Strahan Susan E (Orcid ID: 0000-0002-7511-4577)
Douglass Anne R. (Orcid ID: 0000-0002-5556-9988)

Why do Antarctic Ozone recovery trends vary?

Susan E. Strahan[1,2], Anne R. Douglass[1], and Megan R. Damon[1,3]

[1]Atmospheric Chemistry and Dynamics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

[2]Universities Space Research Association, Columbia, MD, USA

[3]Science Systems and Applications, Inc., Lanham, MD, USA

Corresponding author: Susan Strahan (susan.e.strahan@nasa.gov)

Key Points:

- A simulation that reproduces observed Antarctic ozone identifies the recovery metric with consistent high sensitivity to halogen decline
- Polar cap and ozone mass deficit metrics' sensitivities to halogen decline vary from year to year, affecting their trends
- The greatest O_3 increase from halogen decline occurs just outside the Ozone Hole defined by the area where O_3 is <220 Dobson Units

Accepted Article

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2019JD030996

© 2019 American Geophysical Union. All rights reserved.

Abstract

We use satellite ozone records and Global Modeling Initiative (GMI) chemistry transport model (CTM) simulations integrated with MERRA2 meteorology to identify a metric that accurately captures the trend in Antarctic ozone attributable to the decline in ozone depleting substances (ODSs). The GMI CTM Baseline simulation with realistically varying ODS levels closely matches observed interannual to decadal scale variations in Antarctic September ozone over the past 4 decades. The expected increase or recovery trend is obtained from the differences between the Baseline simulation and one with identical meteorology and fixed 1995 ODS levels. The differences show that vortex-averaged column O_3 has the greatest sensitivity to ODS change from 1-20 Sep. The observed vortex-averaged column O_3 during this period produces a trend consistent with the expected recovery attributable to ODS decline. Trends from dates after 20 Sep have smaller sensitivity to ODS decline and are more uncertain due to transport variability. Simulations show that the greatest decrease in O_3 loss (*i.e.*, recovery) occurs inside the vortex near the edge. The polar cap metrics have vortex size-dependent bias and do not consistently sample this region. Because the 60-90^{o}S 220 DU O_3 mass deficit (OMD) metric does not sample the edge region, its trend is lower than the expected trend; this is improved by area weighting. The 250 DU OMD metric samples more of the edge region, which increases its trend. Approximately 25% of the September Antarctic O_3 increase is due to higher O_3 levels in June prior to winter depletion.

Plain Language Summary

Atmospheric levels of ozone depleting substances (ODSs) are decreasing due to the Montreal Protocol and its amendments. The Ozone Hole is beginning to shrink, but trends derived from Antarctic O_3 observations differ widely. We compared model simulations with different ODS levels to reveal where, when, and how much Antarctic O_3 increases in response to ODS decline. We found that trend studies used O_3 measurements from different dates and regions, which usually did not line up with regions where the greatest O_3 recovery occurs or with the time period when O_3 changes are most sensitive to declining ODSs. This explains why reported trends and their uncertainties vary so much.

Our simulations found the strongest, clearest signal of increasing O_3 due to ODS change between 1-20 Sep inside the Antarctic polar vortex. The observed vortex O_3 trend from 1-20 Sep is 1.2-1.6 DU/yr and is consistent with the simulations' results, indicating that O_3 is increasing at the rate expected due to ODS decrease. The greatest recovery occurs inside the vortex but just outside the O_3 220 DU contour, the traditionally defined boundary of the Ozone Hole. Simulations and observations agree that the September Ozone Hole is shrinking at 0.24-0.36 million km^2/yr.

1 Introduction

Due to the Montreal Protocol and its amendments, stratospheric loading of chlorine and bromine-containing ozone depleting substances (ODSs) have declined since reaching a peak in the 1990s. In the upper stratosphere, observations show a decline in HCl, the end product of chlorofluorocarbon (CFC) destruction (Froidevaux et al., 2006) and an increase in O_3 (Kyrölä et al., 2013; Gebhardt et al., 2014). In the past few years, more than one dozen ozone studies found increasing trends in Antarctic O_3 during September, although not all of the trends were statistically significant (Langematz and Tully in Chapter 4 of the Scientific Assessment of Ozone Depletion, 2018). It is on this basis that the World Meteorological Organization (WMO) assessment concludes that the early signs of the second stage of Antarctic ozone recovery are becoming apparent. Using observations as well as models, some

© 2019 American Geophysical Union. All rights reserved.

Accepted Article

studies have attributed some of the positive trend to the decline in ozone depleting substances (*e.g.*, Solomon et al., 2016; Strahan and Douglass, 2018).

Solomon et al. (2016) and Steinbrecht et al. (2018) note that the expected recovery of Antarctic O_3 attributable to ODS decline is about 0.5%/yr, or roughly 1 DU/yr. But reported Antarctic ozone recovery trends, derived from different data sets and analysis techniques have a wide range of values and uncertainties. Most studies determine a trend by applying linear or multiple linear regression (MLR) to a specific long term data record; the MLR uses proxies to account for dynamical and chemical drivers of O_3 variation. Satellite observations of total column O_3 are often used because they provide excellent coverage of the high latitudes in late winter (September) and spring. Various metrics are used to calculate recovery trends from satellite column O_3 data sets: 60-90^{o}S or 63-90^{o}S polar cap mean O_3 (Solomon et al., 2016; Chipperfield et al., 2017; Weber et al., 2018), the area with column O_3 <220 Dobson Units (DU) (*Solomon et al*., 2016), and ozone mass deficits (OMDs) with respect to 220 DU O_3 (DeLaat et al., 2017; Pazmiño et al., 2018). The WMO Assessment (Langematz and Tully, 2018) points out that many factors play a role in the different trends and uncertainties obtained by these studies, such as possible drifts in the observational data sets, the length of the data record, procedures used to merge data sets, and formulations of the regression models, including the choices for proxies. The report notes that the role of the regression model in trend uncertainty is so great that 2 studies using similar data records and MLR models, Chipperfield et al. (2017) and Weber et al. (2018), produced trends that varied by nearly a factor of 2. The SPARC (2019) report on long-term O_3 trends offers a substantial discussion of the impact of data sets and MLR models on trend calculations.

Even if there were no uncertainties arising from drifts, data set merging, length of the record, or MLR models, Antarctic O_3 recovery trends attributable to ODS decline would not necessarily agree because the metrics chosen for trend calculations use O_3 data with different spatial and temporal domains. Polar O_3 depletion depends on ODS amounts and temperature, but the balance between O_3 changes driven by chemistry and those driven by dynamics varies strongly with day of the year and latitude, and from year to year. *Solomon et al.* (2016) obtained a trend of 2.5 (1.5) DU/yr for 2000-2014 for the 63-90^{o}S September polar cap average, while Chipperfield et al. (2017) obtained a 4.7% (9.1)/decade (equal to ~1 DU/yr) trend using a slightly larger September polar cap (60-90^{o}S) and 1 additional year of observations. (Two standard deviations (2σ) of trend results are shown in parentheses.) Instead of a polar cap average, Pazmiño et al. (2018) averaged September column O_3 observations over the vortex and obtained trends from MLR that ranged from 1.8 (1.0) to 2.8 (1.5) DU/yr for 2001-2017; their trends for 15 Sep-15 Oct were about one third smaller. They attribute the wide range of trend results to their use of different MLR proxies and definitions for the vortex edge. Interestingly, they also fit trends for the OMD from 60^o-90^{o}S, which when converted to Dobson units produces September trends of 1.1 (0.5) DU/yr – only half as large as the vortex trends but with greater statistical significance. Some studies have determined trends for September through November (Salby et al., 2011; Kuttippurath and Nair, 2017), a period that captures all of the seasonal O_3 depletion but is subject to large year-to-year O_3 variations that vary with vortex stability, increasing the possibility that the trend will have a dynamical component. This dizzying array of trend results is a consequence of the specific dates and spatial boundaries that define each O_3 metric, leading each to have its own particular sensitivity to ODS decline and dynamical variability.

The trends reported in Langematz and Tully (2018) provide a preponderance of evidence that Antarctic O_3 is increasing but that the rate of increase is highly uncertain. A quantitative measure of Antarctic O_3 increase attributable to the known ODS decline is essential in order to demonstrate that the Montreal Protocol and its amendments are having

© 2019 American Geophysical Union. All rights reserved.

the intended impact. If Antarctic O_3 in the coming decades does not respond as we predict, then we must assess whether this indicates a change in stratospheric chemical or dynamical processes controlling O_3 or whether it points to unexpected emissions of regulated or unregulated ODSs (*e.g*., Montzka et al., 2018; Hossaini et al., 2019).

In this study we quantify the expected Antarctic ozone trend attributable to ODS decline using a simulation with realistic chemistry and dynamics. We contrast that simulation to 2 model perturbation experiments that use identical meteorology. The differences between the simulations isolate the chemical impact of declining ODSs on O_3, independent of dynamical variability and independent of climate change (*i.e*., upper stratospheric O_3 increases resulting from increasing CO_2/decreasing temperature). The simulations allow us to look under the hood of O_3 depletion to explore how well different observational metrics capture the full extent of ozone depletion and to see where and when recovery is occurring. We will show that some metrics have sensitivity to ODS decline that varies with vortex size, leading to increased trend uncertainty and possible bias. We identify the observed total column ozone metric that has the maximum sensitivity to ODS decline and the minimum sensitivity to dynamics. The trend calculated from this metric best captures the Antarctic ozone response attributable to declining ODSs and produces a statistically significant trend from an ordinary least squares linear fit without need for MLR.

