

Reviewer 1

General comments

1. Comment

This paper presents a substantial effort to compare profiles from multiple models to measurements acquired by aircraft. However, there are many grammatical errors, formatting errors in figures, lack of definitions, and lack of robust conclusion statements that would help a paper of such size.

Reply. We have spent considerable effort to improve the presentation of the paper in all its part, from editing grammar and improving to figures. Conclusions have been completely rewritten. The paper has been proof-read by two native English speakers.

2. Comment

Although I have not made comments for the entire paper, I have a list of suggested edits that should be considered. Do not mistake my lack of comments in the middle of the paper as a pass of prose. By and large, the writing lacks a thorough proof-read, and would benefit greatly from a major revision in terms of writing prose (i.e. proper use of tense, pluralization, sentence structure, etc.). Many paragraphs could be synthesized better to reduce total size and be clearer to the reader.

Reply. The paper has been thoroughly checked for errors. Many parts have been completely reworded and several paragraphs have modified.

3. Comment

Also, standard conventions should be adhered to in the paper (i.e. 18-km horizontal grid vs 24x24 km² grid vs 150km x 150 km grid); or terms like layers vs levels, or hpa vs mb; choose one style and stick with it. Also consider striking "resolution" from model grid size descriptions and replace with a more appropriate term such as "grid spacing" or other more specific term.

Reply. We have harmonised the notation in the manuscript. Although "resolution" is commonly used in the scientific literature to refer to model grid spacing, we have replaced "resolution" to "grid spacing" throughout the text.

Specific Comments:

1. Abstract:

define acronym MOZAIC.

This was done, although in other papers published by GMD acronyms are not expanded in the abstract.

3rd to last sentence should be re-worded. significant findings (numerical quantities) could be included.

We have reworded the sentence and added numerical quantities to summarise the significant findings.

2. Introduction Page 524:

1st Paragraph should mention that policies are written for ground level concentrations, so that the point of the final sentence is clear.

Done as suggested.

2nd sentence should be reworded/structured.

Done as suggested

Many sentences in the paper use too many casual words and unneeded wordage that makes it difficult to determine the main point. For example, 2nd paragraph , 1st sentence reads: "Precise simulation of tropospheric fields is not only crucial from the perspective of emission control, but also to test the capability of models to capture the vertical distribution of pollutants, the exchanges between the boundary layer and the free troposphere, as well as the horizontal fluxes to and from continental domains." Consider the following revision: "Precise simulation of the troposphere is crucial from the perspective of emissions controls, and requires that we test the capability of models to represent the vertical distribution of pollutants, the exchanges between boundary layer and the free troposphere, and the horizontal fluxes between continental domains."

Done as suggested. As explained before, the text has been extensively reworded to make it easy to read.

Line 19: change "transport" to "chemistry and transport".

Done as suggested.

Reword sentence in Lines 21 - 27 to: "It is also important to assess the capability of models to simulate the meteorology that drives the transport and dispersion of pollutants, as errors in meteorology are inherited by AQ models, thereby producing errors in model-predicted pollutant concentrations and leading to potential compensating errors." If you want to discuss the effect of compensating errors as well I would do so at the end of this sentence, not the beginning.

Done as suggested.

Introduction Pages 525-526: Paragraph which starts as "The evaluation of regional AQ...": I would make this the 1st paragraph of the paper.

Done as suggested.

Throughout the paper... do not call model simulations "runs". They should be called simulations, predictions, or model results.

The term "runs" has been replaced with "simulations" as suggested

Line 20: "... and estimates based on ozonesonde profiles..." estimates of what?

Done. We have clarified the sentence in the text

Paragraph that starts on 525 and continues to 526 would benefit greatly from quantities rather than just qualitative summary.

We provided a quantitative discussion and additional references

Page 526

Line 15: remove first instance of "individual" and "in particular". Line 19 change "connecting" to "between".

Done as suggested.

Lines 20-23: sentence should be reworded.

Done as suggested. The majority of sentences have been completely reworded

Line 24 change "trying therefore" to "attempting". Comment: is AQMEII really just a case study?.

