To the Editor:

This document contains my detailed responses to the two reviews of manuscript acp-2014-855, followed by the revised manuscript with changes indicated.

Louisa Emmons

emmons@ucar.edu

Responses to Referee #1

This paper provides an overview of the POLARCAT Model Intercomparison Project (POLMIP) and presents the results concerning ozone and its precursors. The paper is accompanied by two additional manuscripts by Monks et al. and Arnold et al. that contain complementary analysis. The paper addresses the very important question of model comparison focusing on ozone and its precursors in the Arctic region. 9 global and 2 regional models with similar emissions are included in the comparison. Consistent differences exist between different models as well as between models and observations (aircraft observations from POLARCAT mission, ozonesondes, OMI NO2 columns, C2H6 and C3H8 surface measurements). As demonstrated nicely with tracers in the accompanying paper by Monks et al., the differences between models in CO in the Arctic are due mostly to differences in chemistry (OH fields are different) rather than differences in transport from source regions. The paper by Emmons et al. contains an analyses of some of the reasons for the different OH concentrations: different cloud coverage leading to different photolysis rates, differences in H2O mixing ratios. Other possible reasons, such as differences in ozone deposition velocities, are identified but not analysed. The comparisons with observations allow identifying some consistent model biases and lead to suggestions for possible improvements, e.g. emission inventories.

General Comments

This is a well-written article presenting an important inter-comparison of models in the context of a measurement campaign (POLARCAT). Important and interesting insights on the reasons of the inter-model differences are provided.

1) What is not entirely clear is that even though it is stated that the same emission inventories are used in all models, some differences are present. This is due to differences in the chemical mechanisms (some species are not explicitly modelled), but it is not clear if these are the only differences (e.g. the differences in GEOS-CHEM emissions).

All groups were asked to use the same emissions, but the implementation was not perfectly identical. Table 2 was included to document what was actually used in each model.

2) Some of the figures (especially related to the comparison with OMI NO2 columns) are somewhat hard to analyse.

We have increased the font size and increased image sizes for the OMI results and comparisons, along with other improvements in other figures.

3) It could be also helpful if the objectives of this paper with respect to the accompanying papers (especially Monks et al) could be somewhat more detailed in section 1.

We prefer to keep the focus of the introduction on the objectives of this paper. The Monks et al and Arnold et al papers are in the same special issue, so readily accessible to readers.

Some additional suggestions for improving the clarity of the text are provided in the specific comments below.

Specific Comments

Page 29335, line 13: change "slow mixing" to "slow vertical mixing"? changed

Page 29335, line 18: add that the rapid advection follows isentropic surfaces, which can explain the layering

added

Page 29336, line 17: to be more clear, change "than transport does" to " than differences due to different transport in the models".

Changed to "than differences in transport in the models"

Page 29336, line 26: rephrase "following these". It would be also nice to give the corresponding section numbers.

This paragraph has been re-written to more clearly identify the paper sections.

Page 29342, lines 28-29: Only one version of MOZART-4 results is shown starting with figure 7. Which of these versions is used?

The version with the lookup table is used. This is clarified in the MOZART-4 description.

Page 29345, line 20: I would suggest to change the title of section 4, for example: Overview of model characteristics and of the main model differences.

Good suggestion. Changed to 'Overview of model characteristics and differences'.

Page 29346: line 21: remove "in" after "agree on" **corrected**

Page 29347, line 19: should be Table 2? **Yes, corrected.**

Page 29347, line 19: It is not clear why different emissions were used in GEOS-Chem. This is not explained in the model description section.

At the time of running these simulations, GEOS-Chem did not have an easy procedure for reading emissions from external netcdf files, so were unable to use the provided emissions files, but attempted to replicate them in their simulation.

Page 29348: line 18-19: the means were compared?

To determine the biases between model and observations the mean of each altitude bin was used. This has been clarified in the text: "Since the models, with roughly 0.5–1 km vertical layer spacing in the free troposphere cannot reproduce all of the observed structure, the ozonesonde data and model profiles were binned to 100 hPa layers. The mean of each bin between the surface and 300 hPa was used to calculate the bias between model and measurements for each profile."

Page 29348: lines 23-25- Please indicate how the too strong transport from the stratosphere is consistent with the values lower than observations above 300 mb.

And also is the strong negative bias present for some models above 300mb real? Or is this due to the binning of levels across the tropopause (and the exact location of model levels in this zone of high vertical gradients)?

The figures have been re-plotted to not show values above 300 hPa, as the strong gradient across the tropopause makes the comparison difficult. Text modified, given in previous response.

Page 29349: section 5.2: why is the difference between GEOS-Chem and other models so large for ethane? According to table 2, the differences in total fluxes are small.

Probably due to the much lower OH in GEOS-Chem.

Page 29349: section 5.3, first paragraph: Please specify if total NO2 columns are considered or rather mostly tropospheric columns.

Yes, just tropospheric columns are used. This has been added to the text.

Page 29349: line 25: please explain further the sentence starting with: The averaging kernels of the . . .

An explanation of the need for using averaging kernels has been added. I hope this is what the reviewer meant.

Page 29349: section 5.3, first paragraph: it would be very helpful to have some more information on the relative values of the averaging kernels at different levels. How much weight is on average given to the boundary layer as compared to the free troposphere and/or upper troposphere?

An explanation of the effect of the averaging kernel as applied in the DOMINO product is given in Eskes and Boersma (2003) who show that the range of sensitivities is large between 200 and 1000 hPa, with the lowest sensitivity towards the surface and largest sensitivity in the stratosphere. For instance, depending on the surface albedo the sensitivity may increase by roughly a factor 3 when moving from the boundary layer to the upper troposphere.

We have changed the text as follows:

"The transformation of the model profiles with the averaging kernels gives model levels in the free troposphere relatively greater weight in the column calculation. For instance, depending on the surface albedo the sensitivity to the upper free troposphere compared to the surface layer may increase by roughly a factor 3 (Eskes and Boersma, 2003). This means that errors in the shape of the NO₂ profile can contribute to biases in the total column."

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, Atmos. Chem. Phys., 3, 1285-1291, doi:10.5194/acp-3-1285-2003, 2003.

Page 29350: line 6: replace "model bias" by "mean model bias" **changed**

Page 29350: lines 17-20. Split the sentence by ending the first sentence after respectively. Please explain how it was decided whether 90% of NOx emissions originate from anthropogenic or biomass burning emissions. This is based on model

inventories used? Also should a threshold be used to filter the data in the plot? The ship emissions do not bring much information to these plots.

Sentence split. Added to the text that the 90% limit is based on the emissions inventory. Fig. 13 has been remade (see new figure below).

Page 29350, lines 23-25 – why not restrict the zone over NW Europe so that it does not include the North Sea? The strong emissions are only over land.

The European region has been moved to over land (see new Fig. 13 below).

Page 29350, lines 25-29: a conclusion seems to be missing. Is my understanding correct: model NO2 columns are underestimated over pollution hotspots, but the spread in model results does not allow to draw conclusions on the quality of the NOx emissions used. Also it would be helpful to understand what weight is given to the boundary layer NOx (see also comment above).

For most regions the model variance is low enough to draw conclusions – the anthropogenic emissions are too low and the fire emissions are too high. Sentence added:

"Since all models used the same NO emissions, the large variation between models (as seen in Figure 12) indicates differences in the rapid photochemical cycling between NO and NO₂ in the boundary layer."

Page 29351: lines 9-10: Split the sentence: Figure 15 shows . . . of the campaign. The flight tracks have . . .

Don't see that is needed.

Page 29351: line 25: add that data were binned in 1 km bins. **Added.**

Page 29351: line 26: how was the measurement uncertainty calculated for the binned median values?

Sentence changed to: "The thick error bars represent the measurement uncertainty (determined by applying the fractional uncertainty reported in each measurement data file to the median binned value), while the thinner horizontal lines show the variation (25th to 75th percentile) in the observations over the flights."

Page 29352: line 6-7: the sentence starting with 'In the cases...' should probably be moved to the end of the previous paragraph. In Fig 19, only one measurement is considered. Also it might be helpful to add, that "more than one measurement" indicates measurements with different instrument/technique.

No, Fig. 19 shows the average of all measurement techniques for each compound. The observations from each instrument are shown separately in Figs. 16 and 18. Sentence re-written: "In the cases where a compound was measured by more than one instrument, the differences between the model and each observation were averaged over all the measurement techniques."

Page 29352, lines 10-11: I would not put the sentence 'consistent with' between parentheses: this is an important statement.

Parentheses removed.

Page 29352, lines 16 : after 'boundary layer' could add '(ARCTAS-B, fig. 18). Added.

Page 29352, line 20 : missing 'some' before cases. **Added.**

Page 29353: lines 16-18: the fine structure in OH is due mostly to what sub-grid processes? Cloud distribution and resulting photolysis? Fine scale plumes? We believe clouds and photolysis are most frequently the cause of OH structure. Sentence re-written: "The distribution of OH is strongly affected by clouds and their impact on photolysis, which coarse-grid models cannot be expected to reproduce, however, these differences are likely averaged out in the binned vertical profiles."

Page 29354: line 17: the exact location and strength of model sources could be also added to the list of reasons why the modelled and observed plumes do not overlap in space and time.

Added.

Page 29354: lines 19-22: were all pixels in the selected zones affected by fires? If not, why not take only pixels inside the plumes from fires? Will taking pixels not in the plumes affect the calculated slopes?

No, not all grid points on each day are affected by the fires. We neglected to say the calculated slopes were determined for points with CO > 150ppb, and thus the non-fire points are ignored. The text has been updated to add this point.

Page 29355: line 9: "fire emissions" or should it be "fire emission factors"? Not having compared CO in these airmasses (observations vs plumes), it seems that one cannot extend the conclusions to emissions but should talk only about emission factors.

Quite true. 'emissions' has been changed to 'emission factors'.

Page 29355: line 17: as in the previous comment. Should it be emission factors rather than emissions?

Yes, corrected.

Page 29356: line 26: replace "ozonesondes" by "ozone" **Changed.**

Figure 8: The station Narragansett seems not to be used in later figures (9 and 10) Narragansett had much less frequent sampling (4 sondes in April, 3 in June-July), so was not used. An explanation has been added to the text

Figure 9-10: Comments on these figures: showing both: the individual measured profiles and mean biases on the same figures does not really facilitate the analysis of these figures. Can they be split to show the mean profiles (observations + models) and biases separately? This would also allow zooming in on the biases. It would also be nice to have in these figures some information on the standard deviations. It is understood that it would be hard to include this information for these plots for all

models at the same time. But it could be presented at least for the observed data instead of showing the individual ozonesonde profiles. Showing individual observed profiles probably does not bring much useful information compared to mean + standard deviation. Showing mean + standard deviation would however simplify the figures.

The ozonesonde plots have been remade, replacing the individual sonde profiles with the mean and standard deviation for each site. (see below)

Figure 13: indicate in the legend that left column for April and right for June-July. **Added to caption.**

Figure 14: indicate in the legend that the box plots show model results! They include data from all models?

The caption stated "Summary of model mean...", but has been re-written. Observations are also shown.

Figure 15: It seems that Grace and POLARCAT France measurements are not used in this paper. Is there a reason for this? This could be mentioned in the legend. Please add also AP next to ARCPAC.

The GRACE and France experiments had limited gas-phase measurements available so added little to the conclusions already presented. So in the interest of clarity and space they were not included. AP is defined in the Figure 19 caption.

Figures 16-18: not easy to distinguish between thick and thin error bars? Use error bars with vertical lines? Also replace ARCTAS-A by ARCTAS A1? WRF can be removed from the legend in figure 16?

Figures 16-18 have been updated.

Figure 17: nomenclature: "P3" not used in figure 15, for clarity might be helpful to remove it.

OK.

Figure 19: why OH was not included? Limited space.

Table 2: Change title from Emissions to Global emissions. Indicate in the legend that the regional models were not listed, as the global values cannot be provided. **Changed.**

Referee #2

We thank the reviewer for his/her generally favorable view of the paper along with very detailed comments. After removing the duplicated sections of the review due to an apparent error in copy and pasting in the submission of the review, we have reproduced the review below, and included our responses (in bold).

GENERAL REMARKS

Analyzing the content and the context of the manuscript, I would say the study described in the manuscript has fulfilled 3 objectives: 1) Introduction of the models which are used in POLMIP; 2) Showing a comparison of the models with observations in the Arctic, and make conclusions about how good the models reproduce the observations; and 3) Conclusions about the emission data set.

I think the manuscript does a valuable job in comparing the results of model simulations with observations in the Arctic. It is important to estimate the reliability of CTMs in this remote but vulnerable region, which is impacted by pollution from different origins. Therefore, it is very nice that the study uses for this the extensive data set obtained during the International Polar year 2008. The manuscript is well written and agreeable to read.

However, the paper lacks analysis to attribute the differences between model results and observations to specific model components. It does it for some model components (complexity of the chemistry scheme to some degree, photolysis rates and cloud fields), but other possible contributing factors are only mentioned but have not been investigated (impact of dry/wet deposition, tracer transport scheme, convection and boundary layer parameterisations, vertical resolution in lowest 5 km, reduced grid near the poles, inclusion of stratospheric chemistry, ...). In the abstract is mentioned : "to quantify the differences in model chemistry and transport schemes." I read this as a focus of POLMIP (broader than this manuscript), and not only of this study. However, it gives the suggestion that it might all be investigated in this study. The abstract continues : "Differences in a number of model parameters are identified as contributing to differences in the modeled chemical species, including cloud fields and photolysis rates." But this makes arise the question : which others have been identified? Also, the analysis of cloud fields and photolysis rates feels a bit limited to an illustration. Further, in the conclusions is written : "However, numerous differences occurred among the model outputs due to the different chemical schemes and physical parameterizations such as convection, boundary layer mixing and ventilation, wet and dry deposition." However, at the end I have the impression that the contribution of these different parameterisations to the actual model differences is not quantified. Finally, in the first paragraph of the conclusions is already written : "Additional model diagnostics are required to completely understand the differences among models." I think such suggestions should be made at the end of the conclusions.

I write this because, at the moment, as a reader one gets the impression that certain things will be investigated or analyzed, but in the end they are not. It would therefore be nice if the analysis could be improved and extended. What type of model components play a large role in the modeling of the chemical composition of the Arctic atmosphere? If this extension is not possible, then it should be explained why only the clouds and photolysis rates are explicitly illustrated. In that case also, make the text more clean and sober: focus only on the aspects you really investigate in the study. Just mention the ones you don't investigate (and possibly estimate their relevance based on other studies for the Arctic), but state from the beginning that they are not the focus of the study. Therefore the abstract, Section 4, and the Conclusions (Section 7) should be thoroughfully rewritten.

Attempts have been made to revise the abstract and conclusions so as to clarify the limited set of comparisons and evaluation performed in this paper. Discussion of additional diagnostics that would have been useful, etc., have been moved to the end of the Conclusions.

There are valuable conclusions about the emissions data set. However, these conclusions only have a value if this is a publicly available dataset, and which is intended to be used by others, or which at least can be investigated. Although the webpage http://bio.cgrer.uiowa.edu/arctas/emission.html was accessible, further links which would guide to the data did not work. The links on that web-page which did not work were : (i) "This inventory is available for download", and (ii) "Gridded emissions can be accessed at the University of Iowa ACESS web- site http://www.cgrer.uiowa.edu/EMISSION_DATA/index_16.htm, under the direction of Gregory Carmichael, Principal Investigator of this project."

We believe that this emission inventory is intended for use by any who are interested. It is freely available (without password control) from a website. It was developed for the ARCTAS and POLARCAT modeling groups, and anyone else interested. I don't believe the University of Iowa intends to prevent access to the inventory by having broken links and we have contacted them to ask that the links be repaired.

The direct link to the data is:

http://bio.cgrer.uiowa.edu/arctas/arctas/07222009/.

Further, I think the manuscript contains too many figures.

We believe the figures presented are all useful. The other reviewer suggested a number of improvements to the figures (which we have made), implying they saw value in all of the figures.

The structure of the paper could be improved. Now the Sections 4, 5, and 6 all contain results, so they might be brought together (as Sections 4 and 6 are short). Section 3 should have a more general title.

Sections 4, 5 and 6 were separated to highlight the rather different results being presented. Section 3 is renamed: Model configurations and inputs. We believe this organization of the document helps the reader.

Below you can find : (i) a list of smaller general remarks, (ii) a list of detailed comments on the manuscript text, (iii) a list of detailed comments on the tables, and (iv) a list of detailed comments on the figures

Each of these comments is addressed below and in the revised manuscript.

SMALLER GENERAL REMARKS

Figures There are too many figures. In some figures, too small fonts are used.
 Some plots are too small. The layout of some of the figures should be improved.
 We find all of the figures valuable (as addressed above) and have improved the layout and font sizes.

2. For MOZART, 2 things are unclear : (i) which photolysis rate calculation is actually used for the principal simulations? (ii) how are the clouds and the humidity calculated? Is CAM (nudged to GEOS-5) also used?

(i) The lookup table for photolysis is used for most simulations (this is clarified as explained below). (ii) MOZART-4 CTM contains a cloud parameterization, as described in the MOZART description papers. It is completely unrelated to the CAM simulations. CAM4-chem and CAM5-chem simulations for POLMIP are nudged with GEOS-5 meteorology.

3. Maybe mention a bit more explicitly which papers are companioning papers within POLMIP. I presume that Monks et al. [2014] and Arnold et al. [2014] are, but this is not explicitly said. Are there others?

These are the only 3 POLMIP papers to date and it is unlikely there will be others.

4. I would consequently use "wildfire" instead of "fire".

The FINN fire emissions include all open burning - both wildfires and prescribed agricultural burns that are detected by satellite, thus, we prefer to use the more general term.

5. Abbreviations Be consequent in how abbreviations are defined : first the long expression, followed by the abbreviation between brackets. It is probably not possible to always follow this rule, but try to be as consequent as possible. What strategy is followed for the first letter in the full expression : CTM (small letter in full expression), GCM (big letters in full expression), SLCP (small letters in full expression)?

We have tried to modify the text to follow this rule.

6. Homogenize the description of the different models (see DETAILED COMMENTS on TEXT)

Answered below.

