Overall Response – big issues

Two big issues with this long and complex paper have been raised by the reviews: (1) what is the science here?, and (2) can you be more objective in comparing the probability distributions. Both points are valid. Thus, we returned to the paper to revise the conclusions to summarize what we learned from the multi-model comparison using our new diagnostics. Second we added a Supplement giving reduced/abbreviated statistics for the many plots in the main paper. We compared both standard deviation and equivalent percentiles, and developed a fitted ellipse approach to summarize the 2D plots. All of this has greatly improved the manuscript. These revisions took almost three weeks. We ask for a reasonably prompt response from the editor and reviewers, since this manuscript was submitted on December 7th, was clearly lost in the holiday season, and thus has spent nearly 5 months in review, before any revisions.

Response to Anonymous Referee #1

> General comments. The manuscript presents an interesting method to test chemistry-climate models (CCMs) based on aircraft measurements. The method is motivated by the limited representativeness of aircraft measurements for the large grid cells of CCMs. While the proposed method can contribute to the evaluation of CCMs, alternative methods should be considered as well. For example, CCMs can assimilate meteorological analyses (model nudging) to realistically represent atmospheric conditions, so that the output can be directly compared to measurement data. And even though the grid size of models is course compared to aircraft measurements, some of the sub-grid scale variability can be accounted for by interpolating model results across the flight tracks.

The authors thank referee #1 for the prompt, thoughtful, and constructive comments (RC1). Many of these are technical and are corrected below as noted, but some are stylistic and the authors choose to retain their own style. This manuscript contains enough new material and approaches in developing chemical climatologies, and we cannot really evaluate too much more. The goal of a chemical climatology and its use for CCMs is already discussed and highlighted.

> -The manuscript includes chemistry-transport models (CTMs) for which it should be unproblematic to represent realistic meteorology. Therefore, the method is particularly useful for CCM groups who have not implemented nudging techniques. Please mention this. Nevertheless, the new method is a practical and interesting addition to the techniques used to evaluate CCMs and CTMs. A more extensive discussion of pros and cons of different methods is recommended.

Lines 100-114 directly address this issue. Nudged CCMs or CCMs running with 'specified dynamics' are truly different models than the free-running climate versions. The nudging is an acceleration term that creates different residual tracer circulations and even water cycles. We do not see a need to expand this 15-line explanation.

> -The manuscript generally reads well, but the abstract and introduction need improvement. The introduction presents confusing statements, and the reader is guessing what the work actually focuses on up to L123. Please consider a more traditional structure of the introduction and present the concepts of representativeness (measurements vs model output) and chemical reactivity early on. I found the text partly confusing and unclear, often not to the point and sometimes not relevant. Examples and details are given below.

Yes, the abstract and introduction have been edited to reflect many of the referee's detailed comments, and this greatly improves the presentation. There appears to be a stylistic issue with this referee and even a misunderstanding of the importance and timing of the discussion of representativeness. Representativeness comes later in the discussion, first we need to address heterogeneity and parcel sampling and how to build up statistics. Then we need to discuss reactivity weighting of the parcels. Figure 1 is needed to demonstrate the filamentary nature of reactivity. Representativeness applies first to the measurements to ensure that we have a reproducible climatology – see Figure 8 using models as an example for the sampling.

Overall we have gone through the introduction and tightened it based on the referee comments below and the apparent misunderstanding.

> -Provided that the presentation, notably of the abstract and introduction, are improved, and the application of the method is described in the context of alternative methods, the manuscript will be acceptable for publication in ACP. However, since a methodology rather than new science is presented, it may also be considered to refer this manuscript to GMD, or add to the title that it is a technical note.

We disagree with the referee's view of this manuscript: we believe it does provide a totally new model comparison across 6 major CCM/CTMs; and it does identify unique patterns of differences. In term of alternative methods, the referee is not clear, but we have added a Supplement that tabulates all the standard statistics for the plotted probability densities. We hope this helps. In addition, the introduction (lines 146 & 167) have been revised to point to what we found with the multi-model comparison.

> -Sect. 5 seems to suggest that the NASA ATom missions are unique in providing high-frequency measurements across the Pacific and Atlantic Oceans. Such measurements have been made since decades, which is a great asset to the atmospheric chemistry community and in particular modelers (probably underused), and it should be mentioned that these datasets can also be applied in the presented methodology. The focus on the NASA ATom missions in the manuscript are distinctive, which does not do justice to the more general applications of the method.

While we note some results from the incredibly valuable MOZAIC/IAGOS aircraft measurements and reference these papers, we failed to emphasize their importance in providing an unbiased climatology for O_3 and CO at cruise levels and for profiles over airports. We also refer to the early exploratory missions that were also fairly objective searches of atmospheric composition. Starting at line 136 we have augmented the Introduction to add this discussion. The major extension with ATom is the near-complete chemical package and the regular near

global profiling. If the referee knows of other objective sampling chemical climatologies, please let us know and we will include in the revised discussion.

> -The ending of the paper is unsatisfactory, with a promise of interesting things to come. Is it possible to present a sneak preview? This could make the manuscript more scientifically interesting rather than only technical, and help justify that it is published in ACP.

On this we agree with the referee and have strengthened the conclusion section, including new information and findings about the tropospheric chemistry of the 6 models, while leaving in a mention of the opportunities with Atom.

> Specific comments. The title is not very informative about the paper contents. Please reconsider.

We reconsidered, and believe the title is sound and accurately describes what is new here. There are many possible titles one could come up with, and this is a stylistic choice of the authors.

> L27: Few tenths of km?

Thanks, done.

> L29-34 present technical details, before the reader knows what the work is about. Better first explain the objectives. And/or consider deleting these sentences.

Good point. We have inserted a phrase at Line 29 to explain the objectives of the research. The other sentences are needed to flesh out the objectives.

> L45: Understanding chemical heterogeneity is important, but not a sole prerequisite to understand global atmospheric chemistry. Please reformulate.

The text as is makes only the point that heterogeneity is one aspect that must be understood, if one wants to understand global tropospheric chemistry. We did not say it, and it is certainly NOT the sole issue, we agree there. No change is needed, as style, putting this first is the authors' choice.

> L48: Depends on where you are and what you are looking at. Please reformulate.

This statement is basically correct – it refers to the chemical mix from the previous sentence. It is hard to imagine a place with truly isolated air parcels for long periods of time: there are always gradients at some levels and mixing across shear zones. The only example of quasi-seasonal isolation occurs in the summer stratosphere, and that does not last through the year. This seems to be a stylistic issue, not changed.

> L52: "we ask which air is more important for the chemical evolution of human-driven air pollution". More important in view of what?

Thanks, revised to explicitly say chemical evolution of the global tropospheric pollution. (instead of pollution in general). Methane and ozone are introduced later, they do not fit here.

L55-56: What do you mean?

We believe that the sentence can be understood simply as is: If we know the chemical mix of all the air parcels, then we can use fingerprinting or trajectories or other methods to identify the level of human influence. Have added "and where they come from".

L57: Do you aim to understand the impact of heterogeneity?

Yes, we do. We will look at how the high/low ends of the statistical distribution affect the net rates. For example are they more important than their frequency of occurrence. This is a non-linear system, especially when one considers the co-variance of other key species. Can we just average all he species and get the same rates?

L58: Thus, the integration of chemical reactivity controls the residence time of ozone? Please reformulate.

Yes, this was cryptic. We have added “atmosphere’s” to describe the integration. It is not our integration but the atmosphere’s integration. We also changed the “residence time” to “evolution”.

L67-68: see Chatfield and Delany, 1990.

Thanks, yes, a good choice. We have added the reference to the non-linear discussion here.

L71: It may be helpful to define “heterogeneity” more precisely. Related to sub-grid scale processes? Dependence on grid size and compound considered? First you suggest that we simplify heterogeneity in view of modeling (L63-66) and now you use a model result to illustrate heterogeneity.

We think this is clear. We do NOT suggest that we simplify heterogeneity, but rather than many modeling efforts do so.

L76: Atmospheric rivers are narrower and related to tropical-extra-tropical interactions.

Atmospheric rivers in a column sense may not be narrower. The text has been edited to clarify that atmos-rivers are not LCH4-rivers. And indeed , ti makes it even more interesting to discover what causes the LCH4-rivers.

L78-79: Why/how does this indicate that chemical heterogeneity plays a role?

Obviously the chemical composition across the rivers of LCH4 is not homogeneous. Clearly there are some interesting heterogeneous patterns out there, at the reactivity level.

L79: Figure 1 seems trivial and could be omitted.

Sorry Figure 1 is distinct and shows an aspect of atmospheric chemistry that remains largely unnoted and unpublished. While atmospheric rivers of water have many references, we could not find any to these rivers of chemical reactivity.

L84: Depends on scale.

Thanks, yes, have replaced 'ways' with "patterns" which includes the space-time scales of emissions.

> L110-113: Nudging to meteorological analyses is typically applied to avoid this problem.
>
> L114-119: Please also consider comparing CTM or nudged-CCM output at the location and time of the measurements. See S4D routine of Jöckel et al., 2010 (http://www.geosci-model-dev.net/3/717/2010/).

See earlier comments. A free-running and a nudged CCM are entirely different models. Convection, clouds and everything changes. There are usually hidden residual transports with such nudging, sicne one has to correct the model's trends with winds and rain to the observations. The reference Joeckel et al work is one way to assimilate met data, but that makes it a different model. Moreover, the models that nudge/assimilate/forecast meteorology never seem to agree anyway.

> L123, L128: Now the central issue becomes clear: representativeness. It would be helpful to present this earlier in the abstract and introduction.

See earlier comments. Representativeness is but one of the problems. If there is sufficient sampling (e.g., from a model) then R is not a problem. With limited observations we must question R. All the ground work must be laid before we discuss R.

> L189: Please explain why.

Yes, good point. We have added the explanatory sentence at line 191. " L-CH4 is highly variable across parcels, and the integral of L-CH4 determines the atmospheric lifetime of CH_4 and the buildup of its emissions in the atmosphere."

Response to Anonymous Referee #2

> *General comments The manuscript presents an interesting new approach for the comparison of observed atmospheric in-situ data alongside an aircraft flight track and results from chemistry-climate models. Such a comparison has the intrinsic problem that the observational data are related to specific conditions of time and geographical location which cannot be reproduced adequately within a global model due to their coarse grid and difficulties to represent real transport, emission, and deposition processes. This is related to the question of representativeness of the observations and the full variability of atmospheric constituents.*

We concur and these aspects are discussed in the paper.

> *The manuscript tries to add a new approach to overcome these problems by including more of the climatology of the model data and by shifting the comparison of single constituents towards the comparison of relationships between constituents including an additional weighting by processes like methane oxidation and ozone production. My feeling here is that both approaches (together or each of them separately) are very promising but the final concept how to do the comparison between observations and models is missing.*

We agree in part, but as the reviewer notes, we have begun pushing new types of promising comparisons that may yet prove fruitful. Even with my best optimism, I could never presume that we could come up with a 'final concept.'

Nevertheless, the point is taken and we have spent considerable time developing a large Supplement to the manuscript that summarizes the plots in simpler statistics (percentiles, means, standard deviations, even 2D fitted ellipses), which allow for more ready comparisons. The model comparisons and case studies shown have not changed and thus the basic conclusions remain, but the more objective statistics as suggested by the referee have made the distinctions between models more objective. Thanks. Changes to the text occur throughout as we link to the Supplement, and Fig 8 has added a plot of the fitted ellipses to the transect examples. See further discussion in next comment.

> *Since the title, abstract, and introduction of the manuscript builds up a certain expectation about the presentation of a new and better tool for this kind of comparison I did feel quite disappointed at the end. The manuscript explains how to generate relationships between constituents and the additional weighting processes, both for the observations and for the model data, but the comparisons are then done by visual inspection and their subjective interpretation. Since I myself at some occasions would come to different interpretations looking at the presented figures of relationships between constituents, both weighted or unweighted, it is clear that this new approach is not facilitating the development of objective arguments, at least at the moment.*

This comment is true and the community has been wrestling with comparing complex structures of chemical constituents (e.g., color contours) and then decreeing that "B looks pretty much like A" or the counterpart. We did not have an easy go-to method, but took on the task of converting the figures into fewer statistics in the Supplement described above. These are now introduced in the paper and for most figures and at least represent a first step. For 1D PDs we tabulate both the percentiles (16%-50%-84%) and compare with the mean plus or minus one standard deviation. All of these statistics are calculated in $\log_{10}$ ([X]) where [X] is the mole fraction in ppt. The bin width is set to 0.05. This is the correct choice as we find most PDs to be log normal in that the percentiles match the standard deviation in $\log_{10}$ space. For 2D PDs we could not easily find a percentile approach and so stick to the Gaussian statistics. The diagnostics here are a bit more exotic: the mean centroid and a single 2D combined variance (sigma-x^2 + sigma-y^2) are obvious; next we rotate the data about the centroid (conserves the 2D variance) until the ratio sigma-y/sigma-x is at a minimum (i.e., the flattest ellipse); and we record the angle of rotation and the minimum ratio. This now gives us just 5 numbers to describe the 2D PDs; it also allows us to clearly quantify tight quasi-linear relationships; and it opens up methods to calculate an overlap fraction of the ellipses between two PDs. A new didactic figure of how this works is supplied as Fig S1.

> *Detailed comments: As I see it, the new approach in the manuscript contains three different parts. The first one shows how to compare between 3-D model results and observed data. The traditional way would be to compare observed data obtained at a specific time and location to model results from matching grid boxes for the same time of*

day and year. The new way is to compare model data from a broader region and time so that the exact match to the location and time of the observation is no intrinsic requirement any longer. This broadens the view because the model might be able to reproduce observed phenomena not at the exact coordinates but at a slightly different time or location. The key point here is: are the model climatology of a species and the unbiased climatology of observations similar or not. The second part demonstrates how to calculate important chemical reaction pathways with respect to methane loss and ozone loss/production (called reactivities) for a list of observed key species without the observation of reaction partners like HOx radicals. This is done by preparing a spin-up time of the involved CCM models, initializing the observed key species in the appropriate grid boxes of the model, switching-off model processes like transport, emission, deposition, and then propagation the single boxes in a 0-D fashion 24h into the future. The reactivities are then calculated as 24 hour averages of the processes methane oxidation, ozone production and destruction, and attributing these results to transport, emission, and deposition without initializing the key species to observed values.

Yes, this was our intent. We have already achieved a major threshold in this paper with 6 major chemistry models having implemented the A-run format successfully that will allow them to ingest an observational data stream of key species when they become available (e.g., ATom-1 will become public in the latter part of this year). Such measurements should be able to establish constraints on large-scale "reactivities" as calculated by different models, and thus the agreement/spread of model results with measurements can place some constraints on model predictions for future atmospheres.

The manuscript tries to provide evidence that the switching-off of transport, emission, and deposition does not alter the reactivities too much by comparing full and switched-off model runs. But since the reactivities are 24h averages, the concentrations of the species are the same at the beginning of this 24h period shifting away only slowly afterwards, and since all species are in a kind of balance to each other due to the spin-up period, the real information content of this exercise might be less than anticipated. The exercise by initializing the key species to different values than the models own values is not done. Since such new values of the key species might not be in balance with the other species calculated during the model in the spin-up period, there might be discrepancies when the system tries to adjust to a new balance within the following 24h.

As the reviewer will probably agree, there is enough new material in this paper that needs to be examined, played with, and understood. The final point above is already underway as we have distributed a simulated observing transect from an independent model. Half of the models have completed this exercise and it will be a short, GRL-length paper this summer. We await the public release of the ATom data when the quality controls and calibrations are complete. This is promised by July, and that will begin a new stage of investigation along these lines.

The third part of the approach is to compare relationships between key species, either weighted or unweighted by the reactivities, rather than the key species themselves. I find this kind of approach very interesting since the underlying atmospheric processes introduce complex couplings between different species. Looking at a single species might give a wrong impression because good agreement between model and observation can be caused by compensating errors. But how to compare these

relationships prepared from model data and from observations? There is no objective tool proposed to do that. The interpretation of a match or mismatch looking at relationships in the observations or in the model results still remains to be a subjective choice of the authors or readers of a manuscript. This gap is mentioned in the paper itself at line 645.

Yes, we have introduced quantitative values for the PDs that allow one to address overlap of the weighted parcels in both 1D and 2D. We cannot full flesh this out here (the paper is gaining weight rapidly), but we expect to do this with the upcoming simulated observations paper and then the first paper using all of the ATom data stream in all the models.

Overall, I find that the proposed approach might have a lot of advantages, but its real capability is not proven or even is not illustrated.

The new brief statistics describing the distributions should help with illustration (see Fig S1), but the real capability of these new approaches is just beginning. I believe that we are still in exploration mode.

At the moment, the authors only can show results of climatologies of data calculated by six different CCMs. I think that one of them, model D, is quite different to the other five, for example looking at the probability distribution of HCHO in figure 5, the large values of J-O3(1D) (together with model B) in figure 4, the large values of P-O3 at 20N-60N in figure 3, the large values of L-O3 at 20N-60N in figure 3, or the small values of L-CH4 at 20S-20N in figure 3. If one would take the results of this model D as kind of observed dataset along the fight track shown in figure 9 and would follow the proposed approach, can one expect to reveal the difference of model D compared to the other five? This kind of question is already started at line 547 going to line 561.

Yes, and with the next (pre-ATom release) paper we will see if model D's differences are from the initial species as predicted by model D or from its photochemistry module. From the 1D statistics alone, it is clear that model D has a very different distribution of reactivity since the weighted means shift differently than the other models (Table S5).

But besides mentioning some of the differences of model D there is no clue to explain its different behavior. If such an explanation is not possible for model D, can we really expect to learn something once observed data from ATom are available?

We thank this reviewer for a careful read about what is new in this manuscript and what is missing. We have added a major new component (the statistics supplement) but cannot expand to include all of the useful suggestions. Identifying major differences in model results is a first-order question for the community. This manuscript starts this process with the comparison of six global models. This reviewer clearly understands the complexity of current 3D chemistry models, and must realize how hard it is to identify a cause of the model differences. This thankless task is really up to the originator and developer of that model who can understand the inner tickings. What we provide here are valuable diagnostics (esp. with paper #2) to allow more educated tracing of model differences (whether against other models or observations, and when constrained by initialization of key species).

Response to R. J. Salawitch Comments

> *I am writing to note the submission by Prather et al., to examine probability distribution functions weighted by chemical reactivity, seems to be a promising complement to other approaches being developed to *quantify* differences in the representation of OH and CH4 lifetime within CCMs and CTMs.*

Thanks for the encouragement.

> *The community eagerly awaits the data from ATom, and the approach outlined in the submitted paper will likely advance our understanding of not only why models differ in their treatment of OH and CH4 lifetime, but which models might actually be closer to the truth.*

Yes, the models here are also awaiting the first archival release of ATom data later this summer.

> *At the same time, I am sympathetic to the comment of the reviewer who stated:*
>
> *Sect. 5 seems to suggest that the NASA A Tom missions are unique in providing high-frequency measurements across the Pacific and Atlantic Oceans. Such measurements have been made since decades, which is a great asset to the atmospheric chemistry community and in particular modelers (probably underused), and it should be mentioned that these datasets can also be applied in the presented methodology. The focus on the NASA ATom missions in the manuscript are distinctive, which does not do justice to the more general applications of the method.*

The original contributions and continued value of the other chemistry missions is very important, it is noted up front in the manuscript. ATom does have unique features that we highlight in Section 5. We believe it is the only such transect like mission with a chemical package that includes all the key species for reactivity and regular profiling from 60N to 60S. ATom avoids the subjective sampling which prevents us from using their data without a full hindcast-matchng with the chemical models (not possible with climate models).

> *For instance, the recently completed CONTRAST campaign is in the long-line of missions that have reported publicly available, high-frequency measurements in remote regions of the troposphere.*

CONTRAST was a great mission, but it was admittedly a "process study" for measuring convection from the marine boundary layer to the UT/LS. As such it spent more time in the boundary layer and upper troposphere, rather than the middle troposphere (0.5 – 8 km) where most of the chemical reactivity occurs. With a well thought out adjustment of the sampling patterns, it could be usefully merged with the ATom data. A reference to the CONTRAST mission and to the box modeling is now in the manuscript.

> *Should this paper proceed, Prather et al. might want to cite the usefulness of chemical fingerprinting via use of emission ratios as well as trajectory-based analyses for typing filaments to specific source regions, such as recently published by Anderson et al.*

(Nature Communications, 2016 http://www.nature.com/articles/ncomms10267) as another, complimentary means to look at this type of measurements.

The work here is to identify which parcels are highly reactive and controlling the O3 and CH4 abundances. After these have been identified along with the frequency of their occurrence, then someone else can work on attributing sources for these hot spots.

Finally, the use of the CTM/CCMs in the A-run mode is a fascinating idea. Perhaps this will break the log-jam the community presently faces, driven by the difficulty in separating differences between OH precursors and chemical mechanism, with regards to model differences in OH and CH4 lifetime. At the same time should the editor decide that this paper will proceed, Prather et al. might want to note another new, recently published, promising approach: the use of neural networks (NNs) trained using archived, model output to simulated the chemical mechanism of each global model (i.e., Nicely et al., JGR, 2017 http://onlinelibrary.wiley.com/doi/10.1002/2016JD026239/full)

We acknowledge that neural networks can provide a very useful abstraction of the complex and certainly obscure workings of these CCMs. Indeed, the dataset collected from the six models for ATom could provide a sample of how the models calculate O3 and CH4 production and loss. So, the NNs provide a transfer standard for the chemical model but not for the resulting distribution of key species, whether this is better than running the models themselves, even in A-run mode as demonstrated here is an open question. We put in a reference to this paper in the A-run discussion (lines ~ 450ff).

Personally, I hope this paper does proceed because I think the use of NNs (which require groups to archive specific quantities) versus special, new runs such as the A- run mode (which require groups to "disable processes that connect and mix air parcels" (line 318) will be a ripe discussion point among modelers at future meetings such as the CCMI meeting being held 13-15 June 2017 in Toulouse, France.

Yes, it could.

