This manuscript reports on the first year-round measurements of NOx, PAN and NOy made on the Greenland ice sheet, plus select NMHC measurements made with an automated GC system, and combines them with ongoing measurements of ozone provided by NOAA. The nitrogen oxide data are very interesting by virtue of being unique. Data interpretation focuses on the seasonality of the different N oxides, the speciation of NOy, and the sources of short-lived plumes that are superimposed on the relatively smooth seasonal variations. Source attribution relies heavily on FLEXPART "retroplume" analysis to identify enhancements in the N-oxides and some NMHC that can be linked to anthropogenic emissions and/or biomass burning in north America versus Europe versus Asia. All of the N oxides show broad peaks from late winter into spring and then decrease through summer into fall, with NOx staying elevated later into summer than PAN and NOy. PAN is the dominant fraction of NOy throughout the year, with the PAN/NOy ratio peaking near 0.8 in April and remaining ~ 0.5 even at its minimum in July. These findings are qualitatively similar with previous results from long-term stations in the Arctic basin, so probably were broadly expected (certainly not surprising to me), but confirmation is still valuable. Authors note that NOx/NOy increases from ~ 0.04 in winter to a little over 0.1 in summer, with the summer time increase tentatively ascribed to release of NOx from sunlit snow. They also point out that the sum of NOx plus PAN never accounts for all of measured NOy, with unmeasured species contributing 20-38% of NOy throughout the year. This shortfall is termed "odd NOy" and amounts to nearly 50 ppt in summer and nearly 100 ppt in winter (for monthly averages). Authors suggest that odd NOy probably represents alkyl nitrates, nitric, pernitric, and nitrous acids; with the first 3 potentially transported to Summit from remote sources at lower latitudes, while snowpack emissions may account for some fraction of the 3 acids during summer. One previous study found that the sum of C1 to C4 alkyl nitrates reaching Summit increased from ~ 10 ppt in July/Aug to a max < 35 ppt in Feb. This would imply that the sum of nitric, pernitric and nitrous acids ranges from 65 ppt in winter down to 30-40 ppt in summer. There are no wintertime measurements, but several campaigns have measured all three acids at Summit during summer and 30 ppt for monthly averaged sum would be quite a bit higher than reported values (Grannas et al., 2007 summarize most of these studies and provide references to original studies). Perhaps the calculated odd NOy is not inconsistent with these results if the uncertainties are considered, or perhaps 2008-2010 captured different conditions than these other studies. Authors suggest that it would be worthwhile to measure all of the candidate N oxides simultaneously to clarify the NOy budget, but should acknowledge that this was the objective of the 1998 and 1999 campaigns when the MTU group first went to Summit (there were no measurements of pernitric acid in that study, but the sum of alkyl nitrates plus nitric and nitrous acids was in the 15-20 ppt range and it is not likely that average pernitric exceeded 10 ppt). The NMHC observations are very consistent with previous results based on canister samples reported in Swanson et al. (2003). This is encouraging, but it appears that these results have been highlighted in 2 papers from the Helmig group already, so I am not sure they need to be included in this manuscript as well. To me, they add very little to the present story and none of the main points of this paper would be impacted if the NMHC were removed. The source attribution effort for plumes reaching Summit relies heavily on FLEXPART, which is an appropriate and well validated tool. However, I find that this section of the manuscript (3.2) is not very well presented, verging on confusing in places. Specific examples will be pointed out below, but at a high level I find it disturbing that several times the authors seem to suggest that errors in FLEXPART transport can be so large that they mistrust it (in general) juxtaposed to a paragraph where they are interpreting changes from one time step to the next as valid. Following is a commingled list of specific and technical comments, keyed to the page and line numbers in the "printer friendly" version downloaded from the ACPD site.

The authors would like to thank the anonymous reviewer for the helpful comments and providing detailed suggestions. The manuscript will be revised based on the suggestions given. Responses to each comment are given below, with details on revisions to the manuscript.

13819/12 as noted above, 40 ppt for the sum of N acids plus alkyl nitrates seems high for summer at Summit. Consider confidence level of these calculated differences, and perhaps whether they should be highlighted in the abstract.

Section 3.1.1 will be updated to include a discussion on the uncertainty in the determination of odd $NO_y$ by propagation of errors from the PAN, $NO_x$ and $NO_y$ data. Figure 2 has been updated to include these error estimates (see below). Figure 2 shows that the sum of PAN and $NO_x$ can account for the monthly averaged $NO_y$ within the uncertainty range during spring and fall months. However, during July 2008 and 2009, the deficit in $NO_y$ cannot be explained by measurement uncertainties alone. A similar result is observed during the winter months.

