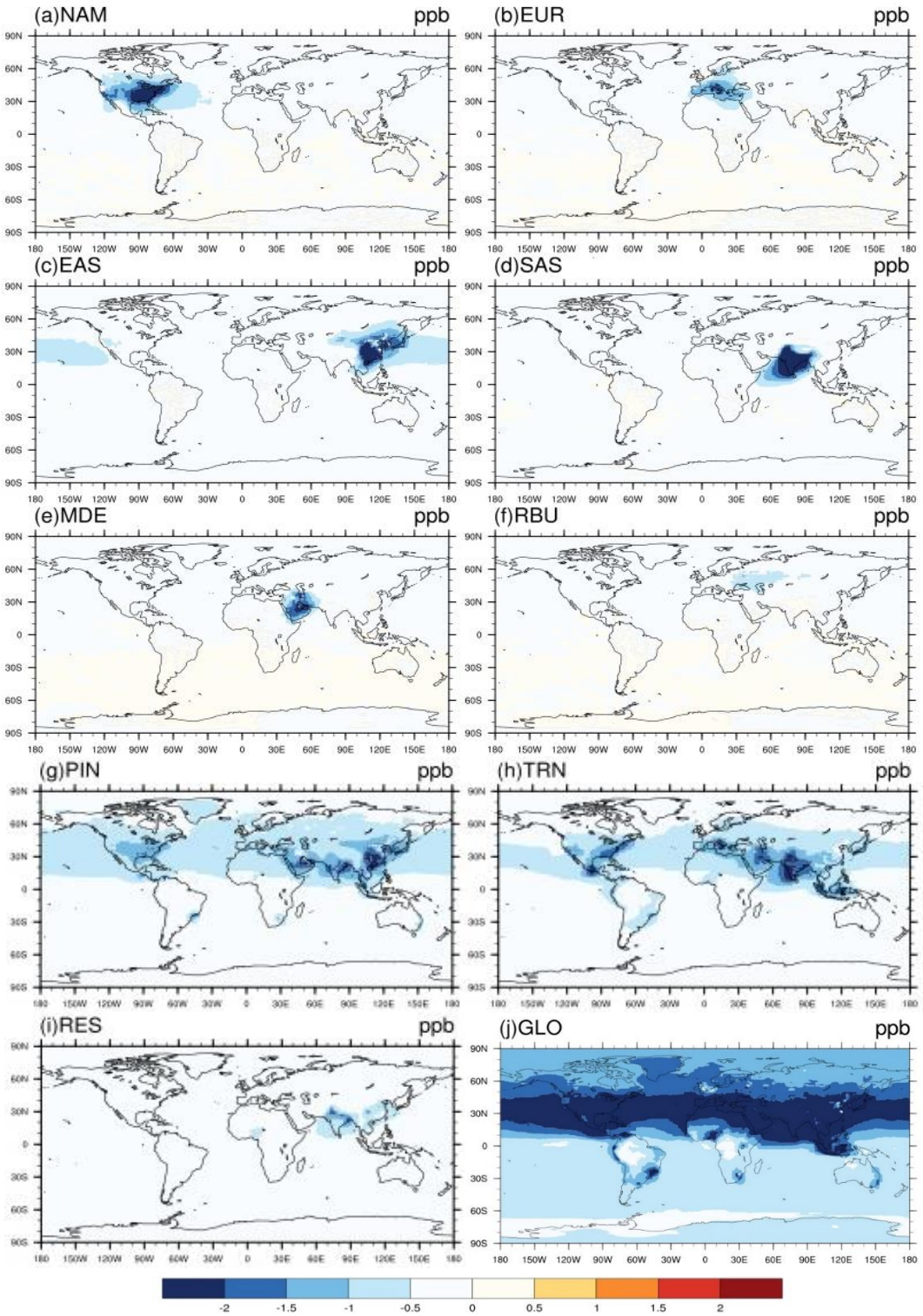

Figure 1– Global difference in multi-model mean O_3 concentrations (ppb) in 20% emission reduction scenarios relative to the baseline for the year 2010 in a) North America (NAM), b) Europe (EUR), c) East Asia (EAS), d) South Asia (SAS), e) Middle East (MDE), f) Russia/Belarus/Ukraine (RBU), g) Power and Industry (PIN), h) Transportation (TRN), i) Residential (RES) and j) Global (GLO), shown for the 6-mo. O_3 season average of 1-hr. daily maximum health relevant metric.

