nature portfolio

Peer Review File

Severe atmospheric pollution in the Middle East

Web links to the author's journal account have been redacted from the decision letters as indicated to maintain confidentiality.

24th Mar 22

Dear Professor Lelieveld,

Your manuscript titled "Severe atmospheric pollution in the Middle East" has now been seen by 4 reviewers, and I include their comments at the end of this message. They find your work of interest, but some important points are raised. We are interested in the possibility of publishing your study in Communications Earth & Environment, but would like to consider your responses to these concerns and assess a revised manuscript before we make a final decision on publication.

We therefore invite you to revise and resubmit your manuscript, along with a point-by-point response that takes into account all of the points raised. In particular, we ask that you remove the comparison to COVID-19, provide comprehensive method details and robust support for your claims, including appropriate statistical analysis. Please note that we offer a longer format than some of the other Nature Portfolio journals (up to ~5000 words for the main text and unlimited space for Methods), which should allow for greater context and discussion of your findings, and better integration of important material currently presented in the Supplemental. Please highlight all changes in the manuscript text file.

We are committed to providing a fair and constructive peer-review process. Please don't hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the referees' comments (which should be in a separate document to any cover letter) and the completed checklist:
[link redacted]
** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to receive your revised paper within six weeks; please let us know if you aren't able to submit it within this time so that we can discuss how best to proceed. If we don't hear from you, and the revision process takes significantly longer, we may close your file. In this event, we will still be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

We understand that due to the current global situation, the time required for revision may be longer than usual. We would appreciate it if you could keep us informed about an estimated timescale for resubmission, to facilitate our planning. Of course, if you are unable to estimate, we are happy to accommodate necessary extensions nevertheless.

Please do not hesitate to contact me if you have any questions or would like to discuss these revisions further. We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Best regards,

Clare

Dr Clare Davis
Senior Editor
Communications Earth & Environment

www.nature.com/commsenv/
@CommsEarth

EDITORIAL POLICIES AND FORMATTING

We ask that you ensure your manuscript complies with our editorial policies. Please ensure that the following formatting requirements are met, and any checklist relevant to your research is completed and uploaded as a Related Manuscript file type with the revised article.

Editorial Policy: Policy requirements

Furthermore, please align your manuscript with our format requirements, which are summarized on the following checklist:
Communications Earth & Environment formatting checklist

and also in our style and formatting guide Communications Earth & Environment formatting guide .

*** DATA: Communications Earth & Environment endorses the principles of the Enabling FAIR data project (http://www.copdess.org/enabling-fair-data-project/). We ask authors to make the data that support their conclusions available in permanent, publically accessible data repositories. (Please contact the editor if you are unable to make your data available).

All Communications Earth & Environment manuscripts must include a section titled "Data Availability" at the end of the Methods section or main text (if no Methods). More information on this policy, is available at http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf.

In particular, the Data availability statement should include:
- Unique identifiers (such as DOIs and hyperlinks for datasets in public repositories)
- Accession codes where appropriate
- If applicable, a statement regarding data available with restrictions
- If a dataset has a Digital Object Identifier (DOI) as its unique identifier, we strongly encourage including this in the Reference list and citing the dataset in the Data Availability Statement.

DATA SOURCES: All new data associated with the paper should be placed in a persistent repository where they can be freely and enduringly accessed. We recommend submitting the data to discipline-

specific, community-recognized repositories, where possible and a list of recommended repositories is provided at http://www.nature.com/sdata/policies/repositories.

If a community resource is unavailable, data can be submitted to generalist repositories such as figshare or Dryad Digital Repository. Please provide a unique identifier for the data (for example a DOI or a permanent URL) in the data availability statement, if possible. If the repository does not provide identifiers, we encourage authors to supply the search terms that will return the data. For data that have been obtained from publically available sources, please provide a URL and the specific data product name in the data availability statement. Data with a DOI should be further cited in the methods reference section.

Please refer to our data policies at http://www.nature.com/authors/policies/availability.html.

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

While the topic is good and the analysis is well-done, the importance of the research needs to be better explained.
Also, the authors should focus more on the critical analysis of the data obtained, use more quantitative and statistical analysis, and also focus more on discussions of the causes of pollutant change.

Reviewer #2 (Remarks to the Author):

The paper presents findings related to the composition of air pollution in the Middle East region. It shows results concluding that anthropogenic origin of air pollution is important, unlike the popular belief that desert dust is dominant in aerosol composition. The paper therefore concludes that there is room for action to tackle air pollution of human origin in the region and its impact on health, ecosystems, and climate change.
The paper is important in addressing critical data gaps and understanding of the sources of air pollution in the Middle East region. Triggering efficient action to reduce anthropogenic air pollution is especially important in regions affected by natural dust.
I would propose 3 points for improvements :
1) Putting the results in perspective: there is already a body of evidence from different fields showing that while natural/desert dust might be dominant, there is nevertheless a non-negligible amount of anthropogenic origin. This is an important message and building on the different angles from which this message has come across would be useful as it can help build the case and push forward. Suggestion to add references from the various fields, e.g. McDuffie et al, Nature Communication, 2021, 12:3594; Hopke et al, STOTE, 2020, 740 (140091).

2) The impact on health section :
While it is of added values to put some health impacts figures, the COVID-19 comparison is somehow inappropriate for different reasons. First of all, COVID-19 is a disease (cause of death), not a risk factor. If the authors would like to compare air pollution attributable deaths, they should do it with another risk factor (environmental or individual) – with a text mention, and not in a table would be sufficient. Second, the timeframe is also not comparable and while this is acknowledged in the text, this just should not be compared.
On table S2. If the death calculation have been done with the anthropogenic share of PM2.5, this should be stated more explicitly.
3) Providing some sort of recommendations for better monitoring and modelling from the region:
The recently updated WHO Air Quality Guideline has included a good practice statement on sand and dust storms, that the authors may want to refer to (chapter 4.4).

Reviewer #3 (Remarks to the Author):

Review attached as PDF.

Reviewer #4 (Remarks to the Author):

The manuscript by Osipov et al. presents new measurements and model simulations evaluating PM2.5 and O3 distribution in the middle east. The measurements were collected during the ship-based AQABA campaign. Overall the paper emphasizes the anthropogenic contributions to these species and their impacts on human health and climate. This is interesting, as the region is generally assumed to be highly impacted by wind-blown desert dust, which is considered a “natural” source of PM2.5 in this context.
While the combination of new modeling and field campaign data lead to several interesting conclusions about the region, overall I found this paper was not ready for publication for the following reason — the main manuscript text is too often insufficient in terms of explanation and clarity, and in many instances makes claims that are not well supported by the data provided. This paper requires substantial rewriting and revision, at a minimum. It’s not clear to me if it would be a good fit for this journal unless the authors choose to focus the work on a subset of aspects that can be more fully, and more carefully, explained.

A fundamental concern is that the authors have done a rather poor job of separating text between the main manuscript and the supplemental. I was often left confused or surprised by the former only to spend time going through the latter to find out that the data int the supplement itself doesn’t support the statements of the main text, or that the basis for statements in the main text are not provided at all.

For example, in the main text the authors directly state that the major fraction of severe air pollution health impacts in the Middle East, and Kuwait in particular, is of anthropogenic origin. However, the results provided in the supporting information show this is more likely 20%. My

comments below expand on this, and many other instances.

That being said, there are many elements of the main text that are supported by their results, and there are many interesting findings. I just wish they were presented more completely, and fear that a brief-format article like this is not well suited for an in-depth analysis of a myriad of topics (deposition, climate, emissions, health impacts) stemming from this field campaign. The authors just need to be a lot more careful about backing up their claims and explaining their results precisely, and not glossing over some parts of their datasets or modeling that runs counter to the overall message they wish to get across. Below are many comments and questions on aspects that are mostly claims made by the authors where the evidence is not provided, or where the details run counter to their summary.

Major comments (page.line number):

Fig 1: The large anthropogenic faction of fine PM over the Mediterranean and extending over southern Europe is surprising. Is there no biogenic or sea-salt PM2.5 in Greece? Or is this a consequence of the boundary conditions used, or the treatment of natural aerosol sources? Looking at page 4 of the supplemental, I don't see any description of biogenic emissions or sea salt aerosols. Without these components, fig 1(d) is potentially highly misleading in some areas. I suggest the authors remove or limit the scope of this figure. While analysis in Fig 2 likely indicates these components are negligible along the ship measurement track, with perhaps the exception of significant sea salt in the Arabian Gulf and Arabian Sea, the map in fig 1(d) extends much beyond this, spatially.

4.17: This glosses over some areas where it seems the IC results indicate sea salt could be important (Arabian Gulf? it's a bit hard to tell from the figure) or where there are large discrepancies between total PM and spectated PM (Arabian Sea). While the overall conclusions aren't in question, these other features do warrant consideration and discussion.

5.12: Regarding the areas impacted by deposition, is there a figure or quantitative result that can be provided here? How much sulfate is deposited? What is meant by "intensely" or "particularly affected"? If the details on this topic can't be provided, maybe it should be omitted?