Done as suggested. We no longer refer to AQMEII as case study

Page 527

Line 1: remove "existence, usability, and" Line 5-6: remove "not because they are assumed to be unimportant but"

Done as suggested

Line 7: Eliminate or expound the statement "This is the first study of its kind."

Done as suggested

Final sentence of Introduction: change to "Thus 4-D datasets of ozone are effectively used to assess seasonal and boundary condition errors."

Done as suggested

3. Section 2

First sentence: remove first comma, replace "in the context of" with "by", and turn the "c" in aircraft to uppercase as defined by the MOZAIC acronym.

Done as suggested

Page 528

First (continued from previous page) paragraph and second paragraph should be blended together to deliver the main points.

Done as suggested

Line 16: remove "The strategy adopted within AQMEII for the comparison with the" and replace with "We plotted the".

Done as suggested

Line 17: remove "consisted of plotting" and replace with "of".

Done as suggested

Line 18: Domains were identified? Are these the same as the trajectory domains mentioned further down in the paragraph? What is a trajectory domain?

Done as suggested. We have clarified the meaning of "Mozaic Domain" and simplified the discussion in the revised text.

Lines 22-23: replace "the identification of" with "we identified".

Done as suggested

Line 25: replace "which are" with "to be".

Done as suggested

Line 25: remove "Extraction has been made taking" and add "was extracted" to the end of the

sentence.

Done as suggested

Line 26-28: Change sentence to: "To simplify the data request and extraction for modelers, four areas in North America (NA) and one in Europe (EU) were identified, allowing the trajectory projections of several airports to be grouped together.

Done as suggested

Line 29: remove "All modelers were then asked to deliver" and combine with following sentence.

Done as suggested

Page 529: sentence explaining the ENSEMBLE web-interface should be moved to the paragraph where ENSEMBLE is defined.

Done as suggested. Description of the ENSEMBLE system has been moved to the first paragraph of Section 2.1

Line 13: replace "node" with "cell".

Done as suggested

Line 14: Wording is bad; how does a model value represent a measurement?

Clarification provided in the text

Line 15: Remove "For all the airports included in this study" and reword the sentence (especially the last part).

We have reworded the sentence

Line 19: Move "for each of the five areas (four in NA and one in EU)" to the end of the sentence.

Done as suggested

Line 22: insert "the" between "represented" and "two".

Done as suggested

Line 23: replace "has" with "had".

Done as suggested

4. Section 2.1

WMO should be defined. Is a whole section necessary for 2 sentences?

Done as suggested. The section dedicated to ozonesondes was merged with the main section.

remove "these data too" on page 530.

Done as suggested

5. Page 530, Section 3

Acronyms need to be defined.

Acronyms have been defined where possible

Are bulleted lists really necessary? I would think the models can be listed within the paragraphs.
We have removed the bullet list

Line 4: remove "models and" and insert "and models used" after "groups".

Done as suggested

Remove 2nd sentence.

Done as suggested

Line 13: insert "respectively" at the end of the sentence.

Done as suggested

Line 17: replace "providing" with "that provided".

Done as suggested

Page 531

Line 4: replace "These" with "The EU".

Done as suggested

Line 5: respectively?

Done as suggested

Line 9: replace "provided" with "obtained"... this sentence is unclear, especially the 2nd half. Avoid using "in this light" or "in light of" e.g. change "can be examined in this light" to "can also be examined this way".

Done as suggested

Line 24: remove "even if standard emissions were used".

Done as suggested

Page 532

Line 3: replace "made available" with "provided".

Done as suggested

Line 6: insert "the" between "by" and "European".

Done as suggested

Line 7: move Schere reference to the GEMS parentheses beforehand.

Done as suggested

Remove "see" and "for more details".

Done as suggested

Line 8: remove "but not all" and combine with the following sentence.

Done as suggested

Line 16: move "CMAQ" to sit between "utilizing" and "version".

Done as suggested

Line 16: replace "run" with "used".

Done as suggested

Line 18: remove "of the model".

Done as suggested

Line 20: it isn't necessary to describe what is in North America.

We prefer to leave the countries covered by the domain, it is more informative.

