18th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
9-12 October 2017, Bologna, Italy

THE ROLE OF MULTI-MODEL ENSEMBLES IN ASSESSING THE AIR QUALITY IMPACT ON CROP YIELDS AND MORTALITY

Efisio Solazzo[1], Angelo Riccio[2], Rita Van Dingenen[1] and Stefano Galmarini[1]

[1] European Commission, Joint Research Centre (JRC), Ispra (VA)
[2] University of Naples Parthenope, Naples (Italy)

Abstract:
This work promotes a critical use of modelling information on air-pollution health and agriculture impacts, with the primary goal of providing more reliable estimates to decision makers and stakeholders. To date, the accuracy of air quality (AQ) models and the quantification of the uncertainty of their results have rarely been quantified explicitly in impact assessment studies, therefore without giving information on the robustness of the information used in the decision making process and undermining the confidence in the results obtained. A suite of twelve regional-scale chemistry transport AQ models produced in the third phase of the Air Quality Model Evaluation International Initiative (AQMEII) is used here to calculate the impact of $PM_{2.5}$ and ozone on human health and crop yields and the associated uncertainties over Europe. A novel methodology is developed and applied to remove the offsetting bias from the models, which are then combined in multi-model (MM) ensembles. The application of unbiased MM ensembles offers an unprecedented attempt to *i)* establish and *ii)* mitigate the uncertainty due to AQ modelling on impact calculations.
We use the FASST (FAst Scenario Screening Tool) impact assessment tool to demonstrate that the accuracy of assessment of ozone-induced crop loss of wheat and maize and impact on human health (mortality) can improve dramatically when using accurate MM ensembles in place of single model realizations, as it is commonly assumed.

Key words: *AQMEII, Impact assessment, Uncertainty, FASST*

INTRODUCTION

As air quality (AQ) models are routinely consulted by decision makers for compliances against regulatory targets, accurateness and reliable results are mandatory. Further, AQ models are at the core of accountability research to demonstrate the extent to which regulations causally impacted emissions, air quality, and public health, aiming at integrating the assessment of societal and economic impact of air pollution on the biosphere, in particular on human health, agriculture, and ecosystems. Because of their intrinsic limitations and uncertainties, it is necessary to estimate the uncertainty due to AQ modelling, since the misestimate of the uncertainty could deteriorate the range of potential impact outcomes.

The cause-effect chain from emission, to dispersion and exposure (of people, crops, etc.) to quantification of impacts and economic valuation, forms the base of the AQ impact calculation. Each element of the chain involves a specific kind of model and relies on different input data. Both elements are affected by errors and more or less sophisticated calculation methods produce uncertainties that propagate and interact throughout the chain. The uncertainties associated with AQ predictions are often overlooked and even more rarely accounted for explicitly, as more emphasis is given to the uncertainty associated with the dose-response estimates and population exposure, gravely overlooking the fact that air concentration levels are primary information that will affect all subsequent calculation.

Understanding how the overall AQ uncertainty is conveyed to the health and crops impact calculations and how to mitigate it are the main aims of this study. A suite of 12 regional models (described in Solazzo et al., 2017) participating to the third phase of the Air Quality Model Evaluation International Initiative (AQMEII3) is used here to calculate the impact of $PM_{2.5}$ and ozone on mortality and crop yields and the associated uncertainties over Europe for the year of 2010. A novel technique is applied to remove the offsetting bias from the models by using the spatially distributed time series of measurements obtained by the regulatory AQ networks. The unbiased models are then statistically combined in multi-model (MM) ensembles to quantify explicitly the physical range of variability associated with AQ modelling.

The unbiased MM ensemble has been used to calculate the impacts of $PM_{2.5}$ and ozone on human health and the impact of ozone on crop yield by adopting the impact formulations implemented in the TM5-FASST (FAst Scenario Screening Tool) tool (Rao et al., 2016). FASST consists of *i)* a module to go from emissions to concentrations/metrics using source-receptor relationships (eg from TM5), and *ii)* modules to calculate impacts from concentrations/exposure metrics (eg from external models like MM ensemble). The MM-FASST impact calculations have been compared against the 'standard' FASST outcome, based on the source-receptor relationships calculated by means of the TM5 global AQ model and by the EMEP regional AQ model.