5.18-25 and Fig S7: The authors have not carefully considered the issues with their NHx simulations or what this may tell them about emissions. The should consider the following:
- How did the authors determine that ammonia emissions are underestimated by a factor of 15 (as opposed to some other factor)?
- Why did the authors choose to increase NH3 emissions from agricultural sources and not industrial sources? In the supplemental, they state that "sensitivity analysis" shows that point sources could not explain the missing NH3. What sensitivity analysis was conducted, and how was this concluded? Is there another set of modeling results where the point sources were increased but the area sources were not? The authors state that increasing the area sources lead to "capturing the bulk of the ammonia variability" — can this statement be quantified? A proper comparison would be to e.g. present simulation results for different cases (increasing area sources, increasing point sours, or some combination of both) along with quantitative statistical evaluation of correlation and error compared to observations, and then draw conclusions from this.
- NHx deposition is also highly uncertain given the dependence on precipitation which is often not

well simulated — how much could this contribute to the discrepancies noted for NHx between the observations and model?
- Also, if the model's NOx and SO2 levels are not correct, then partitioning of NHx can be impacted; since NH3 and NH4 have different deposition rates, incorrect partitioning can also contribute to errors in simulated NHx.
- In Fig S7, right, panel, if the grey is the default simulation and the brown in the boosted simulation, what do the orange markers represent? They are labeled as "Model", but they don't match either the default or boosted model simulations. They are not directly explained in the figure caption or text. Also, why are there no model values to compare to observed values during the first ~week of the campaign and the last ~week? For these periods only the blue "Obs" markers are shown. It's odd that the timeseries in this figure, in terms of dates, does not match that of the previous figures (S4-S6). Where was this data collected, why is it not modeled, and why is it not shown in the other figures (S4 - S6)?

6.1-5: How do the authors distinguish, quantitatively, between BC from shipping vs BC from urban sources? It seems that additional model simulations with tagged or perturbed emissions would be necessary to make this distinction, or evaluation of ratios of species with known sources. This is another example of claims being made that, while reasonable, are not backed up with any analysis or quantitative evidence. ….. Ok, after reading the supplemental in full, I see where the evidence for some of these claims are provided. Still, the summary in the main text doesn't do a good job of providing quantitative evidence or explanation, even if only briefly, of how conclusions were drawn. It seems overall that the split between main text and supplemental text isn't working well for this paper, and I'd suggest a different, longer, format and thus possibly a different journal. Regardless, the supplemental provides details such as the contribution of biomass burning to background BC being up to 40% in the seas surrounding the Arabian Peninsula, whereas the main text says this is from shipping. Lastly, it's not explained how spectral measurements are used to distinguish between fossil fuel and biomass burning BC — please explain this technique.

S10.7: How are dust mass fluxes "optimized"? This isn't explained at all. I'm guessing from the subsequent discussion of dust emission size distributions that here they referring to using a uni-modal size distribution (coarse mode only) and omitting accumulation mode dust. But this is a single selection, not an optimization. They don't show model performance for different choices regrading the dust size distribution. They also don't adopt the coarse mode shape from Kok (2011) which has been found to improve the representation of coarse mode dust in other models.

6.5-15: While I appreciate the result that in some areas the 500 nm AOD is dominated by anthropogenic PM is interesting, I think further discussion is warranted. First, it is commonly known that dust is primarily coarse and thus not the main contributed to AOD at 500 nm. Coarse-mode AOD from e.g. MODIS is derived using other wavelengths. So this result should be put a bit more in context. Second, the feature isn't universal in their reported observations. There are parts of the campaign (e.g. Arabian Gulf) where the modeled AOD is significantly larger than the observations, and the reason isn't explained.

7.5-7: I'm not sure what they are trying to say here. If half of PM2.5 is desert dust, then how can the health impacts be mostly owing to anthropogenic PM2.5? Wouldn't that also be half? Or is this based on the number of people exposed in areas that have a higher anthropogenic fraction? I looked in the supplemental, and it doesn't appear they have calculated the attributable fraction specifically for anthropogenic PM2.5, so this is a bit hard to say. They do provide the PM2.5 concentrations by

total vs anthropogenic, in table S2. While there are some countries for which anthropogenic dominates (Egypt) there are others where it doesn't (Saudi Arabia). Several in the latter category are those with severe health impacts as noted non line 11/12, yet line 13/14 misleadingly implies that a major fraction of these health impacts are controllable. In fact, for the country with the most severe %-based health impacts (Kuwait), it is likely only ~20% is anthropogenic (assuming direct relationship between PM2.5 and deaths, which is an oversimplification and the authors should be able to estimate this more precisely). Hence, a more thorough analysis and careful discussion is needed.

6.20 and 7.24: The authors have conflated AOD at 500 nm directly with radiative forcing, which is incorrect (one needs to consider the entire short-wave spectrum). Also, by definition "radiative forcing" is the impact of anthropogenic activity on the radiative balance, and thus to conclude that it is mostly anthropogenic is somewhat nonsensical. The authors should more correctly refer to the "radiative effects" not the "radiative forcing" of PM it is the former which includes both anthropogenic and natural causes. Their analysis is also limited direct effects; indirect effects are not discussed nor considered.

General: The authors should take a look at the recent Nature Communications paper from McDuffie et al. (https://www.nature.com/articles/s41467-021-23853-y.pdf), which I consulted for comparison to their claims that anthropogenic sources are not a known major impact in this region, and for the evaluation of the PM2.5 concentrations and associated health burden. On the former topic, be aware that they show, for example, that slightly more than 50% of the PM2.5 in Egypt is of anthropogenic origin. They also estimate a similar number of premature deaths as the present study (about 90,000). They report a population-weighted PM2.5 concentration of 66 ug/m3 whereas the present work reports a non-weighted concentration of 40 ug/m3.

Minor comments and corrections (page.line number):

3.12 and elsewhere: I don't think it is advisable to refer to specific features of supplemental figures, as they need to be literally supplemental. If there is a specific feature that warrants discussion, it should be fit into the main manuscript.

2.28: This phrase is a bit odd/misleading as the strongest associations of fine PM and health are based on mass concentration, not number concentration. There are though stronger associations between fine than coarse PM mass concentrations. I suggest the text be rewritten to be more accurate in this regard.

3.1 Should the title of this section be O3 instead? Components of PM2.5 are produced by photochemical reactions.

3.9: I don't think "indigenous" is the right word here. Perhaps they meant anthropogenic?

3.14: How is it known that O3 isn't exported in the form of precursors (e.g., PAN) rather than directly as O3?

Fig S3: Provide statistical evaluations for each (gases, aerosols, optical properties), separated by species or property.

Fig S4: Please define acronyms used in the figure (e.g., NRS).

3: How is net O3 production diagnosed, both from the observations and from the model?

4: How is it diagnosed that the region is NOx-limited? Were they considering ratios of particular species to make this determination empirically, is it based on model evaluation, or other means?

S11.23: What is the “spectral loss function”?

S12.22: The sentence “The previously underrated…” I’m not sure what emissions they are referring to. Could this be re-written or explained more fully?

Review of "Severe atmospheric pollution in the Middle East" by S. Osipov et al., Communications Earth & Environment in 2022

General Description and Comments:

The study uses observations of atmospheric composition from a recent ship-borne campaign and a regional chemistry transport model to determine that the anthropogenic component of submicron particles dominates in most locations in the Middle East. This is an important result, as it opposes the current impression that desert dust is the overwhelmingly dominant contributor to health-hazardous particles in the region. It also has implications for developing policies to regulate fine particles.

The focus and research findings are suited to the journal, but the paper in its current form is often missing relevant methodological details. This is the case in the following locations:

- p. 3 line 14: It's not clear whether the ozone production value is from the current or previous studies. If the former, is it derived from the observations or model? Also not discussed is what causes such high ozone production rates.

- p. 4 line 5: Is the NO2 pollution result from the model or measurements? Also unusual is the use of the "~" to indicate approximate for a clearly defined WHO guideline. Presumably this is because the guideline has been converted from μg m^{-3} to ppbv, but why not rather convert the data from ppbv to μg m^{-3} for comparison to the guideline?

- p. 4 lines 5-6: How was it deduced that ozone production in the Middle East is mostly NOx-limited or transitionary? Is this based on the model or the measurements or both?

- p. 4 lines 18-19: How was this 40% value obtained?

- p. 5 lines 12-14: Not indicated is where these simulation results are shown in the text.

Often quantitative information is lacking to support the insight or conclusions drawn in paragraphs in the paper. These occur in the following locations:

- p. 4 line 22: Indicate what the relative humidity is. Presumably this was measured or can be calculated from the measurements?

- p. 5, line 22: Quantify in the text by how much the emissions are "severely underestimated".

- p. 6, line 20: It would be more meaningful to say by how much, as this has implications for anthropogenic radiative forcing estimates.

- p. 7 lines 7-11: Summarize the main premature mortality estimates in the text, rather than leaving the reader to interpret the table, especially given the table presents the %

rather than number of excess deaths from PM2.5 and COVID-19 and the PM2.5 values are annual, whereas COVID-19 values are over a 2-year period.

- p. 5 lines 12-14: No quantitative information is presented to support these findings.

The manuscript would also benefit from a brief discussion of the broad applicability of the measurements obtained over a single season and year to other seasons and years.

Specific Comments:

Figure 1 could be better utilized to show the location of the Basin referred to on p. 3 line 28 and the ship tracks during the campaign. The latter are currently in the supplementary, but are key to the publication, so should really be in the main text.

