Please find below our response to the referee comments on our Manuscript "The Co-benefits of a Low-Carbon Future on Air Quality in Europe"

To guide the guide the review process, the comments from both reviewers are in *italics* and our responses are in **bold**. We have numbered the reviewer comments to make it easier to refer to them in the response. Where we refer to lines or figures in our response to comments, these refer to lines in the tracked changes document.

We would like to thank both referees for the time taken to review our manuscript and for providing helpful and constructive comments. We have responded to comments and have modified our manuscript accordingly. We hope our revised manuscript is now suitable for publication.

Reviewer 1

This work describes the impact of three different SSP scenarios on future PM2.5 and O3 concentrations in Europe. The WRF-Chem model is applied over Europe on 30 x 30 km for the year 2050 and the SSP scenarios are compared to the Base Case year 2014. The WRF-Chem model is first evaluated by comparing calculated gas and aerosol species with observations followed by the analysis of the SSP scenarios on calculated PM2.5 and O3 concentrations.

The overall structure of the paper is good and the topic is meaningful, but it lacks a profound explanation for the results. For example, the authors make no attempt to physically and chemically relate the changes in emissions (SSP scenarios) to the observed changes in PM2.5 and O3 concentrations. The manuscript amounts to little more than a statement that where emissions of aerosol precursors have decreased in 2050, i.e. a decrease/increase in calculated PM2.5 and O3 are found.

Therefore, the article needs substantial revisions before it can be accepted for publication.

We thank the reviewer for the comments, they have been very helpful particularly in providing sources to extrapolate some of the trends. We have improved our analysis significantly to add information on the underlying chemistry behind our results.

Major comments.

1. The Introduction is quite long. Please try to shorten it.

We agree and have taken steps to reduce unneeded background in the introduction. We have cut the introduction from 1970 to 1424 words.

2. Recent studies (e.g. Thunis et al. (2021, 2022) and Clappier et al. (2021)) showed that the relationship between changing the emissions (e.g. 10%, 25% and 50%) and resulting concentration changes on PM2.5 can be non-linear. Furthermore, they also showed contrasting chemical regimes over a distance of a few hundreds of kilometers. PM2.5 chemical regimes are mainly determined by the relative importance of NOx versus NH3 responses to emission reductions and show large variations seasonally and spatially. For example, in the Po Valley, responses of PM2.5 concentrations to NOx emission reductions beyond 25% become non-linear mainly during wintertime. The seasonality of emission reductions is not well addressed in the paper.

We have reworked our discussion section using most of the suggested references to discuss non-linear chemical responses to emissions changes, with a focus on $PM_{2.5}$ and O_3 responses to NH_3, NO_x and SO_2 emissions changes and how these may explain some of the trends we see. We also plot these emissions reductions spatially. Please see our responses to the comment number 37 for more details.

Also, the size of the receptor (city definition) is important on the sensitivity of the emission reductions (especially in city centres) and 30 x 30 km could be too coarse to address this issue properly.

Therefore, I would recommend to add to this study 6 simulations (3 SSP scenarios for PM2.5 and 3 for O3) on a higher resolution over a smaller domain (e.g. Po Valley area, Poland, Benelux or Ruhr area) to investigate the impact of the SSP scenarios on PM25 and O3, and describe the underlying chemical mechanisms on the calculated PM2.5 and O3 concentrations and the differences between the three scenarios.

We specifically designed our methodology with country/regional policy relevance in mind and based on what is feasible with the computing resources we have available. We aim to simulate at scales that could improve understanding of air quality co-benefits at the national level that policymakers in Europe often consider. We are aware the current resolution is not optimal for understanding the chemical mechanisms at city level and as such, we have not focused on this level. This paper is intended to be similar in focus as other ACP publications such as Turnock et al. (2020) and Silva et al. (2016). Our intended niche was using a mid-point between global and local scale models, hence the comparison we provide with Turnock et al. (2020).

Additionally, to set up local scale simulations would require effectively a whole new model setup. The emissions provided for CMIP6 are designed for global models and are not at a resolution appropriate for local-scale simulations. The computational expense of WRF-Chem also means that adding additional simulations would be unfeasible (each of our simulations took upwards of a month and produced several terabytes of output).