7. LMDZ-INCA, LMDZ both are used in the text. Try to just choose one. **Corrected to LMDz-INCA**.

DETAILED COMMENTS on TEXT

Please find below a list of detailed comments. Among these, there are quite some comments and suggestions about language use. Feel free not to follow these suggestions on language, but please give a good reason why you do not do so. We have accepted the suggestions where they improve the precision of the grammar and readability, however, in some cases we kept our original language as we felt it conveyed our meaning more clearly.

p 29332 : Norrkoping -> Norrk\"oping (\" on the o) Corrected

p 29332 : University Innsbruck -> University of Innsbruck **Corrected**

p 29333, l 2-3 : "atmospheric chemistry observations" : this is a bit vague. **Change to "observations of atmospheric composition"**

p 29333, l 5-11 : 5 times "differences" on only a few lines **rewritten**

p 29333, l 15-23 : the order of these last three sentences is a bit strange. First, a sentence about usefulness of aircraft observations without conclusions. Second, a sentence about satellite observations with conclusions. Third, a sentence about aircraft observations (although) with conclusions. Maybe sentence one and three can be put together.

rewritten

new Abstract:

A model intercomparison activity was inspired by the large suite of observations of atmospheric composition made during the International Polar Year (2008) in the Arctic. Nine global and two regional chemical transport models participated in this intercomparison and have performed simulations for 2008 using a common emissions inventory to assess the differences in model chemistry and transport schemes. This paper summarizes the models and compares their simulations of ozone and its precursors, and presents an evaluation of the simulations using a variety of surface, balloon, aircraft and satellite observations. Each type of measurement has some limitations in spatial or temporal coverage or in composition, but together they assist in quantifying the limitations of the models in the Arctic and surrounding regions. Despite using the same emissions, large differences are seen among the models. The cloud fields and photolysis rates are shown to vary greatly among the models, indicating one source of the differences in the simulated chemical species. The largest differences among models, and between models and observations, are in NO_V partitioning (PAN vs. HNO₃) and in oxygenated volatile organic compounds (VOCs) such as acetaldehyde and acetone. Comparisons to surface site measurements of ethane and propane indicate that the emissions of these species are significantly underestimated. Satellite observations of NO₂ from OMI have been used to evaluate the models over source regions, indicating anthropogenic emissions are underestimated in East Asia, but fire emissions are generally overestimated. The emission factors for wildfires in Canada are evaluated using the correlations of VOCs to CO in the model output in comparison to enhancement factors derived from aircraft observations, showing reasonable agreement for methanol and acetaldehyde, but

underestimate ethanol, propane and acetone, while overestimating ethane emission factors.

p 29334, l 3-5 : Two aspects are mentioned ((i) heat transport from lower latitudes, (ii) local radiative forcing). The second one is really about climate change, while the first one is even true without climate change. Try to formulate more precisely. Sentence changed to: "Arctic temperatures are affected by increased heat transport from lower latitudes and by local in-situ response to radiative forcing due to changes in greenhouse gases and aerosols."

p 29334, l 8 : aerosol -> aerosols (as on line 5) Changed

p 29334, l 27-28 : twice "significant" in the same sentence **Changed**

p 29335, l 4 : "it" refers to "Arctic" I presume. It would be clearer to explicitly write it.

Corrected

p 29335, l 12-16 : maybe add "only" before "in close proximity", and "mainly" before "retaining only". Otherwise the last 3 parts of the sentence do not fit together well. **Changed**

p 29335, l 21 : "slow" before "mixing"? Changed

p29335, l22: and more->or more **Changed**

p 29335, l 25 : "Climate" -> "of Climate" in the definition of POLARCAT **Corrected**

p 29336, l 5 : "focused" -> "focuses" **Corrected**

p 29336, l 5 : "this comparison" : does this refer to the whole POLMIP, or just to this paper? Is POLMIP limited to gas phase chemistry evaluation?

The comparisons performed within POLMIP are primarily of gas phase compounds. Text clarified.

p 29336, l 9-10 : "evaluate ... with ... observations" : is this correct language use? **Kept original language.**

p 29336, l 13-14 : "by methyl chloroform observations and emissions" : maybe "by methyl chloroform observations and its emission estimates" **Changed**

p 29336, l 17 : "than transport does" -> "than differences in transport"

Changed to "than differences in transport in the models"

p 29336, l 23 : "efficiency in" -> "efficiency of"

Changed

p 29336, l 24 - p 29337, l 2 : I would suggest to add section numbers in this paragraph

p 29336, l 26 : "of all of the models to observations" -> "of model results with observations"

Changed

p 29336, l 26-28 : This sentence is a bit strange as an ozonesondes is more an instrument, while "NMHC" and "compounds" are species. I would suggest to homogenize the sentence.

Changed

p 29336, l 28 : "emissions" -> "the emissions" not changed

p 29337, l 4 : "collaboration of experiments" : is this the correct description? "Consortium" might be better.

p 29337, l 12-13 : "of each mission" -> "for each mission"

Changed

p 29337, l 13-14 : "transported to the Arctic" : shouldn't it be after "wildfire plumes" (if the measured wildfire plumes are not the ones going to the Arctic, than the sentence is correct I think)

Corrected

p 29337, l 17-18 : abbreviation precedes the full name, while different for ARCTAS **Changed – ARCPAC placed in parentheses.**

p 29337, l 20 : is the word "sources" needed? **'sources' removed.**

p 29337, l 21 : fire plumes -> wild fire plumes **Changed**

p 29337, l 24 : "Spring" -> "spring", maybe cancel "in spring" as the dates make this clear (between 30 March and 11 April).

Removed 'in spring'.

p 29338, l 1 : shouldn't GRACE be explained : Greenland Aerosol and Chemistry Experiment?

Yes, thank you.

p 29338, l 2 : "Greenland" maybe not needed as already mentioned on page 29337, l 25. On the other hand, it is maybe good to repeat it.

Not changed

p 29338, l 5 : here again the explanation follows the abbreviation. Maybe it is unavoidable due to the YAK-part.

Not changed

p 29338, l 10 : the section "3 Models" seems to be more general than just about models. Another title might be more appropriate.

True. Section title changed to 'Model configurations and inputs'

p 29338, l 12 : twice "output" -> maybe change the second one into "monthly mean species distributions and diagnostics"

Good suggestion.

p 29338, l 14-15 : there is some tension between "All the models" and "with a few exceptions". Maybe change in "Most of the models"

Changed to: "A single emissions inventory was specified for use by all of the models"

p 29338, l 16 : "global" - except WRF?

Re-worded: "Each global model was run at its standard resolution ..."

p 29338, l 16 : "meteorology" is vague; maybe "meteorological forcing". In addition to differences in "chemistry scheme", "meteorology", and "deposition schemes", the models probably also differ in vertical distribution of BB emission distribution, tracer transport schemes, detrainment/entrainment rates in convection, etc. **True. 'meteorology and deposition schemes' changed to 'meteorological forcing and other parameterizations'.**

p 29338, l 17-18 : twice "output", and "a number" should be avoided. Maybe : "... included monthly mean distributions of mixing ratios and some other diagnostics" **Changed to 'monthly mean species distributions and diagnostics'**

p 29338, l 18 : "evaluation" -> "an evaluation" **not changed**

p 39338, l 20 : "a smaller number" -> "a limited number" **not changed**

```
p 29339, l 2 : www.ceip.at -> http://www.ceip.at as on p29335, l 28 corrected
```

p 29339, l 3-5 : maybe the sentence can be rewritten to have "speciation" only once **First 'speciation' changed to: 'emissions for specific hydrocarbons'**

p 29339, l 4 : is in this paper VOCs used to describe the same set as NMHC (defined on p 29336, l 27)?

No, NMHC generally refers to alkanes, alkenes, etc. (compounds with only carbon and hydrogen atoms) while VOCs includes oxygenated hydrocarbons (NMHC plus acetone, formaldehyde, acetaldehyde, etc., etc.). Thus, it seems useful to make the distinction.

p 29339, l 6-8 : "provided daily" : maybe also mention that they are given to/used in the models on daily basis too.

Changed to start sentence with 'Daily biomass burning ...'

p 29339, l 7 : INventory to agree with FINN (maybe put N as a capital letter) **Must have been changed by the typesetter.** Corrected.

p 29339, l 10 : "these" -> "the ARCTAS", because "these" is confusing as different emission datasets are mentioned just above **Changed.**

p 29339, l 11 : "showed" -> "showed that" not changed.

p 29339, l 17 : "usual" -> "standard" not changed.

p 29339, l 19 : "intercomparison" -> "POLMIP intercomparison" to make clear that it is not just about this paper **Changed.**

p 29339, l 19 : "dynamics" -> "tracer transport" **Changed.**

p 29339, l 26 : "far removed" -> far away l like my words.

p 29339, l 28 : "This offset in location produces differences in atmospheric composition" : I think I understand what is meant, but it should be expressed more clearly

Re-written: Since anthropogenic and biomass burning emissions have different relative amounts of CO, NOx and VOCs, the offset in location of the two source types leads to significant differences in atmospheric composition within these regions. These differences have particular relevance in the analyses of Monks et al. and Arnold et al. that use these tracers.

p 29340, l 4 : "March through August" -> looking at Fig. 1, I would rather say "March through July"

Yes.

p 29340, l 10-11 : meteorology -> "origin of meteorological data" Changed to: 'origin of meteorological analyses'

p29340,l12: Table1->Table2

Removed that sentence (about lightning emissions, as it was mentioned above).

p 29340, l 27 : "integrated forecasting system" -> "Integrated Forecasting System" **Corrected**

p 29340, l 27-28 : I would add "(ECMWF)" after "European Centre for Medium Range Weather Forecasting" (it is later used, e.g., on p 29342, l 20) **Added**

p 29341, l 4-5 : is it necessary to mention "and applies the emission and dry deposition fluxes as part of the vertical diffusion scheme"? **No. removed.**

p 29341, l 5-6 : is this the best way to describe this?

It is technically correct.

p 29341, l 2 : CTM is defined here while "chemical transport model" is already used earlier (e.g. p 29334, l 27). The definition should be given there. **Changed.**

p 29341, l 7 : "model convective precipitation" -> "convective precipitation" Not changed as it is important to emphasize the lightning parameterization is not based on observations and is consistent with the model physics.

p 29341, l 7 : "the C-shaped profile" -> "a C-shaped profile" Not changed. A specific profile is given in Pickering et al.

p 29341, l 10 : 3-D is defined later (page 29343, l 26) It seems obvious that these are all 3D models, so removed in both places.

```
p 29341, l 19 : "includes an" -> "includes " ok.
```

p 29341, l 24 : GMI : abbreviation first **not changed.**

p 29341, l 25 : I would replace "chemical transport model" by "CTM" **ok.**

p 29341, l 27-28 : is it necessary to mention "with all the emissions from the specified inventory". I would rather only mention the exceptions to this rule. **Removed.**

p 29342, l 1 : "several" : are there more than these 2? If so add "e.g.," at the beginning, or ", ..." at the end.

Added e.g.

p 29342, l 5 : "but" -> "and" **not changed.**

p 29342, l 9 : LMDz should be after the full expression **changed**.

p 29342, l 9 : the definition of GCM, should not use capital letters to be in agreement with the definition of CTM : so general circulation model **changed.**

p 29342, l 11 : ORCHIDEE should be after the full expression **kept the original.**

p 29342, l 16 : "gas phase" versus "gas-phase" (both used in the text) **2 words replaced with hyphenated.**

p 29342, l 18-19 : I would write "sulfate" instead of "sulfates" **changed.**

p 29342, l 20 : "6h" too cryptic -> "6-hourly" changed.

p 29342, l 23 : "global annual total" -> "global annual total emission"

changed to 'annual lightning emissions total'

p 29342, l 24 : abbreviation before long expression **not changed.**

p 29342, l 25 : use the abbreviation CTM (because it is introduced earlier on p 29341, l2)

changed.

p 29342, l 26-28 : should it be mentioned that the specified emissions have been used?

Replaced 'For these simulations all of the emissions were from the specified inventory (i.e., online MEGAN was not used for biogenic emissions).' with 'While MOZART-4 includes the capability to calculate biogenic isoprene and terpenes using the MEGAN algorithms, the specified monthly mean emissions were used for POLMIP.'

p 29342, l 28-29 : which one is actually used for the analysis here? In Fig. 4, they are presented both: but what with the rest of the analysis?

Generally the LUT run was used. Added at this point: 'Unless otherwise stated, the results shown here are from the LUT simulation.'

p 29343, l 3-5 : this sentence is identical to the one for GMI, except for the list of oxygenated hydrocarbons. Maybe try a slightly different formulation. **Sentence re-written.**

p 29343, l 5 : "is the same"; but the CAM-chem explication mentions stratospheric chemistry"? Are the aerosols identical?

Re-written to clarify that the tropospheric gas-phase chemistry is the same in MOZART-4 and CAM-chem (there are other differences).

p 29343, l 9 : abbreviation before definition **not changed.**

p 29343, l 16 : "It includes ..." -> "In total, the TM5 chemical mechanism ..." changed

p 29343, l 18-19 : is it worth mentioning this specificity, as other models might also do it (but just don't mention it in their description)?

True. Sentence removed.

p 29343, l 22 : "as" -> "than" changed

p 29343, l 23 : "by a fourth-order polynomial function" -> "as a function of" **not changed**

p 29343, l 20-25 : is the lightning parameterization of TM5 the same as the one of C-IFS? But apparently different aspects are stressed.

The lightning parameterization in C-IFS as applied for the POLMIP evaluations was in fact very similar to what was applied in TM5, except for a different scaling factor between convective precipitation and lightning flashes. Slight changes have been made to the C-IFS description: "Lightning emissions in C-IFS are based on the model convective precipitation (Meijer et al., 2001) and use the C-shaped profile suggested by Pickering et al. (1998), and follows the same implementation as TM5, except that the lightning emissions are scaled to give a global annual total of 4.9

Tg N yr^{−1}."

and the TM5 description:

"NO_X production from lightning is calculated using a linear relationship between lightning flashes and convective precipitation (Meijer et al., 2001), using a C-shaped profile suggested by Pickering et al. (1998). ..."

p 29343, l 25 : has TOMCAT a name definition? 3-D should be defined earlier. **No definition for TOMCAT. 3-D removed (see above)**

p 29344, l 1 : "Extended Tropospheric chemistry" : shouldn't chemistry start with a capital letter?

changed

p 29344, l 3 : aerosol -> aerosols not changed

p 29344, l 5 : the ";" between the references -> "and" changed

p 29344, l 9 : full expression for MATCH?

"Multiple-scale Atmospheric Transport and Chemistry Modeling System" added.

p 29344, l 9 : chemistry transport model -> CTM changed

p 29344, l 16: "in (Andersson et al. 2007)" -> "in Andersson et al. (2007)" changed

p 29344, l 16 : "evaluation" -> an "evaluation" **not changed**

p 29344, l 17 : this should not be "we" as it does not refer to all authors changed to '... data ... were used ...'

p 29344, l 20 : I would add "only" after "but" changed

p 29344, l 20 : "reaching about 16 km" -> "reaching about 16 km high" changed to 'reaching to about 16 km'

 $p\ 29344, l\ 21: I\ would\ skip$ "In addition to the standard daily POLMIP emissions" not changed

p 29344, l 23 : "annual global total" -> "annual global total emission" **not changed as it is clear the sentence is about emissions.**

p 29344, l 23-25 : why specifically mentioning the DMS emissions? **Because they are different from the specified inventory.**

p 29344, l 25 : Maybe something can be mentioned about the boundary conditions for tracer concentrations at the 20N boundary?

Added: "Monthly average results for 2007 and 2008 from global model runs using MOZART at ECMWF in the MACC (Monitoring Atmospheric Chemical Composition) project were used as both upper and 20N chemical boundary conditions."

p 29344, l 26 : has WRF a full name?

Full name added: The Weather Research and Forecasting model with Chemistry

p 29345, l 5 : fire -> wildfire

not changed – FINN includes all open burning, including prescribed agricultural fires

p29345,l5: 1->1x1

not changed. Seems clear to me.

p 29345, l 13 : GOCART (definition after abbreviation) **not changed**

p 29345, l 17 : N. America -> North America not changed

p 29345, l 24 : "meteorology fields -> "meteorological fields" **not changed**

p 29345, l 25-27 : shouldn't it be LMDz instead of LMDZ? changed

p 29346, l 3-4 : How can you calculate water vapour only based on surface water fluxes? Is there a cloud parameterization in MOZART? Is there CAM behind? **MOZART does have a cloud parameterization. It is based on the MATCH**

CTM developed by P.Rasch and M.Lawrence at NCAR many years ago (not SMHI-MATCH used in this paper).

p 29346, l 15 : "that are" -> "which are" ; "source of" -> "source for" not changed

p 29346, l 21 : "agree on in the location" -> "agree on the location" or "agree in the location"

changed

29346, l 26 - p 29347, l 2 : by mentioning explicitly dry deposition, one can give the impression that dry deposition is determining for the tropospheric ozone differences. Especially the April differences in the NH will only by slightly impacted by the deposition scheme, I presume. Isn't the influx from the stratosphere determining: is there a difference in performance among models which prescribe 03 at the top of the model, and those using explicit stratospheric chemistry? The discussion of differences in ozone distributions has been expanded to mention strat-trop exchange and chemistry, as well as identifying deposition as important to the lower troposphere. There does not seem to

be a consistent difference between models with simulated or specified stratosphere.

p 29346, l 28 : "surface layer" : or is it meant "boundary layer"? The surface layer is often just a fraction (1 tenth) of the boundary layer.

'surface layer' changed to 'boundary layer'

p 29347, l 4 : magnitude -> should be little bit more specific like "concentration/mixing ratio/value" added 'concentration'.

p 29347, l 5-6 : I would think that the other models possibly also show a maximum in the tropical mid- to upper-troposphere. But when that maximum is lower than 2x10-6 it will not be visible in this type of plots. Maybe change : "maximum" -> "a maximum higher than 2x10e-6".

Added 'greater than 2x10-6'.

p 29347, l 7 : "have" -> "reach OH concentrations of" changed

p 29347, l 11 : "a number of compounds" -> I would suggest to be more specific changed to ' the time series of ozone and its precursors'

p 29347, l 14 : CO was already used on page 29334, so the definition should come earlier

definition added in Introduction

p 29347, l 17 : C2H6 already mentioned earlier, so should be defined earlier it was mentioned only in the model descriptions; not changed

p 29347, l 17-19 : can differences in transport contribute to these differences? **Possibly, but the long lifetime of ethane makes that less important, l think.**

p29347,l19: Table1->Table2? corrected

p 29347, l 21-22 : inverse order of H2O2 and hydrogen peroxide **I don't understand why this is so important.**

p 29348, l 2-3 : strange sentence, improve it. Strange? I see nothing wrong with: The differences among models are further explained below with regard to comparisons to observations.

p 29348, l 15 : "The hourly model output" -> "Hourly model output" **not changed**

p 29248, l 21 : can one learn something from showing additionally RMSE-profiles? I think the direct bias (preserving the sign) is more informative in evaluating the model performance. p 29248, l 24-25 : what type of upper boundary conditions have been used in SMATCH? Maybe also refer to Figs. 5 and 7 where the different behavior of SMATCH was noticeable.

The source of the upper boundary conditions was added to the description of SMHI-MATCH.

p 29348, l 25 - p 29349, l 1 : GEOS-chem : can this be linked to the different emissions? Is there a reason for this?

It is likely that the low ozone in GEOS-Chem is related to the high uptake of HO2 on aerosols.

p 29349, l 10 : provide -> provides **corrected**

p 29349, l 12 : "over a range of latitudes" : "over a range of mid- to high latitude stations"

changed to 'over a range of northern mid- to high latitudes'

p 29349, l 15 : "altitude" : has for some of the stations a model level different from the lowest level been chosen?

Yes, to account for the surface elevation of the coarse model grids being lower than the sampling site.

p 29350, l 1-3 : While the former sentence indicates an advantage of using averaging kernels, this sentence mentions a possible disadvantage. Therefore I would replace "also" by something like "on the other hand", or start with "however".