Abstract: "Hitherto" and "Furthermore" are unnecessary and sound odd.

p. 2 line 8: "covered" is more appropriate than "addressed".

p. 2 line 10: "these obstructing factors" suggests there are many, but the impression from the previous sentence is that there is only one (poor representation of Middle East emissions in inventories).

p. 2 line 17: "various" is vague.

p. 3 line 7: Where is the transported pollution coming from?

p. 3, lines 7-10: This sentence is long and difficult to comprehend.

p. 4 line 26: Please specify the source of petrochemical industry SO2 and whether there are emission control technologies or other methods to address this source. This seems pertinent, given the findings in the study that anthropogenic emissions dominant submicron particles.

p. 4 line 26: What contribution do ships make to SO2 and submicron sulfate?

p. 5 lines 1-2: "much more spatially uniform" than what?

p. 6 lines 15-22: Surprising to not see AERONET observations used to support the Angstrom exponent analysis. Especially given that AERONET is long-term, year-round, publicly available, and has many sites in the Middle East (https://aeronet.gsfc.nasa.gov/).

p. 7, lines 23-24: Radiative forcing is determined by integrating over all wavelengths, so this conclusion doesn't seem appropriate.

Figure S7: The data could be more clearly labelled. Grey and brown are modelled NHx. Is orange just ammonium? If so, is this from the default or boosted model? Presumably blue "Obs" is ammonium only, given that NH3 wasn't measured?

We thank the referees for reviewing our manuscript. To address the comments and recommendations, we have revised the manuscript as summarized below:

- To improve the clarity and data analysis, and better explain the importance of the research, we used a large part of the supplementary information and integrated and adjusted it to the main text.
- We added quantitative diagnostics and use precise language where possible. For example, we use probability density functions to describe the observations during the campaign, statistical metrics for the model evaluation against the observations and population-weighted diagnostics to describe the exposure to air pollution in the region. We shifted more towards generalizing and discuss region-wise rather than country-level results.
- We updated the sections of the text and figures that were critically reviewed. Updated text and figures should be easier to read. We added more figures in the supplement to support every claim in the manuscript. We edited the text and explicitly state where to find the details on the ozone chemistry during the campaign.

Below we describe the improvements in detail and provide the point-by-point responses.

Reviewer #1 (Remarks to the Author):

While the topic is good and the analysis is well-done, the importance of the research needs to be better explained.
Also, the authors should focus more on the critical analysis of the data obtained, use more quantitative and statistical analysis, and also focus more on discussions of the causes of pollutant change.

Thank you. Please, refer to our summary above and responses below, where we explain in detail how the revised manuscript addresses your suggestions.

Reviewer #2 (Remarks to the Author):

The paper presents findings related to the composition of air pollution in the Middle East region. It shows results concluding that anthropogenic origin of air pollution is important, unlike the popular belief that desert dust is dominant in aerosol composition. The paper therefore concludes that there is room for action to tackle air pollution of human origin in the region and its impact on health, ecosystems, and climate change.
The paper is important in addressing critical data gaps and understanding of the sources of air pollution in the Middle East region. Triggering efficient action to reduce anthropogenic air pollution is especially important in regions affected by natural dust.
I would propose 3 points for improvements :

1) Putting the results in perspective: there is already a body of evidence from different fields showing that while natural/desert dust might be dominant, there is nevertheless a non-negligible amount of anthropogenic origin. This is an important message and building on the different angles from which this message has come across would be useful as it can help build the case and push forward. Suggestion to add references from the various fields, e.g. McDuffie et al, Nature Communication, 2021, 12:3594; Hopke et al, STOTE, 2020, 740 (140091).

Thank you for your supportive comment and for sharing these great references. McDuffie et al, 2021 is now cited in the revised manuscript.

2) The impact on health section :
While it is of added values to put some health impacts figures, the COVID-19 comparison is somehow inappropriate for different reasons. First of all, COVID-19 is a disease (cause of death), not a risk factor. If the authors would like to compare air pollution attributable deaths, they should do it with another risk factor (environmental or individual) – with a text mention, and not in a table would be sufficient. Second, the timeframe is also not comparable and while this is acknowledged in the text, this just should not be compared.

Thank you for pointing this out. Following your recommendation, we added the comparison of our results with the GBD assessment and report comparable risk factors that contribute to mortality (e.g. tobacco smoking).

Following your recommendation, we removed all the discussions/comparisons with COVID-19 except a single case in the “Air pollution health burden” section. We agree that COVID-19 and air pollution are entities of different categories (disease and risk factor). Nevertheless, both are expressed in the same units of excess mortality (allowing them to be compared to some degree). Since the newly emerged COVID-19 has received a lot of attention recently, by comparing our results with the COVID-19 we are putting the importance of air pollution further into perspective, underscoring the attention that air quality deserves. In the text, we inform the reader about the differences between the two.

Regarding the timeframe, please note that we report annual mortality in both cases (air pollution and COVID-19), which are suitable for comparison. We acknowledge that the timeframe, during which the COVID-19 statistics were collected, is indeed limited.

On table S2. If the death calculation have been done with the anthropogenic share of PM2.5, this should be stated more explicitly.

We revised the caption of Table S2 explicitly indicating that the excess mortality calculations are due to total (natural and anthropogenic) PM2.5.

3) Providing some sort of recommendations for better monitoring and modeling from the region: The recently updated WHO Air Quality Guideline has included a good practice statement on sand and dust storms, that the authors may want to refer to (chapter 4.4).

Thank you for pointing out the new WHO guidelines. We feel that specific recommendations for monitoring and modeling would be more appropriate within the framework of an upcoming, more detailed climate and air quality projection analysis where we will present projections that are consistent with IPCC. This work is underway and we will include these recommendations for the region in a manuscript in preparation.

Reviewer #3 (Remarks to the Author):

General Description and Comments:
The study uses observations of atmospheric composition from a recent ship-borne campaign and a regional chemistry transport model to determine that the anthropogenic component of submicron particles dominates in most locations in the Middle East. This is an important result, as it opposes the current impression that desert dust is the overwhelmingly dominant contributor to health-hazardous particles in the region. It also has implications for developing policies to regulate fine particles.
The focus and research findings are suited to the journal, but the paper in its current form is often missing relevant methodological details.

Thank you for your support and constructive suggestion. We have revised our manuscript and included further details to support our conclusions.

In particular, many questions below are related to ozone chemistry asking for the details related to ozone analysis. Many of these details are available in Tadic et al., 2020, who performed the analysis largely based on observations and the global EMAC model, which is a good numerical tool to represent stratosphere-troposphere exchange and hemispheric transport. Here we add results with a regional, high-resolution model which is more appropriate to represent localized emissions and photochemistry. For ozone, we revised the manuscript and added clarifications to distinguish between our own calculations and Tadic et al., 2020 or similar previously published studies related to the observational AQABA campaign.

This is the case in the following locations:
• p. 3 line 14: It's not clear whether the ozone production value is from the current or previous studies. If the former, is it derived from the observations or model? Also not discussed is what causes such high ozone production rates.

We added the clarification citing Tadic et al., 2020, where the exact value of the net ozone production was estimated (i.e. up to 32 ppbv day^{-1}).

The conditions contributing to high ozone levels (i.e. high net ozone production rates) are discussed and summarized in the previous paragraph, citing previous research and highlighting the role of stratosphere-troposphere exchange, intensive photochemistry due to a tropical location, high concentrations of O_3 precursors due to strong local sources and transport.

Detailed analysis of the photochemical ozone production in the Middle East is available in Tadic et al., 2020, which is now cited several times in our manuscript. We also interpret our results to provide overall support for an earlier modeling study by Lelieveld et al. (2009) who performed their analysis with a course grid, global model, and did not have measurement data available at the time, hence their results should be considered as more qualitative. With the present, more detailed model setup and the comprehensive measurement data we are now able to provide quantitative evidence.

• p. 4 line 5: Is the NO2 pollution result from the model or measurements? Also unusual is the use of the "~" to indicate approximate for a clearly defined WHO guideline. Presumably this is because the guideline has been converted from g m-3 to ppbv, but why not rather convert the data from ppbv to g m-3 for comparison to the guideline?

We added clarification that NO2 pollution is modeled and added Figure S15 showing NO2 map. We added the exact value of the WHO guideline in µg m^{-3} and equivalent volume mixing ratio assuming STP conditions.

• p. 4 lines 5-6: How was it deduced that ozone production in the Middle East is mostly NOx-limited or transitionary? Is this based on the model or the measurements or both?

We clarified the text by citing Tadic et al., 2020, who established the NOx or VOC limitation regimes based on the analysis of the observations.

• p. 4 lines 18-19: How was this 40% value obtained?

We clarified the text stating that 40% is the median value during the campaign, which is shown in Figure 2. According to the caption, Figure 2 shows the fractions of the speciated aerosols (and others) according to the mass spectrometer (AMS) and size distribution (OPC) measurements along the ship track. The 40% value represents the mass fraction of the organic aerosols within the PM1.

• p. 5 lines 12-14: Not indicated is where these simulation results are shown in the text.

We added Figure S16 showing the dry deposition map.

Often quantitative information is lacking to support the insight or conclusions drawn in paragraphs in the paper. These occur in the following locations:

Thank you for pointing to these omissions. We added relevant information to support each case listed below.

• p. 4 line 22: Indicate what the relative humidity is. Presumably this was measured or can be calculated from the measurements?

We added references to Figure S13, which shows the observed relative humidity and other meteorological parameters.

• p. 5, line 22: Quantify in the text by how much the emissions are “severely underestimated”.

Our estimate is given in the same line. We quote: “likely by more than factor of 15”.

• p. 6, line 20: It would be more meaningful to say by how much, as this has implications for anthropogenic radiative forcing estimates.