2 Observations, Model Simulations, and Model Performance

All analyses and model evaluations in this study use Level 3 daily column O_3 measurements, 1980-2018, from the following satellite instruments: Nimbus-7 and EP Total Ozone Mapping Spectrometers (TOMS), Aura Ozone Measuring Instrument (OMI) (Bhartia, 2012), and Ozone Measurements Profiler Suite (OMPS) (Jaross, 2017). Column O_3 is measured only in the sunlit atmosphere. Version 4.2 pressure level O_3 measurements from the Aura Microwave Limb Sounder (MLS), 2004-2018, are also used; daily MLS observations 82^oS-82^oN are available in all months of the year. These data sets are available from the NASA Goddard Earth Science Data Center (https://disc.gsfc.nasa.gov/datasets). The Level 3 gridded O_3 record has 1.25^o longitude x 1^o latitude resolution from 1980-2004 and 1^ox1^o resolution from 2005 onward. Our observational analyses use Modern Era Retrospective for Research and Analysis 2 (MERRA2) assimilated ozone (which assimilates MLS O_3) to fill data gaps between satellite orbits and in polar night in the TOMS/OMI/OMPS data set (Wargan et al., 2015). The polar night area shrinks rapidly in September and an average 13% of the vortex from September 1-20 is filled by assimilated ozone; analyses using the TOMS/OMI/OMPS data with and without MERRA2 ozone filling produce the same trends and significance. MLS O_3 data from 261-0.46 hPa are used to calculate stratospheric columns; column accuracies are reported to be ±4% (Livesey et al., 2017). Tropospheric ozone columns from the surface to 268 hPa from the chemistry transport model (CTM) simulation described below are added to the MLS stratospheric columns to create total columns for MLS data for comparisons with the CTM and the TOMS/OMI/OMPS data set. The simulated Antarctic tropospheric columns typically vary from 16-18 DU. The CTM tropospheric O_3 columns were evaluated against global sonde observations and found to have no bias (Ziemke et al., 2014).

We use 3 simulations from the Global Modeling Initiative (GMI) CTM integrated with MERRA2 meteorology 1980-2018 (Gelaro et al., 2017) to separate the roles of transport and chemical impacts of ODSs in Antarctic O_3 trends. All simulations have 1^o x 1.25^o horizontal resolution and 72 hybrid sigma-pressure levels from the surface to 0.01 hPa. The

© 2019 American Geophysical Union. All rights reserved.

CTM uses an updated version of chemical mechanism described in Strahan et al. (2013); stratospheric sulfate surface area densities (Arfeuille et al., 2013) are the same as used in Chemistry-Climate Model Initiative (CCMI) simulations. Previous studies demonstrated the realism of polar chemistry and transport in GMI CTM MERRA and MERRA2 simulations (Strahan et al., 2013; Strahan et al., 2016; Strahan and Douglass, 2018). Changes in ODS-driven O_3 depletion are coupled to dynamics through radiation. Ozone depletion in Antarctic spring leads to reduced lower stratospheric temperatures, strengthening of the polar jet, delayed vortex breakup, and increased high latitude downwelling (Waugh et al., 1999; Stolarski et al., 2006). These impacts are imbedded in reanalysis meteorology and not separable in CTM simulations. The results here reflect only the chemical impacts on O_3 due to ODS decline.

The Baseline simulation is forced at the surface by realistic time-dependent mixing ratios for long-lived source gases (ODSs, N_2O, and CH_4). The High ODS simulation is initialized with the Baseline simulation on 1 Jan 1995 and is integrated from 1995-2018 with the same meteorology as Baseline, but forced at the surface with constant ODS mixing ratios from 1995. The only differences between the Baseline and High ODS simulations are the ODS levels. The O_3 differences between these simulations quantify O_3 recovery due to ODS decline independent of dynamical impacts on O_3. The 'No Het' simulation has the rates for heterogeneous halogen reactions with O_3 set to zero (*i.e.*, no polar O_3 depletion). It is integrated from May to December each year using initial conditions from the corresponding year of the Baseline simulation on 1 May. Changes in Antarctic O_3 in the No Het simulation are due only to transport and gas phase chemistry; the latter is insignificant in winter. Because No Het and Baseline have the same transport, the difference between them quantifies each year's heterogeneous O_3 depletion. The usefulness of vortex-averaged quantities for diagnosing polar chemical and transport processes was shown in our aforementioned studies. We use potential vorticity (PV) fields from MERRA2 to define the vortex edge on the 450 K potential temperature surface in the lower stratosphere (~70 hPa) where the greatest O_3 depletion occurs. All observations and model outputs in this study use PV of -3.0×10^{-5} kg $m^2 s^{-1}$ to define the vortex edge because it consistently locates the region of the steepest PV gradients in winter and is usually co-located with steep column O_3 gradients. A ±10% change in the PV value for the edge produces an approximate ±10% change in vortex area, but the calculated trend of the O_3 mean varies by less than ±1%.

September Antarctic vortex column O_3 values reflect seasonal depletion and transport changes that occurred since early winter. Figure 1a shows that simulated September vortex-averaged column O_3 from 1980-present matches the variability at interannual to decadal time scales in the O_3 data, with a consistent 13 DU low bias (~6%) except for the few years following the Mt. Pinatubo eruption in 1991. The black line showing MLS total column O_3 since 2004 falls on top of the TOMS/OMI/OMPS data set. Figure 1b compares MLS and simulated vortex column O_3 in June, after the vortex has formed but before significant seasonal depletion has begun. Their RMS difference is 3 DU, which is much smaller than the 4% (11 DU) MLS column measurement uncertainty. That the simulation shows no bias in June and develops a mean low bias of 13 DU by September suggests that depletion may be too great in winter, although errors in transport may contribute. (The bias is in the simulated stratospheric column because the MLS total columns use the GMI tropospheric columns.) Fig. 1c shows the year-to-year variations in vortex-averaged depletion and transport during winter; the 14-year seasonal mean change is removed from both model and data to emphasize the agreement in the year-to-year O_3 change. Both the simulation and observations show a roughly 20 DU reduction in ozone depletion during the Aura period (dashed lines). This is

© 2019 American Geophysical Union. All rights reserved.

consistent with the reduction in depletion derived from observations reported by Strahan and Douglass (2018).

3 Identifying Antarctic O_3 Recovery Attributable to ODS Decline

Ozone observations inside the Antarctic vortex will provide the strongest signal of recovery attributable to ODS decline. Ozone depletion occurs almost entirely inside the polar vortex because temperatures there are low enough to form the polar stratospheric clouds (PSCs) that transform halogen reservoir gases (*i.e.*, $ClONO_2$ and HCl) into reactive species while denitrifying and dehydrating the air. The PSC-processed air is generally confined by the vortex, allowing rapid catalytic ozone depletion when sunlight returns in late winter (Solomon, 1990; Schoeberl and Hartmann, 1991). Most of the ozone-depleted air remains inside the vortex because sharp gradients in potential vorticity inhibit mixing with air masses outside the vortex (Hartmann et al., 1989). The potential vorticity gradients weaken and mixing increases in September as the tropospheric waves that drive stratospheric dynamics increase.

We show the sensitivity of ozone depletion to ODS change by examining the differences between the GMI High ODS and the Baseline simulations. Dynamics affects Antarctic O_3 through transport and through the temperature dependence of PSC processing, but because Baseline and High ODS use the same meteorology, their differences are due only to changes in ODS-driven O_3 loss and not temperature or dynamical variability. Figure 2a shows the differences in vortex-averaged column O_3 in the two simulations as a function of day of the year 1995-2018. The maximum sensitivity of O_3 loss to ODS decline occurs between 1-20 September (dashed lines), dates which are centered in the 35-day maximum loss rate period determined by Hassler et al. (2011) using South Pole ozone sondes. Figure 2b explains why. The rate of O_3 loss in High ODS before mid-September is faster than in Baseline, increasing the O_3 difference between them. After that, the loss practically ceases in the High ODS lower stratosphere from 100-50 hPa because there is little O_3 left. But Baseline O_3 loss continues into October because there is still O_3 left to destroy; the continued loss in Baseline decreases the difference between the simulations. Although lower stratospheric depletion continues in Baseline, column O_3 increases in October in both simulations due to transport; this will be discussed in Section 4.2. The October sensitivity of O_3 loss to ODS decline is only 2/3 that of 1-20 Sep, making trend attribution more challenging for metrics using October data. Recovery metrics using 1-30 Sep data (*e.g.*, Solomon et al. (2016), Pazmiño et al. (2018)) will have ~10% lower sensitivity to ODS decline than metrics using data from 1-20 Sep.

Figure 2c shows the rate of Antarctic O_3 recovery in June and September using the simulation differences in Fig. 2a. The difference in Antarctic stratospheric ODS levels between the simulations in September is shown by the Equivalent Effective Stratospheric Chlorine (EESC) at 50 hPa (blue line). EESC is calculated using age-of-air dependent fractional release values and an age-of-air spectrum (Newman et al., 2007). Weak recovery in September is seen in the first 3-5 years after ODS emissions began to decline, reflecting the mean transit time from the surface to the Antarctic lower stratosphere. In this study we calculate trends starting in 2000 when the full impact of reduced ODS emissions has reached the Antarctic stratosphere. The reduction in depletion in the 5 years before 2000 is 1.7 DU. The rate of O_3 increase after 2000 is linear at 1.22 DU/yr, with a one-sided 95% confidence interval of 0.01. The change in EESC after 2000 is also linear (26 ppt/yr). The ratio of these

© 2019 American Geophysical Union. All rights reserved.

lines is the vortex-averaged column O_3 sensitivity to ODSs. The simulation's sensitivity is 46.9 DU (1.3) O_3/ppb EESC from 2000-2018. Notice that about 25% of the O_3 increase occurs in June prior to winter depletion (6.5 DU in 2018), suggesting that global O_3 recovery has increased Antarctic O_3 at the time of vortex formation in late fall.

Figure 2a indicates that vortex averaged column O_3 from 1-20 Sep is an optimal metric for identifying the trend attributable to ODS decline because of the maximum sensitivity to ODSs during this period. The TOMS/OMI/OMPS data are suitable for this metric because sampling reaches 77°S on 1 Sep and 84°S by 20 Sep, covering an average of 87% of the vortex during this period. Unlike the highly significant trend determined by model differences, trends calculated from observations have greater uncertainty because of the large interannual O_3 variability (IAV) that comes from dynamics; in addition, the trend may include bias if there is a trend in dynamics.

We test whether this metric accurately identifies the trend attributable to ODSs in a record with large IAV by comparing the trend in Baseline with the trend determined by subtracting High ODS from Baseline O_3 (Fig. 2a). Fig. 3a shows the Baseline trend is 1.21 (0.20) DU/yr (95% confidence interval (CI)) during this period, in close agreement with the known model trend of 1.22 DU/yr discussed above. The uncertainty is much larger because Baseline, like the observations, has large dynamically-driven IAV. All trends calculated in this study are reported with 95% confidence intervals.