METHODOLOGY AND RESULTS

The removal of the bias, or more correctly its adjustment, operates in the direction of compensating the common errors but leaving the portion of uncertainty due to internal model components. Since all AQMEII3's models rely on the same set of emissions and boundary conditions (while are free on the choice of the meteorological drivers), the benefits of combining the models into ensemble is negligible (all the bias associated to these fields having the same sign). Thus, once the common bias is removed, the variability left is largely dominated by the intrinsic diversity of the models, which is based on the way physical processes are described by each member (model) of the ensemble. We do not want to remove these differences as they are carriers of very relevant information, which is the basis for defining ranges of variability of the models results.

First, the bias at a specific grid cell is constrained to be lower than the corresponding modelled result, to avoid meaningless negative corrected values. Moreover, model values are assumed to be affected by a slowly varying bias, encoded through a smoothness constraint, and large model-to-data deviations are penalised to preserve the original model estimation as much as possible through a simple Tikhonov regularization, constraining the squared length of the solution. These conditions are mathematically equivalent to the minimization of the cost function $J=J_1+J_2+J_3$, with the three contributions given by:

$$\begin{cases} J_1 = \omega(\boldsymbol{M}(\boldsymbol{x}-\boldsymbol{b})-\boldsymbol{y})^T(\boldsymbol{M}(\boldsymbol{x}-\boldsymbol{b})-\boldsymbol{y}) \\ J_2 = \epsilon(\boldsymbol{Db})^T(\boldsymbol{Db}) \\ J_3 = \delta\epsilon\boldsymbol{b}^T\boldsymbol{b} \end{cases} \qquad \text{Eq 1}$$

$\boldsymbol{x}$ and $\boldsymbol{y}$ are the model results and observations, respectively; $\boldsymbol{b}$ is the vector of bias corrections. **M** is a matrix mapping model results onto observation sites. J_1 measures the misfit between model values and observations. ω weights the distance of corrected model values from observations. Smoothness is implied by J_2, where **D** is a tridiagonal matrix with elements on the main diagonal equal to -2 and elements of the diagonals above and below equal to 1 (i.e. the discrete representation of the second derivative), and ε a regularisation parameter determining the weight of this smoothness constraint. The preference for small bias corrections is given by J_3, where δ measures the weight of this constraint relative to the previous one. The accuracy of hourly ozone (median values over 1061 monitoring stations) has improved by ~4 times due to the bias adjustment (not shown).

The MM mean (hereafter we refer to $\widehat{MM_{mean}}$ to indicate the unbiased MM mean in contrast to the uncorrected, biased MM mean indicated as MM_{mean}) is used here to assess:

1. the yield reduction due to surface ozone exposure for the year 2010 is estimated for wheat and maize (**Figure 1**) using the accumulated ozone above 40 ppbv threshold (AOT40) metric. The concentration-response functions implemented within FASST, derived from field studies (details in Van Dingenen et al., 2009 and reference therein), have been adopted;
2. the mortality due to exposure to $PM_{2.5}$ (**Figure 2**) and ozone (**Figure 3**). The hourly modelled concentrations are aggregated according to the underlying epidemiological studies implemented in FASST to estimate mortality (Anenberg et al., 2010; Burnett et al., 2014). For $PM_{2.5}$, the annual average concentration is used, whereas for ozone the maximal 6-month ozone average of the 1-h daily maximum ozone concentration (M6M) is used. The results for these two metrics are made available

on a regular grid of 25 x 25 km horizontal spacing covering the European continent with 65341 cells. The AQ models used in this study included the contribution of natural sources to the total concentration. A zero-risk threshold of 5.8 µg m^{-3} is set for $PM_{2.5}$ and of 33.5 ppb for ozone.