We have planned to quantify the exact contributions of the natural and anthropogenic aerosols to the radiative forcing upcoming work where we discuss the climate implications more thoroughly, including scenario calculations. In the section “Aerosol optical properties” we provide the partitioning of the visible AOD due to natural and anthropogenic aerosols. We think that this analysis is sufficient for the current manuscript (especially given the words limit constraint) and serves as the first-order estimate of the anthropogenic fraction in radiative forcing. The overall aerosol radiative forcing in the region was reported previously by Osipov et al., 2018 (https://doi.org/10.1002/2017JC013335) and Brindley et al., 2015 (http://dx.doi.org/10.1002/2015JD023282).

• p. 7 lines 7-11: Summarize the main premature mortality estimates in the text, rather than leaving the reader to interpret the table, especially given the table presents the % rather than number of excess deaths from PM2.5 and COVID-19 and the PM2.5 values are annual, whereas COVID-19 values are over a 2-year period.

Thank you. We added region-wide statistics and summarize them in the revised manuscript. Please note that the excess mortality is always annual in our manuscript and, in case of Table S2, this is indicated in the caption. Excess mortality due to COVID-19 was removed from Table S2.

• p. 5 lines 12-14: No quantitative information is presented to support these findings.

We added Figure S16 showing the dry deposition maps.

The manuscript would also benefit from a brief discussion of the broad applicability of the measurements obtained over a single season and year to other seasons and years.

Thank you.

Specific Comments:
Figure 1 could be better utilized to show the location of the Basin referred to on p. 3 line 28 and the ship tracks during the campaign. The latter are currently in the supplementary, but are

key to the publication, so should really be in the main text.

Thank you. We added Figure 1, showing the ship track and region of interest. We also added Fig. S14 showing the FLEXPART retroplumes and exact transport pathways of air masses.

Abstract: “Hitherto” and “Furthermore” are unnecessary and sound odd.

Thank you. We think that “hitherto” and “furthermore” are useful in separating the state of knowledge before and after the AQABA campaign.

p. 2 line 8: “covered” is more appropriate than “addressed”.

Thank you.

p. 2 line 10: “these obstructing factors” suggests there are many, but the impression from the previous sentence is that there is only one (poor representation of Middle East emissions in inventories).

Thank you. Here we refer to two main factors: the paucity of observations (such as AQABA measurements of the ambient concentrations of the trace gases and aerosols) and poorly represented strong emissions.

p. 2 line 17: “various” is vague.

Thank you. The range of the conditions is further specified in the same sentence.

p. 3 line 7: Where is the transported pollution coming from?

The Eastern Mediterranean is influenced by a mixture of European, North African, Middle East, and Asian sources. The latest research points out that northeastern Europe and northwestern Asia are likely the strongest sources. We refer the referee to a recent review by Dayan et al. 2017 (Section 2.3, https://doi.org/10.5194/acp-17-13233-2017) and to Lelieveld et al., 2002 (https://doi.org/10.1126/science.1075457) for further details and concepts.

p. 3, lines 7-10: This sentence is long and difficult to comprehend.

Thank you.

p. 4 line 26: Please specify the source of petrochemical industry SO2 and whether there are emission control technologies or other methods to address this source. This seems pertinent, given the findings in the study that anthropogenic emissions dominant submicron particles.

We agree that it would be helpful to add this information, but unfortunately the national and local authorities do not share it, hence the emission estimates remain uncertain, even though we

have been able to bring them in better agreement with measurements than was possible previously. The need to continuously improve emission inventories is of high priority and mentioned in the text.

p. 4 line 26: What contribution do ships make to SO2 and submicron sulfate?

We clarified the text specifying that marine traffic emissions establish the background sulfate pollution in remote and pristine areas. Revised Figure S6 underscores the contribution of the shipping emissions by showing an additional sensitivity simulation in which shipping emissions are switched off.

p. 5 lines 1-2: “much more spatially uniform” than what?

We added the “gas phase” for clarification.

p. 6 lines 15-22: Surprising to not see AERONET observations used to support the Angstrom exponent analysis. Especially given that AERONET is long-term, year-round, publicly available, and has many sites in the Middle East (https://aeronet.gsfc.nasa.gov/).

The AOD observations presented in this manuscript are, in fact, part of the AERONET network and available from the URL specified above. A “standard” CIMEL sunphotometer was employed during the AQABA campaign, except it was additionally coupled to a gimbal to compensate for the movement of the ship. Unlike stationary measurements at the dedicated Aeronet site, AQABA optical measurements are advantageous due to the wide spatial coverage and collocation with the rest of the observational suite. We see relatively limited useful information by adding the AOD comparison outside of the AQABA campaign.

p. 7, lines 23-24: Radiative forcing is determined by integrating over all wavelengths, so this conclusion doesn’t seem appropriate.

Thank you. We agree that the exact radiative forcing is determined by a number of factors including the spectral optical properties of the aerosols (extinction, single scattering albedo, phase function), surface albedo, cloud coverage, incident solar radiation at TOA, and illumination geometry. Nevertheless, the column AOD in visible alone largely regulates the magnitude of the energy balance perturbations exerted by the aerosols (e.g., see Figure 12 in Osipov et. al., 2015, doi:10.5194/acp-15-9537-2015). We think that our conclusion is appropriate since dust and anthropogenic aerosols contribute approximately equal to the AOD in the region.

Figure S7: The data could be more clearly labelled. Grey and brown are modelled NHx. Is orange just ammonium? If so, is this from the default or boosted model? Presumably blue “Obs” is ammonium only, given that NH3 wasn’t measured?

Thank you. We revised the legend in the figure. Figure S7 presents observations and two simulations (default and boosted NH3 emissions), as well as two diagnostics (NH4 and NHx). By definition NH4 <= NHx.
We use NHx to show that default EDGAR NH3 emissions are too low to explain observed NH4 concentrations. At the same time, simulation with boosted NH3 emissions represents the lower possible boundary of the emissions, since the majority of the time simulated NHx ~= observed NH4.

Reviewer #4 (Remarks to the Author):

The manuscript by Osipov et al. presents new measurements and model simulations evaluating PM2.5 and O3 distribution in the middle east. The measurements were collected during the ship-based AQABA campaign. Overall the paper emphasizes the anthropogenic contributions to these species and their impacts on human health and climate. This is interesting, as the region is generally assumed to be highly impacted by wind-blown desert dust, which is considered a "natural" source of PM2.5 in this context.
While the combination of new modeling and field campaign data lead to several interesting conclusions about the region, overall I found this paper was not ready for publication for the following reason — the main manuscript text is too often insufficient in terms of explanation and clarity, and in many instances makes claims that are not well supported by the data provided. This paper requires substantial rewriting and revision, at a minimum. It's not clear to me if it would be a good fit for this journal unless the authors choose to focus the work on a subset of aspects that can be more fully, and more carefully, explained.

Thank you for your supportive comment and the constructive criticism. We agree that more explanation was desirable, and actually the Editor has given us additional text space to meet this requirement. Following your recommendation, we revised the manuscript and transferred/adapted a large portion of the supplementary information into the main text. Please also refer to the top of the document for the summary of the improvements.

A fundamental concern is that the authors have done a rather poor job of separating text between the main manuscript and the supplemental. I was often left confused or surprised by the former only to spend time going through the latter to find out that the data int the supplement itself doesn't support the statements of the main text, or that the basis for statements in the main text are not provided at all.

Thank you. We revised the manuscript and the reader should be able to find the needed evidence in the main text supporting every claim.

For example, in the main text the authors directly state that the major fraction of severe air pollution health impacts in the Middle East, and Kuwait in particular, is of anthropogenic origin.

However, the results provided in the supporting information show this is more likely 20%. My comments below expand on this, and many other instances.

We revised the text to avoid possible misinterpretation of our results. The manuscript now clearly states the total and anthropogenic PM2.5 concentrations region-wise (in the Middle East). On the country level, we clarified on p11 line 20 that severe health effects in a number of countries (including Kuwait) are due to the overall (natural and anthropogenic, rather than anthropogenic only) air pollution.

That being said, there are many elements of the main text that are supported by their results, and there are many interesting findings. I just wish they were presented more completely, and fear that a brief-format article like this is not well suited for an in-depth analysis of a myriad of topics (deposition, climate, emissions, health impacts) stemming from this field campaign. The authors just need to be a lot more careful about backing up their claims and explaining their results precisely, and not glossing over some parts of their datasets or modeling that runs counter to the overall message they wish to get across. Below are many comments and questions on aspects that are mostly claims made by the authors where the evidence is not provided, or where the details run counter to their summary.

Thank you. We revised the manuscript focusing more on our main results: characterizing the atmospheric chemistry in the region and underscoring the role of the anthropogenic emissions in regulating the air quality and the radiative impacts on climate in the region. We have addressed a number of cases in the text, which were causing the misinterpretation. The revised text should be unambiguous in the interpretation of our results.

Major comments (page.line number):

Fig 1: The large anthropogenic faction of fine PM over the Mediterranean and extending over southern Europe is surprising. Is there no biogenic or sea-salt PM2.5 in Greece? Or is this a consequence of the boundary conditions used, or the treatment of natural aerosol sources? Looking at page 4 of the supplemental, I don't see any description of biogenic emissions or sea salt aerosols. Without these components, fig 1(d) is potentially highly misleading in some areas. I suggest the authors remove or limit the scope of this figure. While analysis in Fig 2 likely indicates these components are negligible along the ship measurement track, with perhaps the exception of significant sea salt in the Arabian Gulf and Arabian Sea, the map in fig 1(d) extends much beyond this, spatially.