Global Atmospheric Chemistry – Which Air Matters

Michael J. Prather[1], Xin Zhu[1], Clare M. Flynn[1], Sarah A. Strode[2,3], Jose M. Rodriguez[3,2], Stephen D. Steenrod[2,3], Junhua Liu[2,3], Jean-Francois Lamarque[4], Arlene M. Fiore[5], Larry W. Horowitz [6], Jingqiu Mao[7], Lee T. Murray[8], Drew T. Shindell[9], Steven C. Wofsy[10]

[1]Department of Earth System Science, University of California, Irvine, CA 92697-3100, USA
[2]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[3]Universities Space Research Association (USRA), GESTAR, Columbia, MD, USA
[4]Atmospheric Chemistry, Observations & Modelling Laboratory, National Center for Atmospheric Research, Boulder, CO 80301, USA
[5]Department of Earth and Environmental Sciences and Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York, USA
[6]Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration, Princeton, NJ, USA
[7]Geophysical Institute and Department of Chemistry, University of Alaska Fairbanks, Fairbanks, Alaska, USA
[8]Department of Earth and Environmental Sciences, University of Rochester, Rochester, NY 14627-0221 USA
[9]Nicholas School of the Environment, Duke University, Durham, NC, USA
[10]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, 02138 USA

Abstract. An approach for analysis and modeling of global atmospheric chemistry is developed for application to measurements that provide a tropospheric climatology of those heterogeneously distributed, reactive species that control the loss of methane and the production and loss of ozone. We identify key species (e.g., O_3, NOx, HNO_3, HNO_4, $C_2H_3NO_5$, H_2O, HOOH, CH_3OOH, HCHO, CO, CH_4, C_2H_6, acetaldehyde, acetone), and presume that they can be measured simultaneously in air parcels on the scale of a few km horizontally and a few tenths of km vertically. As a first step, s~~S~~ix global models have prepared such climatologies sampled ~~(~~at the model~~ed~~ resolution~~)~~ for August with emphasis on the vast central Pacific ~~and Atlantic~~ Ocean basin~~s~~. Objectives of this paper are to identify and characterize ~~We show clear~~ differences in model generated reactivities as well as species covariances that could readily be discriminated with an unbiased climatology. A primary tool is comparison of multi-dimensional probability densities of key species weighted by the mass of such parcels or frequency of occurrence as well as by the reactivity of the parcels with respect to methane and ozone. The reactivity-weighted probabilities tell us which parcels matter in this case, and this method ~~.~~ shows skill in differentiating among the models' chemistry. Testing 100-km scale models with 2-km measurements using these tools also addresses a core question about model resolution and whether fine-scale atmospheric structures matter to the overall ozone and methane budget. A new method enabling these six global chemistry-climate models to ingest an externally-sourced climatology and then compute air parcel reactivity is demonstrated. Such an ~~observed~~ objective climatology containing these key species is anticipated from the NASA Atmospheric Tomography (ATom) aircraft mission (2015-2020), ~~measuring the key species,~~ executing profiles over the Pacific and Atlantic Ocean basins. This modeling study addresses ~~work is~~ a core part of the design of ATom.

1. Introduction

To understand global atmospheric chemistry is to understand the chemical heterogeneity of air parcels across the vastness of the troposphere (e.g., Fishman et al., 1996; Ehhalt et al., 1997; Marenco et al., 1998; Jacob et al., 2003; Olson et al., 2004; Kunz et al., 2008; Jacob et al., 2010; Nicely et al., 2016). These air parcels are ephemeral, being continually created, evolving, and mixed with others. Even the concept of discrete as opposed to a continuum of air parcels is a conceit based in part on our modeling of the atmosphere in quantized units such as gridded cells or 1-second averages. Yet, the concept of distinct air parcels remains useful for parsing in situ aircraft measurements and for the analysis presented here in which we ask which air is more important for the chemical evolution of global tropospheric ~~human-driven air~~ pollution.

To understand the mix of chemicals in the atmosphere and where they come from is to recognize how humans have perturbed the common air we breathe. We seek knowledge of the photochemical evolution in each air parcel to understand the overall impact of this heterogeneity and to interpret human impact on past changes and predict future ones. The atmosphere's integration of this chemical reactivity over the ensemble of such heterogeneous air masses controls the evolution of ~~atmospheric residence times~~ of air pollutants and reactive greenhouse gases, particularly methane and ozone. Hence, it allows us to evaluate the consequences of many atmospheric pollutants as regards global air quality and climate.

We have a tendency to simplify this heterogeneity as global, hemispheric, or even regional averages that can be represented with an average chemical composition. This holds true especially when diagnosing the sources and sinks of critical pollutants, or when comparing models with atmospheric measurements. Yet, chemistry inherently involves quadratic reactions of two or more species and hence is non-linear – viz. the chemistry integrated over a mix of parcels is not necessarily the same as that over the average of the mix (e.g., Chatfield & Delaney, 1990). We have progressed in modeling atmospheric chemistry over the past four decades from a few boxes (e.g., stratosphere and troposphere, northern and southern hemispheres) to high-resolution gridded models with many millions of cells. These models simulate myriads of air parcels that at times represent the observed atmospheric heterogeneity of species composition.

For example, Figure 1a presents a single-day snapshot of the column loss of methane as simulated by the UC Irvine chemistry-transport model (CTM) at a resolution of 1 degree in latitude and longitude. Even column averages over 24 hours show a filamentary structure with most of the tropospheric loss of methane occurring in sharp synoptic patterns. These chemical patterns have similarities with the atmospheric rivers of column water vapor (Newell et al., 1992; Dacre et al., 2015; Mundhenk et al., 2016) in terms of filamentary appearance~~,~~ and being ~~both patterns are~~ dominated by the lower half of the troposphere. Nevertheless, the methane-loss filaments do not coincide with atmospheric rivers (Figure 1a vs. 1b), indicating that chemical-specific heterogeneity other than tropical water vapor plays a role in these fine-scale structures, (e.g., Ehhalt et al., 1997; Browell et al., 2003; Charlton-Perez et al., 2009).

This heterogeneity of species and chemical reactivity (e.g., methane loss) is clearly structured and not simply Gaussian. Its structure reflects the combined influence of meteorological transport and mixing as well as the ~~ways~~ patterns that different species are co-emitted and transformed around the globe. For example, large plumes from industrial regions or biomass burning when lofted into the free troposphere by deep convection or frontal systems will naturally be sheared into laminae, travel long distances, and appear ubiquitously (Newell et al., 1999; Stoller et al., 1999; Singh et al., 2000; Blake et al., 2003; Heald et al., 2003; Damoah et al., 2004; Hecobian et al., 2011; Wofsy et al., 2011). This shear or random strain in the atmosphere acts to maintain the pollution concentrated within the layer and preserves the sharp gradients relative to the neighboring atmosphere before they dissolve into the surrounding atmosphere, e.g., (Prather and Jaffe, 1990; Thuburn and Tan, 1997; Esler, 2003; Pisso et al., 2009). Characterizing chemical species in the atmosphere as having mean abundances, or even mean vertical profiles, with a standard deviation to represent the observed variability, does not really describe how these models generate heterogeneity and how the different species co-vary. Assuredly, the atmosphere has more processes and structures than are in our current, high-resolution models as seen in Figure 1, but the extent to which these models represent the key processes shaping the observed patterns is understudied.

Characterizing atmospheric measurements of this chemical heterogeneity specifically for testing models is problematic. Simple direct comparisons of atmospheric rivers, pollution or biomass

burning plumes, and other structures in the troposphere or stratosphere are~~is~~ difficult, even with models using the historical meteorology and chemical emissions, because of slight phase errors in the location of large-scale gradients or laminae (e.g., Reid et al., 1998; Manney et al., 1998; Wild et al., 2003; Kiley et al., 2003; Allen et al., 2004; Schoeberl et al., 2007; Elguindi et al., 2010). The other type of chemistry models, the chemistry-climate models (CCMs), are our means of understanding future air pollution (Prather et al., 2003; Mickley et al., 2004; Jacob and Winner, 2009; Fiore et al., 2012; Barnes and Fiore, 2013; Turner et al., 2013; Fang et al., 2013; Schnell et al., 2015), but CCMs describe the chemical climate and not the hindcast of specific chemical measurements. Most large CCM groups have a parallel CTM versions, but these forced-meteorology versions will likely have different clouds, convection and transport, changing the chemical climatology. Aircraft campaigns often use a photochemical box models to provide an observationally constrained check on reactive species (Olson et al., 2004; Apel et al., 2012; Olson et al., 2012; Stone et al., 2012), and more recently these have extended the box model as a transfer standard across CCM/CTMs (Nicely et al., 2017~~6~~) that can integrate reactive chemistry over 24 hours. The problem remains that the 24-hour integration requires a global model's diagnostics for the diurnal cycle of cloud cover and ozone/aerosol influence on photolysis.

We describe a new approach for developing chemical climatologies from atmospheric chemistry measurements, and for using the major global 3D CTM/CCMs as box-models to integrate the 24-hour rates for important species like methane and ozone. Our goal is to provide climatologies that can point to specific patterns of the key chemical species whose initial values control the chemical evolution of the air parcels. Knowing the correct multi-species patterns, and how different models succeed or fail in reproducing them, will give developers the largest leverage in improving the chemical and physical processes in the models. A critical issue in preparing such a chemical climatology is representativeness, i.e., just how well do the observations represent the region in which they were made and how well should the models match the space-time frequency of the observations. There is a growing literature on the issues of representativeness of atmospheric measurements (Nappo et al., 1982; Crawford et al., 2003; Hsu et al., 2004; Ramsey and Hewitt, 2005; Larsen et al., 2014; Eckstein et al., 2017~~6~~) including defining the chemical patterns through cluster analysis (Koppe et al., 2009).

There have been many ~~Most~~ aircraft ~~tmospheric chemistry~~ missions designed to provide a wealth of in situ, high-resolution atmospheric chemistry data, including some with a nearly complete package of key species needed to calculate reactivities ~~have had a specific focus~~ (Jacob et al., 2003; Engel et al., 2006; Jacob et al., 2010; Pan et al., 2017). Unfortunately, many of these ~~and thus~~ produced a biased, non-climatological sampling, for example, by chasing pollution plumes (Hsu et al., 2004) or by measuring only in clear skies (Nicely et al., 2016). The Pacific Exploratory Missions, PEM-Tropics and PEM-West, were notable in providing a mostly unbiased, exploratory sampling of specific regions in the remote Pacific with a full chemical payload measuring most of the key species (Hoell et al., 1996; Hoell et al., 1999; Raper et al., 2001; Davis et al., 2003). The MOZAIC-IAGOS program uses in-service aircraft and has provided a unique multi-year, objective climatology of some key species (O_3, CO, H_2O) but only along major flight routes at cruise altitudes (8-12 km) and at profiles above airports (Marenco et al., 1998; Thouret et al., 1998; Kunz et al., 2008; Elguindi et al., 2010; Logan et al., 2012; Gaudel et al., 2015).

We examine below some aspects of making objective climatologies of chemical observations, in particular the representativeness of atmospheric transects over the remote ocean basins. Our approach was designed specifically as part of the current NASA Atmospheric Tomography (ATom) aircraft mission in which the DC-8 is instrumented to make high-frequency in situ measurements of the most important reactive species and flies down the middle of the Pacific and Atlantic Oceans, profiling as frequently as possible. The resulting climatology represents the heterogeneity of the atmosphere, including the co-variance of key reactive species.

This approach is tested here using six CTM/CCMs described in Table 1. It allows us to identify models that look alike in reactivity statistics and those that are distinctly different. We have seen large uniform anomalies in a specific species as well as different patterns or locations of the most reactive parcels. For example, we list the models' average reactivities for the tropical Pacific and the globe in Table 1c. The tropical Pacific average P-O3 is similar across models and is about ½ that of the global average, which is dominated polluted, near-surface parcels over land. The L-O3 is typically the same for Pacific and the globe; and the L-CH4 is greater over the Pacific than over the globe. Model D stands out in reversing or exaggerating these typical Pacific vs globe

differences, indicating very different locations for the reactivity. We use t~~examples. T~~hese models to ~~y are meant to~~ demonstrate the methodology and the ability to discriminate among them with ATom-like measurements. ~~The m~~Model versions ~~are~~ used here should be considered snapshots in the development cycle. ~~as examples. They are meant to demonstrate the methodology and the ability to discriminate among them with ATom-like measurements.~~ No model tuning or development occurred as part of this work, except to correct where quantities were missed or misdiagnosed. These diagnostics ~~will~~ need to be revisited for the model versions used in upcoming assessments (e.g., Lamarque et al., 2013; Collins et al. 2016)

Typically, the probability of occurrence of a species' abundance is weighted by the air mass of the parcel, but if we are interested in the chemical reactivity, then the parcel should be weighted by the chemical rates in the parcel (e.g., moles per day). Such weighting is an obvious choice in that it tells us which air parcels matter for chemical budgets, including, for example, whether infrequently observed pollution plumes are responsible for a large fraction of ozone production.

In Section 2 we define our use of reactivity in this paper (i.e., the production and loss of ozone, the loss of methane) and identify about a dozen key chemical species and other variables that once initialized determine the chemical evolution of an air parcel. In Section 3 we show how the CTM/CCMs can be altered slightly to calculate the reactivity of air parcels using the native grid cells of the model and a prescribed initialization of the key chemical species. This approach allows the CTM/CCMs to be run using either model data or observations, or a mixture of both. In Section 4 we derive multi-dimension probability distributions for these key variables over a suitable latitude-longitude-pressure domain using grid-cell values from several CTM/CCMs. Tables of simplified statistics describing these probability distributions are presented and discussed in the Supplement to this paper. The~~se~~ full distributions and simple statistics clearly show the basic differences in chemical heterogeneity and reactivity across the six models. We conclude in Section 5 with a summary the model comparisons and what is learned from the new diagnostics. We also discuss ~~a discussion of~~ the ongoing NASA ATom mission (2015-2020), which will provide the air parcel measurements of key species to initialize the models' calculation of reactivity in each parcel and thus provide an observed climatology of the chemical reactivity of the troposphere. This approach moves us towards an understanding of which

species exert the largest influence on the atmosphere, and thus which are thus most crucial for us to establish a global climatology.

Table 1a. Participating models					
model	type	driving meteorology	year	model grid	effective resol @ 500hPa
CAM4-Chem	CCM	SSTs	2000s	0.47°x0.625°x52L	0.47° x 0.625°x 38hPa
GEOS-Chem	CTM	GEOS5-FP	2013	2º x 2.5º x 72L	2º x 2.5º x 38hPa
GFDL-AM3	CCM	NCEP (nudged)	2013	C180L48	0.5° x 0.5° x 71hPa
GISS-E2	CCM	Daily SSTs prescribed, winds nudged to MERRA	2013	2º x 2.5º x 40L	2º x 2.5º x 50hPa
GMI-CTM	CTM	MERRA	2001	1° x 1.25° x 72L	1° x 1.25° x 38hPa
UCI-CTM	CTM	ECMWF IFS Cy38r1	2005	T159N80L60	1.1° x 1.1° x 38hPa

Table 1b.			
model	POC	email	model url
CAM4-Chem	Jean-Francois Lamarque	lamar@ucar.edu	http://www.cesm.ucar.edu/models/current.html
GEOS-Chem	Lee Murray	lee.murray@rochester.edu	http://geos-chem.org_ver 10-01
GFDL-AM3	Arlene Fiore	amfiore@ldeo.columbia.edu	https://www.gfdl.noaa.gov/am3-model/
GISS-E2	Lee Murray	lee.murray@rochester.edu	http://www.giss.nasa.gov/tools/modelE/
GMI-CTM	Sarah Strode	Sarah.A.Strode@nasa.gov	http://gmi.gsfc.nasa.gov
UCI-CTM	Michael Prather	mprather@uci.edu	ftp://halo.ess.uci.edu/public/xzhu/qcode_72c

205

Table 1c. Model average reactivities (ppb/day) for P-O3, L-O3, L-CH4.							
model	code	P-O3		L-O3		L-CH4	
		Tr.Pac.	Global	Tr.Pac.	Global	Tr.Pac.	Global
CAM4-Chem	A	0.979	2.070	1.963	1.802	1.017	0.745
GEOS-Chem	B	0.791	2.252	1.616	1.837	0.765	0.738
GFDL-AM3	C	0.860	2.036	1.550	1.535	0.726	0.599
GISS-E2	D	1.092	3.715	2.589	3.409	0.453	0.693
GMI-CTM	E	0.778	1.513	1.834	1.690	0.848	0.674
UCI-CTM	F	1.088	2.100	1.788	1.990	0.854	0.702
All results are mass-weighted by tropospheric parcels up to 200 hPa from the model C-runs for 16 August.							

Table 1dc.		
model	~~graph label~~code	relevant refs
CAM4-Chem	A	(Lamarque et al., 2012;Tilmes et al., 2016)
GEOS-Chem	B	(Bey et al., 2001;Eastham et al., 2014)
GFDL-AM3	C	(Donner et al., 2011;Naik et al., 2013a;Li et al., 2016)
GISS-E2	D	(Schmidt et al., 2014;Shindell et al., 2013)
GMI-CTM	E	(Strahan et al., 2007;Duncan et al., 2007)
UCI-CTM	F	(Holmes et al., 2013;Holmes et al., 2014;Prather, 2015;Sovde et al., 2012)

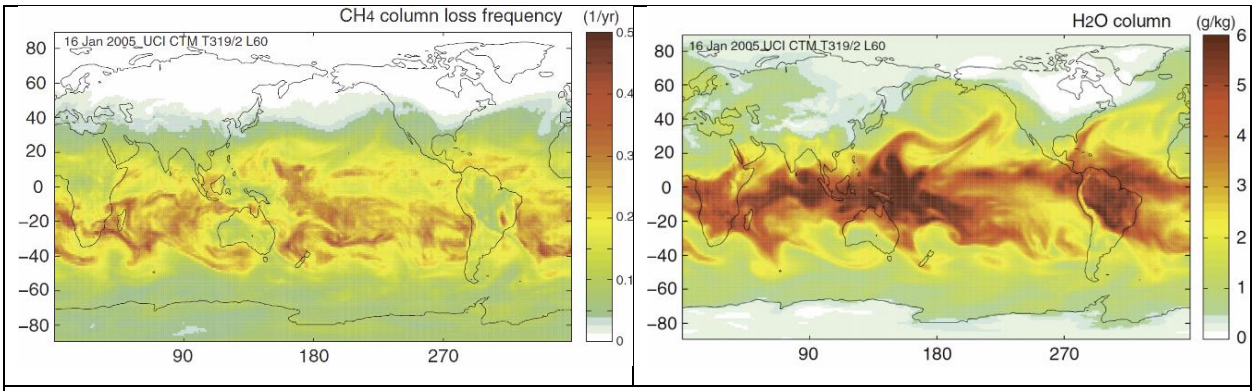


Figure 1. (a) Column tropospheric loss frequency (1/yr) for CH_4 and (b) column average H_2O abundance (g-H_2O/kg-air) taken from a 1-day integration (16 Jan 2005) using the University of California Irvine (UCI) chemistry-transport model (CTM) run at T319N80L57 resolution (~1° horizontal) using forecast meteorology from the European Centre for Medium-Range Weather Forecasts, see Sovde et al. (2012). As expected, the northern winter shows very little CH_4 loss above 40N.

2. *Key chemical species for tropospheric reactivity*

The reactivity of an air parcel is defined here as a daily average of the rates affecting critical species, in this case, ozone (O_3), a greenhouse gas and air quality threat, and methane (CH_4), the second most important emitted greenhouse gas after CO_2. Methane is emitted mostly through human activities but also naturally; and it is lost primarily (>80%) through reaction with the hydroxyl radical (OH) in the troposphere (reaction 1). Other atmospheric losses in decreasing order of magnitude and certainty are reaction with stratospheric OH, surface uptake by biota, and reaction with Cl atoms (Prather et al., 2012; Ciais et al., 2013).

$$CH_4 + OH \rightarrow CH_3 + H_2O \quad (1)$$

The CH_4 abundance varies little throughout the troposphere (~10%), and the destruction of CH_4 occurs with a mean loss frequency of ~0.1 /yr (see Fig. 1a). Here we focus on calculating the tropospheric loss of CH_4 by OH over 24 hours (reaction 1, designated L-CH4) in units of ppb (nanomoles/mole-air) per day. L-CH4 is highly variable across parcels, and the integral of L-CH4 determines the atmospheric lifetime of CH_4 and the buildup of its emissions in the atmosphere.

Tropospheric O_3 has stratospheric sources and surface sinks, which average to about 0.2 - 0.3 ppb per day, and much larger in situ photochemical production and losses that average about 1.1 - 1.5 ppb per day (Stevenson et al., 2006; Stevenson et al., 2013; Young et al., 2013; Hardacre et al., 2015). The O_3 abundance varies greatly throughout the troposphere, by a factor of 10 or more, and its mean residence time is about a month (Stevenson et al., 2006; Wu et al., 2007; Hsu and Prather, 2009). O_3 is an intermediate source of atomic O in many tropospheric reactions, and its net production and loss is determined in the long term by the breaking and reforming of the O-O bond originating with molecular oxygen. Chemical reactions are traditionally grouped into production (P-O3, ppb/day)

$$HO_2 + NO \rightarrow NO_2 + OH \quad (2a)$$

$$RO_2 + NO \rightarrow NO_2 + RO \quad (2b)$$

$$\text{where } NO_2 + h\nu \rightarrow NO + O \text{ and } O + O_2 \rightarrow O_3 \quad (2c)$$

$$O_2 + h\nu \rightarrow O + O \text{ (times 2)} \quad (2d)$$

and loss (L-O3, ppb/day).

$O_3 + OH \rightarrow O_2 + HO_2$ (3a)

$O_3 + HO_2 \rightarrow HO + O_2 + O_2$ (3b)

$O(^1D) + H_2O \rightarrow OH + OH$ (3c)

where $O_3 + h\nu \rightarrow O(^1D) + O_2$ (3d)

In the troposphere, reaction 2d is important only in the tropics above 12 km (Prather, 2009). The true P minus L of O_3 includes a large number of other reactions, particularly involving oxides of nitrogen and hydrocarbons; but throughout the remote troposphere (i.e., away from fresh pollution sources), reactions (2) minus reactions (3) accurately approximate the true P – L that the models calculate using the full set of reactions. One reason for separating P and L in this way is to think of P as independent of O_3 and L as being linearly proportional. Unfortunately, while the P reactions (2) have no obvious O_3 terms, both these reactions and the OH and HO_2 abundances in reactions (3) depend indirectly on O_3; and thus with a true linearization of P–L, the lifetime of O_3 is much shorter than inferred from L (Prather and Holmes, 2013). A similar chemical feedback with opposite sign occurs for CH_4 whereby the lifetime of a CH_4 addition is longer than inferred from the linear relationship of reaction (1) (Prather, 1996). We retain these definitions of P-O3, L-O3, and L-CH4 because they still represent the reactivity in remote regions and the reaction rates, rather than a linearization, are straightforward CTM/CCM diagnostics.

We define the reactivity of an air parcel (reactions 1-3) in terms of 24-hour average rates and hence the units of ppb per day. Reactivity defined here requires sunlight; nighttime sources of OH from alkenes and isoprene via ozonolysis or nitrate radicals (Paulson and Orlando, 1996) are important primarily in continental air over emission sources. This calculation integrates over the diurnal cycle of photolysis rates driven by changing solar zenith angle, clouds, O_3 and aerosol profiles, all of which are simulated in CTM/CCMs.

What key constituents are needed for modeling reactivity? Models simulate many tens to hundreds of chemical species. While many are important for calculating the instantaneous reaction rates, e.g. $O(^1D)$, they are not the key species. Key is defined here as a constituent whose initial value significantly affects the 24-hour reactivity, whereas other species can be initialized to any reasonable value and not affect it. For example, OH and HO_2 are radical HOx

species whose abundances directly determine the rates of reactions (1-3). Nevertheless, these are not key species as their abundances can be initialized to zero and are rapidly reset in seconds to a temporary steady state with first sunlight or changing clouds through reactions 3cd among others (Rohrer and Berresheim, 2006). This argument applies to similar radical species such as CH_3OO, but not to HOx sources like CH_3OOH and HOOH whose initial values will control the abundance of OH and the reactivities over the day.