Minor comments:

3) The paper describes the impact of the different SSP scenarios on PM2.5 and O3 in Europe. Therefore, I suggest to change the title that captures these two pollutants.

We agree and have revised the title.

4) Line 25: Be consistent with the naming of the species Ammonia, Sulfur, Nitrogen (non-capitals) throughout the whole text.

Amended

5) Line 30: remove the text between brackets.

Amended

6) Line 40: Remove the dot after m3.

Amended

7) Line 43: For reducing O3 levels in cities, this is not so 'feasible'. Please rephrase.

We agree and have rephrased this part of the introduction to account for this. Please see lines 44-47 in the tracked changes document.

8) Line 47-49: Note that for some locations the PM10 and O3 concentrations increased, see Putaud et al., (2021, 2023). Please rephrase.

See response to comment 9

9) Line 52: Also background CH4 concentrations have increased over the years.

In response to these comments, the paper now specifies that PM2.5 air quality in Europe has improved over recent years and ozone has worsened. We do not focus on CH_4 or PM_{10}, so have omitted this in order to keep the length of the introduction down. See lines 51-58 in the tracked changes document.

10) Line 55: Is mentioned earlier.

Amended to remove repetition

11) Line 70: Add references here.

We have added a reference to Von Schneidemesser et al. (2015), which describes in detail the chemical processes relevant to climate and air quality policies. (line 75)

12) Line 74-77: This is a repetition of line 30.

Removed with reference to previous discussion

13) Line 79: “Modelling”

Amended

14) Line 90: Can the authors address if future land-use/cover changes are included in the SSP scenarios? If not included, what would be the impact of future land cover changes on the outcome of the scenarios?

The SSP scenarios do have land use projections, and these are used in the gridded emission projections. This information has been added to the paper (lines 166-167)

15) Line 114-119: This part fits better in a section Discussion, not in the Introduction.

The first two sentences have been moved to the methodology (lines 143-145) where the emissions-only simulations are introduced. The second two sentences have been removed.

Section 2.1

16) Are the simulations performed with feedback switched on in WRF-Chem? If so, could you please describe if the meteorology changes between the scenarios and how much this affects the air pollutants? This might be important for SO2 oxidation in clouds (SO4= formation) and O3 photochemistry.

The aerosol-radiation feedback is switched on, however as the simulations are nudged to meteorology, there is no meaningful meteorological difference between the scenarios. We have added this information to our methodology (lines 169-170).

17) Line 166: Remove "The emissions… of futures" and replace by something like that you performed three different emission reduction scenarios.

Amended, line 190

18) Line 175: Can you name the countries where NH3 has different trajectories between SSP2-4.5 and SSP3-7.0?

There are not notably different trajectories in NH_3 emissions between SSP2-4.5 and SSP3-7.0, in fact the same countries seem to show the same trends in these scenarios (increases in France, Spain, Italy, decreases in Western Germany). This line is meant to say that these scenarios show different trajectories compared to SSP1-2.6. I have amended to reflect this and provide examples (lines 202-204).

19) Page 7: Can you please elaborate more what the differences are between the numbers in Figure 1 and Table 2. The numbers in Table 2 do not correspond with the number is Figure 1. You can also put the relative reductions between brackets in each column.

Thank you very much for pointing this out – Erroneously, the global emissions had been put in the table instead of the European emissions. This has been amended and Figure 1 is now representative of the changes in table 1.

20) Page 7: Table 2 should be "2014 emissions".

Amended

21) It would be nice to have this information for the different countries in the model domain. Or only for the bigger countries, e.g. UK, FR, ES, DE, I. Not every country reduces in the same manner their emissions. This would help to understand the emission ratios and the corresponding chemical regimes (e.g. NOx vs. NH3 for PM2.5 formation & VOC/NOX ratios for O3) for the different scenarios.

Also, it would be interesting to see the difference in spatial distribution in the emissions between the SSP scenarios for NOx, NH3 and PM2.5.

We haven't extracted the emissions changes in individual countries, but we now plot the emissions changes spatially in the supplementary material (Supplementary figs A3, A4, A5). This means that we can refer to emissions changes in particular countries/regions (and have done so in describing the $PM_{2.5}$ and O_3 formation).

