Response to Reviewers: Reducing future air pollution-related premature mortality over Europe by mitigating emissions: assessing an 80% renewable energies scenario

Reviewer ♯1

The authors have addressed most of the comments raised by reviewers. However, there are still a couple of points that were not addressed and must be addressed before the publication.

A: We would like to acknowledge Reviewer ♯1 for the very positive view on the manuscript and fruitful comments. Please find below our item-by-item response to the concerns raised.

Major

1. RCP8.5 scenario: The authors presented practical reasons for the RCP8.5 scenario (e.g., REPAIR initiative, simulation time), but did not provide the scientific justification of the use of RCP8.5. In their response, there was no discussion about the paper the reviewer had raised (Hausfather and Peters, 2020). They said it was discussed in the Conclusions section but it was not clear - why an unlikely scenario was chosen and why RCP8.5 was still feasible for this study.

   A: The reviewer is right. We have included the following discussion at the end of the conclusion section:

   In addition, one of the aspects that deserve further attention with respect to this contribution is the election of the future forcing scenario (RCP8.5). This high-emissions scenario is frequently referred to as “business as usual” (BAU), suggesting that is a likely outcome if society does not make concerted efforts to cut greenhouse gas emissions (Pielke and Ritchie, 2021). This RCP8.5 was originally developed to represent an upper limit to climate impacts (Moss et al., 2010), and was intended to explore an unlikely high-risk future (Riahi et al., 2011). In this sense, several authors have highlighted that this worst-case scenario is an extremely-unlikely scenario and should not be treated as a BAU scenario consistent with high CO_2 forcing (Ritchie and Dowlatabadi, 2017; Ho et al., 2019; Peters and Hausfather, 2020). Despite the criticism on the election of RCP8.5 as a reference scenario (e.g. Grant et al. (2020); Peters and Hausfather (2020); Hausfather and Peters (2020); Pielke and Ritchie (2021)), other works keep the debate open. For example, Schwalm et al. (2020a,b) indicate RCP8.5, the most aggressive scenario in assumed fossil fuel use for global climate models, will continue to serve as a useful tool for quantifying physical climate risk, especially over near- to midterm policy-relevant time horizons. These same authors indicate that not only are the emissions consistent with RCP8.5 in close agreement with historical total cumulative CO_2 emissions (within 1%), but RCP8.5 is also the best match out to midcentury under current and stated policies with still highly plausible levels of CO_2 emissions in 2100. Other works, assuming RCP8.5 is not the most-likely scenario, point out that assessing RCP8.5 might be a helpful exercise, since it flags potential risks that emerge only at the extremes (O’Neill et al., 2016).

   Despite the uncertainties and the limitations arising from this work (e.g. election of the forcing scenario, the use of a constant baseline mortality in all scenarios), it becomes clear from this contribution that governments and public entities must take project and clearly implement mitigation policies, which could improve air quality and therefore, the wellness of the European citizens.

2. Although both reviewers raised a question about the model simulation and further the use of incorporating "observed" PM, the revised manuscript still lacks a clear explanation. The authors provided two previous studies (Jerez et al., 2013; Ratola and Jimenez-Guerrero, 2016), but I think these are not appropriate references for the model used in this study. [...] If the authors do not evaluate their model, the authors should provide the previous studies with the "same" model configuration (or at least the same chemical mechanism and aerosol scheme if not the same model).

   A: Following the reviewer’s advice, a full model evaluation against PM2.5 observations available for Europe during the target period (taken from the AirBase database of the European Environment Agency (available at https://discomap.eea.europa.eu/App/AirQualityStatistics/index.html#) was conducted. Overall, PM2.5 data from 108 stations over Europe was taken into account during the period 1991-2010. The results are now presented in the Supplementary Material, and the following text and figure was included in the revised version of the manuscript:

The robustness of this simulation for representing PM2.5 is evaluated in the Supplementary Material (Tables SM2 and SM3), where the model has been compared with data from 108 stations belonging to the AirBase database of the European Environment Agency. The results are summarized in Figure 2, and the numerical results for each station can be found in Table SM3. Briefly, the low errors found (for example, average mean bias under 2 $\mu g\ m^{-3}$ and mean fractional bias $< 9\%$) guarantee the phase accordance (timing) between the simulated and observational series, their similar amplitude and, also, the quantitative accuracy of the simulated climatologies, hence making us confident of the suitability of the modeling system for the purpose of this study.

