Fingerprinting the Recovery of Antarctic Ozone

Peidong Wang[a#], Susan Solomon[a#], Benjamin D. Santer[b,c], Douglas E. Kinnison[d], Qiang Fu[e], Kane A. Stone[a], Jun Zhang[d], Gloria L. Manney[f,g], Luis F. Millán[h]

[a]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

[b]Physical Oceanography Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

[c]Joint Institute for Regional Earth System Science and Engineering, University of California Los Angeles, Los Angeles, CA, USA

[d]Atmospheric Chemistry Observations and Modeling, National Science Foundation, National Center for Atmospheric Research, Boulder, CO, USA

[e]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA

[f]NorthWest Research Associates, Socorro, NM, USA

[g]New Mexico Institute of Mining and Technology, Socorro, NM, USA

[h]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

[#]Corresponding authors: Peidong Wang (pdwang@mit.edu) and Susan Solomon (solos@mit.edu)

The Antarctic ozone "hole" was discovered in 1985[1], and man-made ozone-depleting substances (ODS) are its primary cause[2]. Following reductions of ODSs under the Montreal Protocol[3], signs of ozone recovery have been reported, based largely on observations and broad yet compelling model-data comparisons[4]. While such approaches are highly

valuable, they don't provide rigorous statistical detection of the temporal and spatial structure of Antarctic ozone recovery in the presence of internal climate variability. Here, we apply pattern-based detection and attribution methods as employed in climate change studies[5–11] to separate anthropogenically forced ozone responses from internal variability, relying on trend pattern information as a function of month and height. The analysis uses satellite observations together with single-model and multi-model ensemble simulations to identify and quantify the month-height Antarctic ozone recovery "fingerprint"[12]. We demonstrate that the data and simulations show compelling agreement in the fingerprint pattern of the ozone response to decreasing ODSs since 2005. We also show that ODS forcing has enhanced ozone internal variability during the austral spring, influencing detection of forced responses and their time of emergence. Our results provide robust statistical and physical evidence that actions taken under the Montreal Protocol to reduce ODSs are indeed resulting in the beginning of Antarctic ozone recovery, defined as increases in ozone consistent with expected month-height patterns.

Despite the decline in ozone-depleting substances (ODS)[3], the continuing occurrence of large Antarctic ozone "holes" in recent years – partly due to the 2020 Australian wildfire[13–15] and 2022 Hunga volcanic eruption[16–19] – along with concerns about a significant decrease in October mid-stratospheric ozone[20], have stimulated discussion about the detectability of a robust signal in Antarctic ozone recovery. Formal detection and attribution (D&A) methods are needed to assess the effectiveness of the Montreal Protocol.

Initial-condition large ensembles (LEs) generated with fully coupled global climate models offer a unique opportunity to "fingerprint" the anthropogenic influence. They provide valuable information on both natural internal variability and the characteristic space and time signatures of the climate responses to different external forcings[5–12]. Such ensembles are rarely used in ozone studies, which generally apply multiple linear regression to isolate underlying forced trends by fitting ozone time series to known sources of variability (such as the solar cycle, El Niño-Southern Oscillation, Quasi-Biennial Oscillation, volcanic activities, etc.)[4,21]. In contrast, LEs do not require the assumption that the response is a linear combination of independent predictor variables[22]. Although some D&A methods have been applied to global ozone depletion and the time of emergence of total column recovery using multi-model ensembles[23–25], previous studies have not fully utilized the pattern-based D&A techniques. Furthermore, multiple members from a single-model that is well-evaluated and realistic can reliably quantify modeled forced signals and intrinsic variability, whereas one run from each of a multi-model ensemble samples larger cross-model differences in forcing, response, and variability[26], thus hampering the identification of the beginning of ozone recovery[23,24].

Observed and model-simulated ozone trends

We compare observed and simulated month-height trend patterns over 2005 to 2018 for Antarctic ozone, spatially averaged over 66°-82°S (Figure 1). Observations are from MLS[27] (Microwave Limb Sounder). Model results are from two different sources: 1) the multi-model ensemble mean computed using one run of each of the 19 models that participated in the CCMI-1[28] (phase 1 of the Chemistry–Climate Model Initiative); and 2) the single-model ensemble of 10 different realizations of CESM-WACCM[29,30] (Community Earth System Model-Whole Atmosphere

Community Climate Model; referred to as WACCM for short). The WACCM (and some models in the CCMI) ensemble comprises free-running coupled ocean-atmosphere simulations forced by time-evolving changes in greenhouse gases (GHG) and ODS, referred to as the "refC2" scenario (described in detail in the Methods section). We mainly focus on ozone trends from 2005 to 2018 (results covering longer periods are also considered below) to avoid recent exceptional events that are not represented in these models, including the 2019 sudden stratospheric warming[31] (SSW) and post-2020 wildfire and volcanic eruption mentioned earlier (note that the 2015 Calbuco eruption[32] is also not generally simulated in these models but is included in our trend analysis). Time series used for calculating trends are presented for illustration for four selected months and heights in Figure 1a,b,f,g; the expected post-2018 differences between observations and simulations are marked with dashed lines (in later figures as well).

Compared to the single real-world realization provided by MLS, the modeled mean trends in WACCM and CCMI have smaller amplitudes (Figure 1c,d,e). This is partly due to the fact that the model ensemble means are averages over many different realizations with varying phasing of internal variability (which is uncorrelated, except by chance) superimposed on the forced response. For comparison, Extended Data Figure 1 shows the ozone trends in 2005-2018 from individual WACCM realizations and CCMI model runs. Some individual members display month-height trend patterns that are more similar to MLS while others are less similar, reflecting differences in the phasing of internal variability. Nonetheless, nearly all realizations preserve some common features, reflecting the "fingerprint" of GHG and ODS forcings on ozone trends.