21) Line 197: what do you mean by "For some analysis.." ?

For some figures and tables, we weighted $PM_{2.5}$ and O_3 by population, but not all. The ones where population weighting is used mention it in the captions. This is just to provide the formula where relevant.

22) Line 203: Add "other" between "and aerosol".

Amended

23) Line 206: Please provide reference.

Have amended to "as measurement and modelling of air quality in complex terrain is challenging and frequently less accurate (Giovannini et al. 2020)" (line 235-236)

24) Line 207: Add dot at the end of the sentence.

Amended

25) Line 209: Use the correct syntax for Figure(s) -> Fig.

Amended

26) Line 210: Indicate which panels (a) and (b).

Amended

27) Figure 2: Please provide Titles of the plots, plus add to the plot for example the Bias.

We have updated figure 2 based on these suggestions, and some suggestions from reviewer 2 to improve the readability and have added some more validation sites (K-Puszta, Hungary; Rucava, Latvia) to cover more sites in Eastern Europe.

28) Table 3: 5th Column can be named "Bias".

Amended

29) Table 3: Why is O3 missing in the evaluation?

It was excluded as the purpose of the table was to focus on $PM_{2.5}$ components and diagnose the overall $PM_{2.5}$ bias.

30) Line 218: I don't see the added value of Figure 3.

This figure was included to demonstrate whether it was a systemic bias or seasonal. We have kept it in as we think this is valuable context.

31) Also I don't see the relevance for Figure 4a. Knowing that the emissions are different, source sector allocation, temporal and spatial profiles.

We added this simply as another comparison than the EBAS observations to provide additional validation. We have removed it.

32) Line 249: PM2.5 is overestimated by 8 ug/m3 on average, but I wonder if that's true for the Eastern part of Europe. Can the author say something about the inclusion of condesables in the emissions? I guess they aren't. The condensables have a large fraction to the PM2.5 emissions for the residential heating sector, especially in Eastern Europe. How is the model performing in that area with respect to PM2.5 concentrations?

To our knowledge, the emissions do not factor in condensables. The papers discussing the production of these emissions make no reference to them (Hoesly et al. 2018; Feng et al. 2020). While fewer suitable observations were available in Eastern Europe, those that are included show lower bias than the mean. We have added this detail to our validation section (lines 247-249).

33) Line 250: It's a classical problem that models overestimate ammonium nitrate aerosol when compared to quartz filters, due to the evaporation of ammonium nitrate aerosol when temperatures are higher than 20 degrees Celsius (Schaap et al., 2003, De Meij et al., 2006). Did the authors investigate if for these measurement sites quartz filters are used? This could explain the bias.

Thank you for the suggestion, we have now investigated and do not believe any of the observations we have used to use quartz filters. All EBAS sites use slightly different monitoring methodologies and we see PM2.5 overestimations at all of them. We therefore expect this to be due to the model, rather than the observations. We have added this info to the paper (lines 249-250)

34) Figure 5 caption, remove "of" form "of the annual".

Amended

35) Also, I don't see the added value of Fig.5a. For Fig5. b-d, I propose to show the differences between 2014 and the scenarios in absolute terms. You can put the relative *difference to the Appendix. Changes in low concentrations might lead to large relative differences, while larger differences in absolute values in lower relative differences. Changes in absolute values would make more sense.*

We provide 5a and 8a simply to show the concentrations we simulate for the present day as useful context for our results. We prefer to discuss and display our results as relative changes. We do however, agree that providing the absolute changes is valuable contextual information, so we have included figures spatially showing the absolute change to the appendix (Supplementary Figures A2 and A6).

36) The same applies to Figure 8.

As above

37) Line 265-276: You describe the changes in PM2.5 concentrations for the different scenarios and provide possible explanations for that. Please quantify the emission reductions for PM, NOx and SO2 and describe the chemical mechanisms for the areas that you mention.

We quantified the changes for the whole domain in Table 2 and Figure 1. We now present spatial analysis showing the emissions changes in supplementary figures A3, A4 and A5. We have also enhanced our discussion to describe the relevant chemical mechanisms, focusing on NO_x, NH_3 and SO_2. This can be seen in the paragraphs between lines 325 and 359 in the tracked changes paper.