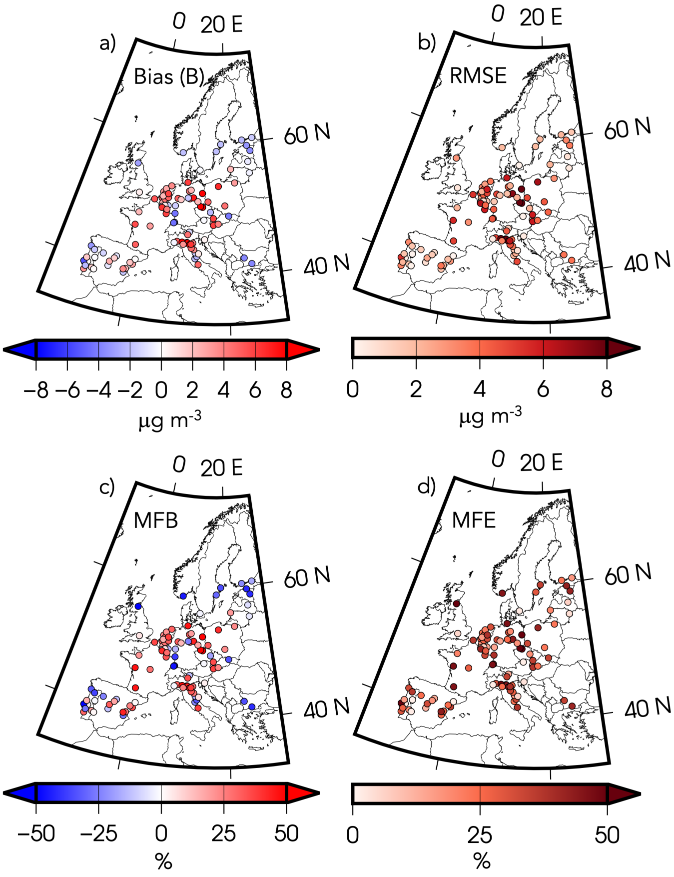

Figure 1. Results of the model validation for PM2.5 simulations: (a) mean bias (B, $\mu g\ m^{-3}$); (b) root mean square error (RMSE, $\mu g\ m^{-3}$); (c) mean fractional bias (MFB, %); (d) mean fractional error (MFE, %).

3. There was also no justification of using model results only, which was not constrained by satellite products or surface observations. I think there are many papers out there that used satellite products and/or surface PM2.5 observations to improve model results (Lee et al., 2015; van Donkelaar et al., 2016; Chem et al., 2020; McDuffie et al., 2021).

A: The reviewer raises a very interesting point here, related to the bias correction of the model simulations for improving the representation of PM2.5 concentrations. The assimilation or use of data to correct the bias present-climate simulations have been widely used, not only in the references mentioned by the reviewer (e.g. Lee et al., 2015; van Donkelaar et al., 2016; Chem et al., 2020; McDuffie et al., 2021), but also by the authors of this contribution in a number of papers covering climatic periods with available observations (e.g. Jiménez-Guerrero and Ratola, 2021). However, one of the main objectives of this contribution has to do with future climate scenarios. The question arising here is: how can we correct the bias for future PM2.5, when there are no available observations to constrain the simulations? Should we use observations/satellite products to constrain present-day simulations, but not apply bias-correction techniques for the future? This could introduce an important source of uncertainty, modifying the change signal.

In order to clarify why this contribution does not use bias-correction techniques, the following discussion has been introduced in the revised version of the manuscript:

Ground-based observations and satellite products are often used to improve modeling results for present-day simulations concerning particulate matter (e.g. Lee et al. (2015); van Donkelaar et al. (2016); Chen et al. (2020); Jiménez-Guerrero and Ratola (2021); McDuffie et al. (2021). However, these bias-correction techniques, widely used in climate impact modeling (Maraun, 2016), are limited when future scenarios are included in the simulations, since no observations can constrain future modeling results. Instead, we have decided to use the so-called "delta method" (Räisänen, 2007) to present the results and the future changes in air pollution, as recommended in Fernández et al. (2019). In the simple terms applied in this contribution, we assume that the results of the evaluation presented in the Supplementary Material point to accurate results (small biases) for present-day PM2.5 simulations, and that the difference in future (2031-2050) minus reference mean climate simulation (1991- 2010) will cancel out likely model errors. This is related to bias correction methods. In particular, delta changes are insensitive to local shift bias correction methods. It is true that more complex bias-correction techniques could have been applied (e.g. quantile mapping), but for those methods, bias corrected and delta change projections differ (Ho et al., 2012; Räisänen and Räty, 2013; Fernández et al., 2019), leading to a new source of uncertainty. Therefore, this contribution uses the delta method (assuming the cancelation of present and future biases), as also implemented in other works related to air pollution impacts on health issues (e.g. Silva et al. (2017), or the contributions of Tarín-Carrasco et al. (2019); Tarín-Carrasco et al. (2021); Guzmán et al. (2022) that rely on these very simulations; among many others).