Increases in lower stratospheric ozone (at altitudes below the pressure level of ~30 hPa) during the austral spring are mainly due to the reduction in ODS concentrations[3], leading to less heterogeneous ozone loss[2]. The seasonal signature of the ozone hole in August-December is apparent in this region. Increases in upper stratospheric ozone (above ~10 hPa) are partly due to cooling caused by increasing GHGs (mainly CO_2)[33] and its effect on temperature-dependent ozone photochemistry[34]. They are also partly due to less reactive chlorine, through its effect on the ClO+O reaction that peaks near 40 km[35] (see Extended Data Figure 2). In the descending circulation of the polar winter, the increased ozone in the upper stratosphere propagates down to the mid-stratosphere. This combined month- and height-resolved pattern characterizes the "fingerprint" of GHG and ODS forcings on Antarctic ozone changes. In this study, we define ozone recovery as statistically significant increases in ozone that display these characteristic time-space fingerprints.

Noise of ozone variability and ODS forcing

Signal-to-noise analyses in D&A climate studies typically use natural internal variability ("noise") estimated from long pre-industrial control runs[5–11]. We rely on several noise estimates here from different WACCM scenarios (described in detail in the Methods section). Figure 2a,b show the month-height patterns of noise, defined as the standard deviations of ozone trends in the two separate 10-member WACCM historical and refC2 ensembles (after first removing the mean forced response). Although generated with the same physical climate model, the magnitude of noise trends in the historical and refC2 simulations exhibits notable differences. This is especially important in the austral spring in the lower stratosphere where the ozone "hole" typically occurs. The historical scenario has low GHGs and low ODSs, representing atmospheric

conditions prior to the onset of large ozone losses in the 1980s. The refC2 scenario has high GHGs and high ODSs comparable to present-day conditions. This indicates that the forcing differences in the two scenarios directly affect internal noise. It is important to account for forced changes in variability when examining ozone recovery – the estimated statistical significance of the observed trends can be affected by changes in noise amplitude.

There is a striking enhancement in ozone variability under the present-day high ODS conditions in austral spring in the lower stratosphere (marked with the white dashed boxes in Figure 2a,b); the standard deviations in refC2 and fGHG scenarios (low GHG but high ODS) are both increased by ~130% compared to the historical scenario. GHG forcing alone (the fODS scenario, with low ODS but high GHG) yields a narrow spread in ozone variability similar to the historical case, confirming that the variability enhancement is primarily driven by ODS forcing. Such enhancement occurs because chemical ozone loss due to ODS acts in the same direction as the natural variations in ozone arising from Brewer-Dobson circulation (BDC) variability. For example, a weak BDC associated with negative ozone and negative temperature anomalies favors more heterogeneous ozone loss in the presence of high ODS, extending the lower tail of the distribution of absolute ozone concentration[36] (Figure 2d) and broadening ozone variability.

For accurate analysis of the statistical significance of ozone changes, it is critical that the model-based noise is realistic. Figure 2c shows the distributions of monthly-mean ozone anomalies from MLS and WACCM (after first removing the mean forced response) relative to the present-day MLS climatology, in the region highlighted by the white dashed boxes in Figure 2a,b. The same data are displayed in Figure 2d in terms of absolute ozone mixing ratios. The distributions

of ozone internal variability in MLS and in the refC2 scenario are virtually identical, suggesting the refC2 noise provides a credible estimate of the real-world internal variability.

Signal-to-noise analysis of ozone changes

Here we use model-simulated internal variability noise to assess whether the forced ozone response has emerged. This assessment is performed for individual months and heights ("local" analysis) and for the "overall" month-height pattern.

Figure 3d displays the local signal to noise (S/N) ratio inferred from WACCM for a trend length of 14 years (2005 to 2018). It represents the local signal in WACCM ensemble mean (Figure 1d) divided by the local noise in the WACCM refC2 run (Figure 2b). In Figure 3c, the mean forced signal from WACCM is replaced by the MLS observed trend, which contains both the forced response and internal variability. A larger local S/N ratio indicates increased likelihood that the ozone trend is anthropogenically forced. Based on the WACCM S/N for the refC2 scenario, the beginning of ozone recovery (as a forced response to combined GHG and ODS forcing) can be detected with high confidence by 2018 in certain months and heights. In the upper stratosphere, recovery is significantly larger than internal variability in every month except during winter, when it propagates to the middle stratosphere due to polar descent. There is also a relative maximum in local S/N in September in the lower stratosphere in MLS and in the WACCM ensemble mean. The overall pattern of local S/N is similar in CCMI, but statistical significance is lower in several key regions (Extended Data Figure 4). This is expected given that the multi-model CCMI noise does not reflect intrinsic variability alone and is larger than in the WACCM single model.

To explore the impact of the post-2018 exceptional events, we performed a local S/N analysis over a longer period (2005-2023; see Extended Data Figure 5). The month-height local S/N pattern over 2005-2023 shows many features similar to those in Figures 3c,d, but also pronounced differences between WACCM and MLS, especially in the mid-stratosphere in October-December. LEs that account for the exceptional forcings[15,16,32], would be expected to provide better agreement with the observed ozone trends.