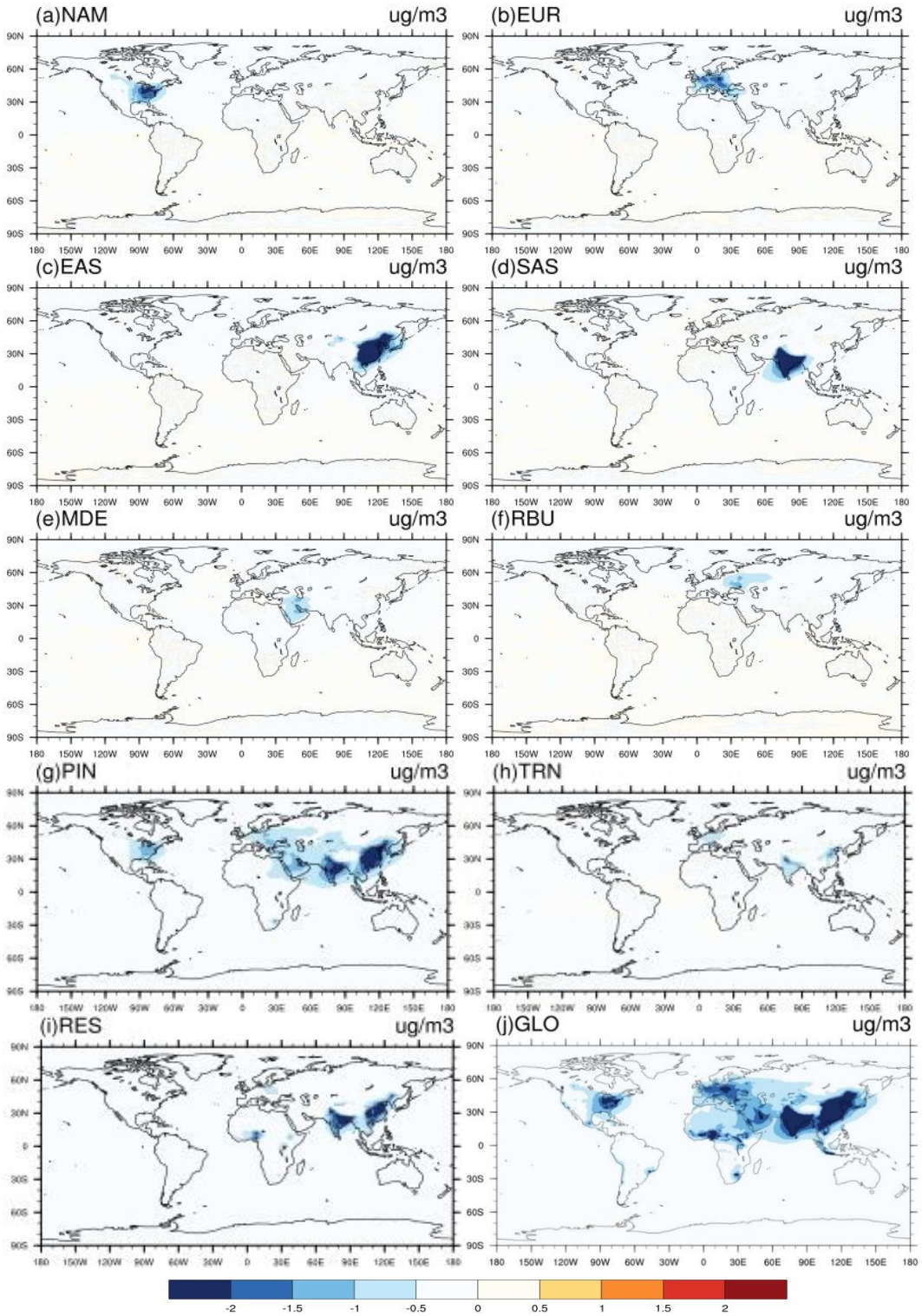

Figure 2– Global difference in multi-model annual mean $PM_{2.5}$ concentrations ($\mu g/m^3$) in 20% emission reduction scenarios relative to the baseline for the year 2010 in a) North America (NAM), b) Europe (EUR), c) East Asia (EAS), d) South Asia (SAS), e) Middle East (MDE), f) Russia/Belarus/Ukraine (RBU), g) Power and Industry (PIN), h) Transportation (TRN), Residential (RES) and j) Global (GLO).