4. Please provide the detailed methodology of model simulation. The paper nicely presented how to calculate premature mortality and emission scenarios in detail, but does not have a model description, especially for the aerosol scheme that is critical to PM2.5 estimation. I suggest the authors include these details but are not limited to: (1) Which aerosols were simulated, by aerosol type (2) Was it sectional, bulk, or modal? How was aerosol size less than 2.5 um calculated? (3) Was nitrate aerosol included explicitly in the simulation? (4) Was secondary organic aerosol simulated? if so, which SOA scheme was used? two-product, volatility basis set, or others? what kinds of VOCs were considered for SOA precursors? (5) Was thermodynamic partitioning of aerosols calculated like Jerez et al. (2013)? If so, was it ISORROPIA or MOSAIC or other? (6) Does aerosol affect cloud and precipitation in the model?
A: The following information has been included in the revised version of the manuscript, as suggested by the reviewer:

The parameterizations implemented in the WRF-Chem model are summarized in Table SM1 of the Supplementary Material. Further details about the methodology of the model simulations are included below. The GOCART aerosol module (Ginoux et al., 2001; Chin et al., 2002), the aerosol scheme used in this work, includes a bulk approach for black carbon (BC), organic carbon (OC), and sulfate, as well as a sectional scheme for mineral dust and sea salt using Kok (2011) brittle fragmentation theory, a simple and cheap computational approach (Palacios-Peña et al., 2020a). In this work, this scheme has been coupled with the RACM-KPP (KPP: kinetics preprocessor; Stockwell et al. (1997); Geiger et al. (2003). ISORROPIA (Nenes et al., 1998) was used for thermodynamic partitioning of aerosols.

In order to isolate the possible effects of climate change on pathologies only due to changes in atmospheric pollutants, constant anthropogenic emissions for all present and future simulations are assumed. Anthropogenic emissions come from the ACCMIP database (Lamarque et al., 2010) for the year 2000 by country and sector with a spatial resolution of 0.1°. This allows possible impacts to be anticipated if mitigation strategies for regulatory pollutants are not carried out and characterizes the climatic penalty on air quality levels. ACCMIP compiled a global emission dataset with annual *official or scientific inventories at the national or regional scale for CH_4, NMVOC, CO, SO_2, NO_x, NH_3, PM10, PM2.5, black carbon and organic carbon. Climate-dependent natural emission sources include desert dust, sea salt aerosols and biogenic volatile organic compounds (VOCs). The emissions were pre-processed according to Freitas et al. (2011).*

As stated in Ukhov et al. (2021), the estimation of the PM2.5 is carried out by the subroutine sum_pm_gocart in module_gocart_aerosols.F. This estimation considers dust and sea salt concentration in their bins 1 (ranges 0.1-1.0 and 0.1-0.5 μm, respectively) and 2 (1.0-1.8 and 0.5-1.5 μm, respectively), black and organic carbon and sulphate. GOCART does not include the treatment of secondary organic aerosols (SOA). The authors are aware of limitation; however, the WRF-Chem version forces to use the GOCART scheme if desert dust and sea salt aerosols are to be included (Palacios-Peña et al., 2020a). Nitrate aerosol are also not explicitly included in the simulations conducted here.

Last, it should be mentioned that the GOCART aerosol scheme in the WRF-Chem simulations presented here does not allow a full coupling of aerosol-cloud interactions (Palacios-Peña et al., 2020b). For instance, convective wet scavenging and cloud chemistry are not available. However, here the Morrison microphysics (Morrison et al., 2009) acts as a double moment scheme. Hence, the configuration of the model here allows a double-moment microphysics with greater flexibility when representing size distributions and hence microphysical process rates (Palacios-Peña et al., 2020a). When the double moment scheme is activated (as here), a prognostic droplet number concentration using gamma functions and mixing ratios of cloud ice, rain, snow, graupel and hail, cloud droplets, and water vapor is estimated (Morrison et al., 2009). Finally, the interaction of cloud and solar radiation with the Morrison microphysics scheme is implemented in WRF-Chem. Therefore, the droplet number will affect both the droplet mean radius and the cloud optical depth calculated by the model, affecting cloud and precipitation in the model.

Minor: Minor comments are mostly clarifying questions.

1. The authors said they used climatological biomass burning emissions in response to the reviewer's comment. I fully agree with the authors' view about biomass burning emissions. It would be helpful if the authors could provide the absolute number of biomass burning emissions, especially for future studies that will compare their results to this study.
   A: As commented in the previous stage of responses to the reviewer's comments, the database for biomass burning emissions available was derived from a climatological database, and was therefore neglected in our simulations. So no biomass burning emissions was used in WRF-Chem simulations for the motives previously presented.

2. It looks like natural emission sources are different between PRE-P2010 and FUT-P2010, although anthropogenic emissions are fixed. If so, please provide the emission total of dust, sea salt, and biogenic VOCs for both present and future conditions.
   A: The reviewer is right. Natural emission sources differ between PRE-P2010 and FUT-P2010, which anthropogenic emissions fixed. The natural emissions are estimated online each timestep of the model and used internally by WRF-Chem for the calculation of PM2.5 concentrations. Unfortunately, these emissions were not stored in the project database since they were an internal input to the chemistry transport module of the model.

3. "COPD, LC, LRI, and Other NCD barely change, since these causes are not too much sensitive to PM2.5 concentration as IHD (Figure 9), [...]": Figure 9 shows similar sensitivities to PM2.5 for LRI and IHD. It needs more discussion.
   A: We agree with the reviewer's comment. The sentence has been rephrased to "On the other hand, COPD, LC, and Other NCD barely change, since these causes are not too much sensitive to PM2.5 concentration as IHD and LRI at low PM2.5 concentrations (Figure 10), as also discussed in Tarín-Carrasco et al. (2021)."