We also performed a S/N analysis using the overall month-height fingerprint pattern of Antarctic ozone trends since 2005 (see Figure 4). The element-wise covariance between the observed trend pattern (Figure 1c) and model ensemble-mean forced response (Figure 1d,e), at varying trend lengths, is divided by the spread of the covariance between individual noise patterns and the ensemble-mean responses (see the Methods section). Our results indicate that for the period 2005-2018, the observed time-space structure of ozone changes over Antarctica is consistent with time-evolving ODS and GHG forcing. The observed changes during this period are inconsistent with natural internal variability alone (at the 5% significance level for the observed MLS pattern projected on both WACCM and on CCMI month-height fingerprints). Although the exceptional years in and after 2020 lower the overall S/N, MLS trends projected on WACCM results (which neglect these events), nonetheless remain significant at the 10% significance level as late as the end of 2023.

We also project the observed pattern onto the GHG-only and ODS-only runs from WACCM. Despite the fact that ozone has a nonlinear response to the coupled GHG and ODS forcing in the

upper stratosphere[34], the linear addition in the covariance between MLS and the separate GHG and ODS fingerprints is close to the coupled GHG+ODS results for the entire month-height domain in our analysis (see Figure 4). This linear additivity enables us to estimate the relative contributions of GHG and ODS forcings in explaining the pattern similarity between observations and the coupled GHG+ODS fingerprint. The higher S/N values for the MLS trends projected onto the ODS fingerprint compared to the GHG fingerprint suggest that the observed pattern is dominated by the forced response to decreasing ODS concentrations. Compared to GHG, ODS-only forcing explains greater than 50% more pattern covariance with MLS after 2014 (also visually illustrated in Extended Data Figure 2). Indeed, the contribution of ODS-only forcing to the observed pattern in explaining the coupled GHG+ODS fingerprint differs significantly from internal variability noise at the 10% (and in some years 5%) significance level after 2014.

Antarctic springtime total ozone recovery

Signs of total column ozone recovery are often sought during the Antarctic spring[4], the season when the ozone hole maximizes in depth and extent. In September, the emergence of ozone recovery occurs around 2018 both in terms of ozone at a single illustrative level (82.5 hPa) in Figure 3e and in terms of the total column ozone (TCO) in Extended Data Figure 6a, where the observed TCO is from the OMI[37] (Ozone Monitoring Instrument). Even with exceptionally low ozone in and after 2020, the total ozone healing signal from the satellite data is still significantly outside of the internal variability noise (at the 10% level) in September.

A recent study raises the concern that October ozone in the mid-stratosphere has significantly decreased[20]. However, this time and location is subject to only a small healing signal and displays substantial internal variability (see Figure 3d,f), implying that trends with weak statistical significance here may well be spurious. Further, we note that because low ozone concentrations are mainly confined within the polar vortex[38], vortex variations (e.g., changes in size, shape, and position) can contribute to ozone internal variability when concentrations are spatially averaged over a fixed latitude range. This is especially important for months and heights when the vortex is often asymmetric (e.g., October in the mid-stratosphere). For example, maps of ozone anomalies and vortex locations in October at 12.1 hPa (where a strong ozone decrease was reported in a recent study[20]) in Extended Data Figure 7 illustrate how a shift of the vortex off the pole affects how it is sampled when using spatial averages calculated with fixed latitudinal boundaries (e.g., for a simple comparison with satellite coverage). The reported large negative ozone trend in October in the mid-stratosphere[20] is significantly reduced when considering the vortex averaged ozone (see Extended Data Figure 8).

The emergence of column ozone recovery in October and November due to combined ODS and GHG forcing (based on the WACCM ensemble mean signal) had been expected around 2021 under typical conditions (Extended Data Figure 6b,c). However, the unusually low ozone years in and after 2020 may have delayed detection in the observations. This underscores the importance of maintaining a long observation record to ensure high confidence in detecting and attributing future ozone changes at this time of year.

Summary and outlook

We performed a pattern-based fingerprint analysis for Antarctic ozone recovery, analogous to fingerprinting anthropogenic climate change[5–12]. Our S/N results for local and overall pattern similarity between MLS and single-model or multi-model ensembles provide high confidence that observed Antarctic ozone trends are primarily responses to ODS forcing rather than natural variability.

Further, we have shown that the amplitude of lower stratospheric ozone variability is markedly enhanced in a present-day simulation relative to the amplitude of ozone variability in a "pre-ozone depletion" simulation. It is crucial to consider this modulation of internal variability when evaluating the statistical significance of ozone trends. Additionally, this enhancement in ozone variability due to ODS forcings sheds light on a potential pathway for external forcing to modulate specific modes of natural internal variability, such as the Southern Annular Mode[39].

A significant ODS-driven signal of local ozone recovery in October and November has yet to emerge in the observations, likely due to the exceptionally low ozone years in and after 2020. These low ozone years are at least partly due to known volcanic and wildfire forcings not included in the available simulations. While October ozone exhibits a decreasing trend in the middle stratosphere[20], this negative trend is notably mitigated by adopting a different coordinate system that accounts for vortex variations.

Some caveats of the current analysis should be noted. Only one model with 10 members is examined in detail in this work. To improve confidence in the detection and attribution of forced responses versus natural variability in future ozone recovery assessments, it would be beneficial

to use larger initial condition ensembles from multiple single models rather than relying on single realizations from many different models[11,23,24]. The forced response in this study considers GHG and ODS only, and does not include known forcings from important volcanoes and major wildfires after 2012; future ensemble runs including these forcings would likely improve consistency of the simulated and observed ozone changes. The projected long-lasting stratospheric water vapor from the Hunga eruption[18,40] or future volcanic or wildfire forcing[41,42] could reduce the future ozone recovery signal. Indeed, even the large S/N ratio we now see in the upper stratosphere could be temporarily obscured by uncertainties in future GHG emissions[43] and solar proton events[44].