This paragraph has been re-worded based on comments by Referee #1 and this sentence has been modified.

p 29350, l 4-5 : I think this introductory sentence should be improved. **I don't see it as deficient.**

p 29350, l 10 : "fires" -> "wild fires" not changed, as explained previously

p 29350, l 14-16 : can this be more specific? **The specifics are listed in the previous sentences.**

p 29350, l 17 : "median", while the figure 13 caption says "mean" The bias plots have been removed from Fig.13, so this is irrelevant now.

p 29350, l 19 : I think this "while"-construction is not so lucky; "regions" : I'd rather use the word "pixel" again, or "grid point" 'not so lucky'? 'while only regions ...' changed to 'and only pixels'

p 29350, l 21-22 : "Figure 13a and c" -> "Figures 13a and c" corrected

p 29350, l 23 : median **correct, not mean**

p 29350, l 24-26 : "Northwest" , while "north east" on line 9 **corrected (one word, lowercase)**

p 29350, l 29 : "East-China", while "East China" on line 13 (but here it is more an adjective) **corrected**

p 29350, 1 29 : "indicating a large uncertainty introduced by the models" : this looks like models introduce uncertainty. I'd rather say that our knowledge is uncertain. That is what we meant – the models all have the same NO emissions but end up with different NO2 column amounts. Added sentence to clarify: 'Since all models used the same NO emissions, the large variation between models (as seen in Figure 14) indicates differences in the chemistry and transport processes affecting NO and NO2.'

p 29351, l 5 : "for forests" -> "for forest fires" not changed

p 29351, l 6 : "the hourly output" -> "hourly output" not changed

p 29351, l 11 : I would put (A1, A2) immediately behind ARCTAS-A seems clear enough as is

p 29351, l 19 : GRACE -> to "-GRACE" not changed

p29351,l20: fire->wildfire **not changed**

p 29351, l 22 : "with" -> "and" not changed

p 29351, l 24-25 : the same -> in the same way **changed**

p 29351, l 25 : "were" -> "was" corrected

p 29351, l 26 : "measurement uncertainty" : is this the measurement uncertainty on one single observation? Or is it a reduced uncertainty as the observations shown are already the mean/median over a large number of individual observations? It is the measurement uncertainty of a single measurement.

p 29351, l 22 - p29352, l 2 : In this short paragraph of text, three large figures are introduced. In the following sentence "To make a more quantitative ...", one temporarily gets the impression that these profiles as such will not be discussed anymore further. However, in the next 2 pages, one refers several times to the profiles. To avoid this initial misconception, I would start the sentence "To make a more ..." by something like "in addition".

Yes, that is better.

p 29352, l 12 : "indicating" -> "indicating that" changed

p29352,l14: 100%->90% corrected

p 29352, l 17 : (Fig. 14) -> (Fig 18.) No, we meant the OMI NO2 comparison in Fig. 14.

p 29352, l 18 : hugely : I would use another word **vastly?**

p29353,l2: "oftoofineascale"->"ofatoofinescale"? of a scale too fine?

p 29353, l 7-8 : I would replace the first "with" by "and" changed

p 29353, l 8 : ethanol is not shown in figure 19? **No, but it is in Figs 16 and 18.**

p 29353, l 9 : a poor job of : can you express this differently? **Maybe...**

p 29353, l 10 : much closer is true for CARB, but not for B **but B is on average better than A1 and A2**

p 29353, l 16 : refer to Figs. 16 and 18, after "For ARCTAS" **added**

p 29353, l 16-18 : however if the conditions are homogeneous, it should not have a large impact. And if the flight legs are long, part of the observations will be in/below/above clouds, and averages can than still make sense. Yes, so that is why it is shown. Sentence has been re-written.

p 29353, l 18 : These average biases -> "The average biases" **changed**

p 29353, l 20 : I would add "(see Fig. 18)" after "In summer" added at end of sentence

p 29254, l 4 : fires -> wild fires changed

p 29354, l 7-8 : is the "," before "back trajectories" correct? Yes, this is a list of the several techniques used by Hornbrook et al.

p 29354, l 15 : "make" -> "makes" corrected

p 29354, l 21 : express 252-258 E as 92-98 W, etc ... Is that necessary?

p 29354, l 21 : between the surface and 850 hPa (to avoid confusion) **changed**

p 29355, l 1 : 252-258E : express as ... W not changed

p 29355, l 19-20 : maybe you can add "POLMIP" **ok**

p 29355, l 23 : "driven to at least some degree by observed ..." -> "driven by, to some degree at least, observed ..."

added commas without rearranging text

p 29355, l 23-24 : meteorology -> meteorological data/fields not changed

p 29355, l 25 : "occurred among the model outputs" -> "occurred in the model results"

not changed

p 29356, l 3 : I would not write "completely", as that is probably a too high expectation

changed to 'better'

p 29356, l 1-6 : this is a very weak conclusion : just illustrating differences, and requiring for "additional model diagnostics in the future". This contradicts with the suggestion in the abstract (line 5-6) "to quantify the differences in model chemistry and transport schemes". Do we learn something about transport schemes? Suggestions for new/other research should be stated at the end of the conclusions. The abstract was perhaps overstated. '... quantify the differences' has been changed to 'assess the differences'. The second sentence of the abstract is referring to POLMIP in general, and Monks et al concludes the transport schemes are similar among the models. The comments about additional diagnostics have been moved to the end of the Conclusions.

p 29356, l 7-9 : Why is this a reason?

'model errors' changed to 'emissions errors'. The full sentence is:

'While the extensive suite of aircraft observations in 2008 at high northern latitudes is extremely valuable for evaluating the models, they cannot uniquely identify the source of emissions errors, as the Arctic is influenced by many sources at lower latitudes.'

p 29356, l 14 : indicate -> indicate that **OK**.

p 29356. l 19 : "dynamics" : this contradicts a bit the sentence on p 29355, l 23-25, where it was suggested that all model represent the "dynamics of the study year". **'dynamics' changed to 'boundary layer parameterizations'.**

p 29356, l 29 : OVOCs should be defined. Or is VOC meant? **Changed to VOCs.**

New Conclusions:

Eleven global or regional chemistry models participated in the POLARCAT Model Intercomparison Project (POLMIP), allowing for an assessment of our current understanding of the chemical and transport processes affecting the distributions of ozone and its precursors in the Arctic. To limit the differences among models, a standard emissions inventory was used. All of the models were driven, to at least some degree, by observed meteorology (GEOS-5, NCEP or ECMWF), and therefore represented the dynamics of the study year (2008).

While the extensive suite of aircraft observations in 2008 at high northern latitudes is extremely valuable for evaluating the models, they cannot uniquely identify the source of emissions errors, as the Arctic is influenced by many sources at lower latitudes. However, several conclusions can be drawn about the emissions inventory used in this study. Based on the comparisons to aircraft observations and the NOAA surface network data, emissions of CO, ethane and propane are clearly too low. The comparisons to satellite retrievals of OMI NO2 show a few regions of consistent model errors that indicate that anthropogenic NO_X emissions are underestimated in East Asia, while fire emissions are overestimated in Siberia. Large differences are seen among the model NO2 tropospheric columns over Europe and China, thus limiting the conclusions that can be drawn regarding the accuracy of the emissions inventory. The large range in modeled NO₂ (where NO_x emissions were the same) also indicates that model chemistry and boundary layer parameterizations can significantly impact NO_X chemistry. More accurate emissions inventories might greatly improve many of the model deficiencies identified in this study. Emissions inventories modified based on inverse modeling results, as well as results of this study, will be used in future work as one step in improving model simulations of Arctic atmospheric composition.

The simultaneous evaluation of the models with observations of reactive nitrogen species and VOCs has illustrated that large differences exist in the model chemical mechanisms, especially in their representation of VOCs and their oxidation. Most of the models showed a negative bias in comparison to ozone observations from son- des and aircraft, with a slightly larger difference in April than in summer. The models frequently underestimated ozone in the free troposphere by 10–20 ppb in the comparison with ozonesondes. In addition, 10– 30% negative model biases were seen in comparison to the mid-troposphere aircraft ozone measurements. Comparisons for ozone precursors such as NO_X , PAN, and VOCs show much greater biases and differences among models. It appears numerous factors are the causes of these model differences. The differences among model photolysis rates and cloud distributions indicate some of the possible causes for differences in modeled OH, which leads to differences in numerous species and ozone production and loss rates.

Some differences among the simulated results are likely due to different physical parameterizations such as convection, boundary layer mixing and ventilation, wet and dry deposition. Additional model diagnostics are required to better understand the differences among models. For example, comparison of the wet deposition rates and fluxes of a number of compounds could be informative in understanding the budgets of NO_V , HO_X and VOCs.

Evaluation of chemical transport models with numerous simultaneous observations, such as those of the POLARCAT aircraft experiments, can assist in a critical assessment of ozone simulations and identify model components in need of improvement. Model representation of the oxidation of VOCs and the NO_y budget can have a significant impact on ozone distributions. Future chemical model comparisons should consider evaluation of VOCs and reactive nitrogen species as an important component of the evaluation of ozone simulations.

DETAILED COMMENTS on the TABLES :

Table 1 : The abbreviation "bb" should be defined in the caption or in the text. BIGALK, BIGENE, ... should be defined (or referenced). **Changed.**

Table 2 : Try to use the same naming for acetone and methanol as in Table 1. **Changed.**

Table 3 : For WRF-Chem, the number of levels is not indicated. In the chemistry description for WRF-Chem is written "MOZART" : is this "MOZART-4"? Number of levels added. 'MOZART' is a standard chemical option in WRF-Chem. It is based on MOZART-4.

DETAILED COMMENTS on the FIGURES :

Figure 2 : The colours for CAM4-chem and CAM5-chem are very similar (as in other figures). Can this be changed? Units for pressure should be mentioned [hPa]. Is the unit mmol/mol for water vapour correct (as often water vapour mixing ratio is expressed in kg/kg or g/kg)? If possible, it would be nice if some extra values where given on the y-axis (e.g. 200, 500, 700 hPa).

The colors for the 2 CAM-chem versions have been changed slightly. The water vapor units are correct. Units added to y-axis labels.

Figures 3-4 : These plots are too small. **Will try to enlarge them.**

Figures 3-6 : Units for Pressure should be mentioned (on the y-axis or in the caption), and if possible extra values should be indicated on the y-axis. **Units added.**

Figure 7 : It would increase the readability of the figure when the names of the species are mentioned on the top of every individual plot (instead of on the y-axis). Only having on the y-axis "mixing ratio [ppbv]" would be ok. The common title "50-70N ZA 700 hPa" is not nicely integrated in the figure. The text in the caption is possibly enough to make this clear.

Will try to make these prettier.

Figure 9 and 10 : It would be nice to have the station latitude (and possibly the longitude) in the top of every individual plot. If possible, it would be nice to have some more pressure values on the y-axis. It would be nice to have the number of sondes indicated in the plots. It is mentioned in the text that there were daily launches during April, but it seems that for some stations there were much less than 30 profiles available.

Some more values added to y-axis, and number of sondes added. The map shows the location of the sites, so have not added latitudes to this plot. The range of dates for the sondes (April 1-19 and June 25-July 12) was amended in the text to be more specific.

Figure 11 : Try to use for the longitude the "... W" notation if the longitude is between 180 E and 359 E. **Not changed.**

Figure 12 : I don't know if the general title for this figure is needed - I would rather use the figure caption for this. Possibly add "upper left panel" after "OMI tropospheric column NO2". The figure is in general too small to read the values and units on the colourbar.

The plots of the individual model biases have been replaced with the multimodel median bias, so there are now just 2 panels (see below).

Figure 13 : The text in this figure is too small. It would be nice to indicate the boxes also in the bias figures (although it is true that the reason for their definitions comes from the OMI plot). The actual link between the individual boxes in Fig. 13 and individual descriptive names used in Fig. 14 and in the text (NW-Europe, NE-US., E-China, Japan, S. Korea, W. Asia, E. Asia, Canada, East. Sib.) are never explicitly made. Maybe one should, or (i) indicate the names in the figures next to the boxes, or (ii) make a table giving the coordinates of the boundaries/corners of the areas. **Removed the panels of the bias, so only 4 panels are shown – anthro- and biomass burning-filtered pixels for each season (see below).**

Figure 14 : A too small font is used in this figure. The word "model" is difficult to read in the upper right blue box. Somewhere should be mentioned that one looks at column values. I would replace "whiskers to" -> "whiskers show" or "and whiskers". **Figure has been remade (see below).**

Figure 15 : Within the figure I would also write "ARCTAS-CARB instead" of "ARCTAS- C" to limit ambiguity. **Changed.**

Figure 16 : Species names should be shown at the top of the figures. There are also x-axis labels winch overlap. One should try to improve the writing of "j o3 o1d" into "j(o3->o1d)" and "j no2" into "j(no2)".

Figure remade (see below).

Figure 17 : The mentioning of "ARCPAC P3 April 11-21" is not very elegant with respect to the rest of the figure (same comment as for Fig. 7). It would also be more practical for the reader if the names of the species should be mentioned at the top of the individual plots instead of below the x-axis.

Figure remade (see below).

Figure 18 : There are overlapping or too close values on the x-axis for some of the plots. Name of the species should be mentioned on the top of each individual plot, instead of on the x-axis. The x-axis should contain "mixing ration [ppbv]" or just "[ppbv]".

Figure remade (see below).

Figure 19 : The way "Campaign" (horizontal) and "Model Bias (%)" (vertical) are added in the figure should be improved. Eventually change the caption to contain the information, e.g. : "Mean bias" -> "Mean bias (%)" and "A1 : ARCTAS ..." -> "The campaigns are ..." . I would also write "ARCTAS-A1" and "ARCTAS-A2" instead of "ARCTAS-A" twice.

Caption changed as suggested.

Figure 20 : In the caption, I think Enhancement Ratio can be written with small letters. Does it make sense to also plot the uncertainty on the estimate from the models? Or does that give no interesting information?

The 1-sigma uncertainty in the derivation of the slopes from the model results are smaller than the symbols in Fig. 20. It is difficult to make further estimates of the model uncertainties.

New Figures



Figure 9. Comparison of models to ozonesondes for April, showing mean and standard deviation of the observations (black line) and the mean bias (colored lines) for each model at each site (Tarasick et al., 2010; Thompson et al., 2011). Results shown for only surface to 300 hPa for clarity. The number of sondes for each site is indicated in the lower right corner of each panel.



Figure 10. As Fig. 9, but for June–July.



Figure 12. a) OMI tropospheric column NO₂ and b) median of the model biases, both for 18 June–15 July.



Figure 13. OMI NO₂ filtered for dominant anthropogenic emissions (**a**, **b**) and fire emissions (**c**, **d**). Boxes indicate the regions for which biases have been calculated in Fig. 14. Panels **a**, **c** show April, panels **b**, **d** show June-July.



Figure 14. Summary of the regional means from each model and the OMI NO₂ tropospheric columns for each region indicated in Figure 13. (a) Anthropogenic emissions in April and (b) June–July, (c) biomass burning in both seasons. Red circles are mean OMI NO₂ observations for the region; box plots show median, 25th and 75th quartiles, whiskers to 5th and 95th percentiles of the model means.



Figure 16.



Figure 17.



Figure 18.

Discussion Paper

Manuscript prepared for Atmos. Chem. Phys. Discuss. with version 2014/09/16 7.15 Copernicus papers of the Later Class copernicus.cls. Date: 20 March 2015

The POLARCAT Model Intercomparison Project (POLMIP): overview and evaluation with observations

L. K. Emmons¹, S. R. Arnold², S. A. Monks², V. Huijnen³, S. Tilmes¹, K. S. Law⁴, J. L. Thomas⁴, J.-C. Raut⁴, I. Bouarar^{4,*}, S. Turquety⁵, Y. Long⁵, B. Duncan⁶, S. Steenrod⁶, S. Strode^{6,**}, J. Flemming⁷, J. Mao⁸, J. Langner⁹, A. M. Thompson⁶, D. Tarasick¹⁰, E. C. Apel¹, D. R. Blake¹¹, R. C. Cohen¹², J. Dibb¹³, G. S. Diskin¹⁴, A. Fried¹⁵, S. R. Hall¹, L. G. Huey¹⁶, A. J. Weinheimer¹, A. Wisthaler¹⁷, T. Mikoviny¹⁷, J. Nowak^{18,***}, J. Peischl¹⁸, J. M. Roberts¹⁸, T. Ryerson¹⁸, C. Warneke¹⁸, and D. Helmig¹⁹

¹Atmospheric Chemistry Division, National Center for Atmospheric Research, Boulder, Colorado, USA

²Institute for Climate and Atmospheric Science, University of Leeds, Leeds, UK

³KNMI, the Netherlands

⁴Sorbonne Universités, UPMC Univ. Paris 06, Université Versailles St-Quentin, CNRS/INSU, LATMOS-IPSL, UMR 8190, Paris, France

⁵Laboratoire de Météorologie Dynamique, IPSL, CNRS, UMR8539, 91128 Palaiseau CEDEX, France

⁶NASA Goddard, Atmospheric Chemistry and Dynamics Laboratory, Code 614, Greenbelt, Maryland, USA

⁷ECMWF, Reading, UK

⁸NOAA GFDL and Princeton University, Princeton, New Jersey, USA

⁹Swedish Meteorological and Hydrological Institute, SE-60176 Nörrkping, Sweden

Discussion Paper

Discussion Paper

¹⁰Environment Canada, Downsview, Ontario, Canada
¹¹Department of Chemistry, University of California-Irvine, Irvine, California, USA
¹²Chemistry Department, University of California-Berkeley, Berkeley, California, USA
¹³University of New Hampshire, Durham, NH, USA
¹⁴NASA Langley Research Center, Chemistry and Dynamics Branch, Hampton, VA, USA
¹⁵University of Colorado, Boulder, CO, USA
¹⁶Georgia Institute of Technology, Atlanta, GA, USA
¹⁷University of Innsbruck, Austria and University of Oslo, Norway
¹⁸NOAA Earth System Research Lab, Boulder, CO, USA
¹⁹INSTAAR, University of Colorado, Boulder, CO, USA
¹⁹INSTAAR, University of Colorado, Boulder, CO, USA
^{**}now at: Max Planck Institute for Meteorology (MPI-M), Hamburg, Germany
^{***}also at: Universities Space Research Association, Columbia, MD, USA

Correspondence to: L. K. Emmons (emmons@ucar.edu)

Discussion Paper

Abstract

A model intercomparison activity was inspired by the large suite of atmospheric chemistry observations observations of atmospheric composition made during the International Polar Year (2008) in the Arctic. Nine global and two regional chemical transport models participated in this intercomparison and have performed simulations for 2008 using a common emissions inventory to quantify assess the differences in model chemistry and transport schemes. This paper summarizes the models and compares their simulations of ozone and its precursors, and presents an evaluation of the simulations using a variety of surface, balloon, aircraft and satellite observations. Each type of measurement has some limitations in spatial or temporal coverage or in composition, but together they assist in quantifying the limitations of the models in the Arctic and surrounding regions. Despite using the same emissions, large differences are seen among the models. Differences in a number of model parameters are identified as contributing to The cloud fields and photolysis rates are shown to vary greatly among the models, indicating one source of the differences in the modelled chemical species, including cloud fields and photolysis ratessimulated chemical species. The largest differences among models, and between models and observations, are in NO_{ν} partitioning (PAN vs. HNO₃) and in oxygenated volatile organic compounds (VOCs) such as acetaldehyde and acetone. Comparisons to surface site measurements of ethane and propane indicate that the emissions of these species are significantly underestimated. While limited in spatial and temporal coverage, the aircraft measurements provide a simultaneous evaluation of many species. Satellite observations of NO₂ from OMI have been used to evaluate the models over source regions, indicating anthropogenic emissions are underestimated in East Asia, but fire emissions are generally overestimated. The emission factors for wildfires in Canada are evaluated using the correlations of VOCs to CO in the model output in comparison to enhancement factors derived from aircraft observations, showing reasonable agreement for methanol and acetaldehyde, but underestimate of ethanol, propane and acetone, while overestimating ethane emission factors.
1 Introduction

Observations show that the Arctic has warmed much more rapidly in the past few decades than global-mean temperature increases. Arctic temperatures are affected by both increased heat transport from lower latitudes and by local in-situ response to radiative forcing from due to changes in greenhouse gases and aerosols (Shindell, 2007). Model calculations suggest that in addition to warming induced by increases in global atmospheric CO₂ concentrations, changes in short-lived climate pollutants (SLCPs), such as tropospheric ozone and aerosol aerosols in the Northern Hemisphere (NH), have contributed substantially to this Arctic warming since 1890 (Shindell and Faluvegi, 2009). This contribution from SLCPs to Arctic heating and efficient local amplification mechanisms (e.g., ice-albedo feedback) put a high priority on understanding the sources and sinks of SLCPs at high latitudes and their climatic effects. Despite the remoteness of the Arctic region, anthropogenic sources in Europe, North America and Asia have been shown to contribute substantially to Arctic tropospheric burdens of SLCPs (e.g., Fisher et al., 2010; Sharma et al., 2013; Monks et al., 2014; Law et al., 2014). The Arctic troposphere is more polluted in winter and spring as a result of long-range transport from northern mid-latitude continents and the lack of efficient photochemical activity or wet scavenging needed to cleanse the atmosphere (Barrie, 1986).