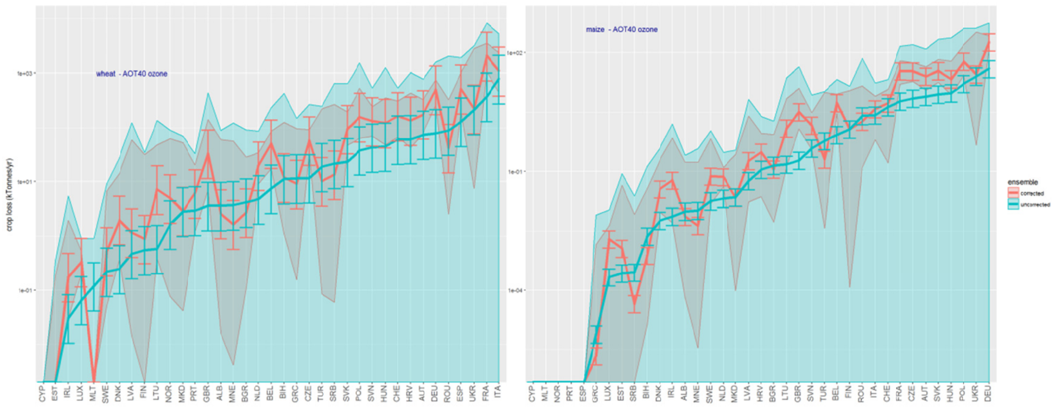

Figure 1. Crop loss (in $\log_{10}$ kTonnes/year) for wheat (left) and maize (right) calculated by using the corrected and uncorrected MM_{mean} for calculating the AOT40 ozone during the growing season. The envelope around the mean (tick lines) represents the wheat yields reduction calculated by taking the minimum and maximum values of the models participating to the ensemble, for each grid cell falling within the country boundaries. The vertical bars represent the confidence interval (at 95th percentile) associated with the response function.

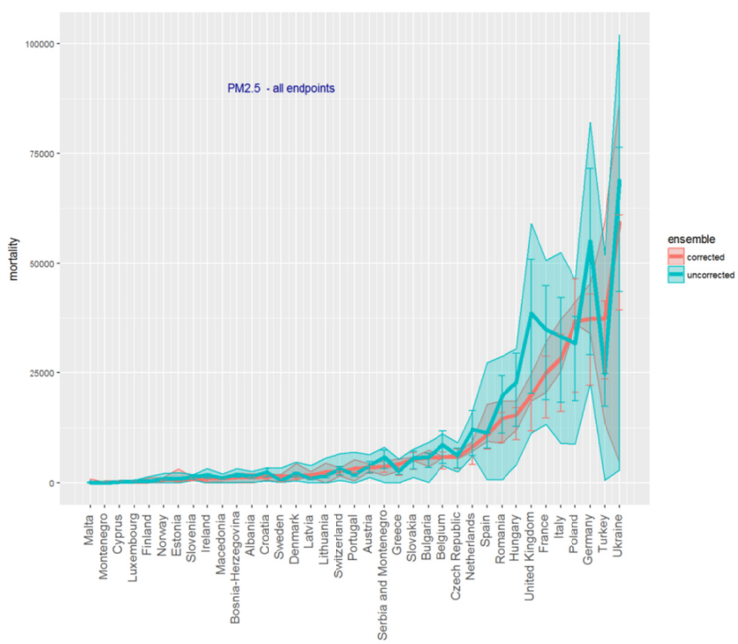

Figure 2. Mortality calculated by using the corrected and uncorrected MM_{mean}. The envelope around the mean (tick lines) represents the mortality calculated by taking the minimum and maximum values of the models participating to the ensemble, for each grid cell. The vertical bars represent the confidence interval (at 95th percentile) associated with the response function. Anthropogenic threshold concentration set to 5.8 µg m^{-3}

The results show that

- the $\widehat{MM_{mean}}$ reduces drastically the AQ uncertainty, well below the uncertainty of the dose-response function for health (**Figure 2** and **Figure 3**) and to approximately comparable levels for maize and wheat yield reduction (**Figure 1**), although country-dependent.