Our work shows how fingerprinting and pattern similarity establish quantitative confidence that Antarctic ozone recovery has begun. It also indicates why it is crucial to maintain global height-resolved observations over extended periods to identify signal patterns that emerge from background noise, raising concerns about the impending satellite data gap in stratospheric measurements[45]. A long observational record can ensure that estimated S/N ratios are less sensitive to short-term episodic perturbations, thereby providing high confidence in detecting and attributing trends[5].

References from main text

1. Farman, J. C., Gardiner, B. G. & Shanklin, J. D. Large losses of total ozone in Antarctica reveal seasonal ClOx/NOx interaction. *Nature* **315**, 207–210 (1985).
2. Solomon, S., Garcia, R. R., Rowland, F. S. & Wuebbles, D. J. On the depletion of Antarctic ozone. *Nature* **321**, 755–758 (1986).

3. Laube, J. C. & Tegtmeier, S. *Scientific Assessment of Ozone Depletion: 2022. Chapter 1: Update on Ozone-Depleting Substances (ODSs) and Other Gases of Interest to the Montreal Protocol*. 509 (2022).
4. Chipperfield, M. P. & Santee, M. L. *Scientific Assessment of Ozone Depletion: 2022. Chapter 4: Polar Stratospheric Ozone: Past, Present & Future*. 509 (2022).
5. Santer, B. D. *et al.* Exceptional stratospheric contribution to human fingerprints on atmospheric temperature. *Proceedings of the National Academy of Sciences* **120**, e2300758120 (2023).
6. Terray, L. *et al.* Near-Surface Salinity as Nature's Rain Gauge to Detect Human Influence on the Tropical Water Cycle. *Journal of Climate* **25**, 958–977 (2012).
7. Stott, P. A., Sutton, R. T. & Smith, D. M. Detection and attribution of Atlantic salinity changes. *Geophysical Research Letters* **35**, (2008).
8. Gillett, N. P., Fyfe, J. C. & Parker, D. E. Attribution of observed sea level pressure trends to greenhouse gas, aerosol, and ozone changes. *Geophysical Research Letters* **40**, 2302–2306 (2013).
9. Christidis, N. & Stott, P. A. Changes in the geopotential height at 500 hPa under the influence of external climatic forcings. *Geophysical Research Letters* **42**, 10,798-10,806 (2015).
10. Shi, J.-R., Santer, B. D., Kwon, Y.-O. & Wijffels, S. E. The emerging human influence on the seasonal cycle of sea surface temperature. *Nature Climate Change* **14**, 364–372 (2024).
11. Santer, B. D. *et al.* Robust Anthropogenic Signal Identified in the Seasonal Cycle of Tropospheric Temperature. *Journal of Climate* **35**, 6075–6100 (2022).

12. Hasselmann, K. Optimal Fingerprints for the Detection of Time-dependent Climate Change. *Journal of Climate* **6**, 1957–1971 (1993).
13. Santee, M. L. *et al.* Prolonged and Pervasive Perturbations in the Composition of the Southern Hemisphere Midlatitude Lower Stratosphere From the Australian New Year's Fires. *Geophysical Research Letters* **49**, e2021GL096270 (2022).
14. Bernath, P., Boone, C. & Crouse, J. Wildfire smoke destroys stratospheric ozone. *Science* **375**, 1292–1295 (2022).
15. Solomon, S. *et al.* Chlorine activation and enhanced ozone depletion induced by wildfire aerosol. *Nature* **615**, 259–264 (2023).
16. Wang, X. *et al.* Stratospheric Climate Anomalies and Ozone Loss Caused by the Hunga Tonga-Hunga Ha'apai Volcanic Eruption. *Journal of Geophysical Research: Atmospheres* **128**, e2023JD039480 (2023).
17. Zhang, J. *et al.* Chemistry Contribution to Stratospheric Ozone Depletion After the Unprecedented Water-Rich Hunga Tonga Eruption. *Geophysical Research Letters* **51**, e2023GL105762 (2024).
18. Wohltmann, I., Santee, M. L., Manney, G. L. & Millán, L. F. The Chemical Effect of Increased Water Vapor From the Hunga Tonga-Hunga Ha'apai Eruption on the Antarctic Ozone Hole. *Geophysical Research Letters* **51**, e2023GL106980 (2024).
19. Manney, G. L. *et al.* Siege in the Southern Stratosphere: Hunga Tonga-Hunga Ha'apai Water Vapor Excluded From the 2022 Antarctic Polar Vortex. *Geophysical Research Letters* **50**, e2023GL103855 (2023).
20. Kessenich, H. E., Seppälä, A. & Rodger, C. J. Potential drivers of the recent large Antarctic ozone holes. *Nature Communications* **14**, 7259 (2023).