Large forest fires in boreal Eurasia and North America also impact the Arctic in the spring and summer seasons (Sodemann et al., 2011). Our understanding of contributions from SLCP sources to present-day Arctic heating is sensitive to the ability of models to simulate the transport and processing of SLCPs en-route to the Arctic from lower latitude sources. This model skill has implications for our confidence in predictions of Arctic climate response to future changes in mid-latitude anthropogenic and wildfire emissions.

Comparisons of model results to long-term surface observations have shown that global chemical transport models have significant (CTMs) have notable limitations in accurately simulating the Arctic tropospheric composition, as well as having significant differences among models (e.g., Shindell et al., 2008). Transport of emissions from lower latitudes to

the Arctic is mainly facilitated by rapid poleward transport in warm conveyor belt airstreams associated with frontal systems of mid-latitude cyclones (Stohl, 2006). The result is that although it the Arctic is remote from source regions, Arctic enhancements in trace gas and aerosol pollution are far from homogeneous. They are instead characterized by episodic import of pollution-enhanced air masses, exported from the mid-latitude boundary layer by large-scale advection in frontal systems (Sodemann et al., 2011; Schmale et al., 2011; Quennehen et al., 2012). Polluted air uplifted from warmer, southerly latitudes (Asia and North America) tends to enter the Arctic at higher altitude, while air near the surface is influenced mainly by low-level flow from colder, more northerly source regions, particularly Europe (e.g., Klonecki et al., 2003; Stohl, 2006; Helmig et al., 2007b; Tilmes et al., 2011; Wespes et al., 2012). The stratification of the large-scale advection and slow vertical mixing leads to fine-scale layering and filamentary air mass structure through the Arctic troposphere, where air masses from different source origins produce distinct layers, and are stirred together only in close proximity, while mainly retaining their own chemical signatures (Engvall et al., 2008; Schmale et al., 2011). Air masses are eventually homogenized by turbulent mixing and radiative cooling, but usually on timescales longer than the rapid advection timescale along isentropic surfaces that is characteristic of intercontinental transport in these systems (Methven et al., 2003, 2006; Stohl, 2006; Arnold et al., 2007; Real et al., 2007). For global Eulerian models, representing this fine-scale structure, subsequent slow mixing and chemical processing is challenging, particularly with characteristic coarse grid-sizes of a hundred kilometers and or more.

A large suite of observations was collected during the International Polar Year (2008) as part of the international POLARCAT (Polar Study using Aircraft, Remote Sensing, Surface Measurements and Models, of Climate, Chemistry, Aerosols and Transport) activity (Law et al., 2014). Numerous papers have been written on these observations and corresponding model simulations (many are in a special issue of Atmospheric Chemistry and Physics: http://www.atmos-chem-phys.net/special_issue182.html).

The POLARCAT Model Intercomparison Project (POLMIP) was organized with the goal of exploiting this large data set to comprehensively evaluate several global chemistry models

and to better understand the causes of model deficiencies in the Arctic. While aerosols are an important component of the Arctic atmospheric composition, this comparison focused on gas phase POLMIP focuses on gas-phase chemistry, primarily CO, carbon monoxide (CO), reactive nitrogen $(NO_v and)$, ozone (O_3) and their precursors. This paper provides an overview of the POLMIP models and their evaluation against observations, as well as an evaluation of the emissions inventories used by the models. Two companion additional papers present more detailed analyses (Monks et al., 2014; Arnold et al., 2014). Monks et al. (2014) comprehensively evaluate the model CO and O_3 distributions with surface, aircraft and satellite observations, as well as compare the effects of chemistry and transport using synthetic tracers. In general, the models are found to underestimate both CO and O₃, while the modeled global mean OH amounts are slightly higher than estimates constrained by methyl chloroform observations and emissionsits emission estimates. suggesting the model errors are not entirely due to low emissions. The comparison of fixedlifetime tracers to idealized OH-loss CO-like species shows that the differences in OH concentrations among models have a greater impact on CO than transport doesdifferences in transport in the models. The tracer analysis also shows a very strong influence of fire emissions on the atmospheric composition of the Arctic. Ozone production in air influenced by biomass burning is evaluated by Arnold et al. (2014). Using tracers of anthropogenic and fire emissions, fire-dominated air was found to have enhanced ozone in the POLMIP models, with the enhancement increasing with airmass age. Differences in NO_{v} partitioning are seen among models, likely due to model differences in efficiency in of vertical transport as well as VOC oxidation schemes.

The next sections section of this paper give gives an overview of the POLARCAT aircraft campaigns that prompted this intercomparison , and (Section 2). Section 3 presents a summary of the models that participated. Following these are comparisons of all of the models to observations, along with a description of the model experiment design and the emissions used. Comparisons of the model results with observations are shown in Section 4, including vertical profiles of ozone from sondes, including ozonesondes, surface layer non-methane hydrocarbons (NMHC), satellite observations of NO₂, and the numerous compounds mea-

sured from research aircraft. Finally, an evaluation of emissions is performed, using satellite observations of , and the fire emission factors is shown in Section 5, using aircraft observations of fire-influenced air masses.

2 POLARCAT observations

POLARCAT is a collaboration consortium of tropospheric chemistry experiments performed during the International Polar Year (IPY) 2008 (Law et al., 2014). A wealth of data on tropospheric ozone and its photochemical precursors were obtained through the depth of the Arctic troposphere during spring and summer. These observations provide an opportunity to evaluate model representations of processes controlling tropospheric ozone in imported pollutant layers above the surface. The NASA Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission (Jacob et al., 2010) was grouped into three parts, ARCTAS-A, ARCTAS-B and ARCTAS-CARB. Three research aircraft took part in this campaign with a slightly different goal of for each mission. ARCTAS-A and ARCTAS-B targeted mid-latitude pollution layers transported to the Arctic and wildfire plumes, respectively, transported to the Arctic. ARCTAS-CARB was focused on California air quality targeting fresh fire plumes in northern California, as well as various anthropogenic sources (Huang et al., 2010; Pfister et al., 2011).

The NOAA ARCPAC (Aerosol, Radiation, and Cloud Processes affecting Arctic Climate (ARCPAC) mission was conducted in spring between the end of March and 21 April using the NOAA P3 aircraft (Brock et al., 2011). It was designed to understand the radiative impacts of anthropogenic pollution and biomass burningsources. The campaign was based in Fairbanks, Alaska, and frequently targeted fire wildfire plumes that were transported from Siberia (e.g., Warneke et al., 2009).

POLARCAT-France, using the French ATR-42 aircraft, was based in Kiruna, Sweden, in Spring, and took place between 30 March and 11 April (de Villiers et al., 2010; Merlaud et al., 2011). The summer mission was based in western Greenland in Kangerlussuaq and

took place between end of June and mid-July (Schmale et al., 2011; Quennehen et al., 2011).

The POLARCAT-GRACE POLARCAT-Greenland Aerosol and Chemistry Experiment (GRACE) mission was conducted during the same time (1–17 July), based also at Kangerlussuaq, Greenland, using the DLR Falcon research aircraft (Roiger et al., 2011). Flights covered latitudes from 57 to 81° N and targeted anthropogenic and fire emissions in the troposphere and lower stratosphere.

YAK-AEROSIB (Airborne Extensive Regional Observations in Siberia) was conducted in July 2008 covering parts of North and Central Siberia using an Antonov-30 research aircraft, operated by the Tomsk Institute of Atmospheric Optics (Paris et al., 2008). This campaign was also performed in collaboration with the POLARCAT program (Paris et al., 2009).

3 ModelsModel configurations and inputs

3.1 Design of model intercomparison

Simulations were run for each model over the same time period, from 1 January 2007 to 31 December 2008. This includes a 1 year spin-up period leading into a full 12 month simulation (January–December 2008) used in the analysis. All models used the same emissions inventory (with a few exceptions) A single emissions inventory was specified for use by all of the models, as described below. Each global model was run at its standard global resolution with its standard chemistry scheme, meteorology, and deposition schemes meteorological forcing and other parameterizations. The requested model output included monthly mean output of a number of species species distributions and diagnostics to allow evaluation of the seasonal cycles of the models using surface and satellite observations. Hourly instantaneous output of a smaller number of species was requested for 30 March–23 April and 18 June–18 July (20–90° N) to allow comparison to the aircraft observations and NO₂ satellite retrievals.

3.2 Description of emissions

For this study, a single set of emissions was specifiedAll modeling groups were asked to use the same emissions inventory. The anthropogenic emissions are from the inventory provided by D. Streets (Argonne National Lab) and University of Iowa for ARCTAS (http://bio. cgrer.uiowa.edu/arctas/emission.html; http://bio.cgrer.uiowa.edu/arctas/arctas/07222009/). This inventory is a composite of regional inventories, including Zhang et al. (2009) for Asia, USNEI 2002 and CAC 2005 for North America, and EMEP 2006 expert emissions (http://www.ceip.at) for Europe. Missing regions and species were filled with EDGAR 3.2FT2000. Only total volatile organic compounds (VOCs) were provided with this inventory, so speciation was the emissions for specific hydrocarbons were based on the VOC speciation of the RETRO inventory as in Lamarque et al. (2010). The anthropogenic emissions do not include any seasonal variation. Biomass Daily biomass burning emissions are from the Fire Inventory INventory of NCAR (FINN), which are based on MODIS fire counts and provided daily (Wiedinmyer et al., 2011). Other emissions (biogenic, ocean, volcano) were derived from the POET inventory (Granier et al., 2005). A preliminary comparison of these the ARCTAS emissions to the MACCity inventory (Granier et al., 2011) showed the ARCTAS inventory has higher emissions and produced results closer to the observations in a MOZART simulation. Due to the different speciation of VOCs in the models, there is some slight difference in emissions totals. Details of these different VOC treatments are given with the model descriptions below. Table 1 gives the emissions totals for each species provided, by sector, while Table 2 gives totals calculated from the supplied output. Each model determined lightning emissions based on their usual formulation, as described below.

3.3 Artificial tracers

One goal of the <u>POLMIP</u> intercomparison was to separately compare dynamics tracer transport and chemistry among the models. Synthetic tracers with a fixed lifetime are a valuable tool for comparing transport only. Artificial fixed-lifetime tracers emitted from anthropogenic and wildfire sources of CO were specified for 3 regions: Europe (30–90° N, 30° W– 60° E), Asia (0–90° N, 60–180° E), and North America (25–90° N, 180–30° W). Each model (except GEOS-Chem) included these tracers in their simulation with a set lifetime of 25 days. Figure 1 shows the anthropogenic CO emissions, with fire emissions overlaid, for April and July monthly averages. The highest fire emissions are generally far removed from the anthropogenic emissions (e.g., northern Canada and Siberia) within each region. This Since anthropogenic and biomass burning emissions have different relative amounts of CO, NO_x and VOCs, the offset in location produces of the two source types leads to significant differences in atmospheric composition within these regionsand should be kept in mind when considering the distributions of these tracers in the. These differences have particular relevance in the analyses of Monks et al. (2014) and Arnold et al. (2014) that use these tracers. Figure 1 also shows the daily variation of the fire emissions averaged over each tracer region. Asia had high fire emissions from March through August July, but at different locations through that period (e.g., farther north in July than April). Biomass burning in eastern Europe began in April, with stronger fires in August. The North America fire emissions were significantly less on average, but were locally important in California and Saskatchewan in June and July.

3.4 Description of POLMIP models

Nine global and two regional models participated in POLMIP. The resolution and meteorology of origin of meteorological analyses of each model is given in Table 3. Additional details are given below. Lightning emissions totals are included in Table 1 or below.

CAM-chem. The Community Atmosphere Model with Chemistry (CAM-chem) is a component of the NCAR Community Earth System Model (CESM). Two versions were used in POLMIP, based on versions 4 and 5 of CAM. The CAM4-chem results shown here are slightly updated from those described in Lamarque et al. (2012), while CAM5-chem includes expanded microphysics and modal aerosols (Liu et al., 2012). Both versions of CAM-chem use the MOZART-4 tropospheric chemistry scheme (see MOZART description below), along with stratospheric chemistry, and are evaluated in Tilmes et al. (2014). For

POLMIP, CAM-chem was run in the "specified dynamics" mode, where the meteorology (temperature, winds, surface heat and water fluxes) is nudged to meteorological fields from GEOS-5, using the lowest 56 levels. Lightning NO emissions are determined according to the cloud height parameterization of Price and Rind (1992) and Price et al. (1997). The vertical distribution follows DeCaria et al. (2005) and the strength of intra-cloud and cloud-ground strikes are assumed equal, as recommended by Ridley et al. (2005).

C-IFS (Composition-IFS). The integrated forecasting system Integrated Forecasting System (IFS) of the European Centre for Medium Range Weather Forecasting (ECMWF) has been extended for the simulation of atmospheric composition in recent years. For the POLMIP runs, the CB05 chemical scheme as implemented in the TM5 chemical transport model (CTM) (Huijnen et al., 2010) has been used (Flemming, 2014). C-IFS uses a semi-Lagrangian advection schemeand applies the emission and dry deposition fluxes as part of the vertical diffusion scheme. The POLMIP runs are a sequence of 24 h forecasts, initialized with the operational meteorological analysis. Lightning emissions in C-IFS are based on the model convective precipitation (Meijer et al., 2001) and use the C-shaped profile suggested by Pickering et al. (1998). The , and follows the same implementation as TM5, except that the lightning emissions are scaled to give a -global annual total of 4.9 Tg N yr⁻¹.

GEOS-Chem. GEOS-Chem is a global 3-D chemical transport model driven by assimilated meteorological observations from the Goddard Earth Observing System (GEOS-5) of the NASA Global Modeling and Assimilation Office (GMAO) (Bey et al., 2001). GEOS-Chem version 9-01-03 (http://www.geos-chem.org) was used for this study. The standard GEOS-Chem simulation of ozone-NO_x-HO_x-VOC chemistry is described by Mao et al. (2010), with more recent implementation of bromine chemistry (Parrella et al., 2012). The chemical mechanism includes updated recommendations from the Jet Propulsion Laboratory (Sander et al., 2011) and the International Union of Pure and Applied Chemistry (http://www.iupac-kinetic.ch.cam.ac.uk). In addition, this simulation includes an HO₂ aerosol reactive uptake with a coefficient of γ (HO₂) = 1 producing H₂O suggested by (Mao et al., 2013). Lightning NO emissions are computed with the algorithm of Price and Rind

(1992) as a function of cloud top height, and scaled globally as described by Murray et al. (2012) to match OTD/LIS climatological observations of lightning flashes.

GMI. GMI (Global Modeling Initiative; http://gmi.gsfc.nasa.gov/) is a NASA offline global chemical transport modelCTM, with a comprehensive representation of tropospheric and stratospheric chemistry (Duncan et al., 2007; Strahan et al., 2007). The simulations for POLMIP were driven by MERRA meteorology, with all of the emissions from the specified inventory. The GMI chemical mechanism treats explicitly the lower hydrocarbons (ethane, propane, isoprene) and has two lumped species for larger alkanes and alkenes following Bey et al. (2001). Several oxygenated hydrocarbons (e.g., formaldehyde, acetaldehyde) are simulated, including direct emissions and chemical production; acetone is specified from a fixed field. The mechanism includes 131 species and over 400 chemical reactions. Flash rates are parameterized in terms of upper tropospheric convective mass flux but scaled so that the seasonally averaged flash rate in each grid box matches the v2.2 OTD/LIS climatology.

LMDZ-INCALMDz-INCA. The LMDz-OR-INCA model consists of the coupling of three individual models. The Interaction between Chemistry and Aerosol (INCA) model is coupled online to the LMDz (Laboratoire de Météorologie Dynamique) General Circulation Model (LMDz) general circulation model (GCM) (Hourdin et al., 2006). LMDz used for the POLMIP exercise is coupled with the ORCHIDEE (Organizing Carbon and Hydrology in Dynamic Ecosystems) dynamic global vegetation for soil–atmosphere exchanges of water and energy (Krinner et al., 2005), but not for biogenic CO₂ or VOC fluxes. INCA is used to simulate the distribution of aerosols and gaseous reactive species in the troposphere. The oxidation scheme was initially described in Hauglustaine et al. (2004) including inorganic and non-methane hydrocarbon chemistry. INCA includes 85 tracers and 264 gas phase gas-phase reactions. For aerosols, the INCA model simulates the distribution of anthropogenic aerosols such as sea salt and dust. LMDz-OR-INCA is forced with horizontal winds from 6h-6-hourly ECMWF ERA Interim reanalysis. Lightning NO emissions are computed interactively during the simulations depending on the convective clouds, according

to Price and Rind (1992), with a vertical distribution based on Pickering et al. (1998) as described in Jourdain and Hauglustaine (2001). The global annual <u>lightning emissions</u> total is 5 Tg N yr^{-1} .

MOZART-4. MOZART-4 (Model for Ozone and Related chemical Tracers, version 4) is an offline global chemical transport modelCTM, with a comprehensive representation of tropospheric chemistry (Emmons et al., 2010). For these simulations all of the emissions were from the specified inventory (i.e., online MEGAN was not used for biogenic emissions) While MOZART-4 includes the capability to calculate biogenic isoprene and terpenes using the MEGAN algorithms, the specified monthly mean emissions were used for POLMIP. Simulations were run with both the an online photolysis calculation (FTUV) and using the lookup table that is a lookup table (LUT), which is the same as that used in CAM-chem (described) The MOZART-4 chemical mechanism treats explicitly the lower hydrocarbons (C₂H₆, C₃H₈, $C_{2}H_{4}$, $C_{3}H_{6}$, $C_{2}H_{2}$, isoprene) and has four lumped species for larger alkanes, alkenes, aromatics and monoterpenes. Several A number of oxygenated hydrocarbons (including formaldehyde, acetaldehyde, acetone, methanol, ethanol) are simulated, including also explicitly treated with direct emissions and chemical production. The mechanism includes 100 species and 200 chemical reactions and the tropospheric gas-phase chemistry is the same as that used in the CAM-chem simulations for this study. Emissions of NO from lightning are parameterized as described above for CAM-chem (Emmons et al., 2010).