A similar situation applies to NO and NO_2 (collectively designated NOx), whereby total NOx changes over the day as it is exchanged with higher oxides of nitrogen but the fraction of NOx in the form of NO is determined rapidly in sunlight by reactions 2abc and 4.

$$NO + O_3 \rightarrow NO_2 + O_2 \qquad (4)$$

In the dark, NOx is almost entirely NO_2, and it is critical to initialize NOx, but not NO and NO_2 separately.

Based on sensitivity tests with the UCI CTM, our list of 18 key species includes:

O_3, NOx, HNO_3, HNO_4, PAN ($C_2H_3NO_5$ = peroxyacetylnitrate), RNO_3, (CH_3NO_3 and all alkylnitrates), HOOH, ROOH (CH_3OOH and smaller contribution from C_2H_5OOH) , HCHO, CH_3CHO (acetaldehyde), C_3H_6O (acetone), CO, CH_4, C_2H_6, alkanes (all C_3H_8 and higher), alkenes (all C_2H_4 and higher), aromatics (benzene + toluene + xylene), C_5H_8 (isoprene + terpenes). (5)

We also add p (hPa), T (K), q (g-H_2O/kg-air), and latitude and longitude to make up the 23 key variables in each air parcel. Some collectives like alkanes may be treated as multiple, separate species in some models, or may be lumped according to their reaction rates. List (5) tends to be inclusive because for much of the troposphere, a smaller list can apply. For some species (e.g., C_5H_8), their role is key only if they are present in large enough abundances, but even when sampling across the Pacific Ocean basin one may find plumes with recent biospheric sources.

This simplification of the chemical system fails in regions of intense emissions of short-lived species or in highly polluted environments such as urban, industrial, or open fires. After pollution plumes have been separated from sources and aged a few days, then our key variables should define the reactivity. Such conditions apply to most of the troposphere, particularly the

air over the vast Pacific and Atlantic Ocean basins. With aged pollution plumes, we expect that some key species (e.g., alkenes, isoprene, aromatics and higher alkanes) will drop off the list because their abundances in much of the remote mid-ocean regions will have fallen below the relevance threshold.

3. Modeling the reactivity of air parcels

Why use the global models instead of single-box models to calculate reactivity statistics? There are several reasons. For one, these CTM/CCMs simulate the full meteorology including cloud cover and its variation over large regions, which is a critical component of reactivity. Second, they usually include self-evaluated ozone and aerosol profiles also needed for the photolysis rates. Third, these models automatically simulate the diurnal cycle in radiation at all seasons, latitudes and longitudes. And fourth, most importantly, these models have built-in chemistry modules that already calculate reactivities, and they are the ones we rely on for climate and air quality assessments. The goal here is to test their simulated chemical heterogeneity. While a box model could be designed (using 3D meteorology) to address the first three needs (e.g., Nicely et al., 2016), it cannot address the last. More simply, all the necessary Earth system components are already built in to the CTM/CCMs, and our approach of testing the modeled climatologies includes that of testing the Earth system components (e.g., emissions, transport, chemistry, scavenging, air-sea exchange, and land-surface interactions).

In a standard CTM or CCM simulation (defined here as a C-run) we calculate the reactivity at a given grid cell, but not that of a parcel. Air parcels move, change location, and mix with neighboring parcels: i.e., there is no way to track quantitatively what might be considered the original parcel. Effectively, we keep integrating the rates in that grid cell as different parcels travel through it and are mixed within it. Let us take a large enough domain of grid cells (e.g., tropical Pacific, 150E to 210E, from surface to 200 hPa) and calculate the statistical distribution of reactivities of all those grid cells. We take these statistics to be equivalent to those we would get from integrating the reactivity over isolated air parcels with the same initialization. Of course the latter is only a thought experiment since the parcels do not remain isolated. In C-runs new air parcels are entering the domain and others are exiting. In a single cell we can start with

a polluted lamina and end with clean air convected from the marine boundary layer, but much of the polluted lamina remains in the larger domain. As long as the domain retains a statistical mix of the key chemical species similar to the initialization, then the reactivity statistics of the C-run should represent the hypothetical reactivity of those initialized parcels.

How can we design a calculation using the CTM/CCMs that allows us to initialize a subset of grid cells with observed air parcels and then calculate a reactivity for those parcels? The goal here is to be able to use the NASA ATom aircraft mission (2015-2020), which was designed to measure those 23 key variables in air parcels profiling from near-surface to 12 km altitude, flying ascents and descents down the middle of the Pacific and Atlantic Ocean basins. Thus ATom data will not fill the global 3D model grid, and thus many cells will be initialized with the model's original chemistry values. The critical design requirement is that we let the model integrate for 24 hours as it normally does in a way that the chemistry in each grid cell depends minimally on any of the grid cells around it.

We thus propose an A-run mode (named after the ATom mission) for the CTM/CCMs in which individual parcel reactivities can be calculated, albeit with some simplifying approximations. Consistent with our definition of reactivity, we consider only ATom parcels that are tropospheric. The A-runs disable processes that connect and mix air parcels. First we drop all calls to the tracer transport sections (advection, convection, diffusion, boundary layer mixing). Second, we must cut all emissions, including lightning and aircraft NOx, because without transport, the emissions would build up unrealistically in the source cells. Third, all tracer scavenging modules must be turned off because in many models the scavenging depends on the vertical distribution of the species.

In this A-mode, the remaining connection of the reactivity calculation with neighboring grid cells is through the photolysis rates, which require profiles of clouds, aerosol layers, and ozone. It is impossible to prescribe all these data over the diurnal cycle for each parcel from observations, and thus we must rely on the CTM/CCM to generate a suitably realistic, diurnal, regional, seasonal climatology for these and hence the photolysis rates. To better average the reactivity

over synoptic variations in clouds, we expect to repeat the same initialization of the A-runs for a range of days over a month containing the observations.

Each ATom parcel (2-8 km along the flight path) will be assigned a unique model grid cell to best match the observation: latitude and pressure grids containing the measurement, and longitude chosen as close as possible but maintaining a unique cell for each parcel. ATom parcels in adjacent grid cells may represent air masses separated by a few km instead of the grid-cell size of order 100 km. A high density of ATom parcels in a region will be placed in the correct latitude and pressure cells but may be strung out in longitude cells. The parcel will use the mole fraction of key species, water vapor (q) and temperature (T) as measured, but will adopt the mean pressure of the grid cell. The model may need to maintain separate storage for the hourly T and q used in the CCM dynamics because it is important to maintain the clouds as they would be done in the C-run, and thus the main-code values of T and q cannot be overwritten with ATom values. The A-run treatment of stratospheric O_3 (i.e., fixed) is unlikely to be identical to the C-run, but it does not appear to drive major changes in the average photolysis rates over a region (see below).

In defining the A-runs thus, we have created some biases in the reactivities relative to the C-runs. Examination of the NOx and HOx budgets of parallel A- and C-runs shows two obvious differences. The A-runs lack emissions and over the remote ocean basins, where the most important emission is NOx (lightning, shipping, aviation). Thus A-runs show a 24-hour decline in NOx abundances compared with the C-runs, resulting in generally lower P-O3. The A-runs also lack scavenging, and thus accumulate more HNO_3 and HOx precursors like HOOH, affecting L-CH4. No other simple objective approach has been found, and we must accept and document these biases in the A-runs.

An examination of how the A- and C-runs differ is shown in Figure 2 using the UCI-CTM's 1D probability distributions of six key species (NOx, HNO_3, HNO4, PAN, HCHO, HOOH) for the central tropical Pacific. The initial distribution for both runs (12h local solar time at 180E, black solid) can be compared with that 24 hours later (36h) for the C- (black dashed) and A-runs (cyan squares, only for 4 species). The number of moles at the beginning and end of the 24-hours in

the C-run (see legend) is a measure of the daily changes in the air parcels entering and leaving the domain. It varies from 0 to 4%, well within the expected representativeness of a given day. With the A-run, however, we see large systematic shifts due to the lack of emissions (NOx) and scavenging (HNO_3, HOOH). For HNO_3 the content increases overall by 9%, with the high-end (>100 ppt) distribution not changing but the low-end (<20 ppt) air gains HNO_3, increasing the middle section (20-100 ppt). This is logical because the low-HNO_3 regions have the most scavenging. This change in distribution over the 24-hour integration of the A-runs is unlikely to change the reactivities as the release of NOx from HNO3 will be more important in the high-HNO_3 regions. For NOx the content decreases overall by 18%, with most air parcels (4-100 ppt) becoming less frequent and an increase in frequency only for parcels with very low NOx, <4 ppt. The 1D distribution of HCHO shifts lightly, but with little overall change in content. The lack of scavenging is even more important for HOOH with an overall increase of 41% and a dramatic shift in the distribution: decreases in 0.3-1.0 ppb appearing as very large increases from 1.0 to 2.5 ppb. The implications for using the A-run bias in computing the reactivities is examined with all 6 models below.

An important assumption in using key species to initialize the reactivity simulations is that the diurnal cycle is not critical, and that ATom measurements can be used without trying to make corrections for the time of measurement. In running these global models, it is not practical to initialize parcels at other than a standard day (i.e., beginning at 00h UT). For some species like HCHO, the daytime loss frequency in the tropics is about 1/2 hr^{-1} (see for example loss photolysis rates for various oxygenated hydrocarbons in Prather (2015)), and thus one might expect it to vary greatly over the sunlight day or with cloud variations. The diurnal change in 1D distributions of the 6 key species is also shown in Figure 2 for the C-runs at 18h (local solar time, red dashed), 24h (dark blue dashed) and 30h (green dashed). The C-runs are in approximate steady-state over the tropical Pacific domain as seen by comparing 12h with 36h, and thus these sunset-midnight-sunrise times show the daily variations. The diurnal cycle does produce visible shifts in the 1D distributions, particularly at the end of night (30h). The shifts in HCHO are small considering its high loss frequency, primarily because both sources and sinks respond similarly to photolysis rates. The seemingly longer-lived HOOH shows larger shifts because production occurs in sunlight but scavenging occurs day and night. PAN and HNO_4 show small

diurnal cycles at the high-abundance end of their distributions where they can be important NOx sources, and initialization errors caused by the diurnal cycle at the low-abundances will have smaller impacts on reactivity.

A test of A- versus C-runs for all 6 CTM/CCMs is shown Figure 3. All models were spun up for a year and stopped at 00H UT on 16 August, with the chemical abundances at this time being used to initialize each model's own C- and A-runs. In this case all species in the model were initialized and not just the 18 key species. Each model ran their own chemistry and meteorology intended to simulate a specific historical year or a typical climate year. All were intended to be typical of the last decade. The models were then run 24 hours and the rates and reactivities diagnosed for both C-runs and A-runs. All models have different resolutions, ranging from 0.5 to 2 degrees. All model statistics (key variables, reactivities, plus 24-hour average photolysis rates) were stored globally. This analysis examines a north-south transect flight over the Pacific Ocean basin as in the NASA ATom flights, but greatly expands the region to include more grid cells: latitude, in 6 domains 60S-40S-20S-0-20N-40N-60N (each region is color keyed in Fig. 3); longitude, in a single broad domain 150E-210E. Vertical profiles (200 hPa–1000 hPa) on the models native grid are shown for the 6 domains as different colors. The standard C-runs with all transport and emissions included are solid lines, while the ATom-like A-runs are dashed.

For L-CH4, the only general agreement is the lesser importance of parcels at altitudes above 500 hPa. For this August test, most models find that the 20N-40N dominates (note that plots are ppb/day and not area weighted), and the 60S-40S and 40S-20S domains are least important (similar to OH structures in Spivakovsky et al. (2000); Lawrence et al. (2001)). Most models show increasing L-CH4 in the first few km above the ocean because of low-level clouds shifting photolysis to the middle troposphere The results for L-O3 show similar patterns of agreement and disagreement among models but emphasize the dominant role of the middle troposphere (500–800 hPa) for O_3 loss. P-O3 has distinct patterns, demonstrating the importance of larger NOx values in the upper (200–500 hPa) and lower troposphere (800–1000 hPa), presumably from lightning NOx. Only GMI-CTM lacks lower troposphere sources of O_3 about 180E. Overall the models show modest, similar amplitudes (but not always sign) in the bias of A-runs relative to C-runs. Thus we can use the model A-runs to tag each parcel in the ATom measured

climatology by its reactivity in the absence of emissions and transport. Clearly these models have largely different chemical climatologies for the middle of the Pacific, and with the ATom climatology to initialize all six models, we will be able to test if these differences reflect the initial key species and/or the photochemical components.

Photolysis rates (J-values) are the driving force for reactivity, and we include also a comparison of the 24-hour average J's (reactions 2d and 3c) in Figure 4. The model spread in J-NO_2 is 20% and likely due to differences in cloud cover as well as the photolysis module in the model. The wide, factor-of-two range in J-O_3(^{1}D) cannot be simply explained through differences in clouds or ozone, for example, a 20% reduction in column O_3 gives only a 33% increase. Such differences will drive a large part of the model differences seen in Figure 3. For example, the large J-O_3 for GISS, and hence large production of OH, can explain in part why GISS has very large L-O3 and P-O3, but not why the L-CH4 (also dependent on OH) matches the other models. Surprisingly GEOS-Chem has even larger J-O_3 but its reactivities are within the range of the other 4 models. A comparison between the A- and C-runs (not shown) confirms that these two runs have almost identical J's as expected since these changes in ozone and aerosols over 24 hours between these two simulations will have small impact on regional average J's.

While the A-run is clearly asking the modeling groups to make some rather uncomfortable code modifications, these tend to be at the very high level of disabling entire components. Other approaches for indirectly comparing chemical models without transport have been developed (e.g., neural networks in ~~{~~Nicely et al., 2017). ~~#4981}~~We ~~recommend~~ choose the A-run~~is~~ approach as it will allow us to more directly compare modeled reactivities based on the primary CTM/CCM coding~~,~~ and still allows ~~including when~~ for all models to be ~~are~~ initialized with the same chemical composition.

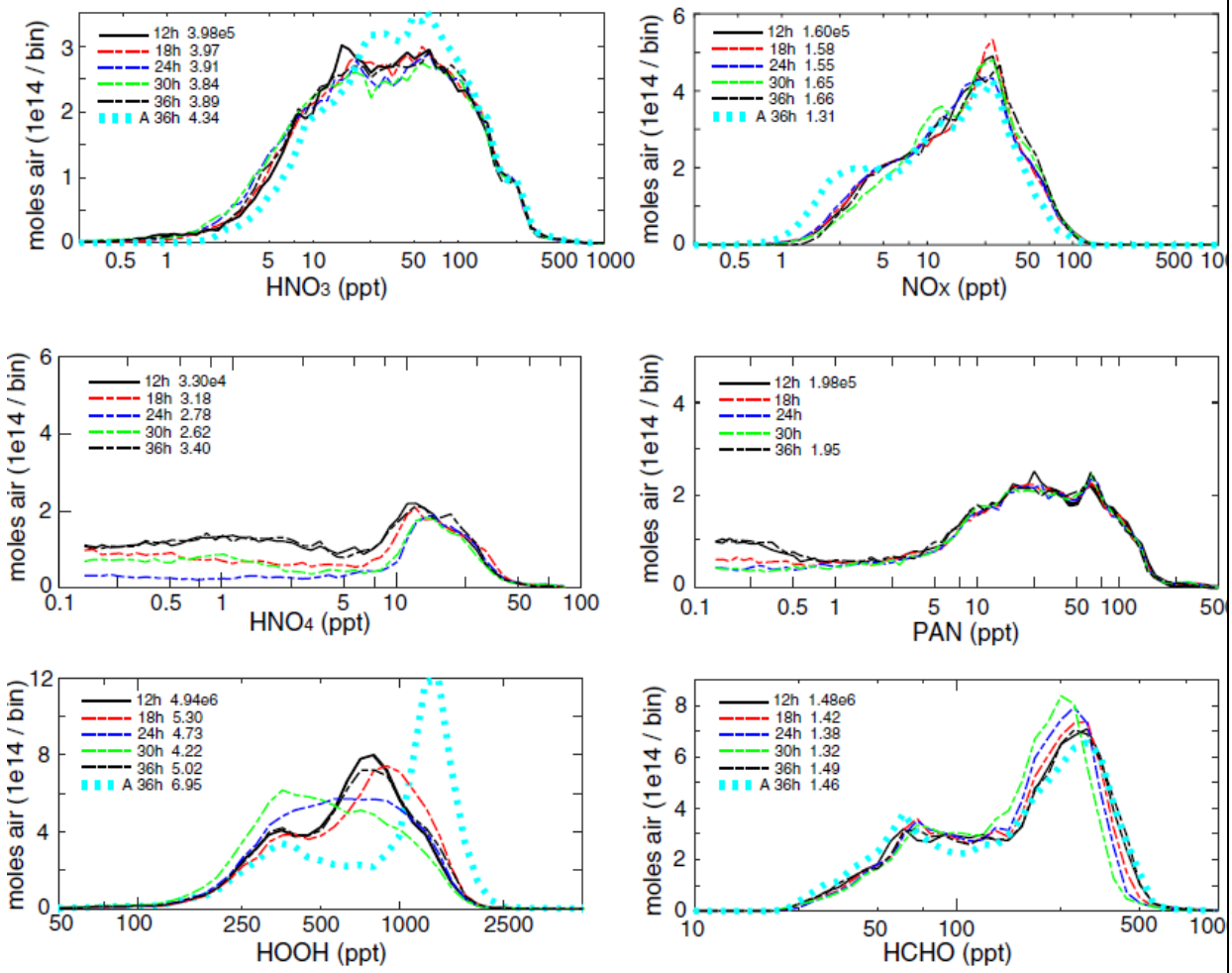
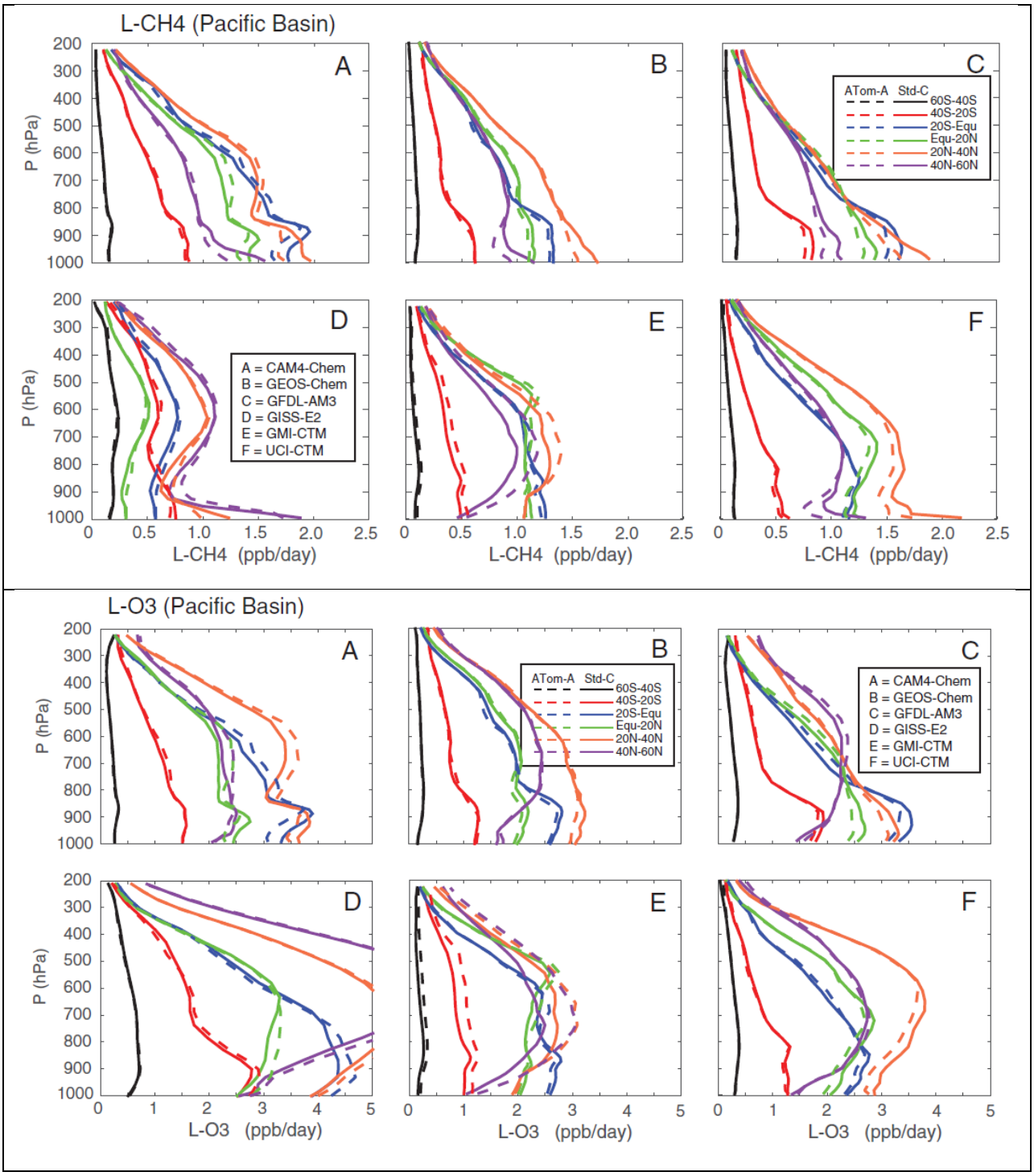


Figure 2. 1D probability distributions for HNO_3, NOx, HNO_4, PAN, HOOH, and HCHO from the UCI-CTM. The domain sampled is the tropical Pacific: 20S-20N, 150E-210E, 0-12km, on 16 August. The units are moles of air per log-scale bin (20 bins per factor of 10). The area under the curve in the log plot is the air mass of the domain, except for HNO_4 and PAN for which there are numerous observations below the cutoff at 0.1 ppt. Five different times are shown for the C-run: local noon (12h), sunset (18h), midnight (24h), sunrise (30h) and the following noon (36h). Also shown is the A-run at noon (12h, same as C-run) and the following noon (A 36h). The numbers of moles of the species in the domain are given in the legend.