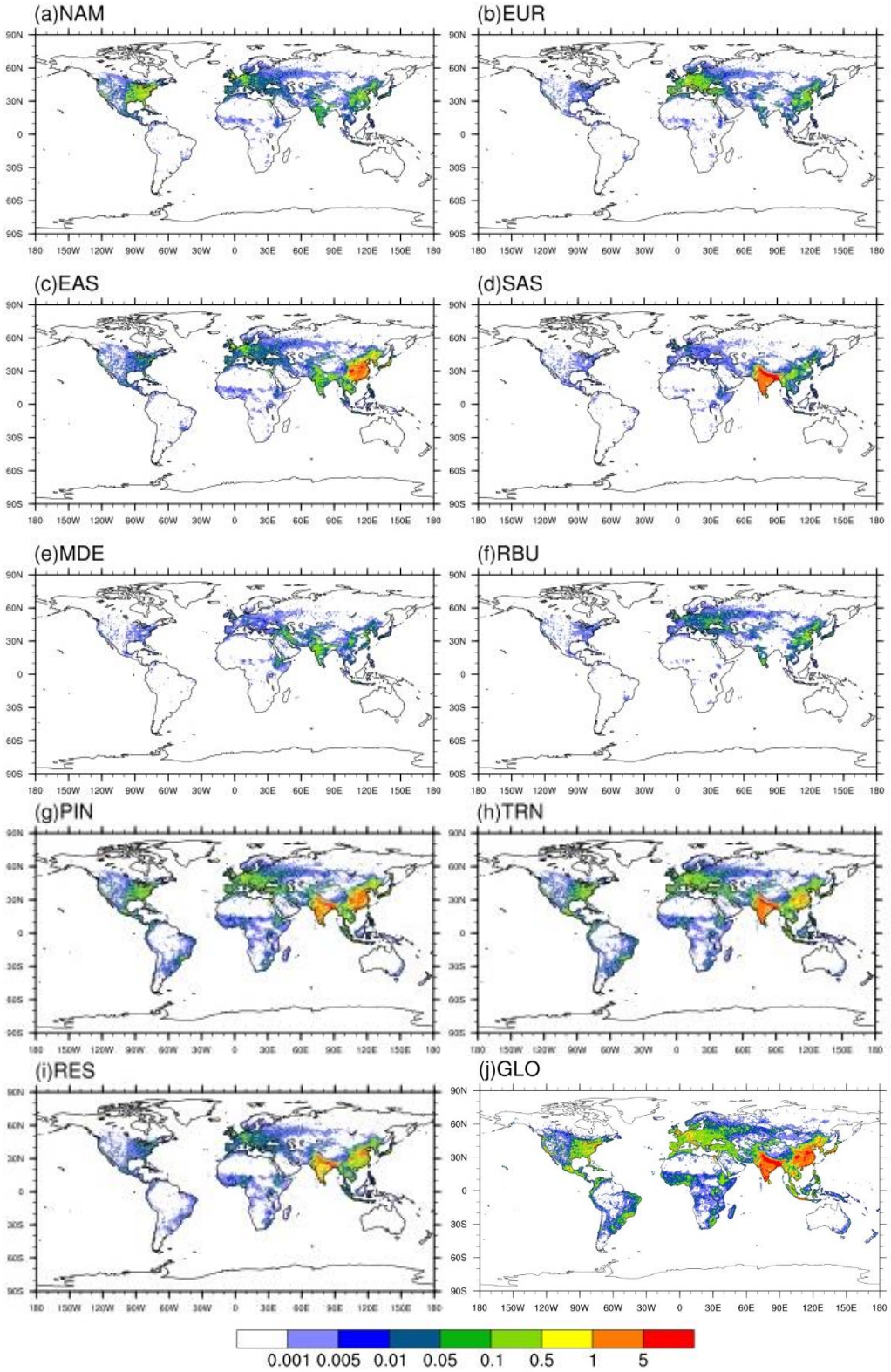

Figure ~~1~~3. Annual avoided O_3-related premature deaths in 2010 per 1,000 km^2 due to 20 % emission reduction scenarios relative to the base case in a) North America (NAM), b) Europe (EUR), c) East Asia (EAS), d) South Asia (SAS), e) Middle East (MDE), f) Russia/Belarus/Ukraine (RBU), g) Power and Industry (PIN), h) Transportation (TRN), i) Residential (RES) and j) Global (GLO).