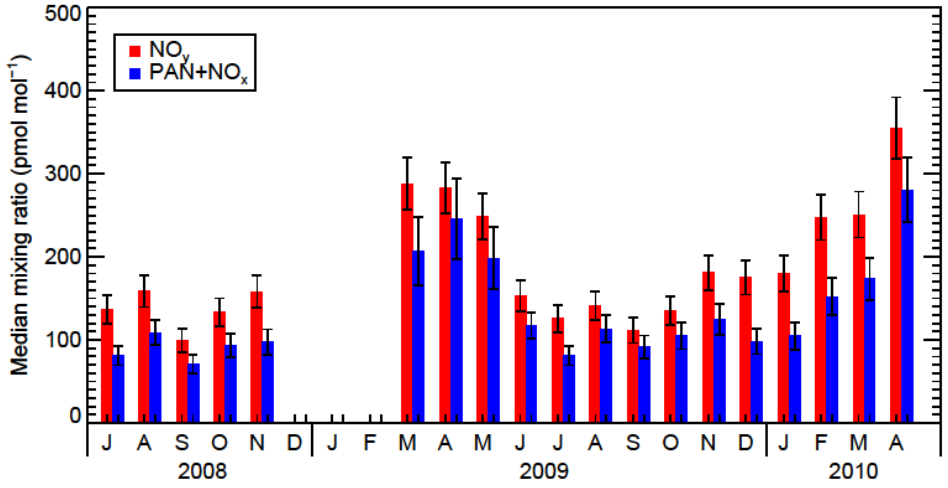


Figure 2: Monthly mean levels of $NO_y$ and PAN+$NO_x$ at Summit calculated from individual 30 min averages. Error bars represent uncertainty in the measurements resulting from measurement accuracy, calibration uncertainty and artifact corrections as discussed in sections 2.2.1 and 2.2.2 and the supplementary material. Uncertainties in PAN+$NO_x$ were determined from the propagation of errors.

Daily averaged values for $NO_y$ and PAN+$NO_x$ (with measurement uncertainties) are shown in the Fig R-1 for June-July 2009 and November 2009- February 2010. The results show that on many days the difference between $NO_y$ and PAN+$NO_x$ (i.e. odd $NO_y$) cannot be explained by measurement uncertainties alone. A statistical analysis of the daily averaged data shows that, of the 439 days when data was available for all species, odd $NO_y$ was greater than the measurement error on 275 days (63%). This information will be provided in the revised manuscript and the abstract updated.

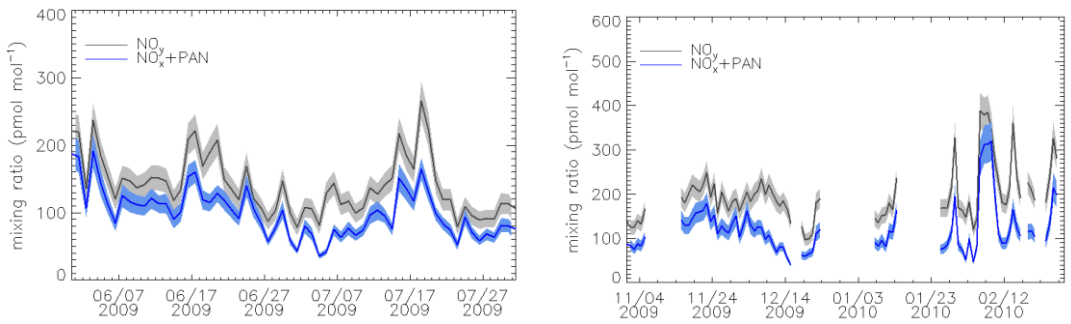


Fig R-1. Daily mean levels of $NO_y$, PAN+$NO_x$ at Summit calculated from individual 30 min averages for June-July 2009 (left) and November 2009- February 2010 (right). Shaded regions represented the measurement uncertainty. Uncertainties in PAN+$NO_x$ were determined from the propagation of errors of PAN, $NO_x$ and $NO_y$.

13821/7-8 quite a few additional TOPSE and POLARCAT references also make the point regarding different source regions near surface versus mid/upper trop, and could be added.

The authors thank the reviewer for pointing out the missing references. Relevant TOPSE and POLARCAT references will be included in the updated manuscript.

13821/11-12 is North Asia a recognized region? Do you mean Siberia? That seems

more common usage

The reviewer is correct, Siberia is the region we are discussing and the manuscript will be updated to reflect this change.

13821/18 emissions from the –→ emissions in the

This sentence will be updated as suggested.

13822/5 should "high altitude Arctic" be "high latitude"? It is true that high altitude Arctic is not impacted strongly by local sources, but also true for high altitudes pretty much everywhere.

This is a good point. Please see the response to the comment below regarding the change of high altitude to high latitude.

13822/19 and 20 here it seems authors did mean altitude above, but one could debate whether 3 km elevation of Summit is really "high altitude", and both of these sentences would also make sense if latitude replaced altitude.

The authors agree that the use of high latitude would make more sense and will change both sentences in the manuscript as suggested.

13823/15 Pretty sure the instrument that went to PICO was built for studies at Summit in 1998 to 2001, so it seems odd not to mention it. Mike Dziobak would know for sure.

The reviewer is correct. The earlier version of the instrument was used during the campaigns at Summit. The PICO instrument was highlighted as it demonstrated the use of the instrument for long-term measurements. However, it is important to mention the instrument at Summit and the sentence has been changed as follows:

*"The system was developed at Michigan Technological University and is based on the same design that was used in Newfoundland in 1996 (Peterson and Honrath, 1999) and subsequently installed at Summit during campaigns in 1998, 1999 and 2000 (Honrath et al., 1999, 2002,Dibb 2002 ) and at the Pico Mountain Site from 2002 to 2010 (Val Martín et al., 2006)."*

Honrath, R. E., M. C. Peterson, S. Guo, J. E. Dibb, P. B. Shepson, and B. Campbell. "Evidence of NOx production within or upon ice particles in the Greenland snowpack." Geophysical Research Letters 26, no. 695-698, 6, 1999.