38) Line 285: Can you please clarify the effect of NH3 emission reductions in NH3-sensitive regions, i.e. NOx abundant and NOx/NH3 ratios?

Performing sensitivity simulations where the impacts of changing emissions of singular species would be out of scope for this paper due to the high computational costs required to do so. We have however enhanced our

discussion of the impacts of NH_3 and NO_x reductions, providing a figure showing NO_2 concentrations in the present and future (Supplementary figure A7) along with the new emissions changes figures we mention previously. We also use many of the suggested references to explain how the trends we see in NH_3 and NO_x emissions reductions Please see the response to comment 37 for this enhanced discussion section.

39) Line 290. Can you indicate for which source sector the largest reductions are found for Slovenia?

Amended – residential emissions show the largest reductions in Slovenia. Added to lines 364-365.

40) Line 297: I understand that sea salt is important for coastal sites, but does sea salt travel so much inland that it has a significant contribution to PM2.5? If so, please quantify it

We now reference literature (Manders et al. 2010) demonstrating that seas salt PM2.5 can have a notable impact on European PM2.5 (up to 5 ug/m3) at up to 300km inland (lines 375-376).

41) Line 310: Figure 7. I wonder if this is relevant, knowing that in Europe, the spatial and sectoral distribution of the emissions varies a lot (e.g. Po Valley vs central France or Poland vs Ireland). I would focus over smaller areas such as the southern part of Poland, Po Valley, Benelux or Ruhr area to highlight the impact of the three SSP scenarios on the yearly distribution of the PM2.5 concentrations and its components.

We think that there is value showing the temporal trends in $PM_{2.5}$ components at the macro scale as this drives much of our discussion. This paper is not written to focus on local areas and is meant to show trends on a regional/continental scale.

42) Line 350: Link this work to Air Quality legislation, it would be interesting to see the impact of these scenarios on the new WHO AQ limit values. The authors can select a few cities to show the impact of the SSP scenarios on WHO AQ limit values.

We cover this in Figure 5, where we have compared population-weighted mean PM2.5 in European countries to AQ limit values. The model is designed to capture trends at country-level and target policymakers at these scales rather than focus on specific cities.

43) Line 354, "such as"

Amended

44) Figure 8 caption: Move the Figs. 8c and d to a new Figure.

We prefer to keep them as a single figure to help the reader understand at a glance the differences between our model setup and Turnock et al. (2020).

45) Line 361. It's true that changes in O3 concentrations caused by changes in NOx emissions over urban areas are small. Reducing NOx increases O3 in city centres. However, the model resolution is 30 x 30km, which is quite coarse. Could the authors perform (as indicated before) additional simulations over a smaller area and see if the statement holds?

We appreciate that the reviewer would like for this to be investigated further, but as discussed above we do not have the capacity to perform these model simulations with a new, as of yet undeveloped model. We think this would be best addressed as further work in a dedicated paper for this purpose.

Reviewer 2

Overall this paper presents a nice evaluation of future scenarios and the implications for air quality using a regional model with higher resolution. That said, I think the manuscript could be improved if some of these minor comments could be addressed.

We thank reviewer 2 for their helpful comments. The paper has been updated to reflect the suggested changes.

50) L25/L195-197: I don't know why the chemical species are capitalized, they should not be.

Amended

51) L27: I would rather say 'has consequences' rather than may have. I think there is enough literature to show that.

Amended

52) L37-38: For clarity it would be good to specify '133 per 100,000' deaths? People?...

Amended to specify 133 deaths per 100,000 people

53) L50-52: Isn't it in urban areas also that we tend to target NOx emissions and so there is less titration of O3 happening? There are papers out there that discuss this and it would be worth mentioning because just saying intercontinental transport seems a bit of an oversimplification.

Amended to include; See new lines at lines 51-57 in the tracked changes paper.

54) L75-76: Aerosols are not only cooling, some are also warming depending on the components and mixing, and the 'masking' is very different depending on the region. Please make sure to communicate this correctly even if you want to just note it in one sentence.

As our simulations are nudged to meteorology, the interactions between aerosols and the climate are not critical for our paper. The introduction has been shortened in response to comment 1 from reviewer one. The section which this comment refers to has been removed as it is not needed.