21. Hassler, B. & Young, P. J. *Scientific Assessment of Ozone Depletion: 2022. Chapter 3: Update on Global Ozone: Past, Present, and Future*. 509 (2022).
22. Santer, B. D. *et al.* Accounting for the effects of volcanoes and ENSO in comparisons of modeled and observed temperature trends. *Journal of Geophysical Research: Atmospheres* **106**, 28033–28059 (2001).
23. Dhomse, S. S. *et al.* Estimates of ozone return dates from Chemistry-Climate Model Initiative simulations. *Atmospheric Chemistry and Physics* **18**, 8409–8438 (2018).
24. Zeng, G. *et al.* Attribution of Stratospheric and Tropospheric Ozone Changes Between 1850 and 2014 in CMIP6 Models. *Journal of Geophysical Research: Atmospheres* **127**, e2022JD036452 (2022).
25. Robertson, F. *et al.* Signal-To-Noise Calculations of Emergence and De-Emergence of Stratospheric Ozone Depletion. *Geophysical Research Letters* **50**, e2023GL104246 (2023).
26. Deser, C. *et al.* Insights from Earth system model initial-condition large ensembles and future prospects. *Nature Climate Change* **10**, 277–286 (2020).
27. Waters, J. W. *et al.* The Earth observing system microwave limb sounder (EOS MLS) on the aura Satellite. *IEEE Transactions on Geoscience and Remote Sensing* **44**, 1075–1092 (2006).
28. Morgenstern, O. *et al.* Review of the global models used within phase 1 of the Chemistry–Climate Model Initiative (CCMI). *Geoscientific Model Development* **10**, 639–671 (2017).
29. Marsh, D. R. *et al.* Climate Change from 1850 to 2005 Simulated in CESM1(WACCM). *Journal of Climate* **26**, 7372–7391 (2013).
30. Garcia, R. R., Smith, A. K., Kinnison, D. E., Cámara, Á. de la & Murphy, D. J. Modification of the Gravity Wave Parameterization in the Whole Atmosphere Community Climate Model: Motivation and Results. *Journal of the Atmospheric Sciences* **74**, 275–291 (2017).

31. Wargan, K., Weir, B., Manney, G. L., Cohn, S. E. & Livesey, N. J. The Anomalous 2019 Antarctic Ozone Hole in the GEOS Constituent Data Assimilation System With MLS Observations. *Journal of Geophysical Research: Atmospheres* **125**, e2020JD033335 (2020).
32. Solomon, S. *et al.* Emergence of healing in the Antarctic ozone layer. *Science* **353**, 269–274 (2016).
33. Manabe, S. & Wetherald, R. T. Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity. *Journal of Atmospheric Sciences* **24**, 241–259 (1967).
34. Haigh, J. D. & Pyle, J. A. Ozone perturbation experiments in a two-dimensional circulation model. *Quarterly Journal of the Royal Meteorological Society* **108**, 551–574 (1982).
35. Molina, M. J. & Rowland, F. S. Stratospheric sink for chlorofluoromethanes: chlorine atom-catalysed destruction of ozone. *Nature* **249**, 810–812 (1974).
36. Solomon, S., Portmann, R. W., Sasaki, T., Hofmann, D. J. & Thompson, D. W. J. Four decades of ozonesonde measurements over Antarctica. *Journal of Geophysical Research: Atmospheres* **110**, (2005).
37. Levelt, P. F. *et al.* The ozone monitoring instrument. *IEEE Transactions on Geoscience and Remote Sensing* **44**, 1093–1101 (2006).
38. Schoeberl, M. R. & Hartmann, D. L. The Dynamics of the Stratospheric Polar Vortex and Its Relation to Springtime Ozone Depletions. *Science* **251**, 46–52 (1991).
39. Thompson, D. W. J. *et al.* Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change. *Nature Geoscience* **4**, 741–749 (2011).
40. Zhou, X. *et al.* Antarctic Vortex Dehydration in 2023 as a Substantial Removal Pathway for Hunga Tonga-Hunga Ha'apai Water Vapor. *Geophysical Research Letters* **51**, e2023GL107630 (2024).

41. Eric Klobas, J., Wilmouth, D. M., Weisenstein, D. K., Anderson, J. G. & Salawitch, R. J. Ozone depletion following future volcanic eruptions. *Geophysical Research Letters* **44**, 7490–7499 (2017).

42. Chim, M. M. *et al.* Climate Projections Very Likely Underestimate Future Volcanic Forcing and Its Climatic Effects. *Geophysical Research Letters* **50**, e2023GL103743 (2023).

43. Revell, L. E., Bodeker, G. E., Huck, P. E., Williamson, B. E. & Rozanov, E. The sensitivity of stratospheric ozone changes through the 21st century to N_2O and CH_4. *Atmospheric Chemistry and Physics* **12**, 11309–11317 (2012).

44. Stone, K. A., Solomon, S. & Kinnison, D. E. On the Identification of Ozone Recovery. *Geophysical Research Letters* **45**, 5158–5165 (2018).

45. Chipperfield, M. P. & Bekki, S. Opinion: Stratospheric ozone – depletion, recovery and new challenges. *Atmospheric Chemistry and Physics* **24**, 2783–2802 (2024).