Honrath, R. E., Y. Lu, M. C. Peterson, J. E. Dibb, M. A. Arsenault, N. J. Cullen, and K. Steffen. "Vertical fluxes of NOx, HONO, and HNO3, above the snowpack at Summit, Greenland." Atmospheric Environment 36, no. 15, 2629-2640, 2002.

Dibb, Jack E., Matthew Arsenault, Matthew C. Peterson, and Richard E. Honrath. "Fast nitrogen oxide photochemistry in Summit, Greenland snow."Atmospheric Environment 36, no. 15, 2501-2511, 2002.

13824/3-16 This paragraph briefly outlining the calibration of the NOx,y instrument is misleading. Details in the supplement clear things up, but need to also be accurate here. Specifically, you cannot assess the LED NO2 convertor efficiency just by challenging it with NO, and you should mention that you challenged the Au convertor with NO2 twice a day, and with HNO3 and NPN every 3 days.

This paragraph has been re-written to clarify the details regarding the $NO_{x,y}$ instrument. The following text will now be included in the main manuscript text:

*"Calibrations were performed every 12 hours to determine the sensitivity of the instrument to NO via standard addition (10 $cm^3$ $min^{-1}$) of ~1 mmol $mol^{-1}$ of NO in nitrogen ($N_2$) (Scott Marrin, Scott Specialty Gases) to the sample flow (650 $cm^3$ $min^{-1}$) at the inlet on the tower. A known amount of $NO_2$, generated via gas phase tritration of NO with $O_3$, was also added to the to the sample flow during the calibration cycle, to determine the conversion efficiencies of the $NO_2$ and $NO_y$ converters. In addition to the standard calibrations, every 3 days the conversion efficiency of the $NO_y$ converter to $HNO_3$ and NPN were determined and artifacts for $NO_y$, NO and $NO_2$ were measured via sampling $NO_x$ free air (Breathing air grade, Airgas, Radnor, PA, USA). The final datasets were corrected for this artifact."*

13824/24-25 The second "additional filter" is confusing. Seems it is based on subjective or arbitrary assessment of too much variability, but not enough to be caught by the objective Poisson filter. Please explain this better.

The second additional filter was utilized to remove potentially polluted measurements or those with unrealistic high values or high variability that was not removed in the first filter. It was noted that enhancements in the $NO_{x,y}$ were occasionally observed during periods coinciding with the logged start and end times of skiway grooming. To be conservative the filter removed all measurements during skiway grooming times. However, on closer inspection of the skiway periods (since the manuscript was submitted), it was observed some of the data did not show large deviations when compared to neighboring measurements (most likely as a result of wind direction), therefore ~ 1% of the data that were originally filtered may not necessarily be impacted by skiway grooming. In the updated data analysis and final manuscript only those measurements made during logged skiway grooming times, with obvious enhancements relative to neighboring data points, have been included in the skiway filter. Changing this filter has a very low impact on the final results, with the campaign mean levels changing by only 1 pmol $mol^{-1}$ for $NO_x$ and $NO_y$.

A second additional filter was also used to remove outliers that were not filtered with the Poission filter, or no instrument malfunction was logged. These outliers were determined via threshold values for NO, $NO_2$ $NO_x$ and $NO_y$ based on very high values or large standard deviations. The filtered points were doubled checked as outliers by comparing to adjacent

observations. Less than 0.2% of the data were filtered as outliers.

We thank the reviewer for noting that the filter techniques were not clear. The revised manuscript will include additional statements to clarify the filters that were used with the percentage of values removed for each filter clearly stated.

13826/24-25 Seems likely that the % uncertainty for PAN measurements varied somewhat depending on the ambient mixing ratio (most likely higher as detection limits approached). To be consistent with 2.2.1 (and to be more useful to reader), should state the ppt levels where the 16 and 22% estimates are valid.

The reviewer is correct. At low PAN ambient mixing ratios measurement precision is an important factor when considering the total uncertainty. At higher mixing ratios errors associated with accuracy of the calibrations and flows are greater. To clarify these points, this section of the manuscript has been re-written with the following text:

*"Similarly to the $NO_x$ and $NO_y$ data, the PAN measurements were averaged over 30 min. The total uncertainty for the 30 min averaged PAN mole fractions was determined from the root sum of the squares of the precision of the instrument (estimated as $2\sigma N^{0.5}$, where N is the number of points averaged in 30 min (N = 3)) and the uncertainty in the calibration standard. The precision is < 57 pmol $mol^{-1}$ with a mean (median) value of 6.7 pmol $mol^{-1}$ (5.6 pmol $mol^{-1}$). Uncertainty in the PAN calibration standard is associated with uncertainties in (a) the calculation of the NO addition, (b) the conversion of NO to PAN from the calibration unit and (c) variability in the PAN sensitivity between calibrations. The uncertainty in the calibration standard is 15% during normal operation. This value increased to 21% during the period in spring 2009 when there were no calibrations. Using the mean precision, the total uncertainty in the PAN measurements at 150 pmol $mol^{-1}$ is estimated to be approximately 16% during normal operation and 22%, during spring 2009.*

13827/8-27 As noted above, consider removing the NMHC from this paper. If that recommendation is not taken, this section needs to describe the duty cycle and time resolution of these measurements. Simple math (6000 measurements over 25 months) suggests nearly 8 data points per day, so maybe 3 hour resolution but how long is sample actually concentrated? Did it work all the time for 25 months? How often calibrated (if 1000 blanks/standards = 500 each, that works out to less than 1/day)? Perhaps most critical, how were the 30 minute N oxide measurements merged with the NMHC measurements that clearly do not have 48 observations/day?

