Response to Anonymous Referee #3

We thank the reviewer for the careful review of the manuscript and study and for their detailed comments. Their critiques, especially their attention to technical details, have helped to significantly strengthen our manuscript.

Here we detail changes made in our revised manuscript and responses to questions addressed by Anonymous Referee #3.

General and Specific Comments

Page 3, line 10. "Models have shown that future climate conditions will likely amplify N_2O production". Expand on this thought. Why is this so?

It is estimated that 44-73% of N_2O emissions originate from land ecosystems (Hirsch et al. 2006; Davidson et al. 2009). A warmer climate will most likely enhance these emissions (Arneth et al. 2010), driving a positive feedback loop. Stocker et al. (2013) investigates the role these feedback loops play in future climate conditions. We have updated the text to the following:

> *Models have shown that future climate conditions will likely amplify N_2O production through positive climate feedback effects, meaning a linear increase in time may under-predict future concentrations based on the current rate of change (Khalil and Rasmussen 1983; Stocker et al. 2013).*

Page 3, line 15. Add Hall et al. 2011

We have updated the manuscript to reflect this change.

Page 3, line 20. Consider adding the following citation for a recent estimate of SF_6 lifetime.

We have revised the manuscript to use updated estimates of the SF_6 lifetime from Kovacs et al. (2017) and Ray et al. (2017).

Page 5, line 4. How did you arrive at the detector temperature of 310 C? Was it optimized for N_2O and or SF_6?

The detector temperature of 310°C was used because we found that the response of the N_2O and SF_6 peaks were well defined at this temperature. Chromatogram output was optimized around N_2O while maintaining a robust SF_6 peak. No significant difference in detector response was found to using a detector temperature of 340°C reported in Hall et al. (2007) and Hall et al. (2011).

Page 5, line 9. Maybe not necessary to the paper. Why sample the archive air 6 times and then the reference gas 6 times instead of alternating between the two types of samples? Wouldn't alternating better track signal drift from injection to injection?

The procedure of running 6 measurements in a row of a sample or standard instead of alternating between standard and sample measurements was used because it was determined that while measuring N_2O, if a run contained an outlier (greater than 2σ), it was statistically more likely to be the first measurements (~30% of the time) compared to being in another position (~15% of the time). Outlier probability was evenly spread across measurement position for SF_6 (~17% for outlier

to be in any position). The cause of this discrepancy for measurements of N_2O is unlikely to be contamination from a previous run as we are purging the sample loop (10 ml) with 9-times the sample loop volume (60 ml min^{-1} for 1.5 minutes). This memory effect is mitigated when running 6 measurements back to back. Drift in the detector response over a set of 6 measurements (~50 min) is assumed to be linear.

Page 7, Results. Can you comment on why there were some large outliers? Problems with the sample or the integrity of a few flasks? Were the outliers the same for both N_2O and SF_6? Why were two different criteria for residual outliers (2-sigma for N_2O and 3-sigma for SF_6) used?

Outliers in the OHSU-PSU air archive are possibly due to several factors including storage integrity or possible contamination during collection. It is unclear what exactly caused each of the far outliers evaluated.

For the initial filtering process, far outliers in the OHSU-PSU air archive were considered to be 6*MAD (median absolute deviation) for N_2O (removes 6 samples) and 7*MAD for SF_6 (removes 2 samples). The far outliers for N_2O are not the far outliers for SF_6. 7*MAD is used for SF_6 as opposed to 6*MAD because the annual increase in measured values for SF_6 (~10% yr^{-1}) is significantly larger than in N_2O (~0.25% yr^{-1}). It is important to note that roughly half of the 159 samples measured from the OHSU-PSU air archive date prior to 1985, meaning the median measured SF_6 mole fraction will be biased towards this early period. By using 7*MAD for SF_6, we only remove values that clearly lie outside of a reasonable measurement.

For the second filter, Polynomial fits (1^{st} degree for N_2O and 2^{nd} degree for SF_6) were applied to the data and residuals outside of 2σ for N_2O and 3σ for SF_6 were removed. 3σ was used for SF_6 because the data points fit tightly to the polynomial fit. The second filter removed another 6 samples for N_2O (for a total of 12) and another 2 samples for SF_6 (for a total of 4). Again, the outliers removed for N_2O are not the same samples removed for SF_6. We found that using 2σ for SF_6 removed data points unnecessarily from the analysis.

Page 7, line 20. It is uncertain to the reviewer what "bootstrapping residual variability 1000 times" means. Did you sample subsample the data 1000 times and re-smooth?

The bootstrap process consists of calculating a new value for each data point from a normal distribution with mean equal to the measurement and standard deviation equal to the residual variability found from the original LOWESS regression. We repeat the LOWESS regression calculation for the new data set. The process is completed 1000 times, from which we calculate the 95% confidence interval for the original LOWESS regression to the measured data.

Page 8, line 5. You could cite Geller et al. as well

We have updated the manuscript to reflect this change.

Technical Corrections

Page 1, line 14. "prior to" to "before"

We have updated the manuscript to reflect this change.

Page 2, line 22. "major" to "primary"

We have updated the manuscript to reflect this change.

Page 4, line 20. "Peak separation is achieved by two Poropak Q 80/100 mesh columns" to "Two Poropak Q 80/100 mesh columns achieve peak separation"

We have updated the manuscript to reflect this change.

Page 4, line 22. "to significantly improve baseline signal stability" to "to improve baseline signal stability significantly"

We have updated the manuscript to reflect this change.

Page 5, line 20. "a two-week period" to "two weeks"

We have updated the manuscript to reflect this change.

Page 6, line 4. "Error" to "The error"

We have updated the manuscript to reflect this change.

Page 6, line 6. "To characterize of the" to "To characterize the" (remove the "of")

We have updated the manuscript to reflect this change.

Page 6, line 8. "a N2O" to "an N2O"

The sentence has been reworded to:

> *To properly account for this interference, SF_6 dilutions at low mixing ratios (0.6 - 6.0 ppt) must have N_2O mole fractions that reflect expected mole fractions in archived samples (300 - 315 ppb).*

Page 7, line 17. "analysis" to "the analysis"

We have updated the manuscript to reflect this change.

Page 7, line 29. "Prinn et al." is missing the period

We have updated the manuscript to reflect this change.

Page 9, line 12. Add a comma after Canada.

We have updated the manuscript to reflect this change.

Page 10, line 19. "have amplitude" to "have an amplitude"

We have updated the manuscript to reflect this change.

Page 10, line 22. “seasonal amplitude” to “the seasonal amplitude”

We have updated the manuscript to reflect this change.

Page 11, line 10. “and minimum amplitude” to “and a minimum amplitude”

We have updated the manuscript to reflect this change.