Figure 3a also shows the vortex averaged O_3 from observations (blue). The observed trend of 1.39 (0.22) DU/yr is slightly higher than the Baseline trend but their confidence intervals overlap, indicating good model agreement with the observations. We calculate the trend in High ODS to see whether a trend in dynamics may contribute to the observed trend. The simulations and the observations have very similar IAV. The High ODS trend in Fig. 3b (black) is 0.0 (0.2), suggesting that any dynamical contribution to the trend cannot be detected from the 19-year time period given the O_3 IAV. The observed vortex-averaged trend is between 1.18 and 1.61 DU/yr (95% CI), consistent with the expected response to declining ODSs determined from the simulations.

We tested the observed trend's sensitivity to the choice of meteorological reanalysis used to define the vortex edge by using ERA-Interim reanalysis PV fields (Dee et al., 2011) to calculate the observed vortex averaged 1-20 Sep column O_3 ('O_3/ERA-I'). For 2000-2018, vortex averaged O_3/ERA-I ranges from 8 DU lower to 2 DU higher than the O_3/MERRA2 record in Fig. 3a (mean difference -1.3 DU (0.6%)) and it has a trend of 1.35 to 1.81 DU/yr (95% CI). This trend is 13% larger than the O_3/MERRA2 trend but the trends agree to within their uncertainties. Severe O_3 depletion in 2000 gives each record a very low starting point, increasing both analyses' trends for shorter data records: from 1.4-1.7 for O_3/MERRA2 and 1.8-2.3 DU/yr for O3/ERA-I for records ending before 2015. For records ending after 2015, trends calculated for both O_3/PV analyses decreased, their differences shrank, and their CIs narrowed. As the data records have lengthened, the impact of the choice of reanalysis product on the observed trend has diminished.

4 Why Different O_3 Metrics Produce Different Trends

Different O_3 recovery metrics have different balances between their sensitivity to ODS decline and dynamical variability because of the spatial and temporal boundaries that

© 2019 American Geophysical Union. All rights reserved.

define them. In this section we use the GMI simulations to investigate how well several commonly used ozone metrics capture ozone loss attributable to ODSs and the extent to which they are affected by dynamics. The metrics examined are the polar cap average column O_3, the area with column O_3 less than 220 DU, and the ozone mass deficit relative to 220 DU O_3 over 60-90^{o}S. Trends calculated with OMD are converted to DU for comparison with other metrics (1 DU = 2.144 x 10^{-2} metric tons (MT) O_3/km^2 and the 60-90^{o}S area is 34.4 million km^2). Section 4.1 compares metrics and their trends using O_3 observations from 1-20 Sep for comparison with the vortex-averaged results in the previous section. Section 4.2 explores the impact of the choice of dates on trend results.

4.1 Trend Sensitivity to Spatial Boundaries

The locations of each metric's spatial boundaries with respect to the distributions of simulated column O_3 and O_3 depletion in a strong (cold) and in a disturbed (warm) vortex reveal weaknesses in the metrics. The left panels of Figure 4 show O_3 averaged over the 2nd week of September in a warm, small vortex (2017) and in a cold, large vortex (2018), but with similar ODS levels. The location of the 2017 vortex edge in this week averaged 2.5^o poleward of the 2018 vortex. The O_3 depletion for the same week (right panels) is the difference between the Baseline and No Het simulations. The 4 boundaries used by the metrics discussed here are: vortex edge (red), 63^{o}S polar cap (yellow dashed), 60^{o}S polar cap (orange dashed), and 220 DU contour (white).

The Antarctic vortex is rarely circular and centered on the pole even in a cold year. Figs. 4b and 4d show that the 60^{o}S and 63^{o}S polar cap boundaries include some non-vortex air whether the year is warm or cold. (Air north of the red contour in Fig. 4 is non-vortex air.) In cold years when the vortex is larger, the 63-90^{o}S cap excludes regions of significant O_3 depletion except where non-vortex air protrudes southward (150-200^{o}E in Fig. 4d); thus, the small polar cap is less sensitive to ODS decline in years with a large vortex. The 60-90^{o}S cap captures more of the large loss area in a cold year but has reduced sensitivity to ODS decline in a warm years because it includes so much non-vortex air. Neither metric reliably captures all regions with large O_3 depletion as well as the vortex edge boundary does. During cold years the 220 DU contour parallels the vortex edge fairly well and captures the areas of greatest O_3 loss (Fig. 4d), but in a dynamically disturbed year that contour is deeper inside the vortex and excludes regions of large depletion (0-150^{o}E in Fig. 4b). Regardless of vortex size, the vortex edge always parallels O_3 loss contours of 50-70 DU during the 2000-2018 period, with recent years following 50-60 DU contours because some recovery has taken place. Unlike polar cap averages and the 220 DU contour, the vortex average consistently samples all locations of O_3 depletion >70 DU while consistently excluding non-vortex air. The polar cap and 220 DU-based metrics have less sensitivity to O_3 depletion than the vortex-averaged metric does, and their sensitivities vary each year depending on vortex size and shape, increasing their trends' uncertainties.

Fig. 5a compares trends in the observed vortex-averaged O_3 data (from Fig. 3a) and the two polar caps. The 60^{o}S cap trend is close to that of the vortex average but has a wider confidence interval; the 63^{o}S cap trend is larger than both and its 95% confidence interval is 1.21-1.95 DU/year. The larger uncertainties in the polar cap trends are explained by Fig. 5b. The differences between both polar cap metrics and the vortex average show a clear dependence on the mean vortex edge latitude (*i.e.*, vortex size). Compared to the vortex-averaged column O_3, the 63^{o}S cap has an average -10 DU low bias while the 60^{o}S metric has a mean +5 DU high bias. The average vortex edge latitude during these dates is 59.3^{o}S,

© 2019 American Geophysical Union. All rights reserved.

explaining why the 60°S cap trend is closer to the vortex-average trend. The dependence of polar cap metrics on vortex size adds uncertainty to their trends.

Fig. 5b shows that the 63-90°S cap is biased up to 20 DU low when the vortex is large. The years 2000 and 2001 had a very large vortex (63°S cap is 20-22 DU low-biased) while the years 2012-2014 had small ones (63°S cap is 6-8 DU low-biased). As a result of its dependence on vortex size, the 63-90°S cap O_3 trend computed for the 2000-2014 period has a high bias of nearly +1 DU/yr. This may explain why the 2.5 DU/yr trend in observed September column O_3 from 2000-2014 reported by Solomon et al. (2016) was twice the size of the 'chemistry only' trend from their chemistry climate model. (This study excluded 2002 data.) The 63-90°S polar cap trend calculated from the 2000-2018 record that includes 2002 is 1.58 DU/yr, about 1/3 lower than the trend from the shorter record. This value falls just inside the confidence interval of the vortex-averaged trend.

Figure 6 examines two frequently used recovery metrics: the area inside the 220 DU contour and the OMD inside the 220 DU contour. Fig. 6a shows that area trends attributable to ODS decline are best calculated using an average of daily September 220 DU O_3 areas rather than the 220 DU area of the monthly averaged September O_3. The trend attributable to ODS change only (not dynamics) is the trend in the area difference between the Baseline and High ODS simulations: -0.29 to -0.27 million km^2/yr (blue line). The trend result for 1-20 Sep daily observed areas is -0.36 to -0.24 million km^2/yr (95% CI), in excellent agreement with the model result. The trend calculated from the 1-30 Sep daily observed areas is the same as for 1-20 Sep (not shown). However, the area inside the 220 DU contour determined from monthly averaged September data (red), the metric used by Solomon et al. (2016), is not the same quantity as the average of the daily area inside the 220 DU contour. The monthly averaged area (with 2002 excluded as in *Solomon et al*.) produces a faster trend, -0.45 to -0.33 million km^2/yr, than the trend from daily observations and is outside of the model trend attributable to ODS change only. Note that the monthly averaged area is much smaller than the area calculated with daily data in 2002 and 2017, the 2 most dynamically disturbed years of the period.

Figure 6b shows the time series of the 60°-90°S OMD for O_3 <220 DU (blue). We calculate OMD over 60°-90°S as in Pazmiño et al. (2018), however, we find that the OMD trend is not substantially different if the latitude range is unlimited. Our OMD trend is 0.81 MT/yr, which we compare to our other trends by dividing by the 60°-90°S area and converting O_3 mass to molecules. The resulting deficit trend, -1.06 (0.15) DU/yr, is 25% lower than the vortex-average O_3 trend (Fig. 3a, blue) and the confidence intervals have only a 0.04 DU/yr overlap. (The negative of the O_3 deficit trend is equivalent to a trend in increasing O_3.) The OMD trend confidence interval is similar to that of the vortex average and both have much smaller CIs than the polar cap average trend. Figure 4 explains why. The 220 DU contour roughly parallels the vortex edge but at a higher latitude. Because O_3 and PV contours behave similarly, this metric is less sensitive to vortex size, eliminating some of the dynamical variability that affects fixed latitude metrics. And like the vortex-averaged metric it never includes non-vortex air. However, Figs. 4b and 4d reveal the downside to the 220 DU OMD: it often excludes areas of substantial O_3 depletion. We will later show that the region of greatest depletion near the vortex edge is also the region where the greatest *change* in O_3 loss (*i.e*., recovery) is occurring.

There are 2 ways to improve the OMD diagnostic: by taking advantage of its correlation with area and by considering an O_3 deficit with respect to a higher O_3 threshold, one that is found closer to the vortex edge. The area IAV in Fig. 6a (black) and the 220 DU

© 2019 American Geophysical Union. All rights reserved.

OMD (Fig. 6b, blue) have a correlation of 0.95. This suggests that weighting the OMD by the area inside the contour, rather than the 60-90°S polar cap area, may reduce dynamical variability as well as increase its magnitude. Fig. 6b shows that the area-weighted 220 DU OMD (black) has an 11% greater O_3 recovery rate with a slightly narrower confidence interval (as a percentage of the trend) compared to the unweighted trend. It is also significantly closer to the vortex-averaged trend. Figure 4 shows that vortex edge falls within 260-300 DU O_3 contours, putting the 250 DU contour inside the vortex but a few degrees closer to the edge than the 220 contour. This allows the 250 DU OMD metric to capture more of the region with the greatest sensitivity to ODSs. The 250 DU OMD trend (red) is slightly larger and has a 95% confidence interval of 1.04-1.38, overlapping well with the 1.17-1.61 DU/yr range for the vortex average. The higher O_3 deficit threshold ameliorates some of the low bias without an increase in uncertainty.