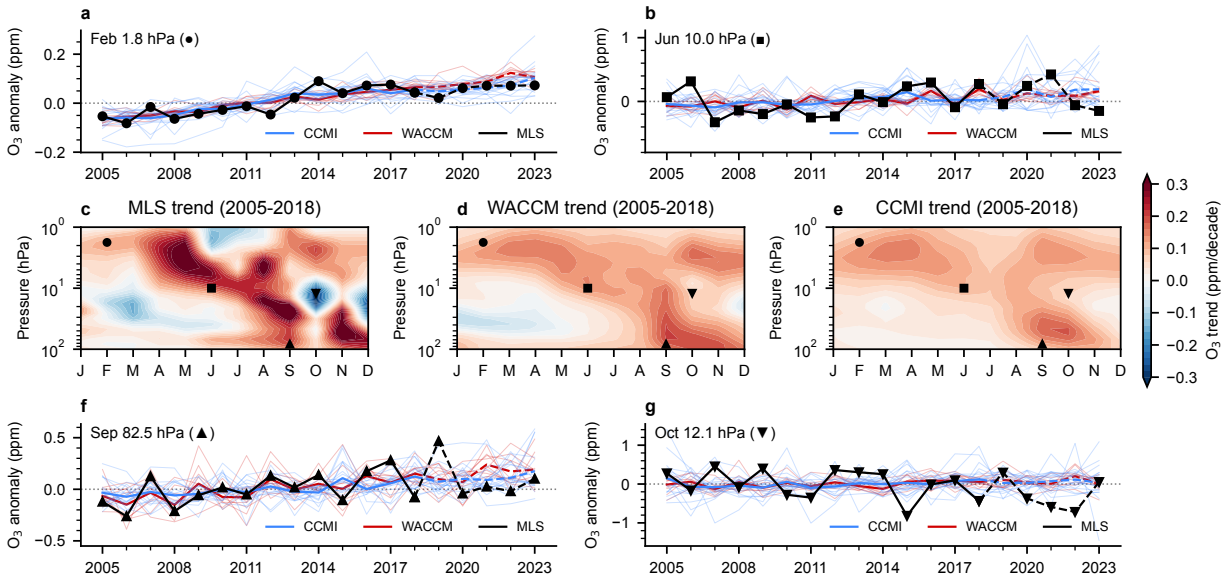
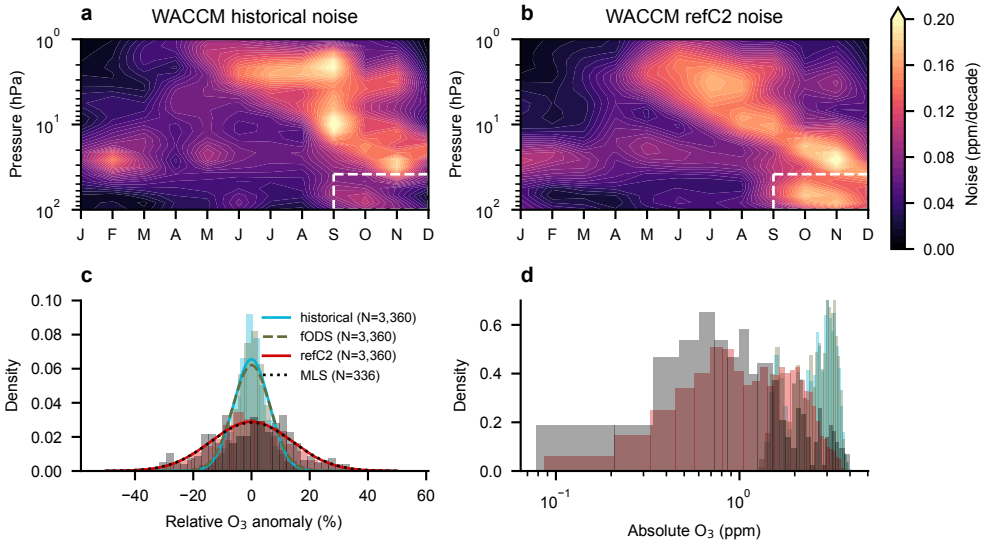
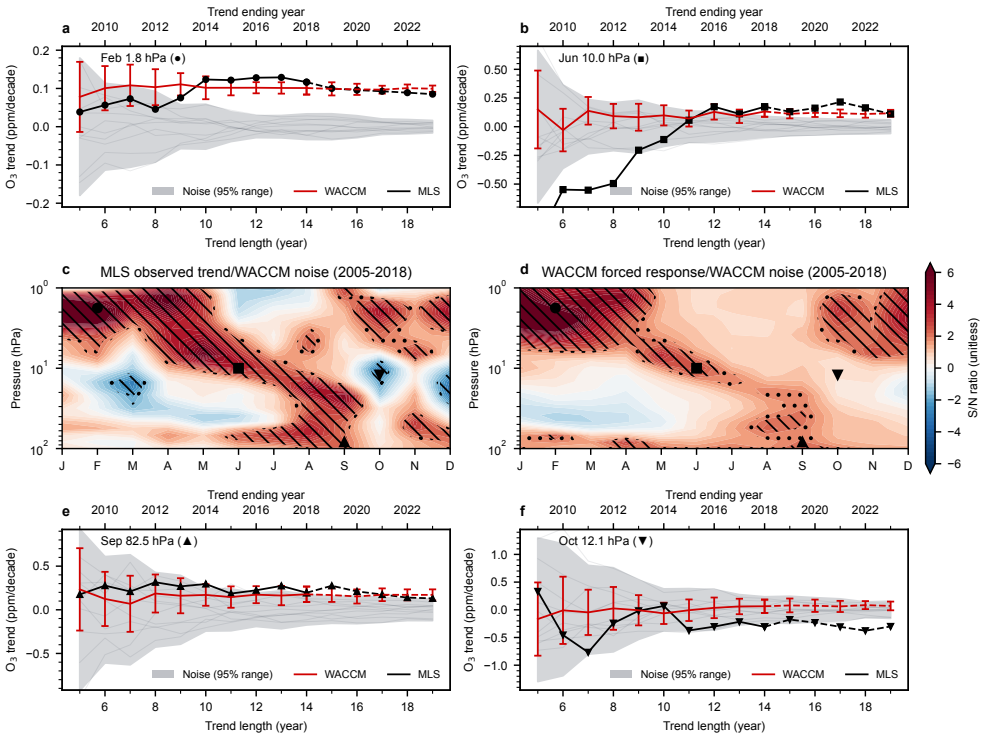