55) L83-98: It is true that using consistent scenarios with the same underlying assumptions improves comparability. But is also has downsides, where various aspects of these scenarios, such as the evolution of air pollutants (in particular for RCPs) are poorly represented and people use them anyway. It also limits the diversity of scenarios considered and evaluated, which provides some limitations on the scientific evaluation.

Thank you for bringing this up, we have added a sentence to ensure this caveat is included at lines 104-105.

56) L127: remove 'of'

Amended – paragraph reworded so no longer in there.

57) L162: do you mean μm?

Microns in diameter – amended to reflect this

58) L168: when describing the middle of the road scenario, you state GHG mitigation and sustainable development does not accelerate or decelerate strongly. I was unaware that we were currently on a sustainable development pathway. Can you explain what this is in reference to? This sounds to me more like BAU not sustainable development.

This is in reference to sustainable behaviours, which are mostly reflected in the emissions, for example there is no great shift to vegetarianism and corresponding reduction in methane emissions. The description has been amended to reflect this better (lines 191-193).

59) L183: Any information on regridding? Or a publication? Or the regridded data provided through an open access link?

The emissions data is freely available on the Input4MIPs website (https://aims2.llnl.gov/search/input4mips/). This has been clarified in the data availability statement. We used standard xarray bilinear regridding methodology.

60) Table 2: The column heading states 2015 emissions, but everywhere else in the text it mentions 2014. Please clarify.

This should be 2014 as in the text, this has been amended.

61) L199-200: Why use population data for 2020 when you are using emissions data for 2014? Why not be consistent and use 2014 population data?

The population data did not have 2014 as an option – we chose 2020 to ensure we could have a consistent source for all population data, reducing the impact of differences between population grid methodologies. The paper has been amended to mention this (lines 228-229).

62) Figure 2: The colors are almost not recognizable for some of the symbols. I would use symbols that are not outlined in black, or make them bigger.

We have removed the black borders, made the symbols bigger and changed the colour scheme in order to make them more readable. See Figure 2 in the tracked changes document or updated manuscript.

63) L266-267: some of the urban areas are self-explanatory. Paris and Madrid stand out. But the industrial regions? Are we considering most of eastern Europe an industrial region? Would be good to be a bit more explicit. In the following sentences where there are e.g., major combustion plants, might be good to put a dot or note those somehow on the map because I think most people don't know where Deax or Belchatow are.

More information has been given on which regions are considered industrial (lines 302-303) . We appreciate the comment about highlighting specific locations, however we decided we did not want to add extra detail to the plots as we think this would make them harder to interpret. As such, we have enhanced the text where specific locations such as Drax and Belchatow are named to explain where they are, (e.g Drax, North Yorkshire, UK) (lines 316-320).

64) L305: If the (average) model overestimation were accounted for, would this put it under the guideline? Or still not? To understand the range being discussed here.

Amended to cover this, see paragraph at lines 379-389

65) L325: starting from the comma in the line above, there seems to be a word missing or something because the sentence is super awkward.

This paragraph has been reworded for clarity (lines 401-407)

66) General: It would be nice if you could add some text to address limitations based on the model performance and the implications thereof. In particular related to PM2.5 since there are not insubstantial differences there and that is also where most health effects come from. If possible, adding some sort of uncertainty estimates would be nice, but if not, discussing whether this is over/under-estimated or higher/lower bounds, etc. would be good. In general, including a limitations section would be helpful. Also aspects such as the mismatch in 'present day' emissions vs population would likely have implications. Some studies have also *found that including information on population demographic change can have a big influence and nothing even close to that is even mentioned here.*

We agree and a limitations section has been added (lines 511-524), we have emphasised the potential overestimations where relevant, noting how this may have impacted our conclusions, although it is hard to quantify this uncertainty. The population issue is a very fair point. We try to focus primarily on the percentage changes in concentration between scenarios, providing the population weighting as an additional source of information for those interested. We are currently working on an additional paper where the scenarios will be used for understanding the health impacts of future air quality changes. In that paper, we consider the impact of population change. We have explained that we do not control for population and the implications of this in the limitations section.