We use the difference between the Baseline and High ODS simulations to show the location of O_3 recovery. Figure 7 shows that the greatest O_3 recovery, 25-35 DU, occurs between the vortex edge (white) and the observed 220 DU contour (red). This region is the most sensitive to changes in ODSs because it receives the most sunlight, which is required for the catalytic cycles that destroy O_3 (Newman et al. 2004). The recovery trend inside the 220 DU contour is smaller than a trend that includes the edge region because it is deeper inside the vortex and losses there are lower. The 220 DU trend will always be low biased relative to the vortex averaged trend. Averaged over 1-20 Sep, the area inside the 207 DU contour in the Baseline simulation is 4.7 million km^2 smaller than the 207 DU area in the High ODS simulation. (We use 207 rather than 220 DU to account for the 13 DU low bias in the Baseline simulation. The Baseline 207 DU contour has an area very close to the observed 220 DU area). This suggests that the 2018 O_3 hole was reduced by 17% as a result of the Montreal Protocol and its amendments

4.2 Trend Sensitivity to Dates

Ozone recovery trends reflect the balance between ozone's sensitivity to ODS-driven loss and to variations in transport, and that balance depends on the dates chosen for the metric. The previous section demonstrated that some metrics' sensitivity to ODS decline depends on vortex size, which increases the trend's uncertainty. Here we show how increased dynamical variability after 20 Sep increases uncertainty in all trends. None of the metrics examined is affected by vortex breakdown, which occurred no earlier than November except in 2002. Data from 2002 are often excluded from trend analyses.

We isolate the effects of interannual variations in O_3 transport by examining vortex-averaged O_3 changes in No Het from September to October and comparing them to changes in Baseline. Fig. 8a shows vortex-averaged O_3 depletion in winter and spring for all years since 1995; this is the difference between No Het and Baseline in each year. White dashed lines mark the 20-day period with the greatest depletion of the entire season, 25 Sep to 14 Oct. This period is 24 days later than the period of greatest sensitivity to ODS change, 1-20 Sep, marked by red dashed lines. Fig. 8b shows the amount of O_3 transported into the vortex since June in the No Het simulation. (Transport is primarily what affects No Het vortex O_3 during this period; there are few DU of gas phase O_3 loss but this opposes the effect of transport.) Transport increases vortex-averaged O_3 by generally less than 30 DU from June to 20 Sep, but by Oct 14 up to 47 DU more O_3 has arrived. The bottom panel quantifies the competing effects of transport and depletion. The red line in Fig. 8c shows the mean O_3 increase due to transport between the earlier period (1-20 Sep) the later period (25 Sep-14 Oct). The black line shows the increase in O_3 depletion between these 2 periods. From 2000-

© 2019 American Geophysical Union. All rights reserved.

Accepted Article

2018 the mean O_3 increase due to transport and the mean decrease due to loss are the same (18 DU), so on average these effects cancel out. But the interannual variability of transport is 3 times greater than that of depletion, causing trends calculated during the latter period to have more dynamical noise while on average any increase in signal due to O_3 depletion will be negated, resulting in no increase in the signal due to O_3 recovery. Worse still, the O_3 sensitivity to ODS decline in October is about 2/3 what it is during 1-20 Sep (see Fig. 2), so the trend attributable to ODS decline will be smaller. Our results here explain why reported October trends are statistically insignificant and smaller than September trends, as noted in Langematz and Tully (2018). The Baseline and No Het results shown in Fig. 8 demonstrate that regardless of metric, the choice of dates after 20 Sep will not improve the accuracy or increase the certainty in O_3 trends attributable to ODS decline.

5 Conclusions

It is essential to have a quantitative and accurate metric for assessing whether Antarctic O_3 recovery is proceeding as expected. Here we used model simulations to identify the metric that best captures O_3 changes attributable to ODS decline. By looking at differences between simulations that had different ODS levels or that had no ODS-driven depletion we identified 1) the dates when O_3 depletion is most sensitive to ODS decline, 2) the dates when seasonal O_3 depletion reaches a maximum, 3) how transport's influence on Antarctic O_3 varies, and 4) the location where the greatest recovery is occurring. These results, summarized in Table 1, allowed us to understand what different O_3 metrics measure and explain why their trends differ.

The simulations revealed that ozone's greatest sensitivity to ODS change occurs from 1-20 Sep and that the vortex-averaged total column O_3 from these dates accurately quantifies the impact of declining ODSs. The 1-20 Sep vortex-averaged column O_3 metric includes essentially all the area where loss occurs and its sensitivity to ODSs does not depend on vortex size. A linear fit to the 2000-2018 TOMS/OMI/OMPS data set produced an O_3 trend of 1.17-1.61 DU/yr (95% CI), in agreement with the expected recovery rate of 1.2 DU/yr calculated from the simulations. This metric's trend was calculated with data from all years since 2000, including 2002, the year excluded from many trend studies due to the rare major stratospheric warming that occurred around 25 Sep (Allen et al., 2003). By differencing the Baseline and the fixed high ODS simulations, we found that the Montreal Protocol and its amendments reduced the area of O_3 depletion from 1-20 Sep 2018 by 17%, or 4.7 million km^2. Because the CTM simulations use reanalysis meteorology and therefore cannot represent or quantify the dynamical effects on O_3 induced by depletion, this result is an estimate of the area change due to ODS decline.

The simulations showed that the largest O_3 recovery is occurring near the vortex edge. Trends calculated from polar cap metrics during the optimal 1-20 Sep period do not have consistent sensitivity to ODS decline from year to year because 1) they do not capture the entire area where depletion occurs, 2) they include O_3 observations where heterogeneous O_3 loss does not occur (*i.e.*, outside the vortex), and 3) their ODS sensitivity depends on vortex size, which has large interannual variability. The year-to-year variations in the sensitivity to ODS decline lead to increased uncertainties in trends calculated from polar cap averages relative to other metrics.

Accepted Article

© 2019 American Geophysical Union. All rights reserved.

The 60-90°S 220 DU OMD trend has less impact from dynamical variability because the 220 DU boundary acts somewhat like a PV contour, which reduces variability in the metric and uncertainty in its trend compared to polar cap averages. Figure 4 showed that O_3 contours are quasi-parallel to the vortex edge, which is often near the 280 DU O_3 contour. The variability in OMD is highly correlated with the area variability inside the 220 DU contour. Area-weighting the 220 DU OMD increases its trend by 11% and slightly reduces its confidence interval. Its trend is lower than the vortex-averaged trend but overlaps the vortex-averaged trend's 95% CI. Because the 220 DU contour is always several degrees inside the vortex, the boundary of this metric always excludes the region with the largest recovery, biasing its trend low compared to the vortex-averaged trend. The simulations showed 25-35 DU O_3 recovery in the region between the 220 DU contour and the vortex edge in 2018, while the area inside the 220 DU contour had 15-25 DU O_3 recovery. The area-weighted OMD inside the 250 DU contour is a slightly better metric than the 220 DU OMD because it includes more of the critical recovery region near the vortex edge while continuing to exclude non-vortex air.

We showed the disadvantages of choosing dates later than 20 Sep for quantifying O_3 recovery attributable to ODS decline, even though O_3 depletion in many years continues past this date. Transport causes O_3 to increase in all metrics after this date; the O_3 increase has large IAV and it always opposes additional O_3 depletion. Choosing dates beyond 20 Sep adds noise to the O_3 time series while on average adding no recovery signal. And because O_3 is less sensitive to ODS decline in October, its trend will be smaller and harder to detect. A longer data record will be required to obtain a statistically significant October trend.

Vortex-averaged column O_3 is an appropriate metric for evaluating recovery trends in chemistry climate models that output daily O_3 and PV fields. For those models that do not save PV fields, the area-weighted 220 or 250 DU OMD for 1-20 Sep is recommended. These 2 metrics have smaller interannual variability than the polar cap metrics and will produce a meaningful trend over shorter intervals; as of 2018 their trends are ~15% lower than the vortex average but within its 95% CI. Antarctic O_3 recovery consists of 2 components: a reduction in heterogeneous chemical loss during winter and an increase in O_3 levels at the beginning of winter, the latter reflecting the impact of global O_3 recovery on the Antarctic. Additional insight into CCMs' performances and reasons for variations in their recovery rates may come from separately evaluating their Antarctic June trends and their trends in winter (June to 1-20 Sep). We found that roughly one quarter of the recovery trend, ~0.3 DU/yr, is present in the June vortex prior to seasonal O_3 depletion. This result, based on simulations, cannot be confirmed by observations (*e.g.*, Fig. 1b) because this trend is small compared to observed interannual variability.

Acknowledgments and Data

This work was supported by the NASA Modeling, Analysis, and Prediction Program and the NASA Atmospheric Composition Modeling and Analysis Program. All Ozone and MERRA2 data sets used here are available from the NASA Goddard Earth Science Data Center (https://disc.gsfc.nasa.gov/datasets). See https://gmi.gsfc.nasa.gov for information on accessing GMI simulations.

Accepted Article

© 2019 American Geophysical Union. All rights reserved.

References

Allen, D., Bevilacqua, R., Nedoluha, G., Randall, C. & Manney, G. (2003). Unusual stratospheric transport and mixing during 2002 Antarctic winter, *Geophysical Research Letters*, *30*, 1599, https://doi.org/10.1029/2003GL017117

Arfeuille, F., Luo, B. P., Heckendorn, P., Weisenstein, D., Sheng, J. X., Rozanov, E., Schraner, M., Brönnimann, S., Thomason, L.W., and Peter, T. (2013). Modeling the stratospheric warming following the Mt. Pinatubo eruption: uncertainties in aerosol extinctions, *Atmospheric Chemical Physics, 13,* 11221–11234, doi:10.5194/acp-13-11221-2013

Bhartia, P.K. (2012), OMI/Aura TOMS-Like Ozone, Aerosol Index, Cloud Radiance Fraction L3 1 day 1 degree x 1 degree V3, NASA Goddard Space Flight Center, Goddard Earth Sciences Data and Information Services Center (GES DISC), 10.5067/Aura/OMI/DATA3001.