Figure 1. Month-height patterns of ozone trends in observations and simulations. The middle row shows the least-squares linear trends in ozone over 2005-2018 as a function of month and pressure in the MLS observations (**c**), the WACCM 10-member ensemble mean (**d**), and the CCMI 19-model ensemble mean (**e**). Results are for spatial averages over 66°-82°S. We also show the time series of ozone anomalies (relative to the 2005-2018 mean) at four illustrative locations (indicated by different markers on the contour figures) in February at 1.8 hPa (**a**), June at 10 hPa (**b**), September at 82.5 hPa (**f**), and October at 12.1 hPa (**g**). Thick blue and red lines are the model ensemble means, while the thin lines are the ozone time series in individual model realizations. The black solid line is from MLS. Time series after 2018 are shown as dashed lines because of likely impacts of exceptional perturbations from the 2019 SSW, 2020 Australian wildfire, and 2022 Hunga eruption.



391

392 **Figure 2. Ozone variability modulated by external forcing.** Panels **a**, **b** show the month-
393 height patterns of the standard deviation of ozone trends (averaged over 66°-82°S) in the
394 WACCM historical and refC2 scenarios, after first removing the ensemble mean. The standard
395 deviation is a measure of the trend uncertainty arising from internal variability among the 10
396 WACCM ensemble members. Both scenarios have the same trend length of 14 years: the
397 historical scenario analysis period is from 1955 to 1968 and the refC2 analysis period is from
398 2005-2018. Panel **c** shows (for the white dashed box in the top panels), the residual internal
399 variability in WACCM and MLS ozone after first removing the mean forced response of
400 monthly-mean ozone. Results are expressed as percent changes relative to the MLS annual-mean
401 climatology. The thick lines are the Gaussian fits to the distributions. Panel **d** displays the same
402 data shown in panel **c**, but in terms of the absolute monthly ozone mixing ratios. The number of
403 data points in each distribution is indicated in the legend. A similar figure for CCMI models is
404 shown in Extended Data Figure 3; while the noise pattern is qualitatively similar to that of
405 WACCM, the CCMI multi-model ensemble also reflects different model responses to forcing
406 which can inflate noise compared to a single-model ensemble.



407

408 **Figure 3. The local signal-to-noise pattern of ozone changes.** The middle panels show the
409 month-height patterns of the observed ozone trends in MLS (**c**) and the WACCM 10-member
410 ensemble mean refC2 ozone trends (**d**) divided by the internal variability noise in refC2 trends
411 (which are given in Figure 2b). Results are for trends over 2005-2018; Extended Data Figure 5
412 similarly displays local S/N for trends calculated over 2005-2023. Hatched regions indicate that
413 the local trend is significantly outside the noise at the 5% (backslashes) and 10% (dots) levels.
414 The top and bottom rows (panels **a**, **b**, **e**, **f**) provide the local signal and noise as a function of
415 trend length at the same four illustrative months/heights shown in Figure 1. The thin gray lines
416 are the trends due to internal variability in each model realization, and the gray shading indicates
417 the 95% range of the realization spread. Thick lines are the trends from MLS (black) and the
418 WACCM ensemble mean (red). The error bars on the red line indicate ±1 standard deviation in
419 the 10-member ensemble.

420

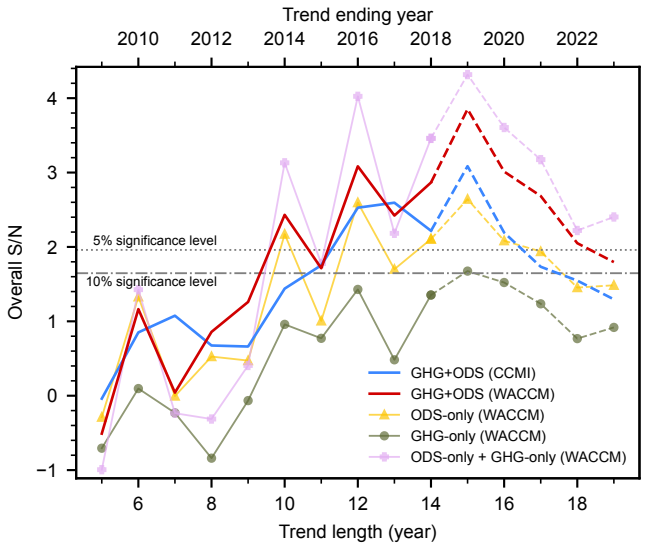


Figure 4. Signal-to-noise characteristics for the overall month-height fingerprint pattern. The numerator of the S/N ratio is the element-wise covariance in the overall month-height ozone trend patterns between MLS and the ensemble means of the coupled GHG+ODS runs from WACCM (red) or CCMI (blue), calculated at trend lengths ranging from 5 to 19 years. The denominator of the ratio is the standard deviation of the element-wise covariance computed between individual noise patterns and the ensemble mean responses. The observed MLS ozone trend pattern is also projected onto the ODS-only and GHG-only fingerprints from WACCM, shown in yellow and dark green lines, respectively (a linear addition of these two separate forcing results is shown in the mauve line). An overall S/N value above the gray dashed or dash-dotted lines indicates that the similarity between the observed month-height pattern and the model-predicted forced response pattern (to GHG, ODS, or combined GHG+ODS changes) is significant at the 5% or 10% level and is therefore unlikely to be explained by internal variability alone.