As the reviewer stated, the NMHC method and data have been presented in Helmig et al., 2014a. However, the Helmig et al. 2014a publication was primarily focused on processes within snowpack measurements of NMHC, with little discussion and interpretation of the ambient air behavior of NMHC. We feel that the NMHC data are a critical part of this study as the NMHC ambient air relative ratios provide important information on the rate of photochemical processing. This analysis yields confirmation of fast transport to the measurement site during periods when low ozone levels were observed in the winter. Additionally, the NMHC provide information on the interannual variability of $O_3$ precursors at the site, which is less clear in the other data due to a gap in the NOx,y measurements and high PAN uncertainties over winter 2008/2009.

The reviewer is correct with the calculations. The NMHC system provide approximately 8

measurements a day, with each measurement representing a collection/sample integration time during the sample pre-focusing step of ~ 45 min. In addition, one blank sample was analyzed ~daily, and a standard every ~ 2 days. The revised manuscript will be updated with this information. A direct comparison was not performed between individual NMHC and N oxide measurements. However, when calculating enhancements due to fire or anthropogenic pollution, both the NMHC and N oxide data (when available) were averaged over the same time period of the event.

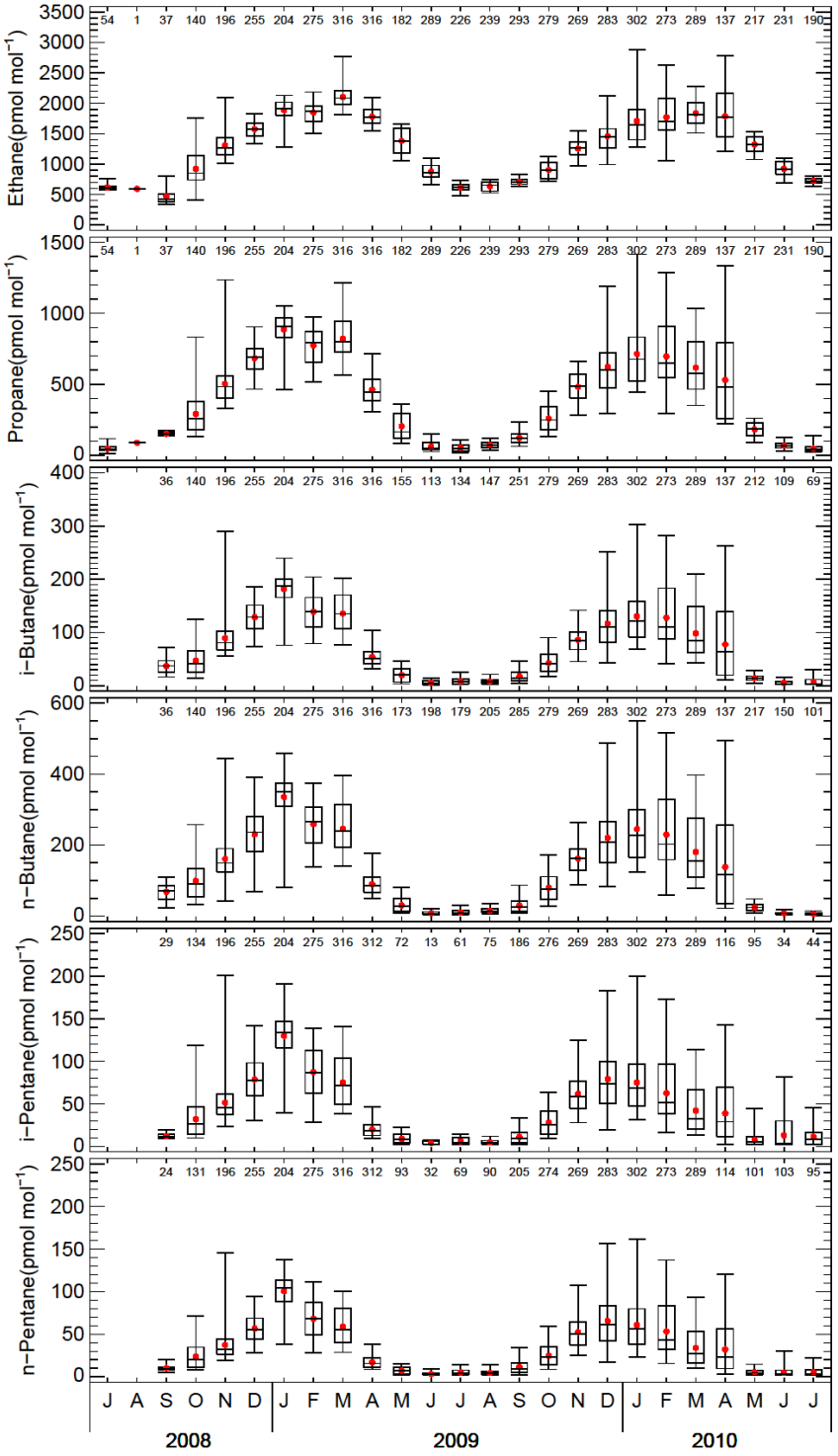


Figure 4. Monthly averages of (a) ethane, (b) propane, (c) i-butane, (d) n-butane, (e) i-pentane and (f) n-pentane, at Summit during from July 2008-July 2010. Median and mean are indicated by a filled red circle and black square, respectively; the box indicates the middle 67% of the data; and the top and bottom of the vertical whiskers indicate the 1st and 99th percentile of all the data. The numbers at the top of each plot represent the number of measurements included in the distribution.