Chipperfield, M., Bekki, S., Dhomse, S., Harris, N.R.P., Hassler, B., Hossaini, R., Steinbrecht, W., Thiéblemont, R., & Weber, M. (2017). Detecting recovery of stratospheric ozone layer, *Nature*, *549*, 211–218, https://doi.org/10.1038/nature23681

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society, 137*, 553-597.

de Laat, A. T. J., van Weele, M. & van der A, R.J. (2017). Onset of stratospheric ozone recovery in the Antarctic ozone hole in assimilated daily total ozone columns, *Journal of Geophysical Research, 122*, https://doi.org/10.1002/2016JD025723

Froidevaux, L., Livesey, N.J., Read, W.G., Salawitch, R.J., Waters, J.W., Drouin, B., et al. (2006). Temporal decrease in upper atmospheric chlorine, *Geophysical Research Letters, 33*, doi:10.1029/2006GL027600

Gebhardt, C., Rozanov, A., Hommel, R., Weber, M., Bovensmann, H., Burrows, J.P., Degenstein, D., Froidevaux, L., & Thompson A.M. (2014). Stratospheric ozone trends and variability as seen by SCIAMACHY from 2002 to 2012, *Atmospheric Chemical Physics, 14*, 831-846.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2), *Journal of Climate, 30(14)*, 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Hassler, B., Daniel, J.S., Johnson, B.J., Solomon, S., & Oltmans, S.J. (2011). An assessment of changing ozone loss rates at South Pole: Twenty-five years of ozonesonde measurements, *Journal of Geophysical Research, 116*, doi:10.1029/2011JD016353

Hartmann, D.L., Chan, K.R., Gary, B.L., Schoeberl, M.R., Newman, P.A., Martin, R.L., Loewenstein, M., Podolske, J.R., & Strahan, S.E. (1989). Potential vorticity and mixing in the south polar vortex during spring, *Journal of Geophysical Research, 94*, 11,625-11,640.

Hossaini, R., Atlas, E., Dhomse, S.S., Chipperfield, M.P., Bernath, P.F., Fernando, A.M., et al. (2019). Recent trends in stratospheric chlorine from very short-Lived substances, *Journal of Geophysical Research, 124*, 2318-2335, doi:10.1029/2018JD029400

© 2019 American Geophysical Union. All rights reserved.

Accepted Article

Jaross, G. (2017). OMPS-NPP L3 NM Ozone (O_3) Total Column 1.0 deg grid daily V2, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC).

Kuttippurath, J. and Nair, P.J. (2017). The signs of Antarctic ozone hole recovery, *Scientific Reports, 7*: 585, doi:10.1038/s41598-017-00722-7

Kyrölä, E., Laine, M., Sofieva, V., Tamminen, J., Päivärinta, S.-M., Tukiainen, S., Zawodny, J., & Thomason, L. (2013). Combined SAGE II–GOMOS ozone profile data set for 1984–2011 and trend analysis of the vertical distribution of ozone, *Atmospheric Chemical Physics, 13,* 10645-10658.

Langematz, U. and Tully, M. (Lead Authors), Calvo, N., Dameris, M., de Laat, A.T.J., Klekociuk, A., Müller, R., & Young, P. (2018). Polar Stratospheric Ozone: Past, Present, and Future, Chapter 4 in Scientific Assessment of Ozone Depletion: 2018, Global Ozone Research and Monitoring Project–Report No. 58, *World Meteorological Organization*, Geneva, Switzerland.

Livesey, N., Read, W.G., Wagner, P.A., Froidevaux, L., Lambert, A., Manney, G.L., et al. (2017). Earth Observing System (EOS) Aura Microwave Limb Sounder (MLS) version 4.2x level 2 data quality and description document JPL D-33509 Rev C.

McPeters, R.D., Frith, S., & Labow, G.J. (2015). OMI total column ozone: extending the long-term data record, *Atmospheric Measurement Techniques, 8*, 4845–4850.

Montzka, S.A., Dutton, G.S., Yu, P., Ray, E., Portmann, R.W., Daniel, J.S., et al. (2018). An unexpected and persistent increase in global emissions of ozone-depleting CFC-11, *Nature, 557*, 413-417.

Nedoluha, G.E., Connor, B.J., Mooney, T., Barrett, J.W., Parrish, A., Gomez, M., et al. (2016). 20 years of ClO measurements in the Antarctic lower stratosphere, *Atmospheric Chemical Physics, 16*, 10725-10734.

Newman, P. A., Daniel, J. S., Waugh, D. W., & Nash, E.R. (2007). A new formulation of equivalent effective stratospheric chlorine (EESC), *Atmospheric Chemical Physics, 7(17)*, 4537–4552, https://doi.org/10.5194/acp-7-4537-2007

Pazmiño, A., Godin-Beekmann, S., Hauchecorne, A., Claud, C., Khaykin, S., Goutail, F., Wolfram, E., Salvador, J., & Quel, E. (2018). Multiple symptoms of total ozone recovery inside the Antarctic vortex during austral spring, *Atmospheric Chemical Physics, 18*, 7557-7572.

Salby, M., Titova, E., & Deschamps, L. (2011). Rebound of Antarctic ozone, *Geophysical Research Letters, 38*, L09702, doi:10.1029/2011GL047266

Schoeberl, M.R. and Hartmann, D.L. (1991). The dynamics of the stratospheric polar vortex and its relation to springtime ozone depletions, *Science, 251*, 46-52.

Solomon, S. (1990). Progress toward a quantitative understanding of Antarctic ozone depletion, *Nature, 347*, 347-354.

Solomon, S., Ivy, D.J., Kinnison, D., Mills, M.J., Neely III, R.R., & Schmidt A. (2016). Emergence of healing in the Antarctic ozone layer, *Nature, 252(6296)*, 269–274. https://doi.org/10.1126/science.aae0061SPARC/IO3C/GAW (2019). SPARC/IO3C/GAW Report on Long-term Ozone Trends and Uncertainties in the Stratosphere. I. Petropavlovskikh, S. Godin-Beekmann, D. Hubert, R. Damadeo, B.

Accepted Article

© 2019 American Geophysical Union. All rights reserved.

Hassler, V. Sofieva (Eds.), SPARC Report No. 9, GAW Report No. 241, WCRP-17/2018, doi: 10.17874/f899e57a20b

Steinbrecht, W., Hegglin, M.I., Harris, N., & Weber, M. (2018). Is global ozone recovering?, *Comptes Rendus Geoscience, 360,* 368-375.

Stolarski, R.S., Douglass, A.R., Gupta, M., Newman, P.A., Pawson, S., Schoeberl, M.R., and Nielsen, J.E. (2006). An ozone increase in the Antarctic summer stratosphere: A dynamical response to the ozone hole, *Geophysical Research Letters, 33,* doi:10.1029/2006GL026820

Strahan, S.E. and Douglass, A.R. (2018). Decline in Antarctic ozone depletion and lower stratospheric chlorine determined from Aura Microwave Limb Sounder observations, *Geophysical Research Letters*, 44, doi:10.1002/2017GL074830

Strahan, S.E., Douglass, A.R., & Newman, P.A. (2013). The contributions of chemistry and transport to low Arctic ozone in March 2011 derived from Aura MLS Observations, *Journal of Geophysical Research, 118,* doi:10.1002/jgrd.50181

Strahan, S. E., Douglass, A.R., & Steenrod, S.D. (2016). Chemical and Dynamical Impacts of Stratospheric Sudden Warmings on Arctic Ozone Variability, *Journal of Geophysical Research, 121,* 11,836–11,851, doi:10.1002/2016JD025128

Wargan, K., Pawson, S., Olsen, M.A., Witte, J.C., Douglass, A.R., Ziemke, J.R., Strahan, S.E., & Nielsen, J.E. (2015). The Global Structure of UTLS Ozone in GEOS-5: A Multi-Year Assimilation of EOS-Aura Data, *Journal of Geophysical Research, 120,* doi: 10.1002/2014JD022493

Waugh, D.W., Randel, W.J., Pawson, S., Newman, P.A., and Nash, E.R. (1999). Persistence of the lower stratospheric polar vortices, *Journal of Geophysical Research, 104,* 27,191-27,201.

Weber, M., Coldewey-Egbers, M., Fioletov, V.E., Frith, S.M., Wild, J.D., Burrows, J.P., Long, C.S., & Loyola, D. (2018). Total ozone trends from 1979 to 2016 derived from five merged observational datasets – the emergence into ozone recovery, *Atmospheric Chemical Physics, 18,* 2097–2117, https://doi.org/10.5194/acp-18-2097-2018

Ziemke, J. R., Olsen, M.A., Witte, J.C., Douglass, A.R., Strahan, S.E., Wargan, K., et al. (2014). Assessment and applications of NASA ozone data products derived from Aura OMI/MLS satellite measurements in context of the GMI chemical transport model, *Journal of Geophysical Research, 119,* 5671–5699, doi:10.1002/2013JD020914

Accepted Article

© 2019 American Geophysical Union. All rights reserved.

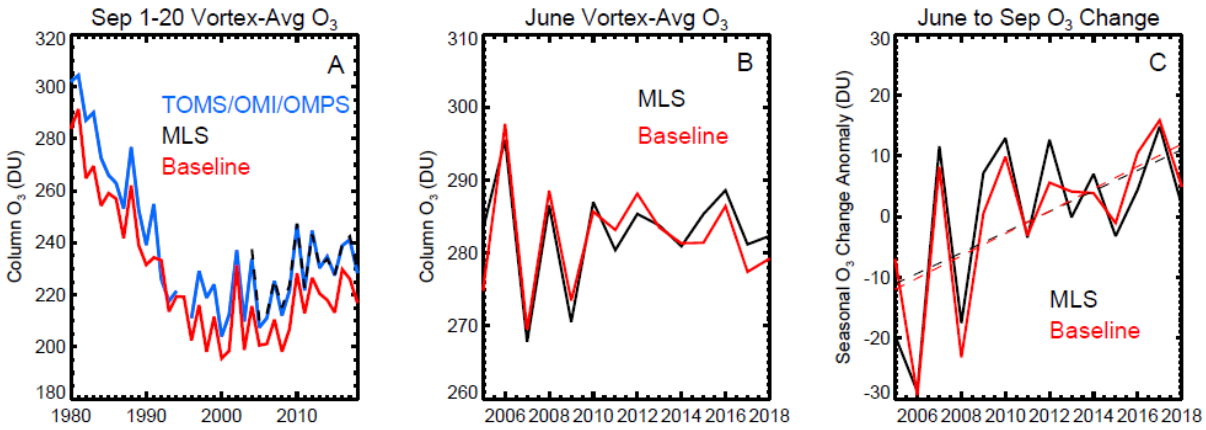


Figure 1. Observed and simulated Antarctic vortex-averaged column O_3 a) averaged over 1-20 Sep, 1980-2018 and b) averaged over 1-30 June, 2005-2018. c) Vortex column O_3 change 2005-2018 between June and Sep 1-20, for MLS and GMI with 14-yr mean removed. The dashed lines are linear fits to MLS and GMI points. The slopes of the lines, which are the same for the simulation and observations, indicate a ~20 DU reduction in winter O_3 loss from 2005-2018.