Page 533

Line 10: insert "the" between "from" and "continuity".

Done as suggested

Line 11: remove "vertical".

Done as suggested

Line 12: shouldn't values be interpolated by log and not linear for atmospheric layers?

We have removed the interpolation method which is not available for all models

Line 17: replace "was set up with" with "included"; remove "simultaneously".

Done as suggested

Line 18: dash and then comma?

Done as suggested

Line 22: reference for Forester filter?

Added

Section 3.1.3 switches to present tense... stay consistent

Done as suggested

Page 537

Line 4: change "gird" to "grid".

Done as suggested

Line 9: move "from the horizontal advection" to after "derived".

Done as suggested

Line 16: replace first "in" with "of"; remove period and replace "It" with "and", continuing the sentence. Line 17: replace "to" with "and".

Done as suggested

Section 4.1 Page 538 Line 15: Remove "It can be firstly observed how".

Done as suggested

Line 17: replace "ones" with "profiles".

Done as suggested

Line 18: replace "most likely descends from" with "is likely because of"; pluralize technique.

Done as suggested

Line 19: remove "that" .

Done as suggested

Line 20: replace "that" with "which"; replace period with comma and remove "This is".

Done as suggested

Line 21 replace "coarser resolution of this model" with "coarse resolution of the model"; remove "Differently".

Done as suggested

Line 23: remove "that being the relative humidity derived from the specific humidity and the temperature,".

Done as suggested

Line 24: insert "the" between "where" and "water".

Done as suggested

Line 25: replace "that is" with "e.g."; remove "even".

Done as suggested

Page 539, Line 1: Remove "relevant".

Done as suggested

Page 554: Lines 11 - 15 have poor wording.

We have revised the sentence there

It is not really appropriate to just state a laundry list of (poorly worded) conclusions as bulleted points as the final statement. There should be a clear summary of the work, a list of impacts and take away messages, and expected future work.

As anticipated to the reply to the first comment, the conclusions section has been modified, addressing reviewer's comments.

Tables and Figures:

Table 1 Caption: remove period between "2006" and "for".

Done as suggested

Figure 1 Caption: replace "of" with "in".

Done as suggested

Figure 1 Legend: Should be moved to be beside the bottom panel so it is clear it is a legend for both

panels.

We have added the legend to figure 1b

Figure 1 images: Use a standard method of ticks for lat/long. There are varying formats and fonts used, and the axes are not labeled. Either format correctly or remove.

We revised the figures as suggested

Figure 2 Legend: Should be moved to be beside the bottom panel so it is clear it is a legend for all panels.

We moved the legend

Figures 3 - 6: inconsistent placement of a,b,c,d,e markers. Fix this.

Done

Figure 8: Portland label is formatted differently than others.

We redid Figure 8 with higher resolution. A script controls the labelling and the alignment. No apparent reason why Portland label should be different.

Figure 8: include the legend.

We added the legend

Figures 9-10: Include both legends for both figures; replace "over" with "for".

Done

Figure 11 Legend: Should be moved to be external of the left panel so it is clear it is a legend for both panels; "nodes"? Any reason these can't be one panel?

We moved the legend between the two panels and removed the term "node".

Figures 12, 14,15,16: Showing half a standard deviation bar is very strange and makes it difficult for the reader. Show both sides of the bar!

Showing half of the standard deviation is not uncommon and makes the graphs clearer. In some cases (portland, philadelphia) the values of standard deviation are so large that showing both sides would cluster the profiles in the figure. That would make even more difficult for the reader to appreciate the discrepancies among the different profiles. After all, the standard deviation is symmetric and we show the half most relevant for the discussion.

Figure 13 Caption: replace "lines derive" with "lines are derived".

Done

Figure 13: axes labels should not fall half way in/out of bounding box.

We removed the bounding box

Figures 14,15,16: Include the Legends!