13829/1 Not clear how the simple treatment of BC scavenging would cause "under-estimation" of the BC tracer reaching Summit. Not including aging effects making BC more hydrophylic would seem to decrease scavenging by any precipitation along the transport path, which ought to lead to overestimates of the concentration still in the airmass(es) reaching Summit.

We thank the reviewer for pointing out how this sentence may be misleading. The underestimation in the BC tracer at Summit is a result of the use of wet scavenging coefficients in the model which are more representative of a hydrophilic aerosol. Close to the source region, the BC would be more hydrophobic, therefore the wet scavenging would be greater than expected. As the BC ages the scattering coefficients are more realistic. However, the overall result is an underestimation. The manuscript has been updated with the following text to clarify the wet scavenging impact on the tracer:

*"The wet scavenging coefficient used in the model is more representative of a hydrophilic aerosol, however, there is no conversion from hydrophobic to hydrophilic properties with aging BC in the model, therefore greater scavenging may occur closer to the source region, resulting in an underestimation of the BC tracer at Summit (Stohl et al., 2013)."*

13829/8 Why choose to create monthly averages of ozone for 3 full years, rather than just the 2 years (or 25 months) with overlapping NOx,y and PAN data?

As a full 3 year data set was available for ozone we included all measurements in the seasonal cycle in Figure 1. However, Figure 1 has now been revised and the 25 months of overlapping data will be presented separately for $NO_{x,y}$, PAN and $O_3$. This figure is presented in the document containing the responses to reviewer 2.

13829/18-22 As written, this sentence seems to imply that Liang et al. (2011) claim that STE is a source of PAN in the Arctic upper troposphere (implying enhanced PAN in the lower stratosphere). What they actually believe they observed was production of PAN in stratospherically influenced air masses in the troposphere above 5 km that had high NOx and HNO3 transferred from the stratosphere. Suggest redrafting this sentence to clarify this point.

This sentence has been revised as suggested. The text now reads as follows:

*"Observations from the Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission in 2008, show that mixed stratospheric-tropospheric air masses, above 5 km, have elevated levels of $NO_x$ and $HNO_3$, which can subsequently be converted to PAN (Liang et al., 2011). Therefore, high mole fractions observed in PAN, $NO_y$ and $NO_x$ may also be the result of sampling air masses mixed with those originating from the stratosphere and upper troposphere."*

13831/5 dectection—→ detection

This typographical error will be corrected.

13831-13833 (end of section 3.1.1) The 2.5 paragraphs discussing the "odd NOy" strike me as more speculative than they need to be. Granted, no one has yet measured nitric, pernitric, and nitrous acids through the winter but they have been measured during campaigns in summers of 2003, 2007, 2008, 2010, and 2011 and during a spring campaign in 2004. (Earlier campaigns in 1998-2002 did not include pernitric acid, but did include canister sampling for NMHC and usually quantified the C1-C4 alkyl nitrates.)

The 2008 GSHOX campaign ended ~9 July so may not overlap the observations discussed herein, but the 2010 campaign lead by M. Hastings extended from mid May till late June. I suspect that the PIs who made measurements of snow sourced NOy compounds during these experiments would be pleased to provide statistical summaries (perhaps even complete data files) that would allow quantitative comparison to your calculated odd NOy in summer 2010 and semi quantitative assessment for the summer and spring seasons more generally.

We are very grateful to the reviewer for providing the information regarding the snow sourced $NO_y$ compounds at Summit. The authors have contacted the PIs of the 2010 and 2011 campaigns at Summit and data has been provided for a comparison to the odd $NO_y$ from this study. Data from 2010 is only available during the months of May and June and although some overlaps with our $NO_x$ and $NO_y$ data in 2010, PAN data was not available after April, so a direct comparison with odd $NO_y$ is not possible. However, the data provides an estimate of the $HNO_3$ and HONO levels during late spring/summer.

Campaign statistics for HONO and $HNO_3$ are given in the table below. The measurements were performed in the clean air sector at Summit with the inlet set at approximately 1.5 m above the snow surface (J.E. Dibb and M.G. Hastings, personal communication, Fibiger et al., in prep.). The third column (HONO + $HNO_3$) only includes periods when both HONO and $HNO_3$ data are available.

| | **HONO (pmol mol$^{-1}$)** | | **$HNO_3$ (pmol mol$^{-1}$)** | | **HONO + $HNO_3$ (pmol mol$^{-1}$)** |
|---|---|---|---|---|---|
| **Period** | **Mean ± SD** | **Min/Max** | **Mean ± SD** | **Min/Max** | **Mean ± SD** |
| **May 18-June 22 2010** | 17.4 ± 14.3 | 0.9/95.9 | 9.3 ±6.8 | 0.5/50.4 | 26.7 ± 18.2 |
| **May 25-June 26 2011** | 8.1 ± 5.5 | -1.1/51.2 | 10.4 ± 14.5 | 0.0/79.6 | 18.6 ± 16.0 |

Figure R-2 shows monthly mean odd $NO_y$ (with calculated uncertainties) for the data obtained at Summit between July 2008 and April 2010. For June 2009, the mean calculated odd $NO_y$ is 36 ± 25 pmol mol$^{-1}$ (mean ± uncertainty). This value is within the measured range of $HNO_3$ + HONO at Summit in 2010 and 2011. The results confirm our initial findings that suggest HONO and $HNO_3$ may contribute to the odd $NO_y$ observed during the summer months. As suggested by the reviewer, the manuscript will be updated to include the results presented here and a discussion on the odd $NO_y$ compounds measured at Summit during previous campaigns to strengthen this section of the manuscript.