- the bias adjustment sensibly affects the mean value producing lower values for mortality due to $PM_{2.5}$ (e.g. in excess of 50% lower in the UK, **Figure 2**) and significantly higher values for crop loss (both cultures, **Figure 1**).
- The bias correction is more effective to the lower bound of the uncertainty, avoiding the underestimation of the impact deriving from the uncorrected MM_{mean} that, except for $PM_{2.5}$ health impact, is of zero. This means that there are modelled values in each cell predicting much lower concentration (zero or close to zero) than the observations. The removal of the bias shifts upwards the lowest values, producing more meaningful and reliable uncertainty ranges. The upper bound of the uncertainty is relatively less affected than the lower one for ozone (**Figure 1** and **Figure 3**) as the metrics used, M6M and AOT40, involve the calculation of maximum and cumulative concentrations above a high threshold, respectively, and AQ models are typically tuned to reproduce such values. In fact, when using the annual average ($PM_{2.5}$, **Figure 2**), the discrepancy between the upper bounds of the uncertainty for $\widehat{MM_{mean}}$ and MM_{mean} is significantly larger.
- The efficacy of the bias correction is limited by the availability of surface measurements. Dense spatial coverage would be ideal and, although spatial heterogeneous, the surface European networks can be suitably used for the purposes of the bias adjustment. For countries like Ukraine and Turkey, for which measurements have not or only partially been retrieved respectively, the uncertainty is not mitigated to the same amount as for the western European countries.

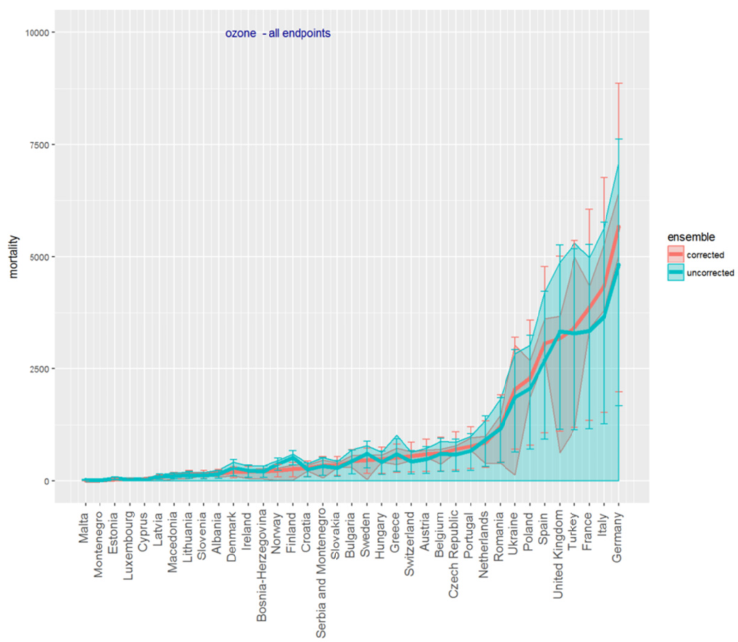

Figure 3. As in Figure 2 for ozone. Anthropogenic threshold concentration set to 33.5 ppb

The plot in **Figure 4** shows the comparison of the $PM_{2.5}$-induced mortality when calculated by using different AQ drivers for the FASST tool: the AQMEII3 MM ensemble, and the source-receptors relationships derived by the global TM5 and by the regional EMEP models. The figures provided by Holland (2014), which are part of the estimates used by the European Union's Thematic Strategy on Air Pollution, are also reported although not directly comparable with FASST as based on slightly different dose-response functions and population maps. The bars represent the uncertainty range as estimated by the $\widehat{MM_{mean}}$. As a general trend, at least for the most populated countries, the MM estimates lie between the TM5 (lower) and the EMEP (upper estimates). The latter is often embedded within the $\widehat{MM_{mean}}$ uncertainty range, while the TM5-FASST is often below the lowermost uncertainty range. Investigations are ongoing to explain such differences.