485


L-CH4 (Pacific Basin)
A
B
C
D
E
F
P (hPa)
200
300
400
500
600
700
800
900
1000
ATom-A Std-C
60S-40S
40S-20S
20S-Equ
Equ-20N
20N-40N
40N-60N
A = CAM4-Chem
B = GEOS-Chem
C = GFDL-AM3
D = GISS-E2
E = GMI-CTM
F = UCI-CTM
0 0.5 1.0 1.5 2.0 2.5
L-CH4 (ppb/day)
L-O3 (Pacific Basin)
0 1 2 3 4 5
L-O3 (ppb/day)

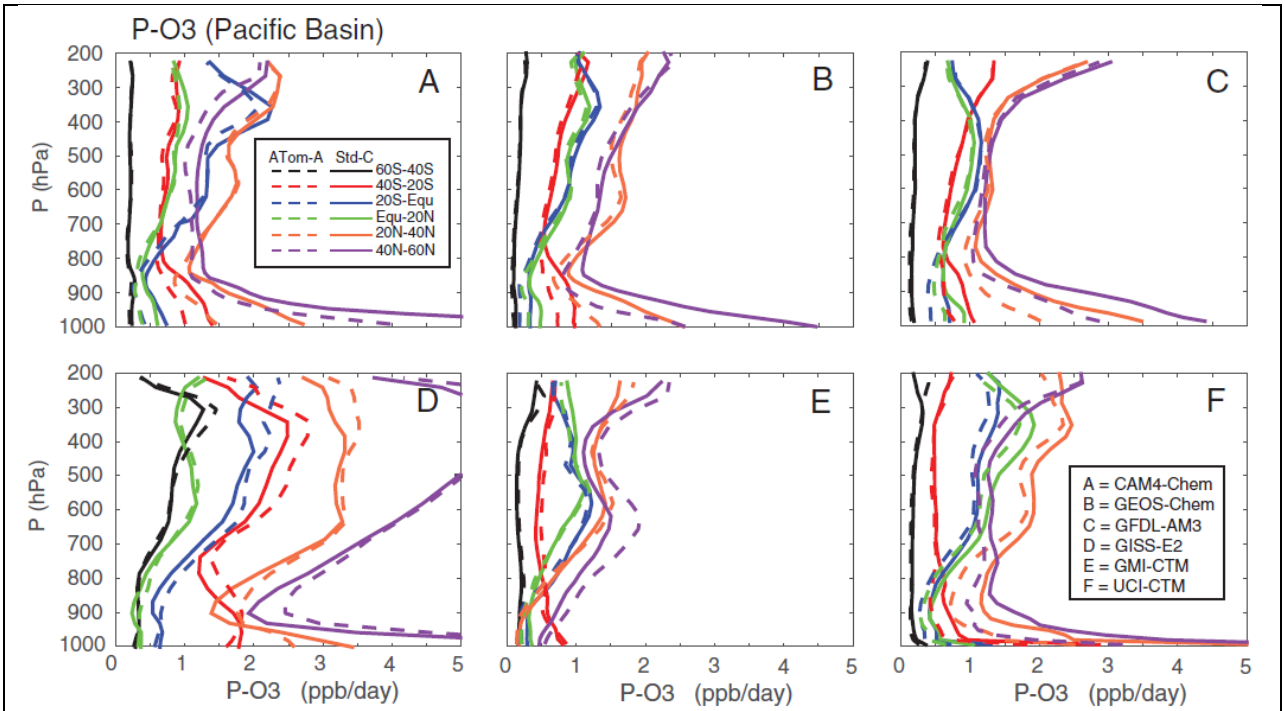


Figure 3. Profiles of reactivity (ppb/day) for loss of CH_4 (L-CH4, top panel), loss of O_3 (L-O3, middle panel), and production of O_3 (P-O3, bottom panel) from 6 global models (Table 1). Cells from each model grid are averaged over 20-deg latitude domains (different colors, see legend), longitudes from 150E to 210E, and for the single day of 16 August. Years vary by model, see text. Solid lines are standard model simulations (C-runs) with the values representing air that passed through the cell over 24 hours. Dashed lines are the no-transport, no-emissions A-runs that keep the initialized chemical values in the same cell over 24 hours.

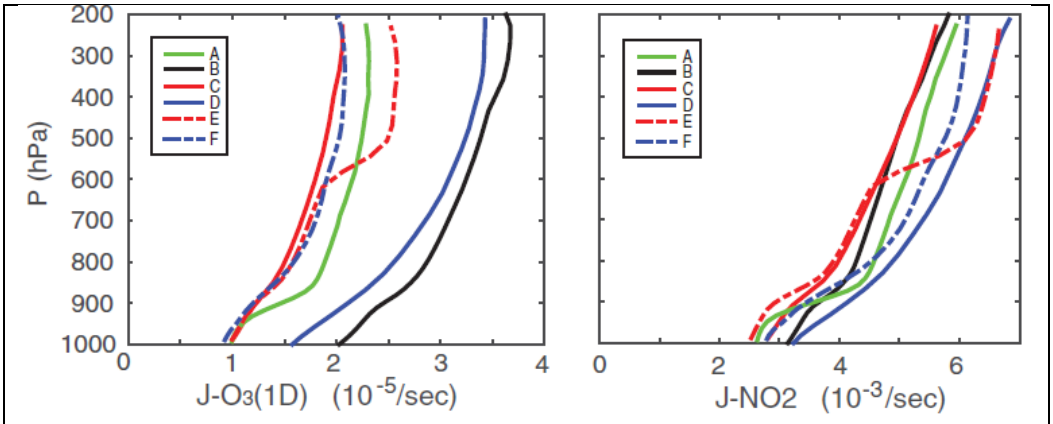


Figure 4. Modeled 24-hour average J-values for O_3+*hv*=>$O(^1D)$+O_2 (left, s^{-1}) and NO_2+*hv*=>NO+O (right, s^{-1}) for the tropical Pacific (20S-20N, 150E-210E). See Fig. 3 and Table 1 for model codes.

4. Probability distributions of species and reactivities

We characterize the heterogeneity in tropospheric chemistry through the joint-probability distributions of the frequency of occurrence of chemical species in air parcels for the six models here. These diagnostics are readily suited to high-frequency in situ observations from an extensive aircraft mission such as ATom, for example see (Koppe et al., 2009). This paper then takes a novel approach by focusing on the chemical budgets for tropospheric ozone and methane. In addition to weighting a parcel according to its occurrence or parcel mass, we include a factor that accounts for the model-calculated reactivities of that parcel. For example, the basic weight of a parcel (moles-air) can be scaled by P-O3 (ppb/day) and the final weight is the moles-O_3/day. In this case the sum of weighted parcels in a region gives the moles of O_3 produced per day in that region. These reactivities can be calculated with A-runs for both models and measurements. Thus, the modeled and measured probability distributions reflect the parcels most important in determining the chemical budgets in these models, and hence the evolution of the atmosphere.

Given the number of key species, the joint probability distributions are multi-dimensional, but for the most part we view them 1D or 2D graphs. There is a history of comparing models and measurements using such graphs (Hoor et al., 2002; Hsu et al., 2004; Engel et al., 2006; Pan et al., 2007; Strahan et al., 2007; Parrington et al., 2013; Gaudel et al., 2015). Often the goal is simply to define a linear correlation, but in many cases a line-fit simply does not describe the heterogeneity (Koppe et al., 2009).

A much more difficult problem is that of representativeness: i.e., How much of the Pacific basin must one sample to get joint probability distributions similar to that of the whole basin? Can aircraft-measured heterogeneity be compared with models that do not follow the exact flight route for the exact period of measurements (e.g., Hsu et al., 2004)? This latter question is critical if we are to use the ATom measurements to test such a wide variety of CTM/CCMs. Here, we consider an idealized test case for representativeness where we sample a model as objectively as possible and then compare with different sampling ‘paths.’

One test of representativeness looks at the reactivities sampled along a single longitude and then integrated over latitude-pressure domains. For example, Figure 1a clearly shows that the instantaneous column integrated L-CH4 varies greatly along longitude transects in the mid-Pacific. The point-to-point variance in 3D will be very large, but if we average over regional domains, can we achieve a representative mean value for reactivity? Based on the profiles of reactivity (Fig. 3) we take three pressure domains (surface – 850 hPa – 500 hPa – 200 hPa, but with stratospheric values screened out by model-designated discriminators) and three latitude domains (60S – 20S – 20N – 60N). The means (ppb/day) and standard deviations (ppb/day) of single-longitude sampling across the mid-Pacific (155E – 233E) on 16 August are shown for the UCI-CTM in Table 2 along with the standard deviation (in %) over the 31 days of August of the daily full-domain average. The standard deviations are a measure of the representativeness of the sampling, by longitude or by day. For L-CH4, the dominant mean loss, >1 ppb/day, is in the surface – 500 hPa in the tropics and summer (northern) mid-latitudes as seen in Fig. 3. For these regions the standard deviation across the longitudinal samples is of order 6-11%; whereas outside of these, it is as large as 20%, but the absolute values are small. A similar pattern holds for L-O3 with standard deviations in dominant regions of 6-14%. Thus any single, fully sampled longitudinal transect through this domain has a 68%-likelihood of being within 6-14% of the mid-Pacific average. The variance of P-O3 is slightly larger, 8-17%, in part because P-O3 depends on the less-frequent high-NOx regions. Assembling a representative sampling of P-O3 at the same % level as L-O3 will be slightly more difficult. Such single-transect representativeness is about as good as we can expect. Thus, model-model differences comparing individual transects from each model would not be significantly unless they exceed these percentages. Averaging over the basin and/or several days should resolve model differences at finer scales. The day-to-day standard deviation for the mid-Pacific averages in Table 2 is shown in percent; it is smaller than across individual longitudinal transects for a given day; and in key regions (surface to 500 hPa, 20S to 60N) it ranges from 1-4% for L-CH4 to 2-8% for P-O3. A remaining question (not resolved with the data sets assembled here) is the year-to-year variance of basin-wide reactivities perhaps associated with the El Nino – Southern Oscillation.

The six models' 1D probability distributions for O_3, CO, NOx, and HCHO over the tropical mid-Pacific basin are shown in Figure 5 and simple statistics (mean ±standard deviation) are

presented in the Supplement Table S5. Modeled data is sampled on the native grid of each model and not interpolated. This approach readily allows us to compare different models. Both 1D and 2D distributions presented here are sorted into 20 log-spaced bins per each factor of 10 (decade) in abundance (ppb or ppt). The dashes in the upper/lower rows of Fig. 5 indicate widths of these bins on each plot. For example, NOx distributions cover more than 3 decades (very small dashes), while the CO covers less than a decade (wide dashes). In the first row labeled "AIR", each grid cell is weighted by its size in moles, and thus the plot shows Petamoles per logarithmic bin. In each subsequent row, the cells are weighted by the reactivity (L-CH4, L-O3, P-O3) in moles/day, plotting thus Megamoles per day per bin.

The AIR plots show clear model differences. Models A and B have much greater frequency of O_3 occurrence from 50 – 150 ppb, and half the models (B, D, E) show a reasonable frequency of O_3 at 10 ppb and less, as might be expected in the tropical Pacific boundary layer (Kley et al., 1996; Singh et al., 1996; Nicely et al., 2016). For CO, model A shows unusually low abundances. For NOx, models C and F lack the NOx below 2.5 ppt that others have. The models are quite similar for HCHO, except for D, which has an unusually symmetric distribution and much lower abundances. When reactivity weighted, new features are found. Note that the area under the AIR-weighted curve is the same for all models, but the area in reactivity-weighted 1D plots is each model's total reactivity (moles/day). Model D has lower values overall for L-CH4 compared with the other models, but it is similar or even slightly higher for L-O3 and P-O3. The high-O_3 abundances in A remain equally important when weighted by any reactivity, but those in B become less important for L-CH4 and L-O3, but even more important for P-O3. This unusual feature adds a new dimension to diagnosing and understanding model differences. The reactivity weighting of the CO distribution does not show anything unusual. The NOx 1D plots show that L-CH4 is more heavily weighed to low NOx values than is L-O3, but P-O3 is weighted strongly to the higher NOx abundances (>10 ppt) as expected. The HCHO reactivity weights in the opposite direction with high abundances (>200 ppt) favoring L-CH4 and L-O3 but lower ones favoring P-O3, probably because the lower ones are from the upper troposphere where colder temperatures suppress both L-CH4 and L-O3 (Fig. 3). The results from the full probability distribution (Fig. 5) are mostly represented in the central statistics of Table S5. The reactivity weighting adds a new dimension to the diagnostics, and after the ATom data set

becomes available it would be productive to make a more detailed comparison that identifies the location and other key species controlling these shifts in reactivity.

These new diagnostics do not instantly identify the cause of model differences, but they do add a new dimension. For example, if we seek to understand why model D is different, we can look at global budgets: both models A & D have P-O3 and L-O3 tropospheric means between 2.5 and 3.5 ppb/day; whereas the other 4 models have values between 1.0 and 2.0. The global L-CH4 -- 0.50 to 0.65 ppb/day -- is similar for all models, with D in the middle. So globally models B & D are similar, but in the mid-Pacific, they are distinct with model D having much lower L-CH4 values in the tropics and especially the lower tropics (Fig. 3, see also Fig. S1 of Naik et al., 2013b). CH_4 loss is a major source of HCHO in the unpolluted atmosphere and this may partly explain D's lower values of tropical HCHO compared with other models. Some of the reduced tropical reactivity in D may be caused by more low clouds in the tropics, and this is apparent in the more rapid fall off in J-O_3(^{1}D) compared with other models (Fig. 4); yet models B & D (not A & D as found in L-O3 and P-O3) have much higher values of J-O_3(^{1}D). With the ATom A-run approach we will be able to remove differences caused by the widely ranging chemical climatologies of species (e.g., seen in Fig. 5, 6, 8) and more directly trace the range of results to the models' basic photolysis and kinetics.

The 2D distributions simply weighted by AIR show remarkable structures that differ significantly across the models, as shown in Figure 6, with summary statistics in Table S6. All 2D plots use the same 20-per-decade log scale as in the 1D analysis, and they are normalized such that if all parcels are distributed uniformly within a 20x20 square (e.g., 0.1-1.0 ppb HOOH, 10-100 ppt NOx) the arbitrary density value would be 1 (a yellow-green color in Figures 6-7). Thus the reactivity-weighted 2D plots are renormalized and do not reflect the individual model's total reactivity. In Figure 6a the AIR-weighted NOx-HOOH plots show a boomerang structure with greatly varying degrees of concentration about some points in the center (reddish regions). For example, models A and D show a very diffuse distribution with a much wider spread in HOOH values at lower NOx. Even for the four models with a central (NOx, HOOH)-line defining a peak frequency of occurrence, this line occurs at different locations. The O_3-H_2O density plots (Fig. 6b) show examples of highly standard and well-measured species with

extreme distributions: O_3 fall within 1 decade throughout most of the troposphere, but H_2O easily spans 3. Several show the bimodality of many parcels with low O_3 with high H_2O (marine boundary layer and above) and a second peak at higher O_3 and dry. For example, C and E look very much alike, but B has these two peaks more separated, and E has a much broader spread in upper tropospheric O_3 abundances.

Simple statistics for the probability distributions in Figure 6 are presented in the Supplement Tables S6abcd. Comparisons of the 1D distributions show that the log-normal distribution in mole fraction (mean μ and standard deviation σ) as represented by ($\mu - \sigma$, μ, $\mu + \sigma$) is for the most part very close to the equivalent percentile distribution (16^{th} %, 50^{th} %, 84^{th} %). For 2D summary statistics, we introduce a fitted ellipse centered at the mean value centroid (X_0, Y_0) with semi-major and semi-minor axes defined as the standard deviation in the two orthogonal axes (σ_X, σ_Y) rotated to find the flattest ellipse (i.e., maximum of σ_X/σ_Y). The values of centroid, semi-major/minor axes, and the degree of rotation are given in Tables S6 for all plots in Figure 6. An example showing a fitted ellipse on top of the 2D probability distribution is given in Figure S1, and the ellipses for all 6 model distributions in Figure 6a and 6b are plotted together in Figures S2a and S2b, respectively. These ellipses can provide a more direct and simple comparison of the central distributions of the models and support the discussion of Figure 6 above.

These plots ~~are only tropical but~~ include all altitudes that can be sampled by the ATom flights. When comparisons with ATom data are made, it will be useful to identify discrepancies in the 2D plots by separating altitude regions~~with altitude or other features~~.

The 2D plots can change the emphasis of certain regions when weighted by reactivity. For example, we take the GMI-CTM modeled NOx-HOOH density (Fig 6a panel E) and show the reactivity weightings in Figure 7. With AIR weighting, the quasi-boomerang has a strong central line with a negative slope. With P-O3, a much broader range is seen and the peak occurrence shifts to lower HOOH values and somewhat even to lower NOx values. With L-CH4, the line disappears and a galaxy-like pattern widens the range of parcels, picking up lower NOx values in two spiral arms. The L-O3 weighting is similar to L-CH4, and differences are discernable only

in small features. Clearly, species other than NOx and HOOH determine the reactivity of parcels, and thus other 2D plots will add new information. We anticipate that ATom measurements will be plotted not only with AIR weightings but also with reactivities calculated for that air parcel with these models (Auvray et al., 2007).

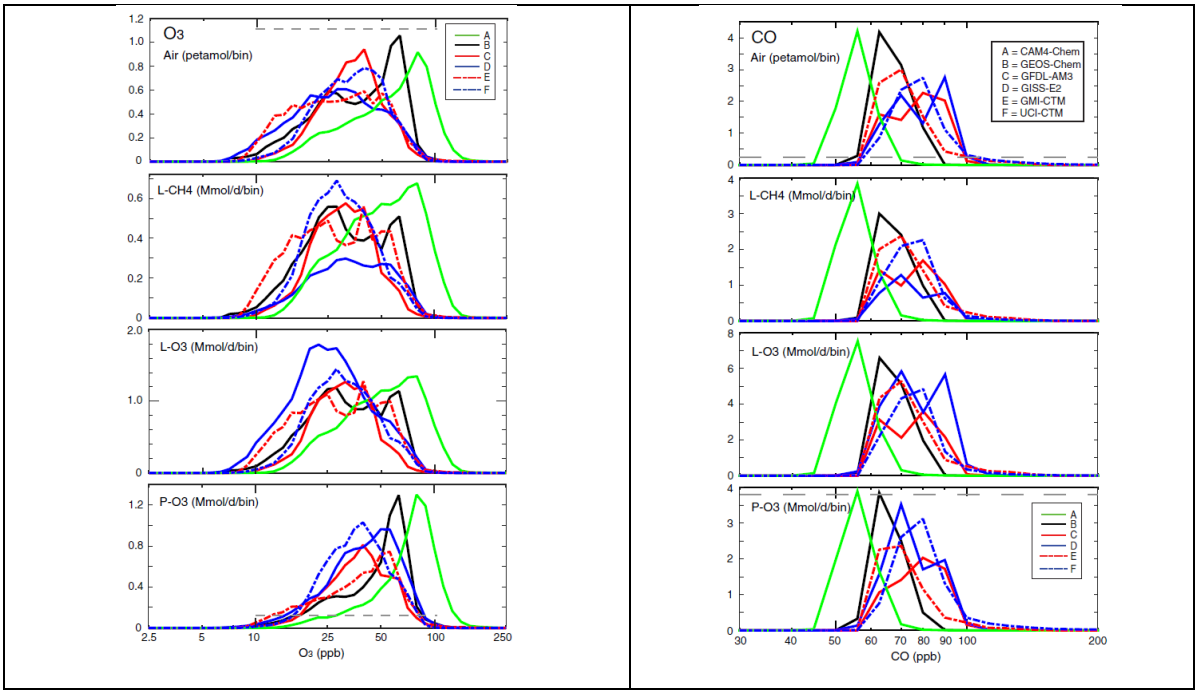
The 2D plots shown here intentionally included all air parcels over the mid-Pacific to ensure that a robust distribution was obtained (see Table 2). If we have only a single longitude slice as in ATom, then will these be so clearly defined? We examine this representativeness test by sub-sampling two models (C: GFDL-AM3 and F: UCI-CTM) at longitudes of 150E, 165E, 180E, 195E, and 210E in Figure 8 to compare with the average over the mid-Pacific domain. The densities are renormalized and show similar peaks and patterns, but of course there is more pixel-level noise and some differences. The transect at 150E is clearly less representative of the mid-Pacific, which is understandable since that longitude includes Papua New Guinea and eastern tropical Australia. Most importantly, the differences for 165E-210E are less than those across the 6 models (Fig. 7a). We need to develop an objective measure for comparing 2D plots between models and ATom measurements, and for judging if their differences are within the range of representativeness. Fortunately, the fitted ellipses provide a remarkably simple approach to evaluating the similarities and differences in these different transects and are plotted in Figure S8cd. For both models C and F we see that the central Pacific single transects (165E, 180E, 195E, 210E) with overlapping ellipses that even match the 150E-201E full block of data. In terms of overlapping area, the single transects overlap the full block at the 86-94% level, whereas the 150E transect is distinctly different with overlap of only 42% (F) and 63% (c). The full-block ellipse from the other model is plotted in Figure 8cd (dashed lines) to show that the models can be distinguished from even single transects (overlap of 60%).

Table 2. Representativeness of reactivities (L-CH4, L-O3, P-O3 all in ppb/day) averaged over 3 latitude and 3 pressure domains over the central Pacific (155E-233E). The first standard deviation (ppb/day) is over the different longitudinal transects on mid-August; and the second (%) is for the average across longitudes sampled over 31 days of August.

L-CH4 (ppb/day)	*60S-20S*	*20S-20N*	*20N-60N*
500-200 hPa	0.08 ± 0.02 ± 8%	0.36 ± 0.06 ± 7%	0.45 ± 0.08 ± 3%
850–500 hPa	0.28 ± 0.04 ± 6%	1.08 ± 0.07 ± 4%	1.26 ± 0.12 ± 2%
surf–850 hPa	0.35 ± 0.03 ± 4%	1.21 ± 0.13 ± 2%	1.44 ± 0.11 ± 1%
L-O3 (ppb/day)			
500-200 hPa	0.24 ± 0.03 ± 7%	0.74 ± 0.15 ± 8%	1.35 ± 0.27 ± 4%

850–500 hPa	0.69 ± 0.08 ± 6%	2.28 ± 0.13 ± 6%	3.01 ± 0.43 ± 3%
surf–850 hPa	0.86 ± 0.06 ± 4%	2.46 ± 0.32 ± 3%	2.60 ± 0.22 ± 3%
P-O3 (ppb/day)			
500-200 hPa	0.35 ± 0.04 ±10%	1.37 ± 0.18 ± 7%	1.78 ± 0.30 ± 4%
850–500 hPa	0.36 ± 0.06 ± 9%	0.92 ± 0.08 ± 8%	1.46 ± 0.13 ± 2%
surf–850 hPa	0.32 ± 0.20 ± 9%	0.43 ± 0.09 ± 2%	2.34 ± 0.33 ± 3%
Results are from the UCI CTM C-runs for 16 August and 1-31 August. The 155E-233E domain includes 69 longitudinal transects. All tropospheric grid cells in the domain are sampled equally and weighted by mass. ~~, but in stratospheric parcels the reactivity is not included.~~ The period 1-31 August shows trends in some domains as the sun moves southward, and this was removed with a line fit to calculate the standard deviation over the month. Results for the A-runs (not shown) differ in mean and standard deviation by a few percent.			