*Fibiger, D.L., M.G. Hastings, J.E. Dibb, D. Chen, L.G. Huey, Source of nitrate to Summit, Greenland: analysis of local snow and atmospheric chemistry, in preparation for Atmospheric Chemistry and Physics*

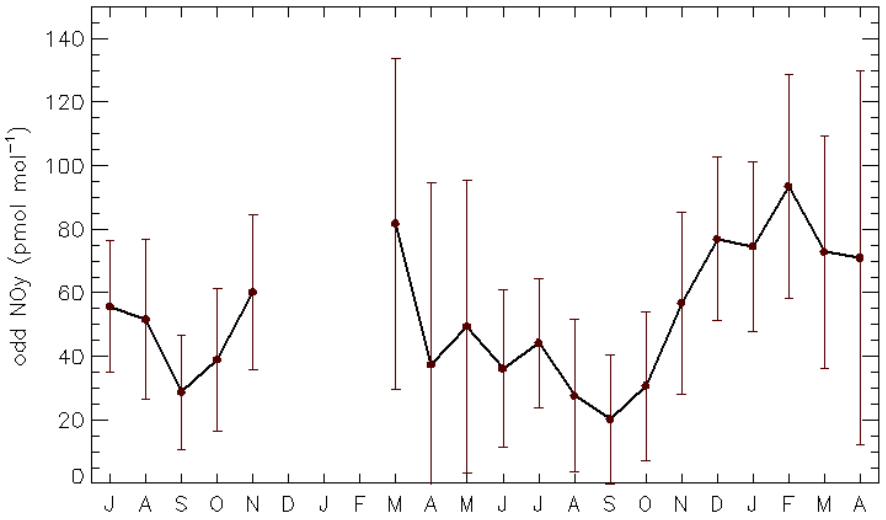


Fig R-2. Monthly levels of odd $NO_y$ (calculated from the 30 min avg $NO_y$, PAN and $NO_x$ measurements) at Summit. Error bars represent the uncertainty in odd $NO_y$, determined from the propagation of errors of PAN, $NO_x$ and $NO_y$.

13833/1-3 Can the larger uncertainty in PAN due to instrument degradation after spring 2009 not be propagated into the uncertainty of calculated odd NOy? This would be particularly important if you are able to compare the sum of measured NOx, PAN, HNO3, HO2NO2, HONO and assumed sum of RONO2 (~ 10 ppt) to measured NOy during summer 2010 as suggested immediately above.

The authors agree with the reviewer and have calculated measurement uncertainties for odd $NO_y$ as shown in Figure R-2 above. The revised manuscript will include a discussion of the odd $NO_y$ calculated uncertainty.

13833-13834 (section 3.1.2) as noted earlier, I do not feel the NMHC add much to this paper, especially if they "have previously been presented in detail (Helmig et al.,2014a,b" (lines 18-19 on 13833)

As mentioned in response to an earlier comment, we feel that the NMHC data are an important part of the study. This statement in the manuscript, however, may be misleading, as the publication by Helmig et al. 2014a focused primarily on processes within the snowpack. This statement will be re-phrased in the revised manuscript to state, "Measurements of primarily firn air conducted with this system were presented by Helmig et al. 2014a"

Figure 4 will also be replaced by a plot that presents the two years of NMHC data separately as a box-and-whisker plot as suggested by reviewer 2.

13834/13-23 Strongly suspect that Figure 4 would show quite similar features if NOx and NOy replaced the NMHC.

The reviewer is correct and similar variability is observed in the $NO_x$ and $NO_y$ measurements as a result of polluted air masses sampled at the site. In this section we focused on NMHC as a comparison could be made between the two winters. This was not possible for $NO_x$ and $NO_y$ as a result of the gap in measurements from November 2008 through February 2009. The text on the variability in the NMHC levels that the reviewer is refereeing to, will be moved to section 3.1.2 below the discussion on NMHC seasonal cycle, in the revised manuscript.

13835/17 to the total,—-→ of the total,

As suggested this sentence will be updated

13833-13838 section 3.2.1 This section intends to focus on winter/early spring (DJFM), which may be fine, but as it unfolds the authors do not stick to their own script and it gets a little confusing at times. Perhaps there should be more material in section 3.2 sort of setting the context and summarizing what FLEXPART thinks about seasonal patterns of transport from different source regions/types, plus any marked differences between the 2 years. Once it is established that urban/industrial plumes dominate DJFM and that 10/13 apparent BB smoke plumes reached Summit in the summer, drilling down into the 2 different seasons could follow.

As suggested by reviewers 1 and 2, section 3.2 will be substantially re-written to improve the manuscript. Subsections 3.2.1 and 3.2.2 will focus on anthropogenic and biomass burning events respectively. An introduction in section 3.2, as suggested by the reviewer, will now focus on the source contribution from FLEXPART. Figure R-3 shown below will be included to demonstrate the seasonal patterns in the transport and sources. The discussions on the contribution from each tracer, currently given individually in sections 3.2.1 and 3.2.2, will be combined and included in this start of this section also. This introduction will help clarify the focus on individual seasons in the subsections.