TM5. TM5 (Tracer Model 5) is an offline global chemical transport model (Huijnen et al., 2010), where tropospheric chemistry is described by a modified carbon bond chemistry mechanism (Williams et al., 2013). The TM5 chemical mechanism includes explicit treatment of the lower hydrocarbons (C_2H_6 , C_3H_8 , C_3H_6) and acetone, while other VOCs are treated in bulk. The mechanism is based on the CB05 scheme with modifications to the ROOH oxidation rate and HO₂ production efficiency from the isoprene + OH oxidation reaction (Williams et al., 2012). Photolysis is modelled modeled by the modified band approach (Williams et al., 2012). It is total, the TM5 chemical mechanism includes 55 species and 104 chemical reactions. Stratospheric O₃ is constrained using ozone columns from the

Multi-Sensor Reanalysis (van der A et al., 2010). For these simulations essentially all of the emissions were taken from the specified inventory, although all NO_x is emitted as NO. production from lightning is calculated using a linear relationship between lightning flashes and convective precipitation (Meijer et al., 2001), using a C-shaped profile suggested by Pickering et al. (1998). Marine lightning is assumed to be 10 times less active as than lightning over land. The fraction of cloud-to-ground over total flashes is determined by a fourth-order polynomial function of the cold cloud thickness (Price and Rind, 1992). The NO_x production for intra-cloud flashes is 10 times less than that for cloud-to-ground flashes, according to Price et al. (1997).

TOMCAT. The TOMCAT model is a Eulerian three-dimensional (3-D)-global CTM (Chipperfield, 2006). This study uses an extended VOC degradation chemistry scheme, which incorporates the oxidation of monoterpenes based on the MOZART-3 scheme and the oxidation of $C_2 - C_4$ alkanes, toluene, ethene, propene, acetone, methanol and acetaldehyde based on the ExTC (Extended Tropospheric chemistryChemistry) scheme (Folberth et al., 2006). Heterogeneous N₂O₅ hydrolysis is included using offline size-resolved aerosol from the GLOMAP model (Mann et al., 2010). The implementation of these two chemistry schemes into TOMCAT is described by and-Monks (2011) and Richards et al. (2013) and has 82 tracers and 229 gas-phase reactions. All anthropogenic, biomass burning and natural emissions were provided by POLMIP, with the exception of lightning emissions, which are coupled to the amount of convection in the model and therefore vary in space and time (Stockwell et al., 1999).

SMHI-MATCH. SMHI-MATCH (Multiple-scale Atmospheric Transport and Chemistry Modeling System) is an offline 3-D chemistry transport model CTM developed at the Swedish Meteorological and Hydrological Institute (Robertson et al., 1999). SMHI-MATCH can be run on both global and regional domains but for POLMIP model runs were performed for the 20–90° N region. The chemical scheme in MATCH considers 61 species using 130 chemical reactions and is based on Simpson (1992) but with extended isoprene chemistry and updated reactions and reaction rates. Information about the implementation of the chemical scheme can be found in Andersson et al. (2007), where evaluation of standard simulations for the European domain is also given. We used ERA-Interim re-analysis data from ECMWF were used to drive SMHI-MATCH for the years 2007 and 2008. Sixhourly data (3hourly 3-hourly for precipitation) were extracted from the ECMWF archives on a $0.75^{\circ} \times 0.75^{\circ}$ rotated latitude–longitude grid. The original data had 60 levels, but only the 35 lowest levels reaching to about 16 km in the Arctic were used in SMHI-MATCH. Monthly average results for 2007 and 2008 from global model runs using MOZART at ECMWF in the MACC (Monitoring Atmospheric Chemical Composition) project were used as both upper and 20° N chemical boundary conditions. In addition to the standard daily POLMIP emissions, NO emissions from lightning were included using monthly data from the GEIAv1 data set, which has an annual global total of $12.2 \,\mathrm{Tg}\,\mathrm{N}\,\mathrm{yr}^{-1}$. DMS emissions were simulated using monthly DMS ocean concentrations and the flux parameterization from Lana et al. (2011).

WRF-Chem. The Weather Research and Forecasting model with Chemistry (WRF-Chem) is a regional CTM, which calculates online chemistry and meteorology (Grell et al., 2005; Fast et al., 2006). For the POLMIP runs the meteorology parameterizations are as described in the WRF-Chem (version 3.4.1) simulations of Thomas et al. (2013). Briefly, the initial and boundary conditions for meteorology are taken from the NCEP Final Analyses (FNL), with nudging applied to wind, temperature, and humidity every 6 h. The MOZART-4 POLMIP run is used for both initial and boundary conditions for gases and aerosols. The POLMIP emissions were used, however, the FINN fire emissions were processed using the WRF-Chem FINN processor, so the fire emissions are at finer resolution than 1° (used by the global models). In addition, an online fire plume rise model was employed (Freitas et al., 2007). Lightning emissions were included using the Price and Rind (1992) parameterization as described in Wong et al. (2013). WRF-Chem was run at two model resolutions (50 and 100 km) during the summer POLARCAT campaigns, with 65 levels from the surface to 50 hPa. Selected chemical species (e.g. ozone) are set to climatological values above 50 hPa and relaxed to a climatology down to the tropopause. For the POLMIP runs, WRF-Chem employs the MOZART-4 gas-phase chemical scheme described in Emmons et al. (2010) and bulk aerosol scheme GOCART (Goddard Chemistry Aerosol Radiation and Transport model, Chin et al., 2002), together referred to as MOZCART. The model was run from 28 June 2008 to 18 July 2008 using a polar-stereographic grid over a domain encompassing both boreal fires and anthropogenic emission regions in N. America to include the ARCTAS-B, POLARCAT-GRACE, and POLARCAT-France flights. Because of the limited temporal and spatial extent of the WRF-Chem results they could not be included in some of the plots and analysis below.

4 Summary Overview of model resultscharacteristics and differences

In order to better understand the differences among models shown later in the comparison to observations, some general model characteristics are illustrated. Figure 2 shows zonal averages of temperature and water vapour vapor for each of the models. As the models are driven or nudged by assimilated meteorology fields, their temperatures are in close agreement. One exception is **LMDZ-INCA**LMDz-INCA, which has only horizontal winds nudged to ECMWF winds and the remaining meteorological fields are calculated. The temperature differences seen here between LMDZ-INCA LMDz-INCA and the other models are comparable to those seen in the ACCMIP comparisons (Lamargue et al., 2013). The models show some variation in water vapourvapor, particularly in the tropics and in the upper troposphere at high northern latitudes. Some models (e.g., MOZART-4, CAM-chem) calculate water vapour vapor and clouds based on surface fluxes, while others use the GEOS-5 or ECMWF provided specific humidity values. Significant differences in cloud distributions are seen among the models, as shown in Fig. 3 (not available from GEOS-Chem, LMDZ-INCA LMDz-INCA or TOMCAT). While MOZART-4, CAM4-chem and CAM5-chem are driven with the same GEOS-5 surface fluxes, the cloud physics, turbulent mixing and convection schemes differ among the models and result in guite different cloud distributions. These differences can lead to significant differences in photolysis rates, as shown in Fig. 4. For example, CAM5-chem has greater cloud fractions in the tropical upper troposphere than CAM4-chem, which leads to lower photolysis rates, particularly noticeable in $J(O_3 \rightarrow O^1D)$. MOZART-FTUV simulations used the online Fast-TUV photolysis scheme that includes the impact of aerosols on photolysis, but also has some outdated cross-sections, that are the larger source of the differences with the MOZART-4 results.

All of these inter-model differences in physical parameters, along with differing transport schemes, lead to differences, to varying degrees, in the modelled modeled ozone and OH distributions. Figure 5 summarizes these model differences by plotting the pressure-latitude location of the 50 and 100 ppb ozone contours of the April and July zonal averages. The $100 \text{ ppbv } O_3$ contour line is one method used to estimate the location of the tropopause. The model results shown here generally agree on in the location of the 100 ppb contour, with two exceptions indicating a lower Arctic tropopause height: MATCH in April and July, and TOMCAT in July. The models vary widely in the distribution of tropospheric ozone. In April at high northern latitudes, the 50 ppb O₃ contour for GEOS-Chem is at the highest altitude (500 hPa at 50° N) while GMI is at the lowest (900 hPa). Great variability is also seen in the tropics in both April and July. Some model differences in the lower troposphere could be due to different ozone dry deposition velocities, which can have a significant impact on ozone in the surface boundary layer (Helmig et al., 2007a). However, ozone deposition rates were not provided for this intercomparison so this impact cannot be assessed. In the upper troposphere, model differences are more likely driven by differences in stratosphere-troposphere exchange. In addition, ozone chemical production and loss rates determine model ozone distributions, as indicated in the comparison of ozone precursors, below.

Figure 6 similarly shows the zonal averages of OH, illustrating the large differences among models in the magnitude of OH concentration. In April, most of the models have values above 2×10^6 molecules cm⁻³ in the northern Tropics from the surface to 500 hPa. GMI is the only model to show a maximum greater than 2×10^6 molecules cm⁻³ also in the upper troposphere. About half of the models have OH concentrations of at least 1×10^6 molecules cm⁻³ throughout the troposphere between latitudes 20° S and 50° N. In July, even greater variability among models is seen in the shapes of both contour levels.

To further illustrate and understand the differences in the modelled ozone, a number of compounds modeled ozone, the time series of ozone and its precursors are plotted in Fig. 7

as monthly zonal averages at 700 hPa over 50–70° N, the latitude range of most of the aircraft observations. As in Fig. 5, wide variation among models is seen for ozone. Here we see disagreement in even the shape of the seasonal cycle. The mixing ratios of carbon monoxide (CO) differ among models by 50 %, largely due to the differences in OH, but also affected by concentrations of hydrocarbons that are precursors of CO. The differences in CO among models are discussed in detail in Monks et al. (2014). Ethane (C_2H_6) is only directly emitted, without any secondary chemical production, so the differences among models are due to OH or emissions. GEOS-Chem used slightly different emissions (see Table +2) and MATCH included acetone (CH₃COCH₃) and acetylene (C₂H₂) emissions in the ethane emissions as they do not simulate those species. The differences in H_2O_2 (hydrogen peroxide) are likely a result of different washout mechanisms in the models, but are also related to the HO₂ differences. In addition, the heterogenous uptake of HO₂ on aerosols may differ significantly among models (e.g., Mao et al., 2013), but was not investigated in this comparison. LMDZ LMDz-INCA and TOMCAT have higher NO₂, PAN and HNO₃ than others. GEOS-Chem has low PAN, but relatively high HNO₃. TM5 and C-IFS have lower formaldehyde (CH₂O) than other models. High variability is seen among the models for acetaldehyde (CH₃CHO) and acetone, with some disagreement in the seasonal cycle. The models have varying complexity in the hydrocarbon oxidation schemes, which contributes to the differences in these oxygenated VOCs, as discussed in Arnold et al. (2014). The differences among models are further explained below with regard to comparisons to observations.

5 Comparison to observations

An overall evaluation of the models is presented here through comparison to ozonesondes, surface network NMHC measurements, satellite retrievals of NO₂, and simultaneous observations of ozone and its precursors from aircraft. A comprehensive evaluation of the CO distributions in the POLMIP models is presented by Monks et al. (2014).

5.1 Ozonesondes

Coincident with the NASA ARCTAS aircraft experiment, daily ozonesondes were launched at a number of sites across North America (Fig. 8) in April and June July during April 1–19 and June 25–July 12 2008 for the Arctic Intensive Ozonesonde Network Study (AR-CIONS; http://croc.gsfc.nasa.gov/arcions) (Tarasick et al., 2010; Thompson et al., 2011). Ozonesondes with their high vertical resolution and absolute accuracy of $\pm (5-10)$ % are extremely valuable for model evaluation. The hourly POLMIP model output was matched to the time and location of each ozonesonde. Since the models, with roughly 0.5-1 km vertical layer spacing in the free troposphere cannot reproduce all of the observed structure, the ozonesonde data and model profiles were binned to 100 hPa layers. The mean of each bin between the surface and 300 hPa was used to calculate the bias between model and measurements for each profile. Figures 9 and 10 show the mean and standard deviation of the observed ozone profiles from each sondeat each site, along with the mean bias for each sonde profile, averaged for each model. model, determined by averaging the difference between each model profile and the corresponding sonde profile. A small number of sondes were launched from Narragansett (4 in April; 3 in June-July), so they have not been used here.

In April, the models generally underestimate the observed ozone profiles throughout the troposphere (negative bias). One consistent exception is SMHI-MATCH, which is higher than observed in the middle and upper troposphere, perhaps indicating that this model has too strong transport of ozone from the stratosphere. At all sites, GEOS-Chem has the lowest ozone values at all altitudes above the boundary layer. TOMCAT also has among the largest negative bias , particularly in the lower troposphereand mid-troposphere, but is higher than most of the models at 300 hPa. All the other models have a fairly uniform (across altitude and sites) negative bias of about 5–10 ppb. The models have slightly lower biases in June–July on average. At Kelowna and Goose Bay, the model biases fall within \pm 10 ppb; however, at several other sites (e.g., Churchill and Bratt's Lake), the model mean

biases are as much as 20 ppb below the observations. These comparisons are consistent with the ozone evaluation using aircraft observations presented by Monks et al. (2014).

5.2 Surface network ethane and propane

The NOAA Global Monitoring Division/INSTAAR network of surface sites provide provides weekly observations of light NMHCs around the globe (Helmig et al., 2009). The model results for ethane and propane are compared to the data over a range of latitudes in the Northern Hemisphere in northern mid- to high latitudes in Fig. 11. Monthly mean model output is used and the nearest grid point (longitude, latitude, altitude) selected for each site. All models (except GEOS-Chem, which used higher ethane emissions and has lower OH concentrations) significantly underestimate the winter-spring observations, indicating the POLMIP emissions are much too low for both C_2H_6 and C_3H_8 , consistent with the conclusion that CO emissions are too low (as discussed in Sect. 5.4 and in Monks et al., 2014).

5.3 Evaluation of NO₂

Satellite observations of NO₂ have been used to evaluate the individual model distributions of NO₂ across the Northern Hemisphere, as well as to evaluate the NO_x emissions used for all the models. Each model was compared to OMI DOMINO-v2 NO₂ tropospheric column densities (Boersma et al., 2011), matching the times of overpasses for each day and filtering out the pixels with satellite-observed radiance fraction originating from clouds greater than 50 %. The averaging kernels of the retrievals are In order to make a quantitative comparison between model results and satellite retrievals (of any kind), the sensitivity of the retrievals to the true atmospheric profile must be taken into account. This is done by transforming each model profile with the corresponding retrieval averaging kernel and a priori information (e.g., Eskes and Boersma, 2003), hence making the evaluation independent of the a priori NO₂ profiles used in DOMINO-v2. Using the averaging kernels also The transformation of the model profiles with the averaging kernels gives model levels in the free troposphere relatively greater weight in the column calculation, which. For instance, depending on the surface albedo the sensitivity to the upper free troposphere compared to the surface layer may increase by roughly a factor 3 (Eskes and Boersma, 2003). This means that errors in the shape of the NO₂ profile can contribute to biases in the total column.

While there are differences among the model distributions of The statistics of the biases between the model results and the OMI NO₂, the multi-model mean is tropospheric columns are used to evaluate the NO_x emissions inventory used in this study. Figure 12 shows the NO₂ tropospheric column from OMI with the model bias for each model for mid-June to mid-Julyfor June 18-July 15 (the period that hourly model output was provided) and the median of the model biases for that period. The models generally underestimate NO₂ over continental regions with high levels of anthropogenic pollution $\frac{1}{2}$ (e.g., California, northeastern United States, Europe, China), however a few models overestimate NO₂ over North America (not shown). All models overestimate NO₂ over north east northeast Asia, in the region of fires (quantified below). OMI NO₂ retrievals have low signal to noise ratios over oceans and continental regions with low pollutant levels, therefore conclusions should not be drawn by the model comparisons for those regions. GEOS-Chem is relatively low over Europe and East China compared to other models. GMI shows often quite high levels of , while TOMCAT shows a negative bias over pollution regions. This all indicates a relatively large variation in behavior among models in columns for different regimes.

In Fig.Figure 13 the observations and associated multi-model median bias OMI tropospheric columns are screened on a daily basis for pixels where at least 90% of the total NO_x emissions, based on the emissions inventory, originate from anthropogenic or biomass burning emission, respectively, while only regions emissions, respectively. Only pixels with significant emission levels are shown. In this way, dominating source regions that are either primarily anthropogenic or biomass burning can be identified. Figure 13a and c includes boxes Boxes are drawn around the highest concentrations in Figure 13, and the biases for these regions are summarised in Fig.regional mean for each model, along with the mean observed columns, are summarized in Figure 14. The bias of the model median for anthropogenic sources is on average positive over Northwest Europe, although

this is mostly related to a positive bias of the model median over pixels that contain the North Sea. The bias is negative over the pollution hotspot of Northwest Europe, as well as over many other regions with high population density. model results are lower than the observations for most of the anthropogenic region comparisons. For instance, NO₂ columns over South Korea are considerably underestimated. Also the inter-quartile range is relatively large for the European and East-China Europe and East China regions indicating a large uncertainty introduced by the models. Since all models used the same NO emissions, the large variation among models (as seen in Figure 14) indicates differences in the chemistry and transport processes affecting NO and NO₂. The large region of biomass burning in western Asia (April) is well captured, but over eastern Asia the models are typically too high. Also, the NO₂ from Siberian fires in July are quite is greatly overestimated. The NO₂ column amounts are much lower for the fires in Canada than in Asia, but the models also overestimate the concentrations of this region, suggesting the NO_x emission factor is too high for forests in the FINN emissions.

5.4 Comparison to aircraft observations

For each aircraft campaign, the hourly output from each model was interpolated to the location and time of the flight tracks. These results have been compared directly to the corresponding observations for as many compounds as available. Figure 15 shows the flight tracks of the campaigns and have been colored to indicate the grouping used in the following comparisons. The ARCTAS-A and ARCPAC (A1, A2, AP) campaigns took place in April and were based in Alaska. The A1 group of flights surveyed the Arctic between Alaska and Greenland at the beginning of April, while A2 and AP were primarily over Alaska in mid-April, which was after significant wildfires began in Siberia and influenced the observations (e.g., Warneke et al., 2009). The ARCTAS-CARB flights focused on characterizing urban and agricultural emissions in California, but also sampled the wildfire emissions present in the state. ARCTAS-B, based in central Canada, sampled fresh and aged fire emissions over Canada and into the Arctic. The POLARCAT-France and GRACE experiments, based

in southern Greenland, sampled downwind of anthropogenic and fire emissions regions and included observations of air masses from North America, Asia, as well as Europe.