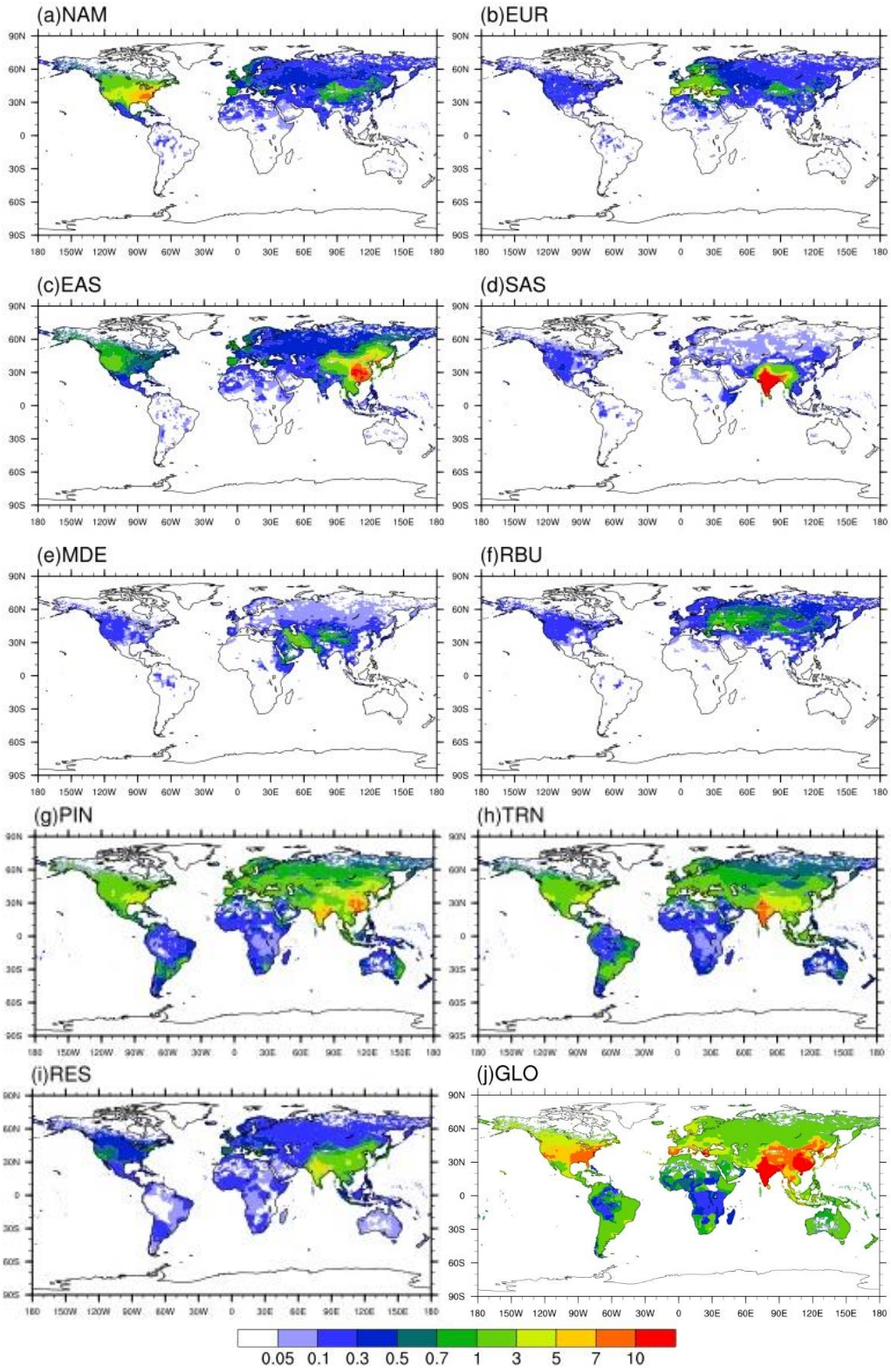

Figure ~~2~~4. Annual avoided O_3-related premature deaths in 2010 per million people due to 20 % emission reduction scenarios relative to the base case in a) North America (NAM), b) Europe (EUR), c) East Asia (EAS), d) South Asia (SAS), e) Middle East (MDE), f) Russia/Belarus/Ukraine (RBU), g) Power and Industry (PIN), h) Transportation (TRN), i) Residential (RES) and j) Global (GLO)