670



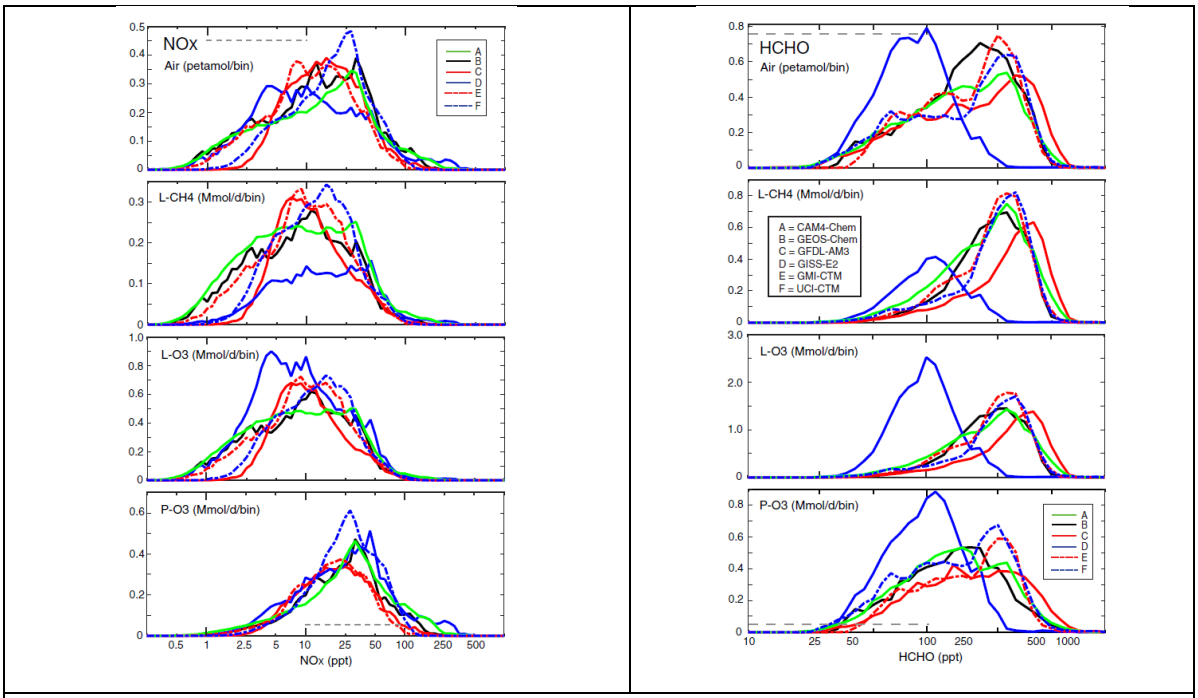


Figure 5. Six modeled 1D probability distributions for O_3, CO, NOx, and HCHO, where the air parcels have been weighted by air mass (row 1), L-CH4 (row 2), L-O3 (row 3), and P-O3 (row 4). The domain being sampled in the models is the tropical Pacific: 20S-20N, 150E-210E, 0.5-12km. Units for the air-weighting are petamoles per bin where the bins are set at 20 per decade (sizes marked by dashed lines in upper or lower panels) and Mmoles per bin per day for the reactivity weighted plots (rows 2-4).

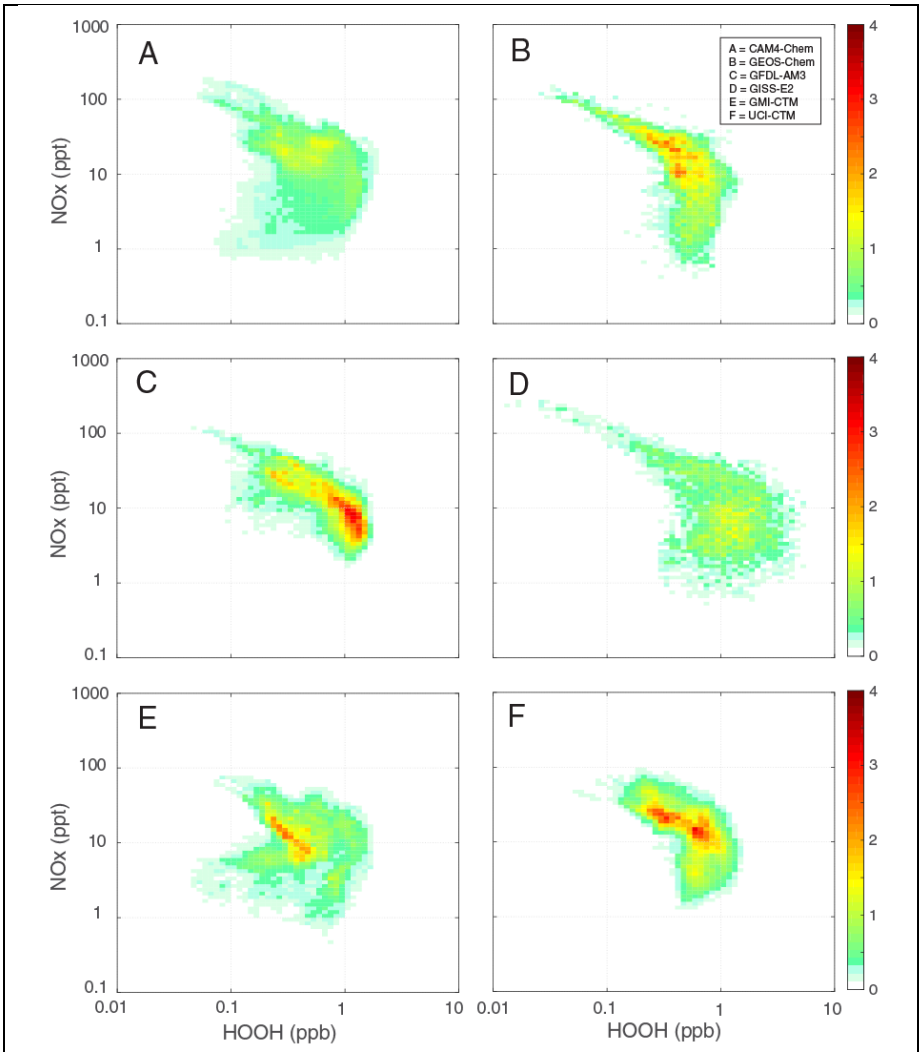


Figure 6a. Six modeled 2D probability distributions for NOx vs. HOOH as weighted by air mass. These are the initial chemical abundances for each model and hence the same for A- and C-runs. All grid cells were binned at 20 per decade in species abundance (mole fraction, ppt for NOx, ppb for HOOH). The density value for each plot is scaled so that a uniform distribution over exactly one decade in both species would give the yellow-green color of 1.0. The domain being sampled in the models is the tropical Pacific: 20S-20N, 150E-210E, 0.5-12km. Model A = CAM4-Chem; B = GEOS-Chem; C = GFDL-AM3; D = GISS-E2; E = GMI-CTM; F = UCI-CTM.

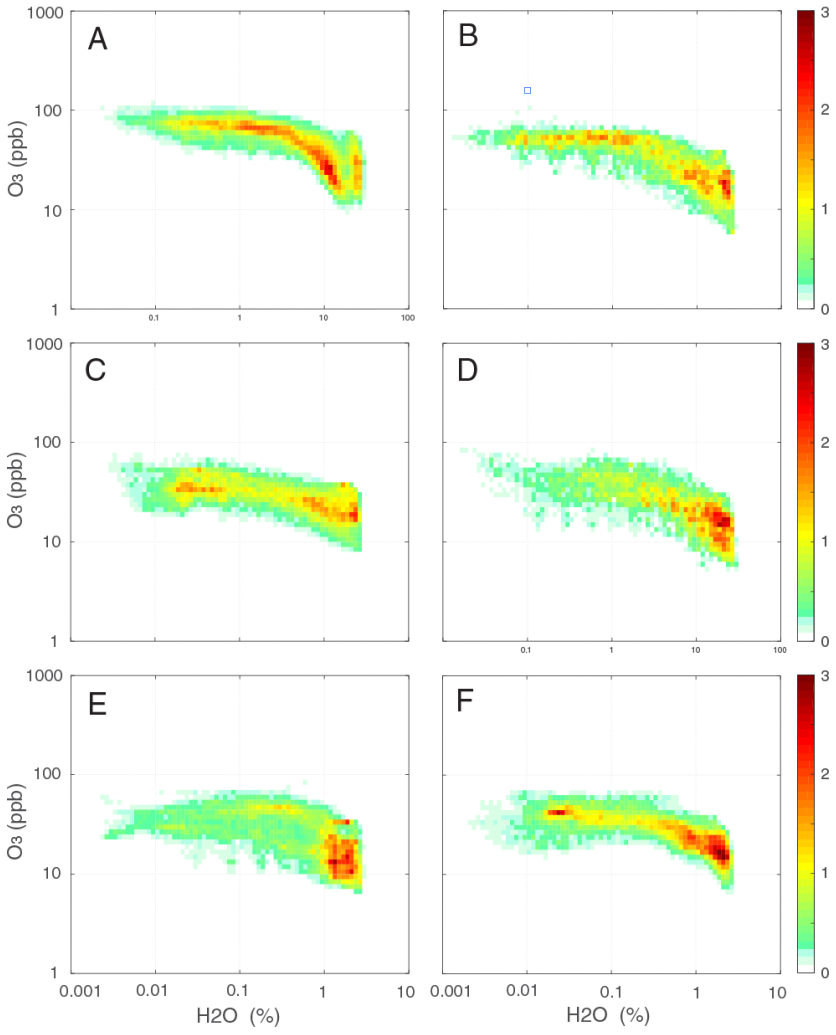


Figure 6b. Six modeled 2D probability distributions for O_3 vs H_2O as weighted by air mass. This color-bar scale differs slightly from other 2D plots. See Fig 6a.

675

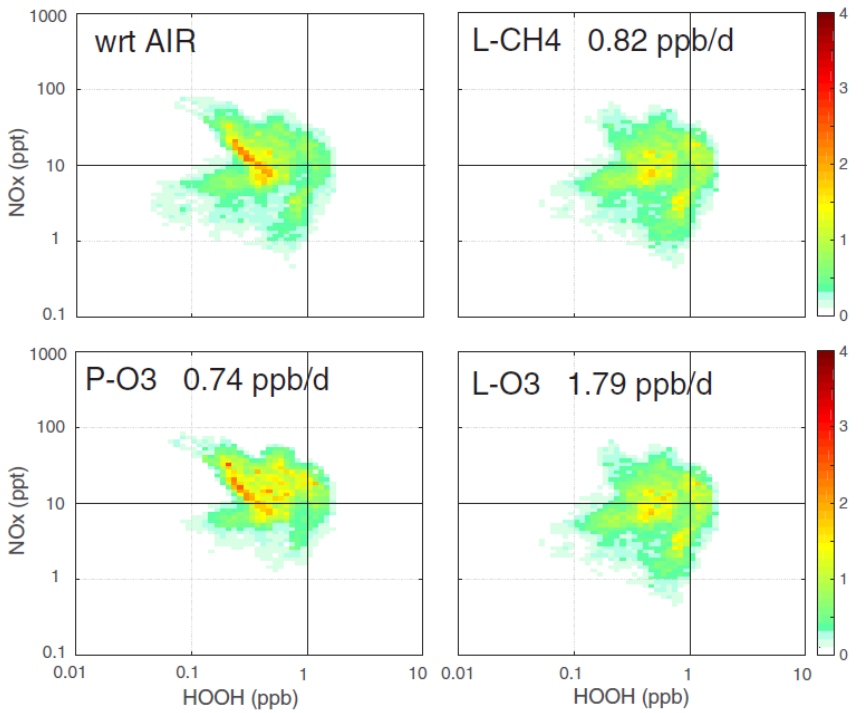


Figure 7. Model E (GMI CTM) 2D probability distributions from A-run for NOx vs. HOOH as weighted by air mass, L-CH4, L-O3 and P-O3. The domain being sampled is 20S-20N, 150E-210E, 0.5-12km, see Fig 6a.

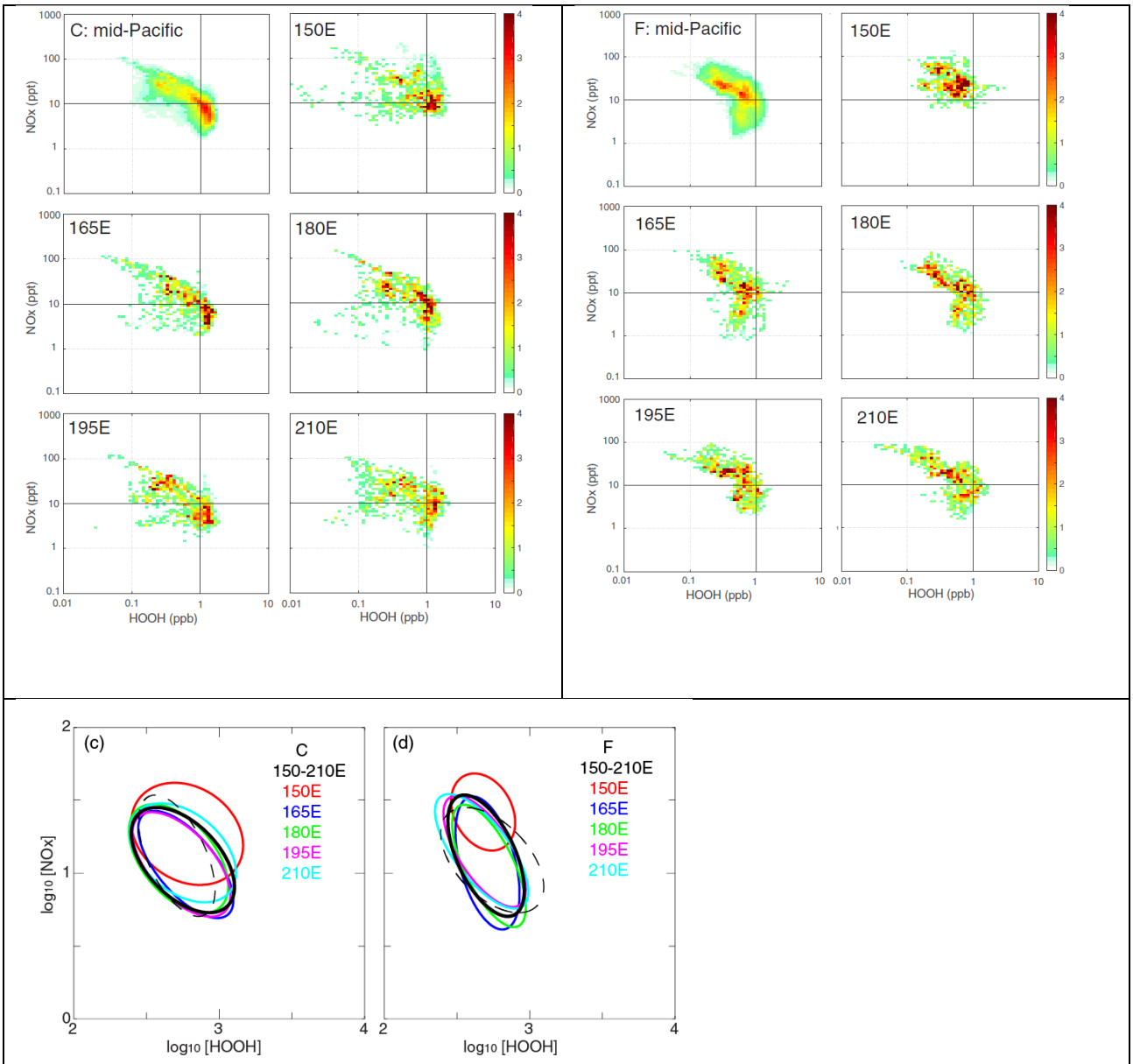


Figure 8. AIR-weighted 2D probability distributions for NOx vs. HOOH averaged over tropical ~~mid-~~Pacific block (150E-210E, 20S-20N, 0-12 km) and for ~~at~~ different single ~~-~~ longitude transects from 150E to 210E, shown for models (a) C (GFDL-AM3) & (b) F (UCI-CTM). ~~The region sampled is 20S-20N, 0.5-12km.~~The fitted 2D ellipses are shown for the full block (thick black line) and 5 longitude transects (colored lines) for models (c) C and (d) F. The block ellipse for the other model is shown as a thin black dashed line.

5. *Discussion and preparation for the ATom data set*

This paper is based on the underlying premise ~~of the NASA ATom missions~~ that high-frequency profiling ~~measurements~~ of the key species controlling the daily-average reactivity of individual air parcels ~~can be made on regular profiling flights~~ throughout the remote ~~middle of the Pacific and Atlantico~~ Ocean basins can ~~, and that these air parcels will~~ provide a unique, objectively sampled chemical climatology identifying those air parcels that matter, i.e., are most important in controlling methane and tropospheric ozone. Such data will further ~~This data set will~~ provide the most rigorous testing and diagnosis of the global chemistry models, in particular~~ly~~ the chemistry-climate models, which require a climatology~~, not a campaign data set~~.

Here we ~~have~~ present a six-model comparison using this new approach. We outline~~d~~ the model development (i.e., the A-runs) that ~~will~~ enables ~~the~~ global chemistry models to readily use high-frequency ~~the~~ measurements from aircraft campaigns like the NASA ATom mission ~~measurements and~~to calculate the chemical reactivity in individual~~the~~ parcels ~~(i.e., the A-runs)~~and over chemical regimes. The multi-model comparison has already identified some commonalities and highlighted several differences among the models in their calculation of tropospheric ozone and methane tendencies. For models that are outliers in particular diagnostics, it is a challenge for them to identify the cause within their own model and perhaps explain why the more common results are the ones in error. A test of these models, isolating the photochemical module by using A-runs with the same string of simulated measurements, is underway.

The multi-model comparison has provided a range of scientific results:

- All 6 models show distinctly different reactivity profiles in the Pacific basin, with model-model differences much larger than the A-run and C-run differences; models that look similar in one reactivity can appear different in another (e.g., L-CH4 in B & C vs. L-O3).
- It is hard to find a consistent pattern in P-O3; we attribute these model differences to wide variations in NOx abundances over the remote Pacific.
- J-values in the tropics, particularly J-O(1D) (reaction 3d), differ widely across the 6 models; this is unexpected considering the general agreement with photolysis model

comparisons (PhotoComp, 2010) and indicates that implementation of the photolysis codes in different models may be inconsistent.

- Probability distributions for the tropics show robust differences with clear outliers; different models are singled out for different species (model A for CO & O3, D for HCHO and NOx); and surprisingly the water vapor shows a large range across models.
- Reactivity-weighted probability distributions show shifts that might be expected, based on L-CH4 and L-O3 occurring primarily in the lower mid-troposphere and P-O3 occurring near the surface and in the upper troposphere; however, not all models show the same shift, implying a very different distribution of reactivity and/or dependence on the key species.
- Representativeness, specifically the ability of a few Pacific transects to provide a chemical climatology for the entire basin, was tested extensively in model F for average reactivities across different longitudes and days and showed modest variability; when compared in terms of 2D probability densities and fitted ellipses, two models showed that longitudinal transects from 165E to 210E were nearly identical, yet distinct from the other model.

The 1D and 2D probability distributions of key species~~, shown here as a sample,~~ are sufficiently diverse across the models that ~~ATom~~ climatology measurements, like those from ATom, will ~~clearly~~ easily be able to differentiate among them and likely identify specific model discrepancies. For example, in Figure 6a models A and E are alone in identifying a population of parcels with low-HOOH that also have low-NOx. If this is not found in the observations, then we have some clues (also looking at other key variables like in Figure 6b) that will identify locations and processes. Further, by looking at the reactivity of these parcels (Figure 7), we can find that this region is important for methane and ozone loss. Some work remains in establishing just how close is good enough in matching 2D (and multi-D) probability distributions of the key species, although the overlap of the 2D fitted ellipses begins to address this.

There are other ATom measurements beyond just key species that might prove useful as climatological tests for the models. The OH loss frequency (L-OH, Sinha et al. (2008); Mao et al. (2009)) is primarily determined by the longer-lived reactive species listed here, can be derived

from the key species, but it is not really a product of the 3D models. Effectively, L-OH provides a climatology of a weighted basket of species. The models' predicted L-OH using their own key species could be tested with the L-OH observations, but then we are just testing the model's key species and our direct comparisons are more useful. Actinic fluxes and thus J-values are being measured by ATom and can be analyzed on a case-by-case basis (Palancar et al., 2011) to assess the role of clouds in determining instantaneous reactivity. To be useful as a climatolog yical, ~~test of~~ the models would need to develop statistics on how the observed J-values (with clouds) deviated from clear-sky (modeled) values, thus checking if the photolysis effect of the cloud statistics in the models is similar that that observed. In this case ATom is probably one of the only useful datasets because flight plans were made independent of clear or cloudy conditions (except for aircraft safety).~~24-hour modeled J-values, one would need to build up 3D cloud statistics from the measured J-values over the large range of solar zenith angles and altitudes that ATom samples.~~ At present there is no clear path to use either L-OH or J's ~~these~~ to improve the climatologies of L-CH4, L-O3, and P-O3.

ATom involves four ~~The first~~ deployments: ~~of four,~~ ATom-1~~,~~ completed in August 2016; ~~, with~~ ATom-2 completed in February ~~flights to begin in January~~ 2017; ATom-3 scheduled for October 2017; and ATom-4 completes in May 2018. ATom was successful in completing all flights with instruments working, and acquiring well over 90% of the proposed data set, and measuring more than 30,000 10-second air parcels A quick look at the pre-ATom planned flight tracks and sampling in Figure 9 shows the coverage of the ocean basins, the large numbers of profiles, and the sampling frequency as a function of altitude. The expected release of ATom-1 data is mid-2017 and will include the global chemical model products discussed here. These measurements and analysis will provide a new approach for understanding which air matters.

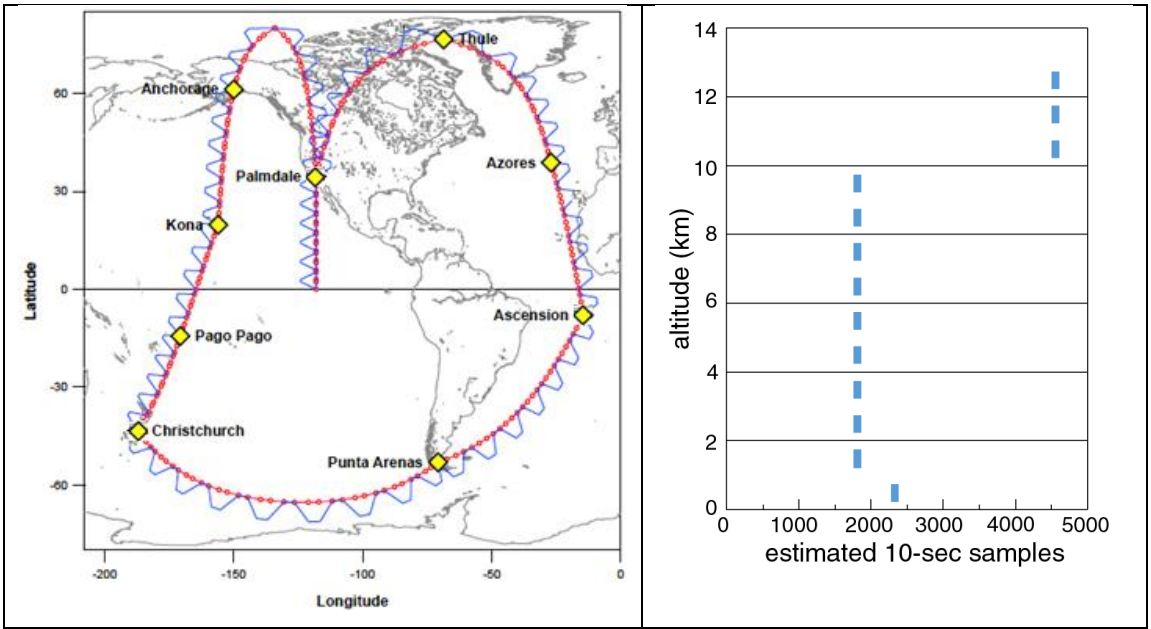


Figure 9. ATom proposed flight tracks (left) and estimated sampling frequency by 1-km altitude bins (right). The actual flights are somewhat altered. The altitude sampling is based on the proposed ~90 hours of flight time, ~180 profiles taking ~35 min for each pair of climb-descend, and 5 min spent each profile in the marine boundary layer. For up to date information on the ATom mission and deployments, see https://espo.nasa.gov/missions/atom/content/ATom.