Figures 16, 17 and 18 show vertical profiles of the observations with model results for the flights during ARCTAS-A1, ARCPAC and ARCTAS-B, respectively. For these plots the observations and the model results along the flight tracks were treated the same in the same way: each group of flights were was binned according to altitude and the median value of that each 1-km bin has been plotted. The thick error bars represent the measurement uncertainty (determined by applying the fractional uncertainty reported in each measurement data file to the median binned value), while the thinner horizontal lines show the variation (25th to 75th percentile) in the observations over the flights. In general the measurement uncertainty is much less than the atmospheric variability, however for ARCPAC, several measurements have relatively large uncertainties (such as SO₂, NO₂ and HNO₃).

To make a more quantitative evaluation of the modelsIn addition, the difference between each model and the observations was determined for each data point along the flight tracks, and then an average bias was determined for the altitude range 3–7 km, as shown in Fig. 19. In the cases where there was a compound was measured by more than one measurement of a compoundinstrument, the differences between the model and each measurement were averaged observation were averaged over all the measurement techniques. The uncertainties shown in Figs. 16–18 need to be kept in mind when considering the biases shown in Fig. 19.

Several models, but not all, under-predict ozone in spring by more than 10%(, consistent with the ozonesonde comparison shown in Fig. 9). All models (except GEOS-chem) underpredict CO and hydrocarbons in spring and summer, likely indicating that the emissions used for POLMIP are too low. NO and NO₂ are generally underestimated in spring, with NO₂ biases ranging from 20 to 10090% too low. In summer, all of the models match well the NO and NO₂ observations in the mid-tropopshere, but NO₂ is generally overestimated in the boundary layer -(ARCTAS-B, Figure 18), consistent with the OMI NO₂ comparisons for the Canada fire regions (Fig. 14).

NO_v partitioning between PAN and HNO₃ is hugely vastly different among the models (see Arnold et al., 2014). Many models significantly overestimate HNO₃ (by a factor of 10 in some cases), which could be primarily due to differences in washout and missing loss processes. A new version of LMDZ-LMDz-INCA includes the uptake of nitric acid on sea salt and dust, accounting for 25% of the total sink of nitric acid (Hauglustaine et al., 2014). GMI includes otherwise unaccounted for nitrogen species in HNO₃, partially explaining its overestimate. The simulated PAN values also vary significantly across models, which may be due to the differences in PAN precursors (NO_x and acetaldehyde) at anthropogenic and fire source regions. Alkyl nitrates were found to be a significant contribution to the NO_v budget of the ARCTAS observations, particularly in low-NO_v environments, and the poor (or lack of) representation of them in the models, could also lead to model errors in NO_v partitioning (Browne et al., 2013). The PAN measurements during ARCPAC are only available for the last half of the campaign, during which numerous fire plumes were sampled that were of too fine a scale too fine to be reproduced in the models, resulting in an apparent underestimate by all of the models in the free troposphere (Fig. 17). The observed PAN values during ARCPAC are significantly higher than the ARCTAS-A1 observations, which were made before the Siberian fire plumes began influencing the Alaskan region.

The models show very different concentrations in various oxygenated VOCs with and very little agreement with observations. Methanol and ethanol are generally underestimated by the models. The models do a poor job of simulating formaldehyde in spring, but are much closer to the observations in summer (during ARCTAS-B and -CARB). In April, acetaldehyde is underestimated by all of the models throughout the troposphere, but with large differences among the models (10–95% biases). In summer the models are more uniformly far below (80–100%) the observations. Acetone is also poorly simulated by the models, with large differences among models in both spring and summer. Acetone in TM5 is particularly low, likely due to excessive dry deposition.

For ARCTAS (Figures 16 and 18), the comparison to OH observations is shown, however, it is unlikely the large model grid boxes can. The distribution of OH is strongly affected by clouds and their impact on photolysis, which coarse-grid models cannot be expected

24

to capture the fine structure of the short-lived OH. These reproduce, however, these differences are likely averaged out in the binned vertical profiles. The average biases indicate that in April, most of the models underestimate OH, particularly in the lower troposphere. The underestimate of ozone by some of the models will also lead to lower OH. In summer, the biases are smaller (see Figure 18). The wide range of results in comparison to H_2O_2 indicate that there is great uncertainty in the simulation of the HO_x budget.

Photolysis rates, calculated from actinic flux measurements on the NASA DC-8, are available for the ARCTAS flights. The photolysis rates $J(O_3 \rightarrow O^1D)$ and $J(NO_2)$ from a few models are compared to the observations in Figs. 16 and 18. MOZART and CAM-Chem, which use the same photolysis parameterization (fully described in Lamarque et al., 2012; Kinnison et al., 2007), agree fairly well with observations, while TOMCAT and SMHI-MATCH generally underestimate the photolysis rates. Some differences are expected due to the difficulty of representing clouds in the models.

6 Enhancement ratios of VOCs in fires

The measurements of numerous compounds and the frequent sampling of air masses influenced by fires wildfires by the DC-8 aircraft during ARCTAS allowed for a derivation of enhancement factors of VOCs relative to CO for several sets of fires, as cataloged and summarized by Hornbrook et al. (2011). In that analysis, Hornbrook et al. (2011) used a variety of parameters to identify fire-influenced air masses, their origin, and age, including acetonitrile and hydrogen cyanide (CH₃CN and HCN, which have primarily biomass burning sources), back trajectories from the aircraft flight tracks, and NMHC ratios (to determine photochemical age). During ARCTAS-B, numerous observations were made of fresh plumes from the fires burning in Saskatchewan, providing good statistics of the enhancement ratios. Since the photochemical age of these sampled plumes was generally less than 2 days, the error introduced due to chemical processing of the plumes is much less than for the older plumes from Asia, for example. The sampling of fresh plumes from the fires in Saskatchewan, with little influence of local anthropogenic sources, make makes this a good period and location for the evaluation of fire emissions in the models.

Due to the coarse resolution of the models, and the targeted sampling by the aircraft flightsalong with the uncertainties in location, vertical distribution and strength of the sources, it is not expected that the models will capture the magnitude or exact location of plumes that were sampled by the aircraft. Therefore, instead of using the model results interpolated to the flight tracks, all of the grid points for each model with CO mixing ratios greater than 150 ppb within the region of the fires (54–58° N, 252–258° E, model levels below between the surface and 850 hPa) were used from the hourly model output output from each model. This model output was used to derive enhancement ratios of VOCs relative to CO, comparable to those derived by Hornbrook et al. (2011) (given in their Table 2 and Fig. 7). Figure 20 shows the enhancement ratios derived from the aircraft measurements, giving the mean and standard deviation of all observed Saskatchewan fire plumes. Also shown in Fig. 20 are the emission factors (EF) determined from the emissions inventory used by the models, averaged over 28 June–5 July and 54–58° N, 252–258° E. For each model, the enhancement ratio was determined as the slope of a linear fit to the correlation of each VOC to CO.

For the VOCs with direct emissions and little or no secondary production (ethane, propane, methanol, ethanol), the VOC/CO ratios of the model mixing ratios are very close to the emission factors of the inventory used by the models. This indicates the chemical processing in the vicinity of the fires is slow enough that the observations are a good indicator of the actual fire emission factors. This also means the model ratios can be quantitatively compared to the observations. Thus, we can conclude for the Saskatchewan fires that the fire emission factors used are too high for ethane, too low for propane, about right for methanol and much too low for ethanol. However, the compounds that have significant chemical production in addition to emissions (i.e., formaldehyde, acetaldehyde and acetone) have very different mixing ratio VOC/CO ratios from the emission ratios. The model enhancement ratios of CH_2O and CH_3CHO are significantly higher than the inventory emission factors due to chemical production, but they agree well with the observations. The

model ratios for acetone, however, are lower than the observations, but not very different from the emission factor, implying the emissions emission factors are too low.

7 Conclusions

Eleven global or regional chemistry models participated in the POLARCAT Model Intercomparison Project (POLMIP), allowing for an assessment of our current understanding of the chemical and transport processes affecting the distributions of ozone and its precursors in the Arctic. To limit the differences among models, a standard emissions inventory was used. All of the models were driven, to at least some degree, by observed meteorology (GEOS-5, NCEP or ECMWF), and therefore represented the dynamics of the study year (2008). However, numerous differences occurred among the model outputs due to different chemical schemes and physical parameterizations such as convection, boundary layer mixing and ventilation, wet and dry deposition. The simultaneous evaluation with observations of reactive nitrogen species and VOCs has illustrated that large differences exist in the model chemical mechanisms, especially in their representation of VOCs and their oxidation. Additional model diagnostics are required to completely understand the differences among models. For example, comparison of the wet deposition rates and fluxes of a number of compounds could be informative in understanding the budgets of , and VOCs.

While the extensive suite of aircraft observations in 2008 at high northern latitudes is extremely valuable for evaluating the models, they cannot uniquely identify the source of model emissions errors, as the Arctic is influenced by many sources at lower latitudes. However, several conclusions can be drawn about the emissions inventory used in this study. Based on the comparisons to aircraft observations and the NOAA surface network data, emissions of CO, ethane and propane are clearly too low. The comparisons to satellite retrievals of OMI NO₂ show a few regions of consistent model errors that indicate that anthropogenic NO_x emissions are underestimated in East Asia, while fire emissions are overestimated in Siberia. Large differences are seen among the model NO₂ tropospheric columns

over Northwest Europe and China, thus limiting the conclusions that can be drawn regarding the accuracy of the emissions inventory. The large range in modelled modeled NO_2 (where NO_x emissions were the same) also indicates that model chemistry and dynamics boundary layer parameterizations can significantly impact NO_x chemistry. More accurate emissions inventories might greatly improve many of the model deficiencies identified in this study. Emissions inventories modified based on inverse modeling results, as well as results of this study, will be used in future work as one step in improving model simulations of the Arctic Arctic atmospheric composition.

The simultaneous evaluation of the models with observations of reactive nitrogen species and VOCs has illustrated that large differences exist in the model chemical mechanisms, especially in their representation of VOCs and their oxidation. Most of the models showed a negative bias in comparison to ozone observations from sondes and aircraft, with a slightly larger difference in April than in summer. The models frequently underestimated ozonesondes ozone in the free troposphere by 10–20 ppb and in the comparison with ozonesondes. In addition, 10–30 % negative model biases were seen in comparison to the mid-troposphere aircraft ozone measurements. Comparisons for ozone precursors such as NO_x , PAN, and OVOCs VOCs show much greater biases and differences among models. It appears numerous factors are the causes of these model differences. The differences among model photolysis rates and cloud distributions indicate some of the possible causes for differences in modelled modeled OH, which led leads to differences in numerous species and ozone production and loss rates.

Some differences among the simulated results are likely due to different physical parameterizations such as convection, boundary layer mixing and ventilation, wet and dry deposition. Additional model diagnostics are required to better understand the differences among models. For example, comparison of the wet deposition rates and fluxes of a number of compounds could be informative in understanding the budgets of NO_y, HO_x and VOCs.

Evaluation of chemical transport models with numerous simultaneous observations, such as those of the POLARCAT aircraft experiments, can assist in a critical assessment of ozone simulations and identify model components in need of improvement. Model representation Discussion Paper

Discussion Paper

Discussion Paper

of the oxidation of VOCs and the NO_y budget can have a significant impact on ozone distributions. Future chemical model comparisons should consider evaluation of VOCs and reactive nitrogen species as an important component of the evaluation of ozone simulations.

Acknowledgements. The numerous individuals who provided observations used in this study are gratefully acknowledged including: William H. Brune, Jingqiu Mao, Xinrong Ren, and David Shelow of Pennsylvania State University for the ARCTAS DC8 LIF OH measurements; Paul Wennberg and John Crounse of California Institute of Technology for the ARCTAS DC8 CIT-CIMS data (supported by NASA Award NNX08AD29G); Steve Montzka of NOAA/ESRL/GMD for NOAA P3 flask samples of propane during ARCPAC; Joost de Gouw of NOAA/ESRL/CSD for ARCPAC PTRMS VOC observations; John Holloway of NOAA/ESRL/CSD for ARCPAC CO and SO₂ (UV fluorescence) measurements.

The GEOS-5 data used with CAM-chem in this study have been provided by the Global Modeling and Assimilation Office (GMAO) at NASA Goddard Space Flight Center. We acknowledge the free use of tropospheric NO₂ column data from the OMI sensor from www.temis.nl. French co-authors acknowledge funding from the French Agence National de Recherche (ANR) CLIMSLIP project and CNRS-LEFE. POLARCAT-France was supported by ANR, CNRS-LEFE and CNES. This work was performed in part using HPC resources from GENCI-IDRIS (Grant 2014-017141). VH acknowledges funding from the European Commission under the Seventh Framework Programme (contract number 218793). Contributions by SMHI were funded by the Swedish Environmental Protection Agency under contract NV-09414-12 and through the Swedish Climate and Clean Air research program, SCAC. AW acknowledges support from BMVIT-FFG/ALR. ARCPAC was supported in part by the NOAA Climate and Health of the Atmosphere programs. JM acknowledges the NOAA Climate Proaram Office's grant NA13OAR4310071. LKE acknowledges support from the National Aeronautics and Space Administration under Award No. NNX08AD22G issued through the Science Mission Directorate, Tropospheric Composition Program. The CESM project is supported by the National Science Foundation and the Office of Science (BER) of the US Department of Energy. The National Center for Atmospheric Research is funded by the National Science Foundation.

Discussion Paper

References

- Andersson, C., Langner, J., and Bergstrom, R.: Interannual variation and trends in air pollution over Europe due to climate variability during 1958-2001 simulated with a regional CTM coupled to the ERA40 reanalysis, Tellus B, 59, 77–98, doi:10.1111/j.1600-0889.2006.00196.x, 2007.
- Arnold, S. R., Methven, J., Evans, M. J., Chipperfield, M. P., Lewis, A. C., Hopkins, J. R., Mc-Quaid, J. B., Watson, N., Purvis, R. M., Lee, J. D., Atlas, E. L., Blake, D. R., and Rappenglueck, B.: Statistical inference of OH concentrations and air mass dilution rates from successive observations of nonmethane hydrocarbons in single air masses, J. Geophys. Res.-Atmos., 112, D10S40, doi:10.1029/2006JD007594.2007.
- Arnold, S. R., Emmons, L. K., Monks, S. A., Law, K. S., Ridley, D. A., Turguety, S., Tilmes, S., Thomas, J. L., Bouarar, I., Flemming, J., Huijnen, V., Mao, J., Duncan, B. N., Steenrod, S., Yoshida, Y., Langner, J., and Long, Y.: Biomass burning influence on high latitude tropospheric ozone and reactive nitrogen in summer 2008: a multi-model analysis based on POLMIP simulations, Atmos. Chem. Phys. Discuss., 14, 24573–24621, doi:10.5194/acpd-14-24573-2014, 2014.
- Barrie, L. A.: Arctic air pollution: an overview of current knowledge, Atmos. Environ., 20, 643-663, doi:10.1016/0004-6981(86)90180-0, 1986.
- Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: model description and evaluation, J. Geophys. Res.-Atmos., 106, 23073-23095, doi:10.1029/2001JD000807, 2001.
- Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO₂ column retrieval algorithm for the Ozone Monitoring Instrument, Atmos. Meas. Tech., 4, 1905–1928, doi:10.5194/amt-4-1905-2011, 2011.
- Brock, C. A., Cozic, J., Bahreini, R., Froyd, K. D., Middlebrook, A. M., McComiskey, A., Brioude, J., Cooper. O. R., Stohl, A., Aikin, K. C., de Gouw, J. A., Fahey, D. W., Ferrare, R. A., Gao, R.-S., Gore, W., Holloway, J. S., Hübler, G., Jefferson, A., Lack, D. A., Lance, S., Moore, R. H., Murphy, D. M., Nenes, A., Novelli, P. C., Nowak, J. B., Ogren, J. A., Peischl, J., Pierce, R. B., Pilewskie, P., Quinn, P. K., Ryerson, T. B., Schmidt, K. S., Schwarz, J. P., Sodemann, H., Spackman, J. R., Stark, H., Thomson, D. S., Thornberry, T., Veres, P., Watts, L. A., Warneke, C., and Wollny, A. G.: Characteristics, sources, and transport of aerosols measured in spring 2008 during

the aerosol, radiation, and cloud processes affecting Arctic Climate (ARCPAC) Project, Atmos. Chem. Phys., 11, 2423–2453, doi:10.5194/acp-11-2423-2011, 2011.

- Browne, E. C., Min, K.-E., Wooldridge, P. J., Apel, E., Blake, D. R., Brune, W. H., Cantrell, C. A., Cubison, M. J., Diskin, G. S., Jimenez, J. L., Weinheimer, A. J., Wennberg, P. O., Wisthaler, A., and Cohen, R. C.: Observations of total RONO₂ over the boreal forest: NO_x sinks and HNO₃ sources, Atmos. Chem. Phys., 13, 4543–4562, doi:10.5194/acp-13-4543-2013, 2013.
- Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B., Duncan, B., Martin, R., Logan, J., Higurashi, A., and Nakajima, T.: Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements, J. Atmos. Sciences, 59, 461–483, doi:10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2, 2002.
- Chipperfield, M. P.: New version of the TOMCAT/SLIMCAT off-line chemical transport model: intercomparison of stratospheric tracer experiments, Q. J. Roy. Meteor. Soc., 132, 1179–1203, doi:10.1256/qj.05.51, 2006.
- de Villiers, R. A., Ancellet, G., Pelon, J., Quennehen, B., Schwarzenboeck, A., Gayet, J. F., and Law, K. S.: Airborne measurements of aerosol optical properties related to early spring transport of mid-latitude sources into the Arctic, Atmos. Chem. Phys., 10, 5011–5030, doi:10.5194/acp-10-5011-2010, 2010.
- DeCaria, A. J., Pickering, K. E., Stenchikov, G. L., and Ott, L. E.: Lightning-generated NO_x and its impact on tropospheric ozone production: a three-dimensional modeling study of a Stratosphere– Troposphere Experiment: Radiation, Aerosols and Ozone (STERAO-A) thunderstorm, J. Geophys. Res.-Atmos., 110, D14303, doi:10.1029/2004JD005556, 2005.
- Duncan, B. N., Strahan, S. E., Yoshida, Y., Steenrod, S. D., and Livesey, N.: Model study of the cross-tropopause transport of biomass burning pollution, Atmos. Chem. Phys., 7, 3713–3736, doi:10.5194/acp-7-3713-2007, 2007.
- Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S. L., and Kloster, S.: Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), Geosci. Model Dev., 3, 43–67, doi:10.5194/gmd-3-43-2010, 2010.
- Engvall, A.-C., Krejci, R., Strom, J., Minikin, A., Treffeisen, R., Stohl, A., and Herber, A.: In-situ airborne observations of the microphysical properties of the Arctic tropospheric aerosol during late spring and summer, Tellus B, 60, 392–404, doi:10.1111/j.1600-0889.2008.00348.x, 2008.