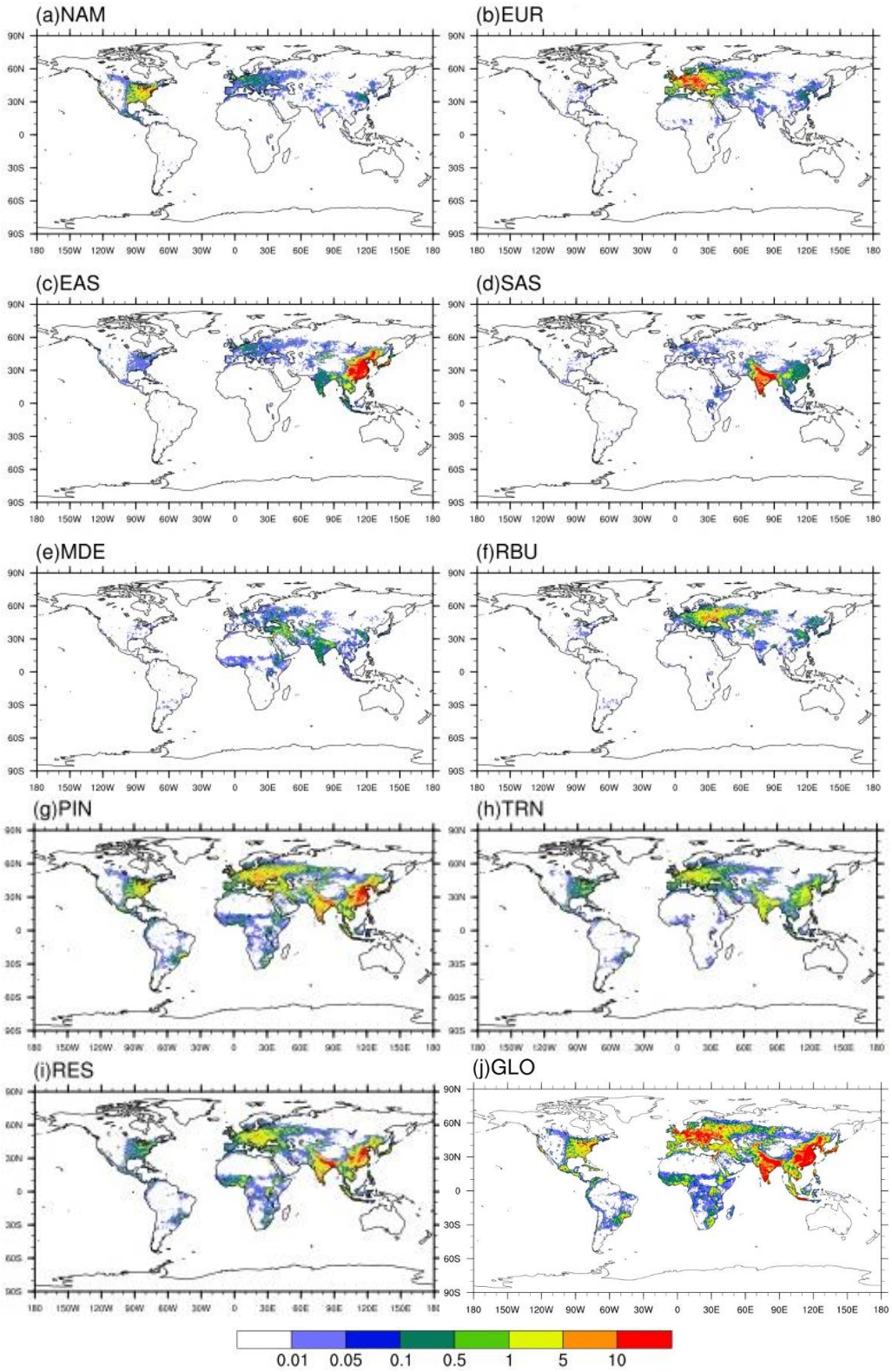

Figure ~~3~~5. Annual avoided $PM_{2.5}$-related premature deaths in 2010 per 1,000 km^2 due to 20 % emission reduction scenarios relative to the base case in a) North America (NAM), b) Europe (EUR), c) East Asia (EAS), d) South Asia (SAS), e) Middle East (MDE), f) Russia/Belarus/Ukraine (RBU), g) Power and Industry (PIN), h) Transportation (TRN), i) Residential (RES) and j) Global (GLO).