Legends have been added

Reviewer 2

General comments

1. Comment

The manuscript of Solazzo et al. is an evaluation study including AQ model results from a well coordinated modeling exercise (AQMEII) and observations from the MOZAIC dataset. As stated by the authors the primary aim of this study is to illustrate the potential for using MOZAIC data for regional scale evaluation and the capabilities of models to simulate concentration and meteorological fields in the vertical. The scientific approach and applied methods are valid. The capabilities of models are tested and evaluated by using some standard metrics and their performance is discussed, without performing any in-depth analysis, which is out of the scope of this work, as stated by the authors. In my opinion, the extent of this material could be used in a more sophisticated way, in order to highlight the common problems of AQ models and suggest pathways for further analysis and model improvements.

The paper is relevant for publication in GMD since the dataset used is quite extended, however major revisions concerning the analysis and presentation of the material would be necessary prior to publication. The language should definitely be improved in order to be more clear and precise.

Reply. We have revised large portions of the manuscript to incorporate reviewer's suggestions. Analysis has been improved by including new references and further explaining the results. We have put the study more into the framework of the model evaluation components of Dennis et al., (2010), explaining that we present here only the first component of model evaluation and assessment (operational evaluation). We highlighted that the effort expanded here should be used as basis for future analyses.

Specific comments

The title of the paper is too generic; the authors could consider revisiting the proposed title including more specific keywords in order to make it more informative.

The title has been changed

The abstract is not very informative. It would be preferable to include more specific comments on model performance based on the calculated metrics. Preferably avoid expressions like "some success" which are not objective. Highlight the major findings and conclusions, including model performance, impact of boundaries, analysis altitudinal correlations/errors and suggestions for future work.

We have reworded and improved the abstract to take on board the suggestion of the reviewer.

The authors state the "...set the stage for future process oriented studies and in depth diagnostic analysis". Highlight which suggestions for future process oriented studies rise as a result of the current analysis.

We have indicated the analysis that could descend from the results presented in the current study both in the conclusions and in the text, where appropriate. More in general, we have put the paper more into the context of the four pillars of model evaluation as defined by Dennis et al. (2010), the first being operational model evaluation, which is necessary to set the stage for

successive diagnostic model evaluations (see last paragraph of Section 1 and Conclusions)

The paragraph “Wind direction” 4.2.2. is a mere description of bias on every airport. What do we learn from this description? The main conclusions one can draw? A more sophisticated synthesis of results would be preferable.

We shortened the discussion on wind direction, focusing on the main points. We agree with the reviewer that it was a discussion on the bias. For the screening operational evaluation presented in this study, the objective is discerning the discrepancies between an observed and modelled quantity. Moreover, wind direction is rarely investigated. Given the length of the manuscript and the difficulty to synthesise the discussion of wind direction for 5 airports, 15 models and 13 vertical layers, we believe a concise discussion stating that, among other fields, models are also biased for wind direction is a point worth making. Then, of course, the next step should be to quantify the effect such a bias has on the concentration of pollutants, but this could not be achieved in this study.

In paragraph 4.2.4 a reference to a figure is missing.

A reference to Figure 5 has been added

“Ozone skills are clustered by height and by modelling group, rather than simply by AQ model” (4.2.5). It would be better to rephrase the statement to convey the right message.

We have reworded the sentence there as suggested

In general, 4.2.5 needs to be edited by a native speaker.

The paper has been proof-read by two native English speakers

Page 555, line 10 “Results obtained by correlating the bias... for example establishing some indicators that could be compared to observed quantities” The suggestion is rather vague, could it become more specific?

We removed the sentence there

May be references to previous evaluations work using MOZAIC data would be necessary in order to give proper credit to previous related work and clearly indicate which is your own original contribution. (E.g. Elguindi et al, 2010 Geosci. Model Dev., 3, 501-518, 2010).

The reference has been included (along with others)

Technical corrections

Some parts of the paper are not fluent in expression, editing by a native speaker is strongly suggested.

The paper has been proof-read by two native English speakers.

Figures composites and legends are too difficult to follow (Fig 2-6)

We tried to make figures clearer. However, given the size the dataset analysed, we had to compromise between length and “density of information” conveyed in each graph.

Several typos scattered through the text. Restructure the main body of paragraphs to make it more easily readable.

In general, we revised and rewritten many parts of the manuscript. The conclusions section has been completely edited and revised.