Acknowledgments. This work was supported by NASA funding of the EVS2 Atmospheric Tomography (ATom) mission through a range of specific funding mechanisms to UC Irvine, NASAS GSFC, Columbia U, NCAR, and Harvard U.

References

Allen, D., Pickering, K., and Fox-Rabinovitz, M.: Evaluation of pollutant outflow and CO sources during TRACE-P using model-calculated, aircraft-based, and Measurements of Pollution in the Troposphere (MOPITT)-derived CO concentrations, J Geophys Res-Atmos, 109, Artn D15s03, 10.1029/2003jd004250, 2004.

Apel, E. C., Olson, J. R., Crawford, J. H., Hornbrook, R. S., Hills, A. J., Cantrell, C. A., Emmons, L. K., Knapp, D. J., Hall, S., Mauldin, R. L., Weinheimer, A. J., Fried, A., Blake, D. R., Crounse, J. D., St Clair, J. M., Wennberg, P. O., Diskin, G. S., Fuelberg, H. E., Wisthaler, A., Mikoviny, T., Brune, W., and Riemer, D. D.: Impact of the deep convection of isoprene and other reactive trace species on radicals and ozone in the upper troposphere, Atmos Chem Phys, 12, 1135-1150, 10.5194/acp-12-1135-2012, 2012.

Auvray, M., Bey, I., Llull, E., Schultz, M. G., and Rast, S.: A model investigation of tropospheric ozone chemical tendencies in long-range transported pollution plumes, J Geophys Res-Atmos, 112, Artn D05304, 10.1029/2006jd007137, 2007.

Barnes, E. A., and Fiore, A. M.: Surface ozone variability and the jet position: Implications for projecting future air quality, Geophys Res Lett, 40, 2839-2844, Doi 10.1002/Grl.50411, 2013.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q. B., Liu, H. G. 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.

Blake, N. J., Blake, D. R., Simpson, I. J., Meinardi, S., Swanson, A. L., Lopez, J. P., Katzenstein, A. S., Barletta, B., Shirai, T., Atlas, E., Sachse, G., Avery, M., Vay, S., Fuelberg, H. E., Kiley, C. M., Kita, K., and Rowland, F. S.: NMHCs and halocarbons in Asian continental outflow during the Transport and Chemical Evolution over the Pacific (TRACE-P) Field Campaign: Comparison with PEM-West B, J Geophys Res-Atmos, 108, Artn 8806, Doi 10.1029/2002jd003367, 2003.

Browell, E. V., Fenn, M. A., Butler, C. F., Grant, W. B., Brackett, V. G., Hair, J. W., Avery, M. A., Newell, R. E., Hu, Y. L., Fuelberg, H. E., Jacob, D. J., Anderson, B. E., Atlas, E. L., Blake, D. R., Brune, W. H., Dibb, J. E., Fried, A., Heikes, B. G., Sachse, G. W., Sandholm, S. T., Singh, H. B., Talbot, R. W., Vay, S. A., Weber, R. J., and Bartlett, K. B.: Large-scale ozone and aerosol distributions, air mass characteristics, and ozone fluxes over the western Pacific Ocean in late winter/early spring, J Geophys Res-Atmos, 108, Artn 8805, doi 10.1029/2002jd003290, 2003.

Charlton-Perez, C. L., Evans, M. J., Marsham, J. H., and Esler, J. G.: The impact of resolution on ship plume simulations with NOx chemistry, Atmos Chem Phys, 9, 7505-7518, 10.5194/acp-9-7505-2009, 2009.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles, in: Climate Change 2013: The Physical Science Basis, IPCC WGI Contribution to the Fifth Assessment Report, edited by: Stocker, T. F., Qin, D., and al., e., Cambridge, United Kingdom, 465-570, 2013.

Collins, W.J., et al.: AerChemMIP - Quantifying the effects of chemistry and aerosols in CMIP6, gmd-2016-139, 2016.

Crawford, J., Olson, J., Davis, D., Chen, G., Barrick, J., Shetter, R., Lefer, B., Jordan, C., Anderson, B., Clarke, A., Sachse, G., Blake, D., Singh, H., Sandolm, S., Tan, D., Kondo, Y., Avery, M., Flocke, F., Eisele, F., Mauldin, L., Zondlo, M., Brune, W., Harder, H., Martinez, M., Talbot, R., Bandy, A., and Thornton, D.: Clouds and trace gas distributions during TRACE-P, J Geophys Res-Atmos, 108, Artn 8818, Doi 10.1029/2002jd003177, 2003.

Dacre, H. F., Clark, P. A., Martinez-Alvarado, O., Stringer, M. A., and Lavers, D. A.: How Do Atmospheric Rivers Form?, B Am Meteorol Soc, 96, 1243-1255, 10.1175/Bams-D-14-00031.1, 2015.

Damoah, R., Spichtinger, N., Forster, C., James, P., Mattis, I., Wandinger, U., Beirle, S., Wagner, T., and Stohl, A.: Around the world in 17 days - hemispheric-scale transport of forest fire smoke from Russia in May 2003, Atmos Chem Phys, 4, 1311-1321, 2004.

Davis, D. D., Chen, G., Crawford, J. H., Liu, S., Tan, D., Sandholm, S. T., Jing, P., Cunnold, D. M., DiNunno, B., Browell, E. V., Grant, W. B., Fenn, M. A., Anderson, B. E., Barrick, J. D., Sachse, G. W., Vay, S. A., Hudgins, C. H., Avery, M. A., Lefer, B., Shetter, R. E., Heikes, B. G., Blake, D. R., Blake, N., Kondo, Y., and Oltmans, S.: An assessment of western North Pacific ozone photochemistry based on springtime observations from NASA's PEM-West B (1994) and TRACE-P (2001) field studies, J Geophys Res-Atmos, 108, Artn 8829, 10.1029/2002jd003232, 2003.

Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., Golaz, J. C., Ginoux, P., Lin, S. J., Schwarzkopf, M. D., Austin, J., Alaka, G., Cooke, W. F., Delworth, T. L., Freidenreich, S. M., Gordon, C. T., Griffies, S. M., Held, I. M., Hurlin, W. J., Klein, S. A., Knutson, T. R., Langenhorst, A. R., Lee, H. C., Lin, Y. L., Magi, B. I., Malyshev, S. L., Milly, P. C. D., Naik, V., Nath, M. J., Pincus, R., Ploshay, J. J., Ramaswamy, V., Seman, C. J., Shevliakova, E., Sirutis, J. J., Stern, W. F., Stouffer, R. J., Wilson, R. J., Winton, M., Wittenberg, A. T., and Zeng, F. R.: The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3, J Climate, 24, 3484-3519, Doi 10.1175/2011jcli3955.1, 2011.

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, 2007.

Eastham, S. D., Weisenstein, D. K., and Barrett, S. R. H.: Development and evaluation of the unified tropospheric-stratospheric chemistry extension (UCX) for the global chemistry-transport model GEOS-Chem, Atmos Environ, 89, 52-63, 10.1016/j.atmosenv.2014.02.001, 2014.

Eckstein, J., Ruhnke, R., Zahn, A., Neumaier, M., Kirner, O., and Braesicke, P.: An assessment of the climatological representativeness of IAGOS-CARIBIC trace gas measurements using EMAC model simulations, Atmos. Chem. Phys. Disc, 1617, 2775-2794, doi: 0.5194/acp-17-2775-201710.5194/acp-2016-179, 20176.

Ehhalt, D. H., Rohrer, F., Kraus, A. B., Prather, M. J., Blake, D. R., and Rowland, F. S.: On the significance of regional trace gas distributions as derived from aircraft campaigns in PEM-West A and B, J Geophys Res-Atmos, 102, 28333-28351, 1997.

Elguindi, N., Clark, H., Ordonez, C., Thouret, V., Flemming, J., Stein, O., Huijnen, V., Moinat, P., Inness, A., Peuch, V. H., Stohl, A., Turquety, S., Athier, G., Cammas, J. P., and Schultz, M.: Current status of the ability of the GEMS/MACC models to reproduce the tropospheric CO vertical distribution as measured by MOZAIC, Geosci Model Dev, 3, 501-518, 10.5194/gmd-3-501-2010, 2010.

Engel, A., Bonisch, H., Brunner, D., Fischer, H., Franke, H., Gunther, G., Gurk, C., Hegglin, M., Hoor, P., Konigstedt, R., Krebsbach, M., Maser, R., Parchatka, U., Peter, T., Schell, D., Schiller, C., Schmidt, U., Spelten, N., Szabo, T., Weers, U., Wernli, H., Wetter, T., and Wirth, V.: Highly resolved observations of trace gases in the lowermost stratosphere and upper troposphere from the Spurt project: an overview, Atmos Chem Phys, 6, 283-301, 2006.

Esler, J. G.: An integrated approach to mixing sensitivities in tropospheric chemistry: A basis for the parameterization of subgrid-scale emissions for chemistry transport models, J Geophys Res-Atmos, 108, Artn 4632, 10.1029/2003jd003627, 2003.

Fang, Y. Y., Mauzerall, D. L., Liu, J. F., Fiore, A. M., and Horowitz, L. W.: Impacts of 21st century climate change on global air pollution-related premature mortality, Climatic Change, 121, 239-253, 10.1007/s10584-013-0847-8, 2013.

Fiore, A. M., Naik, V., Spracklen, D. V., Steiner, A., Unger, N., Prather, M., Bergmann, D., Cameron-Smith, P. J., Cionni, I., Collins, W. J., Dalsoren, S., Eyring, V., Folberth, G. A., Ginoux, P., Horowitz, L. W., Josse, B., Lamarque, J. F., MacKenzie, I. A., Nagashima, T., O'Connor, F. M., Righi, M., Rumbold, S. T., Shindell, D. T., Skeie, R. B., Sudo, K., Szopa, S., Takemura, T., and Zeng, G.: Global air quality and climate, Chem Soc Rev, 41, 6663-6683, Doi 10.1039/C2cs35095e, 2012.

Fishman, J., Hoell, J. M., Bendura, R. D., McNeil, R. J., and Kirchhoff, V. W. J. H.: NASA GTE TRACE A experiment (September October 1992): Overview, J Geophys Res-Atmos, 101, 23865-23879, Doi 10.1029/96jd00123, 1996.

Gaudel, A., Clark, H., Thouret, V., Jones, L., Inness, A., Flemming, J., Stein, O., Huijnen, V., Eskes, H., Nedelec, P., and Boulanger, D.: On the use of MOZAIC-IAGOS data to assess the ability of the MACC reanalysis to reproduce the distribution of ozone and CO in the UTLS over Europe, Tellus B, 67, Artn 27955, 10.3402/Tellusb.V67.27955, 2015.

Hardacre, C., Wild, O., and Emberson, L.: An evaluation of ozone dry deposition in global scale chemistry climate models, Atmos Chem Phys, 15, 6419-6436, 10.5194/acp-15-6419-2015, 2015.

Heald, C. L., Jacob, D. J., Fiore, A. M., Emmons, L. K., Gille, J. C., Deeter, M. N., Warner, J., Edwards, D. P., Crawford, J. H., Hamlin, A. J., Sachse, G. W., Browell, E. V., Avery, M. A., Vay, S. A., Westberg, D. J., Blake, D. R., Singh, H. B., Sandholm, S. T., Talbot, R. W., and Fuelberg, H. E.: Asian outflow and trans-Pacific transport of carbon monoxide and ozone pollution: An integrated satellite, aircraft, and model perspective, J Geophys Res-Atmos, 108, Artn 4804, doi 10.1029/2003jd003507, 2003.

Hecobian, A., Liu, Z., Hennigan, C. J., Huey, L. G., Jimenez, J. L., Cubison, M. J., Vay, S., Diskin, G. S., Sachse, G. W., Wisthaler, A., Mikoviny, T., Weinheimer, A. J., Liao, J., Knapp, D. J., Wennberg, P. O., Kurten, A., Crounse, J. D., St Clair, J., Wang, Y., and Weber, R. J.: Comparison of chemical characteristics of 495 biomass burning plumes intercepted by the NASA DC-8 aircraft during the ARCTAS/CARB-2008 field campaign, Atmos Chem Phys, 11, 13325-13337, 10.5194/acp-11-13325-2011, 2011.

Hoell, J. M., Davis, D., Liu, S. C., Newell, R., Shipham, M., Akimoto, H., McNeal, R. J., Bendura, R. J., and Drewry, J. W.: Pacific exploratory Mission-West A (PEM-West A): September October 1991, J Geophys Res-Atmos, 101, 1641-1653, Doi 10.1029/95jd00622, 1996.

Hoell, J. M., Davis, D. D., Jacob, D. J., Rodgers, M. O., Newell, R. E., Fuelberg, H. E., McNeal, R. J., Raper, J. L., and Bendura, R. J.: Pacific Exploratory Mission in the tropical Pacific: PEM-Tropics A, August-September 1996, J Geophys Res-Atmos, 104, 5567-5583, Doi 10.1029/1998jd100074, 1999.

Holmes, C. D., Prather, M. J., Sovde, O. A., and Myhre, G.: Future methane, hydroxyl, and their uncertainties: key climate and emission parameters for future predictions, Atmos Chem Phys, 13, 285-302, DOI 10.5194/acp-13-285-2013, 2013.

Holmes, C. D., Prather, M. J., and Vinken, G. C. M.: The climate impact of ship NOx emissions: an improved estimate accounting for plume chemistry, Atmos Chem Phys, 14, 6801-6812, 10.5194/acp-14-6801-2014, 2014.

Hoor, P., Fischer, H., Lange, L., Lelieveld, J., and Brunner, D.: Seasonal variations of a mixing layer in the lowermost stratosphere as identified by the CO-O-3 correlation from in situ measurements, J Geophys Res-Atmos, 107, Artn 4044, 10.1029/2000jd000289, 2002.

Hsu, J., Prather, M. J., Wild, O., Sundet, J. K., Isaksen, I. S. A., Browell, E. V., Avery, M. A., and Sachse, G. W.: Are the TRACE-P measurements representative of the western Pacific during March 2001?, J Geophys Res-Atmos, 109, -, Artn D02314, doi 10.1029/2003jd004002, 2004.

Hsu, J., and Prather, M. J.: Stratospheric variability and tropospheric ozone, J Geophys Res-Atmos, 114, Artn D06102, Doi 10.1029/2008jd010942, 2009.

Jacob, D. J., Crawford, J. H., Kleb, M. M., Connors, V. S., Bendura, R. J., Raper, J. L., Sachse, G. W., Gille, J. C., Emmons, L., and Heald, C. L.: Transport and Chemical Evolution over the Pacific (TRACE-P) aircraft mission: Design, execution, and first results, J Geophys Res-Atmos, 108, 1-19, Artn 9000, 10.1029/2002jd003276, 2003.

Jacob, D. J., and Winner, D. A.: Effect of climate change on air quality, Atmos Environ, 43, 51-63, DOI 10.1016/j.atmosenv.2008.09.051, 2009.

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, 10.5194/acp-10-5191-2010, 2010.

Kiley, C. M., Fuelberg, H. E., Palmer, P. I., Allen, D. J., Carmichael, G. R., Jacob, D. J., Mari, C., Pierce, R. B., Pickering, K. E., Tang, Y. H., Wild, O., Fairlie, T. D., Logan, J. A., Sachse, G. W., Shaack, T. K., and Streets, D. G.: An intercomparison and evaluation of aircraft-derived and simulated CO from seven chemical transport models during the TRACE-P experiment, J Geophys Res-Atmos, 108, Artn 8819, 10.1029/2002jd003089, 2003.

Kley, D., Crutzen, P. J., Smit, H. G. J., Vomel, H., Oltmans, S. J., Grassl, H., and Ramanathan, V.: Observations of near-zero ozone concentrations over the convective Pacific: Effects on air chemistry, Science, 274, 230-233, DOI 10.1126/science.274.5285.230, 1996.

Koppe, M., Hermann, M., Brenninkmeijer, C. A. M., Heintzenberg, J., Schlager, H., Schuck, T., Slemr, F., Sprung, D., van Velthoven, P. F. J., Wiedensohler, A., Zahn, A., and Ziereis, H.: Origin of aerosol particles in the mid-latitude and subtropical upper troposphere and lowermost stratosphere from cluster analysis of CARIBIC data, Atmos Chem Phys, 9, 8413-8430, 2009.

Kunz, A., Schiller, C., Rohrer, F., Smit, H. G. J., Nedelec, P., and Spelten, N.: Statistical analysis of water vapour and ozone in the UT/LS observed during SPURT and MOZAIC, Atmos Chem Phys, 8, 6603-6615, 2008.

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.

Lelieveld, J., and Crutzen, P. J.: The Role of Clouds in Tropospheric Photochemistry, J Atmos Chem, 12, 229-267, Doi 10.1007/Bf00048075, 1991.

Larsen, M. L., Briner, C. A., and Boehner, P.: On the Recovery of 3D Spatial Statistics of Particles from 1D Measurements: Implications for Airborne Instruments, J Atmos Ocean Tech, 31, 2078-2087, 10.1175/Jtech-D-14-00004.1, 2014.

Lawrence, M. G., Jockel, P., and von Kuhlmann, R.: What does the global mean OH concentration tell us?, Atmos Chem Phys, 1, 37-49, 2001.

Li, J. Y., Mao, J. Q., Min, K. E., Washenfelder, R. A., Brown, S. S., Kaiser, J., Keutsch, F. N., Volkamer, R., Wolfe, G. M., Hanisco, T. F., Pollack, I. B., Ryerson, T. B., Graus, M., Gilman, J. B., Lerner, B. M., Warneke, C., de Gouw, J. A., Middlebrook, A. M., Liao, J., Welti, A., Henderson, B. H., McNeill, V. F., Hall, S. R., Ullmann, K., Donner, L. J., Paulot, F., and Horowitz, L. W.: Observational constraints on glyoxal production from isoprene oxidation and its contribution to organic aerosol over the Southeast United States, J Geophys Res-Atmos, 121, 9849-9861, 10.1002/2016JD025331, 2016.

Manney, G. L., Bird, J. C., Donovan, D. P., Duck, T. J., Whiteway, J. A., Pal, S. R., and Carswell, A. I.: Modeling ozone laminae in ground-based Arctic wintertime observations using trajectory calculations and satellite data, J Geophys Res-Atmos, 103, 5797-5814, Doi 10.1029/97jd03449, 1998.

Mao, J., Ren, X., Brune, W. H., Olson, J. R., Crawford, J. H., Fried, A., Huey, L. G., Cohen, R. C., Heikes, B., Singh, H. B., Blake, D. R., Sachse, G. W., Diskin, G. S., Hall, S. R., and Shetter, R. E.: Airborne measurement of OH reactivity during INTEX-B, Atmos Chem Phys, 9, 163-173, 10.5194/acp-9-163-2009, 2009.

Marenco, A., Thouret, V., Nédélec, P., Smit, H., Helten, M., Kley, D., Karcher, F., Simon, P., Law, K., Pyle, J., Poschmann, G., Von Wrede, R., Hume, C., and Cook, T.: Measurement of ozone and water vapor by Airbus in-service aircraft: The MOZAIC airborne program, an overview, Journal of Geophysical Research: Atmospheres, 103, 25631-25642, 10.1029/98jd00977, 1998.

Mickley, L. J., Jacob, D. J., Field, B. D., and Rind, D.: Effects of future climate change on regional air pollution episodes in the United States, Geophys. Res. Lett., 31, L24103, 10.1029/2004gl021216, 2004.

Mundhenk, B. D., Barnes, E. A., and Maloney, E. D.: All-Season Climatology and Variability of Atmospheric River Frequencies over the North Pacific, J Climate, 29, 4885-4903, 10.1175/Jcli-D-15-0655.1, 2016.

Naik, V., Horowitz, L. W., Fiore, A. M., Ginoux, P., Mao, J. Q., Aghedo, A. M., and Levy, H.: Impact of preindustrial to present-day changes in short-lived pollutant emissions on atmospheric composition and climate forcing, J Geophys Res-Atmos, 118, 8086-8110, 10.1002/jgrd.50608, 2013a.

Naik, V., Voulgarakis, A., Fiore, A. M., Horowitz, L. W., Lamarque, J. F., Lin, M., Prather, M. J., Young, P. J., Bergmann, D., Cameron-Smith, P. J., Cionni, I., Collins, W. J., Dalsoren, S. B., Doherty, R., Eyring, V., Faluvegi, G., Folberth, G. A., Josse, B., Lee, Y. H., MacKenzie, I. A., Nagashima, T., van Noije, T. P. C., Plummer, D. A., Righi, M., Rumbold, S. T., Skeie, R., Shindell, D. T., Stevenson, D. S., Strode, S., Sudo, K., Szopa, S., and Zeng, G.: Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos Chem Phys, 13, 5277-5298, DOI 10.5194/acp-13-5277-2013, 2013b.

Nappo, C. J., Caneill, J. Y., Furman, R. W., Gifford, F. A., Kaimal, J. C., Kramer, M. L., Lockhart, T. J., Pendergast, M. M., Pielke, R. A., Randerson, D., Shreffler, J. H., and Wyngaard, J. C.: The Workshop on the Representativeness of Meteorological-Observations, June 1981, Boulder, Colo., B Am Meteorol Soc, 63, 761-764, 1982.

Newell, R. E., Newell, N. E., Zhu, Y., and Scott, C.: Tropospheric Rivers - a Pilot-Study, Geophys Res Lett, 19, 2401-2404, Doi 10.1029/92gl02916, 1992.

Newell, R. E., V, T., Cho, J. Y. N., Stoller, P., Marenco, A., and Smit, H. G.: Ubiquity of quasi-horizontal layers in the troposphere, Nature, 398, 316-319, Doi 10.1038/18642, 1999.

Nicely, J. M., Anderson, D. C., Canty, T. P., Salawitch, R. J., Wolfe, G. M., Apel, E. C., Arnold, S. R., Atlas, E. L., Blake, N. J., Bresch, J. F., Campos, T. L., Dickerson, R. R., Duncan, B., Emmons, L. K., Evans, M. J., Fernandez, R. P., Flemming, J., Hall, S. R., Hanisco, T. F., Honomichl, S. B., Hornbrook, R. S., Huijnen, V., Kaser, L., Kinnison, D. E., Lamarque, J. F., Mao, J. Q., Monks, S. A., Montzka, D. D., Pan, L. L., Riemer, D. D., Saiz-Lopez, A., Steenrod, S. D., Stell, M. H., Tilmes, S., Turquety, S., Ullmann, K., and Weinheimer, A. J.: An observationally constrained evaluation of the oxidative capacity in the tropical western Pacific troposphere, J Geophys Res-Atmos, 121, 7461-7488, 10.1002/2016JD025067, 2016.