- Fast, J. D., Gustafson, W. I., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E. G., Grell, G. A., and Peckham, S. E.: Evolution of ozone, particulates, and aerosol direct radiative forcing in the vicinity of Houston using a fully coupled meteorology-chemistry-aerosol model, J. Geophys. Res.-Atmos., 111, D21305, doi:10.1029/2005JD006721, 2006.
- Fisher, J. A., Jacob, D. J., Purdy, M. T., Kopacz, M., Le Sager, P., Carouge, C., Holmes, C. D., Yantosca, R. M., Batchelor, R. L., Strong, K., Diskin, G. S., Fuelberg, H. E., Holloway, J. S., Hyer, E. J., McMillan, W. W., Warner, J., Streets, D. G., Zhang, Q., Wang, Y., and Wu, S.: Source attribution and interannual variability of Arctic pollution in spring constrained by aircraft (ARCTAS, ARCPAC) and satellite (AIRS) observations of carbon monoxide, Atmos. Chem. Phys., 10, 977– 996, doi:10.5194/acp-10-977-2010, 2010.
- Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Josse, B., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric chemistry in the integrated forecasting system of ECMWF, Geosci. Model Dev. Discuss., 7, 7733–7803, doi:10.5194/gmdd-7-7733-2014, 2014.
- Folberth, G. A., Hauglustaine, D. A., Lathière, J., and Brocheton, F.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: model description and impact analysis of biogenic hydrocarbons on tropospheric chemistry, Atmos. Chem. Phys., 6, 2273–2319, doi:10.5194/acp-6-2273-2006, 2006.
- Freitas, S. R., Longo, K. M., Chatfield, R., Latham, D., Silva Dias, M. A. F., Andreae, M. O., Prins, E., Santos, J. C., Gielow, R., and Carvalho Jr., J. A.: Including the sub-grid scale plume rise of vegetation fires in low resolution atmospheric transport models, Atmos. Chem. Phys., 7, 3385–3398, doi:10.5194/acp-7-3385-2007, 2007.
- Granier, C., Lamarque, J., Mieville, A., Müller, J., Olivier, J., Orlando, J., Peters, J., Petron, G., Tyndall, G., and Wallens, S.: POET, a database of surface emissions of ozone precursors, available at: http://www.pole-ether.fr/eccad (last accessed 19 Nov 2014), 2005.
- Granier, C., Bessagnet, B., Bond, T., D'Angiola, A., Denier van der Gon, H., Frost, G. J., Heil, A., Kaiser, J. W., Kinne, S., Klimont, Z., Kloster, S., Lamarque, J.-F., Liousse, C., Masui, T., Meleux, F., Mieville, A., Ohara, T., Raut, J.-C., Riahi, K., Schultz, M. G., Smith, S. J., Thompson, A., van Aardenne, J., van der Werf, G. R., and van Vuuren, D. P.: Evolution of anthropogenic and biomass

burning emissions of air pollutants at global and regional scales during the 1980–2010 period, Clim. Change, 109, 163–190, doi:10/1007/s10584-011-0154-1, 2011.

- Grell, G., Peckham, S., Schmitz, R., McKeen, S., Frost, G., Skamarock, W., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmos. Environ., 39, 6957–6975, doi:10.1016/j.atmosenv.2005.04.027, 2005.
- Hauglustaine, D. A., Hourdin, F., Jourdain, L., Filiberti, M.-A., Walters, S., Lamarque, J.-F., and Holland, E. A.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: description and background tropospheric chemistry evaluation, J. Geophys. Res.-Atmos., 109, D04314, doi:10.1029/2003JD003957, 2004.
- Hauglustaine, D. A., Balkanski, Y., and Schulz, M.: A global model simulation of present and future nitrate aerosols and their direct radiative forcing of climate, Atmos. Chem. Phys., 14, 11031–11063, doi:10.5194/acp-14-11031-2014, 2014.
- Helmig, D., Ganzeveld, L., Butler, T., and Oltmans, S. J.: The role of ozone atmosphere-snow gas exchange on polar, boundary-layer tropospheric ozone a review and sensitivity analysis, Atmos. Chem. Phys., 7, 15–30, doi:10.5194/acp-7-15-2007, 2007a.
- Helmig, D., Oltmans, S., Morse, T., and Dibb, J.: What is causing high ozone at Summit, Greenland, Atmos. Environ., 41, 5031–5043, 2007b.
- Helmig, D., J. Bottenheim, I. E. Galbally, A. Lewis, M. J. T. Milton, S. Penkett, C. Plass-Duelmer, S. Reimann, P. Tans, and S. Thiel: Volatile organic compounds in the global atmosphere, EOS T. Am. Geophys. Un., 90, 513–514, 2009.
- Hornbrook, R. S., Blake, D. R., Diskin, G. S., Fried, A., Fuelberg, H. E., Meinardi, S., Mikoviny, T., Richter, D., Sachse, G. W., Vay, S. A., Walega, J., Weibring, P., Weinheimer, A. J., Wiedinmyer, C., Wisthaler, A., Hills, A., Riemer, D. D., and Apel, E. C.: Observations of nonmethane organic compounds during ARCTAS – Part 1: Biomass burning emissions and plume enhancements, Atmos. Chem. Phys., 11, 11103–11130, doi:10.5194/acp-11-11103-2011, 2011.
- Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., Grandpeix, J.-Y., Krinner, G., LeVan, P., Li, Z.-X., and Lott, F.: The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection, Clim. Dynam., 27, 787–813, doi:10.1007/s00382-006-0158-0, 2006.
- Huang, M., Carmichael, G. R., Adhikary, B., Spak, S. N., Kulkarni, S., Cheng, Y. F., Wei, C., Tang, Y., Parrish, D. D., Oltmans, S. J., D'Allura, A., Kaduwela, A., Cai, C., Weinheimer, A. J., Wong, M., Pierce, R. B., Al-Saadi, J. A., Streets, D. G., and Zhang, Q.: Impacts of transported background

ozone on California air quality during the ARCTAS-CARB period – a multi-scale modeling study, Atmos. Chem. Phys., 10, 6947–6968, doi:10.5194/acp-10-6947-2010, 2010.

- Huijnen, V., Williams, J., van Weele, M., van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., de Laat, J., Boersma, F., Bergamaschi, P., van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nédélec, P., and Pätz, H.-W.: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, Geosci. Model Dev., 3, 445–473, doi:10.5194/gmd-3-445-2010, 2010.
- Jacob, D. J., Crawford, J. H., Maring, H., Clarke, A. D., Dibb, J. E., Emmons, L. K., Ferrare, R. A., Hostetler, C. A., Russell, P. B., Singh, H. B., Thompson, A. M., Shaw, G. E., McCauley, E., Pederson, J. R., and Fisher, J. A.: The Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission: design, execution, and first results, Atmos. Chem. Phys., 10, 5191–5212, doi:10.5194/acp-10-5191-2010, 2010.
- Jourdain, L. and Hauglustaine, D.: The global distribution of lightning NO_x simulated on-line in a general circulation model, Phys. Chem. Earth Pt. C, 26, 585–591, doi:10.1016/S1464-1917(01)00051-4, 2001.
- Kinnison, D. E., Brasseur, G. P., Walters, S., Garcia, R. R., Marsh, D. R., Sassi, F., Harvey, V. L., Randall, C. E., Emmons, L., Lamarque, J. F., Hess, P., Orlando, J. J., Tie, X. X., Randel, W., Pan, L. L., Gettelman, A., Granier, C., Diehl, T., Niemeier, U., and Simmons, A. J.: Sensitivity of chemical tracers to meteorological parameters in the MOZART-3 chemical transport model, J. Geophys. Res.-Atmos., 112, D20302, doi:10.1029/2006JD007879, 2007.
- Klonecki, A., Hess, P., Emmons, L., Smith, L., Orlando, J., and Blake, D.: Seasonal changes in the transport of pollutants into the Arctic troposphere-model study, J. Geophys. Res.-Atmos., 108, 8367, doi:10.1029/2002JD002199, 2003.
- Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere–biosphere system, Global Biogeochem. Cy., 19, GB1015, doi:10.1029/2003GB002199, 2005.
- Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, Atmos. Chem. Phys., 10, 7017–7039, doi:10.5194/acp-10-7017-2010, 2010.

- Lamarque, J.-F., Emmons, L. K., Hess, P. G., Kinnison, D. E., Tilmes, S., Vitt, F., Heald, C. L., Holland, E. A., Lauritzen, P. H., Neu, J., Orlando, J. J., Rasch, P. J., and Tyndall, G. K.: CAM-chem: description and evaluation of interactive atmospheric chemistry in the Community Earth System Model, Geosci. Model Dev., 5, 369–411, doi:10.5194/gmd-5-369-2012, 2012.
- Lamarque, J.-F., Shindell, D. T., Josse, B., Young, P. J., Cionni, I., Eyring, V., Bergmann, D., Cameron-Smith, P., Collins, W. J., Doherty, R., Dalsoren, S., Faluvegi, G., Folberth, G., Ghan, S. J., Horowitz, L. W., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Naik, V., Plummer, D., Righi, M., Rumbold, S. T., Schulz, M., Skeie, R. B., Stevenson, D. S., Strode, S., Sudo, K., Szopa, S., Voulgarakis, A., and Zeng, G.: The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics, Geosci. Model Dev., 6, 179–206, doi:10.5194/gmd-6-179-2013, 2013.
- Lana, A., Bell, T. G., Sim, R., Vallina, S. M., Ballabrera-Poy, J., Kettle, A. J., Dachs, J., Bopp, L., Saltzman, E. S., Stefels, J., Johnson, J. E., and Liss, P. S.: An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean, Global Biogeochem. Cy., 25, GB1004, doi:10.1029/2010GB003850, 2011.
- Law, K., Stohl, A., Quinn, P., Brock, C., Burkhart, J., Paris, J., Ancellet, G., Singh, H., Roiger, A., Schlager, H., Dibb, J., Jacob, D., Arnold, S., Pelon, J., and Thomas, J.: Arctic air pollution: new insights from POLARCAT-IPY, B. Am. Meteorol. Soc., doi:10.1175/BAMS-D-13-00017.1, in press, 2014.
- Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A., Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Ekman, A. M. L., Hess, P., Mahowald, N., Collins, W., Iacono, M. J., Bretherton, C. S., Flanner, M. G., and Mitchell, D.: Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5, Geosci. Model Dev., 5, 709–739, doi:10.5194/gmd-5-709-2012, 2012.
- Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J., and Johnson, C. E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, Geosci. Model Dev., 3, 519–551, doi:10.5194/gmd-3-519-2010, 2010.
- Mao, J., Jacob, D. J., Evans, M. J., Olson, J. R., Ren, X., Brune, W. H., Clair, J. M. St., Crounse, J. D., Spencer, K. M., Beaver, M. R., Wennberg, P. O., Cubison, M. J., Jimenez, J. L., Fried, A., Weibring, P., Walega, J. G., Hall, S. R., Weinheimer, A. J., Cohen, R. C., Chen, G., Crawford, J. H., McNaughton, C., Clarke, A. D., Jaeglé, L., Fisher, J. A., Yantosca, R. M., Le Sager, P., and

Carouge, C.: Chemistry of hydrogen oxide radicals (HO_x) in the Arctic troposphere in spring, Atmos. Chem. Phys., 10, 5823–5838, doi:10.5194/acp-10-5823-2010, 2010.

- Mao, J., Fan, S., Jacob, D. J., and Travis, K. R.: Radical loss in the atmosphere from Cu-Fe redox coupling in aerosols, Atmos. Chem. Phys., 13, 509–519, doi:10.5194/acp-13-509-2013, 2013.
- Meijer, E., van Velthoven, P., Brunner, D., Huntrieser, H., and Kelder, H.: Improvement and evaluation of the parameterisation of nitrogen oxide production by lightning, Phys. Chem. Earth Pt. C, 26, 577–583, doi:10.1016/S1464-1917(01)00050-2, 2001.
- Merlaud, A., Van Roozendael, M., Theys, N., Fayt, C., Hermans, C., Quennehen, B., Schwarzenboeck, A., Ancellet, G., Pommier, M., Pelon, J., Burkhart, J., Stohl, A., and De Mazière, M.: Airborne DOAS measurements in Arctic: vertical distributions of aerosol extinction coefficient and NO₂ concentration, Atmos. Chem. Phys., 11, 9219–9236, doi:10.5194/acp-11-9219-2011, 2011.
- Methven, J., Arnold, S. R., O'Connor, F. M., Barjat, H., Dewey, K., Kent, J., and Brough, N.: Estimating photochemically produced ozone throughout a domain using flight data and a Lagrangian model, J. Geophys. Res.-Atmos., 108, 4271, doi:10.1029/2002JD002955, 2003.
- Methven, J., Arnold, S. R., Stohl, A., Evans, M. J., Avery, M., Law, K., Lewis, A. C., Monks, P. S., Parrish, D. D., Reeves, C. E., Schlager, H., Atlas, E., Blake, D. R., Coe, H., Crosier, J., Flocke, F. M., Holloway, J. S., Hopkins, J. R., McQuaid, J., Purvis, R., Rappenglueck, B., Singh, H. B., Watson, N. M., Whalley, L. K., and Williams, P. I.: Establishing Lagrangian connections between observations within air masses crossing the Atlantic during the International Consortium for Atmospheric Research on Transport and Transformation experiment, J. Geophys. Res.-Atmos., 111, D23S62, doi:10.1029/2006JD007540, 2006.
- Monks, S. A.: A model study of chemistry and transport in the Arctic troposphere, Ph.D. thesis, School of Earth and Environment, University of Leeds, Leeds, UK, 2011.
- Monks, S. A., Arnold, S. R., Emmons, L. K., Law, K. S., Turquety, S., Duncan, B. N., Flemming, J., Huijnen, V., Tilmes, S., Langner, J., Mao, J., Long, Y., Thomas, J. L., Steenrod, S. D., Raut, J. C., Wilson, C., Chipperfield, M. P., Schlager, H., and Ancellet, G.: Multi-model study of chemical and physical controls on transport of anthropogenic and biomass burning pollution to the Arctic, Atmos. Chem. Phys. Discuss., 14, 25281–25350, doi:10.5194/acpd-14-25281-2014, 2014.
- Murray, L. T., Jacob, D. J., Logan, J. A., Hudman, R. C., and Koshak, W. J.: Optimized regional and interannual variability of lightning in a global chemical transport model constrained by LIS/OTD satellite data, J. Geophys. Res.-Atmos., 117, D20307, doi:10.1029/2012JD017934, 2012.
- Paris, J.-D., Ciais, P., Nedelec, P., Ramonet, M., Belan, B. D., Arshinov, M. Y., Golitsyn, G. S., Granberg, I., Stohl, A., Cayez, G., Athier, G., Boumard, F., and Cousin, J.-M.: The YAK-AEROSIB

transcontinental aircraft campaigns: new insights on the transport of CO₂, CO and O₃ across Siberia, Tellus B, 60, 551–568, doi:10.1111/j.1600-0889.2008.00369.x, 2008.

- Paris, J.-D., Stohl, A., Nédélec, P., Arshinov, M. Yu., Panchenko, M. V., Shmargunov, V. P., Law, K. S., Belan, B. D., and Ciais, P.: Wildfire smoke in the Siberian Arctic in summer: source characterization and plume evolution from airborne measurements, Atmos. Chem. Phys., 9, 9315–9327, doi:10.5194/acp-9-9315-2009, 2009.
- Parrella, J. P., Jacob, D. J., Liang, Q., Zhang, Y., Mickley, L. J., Miller, B., Evans, M. J., Yang, X., Pyle, J. A., Theys, N., and Van Roozendael, M.: Tropospheric bromine chemistry: implications for present and pre-industrial ozone and mercury, Atmos. Chem. Phys., 12, 6723–6740, doi:10.5194/acp-12-6723-2012, 2012.
- Pfister, G. G., Avise, J., Wiedinmyer, C., Edwards, D. P., Emmons, L. K., Diskin, G. D., Podolske, J., and Wisthaler, A.: CO source contribution analysis for California during ARCTAS-CARB, Atmos. Chem. Phys., 11, 7515–7532, doi:10.5194/acp-11-7515-2011, 2011.
- Pickering, K. E., Wang, Y., Tao, W.-K., Price, C., and Müller, J.-F.: Vertical distributions of lightning NO_x for use in regional and global chemical transport models, J. Geophys. Res.-Atmos., 103, 31203–31216, doi:10.1029/98JD02651, 1998.
- Price, C. and Rind, D.: A simple lightning parameterization for calculating global lightning distributions, J. Geophys. Res.-Atmos., 97, 9919–9933, doi:10.1029/92JD00719, 1992.
- Price, C., Penner, J., and Prather, M.: NO_x from lightning 1. Global distribution based on lightning physics, J. Geophys. Res., 102, 5929–5941, 1997.
- Quennehen, B., Schwarzenboeck, A., Schmale, J., Schneider, J., Sodemann, H., Stohl, A., Ancellet, G., Crumeyrolle, S., and Law, K. S.: Physical and chemical properties of pollution aerosol particles transported from North America to Greenland as measured during the POLARCAT summer campaign, Atmos. Chem. Phys., 11, 10947–10963, doi:10.5194/acp-11-10947-2011, 2011.
- Quennehen, B., Schwarzenboeck, A., Matsuki, A., Burkhart, J. F., Stohl, A., Ancellet, G., and Law, K. S.: Anthropogenic and forest fire pollution aerosol transported to the Arctic: observations from the POLARCAT-France spring campaign, Atmos. Chem. Phys., 12, 6437–6454, doi:10.5194/acp-12-6437-2012, 2012.
- Real, E., Law, K. S., Weinzierl, B., Fiebig, M., Petzold, A., Wild, O., Methven, J., Arnold, S., Stohl, A., Huntrieser, H., Roiger, A., Schlager, H., Stewart, D., Avery, M., Sachse, G., Browell, E., Ferrare, R., and Blake, D.: Processes influencing ozone levels in Alaskan forest fire plumes during long-range transport over the North Atlantic, J. Geophys. Res.-Atmos., 112, D10S41, doi:10.1029/2006JD007576, 2007.

- Richards, N. A. D., Arnold, S. R., Chipperfield, M. P., Miles, G., Rap, A., Siddans, R., Monks, S. A., and Hollaway, M. J.: The Mediterranean summertime ozone maximum: global emission sensitivities and radiative impacts, Atmos. Chem. Phys., 13, 2331–2345, doi:10.5194/acp-13-2331-2013, 2013.
- Ridley, B., Pickering, K., and Dye, J.: Comments on the parameterization of lightning-produced NO in global chemistry-transport models, Atmos. Environ., 39, 6184–6187, 2005.
- Robertson, L., Langner, J., and Engardt, M.: An Eulerian limited-area atmospheric transport model, J. Appl. Meteorol., 38, 190–210, 1999.
- Roiger, A., Schlager, H., Schäfler, A., Huntrieser, H., Scheibe, M., Aufmhoff, H., Cooper, O. R., Sodemann, H., Stohl, A., Burkhart, J., Lazzara, M., Schiller, C., Law, K. S., and Arnold, F.: In-situ observation of Asian pollution transported into the Arctic lowermost stratosphere, Atmos. Chem. Phys., 11, 10975–10994, doi:10.5194/acp-11-10975-2011, 2011.
- Sander, S. P., Friedl R. R., Barker, J. R., Golden, D. M., Kurylo, M. J., Wine, P. H., Abbatt, J. P. D., Burkholder, J. B., Kolb, C. E., Huie, R. E., and Orkin, V. L.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies Evaluation Number 17, JPL Publication 10-6, Jet Propulsion Laboratory, Pasadena, 2011.
- Schmale, J., Schneider, J., Ancellet, G., Quennehen, B., Stohl, A., Sodemann, H., Burkhart, J. F., Hamburger, T., Arnold, S. R., Schwarzenboeck, A., Borrmann, S., and Law, K. S.: Source identification and airborne chemical characterisation of aerosol pollution from long-range transport over Greenland during POLARCAT summer campaign 2008, Atmos. Chem. Phys., 11, 10097–10123, doi:10.5194/acp-11-10097-2011, 2011.
- Sharma, S., Ishizawa, M., Chan, D., Lavoué, D., Andrews, E., Eleftheriadis, K., and Maksyutov, S.: 16 year simulation of Arctic black carbon: transport, source contribution, and sensitivity analysis on deposition, J. Geophys. Res.-Atmos., 118, 943–964, doi:10.1029/2012JD017774, 2013.
- Shindell, D.: Local and remote contributions to Arctic warming, Geophys. Res. Lett., 34, L14704, doi:10.1029/2007GL030221, 2007.
- Shindell, D. and Faluvegi, G.: Climate response to regional radiative forcing during the twentieth century, Nat. Geosci., 2, 294–300, doi:10.1038/NGEO473, 2009.
- Shindell, D. T., Chin, M., Dentener, F., Doherty, R. M., Faluvegi, G., Fiore, A. M., Hess, P., Koch, D. M., MacKenzie, I. A., Sanderson, M. G., Schultz, M. G., Schulz, M., Stevenson, D. S., Teich, H., Textor, C., Wild, O., Bergmann, D. J., Bey, I., Bian, H., Cuvelier, C., Duncan, B. N., Folberth, G., Horowitz, L. W., Jonson, J., Kaminski, J. W., Marmer, E., Park, R., Pringle, K. J., Schroeder, S., Szopa, S., Takemura, T., Zeng, G., Keating, T. J., and Zuber, A.: A multi-model assessment of

pollution transport to the Arctic, Atmos. Chem. Phys., 8, 5353–5372, doi:10.5194/acp-8-5353-2008, 2008.