CONCLUSIONS

This work responds to the need of a more detailed assessment of uncertainty and of an increase in the confidence in impact assessment model results. The focus of the paper is the impact of air pollution and in particular $PM_{2.5}$ and ozone on human health and crop yields. The combination

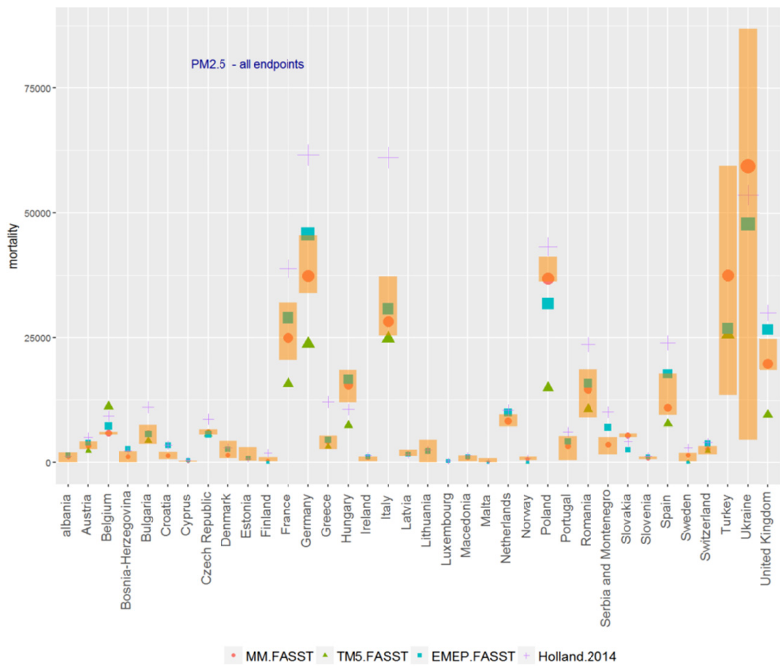

Figure 4. Mortality induced by exposure to surface $PM_{2.5}$ in 2010 estimated by the FASST tool driven by the AQMEII3 corrected MM_{mean} , by the TM5 global model, and by the source-receptor relationships derived by the EMEP model. The data by Holland (2014) are also included for comparison. The bars are the uncertainty associated with $\widehat{MM_{mean}}$, the unbiased MM_{mean}.

of bias correction and multi-model ensemble provides the highest level of quality achievable for the AQ input to an impact assessment model and presented to date in the literature. The results so far do prove that the bias adjustment sensibly reduces the uncertainty and changes the mean value significantly for both crop yield loss and mortality estimates. The results for mortality are provided with uncertainty ranges allowing a more usable and reliable information in a policy making context.

REFERENCES

Anenberg, S. C., and et al. 2010: An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using atmospheric modelling, *Environ. Health Perspect.*, **118**, 1189–1195, doi:10.1289/ehp.0901220.

Burnett, R. T., and et al., 2014: An Integrated Risk Function for Estimating the Global Burden of Disease Attributable to Ambient Fine Particulate Matter Exposure, *Environ. Health Perspect.*, doi:10.1289/ehp.1307049.

Holland, M., 2014: Cost-benefit Analysis of Final Policy Scenarios for the EU Clean Air Package, http://wedocs.unep.org/handle/20.500.11822/16417

Rao, S., and et al., 2016: A multi-model assessment of the co-benefits of climate mitigation for global air quality, *Environ. Res. Lett.*, **11**, 124013, doi:10.1088/1748-9326/11/12/124013, 2016.

Solazzo, E. and et al., 2017: Evaluation and error apportionment of an ensemble of atmospheric chemistry transport modelling systems: multi variable temporal and spatial breakdown. *Atmospheric Chemistry and Physic* **17**, 3001-3054