Nicely, J. M., Salawitch, R. J., Canty, T., Anderson, D. C., Arnold, S. R., Chipperfield, M. P., Emmons, L. K., Flemming, J., Huijnen, V., Kinnison, D. E., Lamarque, J. F., Mao, J. Q., Monks, S. A., Steenrod, S. D., Tilmes, S., and Turquety, S.: Quantifying the causes of differences in tropospheric OH within global models, J Geophys Res-Atmos, 122, 1983-2007, 10.1002/2016JD026239, 2017.

Olson, J. R., Crawford, J. H., Chen, G., Fried, A., Evans, M. J., Jordan, C. E., Sandholm, S. T., Davis, D. D., Anderson, B. E., Avery, M. A., Barrick, J. D., Blake, D. R., Brune, W. H., Eisele, F. L., Flocke, F., Harder, H., Jacob, D. J., Kondo, Y., Lefer, B. L., Martinez, M., Mauldin, R. L., Sachse, G. W., Shetter, R. E., Singh, H. B., Talbot, R. W., and Tan, D.: Testing fast photochemical theory during TRACE-P based on measurements of OH, HO2, and CH2O, J Geophys Res-Atmos, 109, Artn D15s10, Doi 10.1029/2003jd004278, 2004.

Olson, J. R., Crawford, J. H., Brune, W., Mao, J., Ren, X., Fried, A., Anderson, B., Apel, E., Beaver, M., Blake, D., Chen, G., Crounse, J., Dibb, J., Diskin, G., Hall, S. R., Huey, L. G., Knapp, D., Richter, D., Riemer, D., Clair, J. S., Ullmann, K., Walega, J., Weibring, P., Weinheimer, A., Wennberg, P., and Wisthaler, A.: An analysis of fast photochemistry over high northern latitudes during spring and summer using in-situ observations from ARCTAS and TOPSE, Atmos. Chem. Phys., 12, 6799-6825, 10.5194/acp-12-6799-2012, 2012.

Palancar, G. G., Shetter, R. E., Hall, S. R., Toselli, B. M., and Madronich, S.: Ultraviolet actinic flux in clear and cloudy atmospheres: model calculations and aircraft-based measurements, Atmos Chem Phys, 11, 5457-5469, DOI 10.5194/acp-11-5457-2011, 2011.

Pan, L., E. Atlas, R. Salawitch, S. Honomichl, J. Bresch, W. Randel, E. Apel, R. Hornbrook, A. Weinheimer, D. Anderson, S. Andrews, S. Baidar, and S. Beaton, T. C., L. Carpenter, D. Chen, B. Dix, V. Donets, S. Hall, T. Hanisco, C. Homeyer, L. Huey, J. Jensen, L. Kaser, D. Kinnison, T. Koenig, J. Lamarque, C. Liu, J. Luo, Z. Luo, D. Montzka, J. Nicely, R. Pierce, D. Riemer, T. Robinson, P. Romashkin, A. Saiz-Lopez, S. Schauffler, O. Shieh, M. Stell, K. Ullmann, G. Vaughan, R. Volkamer, and G. Wolfe: The Convective Transport of Active Species in the Tropics (CONTRAST) Experimen, Bull. Amer. Meteor. Soc., doi:10.1175/BAMS-D-14-00272.1, 2017.

Pan, L. L., Wei, J. C., Kinnison, D. E., Garcia, R. R., Wuebbles, D. J., and Brasseur, G. P.: A set of diagnostics for evaluating chemistry-climate models in the extratropical tropopause region, J Geophys Res-Atmos, 112, Artn D09316, 10.1029/2006jd007792, 2007.

Parrington, M., Palmer, P. I., Lewis, A. C., Lee, J. D., Rickard, A. R., Di Carlo, P., Taylor, J. W., Hopkins, J. R., Punjabi, S., Oram, D. E., Forster, G., Aruffo, E., Moller, S. J., Bauguitte, S. J. B., Allan, J. D., Coe, H., and Leigh, R. J.: Ozone photochemistry in boreal biomass burning plumes, Atmos Chem Phys, 13, 7321-7341, 10.5194/acp-13-7321-2013, 2013.

Paulson, S. E., and Orlando, J. J.: The reactions of ozone with alkenes: An important source of HOx in the boundary layer, Geophys Res Lett, 23, 3727-3730, Doi 10.1029/96gl03477, 1996.

PhotoComp: Chapter 6 - Stratospheric Chemistry SPARC Report No. 5 on the Evaluation of Chemistry-Climate Models 194-202, 2010.

Pisso, I., Real, E., Law, K. S., Legras, B., Bousserez, N., Attie, J. L., and Schlager, H.: Estimation of mixing in the troposphere from Lagrangian trace gas reconstructions during long-range pollution plume transport, J Geophys Res-Atmos, 114, Artn D19301, 10.1029/2008jd011289, 2009.

Prather, M., and Jaffe, A. H.: Global Impact of the Antarctic Ozone Hole - Chemical Propagation, J Geophys Res-Atmos, 95, 3473-3492, 1990.

Prather, M., Gauss, M., Berntsen, T., Isaksen, I., Sundet, J., Bey, I., Brasseur, G., Dentener, F., Derwent, R., Stevenson, D., Grenfell, L., Hauglustaine, D., Horowitz, L., Jacob, D., Mickley, L., Lawrence, M., von Kuhlmann, R., Muller, J.-F., Pitari, G., Rogers, H., Johnson, M., Pyle, J., Law, K., van Weele, M., and Wild, O.: Fresh air in the 21st century?, Geophys. Res. Lett., 30, 1100, 10.1029/2002gl016285, 2003.

Prather, M. J.: Time scales in atmospheric chemistry: Theory, GWPs for CH4 and CO, and runaway growth, Geophys Res Lett, 23, 2597-2600, 1996.

Prather, M. J.: Tropospheric O_3 from photolysis of O_2, Geophys Res Lett, 36, L03811, Artn L03811, Doi 10.1029/2008gl036851, 2009.

Prather, M. J., Holmes, C. D., and Hsu, J.: Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry, Geophys Res Lett, 39, Artn L09803, doi 10.1029/2012gl051440, 2012.

Prather, M. J., and Holmes, C. D.: A perspective on time: loss frequencies, time scales and lifetimes, Environmental Chemistry, 10, 73-79, Doi 10.1071/En13017, 2013.

Prather, M. J.: Photolysis rates in correlated overlapping cloud fields: Cloud-J 7.3c, Geosci Model Dev, 8, 2587-2595, 10.5194/gmd-8-2587-2015, 2015.

Ramsey, C. A., and Hewitt, A. D.: A methodology for assessing sample representativeness, Environ Forensics, 6, 71-75, 10.1080/15275920590913877, 2005.

Raper, J. L., Kleb, M. M., Jacob, D. J., Davis, D. D., Newell, R. E., Fuelberg, H. E., Bendura, R. J., Hoell, J. M., and McNeal, R. J.: Pacific Exploratory Mission in the tropical Pacific: PEM-Tropics B, March-April 1999, J Geophys Res-Atmos, 106, 32401-32425, Doi 10.1029/2000jd900833, 2001.

Reid, S. J., Rex, M., Von der Gathen, P., Floisand, I., Stordal, F., Carver, G. D., Beck, A., Reimer, E., Kruger-Carstensen, R., De Haan, L. L., Braathen, G., Dorokhov, V., Fast, H., Kyro, E., Gil, M., Litynska, Z., Molyneux, M., Murphy, G., O'Connor, F., Ravegnani, F., Varotsos, C., Wenger, J., and Zerefos, C.: A study of ozone laminae using diabatic trajectories, contour advection and photochemical trajectory model simulations., J Atmos Chem, 30, 187-207, Doi 10.1023/A:1005836212979, 1998.

Rohrer, F., and Berresheim, H.: Strong correlation between levels of tropospheric hydroxyl radicals and solar ultraviolet radiation, Nature, 442, 184-187, 10.1038/nature04924, 2006.

Schmidt, G. A., Kelley, M., Nazarenko, L., Ruedy, R., Russell, G. L., Aleinov, I., Bauer, M., Bauer, S. E., Bhat, M. K., Bleck, R., Canuto, V., Chen, Y. H., Cheng, Y., Clune, T. L., Del Genio, A., de Fainchtein, R., Faluvegi, G., Hansen, J. E., Healy, R. J., Kiang, N. Y., Koch, D., Lacis, A. A., LeGrande, A. N., Lerner, J., Lo, K. K., Matthews, E. E., Menon, S., Miller, R. L., Oinas, V., Oloso, A. O., Perlwitz, J. P., Puma, M. J., Putman, W. M., Rind, D., Romanou, A., Sato, M., Shindell, D. T., Sun, S., Syed, R. A., Tausnev, N., Tsigaridis, K., Unger, N., Voulgarakis, A., Yao, M. S., and Zhang, J. L.: Configuration and assessment of the GISS ModelE2

contributions to the CMIP5 archive, Journal of Advances in Modeling Earth Systems, 6, 141-184, 10.1002/2013MS000265, 2014.

Schnell, J. L., Prather, M. J., Josse, B., Naik, V., Horowitz, L. W., Cameron-Smith, P., Bergmann, D., Zeng, G., Plummer, D. A., Sudo, K., Nagashima, T., Shindell, D. T., Faluvegi, G., and Strode, S. A.: Use of North American and European air quality networks to evaluate global chemistry-climate modeling of surface ozone, Atmos Chem Phys, 15, 10581-10596, 10.5194/acp-15-10581-2015, 2015.

Schoeberl, M. R., Ziemke, J. R., Bojkov, B., Livesey, N., Duncan, B., Strahan, S., Froidevaux, L., Kulawik, S., Bhartia, P. K., Chandra, S., Levelt, P. F., Witte, J. C., Thompson, A. M., Cuevas, E., Redondas, A., Tarasick, D. W., Davies, J., Bodeker, G., Hansen, G., Johnson, B. J., Oltmans, S. J., Vomel, H., Allaart, M., Kelder, H., Newchurch, M., Godin-Beekmann, S., Ancellet, G., Claude, H., Andersen, S. B., Kyro, E., Parrondos, M., Yela, M., Zablocki, G., Moore, D., Dier, H., von der Gathen, P., Viatte, P., Stubi, R., Calpini, B., Skrivankova, P., Dorokhov, V., de Backer, H., Schmidlin, F. J., Coetzee, G., Fujiwara, M., Thouret, V., Posny, F., Morris, G., Merrill, J., Leong, C. P., Koenig-Langlo, G., and Joseph, E.: A trajectory-based estimate of the tropospheric ozone column using the residual method, J Geophys Res-Atmos, 112, Artn D24s49, 10.1029/2007jd008773, 2007.

Shindell, D. T., Pechony, O., Voulgarakis, A., Faluvegi, G., Nazarenko, L., Lamarque, J. F., Bowman, K., Milly, G., Kovari, B., Ruedy, R., and Schmidt, G. A.: Interactive ozone and methane chemistry in GISS-E2 historical and future climate simulations, Atmos Chem Phys, 13, 2653-2689, DOI 10.5194/acp-13-2653-2013, 2013.

Singh, H. B., Gregory, G. L., Anderson, B., Browell, E., Sachse, G. W., Davis, D. D., Crawford, J., Bradshaw, J. D., Talbot, R., Blake, D. R., Thornton, D., Newell, R., and Merrill, J.: Low ozone in the marine boundary payer of the tropical Pacific Ocean: Photochemical loss, chlorine atoms, and entrainment, J Geophys Res-Atmos, 101, 1907-1917, Doi 10.1029/95jd01028, 1996.

Singh, H. B., Viezee, W., Chen, Y., Bradshaw, J., Sandholm, S., Blake, D., Blake, N., Heikes, B., Snow, J., Talbot, R., Browell, E., Gregory, G., Sachse, G., and Vay, S.: Biomass burning influences on the composition of the remote South Pacific troposphere: analysis based on observations from PEM-Tropics-A, Atmos Environ, 34, 635-644, Doi 10.1016/S1352-2310(99)00380-5, 2000.

Sinha, V., Williams, J., Crowley, J. N., and Lelieveld, J.: The comparative reactivity method - a new tool to measure total OH reactivity in ambient air, Atmos Chem Phys, 8, 2213-2227, 2008.

Sovde, O. A., Prather, M. J., Isaksen, I. S. A., Berntsen, T. K., Stordal, F., Zhu, X., Holmes, C. D., and Hsu, J.: The chemical transport model Oslo CTM3, Geosci Model Dev, 5, 1441-1469, DOI 10.5194/gmd-5-1441-2012, 2012.

Spivakovsky, C. M., Logan, J. A., Montzka, S. A., Balkanski, Y. J., Foreman-Fowler, M., Jones, D. B. A., Horowitz, L. W., Fusco, A. C., Brenninkmeijer, C. A. M., Prather, M. J., Wofsy, S. C., and McElroy, M. B.: Three-dimensional climatological distribution of tropospheric OH: Update and evaluation, J Geophys Res-Atmos, 105, 8931-8980, 2000.

Stevenson, D. S., Dentener, F. J., Schultz, M. G., Ellingsen, K., van Noije, T. P. C., Wild, O., Zeng, G., Amann, M., Atherton, C. S., Bell, N., Bergmann, D. J., Bey, I., Butler, T., Cofala, J., Collins, W. J., Derwent, R. G., Doherty, R. M., Drevet, J., Eskes, H. J., Fiore, A. M., Gauss, M., Hauglustaine, D. A., Horowitz, L. W., Isaksen, I. S. A., Krol, M. C., Lamarque, J. F., Lawrence, M. G., Montanaro, V., Muller, J. F., Pitari, G., Prather, M. J., Pyle, J. A., Rast, S., Rodriguez, J. M., Sanderson, M. G., Savage, N. H., Shindell, D. T., Strahan, S. E., Sudo, K., and Szopa, S.: Multimodel ensemble simulations of present-day and near-future tropospheric ozone, J Geophys Res-Atmos, 111, -, Artn D08301, Doi 10.1029/2005jd006338, 2006.

Stevenson, D. S., Young, P. J., Naik, V., Lamarque, J. F., Shindell, D. T., Voulgarakis, A., Skeie, R. B., Dalsoren, S. B., Myhre, G., Berntsen, T. K., Folberth, G. A., Rumbold, S. T., Collins, W. J., MacKenzie, I. A., Doherty, R. M., Zeng, G., van Noije, T. P. C., Strunk, A., Bergmann, D., Cameron-Smith, P., Plummer, D. A., Strode, S. A., Horowitz, L., Lee, Y. H., Szopa, S., Sudo, K., Nagashima, T., Josse, B., Cionni, I., Righi, M., Eyring, V., Conley, A., Bowman, K. W., Wild, O., and Archibald, A.: Tropospheric ozone changes, radiative forcing and attribution to emissions in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos Chem Phys, 13, 3063-3085, DOI 10.5194/acp-13-3063-2013, 2013.

Stoller, P., Cho, J. Y. N., Newell, R. E., Thouret, V., Zhu, Y., Carroll, M. A., Albercook, G. M., Anderson, B. E., Barrick, J. D. W., Browell, E. V., Gregory, G. L., Sachse, G. W., Vay, S., Bradshaw, J. D., and Sandholm, S.: Measurements of atmospheric layers from the NASA DC-8 and P-3B aircraft during PEM-Tropics A, J Geophys Res-Atmos, 104, 5745-5764, Doi 10.1029/98jd02717, 1999.

Stone, D., Whalley, L. K., and Heard, D. E.: Tropospheric OH and HO2 radicals: field measurements and model comparisons, Chem Soc Rev, 41, 6348-6404, 10.1039/c2cs35140d, 2012.

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, 2007.

Thuburn, J., and Tan, D. G. H.: A parameterization of mixdown time for atmospheric chemicals, J Geophys Res-Atmos, 102, 13037-13049, Doi 10.1029/97jd00408, 1997.

Tilmes, S., Lamarque, J. F., Emmons, L. K., Kinnison, D. E., Marsh, D., Garcia, R. R., Smith, A. K., Neely, R. R., Conley, A., Vitt, F., Martin, M. V., Tanimoto, H., Simpson, I., Blake, D. R., and Blake, N.: Representation of the Community Earth System Model (CESM1) CAM4-chem within the Chemistry-Climate Model Initiative (CCMI), Geosci Model Dev, 9, 1853-1890, 10.5194/gmd-9-1853-2016, 2016.

Turner, A. J., Fiore, A. M., Horowitz, L. W., and Bauer, M.: Summertime cyclones over the Great Lakes Storm Track from 1860-2100: variability, trends, and association with ozone pollution, Atmos Chem Phys, 13, 565-578, DOI 10.5194/acp-13-565-2013, 2013.

Wild, O., Sundet, J. K., Prather, M. J., Isaksen, I. S. A., Akimoto, H., Browell, E. V., and Oltmans, S. J.: Chemical transport model ozone simulations for spring 2001 over the western Pacific: Comparisons with TRACE-P lidar, ozonesondes, and Total Ozone Mapping Spectrometer columns, J Geophys Res-Atmos, 108, Artn 8826, 10.1029/2002jd003283, 2003.

Wofsy, S. C., Team, H. S., Team, C. M., and Team, S.: HIAPER Pole-to-Pole Observations (HIPPO): fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, Philos T R Soc A, 369, 2073-2086, DOI 10.1098/rsta.2010.0313, 2011.

Wu, S. L., Mickley, L. J., Jacob, D. J., Logan, J. A., Yantosca, R. M., and Rind, D.: Why are there large differences between models in global budgets of tropospheric ozone?, J Geophys Res-Atmos, 112, Artn D05302, Doi 10.1029/2006jd007801, 2007.

Young, P. J., Archibald, A. T., Bowman, K. W., Lamarque, J. F., Naik, V., Stevenson, D. S., Tilmes, S., Voulgarakis, A., Wild, O., Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W. J., Dalsoren, S. B., Doherty, R. M., Eyring, V., Faluvegi, G., Horowitz, L. W., Josse, B., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Plummer, D. A., Righi, M., Rumbold, S. T., Skeie, R. B., Shindell, D. T., Strode, S. A., Sudo, K., Szopa, S., and Zeng, G.: Pre-industrial to end 21st century projections of tropospheric ozone from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos Chem Phys, 13, 2063-2090, DOI 10.5194/acp-13-2063-2013, 2013.

Supplement for acp-2016-1105 (revised)

Global Atmospheric Chemistry – Which Air Matters

Michael J. Prather[1], Xin Zhu[1], Clare M. Flynn[1], Sarah A. Strode[2,3], Jose M. Rodriguez[3], Stephen D. Steenrod[2,3], Junhua Liu[2,3], Jean-Francois Lamarque[4], Arlene M. Fiore[5], Larry W. Horowitz [6], Jingqiu Mao[7], Lee T. Murray[8], Drew T. Shindell[9], Steven C. Wofsy[10]

[1]Department of Earth System Science, University of California, Irvine, CA 92697-3100, USA
[2]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[3]Universities Space Research Association (USRA), GESTAR, Columbia, MD, USA
[4]Atmospheric Chemistry, Observations & Modelling Laboratory, National Center for Atmospheric Research, Boulder, CO 80301, USA
[5]Department of Earth and Environmental Sciences and Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York, USA
[6]Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration, Princeton, NJ, USA
[7]Geophysical Institute and Department of Chemistry, University of Alaska Fairbanks, Fairbanks, Alaska, USA
[8]Department of Earth and Environmental Sciences, University of Rochester, Rochester, NY 14627-0221 USA
[9]Nicholas School of the Environment, Duke University, Durham, NC, USA
[10]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, 02138 USA

This Supplement examines some simple statistics that may be useful in providing a quantitative synthesis of the full probability distributions (shown in the main paper) so that different models or weightings can be compared and judged to be similar or be given a measure of their difference. As in the main paper, the case here focusses on the tropical Pacific block (150E-210E, 20S-20N, 0-12 km) for one day in the middle of August. The model key is: A = CAM4-Chem; B = GEOS-Chem; C = GFDL-AM3; D = GISS-E2; E = GMI-CTM; F = UCI-CTM. The reactivities are calculated from the C-runs (standard CCM/CTM simulation). The table numbers in the supplement are keyed to the figure numbers in the main paper and thus are not sequential, but the supplement figures are numbered sequentially.

1D Probabilities. The distributions here are evaluated on log of abundance, expressed as mole fraction. The bin spacing is in units of 0.05 for $\log_{10}$ for all probability densities. In the right-hand columns of Tables S6a and S6b, we calculate the 1D probability statistics for the Pacific block for all 6 models for the species [HOOH], {NOx], [H_2O], and [O_3]. The mean values minus and plus 1 standard deviation are given in the upper part of each model-species cell; those of the 16^{th}-50^{th}-84^{th} percentiles are given immediately underneath to facilitate comparison. In Table S8a we give similar statistics comparing the pacific block with single transects, sampling along single longitudes within the block (150E, 165E, 180E, 195E, 210E). In this case we only show results for 2 models (C & F) and 2 species (HOOH & NOx). Note that results for the block itself in Table S8 are the same as in Table S6a.

For [HOOH], [NOx], and [O_3] the Gaussian statistics and the percentiles are almost identical, and thus the central 68% of the species distribution can be readily characterized. For [H_2O] the extremely large spread in abundance across 0-12 km altitude range results in the median being consistently larger than the mean by about 0.2 and the percentile spread (16^{th} to 84^{th} %ile) being noticeably narrower than ± σ.

The 1D statistics are useful, for example, in demonstrating that model D has about 2x higher HOOH than other models and this difference applies across the ±σ range. For NOx, all models have similar averages, but models C & F have almost 2x higher NOx at the 16^{th} percentile. Five models have a similar value for the 84^{th} %ile of NOx (1.5 = ~30 ppt), but model E is much lower (~20 ppt). Surprisingly, model A has notably higher O_3 across the range: a 50^{th} %ile value of ~45 ppt versus ~25 ppt for the other models. Among the other five models, D and E

have a much lower 16th %ile O_3 (~13 ppt vs ~17 ppt). Water vapor (H_2O) across the 0-12 km altitude range is expected to have a very large spread and indeed the 16th – 84th %ile spread ranges from 40x to 80x. Model C is a climate model with prognostic H_2O based on a full convection scheme. It has 2x less H_2O than the other models and a much narrower ±σ range matching its 16th – 84th %ile range.