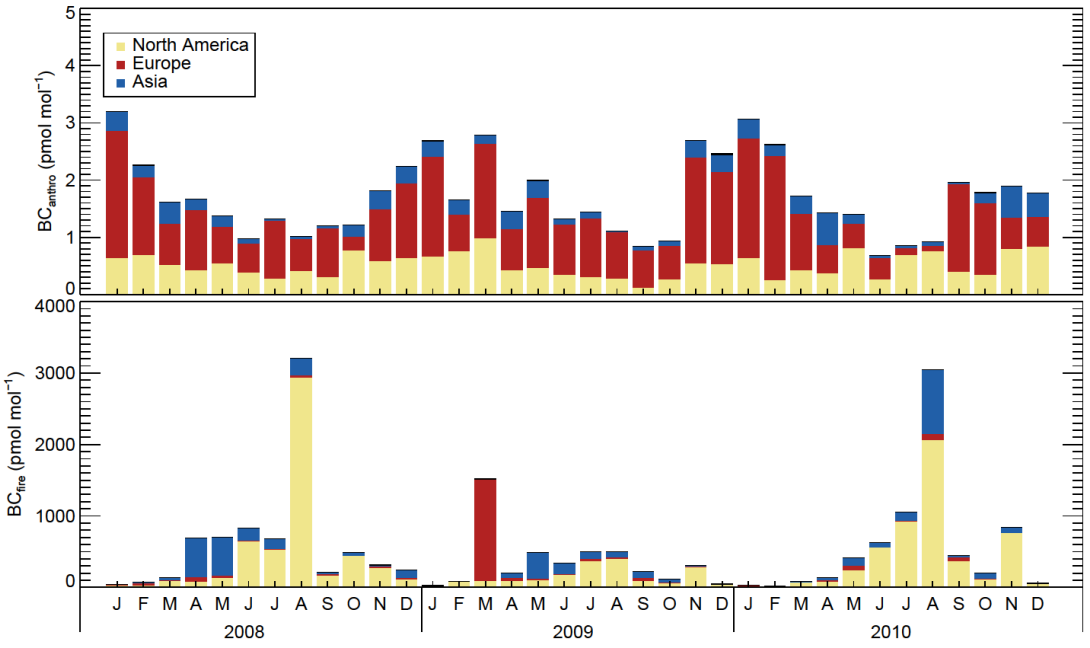


Fig R-3. Barplot showing the total monthly $BC_{anthro}$ and $BC_{fire}$ tracer from FLEXPART. The different colors represent the contributions from North America, Europe and Asia.

The only full paragraph on page 13835 focuses exclusively on FLEXPART analyses in the DJFM season, and specifically on where this model thinks its BCanthro tracer originated in plumes it advects to Summit. This is fine, but it would be useful to compare the FLEXPART source attribution to similar analyses done with a range of different modeling tools as part of the POLARCAT project (especially considering that Stohl was co-leader of POLARCAT and Burkhart provided FLEXPART forecasts to several of the aircraft campaigns that were mounted under the POLARCAT banner). Do the results reported here apply only to Summit/central Greenland, especially the very low impact from Asian sources? Law et al. (in press, and available on-line) provides a nice summary that would facilitate putting Summit into the context of the wider Arctic, probably needing just an additional sentence or two.

The author's agree with the reviewer's comments and the revised manuscript will include a more complete discussion of the source contribution from FLEXPART during different seasons which will be compared to results from studies that performed analyses during the POLARCAT project. A discussion on Summit and the Arctic region in general will also be provided in the revised manuscript.

The next paragraph extends onto the next page, and presents enhancements of NOx, PAN, NOy, ozone and ethane measured at Summit during the pollution episodes identified by FLEXPART. Each of the tracers actually decreases 20-26% of the time that FLEXPART thinks pollution reached Summit, a result which "may be associated with erroneous transport by FLEXPART". Authors need to expand a little on this, else they risk convincing the reader that much of this entire section of the manuscript should be discounted. Did all or most of the tracers decrease in the same episodes? With tracers each showing the "wrong" response 20% of the time it is possible (but not likely) that transport was erroneous nearly all of the time. Could some of the FLEXPART "misses" be "hits" if the time window was expanded a little bit (+/- 3, 6, or 9 hours)? Even a 12 hour offset could be explained if the mean wind speed (in the met fields driving FLEXPART) during 11 days of transport was off by just 4%.

The reviewer has made an important point regarding the time window and errors in FLEXPART. The data has been re-analyzed and results shows that the number of events when the mean trace gas level decreased below the background level are lower when using either a ± 3,6,9 or 12 hour offset in the time window. Care must be taken when choosing an offset, as a window that is too large, may "dampen" the impact of the event through the inclusion of background levels, or overlapping with other pollution events. This section of the manuscript will be revised and updated with analyses performed using different time windows to determine the impact on the event statistics