It is correct that we confine our validation and description efforts to a smaller region than the entire modeling domain, moving away from the influence of the boundary conditions and letting the model develop its own dynamics and chemistry, especially the influence of the anthropogenic emissions within the domain. Nevertheless, in the specific case of the Mediterranean, the model produces results that are consistent with the observations.

The figure below shows the AMS measured aerosol composition averaged over the Mediterranean sea only. The composition is dominated by ammonium sulfate and organics. Further, the organic aerosols are largely primary (the median value of the POA to total OA fraction is 0.99). This excludes the influence of the biogenic emissions, assuming that biogenic emissions are mostly VOC precursors that oxidize into secondary organic aerosols.

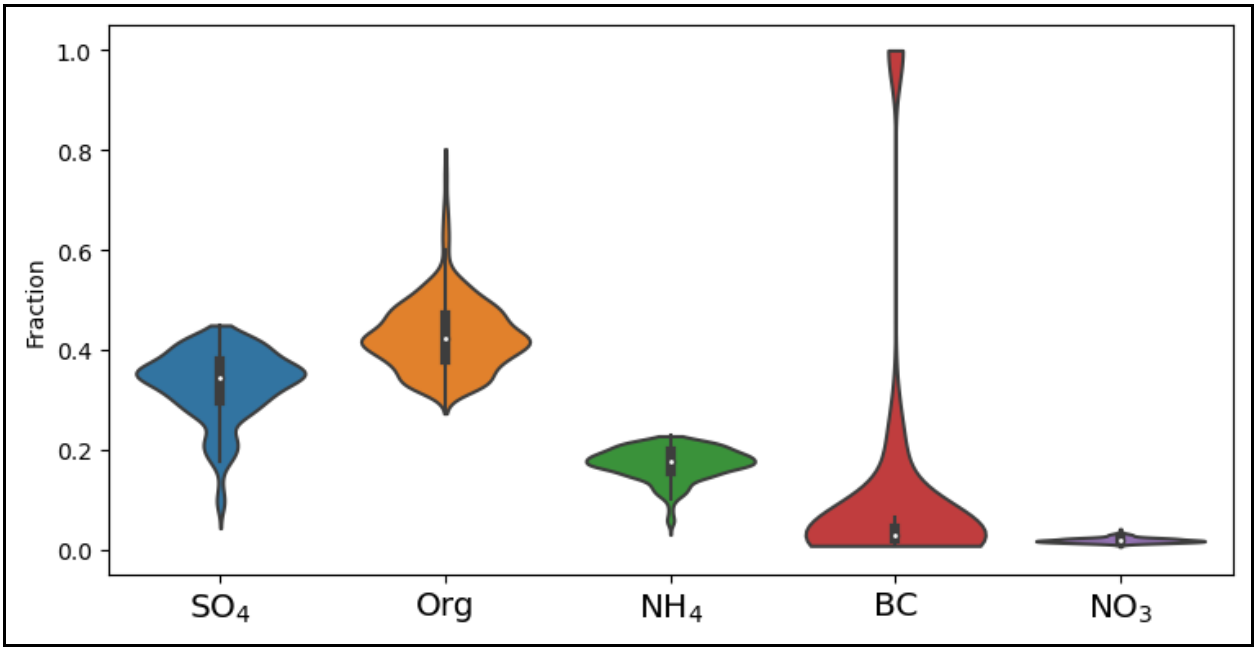

Figure 1. Violin plots of the composition of the AMS PM1 averaged over the Mediterranean Sea.

Regarding sea salt and PM2.5, we would like to note that Figure 1 shows fine PM (i.e. PM1, defined on page 2) and the large concentration of the coarse aerosols in PM2.5 or PM10 (e.g. dust or sea salt) does not imply large concentrations in PM1. In fact, we showed that the dust contribution reduces from approximately 50% in PM2.5 (region-wise) by an order of magnitude to less than ~5% in PM1. A similar conclusion holds for sea salt and is shown in the figure below (PM1000 approximates total sea salt concentration).

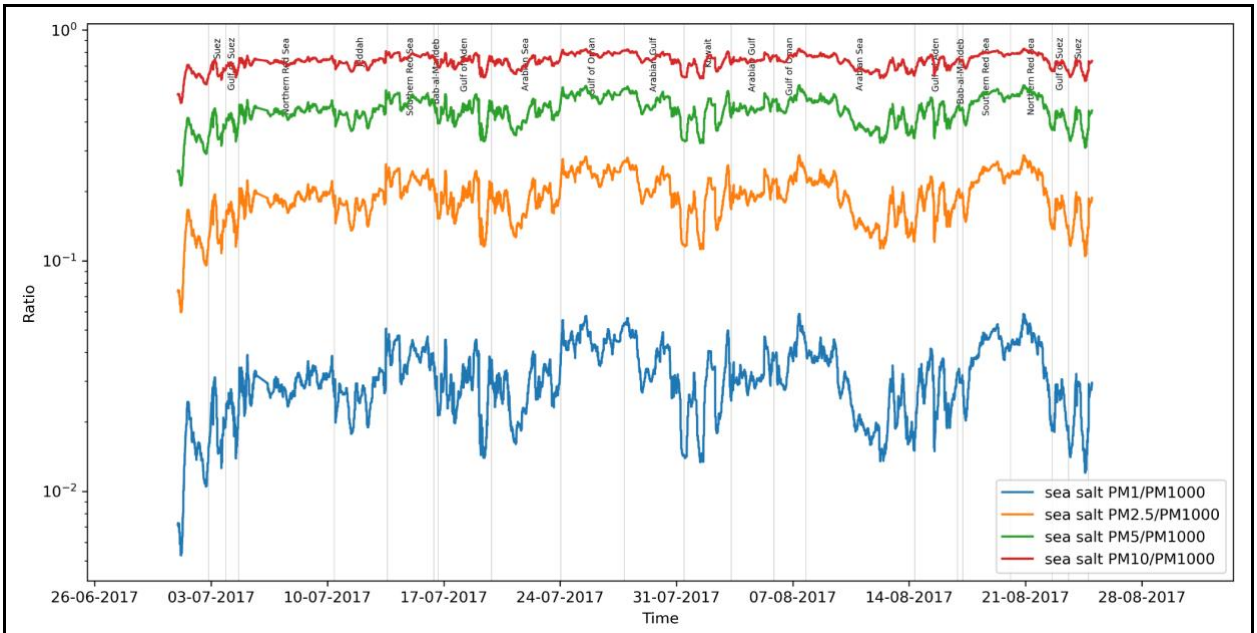

Figure 2. The fraction of the sea salt concentration, integrated from 0 microns to the specified size (i.e. PM1, PM2.5), to the total sea salt concentration.

Finally, the referee does not provide references to support his/her statement that "fig 1(d) is potentially highly misleading". Given our argumentation above, we think that Figure 1 accurately describes the partitioning of the PM1 between anthropogenic and natural components and underlines the various conditions over the seas, e.g. polluted in the Mediterranean Sea and pristine in the Arabian Sea.

We repeated the clarification in the model description section (SI), that we do not model the biogenic emissions.

4.17: This glosses over some areas where it seems the IC results indicate sea salt could be important (Arabian Gulf? it's a bit hard to tell from the figure) or where there are large discrepancies between total PM and spectated PM (Arabian Sea). While the overall conclusions aren't in question, these other features do warrant consideration and discussion.

Thank you. Indeed, there are two marine regions (the Arabian Sea during ship leg 1 and the Arabian Gulf during ship leg 2), where the speciated AMS PM1 does not explain the total PM1 completely, pointing at the contribution of other aerosol species. We employed the IC measurements to show that over the Arabian Gulf aerosol water (not sea salt) fills the gap. Unfortunately, we can not apply the same approach for the Arabian Sea (leg 1) case due to the lack of IC measurements during that period. Dust and sea salt are the most likely candidates to fill the gap over the Arabian Sea, but we do not have the observational evidence to support that, hence the lack of speculation in the text. We would like to emphasize that the Arabian Sea segment (leg 1) is the only and relatively short segment when a considerable difference between total and speciated PM1 was observed.

5.12: Regarding the areas impacted by deposition, is there a figure or quantitative result that can be provided here? How much sulfate is deposited? What is meant by “intensely” or “particularly affected”? If the details on this topic can’t be provided, maybe it should be omitted?

We added Figure S16 showing the dry deposition maps.

5.18-25 and Fig S7: The authors have not carefully considered the issues with their NHx simulations or what this may tell them about emissions. The should consider the following:

Thank you, we have updated figure S7 and revised the manuscript to clarify our NHx analysis. Please note that given the magnitude of the underestimation of the NH3 emissions in the Middle East, simulating the reactive nitrogen accurately does not seem feasible at the moment. Instead, our main goal is to underscore that conventional inventories (i.e. EDGAR) underestimate NH3 emissions in the Middle East. Our second goal is to deduce the lower estimate of the NH3 emissions and to capture the lower envelope of the observed NH4. For this purpose, we use total ammonia NHx and observe that NH4<=NHx by definition.

- How did the authors determine that ammonia emissions are underestimated by a factor of 15 (as opposed to some other factor)?

Please, refer to our comment below.

- Why did the authors choose to increase NH3 emissions from agricultural sources and not industrial sources? In the supplemental, they state that “sensitivity analysis” shows that point sources could not explain the missing NH3. What sensitivity analysis was conducted, and how was this concluded? Is there another set of modeling results where the point sources were increased but the area sources were not? The authors state that increasing the area sources lead to “capturing the bulk of the ammonia variability” — can this statement be quantified? A proper comparison would be to e.g. present simulation results for different cases (increasing area sources, increasing point sours, or some combination of both) along with quantitative statistical evaluation of correlation and error compared to observations, and then draw conclusions from this.