Accepted Article

© 2019 American Geophysical Union. All rights reserved.

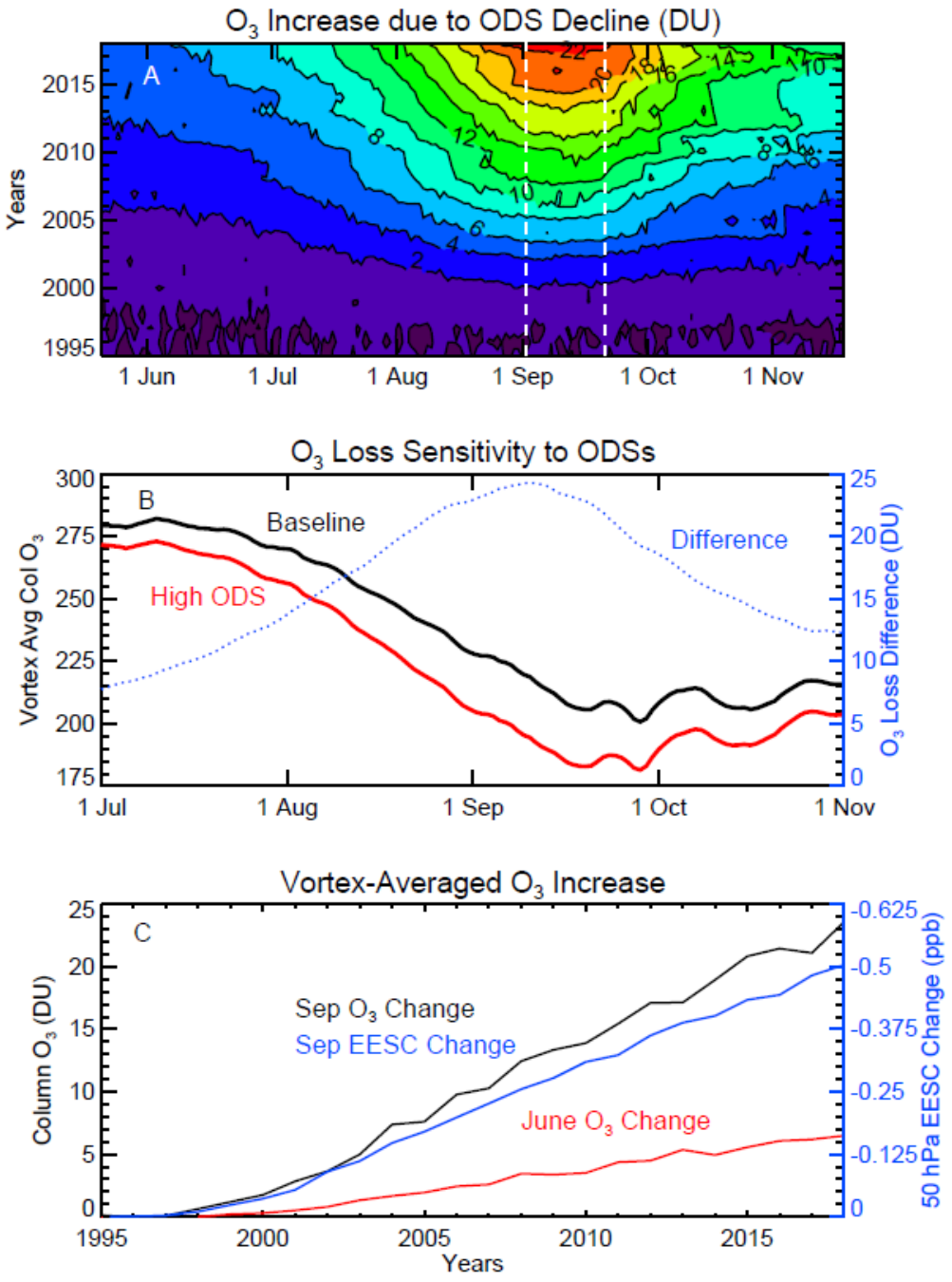


Figure 2. a) The difference between vortex-averaged column O_3 in the High ODS and Baseline simulations as a function of day of the year since 1995. The period of greatest ODS sensitivity is highlighted by the dashed lines. b) Vortex-averaged column O_3 in the Baseline and High ODS simulations in 2018. Their difference (blue) shows that O_3 loss sensitivity to ODS decline peaks around 10 Sep. Before mid-September, the High ODS O_3 loss is faster so the differences increase, but after mid-September loss continues at a greater rate in Baseline because O_3 levels are higher causing the differences to decrease. c) O_3 differences from a) averaged over 1-20 Sep (black) and 1-30 June (red) and EESC change (Baseline-High ODS, blue).

Accepted Article

© 2019 American Geophysical Union. All rights reserved.

Accepted Article

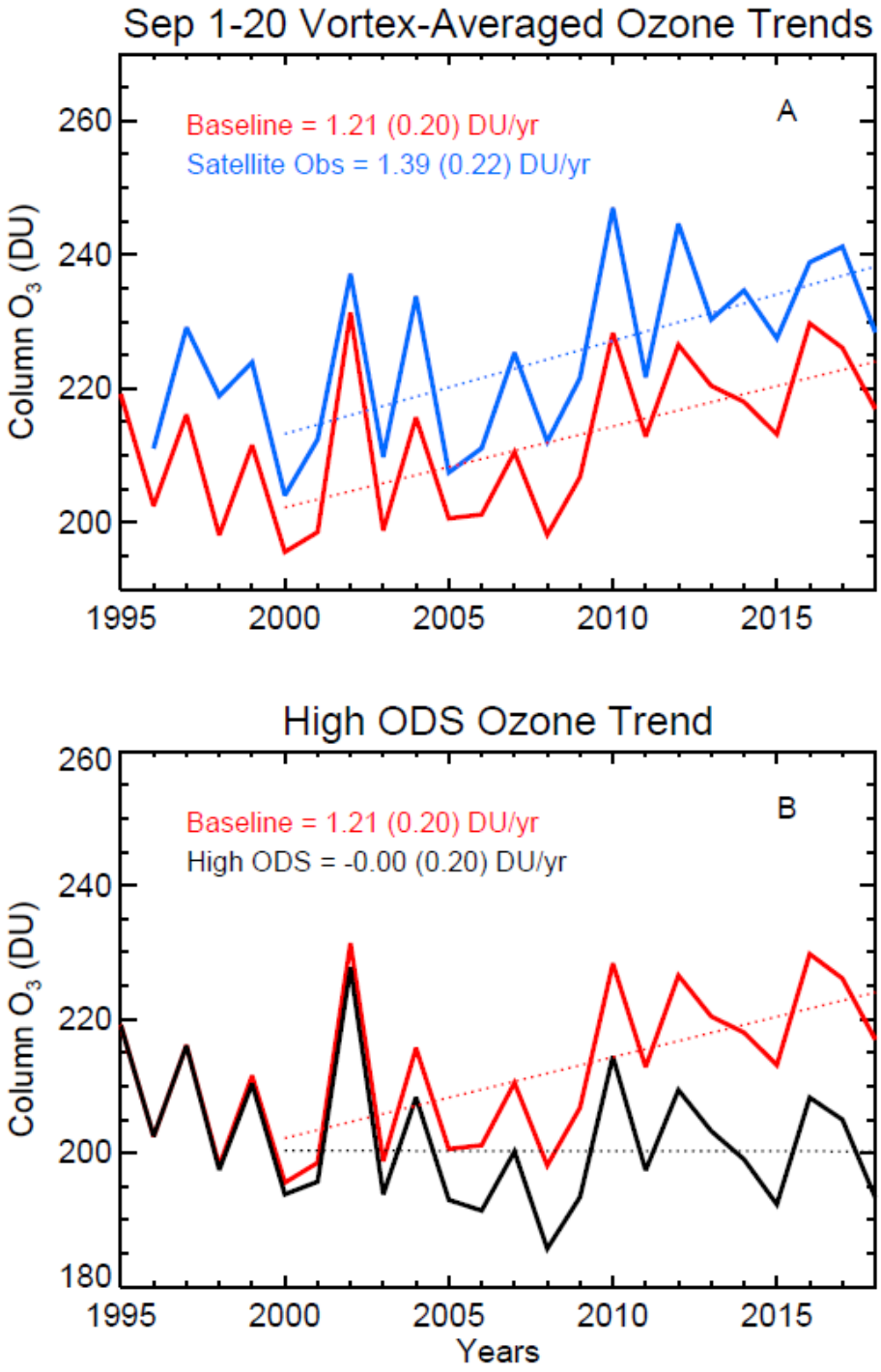


Figure 3. 1-20 Sep Antarctic vortex-averaged column O_3 and their trends with 95% confidence intervals in parentheses, 2000-2018, for a) TOMS/OMI/OMPS (blue) and Baseline (red), and b) the Baseline (red) and High ODS simulations (black).

© 2019 American Geophysical Union. All rights reserved.