The next paragraph (all on 13836) is quite confusing. First 2 sentences start by focusing on detailed time series for the DJFM period in the 2 years, but then analysis of the full 2 year record is interjected, followed by a summary of low ozone events in April-Sept. (Recall that section 3.2.1 is supposed to be focused on winter/spring.) The following paragraph (extending onto 13837) also starts by considering the full record, but then comes back to winter/spring, except that it expands winter to include Oct and Nov along with DJFM (the months that motivated the investigation of anthropogenic plumes). More confusing than jumping back and forth between intervals or seasons of interest are some of the numbers

in these two paragraphs. First we are told that 45 low ozone events were identified in summer months (Apr-Sept), of these 10 lined up with FLEXPART pollution plumes (some anthro, some BB, not specified how many of each). Then it is stated that over the full observation period there were 28 events where ozone decreased when the FLEXPART BCanthro tracer increased, with 21 of these in the expanded Oct-Mar winter. This would seem to imply that the other 7 cases had to occur during Apr-Sept, further implying that 3 of the 10 summer time events had to be linked to BB plumes. However, Table 3 lists all of the BB events and only 2 show ozone decreases > 2 ppb. There is a 3rd BB event with ozone decrease (too small), but that occurs during March, and would also qualify as an anthro plume. Note that it is not easy to try to follow the links between these different numbers, making it that much more disappointing that the simple math does not work out.

The reason for the differences in the statistics is due to the way the analyses is performed when identifying low ozone periods. When focusing only on low ozone events, periods were selected based on when the $O_3$ measurements < 2ppb for at least 12 hours. The event was classified as "polluted" if the mean $BC_{anthro}$ level during the event period was above the 75$^{th}$ percentile. However, the $BC_{anthro}$ may not have been enhanced for at least 12 hours (which is the classification used when identifying anthropogenic pollution events) so may result in different events. The revised manuscript will clarify these points to ensure this is clearer to the reader. Additionally, tables for all events (including event time, event length and trace gas measurement statistics) will be provided in the revised supplementary material.

Coming back to the paragraph bridging pp 13836-13837, the final 4 sentences zoom in on just 8 days in Jan 2010 and focus closely on 23 Jan. Here it is argued that FLEXPART really nails the transport and suggest that decrease in ozone in an anthro plume advected quickly from Europe probably reflects titration by NO close to the source. Could be, but why do you (why should I) believe FLEXPART here but not so much on the top of page13836 (see earlier comment)?

The next paragraph (ends on page13838) continues to basically accept FLEXPART as truth, and starts by presenting evidence (enhanced ln(propane/ethane) ratio) that would be consistent with rapid transport from source region to Summit on both 23 and 31 Jan 2010. However, the authors go on to say that we probably should not be real confident that the propane/ethane ratio is truly reflecting age since emission (13837/14-16). Yet, the remainder of the paragraph implicitly assumes that the retroplume analysis can be accepted and interpreted nearly on a time step by time step basis. However, the final paragraph of section 3.2.1 points out that there were intervals when the tracer measurements at Summit jumped up as if they were impacted by a pollution or BB plume while neither of the FLEXPART BC tracers showed any increase, and again suggest that the FLEXPART transport may be significantly in error.

Clearly, FLEXPART is not perfect, but it is also clear that it is useful. To me, it kind of feels like different people wrote different paragraphs in this section, with some biased to take the model results as gospel and others biased strongly against the model. I think it would help the presentation if an objective assessment of FLEXPART skill capturing plume transport to Summit was included in the expanded section 3.2 suggested above, and if a single author carefully edited section 3.2.1 to have all the paragraphs show appropriate appreciation of both the strengths and limitations of FLEXPART (or any other model).

With any transport model, there will be uncertainties in the transport, as a result of errors in the wind fields and missing/errors in source emissions. As suggested by the reviewer, the revised manuscript will include a more detailed discussion of FLEXPART transport and plume analyses at the start of section 3.2, including a discussion on previous studies that have used FLEXPART to identify pollution events.

13840/line 3 and line 21. First one states that the combination of BCfire and BCanthro tracers identifies 6 smoke plumes with small anthro signature, but second one says there were only 5 "nearly pure" smoke plumes. Note that Table 3 suggests that 2 smoke plumes had high anthro contribution, 6 had medium influence and 5 had low. Seems that line 3 is wrong, and also line 1 that claims 5 medium events (should be 6).

We thank the reviewer for noticing the error here. As shown in table 3, there are 5 plumes with low anthropogenic impacts, 6 plumes with medium anthropogenic impacts and 2 with high anthropogenic contributions. The statements in line 3 and line 21 are wrong in the text and this will be corrected in the revised manuscript.

13840/4-12 This paragraph makes a nice, and convincing, case that most of the BB plumes reaching Summit come from North America. This is consistent with a pretty long list of previous studies based on Greenland firn/ice cores but none of these are mentioned. It also seems that Stohl and Burkhart have earlier paper that reaches similar conclusion. Would be good to add some of these citations to this paragraph.

We thank the reviewer for pointing out that some references are missing and this paragraph will be updated with the relevant citations in the revised manuscript.

13842/4 anlayses–→analyses

This typographical error will be corrected

13842/14-23 Are these really conclusions from this paper? Do they not just confirm Swanson et al. (2003) and repeat findings in Helmig et al. (2014a,b)?

As suggested, these conclusions will be removed from this section.

13843/2-4 and 18-21 These 4 sentences seem to say the same thing. Probably better to kill the first 2 and keep the BB stuff in just one paragraph.

We agree with the reviewer's suggestion and will remove the relevant sentences from lines 2-4.

13844/10-11 HNO3 determine—–→HNO3 to determine would probably want to add RONO2 to this list, and possibly HO2NO2 (comparable to HNO3 in summer)

The grammatical error here will be corrected. The addition of $RONO_2$ and $HO2NO_2$ is an excellent suggestion and these two species have been added to the list of future measurement needs.