There are two types of emissions: point and area sources. We identified which type regulates the NH4 concentrations by performing sensitivity tests. We tested pointed sources first due to SO2 precedent (described on page 5). In the manuscript, we explain that simulated NHx concentrations are weekly sensitive to the strength of the point sources. In the mathematical sense, the inclusion of the van Damme et al. point sources (~10x stronger than in EDGAR), results in NHx concentrations that are much smaller than 10 x NHx in the reference EDGAR run. We concluded that area sources largely regulate the NHx concentrations. Next, we identified the main area source of NH3 emissions (agriculture) and enhanced them. Figure S7 shows that NHx concentrations are indeed sensitive to the strength of the area sources (unlike the point sources). The factor of 15 was chosen to establish simulated NHx concentrations that

serve as the lower envelope of the observed NH4. We observe that NHx >= NH4 by definition and that NHx does not depend on the phase partitioning (gas or aerosol).

We appreciate the referee's motivation to quantify exactly the NH3 emissions in the region. However, given the major uncertainty in the NH3 emissions in the region, we see limited added value in applying exact statistical quantification. We expect that a number of NH3 sources in the Middle East are not represented at all in the EDGAR inventory (similar to black carbon and sulfur emissions) and the problem extends beyond the determination of a single factor. Clearly, further research is required to describe and constrain the reactive nitrogen chemistry in the region.

- NHx deposition is also highly uncertain given the dependence on precipitation which is often not well simulated — how much could this contribute to the discrepancies noted for NHx between the observations and model?

There are no indications that the model simulates NHx deposition incorrectly. Wet deposition does not play a role. We have validated a number of other trace gases and aerosols, establishing that PBL height and turbulent mixing are representative. Further, our conclusion that NH3 emissions are underestimated is supported by an independent study by van Damme et al.

- Also, if the model's NOx and SO2 levels are not correct, then partitioning of NHx can be impacted; since NH3 and NH4 have different deposition rates, incorrect partitioning can also contribute to errors in simulated NHx.

Thank you. In Figure S6 we established that SO2 levels and sulfur chemistry are representative. We employ the total ammonia NHx to minimize the influence of the partitioning on the comparison. Further analysis of the NHx partitioning, as well as pH and effects of the crustal cations, is contingent on the reliable NH3 emissions. These scientific questions are beyond the scope of the current study and will be studied in the future.

- In Fig S7, right, panel, if the grey is the default simulation and the brown in the boosted simulation, what do the orange markers represent? They are labeled as "Model", but they don't match either the default or boosted model simulations. They are not directly explained in the figure caption or text. Also, why are there no model values to compare to observed values during the first ~week of the campaign and the last ~week? For these periods only the blue "Obs" markers are shown. It's odd that the timeseries in this figure, in terms of dates, does not match that of the previous figures (S4-S6). Where was this data collected, why is it not modeled, and why is it not shown in the other figures (S4 - S6)?

We revised Figure S7. Timeseries in Figures S4-S7 are consistent now. We restrict the comparison of the observed and simulated diagnostics to the region of interest, which starts with the ship entering Suez.

6.1-5: How do the authors distinguish, quantitatively, between BC from shipping vs BC from urban sources? It seems that additional model simulations with tagged or perturbed emissions would be necessary to make this distinction, or evaluation of ratios of species with known sources. This is another example of claims being made that, while reasonable, are not backed up with any analysis or quantitative evidence. ….. Ok, after reading the supplemental in full, I see where the evidence for some of these claims are provided. Still, the summary in the main text doesn't do a good job of providing quantitative evidence or explanation, even if only briefly, of how conclusions were drawn. It seems overall that the split between main text and supplemental text isn't working well for this paper, and I'd suggest a different, longer, format and thus possibly a different journal. Regardless, the supplemental provides details such as the contribution of biomass burning to background BC being up to 40% in the seas surrounding the Arabian Peninsula, whereas the main text says this is from shipping. Lastly, it's not explained how spectral measurements are used to distinguish between fossil fuel and biomass burning BC — please explain this technique.

Thank you. We understand that finding the quantitative metrics to support our claims turned out harder than it could be. To address this problem, we integrated/adjusted much of the SI into the main text expanding the manuscript up to the allowed limit of 5000 words. Additionally, we updated the figures and added statistical metrics to ease the interpretation of the results.

We readily distinguish between BC from shipping vs BC from urban sources by switching off the shipping emissions in the simulations and evaluating the corresponding difference in BC concentrations. Model runs without shipping emissions underestimate the BC concentration by a factor of 3 and larger in the pristine marine environments, while BC concentration near the urban centers are not sensitive to shipping emissions. This result is consistent with the source attribution done via aethalometer measurements shown in Figure S9. We do not plot the run without shipping emissions in Figure S8, because the figure is already quite busy. We think that the text description is sufficient. Instead, we updated Figure S9, which now clearly shows the fraction of the fossil fuel BC. According to the updated figure S9, the fossil fuel fraction is 3 times larger (median values) than the biomass burning fraction and represents 75% of the BC concentrations during the campaign. This is consistent with our conclusion that the pronounced BC pollution over the seas that surround Arabian Peninsula is due to shipping activity.

Regarding the spectral technique to distinguish between FF and BB BC, we refer the reviewer to Zotter et al., 2017, which is cited in our manuscript and describes the technique in detail.

S10.7: How are dust mass fluxes "optimized"? This isn't explained at all. I'm guessing from the subsequent discussion of dust emission size distributions that here they referring to using a uni-modal size distribution (coarse mode only) and omitting accumulation mode dust. But this is a single selection, not an optimization. They don't show model performance for different choices regrading the dust size distribution. They also don't adopt the coarse mode shape from Kok (2011) which has been found to improve the representation of coarse mode dust in other models.

We have added Ginoux et al., 2001 citation, which describes the conventional approach to tuning dust emissions. The overall dust flux into the atmosphere is adjusted via a single dimensional factor C. The value of C is chosen to minimize error in simulated AOD and PM concentrations. Our unimodal size distribution of the emitted dust indeed does not rely on and does not require information about the partitioning of the emitted flux between accumulation and coarse modes, but does not eliminate the step of optimizing the C factor.

Regarding the shape describing the emitted dust, our size distribution is qualitatively identical to that of Kok (2011) for the particles up to 10 μm in diameter. In the case of the larger particles, the paucity of the coarse dust measurements limits our ability to validate the simulated concentrations and to alter the emitted size distribution meaningfully.

Regarding the referee's statement "They don't show model performance for different choices regrading the dust size distribution.", we do not think that plotting all model runs is the best way to convey our results clearly. We think that these intermediate steps are not necessary and prefer to present the final results to the reader. Nevertheless, such intermediate results will be made available on request.

6.5-15: While I appreciate the result that in some areas the 500 nm AOD is dominated by anthropogenic PM is interesting, I think further discussion is warranted. First, it is commonly known that dust is primarily coarse and thus not the main contributed to AOD at 500 nm. Coarse-mode AOD from e.g. MODIS is derived using other wavelengths. So this result should be put a bit more in context. Second, the feature isn't universal in their reported observations. There are parts of the campaign (e.g. Arabian Gulf) where the modeled AOD is significantly larger than the observations, and the reason isn't explained.

Thank you. To address the referee's comment, we would like to separate the radiative forcing of climate from remote sensing and spectral aerosol optical properties.

In the manuscript, we discuss the radiative forcing of climate (not remote sensing) and employ AOD at 500 nm as the primary diagnostic to make a point that aerosols (natural or anthropogenic) perturb energy balance in the Middle East. AOD at 500 nm is conventionally used as a proxy for the aerosols radiative forcing (e.g. doi:10.5194/acp-12-5129-2012) due to two main reasons:
1. Shortwave energy flux at the TOA peaks around 500 nm.
2. Normally, the spectral optical depth of aerosols peaks in visible and is much more uniform spectrally compared to trace gases.

Here, we combined the previously established result, that aerosols are the key climate driver in the Middle East due to strong radiative forcing (Osipov et al., 2018, Osipov et al., 2015, described in the introduction) with the AOD source attribution following the AQABA results (~50% AOD is anthropogenic) to establish that anthropogenic aerosols are a key factor in the radiative forcing of climate. We think that our conclusion is scientifically sound.

We find the argument regarding the MODIS is of limited relevance since the choice of the wavelengths is individual for each instrument and optimized for a completely different purpose: AOD retrieval rather than climate forcing.

7.5-7: I'm not sure what they are trying to say here. If half of PM2.5 is desert dust, then how can the health impacts be mostly owing to anthropogenic PM2.5? Wouldn't that also be half? Or is this based on the number of people exposed in areas that have a higher anthropogenic fraction? I looked in the supplemental, and it doesn't appear they have calculated the attributable fraction specifically for anthropogenic PM2.5, so this is a bit hard to say. They do provide the PM2.5 concentrations by total vs anthropogenic, in table S2. While there are some countries for which anthropogenic dominates (Egypt) there are others where it doesn't (Saudi Arabia). Several in the latter category are those with severe health impacts as noted non line 11/12, yet line 13/14 misleadingly implies that a major fraction of these health impacts are controllable. In fact, for the country with the most severe %-based health impacts (Kuwait), it is likely only ~20% is anthropogenic (assuming
direct relationship between PM2.5 and deaths, which is an oversimplification and the authors should be able to estimate this more precisely). Hence, a more thorough analysis and careful discussion is needed.