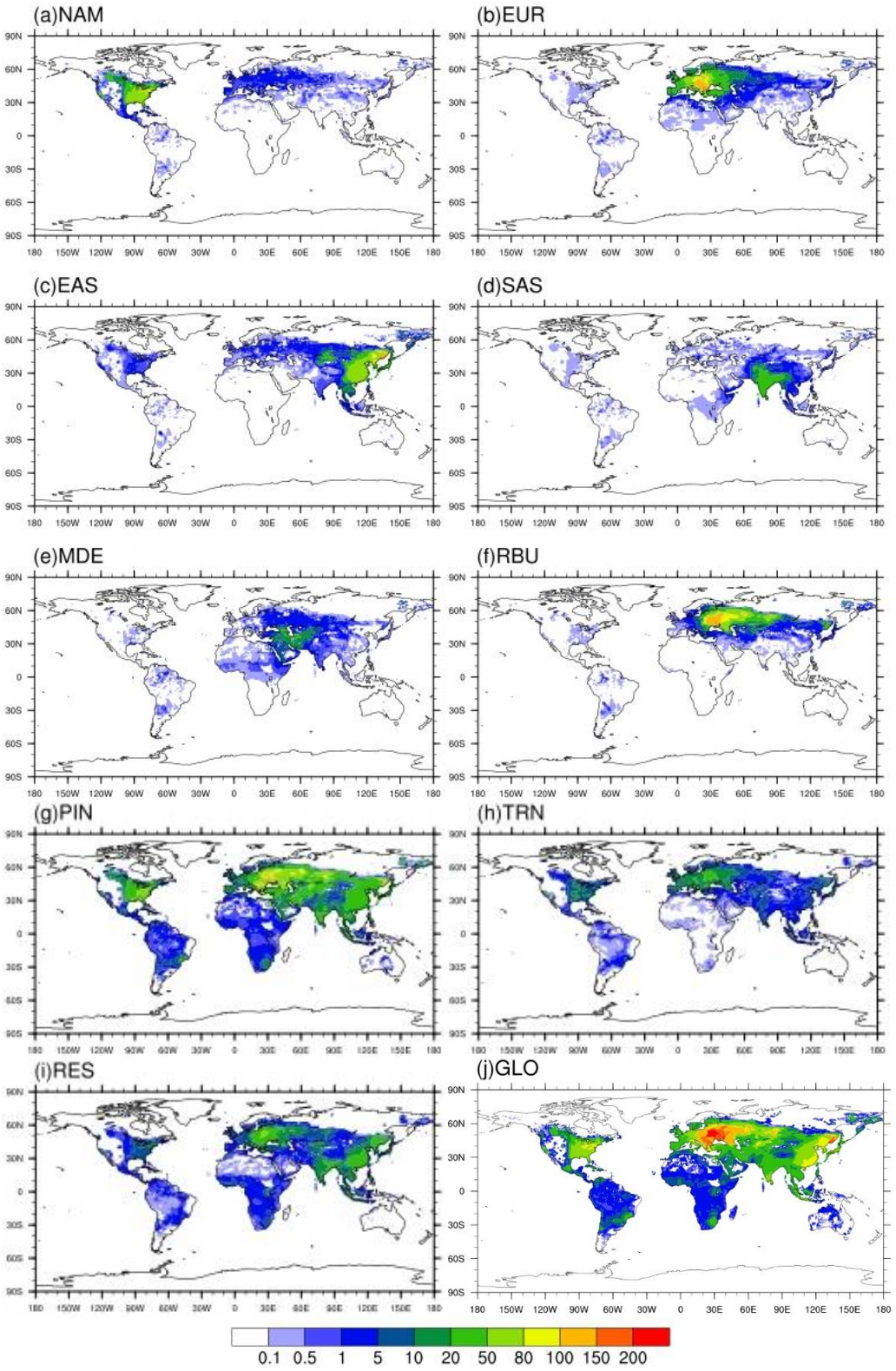

Figure 46. Annual avoided $PM_{2.5}$-related premature deaths in 2010 per million people due to 20 % emission reduction scenarios) relative to the base case in a) North America (NAM), b) Europe (EUR), c) East Asia (EAS), d) South Asia (SAS), e) Middle East (MDE), f) Russia/Belarus/Ukraine (RBU), g) Power and Industry (PIN), h) Transportation (TRN), i) Residential (RES) and j) Global (GLO).