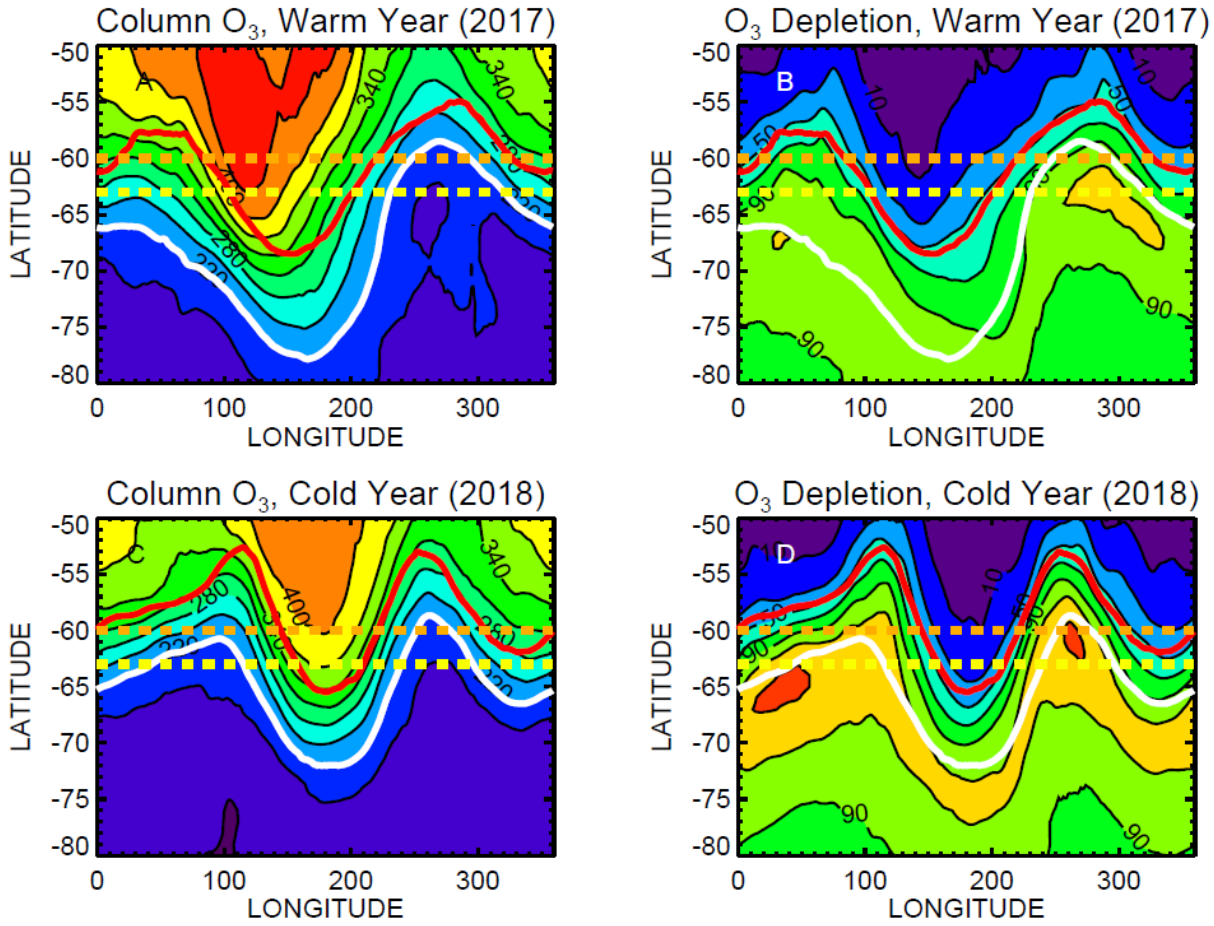


Figure 4. a) Baseline column O_3 and b) O_3 Loss (No Het-Baseline) in DU averaged over 8-14 Sep in 2017, a warmer than average year. c) and d) are the same as a) and b) except for Sep 8-14 2018, a colder than average year. The vortex edge (red), 220 DU contour (white), 63°S boundary (yellow dashed) and 60°S boundary (orange dashed) identify the boundaries of the metrics discussed.

Accepted

© 2019 American Geophysical Union. All rights reserved.

Accepted Article

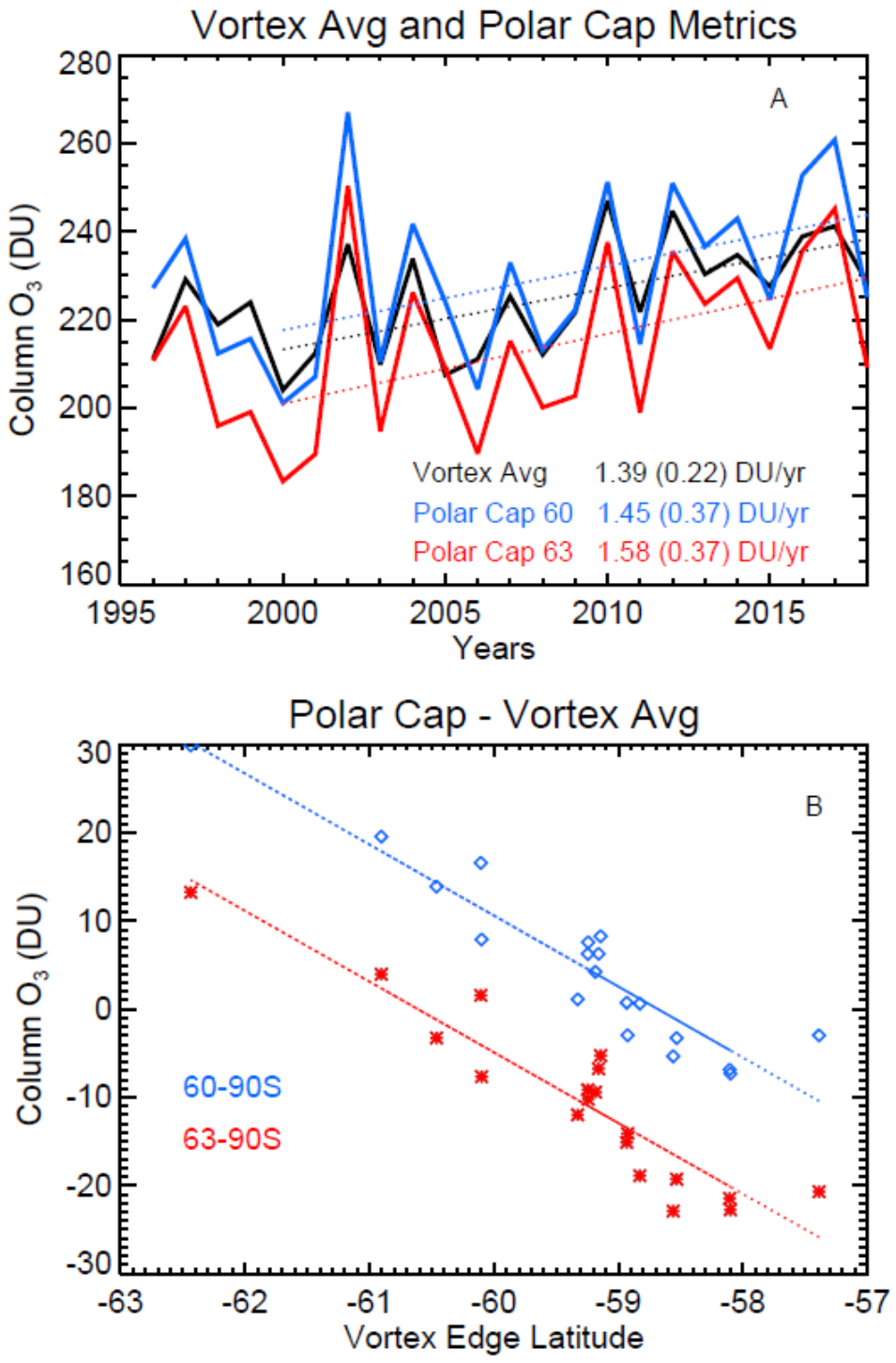


Figure 5. a) Observed polar cap and vortex-averaged O_3 metrics, and their trends with 95% CIs. b) Differences between the polar cap and vortex-averaged metrics, 2000-2018, from a).

© 2019 American Geophysical Union. All rights reserved.

Accepted Article

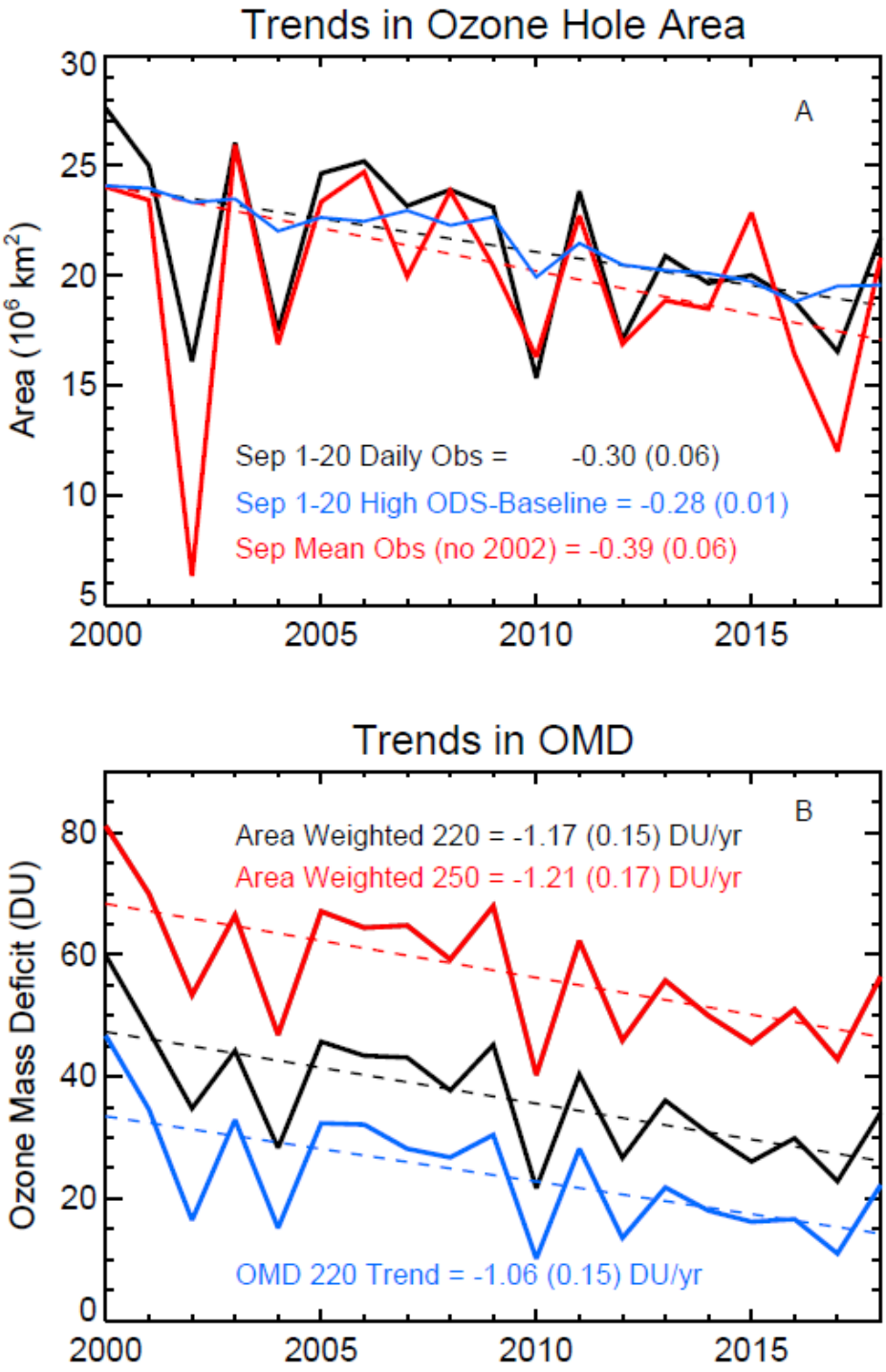


Figure 6. a) Time series, trends, and CIs for area with <220 DU O_3. The areas in black are calculated using daily areas that are averaged over 1-20 Sep, the areas in blue are the daily High ODS-Baseline area differences averaged over 1-20 Sep and offset by $24*10^6$ km^2 (blue), and the areas in red calculate the area <220 DU O_3 after averaging all September O_3 observations together. b) OMD metrics, trends and CIs using 220 DU (black) and 250 DU (red) boundaries, each weighted by the area they cover, and the unweighted 220 DU OMD for 60-90°S (blue).