Thank you for suggesting improvements here. We revised the manuscript and clarified the fraction of the PM2.5 above the recommended threshold is anthropogenic and avoidable. Further, we clarified the text and reference the region-wise diagnostics, that take into account the total population in the Middle East.

6.20 and 7.24: The authors have conflated AOD at 500 nm directly with radiative forcing, which is incorrect (one needs to consider the entire short-wave spectrum). Also, by definition "radiative forcing" is the impact of anthropogenic activity on the radiative balance, and thus to conclude that it is mostly anthropogenic is somewhat nonsensical. The authors should more correctly refer to the "radiative effects" not the "radiative forcing" of PM it is the former which includes both anthropogenic and natural causes. Their analysis is also limited direct effects; indirect effects are not discussed nor considered.

Thank you.
Regarding the AOD-RF (radiative forcing) discussion, please, refer to our comment above (6.5-15).
Regarding the definition of the RF, unfortunately, the reviewer does not provide a reference to support his statement. We use the term "radiative forcing" consistent with the IPCC definition (https://www.ipcc.ch/site/assets/uploads/2018/03/TAR-06.pdf). We quote: "The term "radiative forcing" has been employed in the IPCC Assessments to denote an externally imposed perturbation in the radiative energy budget of the Earth's climate system. Such a perturbation can be brought about by secular changes in the concentrations of radiatively active species (e.g., CO2, aerosols), changes in the solar irradiance incident upon the planet, or other changes that affect the radiative energy absorbed by the surface (e.g., changes in surface reflection properties)". In the context of this discussion, radiative forcing is an external forcing, which can

be natural or anthropogenic. For example, the radiative forcing due to solar variability (natural). We think that the term “radiative forcing” is used appropriately in the manuscript.

General: The authors should take a look at the recent Nature Communications paper from McDuffie et al. (https://www.nature.com/articles/s41467-021-23853-y.pdf), which I consulted for comparison to their claims that anthropogenic sources are not a known major impact in this region, and for the evaluation of the PM2.5 concentrations and associated health burden. On the former topic, be aware that they show, for example, that slightly more than 50% of the PM2.5 in Egypt is of anthropogenic origin. They also estimate a similar number of premature deaths as the present study (about 90,000). They report a population-weighted PM2.5 concentration of 66 ug/m3 whereas the present work reports a non-weighted concentration of 40 ug/m3.

Thank you for the great reference. We included it in the revised manuscript. Please note that PM2.5 is population-weighted in both cases (our manuscript and McDuffie et al.). Regarding the comparison with Egypt, in both cases, the anthropogenic PM2.5 agrees well (~30-35 ug/m^3), while McDuffie et al. report higher dust concentration in PM2.5 (~30 ug/m^3) than our result (~9 ug/m^3). This is consistent with our concussion that the accumulation mode of the dust emitted size distribution produces excessive PM2.5 concentrations.

Minor comments and corrections (page.line number):

3.12 and elsewhere: I don’t think it is advisable to refer to specific features of supplemental figures, as they need to be literally supplemental. If there is a specific feature that warrants discussion, it should be fit into the main manuscript.

Thank you. We merged a large part of the SI with the main text.

2.28: This phrase is a bit odd/misleading as the strongest associations of fine PM and health are based on mass concentration, not number concentration. There are though stronger associations between fine than coarse PM mass concentrations. I suggest the text be rewritten to be more accurate in this regard.

We acknowledge that the representation of the health risks, associated with the exposure to particulate matter, is based on the mass concentrations (i.e. PM2.5). Nevertheless, this representation is statistical and only highlights that mass concentration diagnostics are more readily available compared to the number concentrations. It does not indicate that mass concentrations have a “stronger association” than number concentrations.

We underscore that smaller particles, i.e. PM1, have larger damage potential and quote (https://doi.org/10.1016/j.scitotenv.2021.145509): “*Furthermore, the size (in diameter) of PM_1 is much smaller than $PM_{2.5}$ and PM_{10}, which enables the pollutant to reach lower lobes of a lung and may cause a larger damage to respiratory health."*)“. This is consistent with our statement, that fine PM dominates the number concentration.

Further, we quote (https://pubmed.ncbi.nlm.nih.gov/26993200/): “Unlike PM2.5, UFPs may have much more impacts on human health considering loads of evidences emerging from particulate matters and nanotoxicology research fields.”, where UFP stands for ultrafine particles.

3.1 Should the title of this section be O3 instead? Components of PM2.5 are produced by photochemical reactions.

Thank you, we updated the title.

3.9: I don’t think “indigenous” is the right word here. Perhaps they meant anthropogenic?

We use “indigenous” according to the definition to distinguish local emissions from long-distance transport.

3.14: How is it known that O3 isn’t exported in the form of precursors (e.g., PAN) rather than directly as O3?

We stated that several regions (Gulf of Suez, the northern Red Sea, Arabian Gulf) are the exporters of ozone. In these regions, both the highest ozone concentrations and largest net ozone production rates are observed simultaneously, thus excluding the indirect export of ozone through the precursors as the possible main pathway. But indeed some other pollutants such as organic nitrates have the potential to also contribute to ozone formation outside the area of our analysis, but we feel that this aspect is beyond the scope of the present work.
We also added the Tadic et al. citation, which provides a comprehensive analysis of the ozone chemistry during the AQABA campaign.

Fig S3: Provide statistical evaluations for each (gases, aerosols, optical properties), separated by species or property.

We revised the figure and added statistical evaluations.

Fig S4: Please define acronyms used in the figure (e.g., NRS).

We expanded the acronyms.

3: How is net O3 production diagnosed, both from the observations and from the model?

We cited Tidac et al., 2020 one more time, which describes the net O3 production during the AQABA campaign in detail.

4: How is it diagnosed that the region is NOx-limited? Were they considering ratios of particular species to make this determination empirically, is it based on model evaluation, or other means?

Please refer to the clarification above.

S11.23: What is the “spectral loss function”?

We added clarification in the text explaining that the spectral loss function expresses the loss of the particles at the inlet as a function of aerodynamic diameter.

S12.22: The sentence “The previously underrated…” I’m not sure what emissions they are referring to. Could this be re-written or explained more fully?

These emissions are referenced in the introduction (p2.11)

17th Jun 22

Dear Professor Lelieveld,

Your manuscript titled "Severe atmospheric pollution in the Middle East" has now been seen by our reviewers, whose comments appear below. In light of their advice I am delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment under the open access CC BY license (Creative Commons Attribution v4.0 International License) if you can address the remaining concerns of the reviewers.

We therefore invite you to revise your paper one last time to address the remaining concerns of our reviewers. In particular, you should revisit the comments of Reviewer #3 from the previous round and provide clear details on the changes made to the main text in response to these concerns (see report below for more detail). Similarly, please address the remaining concerns of Reviewer #4 regarding the accuracy of the language around radiative effects and Figure 2. At the same time we ask that you edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table". Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at https://www.nature.com/documents/commsj-file-checklist.pdf .

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication under a CC BY license (Creative Commons Attribution 4.0 International License). This license allows maximum dissemination and re-use of open access materials and is preferred by many research funding bodies.

For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit https://www.nature.com/commsenv/article-processing-charges

At acceptance, you will be provided with instructions for completing this CC BY license on behalf of

all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:
[link redacted]
** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Clare

Dr Clare Davis
Senior Editor
Communications Earth & Environment

www.nature.com/commsenv/
@CommsEarth

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

The authors have done a great job of improving their manuscript. However, I have some small recommendations. Authors should use more statistical information in Abstract and Conclusions sexies.
The authors may cite these works:
http://dx.doi.org/10.1016/j.rse.2013.12.021
https://doi.org/10.1007/s41651-021-00093-8
https://doi.org/10.1007/s00704-017-2362-8

Reviewer #2 (Remarks to the Author):

Thank you for the revised version and addressing my comments in depth. I have no further comments.

Previous general summary :
The paper presents findings related to the composition of air pollution in the Middle East region. It shows results concluding that anthropogenic origin of air pollution is important, unlike the popular belief that desert dust is dominant in aerosol composition. The paper therefore concludes that there

is room for action to tackle air pollution of human origin in the region and its impact on health, ecosystems, and climate change.
The paper is important in addressing critical data gaps and understanding of the sources of air pollution in the Middle East region. Triggering efficient action to reduce anthropogenic air pollution is especially important in regions affected by natural dust.

Reviewer #3 (Remarks to the Author):

The responses to reviewer comments weren't entirely satisfying. It wasn't always clear whether responses to my original reviewer comments (Reviewer 3) had been addressed in the manuscript. Specifically, these include the following:

The only response was "Thank you." to the comment "The manuscript would also benefit from a brief discussion of the broad applicability of the measurements obtained over a single season and year to other seasons and years.". This would strengthen the applicability of the findings from your study.

The response was "Thank you. Here we refer to two main factors: the paucity of observations (such as AQABA measurements of the ambient concentrations of the trace gases and aerosols) and poorly represented strong emissions." to the comment starting "p. 2 line 10". Is wasn't clear whether the manuscript had been updated to inform the reader of this too.

Similar to the above response, the comment starting with "p. 3 line 7" referred the reviewer to additional information, but it wasn't apparent that this same information has been provided for future readers of the manuscript too.