A more complete list of the 1D probability density statistics, based on Figure 5, is presented in Tables S5ab. Six major gases are listed (HOOH, NOx, H_2O, O_3, CO, HCHO), and the effect of reactivity weighting is clearly seen. CO has a very narrow distribution and the mean does not change with reactivity weighting. What is clear here and better quantified is that model A's average CO is about 0.12 (log_{10}) less than the other 5 models. For HOOH the pattern provides clues to the model differences. For models ABCEF, the Air and PO3 weighting show similar statistics, but the LO3 and LCH4 weighting show higher average HOOH by about 0.10. This reflects the greater weighting of the lower troposphere (with greater HOOH) in LO3 & LCH4. As noted above in discussing Fig 6, model D has much larger HOOH, but it also has different shifts across the 4 weighting factors. For H_2O all models show that the average with PO3 drops about 0.2-0.3 (upper troposphere dominates PO3) and that it increases by about 0.3-0.5 with LO3 & LCH4 weighting. All models except D have similarly weighted averages for LO3 & LCH4, and this points to a dramatically different pattern for these two reactivities in model D. For HCHO, models ABCEF have similar patterns with an average and standard deviation of about 2.2 and 0.3 for both Air & PO3 weighting. This average increase to 2.3-2.4 with LO3 & LCH4 weighting but the standard deviation decreases to 0.2. Model D again stands out with a different response to the weightings and a lower HCHO abundance by 0.3 (factor of 2). For O_3, all models show a similar and small shift of the average with the different weightings, and model C is clearly and uniformly across ±σ higher than all other models, as noted in Fig. 6 discussion. NOx has highest ±σ spread (other than H2O) of about 1.0 (factor of 10 in mole fraction). All models agree that PO3 occurs in parcels with much higher NOx (by 0.13 to 0.31) and that LO3 & LCH4 occurs in parcels with lower NOx. Model D again shows that LO3 and LCH4 occur in different parcels. In comparing models C and F, one finds that both have similar amounts of NOx in Air & LO3 & LCH4 weighted parcels, but for PO3 model F has clearly higher NOx by 0.1 across the ±σ range.

2D Probabilities. The colored plots of 2D probability density are harder to characterize simply. There are a range of multivariate statistics methods (e.g., Hotelling T^2 test) that can test whether two different modeled 2D density plots are demonstrably different, but it is easy to see that if we have enough samples, then almost all of the plots here are statistically different. We need to gather some simple statistics from our 2D plots that are similar to the Gaussian statistics or percentiles of the 1D distributions.

We characterize the 2D densities with a fitted ellipse centered at the mean value (X_0, Y_0) with semi-major and semi-minor axes defined as the standard deviation in the two orthogonal axes. These axes are rotated until we find an optimum fit with flattest ellipse (i.e., the lowest ratio of semi-minor to semi-major axis). The long axis of the ellipse is defined by a rotation angle counterclockwise from the X axis. Figure S1 shows an example of how such an ellipse appears on top of the 2D colored density distribution (from Fig 6a model B). Generally, about 40-50% of the weighted parcels fall within this ellipse, which is as expected for two independent normal distributions (0.68 x 0.68). The 2D variance ($\sigma_X^2 + \sigma_Y^2$), calculated with respect to any (X,Y)-

axis located at the mean of the 2D probability distribution, is unchanged by rotation of the axes about the centroid and is defined here as σ_{2D}. All of these quantities, including the rotation angle, are given for the 6 models and the Pacific block in Tables S6a ((X, Y) = (HOOH, NOx)) and S6b ((X, Y) = (H_2O, O_3)).

Figures S2ab show the fitted ellipses for the 2D densities plotted in Figures 6ab. For (HOOH, NOx), models ABE have similar overlapping ellipses; whereas models CF have distinctly smaller but similarly angled and proportioned ellipses. Model D is distinct in terms of offset to higher HOOH values. For (H_2O, O_3), all models have similarly shaped and angled ellipses with rotation angles ranging only from 166 to 173 degrees. The difference in ellipse overlap is characterized simply by Y0, with models BCF having Y0 = 4.45±0.01, models DE having Y0 = 4.35±0.01, the previously identified outlier model A having Y0 = 4.64.

Tables 6cd give the area of each ellipse and the %-overlap with the other ellipses. We define overlap here to be symmetric (i.e., the overlapping area divided by the average of the two ellipse areas) so that a smaller ellipse fitting wholly within a larger ellipse does not result in 100% overlap. For (HOOH, NOx) the %-overlap shows clearly what we diagnosed subjectively from Figure S2a: the 3 %-overlap pairs for ABE and the one %-overlap for CF are all about 80%, much higher than for any other pairs, except for the self-overlap of 100%. The average %-overlap against the other 5 models, clearly identifies the outlier as model D. The question of whether the outlier is the more accurate model will await a similar comparison with observations. For (H_2O, O_3), this diagnostic shows that model A is the clear outlier and identifies the high-overlap pairs seen in the figure.

Figure 8 in the main paper looks at the fundamental uncertainty of representativeness by comparing the 2D densities for (HOOH, NOx) from the over-sampled Pacific block with the smaller samples along five longitudinal transects within the block (150E, 165E, 180E, 195E, 210E). As clearly seen in Figures 8ab for models CF, the single transect sampling is sparser and it is harder to see a clear pattern. The fitted ellipses to these transects in Figures 8cd are able to statistically characterize the distributions and show that the single transects accurately measure the block and that they also clearly distinguish the two models C and F as being similar, but different. The statistics from the Figure 8ab plots are given in Tables S8ab. The %-overlap is a compelling diagnostic: the one transect at 150E is clearly different from all others, and the other 4 transects overlap with modeled block ellipse at the 86% to 94% level. The 150E ellipse is clearly different when plotted (Fig. 8cd) and the %-overlap is only 40-60%, the same overlap between the two models. The 150E transect has the same longitude as Sydney and cuts through Papua New Guinea. This region is directly influenced by deep convection, lightning, and other continental sources, whereas the 165E-210E region looks similar, remote, and presumably representative of much of the Pacific basin. We conclude that fitted ellipses for comparing 2D probability densities is a valuable tool. Perhaps 3D+ distributions could be fitted in a similar manner to ellipsoids for analysis.

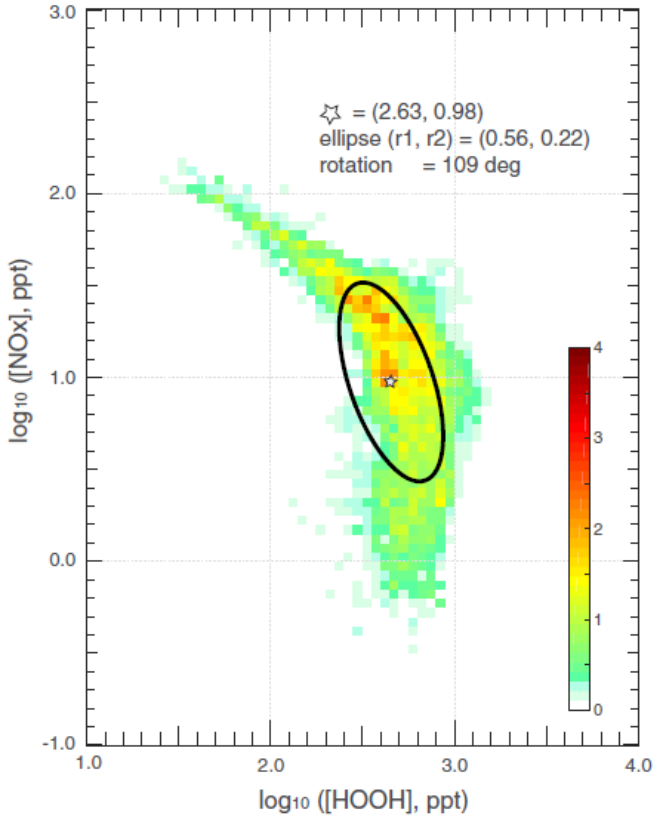


Figure S1. Model B (GEOS-Chem) air-mass weighted tropical Pacific distribution of parcels for the NOx vs. HOOH abundance showing the fitted ellipse. The species abundance is sampled in $\log_{10}$ space with a pixel resolution of 0.05 x 0.05. The color density pixels are the same as for model B in Figure 6a. For summary statistics on this distribution see Table S6a.

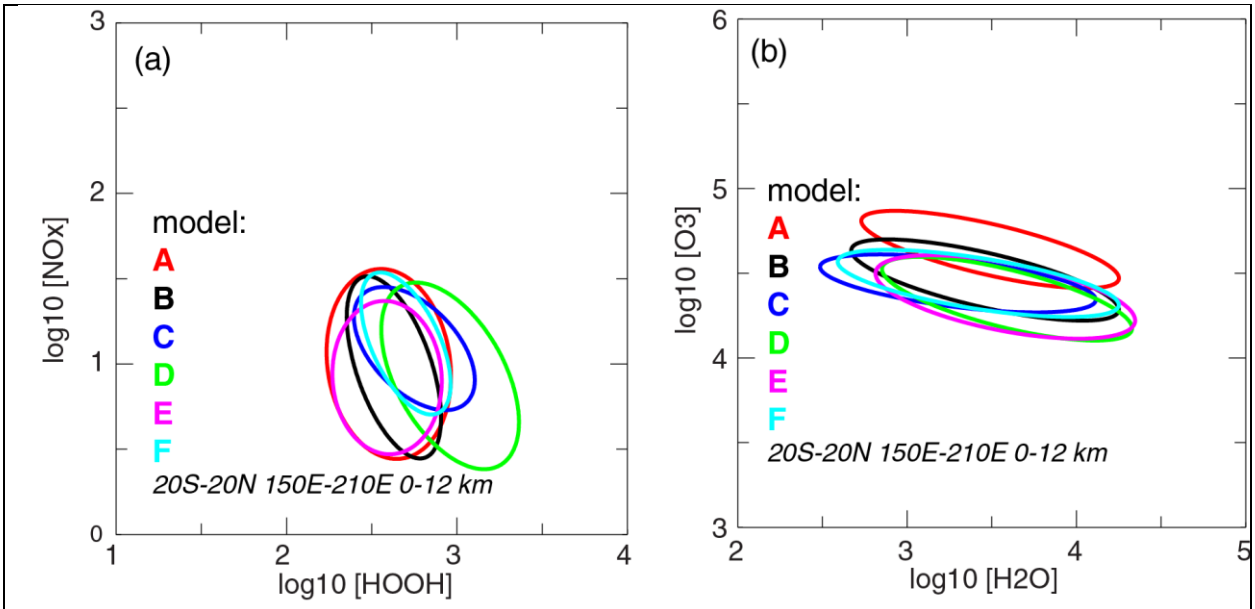


Figure S2. Fitted ellipses to the modeled 2D density distributions in (a) Figure 6a (HOOH, NOx) and (b) Figure 6b (H_2O, O_3). The data defining each ellipse and their %-overlap are given in Table S6ab.

140

Table S6a. Indicators derived from the 2D distribution of (HOOH, NOx) shown in Fig 6a. All values are based on units of $\log_{10}$([X, ppt]).

model	centroid (X_0,Y_0)	σ_{2D}	axes (major,minor)	rotation angle	HOOH X: -σ avg +σ X: 16-50-84 %	NOx Y: -σ avg +σ Y: 16-50-84 %
A	2.60, 1.00	0.665	0.559, 0.361	98	2.24 2.60 2.97 2.22 2.61 2.95	0.44 1.00 1.55 0.35 1.04 1.51
B	2.63, 0.98	0.603	0.560, 0.223	109	2.35 2.63 2.92 2.37 2.66 2.86	0.45 0.98 1.51 0.35 1.01 1.48
C	2.75, 1.09	0.507	0.441, 0.251	134	2.39 2.75 3.10 2.36 2.82 3.06	0.73 1.09 1.46 0.68 1.07 1.44
D	2.96, 0.93	0.682	0.603, 0.318	119	2.56 2.96 3.36 2.60 2.98 3.30	0.38 0.93 1.48 0.37 0.87 1.48
E	2.59, 0.92	0.550	0.448, 0.319	94	2.27 2.59 2.91 2.26 2.58 2.89	0.47 0.92 1.37 0.40 0.92 1.34
F	2.70, 1.12	0.491	0.451, 0.195	115	2.44 2.70 2.96 2.40 2.72 2.92	0.70 1.12 1.54 0.62 1.16 1.51

Table S6b. Indicators derived from the 2D distribution of (H_2O, O_3) shown in Fig 6b. All values are based on units of $\log_{10}$([X, ppt]).

model	centroid (X_0,Y_0)	σ_{2D}	axes (major,minor)	rotation angle	H_2O X: -σ avg +σ X: 16-50-84 %	O_3 Y: -σ avg +σ Y: 16-50-84 %
A	3.49, 4.64	0.799	0.775, 0.149	167	2.74 3.49 4.25 2.58 3.65 4.27	4.42 4.65 4.87 4.36 4.67 4.85
B	3.46, 4.46	0.824	0.808, 0.162	167	2.67 3.46 4.25 2.50 3.62 4.26	4.22 4.46 4.71 4.18 4.47 4.69
C	3.30, 4.44	0.839	0.817, 0.144	173	2.49 3.30 4.11 2.34 3.35 4.21	4.27 4.44 4.62 4.24 4.43 4.59
D	3.59, 4.35	0.780	0.761, 0.171	166	2.85 3.59 4.33 2.74 3.77 4.29	4.11 4.35 4.60 4.07 4.34 4.60
E	3.58, 4.36	0.803	0.778, 0.200	169	2.81 3.58 4.34 2.68 3.82 4.27	4.12 4.36 4.61 4.06 4.34 4.60
F	3.42, 4.44	0.858	0.835, 0.149	171	2.59 3.42 4.24 2.39 3.60 4.27	4.24 4.44 4.64 4.21 4.43 4.62

Table 6c. Overlap percent of SD2 ellipse fit (NOx vs HOOH).

model	area	x A	x B	x C	x D	x E	x F	x others
A	0.633	100%						64%
B	0.387	74%	100%					61%
C	0.346	59%	58%	100%				62%
D	0.603	42%	35%	58%	100%			43%
E	0.452	83%	77%	53%	34%	100%		60%
F	0.283	62%	62%	81%	48%	55%	100%	62%

Overlap % is defined as the overlapping area divided by the average of the two ellipse areas and is thus symmetric. The last column (x others) is the average overlap % against the other 5 models.

Table 6d. Overlap percent of SD2 ellipse fit (O_3 vs H_2O).

model	area	x A	x B	x C	x D	x E	x F	x others
A	0.368	100%						18%
B	0.407	30%	100%					65%
C	0.361	13%	72%	100%				62%
D	0.406	9%	68%	70%	100%			62%
E	0.490	16%	72%	70%	89%	100%		65%
F	0.396	20%	86%	86%	71%	77%	100%	68%

Table S8a. Indicators derived from the 2D distribution of (HOOH, NOx) as shown in Fig 8, including the SD2 ellipse. All values are based on units of $\log_{10}$([X, ppt]). The Block equals all grid cells 150E-201E, 20S-20N, 0-12 km. Individual longitudes include only a single transect sampled.

Model longitude	centroid (X_0,Y_0)	σ_{2D}	axes (major,minor)	rotation angle	HOOH X: -σ avg +σ X: 16-50-84 %	NOx Y: -σ avg +σ Y: 16-50-84 %
C Block	2.75, 1.09	0.507	0.441, 0.251	134	2.39 2.75 3.10 2.36 2.82 3.06	0.73 1.09 1.46 0.68 1.07 1.44
C 150E	2.78, 1.27	0.526	0.414, 0.324	146	2.40 2.78 3.17 2.43 2.86 3.09	0.91 1.27 1.62 0.89 1.16 1.61
C 165E	2.77, 1.06	0.497	0.454, 0.203	130	2.44 2.77 3.10 2.42 2.81 3.08	0.69 1.07 1.44 0.65 1.01 1.42
C 180E	2.72, 1.10	0.502	0.436, 0.248	130	2.39 2.72 3.06 2.35 2.81 3.02	0.73 1.10 1.47 0.67 1.08 1.42
C 195E	2.74, 1.06	0.494	0.444, 0.217	132	2.40 2.74 3.07 2.33 2.77 3.07	0.70 1.06 1.42 0.63 1.03 1.43
C 210E	2.75, 1.14	0.495	0.416, 0.268	140	2.39 2.75 3.11 2.31 2.81 3.08	0.80 1.14 1.48 0.77 1.13 1.45
F Block	2.70, 1.12	0.491	0.452, 0.195	115	2.44 2.70 2.96 2.40 2.72 2.92	0.70 1.12 1.54 0.62 1.16 1.51
F 150E	2.68, 1.42	0.346	0.281, 0.201	118	2.45 2.68 2.90 2.42 2.68 2.86	1.15 1.42 1.68 1.15 1.36 1.69
F 165E	2.71, 1.07	0.508	0.470, 0.194	105	2.49 2.71 2.94 2.44 2.72 2.90	0.61 1.07 1.52 0.52 1.09 1.52
F 180E	2.72, 1.05	0.490	0.460, 0.168	116	2.47 2.72 2.96 2.39 2.74 2.94	0.63 1.05 1.47 0.52 1.09 1.49
F 195E	2.69, 1.15	0.472	0.435, 0.182	123	2.41 2.69 2.97 2.37 2.73 2.93	0.77 1.15 1.53 0.68 1.20 1.50
F 210E	2.67, 1.15	0.504	0.466, 0.192	127	2.35 2.67 2.99 2.32 2.71 2.95	0.76 1.15 1.54 0.70 1.17 1.55

155

Table S8b. Overlap % of SD2 ellipse fit (NOx vs HOOH) for tropical Pacific (150E-210E, 20S-20N, 0-12 km) and single longitude transects shown in Fig.8.

	150E-210E	150E	165E	180E	195E	210E
Model C	*100%*	63%	87%	94%	91%	89%
Models C x F	60%					
Model F	*100%*	42%	88%	86%	90%	86%

Overlap % is defined as the overlapping area divided by the average of the two ellipse areas. The Models C x F denotes the overlap % of the two models for the ALL case of 150E-210E.

Table S5. 1D statistics (mean ± standard deviation) for probability densities shown in Fig 5.													
model		CO				HOOH				H_2O			
		Air	PO3	LO3	LCH4	Air	PO3	LO3	LCH4	Air	PO3	LO3	LCH4
	-1σ	4.69	4.69	4.68	4.68	2.30	2.28	2.43	2.42	2.79	2.51	3.29	3.28
A	**avg**	**4.72**	**4.72**	**4.72**	**4.72**	**2.66**	**2.64**	**2.76**	**2.75**	**3.54**	**3.25**	**3.87**	**3.87**
	+1σ	4.76	4.76	4.76	4.76	3.01	2.99	3.09	3.09	4.30	3.98	4.44	4.46
	-1σ	4.77	4.76	4.77	4.77	2.40	2.29	2.58	2.58	2.72	2.44	3.29	3.29
B	**avg**	**4.81**	**4.80**	**4.81**	**4.81**	**2.68**	**2.62**	**2.78**	**2.78**	**3.51**	**3.20**	**3.86**	**3.88**
	+1σ	4.84	4.83	4.85	4.85	2.97	2.94	2.98	2.99	4.30	3.97	4.44	4.47
	-1σ	4.80	4.81	4.79	4.79	2.44	2.44	2.71	2.70	2.54	2.44	3.28	3.23
C	**avg**	**4.86**	**4.86**	**4.85**	**4.85**	**2.80**	**2.78**	**2.97**	**2.96**	**3.35**	**3.23**	**3.88**	**3.86**
	+1σ	4.92	4.92	4.90	4.90	3.15	3.12	3.22	3.22	4.16	4.02	4.48	4.49
	-1σ	4.80	4.79	4.80	4.79	2.61	2.44	2.84	2.70	2.90	2.56	3.44	3.04
D	**avg**	**4.86**	**4.85**	**4.86**	**4.85**	**3.01**	**2.91**	**3.14**	**3.07**	**3.64**	**3.30**	**3.95**	**3.68**
	+1σ	4.93	4.91	4.92	4.91	3.41	3.37	3.45	3.44	4.38	4.03	4.46	4.33
	-1σ	4.77	4.77	4.77	4.77	2.32	2.39	2.43	2.42	2.86	2.67	3.42	3.39
E	**avg**	**4.84**	**4.83**	**4.84**	**4.84**	**2.64**	**2.69**	**2.74**	**2.73**	**3.63**	**3.40**	**3.93**	**3.93**
	+1σ	4.90	4.90	4.91	4.91	2.96	2.99	3.05	3.05	4.39	4.13	4.44	4.46
	-1σ	4.78	4.78	4.78	4.78	2.49	2.44	2.66	2.66	2.64	2.52	3.31	3.30
F	**avg**	**4.86**	**4.86**	**4.84**	**4.84**	**2.75**	**2.70**	**2.86**	**2.85**	**3.47**	**3.29**	**3.88**	**3.88**
	+1σ	4.93	4.95	4.90	4.89	3.01	2.97	3.05	3.05	4.29	4.06	4.45	4.47

Table S5 (continued).													
model		HCHO				NOx				O_3			
		Air	PO3	LO3	LCH4	Air	PO3	LO3	LCH4	Air	PO3	LO3	LCH4
	-1σ	1.87	1.85	2.07	2.06	0.49	0.89	0.39	0.37	4.47	4.60	4.44	4.43
A	**avg**	**2.17**	**2.14**	**2.32**	**2.31**	**1.05**	**1.36**	**0.90**	**0.89**	**4.70**	**4.80**	**4.66**	**4.65**
	+1σ	2.48	2.43	2.57	2.56	1.61	1.84	1.41	1.42	4.93	4.99	4.88	4.87
	-1σ	1.93	1.85	2.12	2.13	0.50	0.90	0.39	0.36	4.27	4.41	4.24	4.22
B	**avg**	**2.20**	**2.13**	**2.33**	**2.34**	**1.03**	**1.32**	**0.88**	**0.86**	**4.51**	**4.62**	**4.47**	**4.45**
	+1σ	2.47	2.40	2.54	2.55	1.56	1.74	1.37	1.36	4.76	4.83	4.70	4.68
	-1σ	1.89	1.91	2.20	2.19	0.78	0.91	0.66	0.66	4.32	4.36	4.28	4.28
C	**avg**	**2.22**	**2.22**	**2.43**	**2.43**	**1.14**	**1.27**	**1.00**	**1.01**	**4.49**	**4.54**	**4.45**	**4.45**
	+1σ	2.55	2.52	2.66	2.66	1.51	1.64	1.34	1.36	4.67	4.72	4.62	4.62
	-1σ	1.69	1.75	1.79	1.80	0.43	0.88	0.43	0.63	4.16	4.35	4.15	4.24
D	**avg**	**1.89**	**1.97**	**1.97**	**2.00**	**0.98**	**1.36**	**0.89**	**1.12**	**4.40**	**4.56**	**4.38**	**4.47**
	+1σ	2.10	2.19	2.15	2.19	1.53	1.84	1.35	1.61	4.65	4.77	4.60	4.71
	-1σ	1.94	1.93	2.14	2.13	0.52	0.85	0.48	0.46	4.17	4.31	4.19	4.17
E	**avg**	**2.21**	**2.19**	**2.34**	**2.34**	**0.97**	**1.22**	**0.92**	**0.90**	**4.41**	**4.53**	**4.42**	**4.41**
	+1σ	2.47	2.45	2.54	2.54	1.42	1.59	1.36	1.35	4.66	4.76	4.66	4.65
	-1σ	1.89	1.84	2.14	2.13	0.75	1.02	0.65	0.63	4.29	4.36	4.27	4.26
F	**avg**	**2.21**	**2.15**	**2.37**	**2.37**	**1.17**	**1.37**	**1.03**	**1.02**	**4.49**	**4.54**	**4.46**	**4.44**
	+1σ	2.52	2.46	2.60	2.60	1.59	1.73	1.41	1.41	4.69	4.72	4.65	4.63

160