© 2019 American Geophysical Union. All rights reserved.

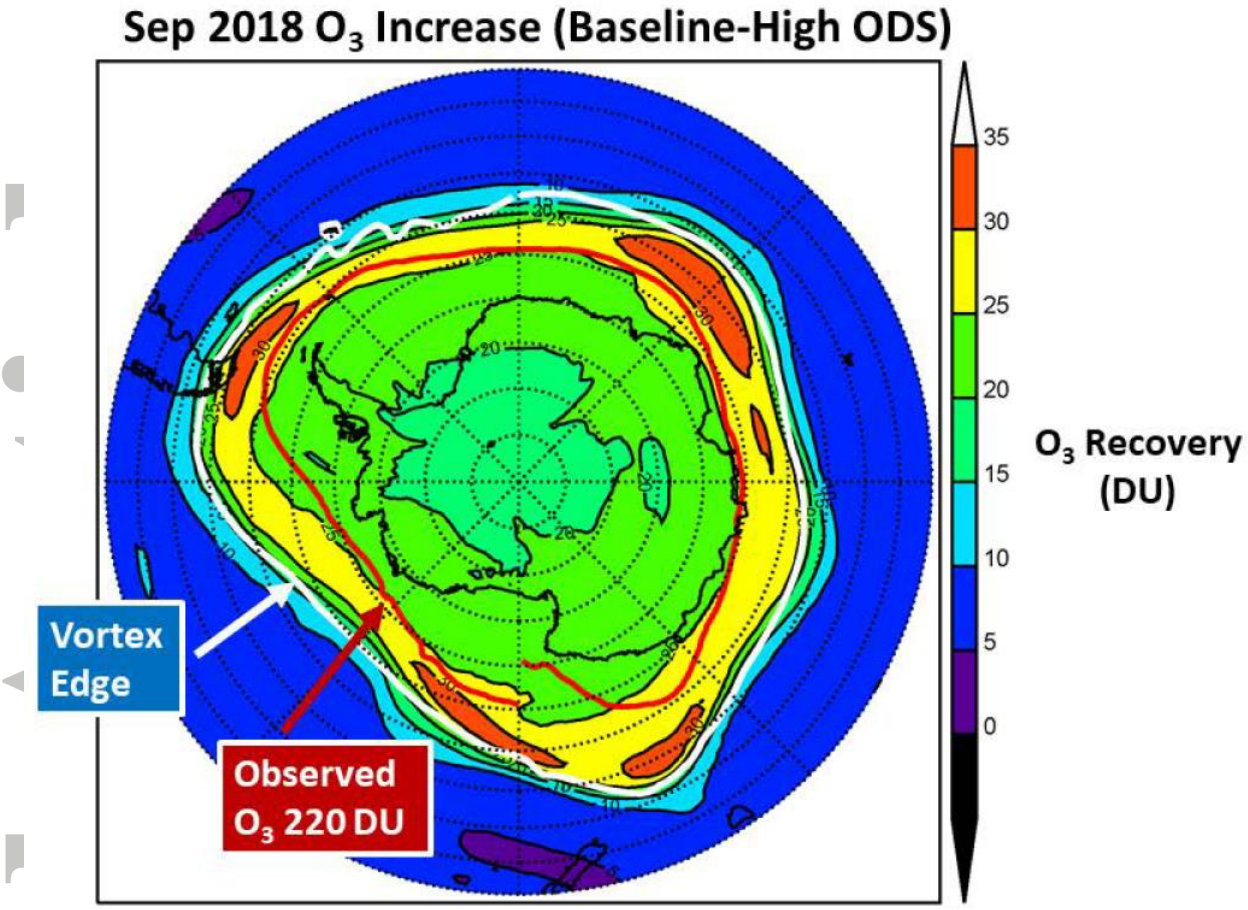


Figure 7. The reduction in O_3 loss (*i.e.*, recovery) due to declining ODSs from Baseline-High ODS column O_3 on Sep. 15, 2018. The regions with the greatest recovery, 25-35 DU, fall outside of the observed 220 DU contour (red) but inside the vortex edge (white). The 220 DU OMD metric excludes the region of largest recovery, which biases its trend low. The 220 DU observed O_3 contour (red) comes from 24 hours of satellite observations. There is a 23 hr lag between observations at 180°E (bottom of figure) that causes a discontinuity in the contour.

Accepted

© 2019 American Geophysical Union. All rights reserved.

Accepted Article

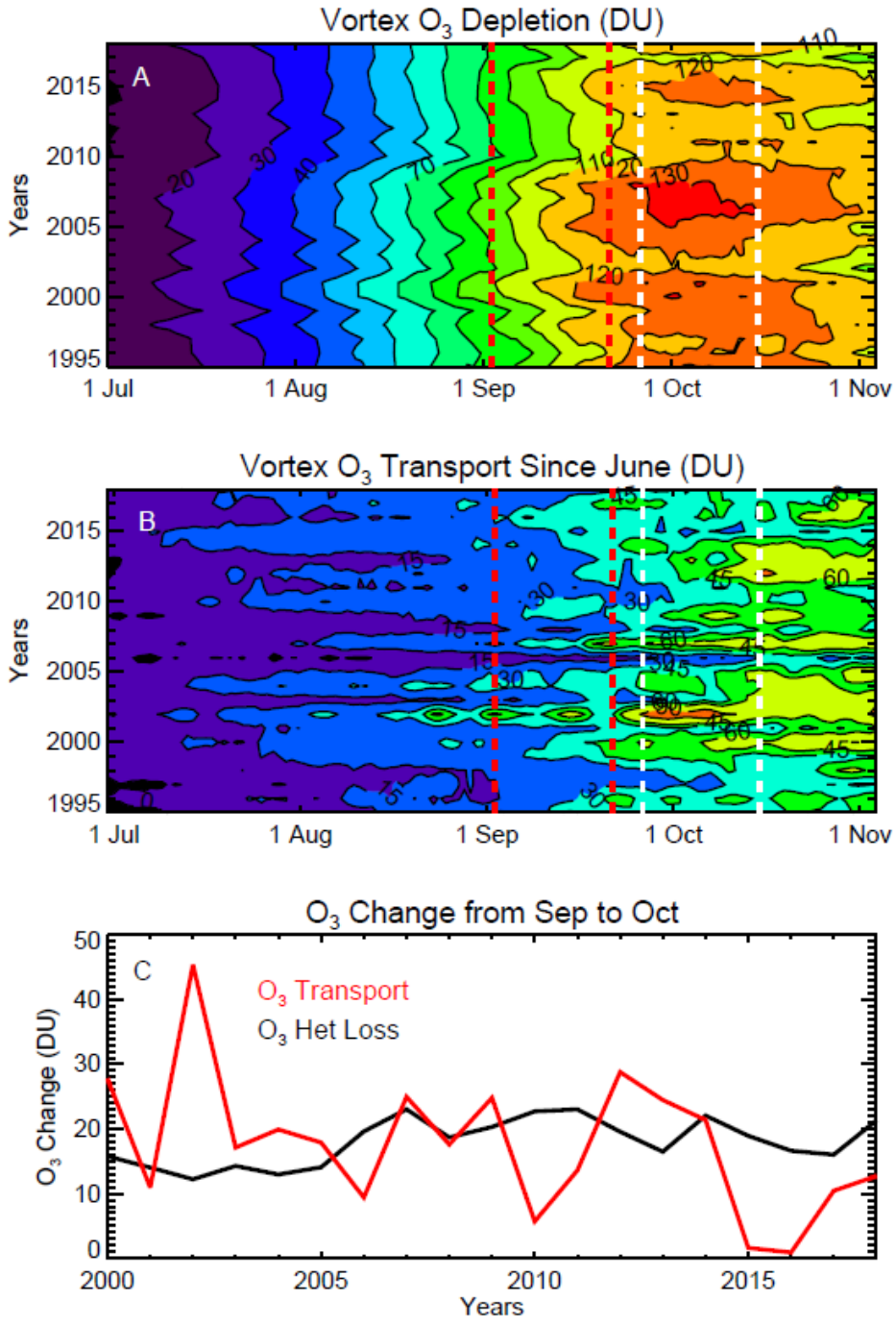


Figure 8. 1995-2018 vortex-averaged a) O_3 Depletion (DU) (No Het-Baseline) and b) O_3 transport since June (No Het) as a function of date. c) the vortex-averaged change in depletion (black) and transport (red) between 1-20 Sep and 25 Sep-14 Oct. These dates are shown by the red and white dashed lines in a) and b). The date ranges are the periods of maximum sensitivity of O_3 loss to ODSs (Sep, red) and maximum seasonal O_3 loss (Oct, white).

© 2019 American Geophysical Union. All rights reserved.

Table 1. Summary of Antarctic O_3 recovery metric differences with respect to the 1-20 Sep Vortex-Averaged Column O_3 metric

Issue	**Impact on Trend**	**Metric Affected**
Inconsistent sensitivity to ODS decline	Increased uncertainty, possible bias that depends on dynamical variability in the record	Fixed latitude polar cap O_3 averages
Lower sensitivity to ODS decline	Low bias, longer record required for trend significance	All O_3 metrics using October observations
Dynamical Variability	Increased uncertainty, longer record required for trend significance	All O_3 metrics using October observations
Regions with greatest O_3 recovery not included	Low bias	220 DU OMD, 250 DU OMD (has less low bias)
Information lost by averaging	Possible bias that depends on dynamical variability in the record	Area inside 220 DU O_3 contour in monthly averaged column O_3

Accepted Article

© 2019 American Geophysical Union. All rights reserved.