Only the response "Thank you." was give for the comment "p. 3, lines 7-10: This sentence is long and difficult to comprehend.". A reader would find it easier to follow what's written if this split into shorter, easier to digest sentences. Currently it's opaque.

Reviewer #4 (Remarks to the Author):

The revisions are commendable and a significant improvement. That being said, prior to publication, please address the following:

- In the abstract, please revise “Hitherto it was understood that dessert dust dominate…” to account for recent work showing that the anthropogenic contribution to PM2.5 in e.g. Egypt is >50% (https://www.nature.com/articles/s41467-021-23853-y.pdf)

- In the abstract, please specify that the following is referring only to 500 nm: “anthropogenic pollution particles account for a major part of the aerosol optical depth.”.

- Fig 2 should explicitly state in the caption that biogenic emissions are not considered in this simulation, rather than have this factor (which is essential to interpreting this figure) provided only in the SI. Otherwise, I still contend that this figure misleadingly represents the anthropogenic

fraction in Greece, where biogenic sources are non-negligible (e.g., https://acp.copernicus.org/articles/19/901/2019/). additionally, lease specify the simulation time period in the caption, and whether or not biomass burning is considered anthropogenic or natural in this context.

- In response to anonymous reviews, the authors should not refer to the reviewer using gendered pronouns.

- The authors provide the correct definition of “radiative forcing” from the IPCC, yet they still miss the point of my previous criticism. As defined by the IPCC, radiative forcing is a perturbation in the energy budget brought about by e.g. changes in concentrations of radiatively active species. Thus, when discussing radiative fording, one must always define the “change” to which the forcing refers. Often this change is the change in concentrations from pre-industrial to present, and this is explicitly stated in numerous studies discussing the radiative forcing of atmospheric constituents. In fact, it is so often this particular change that without any further qualification one typically assumes that the radiative forcing of an atmospheric constituent is in reference to its concentration change since pre-industrial times, i.e., its anthropogenic component. If instead one is considering the change of concentrations relative to a zero-concentration scenario, then the radiative forcing would also include natural contributions to the present day concentrations. There’s even a term for this specific change — “radiative effect”. If the authors need further references on this topic, they should consult Heald et al., 2014 (“Contrasting the direct radiative effect and direct radiative forcing of aerosols”, ACP) the first sentence of the abstract of which is "The direct radiative effect (DRE) of aerosols, which is the instantaneous radiative impact of all atmospheric particles on the Earth's energy balance, is sometimes confused with the direct radiative forcing (DRF), which is the change in DRE from pre-industrial to present-day (not including climate feedbacks)”. Why belabor this point? Well, the authors of the present manuscript seem to make a scientific conclusion out of identifying the anthropogenic contribution to radiative forcing. Within the definition as described here though, it’s evident that such a conclusion is tautological. Further, the authors did not address the other aspect of my criticism, which is that they haven’t distinguished between direct and indirect (cloud-interaction) radiative forcings. In short, the authors should change the phrase “radiative forcing” to “direct radiative effects” on page 10. In practice this is a very small revision to make, which would improve the scientific accuracy of their paper, so I'm not why there's any resistance.

The authors have done a great job of improving their manuscript. However, I have some small recommendations. Authors should use more statistical information in Abstract and Conclusions sexies.

The authors may cite these works:

http://dx.doi.org/10.1016/j.rse.2013.12.021

https://doi.org/10.1007/s41651-021-00093-8

https://doi.org/10.1007/s00704-017-2362-8

We thank the referees for reviewing our manuscript. Below we provide the point-by-point responses.

Reviewer #1 (Remarks to the Author):

The authors have done a great job of improving their manuscript. However, I have some small recommendations. Authors should use more statistical information in Abstract and Conclusions sexies.
The authors may cite these works:
http://dx.doi.org/10.1016/j.rse.2013.12.021
https://doi.org/10.1007/s41651-021-00093-8
https://doi.org/10.1007/s00704-017-2362-8

We added further statistical information in the Abstract indicating the anthropogenic fraction in PM1 and visible AOD.

The latest version of the manuscript already contains a number of references larger than recommended (50), which, unfortunately, prevents us from citing additional works.

Thank you for reviewing our manuscript.

Reviewer #2 (Remarks to the Author):

Thank you for the revised version and addressing my comments in depth. I have no further comments.

Thank you for reviewing our manuscript.

Previous general summary :
The paper presents findings related to the composition of air pollution in the Middle East region. It shows results concluding that anthropogenic origin of air pollution is important, unlike the popular belief that desert dust is dominant in aerosol composition. The paper therefore concludes that there is room for action to tackle air pollution of human origin in the region and its impact on health, ecosystems, and climate change.
The paper is important in addressing critical data gaps and understanding of the sources of air pollution in the Middle East region. Triggering efficient action to reduce anthropogenic air pollution is especially important in regions affected by natural dust.

Reviewer #3 (Remarks to the Author):

The responses to reviewer comments weren't entirely satisfying. It wasn't always clear whether responses to my original reviewer comments (Reviewer 3) had been addressed in the manuscript. Specifically, these include the following:

The only response was "Thank you." to the comment "The manuscript would also benefit from a brief discussion of the broad applicability of the measurements obtained over a single season and year to other seasons and years.". This would strengthen the applicability of the findings from your study.

We updated the manuscript (page 11) and included a remark about the need to extend the measurements, the data analyses and model optimization to other seasons and years.

The response was "Thank you. Here we refer to two main factors: the paucity of observations (such as AQABA measurements of the ambient concentrations of the trace gases and aerosols) and poorly represented strong emissions." to the comment starting "p. 2 line 10". Is wasn't clear whether the manuscript had been updated to inform the reader of this too.

We revised the manuscript and explicitly stated the factors: poor representation of the emissions and lack of observations.

Similar to the above response, the comment starting with "p. 3 line 7" referred the reviewer to additional information, but it wasn't apparent that this same information has been provided for future readers of the manuscript too.

We revised the manuscript and added clarification about the sources of the transported pollution.

Only the response "Thank you." was give for the comment "p. 3, lines 7-10: This sentence is long and difficult to comprehend.". A reader would find it easier to follow what's written if this split into shorter, easier to digest sentences. Currently it's opaque.

The sentence was further simplified.

Thank you for reviewing our manuscript.

Reviewer #4 (Remarks to the Author):

The revisions are commendable and a significant improvement. That being said, prior to publication, please address the following:

- In the abstract, please revise “Hitherto it was understood that dessert dust dominate…” to account for recent work showing that the anthropogenic contribution to PM2.5 in e.g. Egypt is >50% (https://www.nature.com/articles/s41467-021-23853-y.pdf)

We replaced “Hitherto” with the more appropriate “Conventionally”.

- In the abstract, please specify that the following is referring only to 500 nm: “anthropogenic pollution particles account for a major part of the aerosol optical depth.”.

Thank you. We revised the sentence and added “visible”.

- Fig 2 should explicitly state in the caption that biogenic emissions are not considered in this simulation, rather than have this factor (which is essential to interpreting this figure) provided only in the SI. Otherwise, I still contend that this figure misleadingly represents the anthropogenic fraction in Greece, where biogenic sources are non-negligible (e.g., https://acp.copernicus.org/articles/19/901/2019/). additionally, lease specify the simulation time period in the caption, and whether or not biomass burning is considered anthropogenic or natural in this context.

Thank you.
We clarified the caption indicating the biogenic emissions are not considered in the simulation.
We specified the time period in the caption.
We specified that biomass burning is considered anthropogenic. Thank you.

- In response to anonymous reviews, the authors should not refer to the reviewer using gendered pronouns.

We agree with the reviewer and apologize that this has happened, though unintentionally.

- The authors provide the correct definition of “radiative forcing” from the IPCC, yet they still miss the point of my previous criticism. As defined by the IPCC, radiative forcing is a perturbation in the energy budget brought about by e.g. changes in concentrations of radiatively active species. Thus, when discussing radiative fording, one must always define the “change” to which the forcing refers. Often this change is the change in concentrations from pre-industrial to present, and this is explicitly stated in numerous studies discussing the radiative forcing of atmospheric constituents. In fact, it is so often this particular change that without any further qualification one typically assumes that the radiative forcing of an atmospheric constituent is in reference to its concentration change since pre-industrial times, i.e., its anthropogenic component. If instead one is considering the change of concentrations relative to a zero-concentration scenario, then the radiative forcing would also include natural contributions to the present day concentrations. There’s even a term for this specific change — “radiative effect”. If the authors need further references on this topic, they should consult Heald et al., 2014 (“Contrasting the direct radiative effect and direct radiative forcing of aerosols”, ACP) the first sentence of the abstract of which is "The direct radiative effect (DRE) of aerosols, which is

the instantaneous radiative impact of all atmospheric particles on the Earth's energy balance, is sometimes confused with the direct radiative forcing (DRF), which is the change in DRE from pre-industrial to present-day (not including climate feedbacks)". Why belabor this point? Well, the authors of the present manuscript seem to make a scientific conclusion out of identifying the anthropogenic contribution to radiative forcing. Within the definition as described here though, it's evident that such a conclusion is tautological. Further, the authors did not address the other aspect of my criticism, which is that they haven't distinguished between direct and indirect (cloud-interaction) radiative forcings. In short, the authors should change the phrase "radiative forcing" to "direct radiative effects" on page 10. In practice this is a very small revision to make, which would improve the scientific accuracy of their paper, so I'm not why there's any resistance.

Following the reviewer's recommendation, the phrase "radiative forcing" has been replaced by the "direct radiative effects" on page 10.

Thank you for reviewing our manuscript.