- Simpson, D.: Long-period modelling of photochemical oxidants in Europe: model calculations for July 1985, Atmos. Environ., 26A, 1609–1634, 1992.
- Sodemann, H., Pommier, M., Arnold, S. R., Monks, S. A., Stebel, K., Burkhart, J. F., Hair, J. W., Diskin, G. S., Clerbaux, C., Coheur, P.-F., Hurtmans, D., Schlager, H., Blechschmidt, A.-M., Kristjánsson, J. E., and Stohl, A.: Episodes of cross-polar transport in the Arctic troposphere during July 2008 as seen from models, satellite, and aircraft observations, Atmos. Chem. Phys., 11, 3631–3651, doi:10.5194/acp-11-3631-2011, 2011.
- Stockwell, D., Giannakopoulos, C., Plantevin, P.-H., Carver, G., Chipperfield, M., Law, K., Pyle, J., Shallcross, D., and Wang, K.-Y.: Modelling NO_x from lightning and its impact on global chemical fields, Atmos. Environ., 33, 4477–4493, doi:10.1016/S1352-2310(99)00190-9, 1999.
- Stohl, A.: Characteristics of atmospheric transport into the Arctic troposphere, J. Geophys. Res., 111, D11306, doi:10.1029/2005JD006888, 2006.
- Strahan, S. E., Duncan, B. N., and Hoor, P.: Observationally derived transport diagnostics for the lowermost stratosphere and their application to the GMI chemistry and transport model, Atmos. Chem. Phys., 7, 2435–2445, doi:10.5194/acp-7-2435-2007, 2007.
- Tarasick, D. W., Jin, J. J., Fioletov, V. E., Liu, G., Thompson, A. M., Oltmans, S. J., Liu, J., Sioris, C. E., Liu, X., Cooper, O. R., Dann, T., and Thouret, V.: High-resolution tropospheric ozone fields for INTEX and ARCTAS from IONS ozonesondes, J. Geophys. Res.-Atmos., 115, D20301, doi:10.1029/2009JD012918, 2010.
- Thomas, J. L., Raut, J.-C., Law, K. S., Marelle, L., Ancellet, G., Ravetta, F., Fast, J. D., Pfister, G., Emmons, L. K., Diskin, G. S., Weinheimer, A., Roiger, A., and Schlager, H.: Pollution transport from North America to Greenland during summer 2008, Atmos. Chem. Phys., 13, 3825–3848, doi:10.5194/acp-13-3825-2013, 2013.
- Thompson, A. M., Oltmans, S. J., Tarasick, D. W., von der Gathen, P., Smit, H. G. J., and Witte, J. C.: Strategic ozone sounding networks: review of design and accomplishments, Atmos. Environ., 45, 2145–2163, doi:10.1016/j.atmosenv.2010.05.002, 2011.
- Tilmes, S., Emmons, L. K., Law, K. S., Ancellet, G., Schlager, H., Paris, J.-D., Fuelberg, H. E., Streets, D. G., Wiedinmyer, C., Diskin, G. S., Kondo, Y., Holloway, J., Schwarz, J. P., Spackman, J. R., Campos, T., Nédélec, P., and Panchenko, M. V.: Source contributions to Northern Hemisphere CO and black carbon during spring and summer 2008 from POLARCAT and
START08/preHIPPO observations and MOZART-4, Atmos. Chem. Phys. Discuss., 11, 5935–5983, doi:10.5194/acpd-11-5935-2011, 2011.

- Tilmes, S., Lamarque, J.-F., Emmons, L. K., Kinnison, D. E., Ma, P.-L., Liu, X., Ghan, S., Bardeen, C., Arnold, S., Deeter, M., Vitt, F., Ryerson, T., Aikin, K., Elkins, J., Moore, F., Gaoand Spackman, R., Perring, A., Schwarz, J., and Wofsy, S.: Description and evaluation of tropospheric chemistry and aerosols in the Community Earth System Model (CESM1.2), Geosci. Model Dev Discuss., to be submitted, 7, 8875–8940, doi:10.5194/gmdd-7-8875-2014, 2014.
- van der A, R. J., Allaart, M. A. F., and Eskes, H. J.: Multi sensor reanalysis of total ozone, Atmos. Chem. Phys., 10, 11277–11294, doi:10.5194/acp-10-11277-2010, 2010.
- Warneke, C., Bahreini, R., Brioude, J., Brock, C. A., de Gouw, J. A., Fahey, D. W., Froyd, K. D., Holloway, J. S., Middlebrook, A., Miller, L., Montzka, S., Murphy, D. M., Peischl, J., Ryerson, T. B., Schwarz, J. P., Spackman, J. R., and Veres, P.: Biomass burning in Siberia and Kazakhstan as an important source for haze over the Alaskan Arctic in April 2008, Geophys. Res. Lett., 36, L02813, doi:10.1029/2008GL036194, 2009.
- Wespes, C., Emmons, L., Edwards, D. P., Hannigan, J., Hurtmans, D., Saunois, M., Coheur, P.-F., Clerbaux, C., Coffey, M. T., Batchelor, R. L., Lindenmaier, R., Strong, K., Weinheimer, A. J., Nowak, J. B., Ryerson, T. B., Crounse, J. D., and Wennberg, P. O.: Analysis of ozone and nitric acid in spring and summer Arctic pollution using aircraft, ground-based, satellite observations and MOZART-4 model: source attribution and partitioning, Atmos. Chem. Phys., 12, 237–259, doi:10.5194/acp-12-237-2012, 2012.
- Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., and Soja, A. J.: The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning, Geosci. Model Dev., 4, 625–641, doi:10.5194/gmd-4-625-2011, 2011.
- Williams, J. E., Strunk, A., Huijnen, V., and van Weele, M.: The application of the Modified Band Approach for the calculation of on-line photodissociation rate constants in TM5: implications for oxidative capacity, Geosci. Model Dev., 5, 15–35, doi:10.5194/gmd-5-15-2012, 2012.
- Williams, J. E., van Velthoven, P. F. J., and Brenninkmeijer, C. A. M.: Quantifying the uncertainty in simulating global tropospheric composition due to the variability in global emission estimates of Biogenic Volatile Organic Compounds, Atmos. Chem. Phys., 13, 2857–2891, doi:10.5194/acp-13-2857-2013, 2013.

- Wong, J., Barth, M. C., and Noone, D.: Evaluating a lightning parameterization based on cloudtop height for mesoscale numerical model simulations, Geosci. Model Dev., 6, 429–443, doi:10.5194/gmd-6-429-2013, 2013.
- Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z., Park, I. S., Reddy, S., Fu, J. S., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.: Asian emissions in 2006 for the NASA INTEX-B mission, Atmos. Chem. Phys., 9, 5131–5153, doi:10.5194/acp-9-5131-2009, 2009.

H	-
2	~
- 27	n
č	5
È	
E	-
2	5
Ĕ	Ξ.
C	5
Ē	Ę.
-	-
H	-
1	V
2	2
FC	5
à	5
Ĥ	÷

Species	anthro. $T_{\sigma} vr^{-1}$	$\frac{bb-fires}{T\sigma vr^{-1}}$	biogenic Toyr ⁻¹	soil To vr ⁻¹	ocean Toyr ⁻¹	volcano $T\sigma vr^{-1}$	Total Total
	1831	1871	167	18,1	1871	187	1831
CO	591.95	329.7	76.15	0	19.9	0	1017.7
NO	69.88	5.2	0	10.58	0	0	85.7
NO ₂	0	11.32	0	0	0	0	11.3
C ₂ H ₂	2.12	0.39	0	0	0	0	2.5
C_2H_6	6.31	1.66	0.14	0	0.98	0	9.1
C_2H_4	6.77	2.82	16.61	0	1.4	0	27.6
C ₃ H ₈	5.64	0.37	0.02	0	1.29	0	7.3
C ₃ H ₆	3.02	1.56	6.06	0	1.52	0	12.2
BIGALK Alkanes (>C3)	51.22	0.74	0	0	0	0	52
BIGENE Alkenes (>C ₃)	6.47	1.83	0	0	0	0	8.3
TOLUENE-Lumped Aromatics	25.2	10.6	0.25	0	0	0	36.1
ISOP Isoprene	0	0.79	522.99	0	0	0	523.8
$C_{10}H_{16}$	0	0.27	96.57	0	0	0	96.8
Methanol	0.92	5.35	158.99	0	0	0	165.3
Ethanol	5.23	0.04	0	0	0	0	5.3
CH ₂ O	2.97	4.11	4.01	0	0	0	11.1
CH₃CHO	1.99	4.53	11.14	0	0	0	17.6
Acetone	0.53	1.85	28.42	0	0	0	30.8
MEK	2.14	4.65	0.53	0	0	0	7.3
HCOOH	6.63	1.67	0	0	0	0	8.3
CH₃COOH	6.63	7.69	0	0	0	0	14.3
BIGALD Other Aldehydes	0	0.08	0	0	0	0	0.1
CH ₃ COCHO	0	1.9	0	0	0	0	1.9
CRESOL Cresol	0	2.28	0	0	0	0	2.3
GLYALD Glycoaldehyde	0	3.81	0	0	0	0	3.8
HYAC Hydroxyacetone	0	3.88	0	0	0	0	3.9
MACR Methacrolein	0	0.21	0	0	0	0	0.2
MVK	0	0.55	0	0	0	0	0.6
SO ₂	124.21	2.26	0	0	0	9.57	136
NH ₃	41.84	4.33	0	2.34	8.1	0	56.6
BC	5.2	1.92	0	0	0	0	7.1
OC	10.57	20.78	0	0	0	0	31.4
HCN	1.71	1.37	0	0	0	0	3.1
CH ₃ CN	0.87	1.04	0	0	0	0	1.9

Table 1. Emissions provided for POLARCAT model intercomparison.

Table 2. Emissions Global emissions actually used in each model (Tg yr⁻¹, except lighting: Tg N yr⁻¹).

Species	CAM4-chem	CAM5-chem	GEOS-Chem	GMI-GEOS5	MOZART-4	TM5	TOMCAT
CO	1018	1018	908	1062	1019	1018	1020
NO	85	85	85	85	85	93	
NO_2	11	11		11	11		142
C_2H_6	9	9	10	9	9	9	9
C ₃ H ₈	7	7	14	7	7	7	7
CH ₂ O	11	11	5	11	11	11	11
CH₃CHO	17	17	2	17	17	20	17
acetone Acetone	30	30	32		30	30	30
methanol Methanol	165	165		166	165	155	165
isoprene Isoprene	524	524	499	523	524	523	530
Lightning NO	4.6	5.0	6	6.6	6.5	6.8	3.8

Notes: SMHI-MATCH included acetone and C_2H_2 emissions as ethane; GEOS-Chem specifies methanol concentrations and has used slightly different anthropogenic emissions; GMI specifies acetone concentrations; TOMCAT reads NO₂ emissions into the NO_x family tracer, which is then split into NO and NO₂; The files provided for GMI and LMDz-INCA did not include emissions; The regional model totals are not included as global values cannot be provided.

Discussion Paper

Table 3. Summary of POLMIP models.

Model	Resolution	Meteorology	Chemistry
CAM4-chem	$1.9^{\circ} \times 2.5^{\circ}$, 56 levels	GEOS-5	MOZART-4, bulk aerosols
CAM5-chem	$1.9^{\circ} imes 2.5^{\circ}$, 56 levels	GEOS-5	MOZART-4, modal aerosols
C-IFS	$1.125^{\circ} imes 1.125^{\circ}$, 60 levels	ECMWF	tropospheric, CB05
GEOS-Chem	$2^{\circ} \times 2.5^{\circ}$, 47 levels	GEOS-5	tropospheric, 100 species -
GMI-GEOS5	$2^{\circ} \times 2.5^{\circ}$, 72 levels	GEOS-5	stratospheric and tropospheric, 154
			species, GOCART aerosols
LMDZ-INCA LMDz-INCA	$1.9^\circ imes 3.75^\circ$, 39 levels	ERA-Interim	tropospheric, 85 species, aerosols
MOZART-4	$1.9^{\circ} imes 2.5^{\circ}$, 56 levels	GEOS-5	tropospheric, 103 species, bulk aer
TM5	$2^{\circ} \times 3^{\circ}$, 60 levels	ECMWF	tropospheric, CB05
TOMCAT	$2.8^{\circ} \times 2.8^{\circ}$, 31 levels	ERA-Interim	tropospheric, 82 species
SMHI-MATCH	$0.75^{\circ} imes 0.75^{\circ},$ 35 levels, NH	ERA-Interim	tropospheric, 61 species
WRF-Chem	100 and 50 km, 65 levels, Canada	WRF/NCEP FNL	MOZART, GOCART aerosols



Figure 1. CO emissions used for the fixed lifetime tracers. (top) Map of anthropogenic (blue boxes) and fire emissions (yellow-red contours) for April and July monthly averages. (bottom) Time series of daily fire and anthropogenic emissions averaged over each tracer region. Anthropogenic emissions have no temporal variation.



Figure 2. Zonal averages of temperature and water vapour vapor from each of the models, for April and July.

Cloud Fraction - Apr CAM4-cl CAM5-c 0 Latitude C-IFS GMI-GEOS5 0 Latitude 0 Latitude MOZART-4 TM 0 Latitude 0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 Cloud Fraction - Jul CAM CAM5-chen Latitude GMI-GEOS5 0 Latitude C-IFS 0 Latitude 0 Latitude MOZART-4 TMS 0 Latitude 0 Latitude 0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60

Figure 3. Zonal average of cloud fraction for April (top) and July (bottom).





Figure 4. Zonal averages of photolysis rates for each model for April of $J(O_3 \rightarrow O^1D)$ (top) and $J(NO_2)$ (bottom).



Figure 5. Location of the 50 and 100 ppb contours of O_3 for the zonal averages of each model in April and July.



Figure 6. Location of the 1.0 and 2.0×10^6 molecules cm⁻³ contours of OH for the zonal averages of each model in April and July.



Figure 7. Seasonal variation of zonal averages for various compounds at 700 hPa, averaged over $50-70^{\circ}$ N latitude band.



Figure 8. Location of ARCIONS ozonesonde sites used in April and June–July 2008 in coordination with ARCTAS.



Figure 9. Comparison of models to ozonesondes for April, showing <u>each sonde mean and standard</u> deviation of the observations (black <u>lines</u>line) and the mean bias (<u>coloured colored</u> lines) for each model at each site (Tarasick et al., 2010; Thompson et al., 2011). <u>Sonde data are truncated at 110</u>Results shown for only surface to 300 hPa for clarity. The number of sondes for each site is indicated in the lower right corner of each panel.



Figure 10. As Fig. 9, but for June–July.



Figure 11. Ethane (top 6 panels) and propane (lower 6 panels) at several Northern Hemisphere NOAA GMD network sites. Monthly mean model output (<u>coloured colored lines</u>) is plotted with 2008 weekly observations (black circles). Station codes: ALT: Alert, Canada; SUM: Summit, Greenland; BRW: Barrow, Alaska; MHD: Mace Head, Ireland; SHM: Shemya, Alaska; LEF: Wisconsin.



Figure 12. a) OMI tropospheric column NO_2 and b) median of the model biases, both for 18 June–15 July.



Figure 13. OMI NO₂ (a) and multi-model mean bias (b) filtered for dominant anthropogenic emissions (a, b) and fire emissions (c, d). Boxes shown in (a) indicate the regions for which biases have been calculated in Fig. 14. (c, d) As in (a, b) but for fire emissions Panels a, c show April, panels b, d show June-July.



Figure 14. Summary of the regional means from each model mean and the OMI NO₂ filtered tropospheric columns for pixels dominated by anthropogenic each region indicated in Figure 13. (a) Anthropogenic emissions in April (a) and June–July (b) or June–July, (c) biomass burning in both seasons(c). Red stars circles are mean OMI NO₂ observations for the region; box plots show median, 25th and 75th quartiles, whiskers to 5th and 95th percentiles of the model means. (See text and Fig. 13.)

58

Discussion Paper



Figure 15. Location of aircraft flight tracks. ARCTAS-A1: 4–9 April; ARCTAS-A2: 12–17 April; ARC-PAC: 11–21 April; ARCTAS-CARB: 18–24 June; ARCTAS-B: 29 June–10 July; GRACE: 30 June–18 July; France: 30 June–14 July.



Figure 16. Comparison of model results interpolated to flight tracks to observations during the first four DC-8 flights of ARCTAS-A. Each profile is the median of 1 km altitude bins. Medians of observations in each 1 km altitude bin are shown as black circles and crosses (if a second measurement exists for a given parameter). Thick error bars represent the measurement uncertainty; thin error bars span inter-quartile range of all observations.

Discussion Paper



Figure 17. As Fig. 16, for ARCPAC NOAA P3 observations.



Figure 18. As Fig. 16, for ARCTAS-B DC-8 observations.



Figure 19. Mean bias (%) between models and aircraft observations, averaged over 3–7 km. Bias for multiple measurements of a single compound are also averaged -over each campaign, which are A1: ARCTAS-A 4–9 April; A2: ARCTAS-A 12–17 April; AP: ARCPAC 11–21 April; C: ARCTAS-CARB 18–24 June; B: ARCTAS-B 29 June–10 July.

Discussion Paper

Discussion Paper



Figure 20. Correlations of VOCs to CO for the POLMIP models, compared to those derived from the DC-8 observations for the fires in Saskatchewan. The filled circle shows the Enhancement Ratio enhancement ratio derived from DC-8 observations (Hornbrook et al., 2011). The asterisk shows the emission factor (EF) of the model emissions. The colored diamonds are the enhancement ratios determined for each model.

64