436 **Methods**

437 **Satellite data**

438 MLS provides daily observations since August 2004, and the data has been extensively
439 validated[46,47]. Here we use MLS version 5 monthly-mean level 3 ozone mixing ratios on pressure
440 coordinates from 100 hPa to 1 hPa. The level 3 product covers latitudes from 82°S to 82°N,
441 using a 4° latitude bin. Ozone is averaged over 66°-82°S in this paper, weighted by cosine
442 latitude to account for the reduction in area further poleward. We use the monthly-mean values
443 starting from 2005 (excluding the latter half of 2004) so that every month in the trend analysis
444 has the same number of time samples.

445

446 Satellite TCO observations are from the OMI[37] version 3 daily level 3 product, which is onboard
447 the same satellite as MLS. TCO from OMI is also averaged by month starting from 2005 and in
448 the latitude range from 66°-82°S (weighted by cosine latitude).

449

450 **Model and scenario descriptions**

451 A total of 19 different models participated in the CCMI-1, with a total of 33 realizations for the
452 refC2 scenario[23,28]. This scenario characterizes ODS emissions following WMO (2011)[48], and
453 other GHG emissions following RCP6.0[49] from 1960 to 2100. To prevent biasing towards
454 models with more ensemble members, we only use the first realization from each model.

455

456 We also used the fully coupled CESM1-WACCM4[29,30] in this analysis, which incorporates
457 coupled ocean-atmosphere processes with interactive chemistry. Our primary focus is on a 10-
458 member WACCM initial condition ensemble generated with the same refC2 scenario employed

459 by the CCMI-1 models. In addition to refC2, we consider three other WACCM initial condition
460 ensembles, referred to as fODS, fGHG, and historical. fODS (also referred to as the GHG-only
461 runs) fixes ODS forcing at the 1960 level, while GHG concentrations evolve as in the refC2 runs.
462 Alternately, fGHG (or ODS-only runs) fixes GHG concentrations in 1960, while ODS levels
463 evolve as in refC2 runs. The historical scenario involves temporal changes in both GHG and
464 ODS from 1955 to 1979[50]. The CCMI and WACCM simulations are vertically interpolated to
465 MLS pressure levels (linear interpolation in log pressure), and are also averaged over 66°-82°S
466 and cosine-weighted for consistency with MLS.
467
468 Although the WACCM runs analyzed here are less than 30 years in length (from 1995 to 2024
469 for refC2, fODS, and fGHG, and from 1955 to 1979 for the historical scenario), an advantage of
470 the set of simulations is that each scenario has 10 realizations that are slightly perturbed in their
471 initial conditions[51]. This facilitates reliable estimation of both the underlying forced response (the
472 ensemble-mean) and internal variability. In contrast, multi-model ensembles convolve internal
473 variability estimates with inter-model differences or errors in forced responses, and/or with
474 model differences in the amplitude and patterns of internal variability[26,51,52]. For example, not all
475 of the 19 CCMI models are fully coupled to an interactive ocean[28], likely introducing large cross-
476 model differences in forced responses and natural variability.
477
478 As shown here, the WACCM historical and refC2 initial condition ensembles can also be used to
479 explore whether external forcing modulates internal variability – a key issue in signal detection.
480 Relative to multiple linear regression approaches for estimating the anthropogenic component of
481 ozone trends, using the WACCM fully-coupled chemistry climate model has the advantage that

no prior assumptions are required regarding the relationships between different predictor variables (such as the El Niño-Southern Oscillation, Southern Annual Mode, Quasi-Biennial Oscillation, solar cycle, etc.). As in the real world, nonlinear interactions between ozone and climate internal variability can be present in the model simulations.

Although large initial condition ensembles from multiple single-models are preferred for analyzing the interactions between atmospheric chemistry and natural internal variability, the high computational cost may be a barrier to generating such ensembles, at least for some models[26].

Signal and noise definition and uncertainty estimation

The "local" signal and noise denotes an analysis at individual months and heights. We define the local signal as the linear trend in ozone (starting in 2005, ending years can vary from 2009 to 2023) at each month and pressure level, derived from a linear fit to ensemble-mean ozone data from forced model simulations. To calculate the "local" noise, we first subtract the ensemble-mean ozone time series from each individual model realization; the resulting residuals then characterize the internal variability[26]. Noise is defined as the standard deviation of the ozone trends (with the same trend length as the signal) in these residuals. The noise represents the spread in ozone trends that is primarily due to internal variability. Both the local signal and the local noise have units of ppm/decade (parts per million by volume per decade). The statistical significance of ozone trends is determined by the local signal-to-noise ratio, S/N. For two-tailed tests, a 5% (10%) significance level is associated with local S/N values larger than 1.96 (1.645).

505 In addition to the local S/N analysis at individual months and heights, we also applied a
506 conventional "fingerprint" method[5] to the overall simulated and observed month-height patterns
507 of ozone changes. The key point here is that the entire month-height pattern is employed to
508 distinguish a forced response from internal variability. The overall signal is the element-wise
509 uncentered covariance between the month-height ozone trend patterns in MLS (e.g., Figure 1c)
510 and in the WACCM ensemble-mean (e.g., Figure 1d), or between the trend pattern in MLS and
511 the CCMI multi-model mean (e.g., Figure 1e), at varying trend lengths. This is essentially
512 equivalent to projecting the observed month- and height-resolved trend pattern onto the forced
513 response[5]. The overall noise is the standard deviation in the similarly computed uncentered
514 covariance but between internal variability in individual realizations (obtained from the different
515 panels in Extended Data Figure 1 after removing the ensemble mean) and the mean forced
516 response. Fingerprinting is performed over the same space-time ranges used in the local S/N
517 analysis: i.e., using spatially averaged ozone changes between 66°-82°S at altitudes from 100
518 hPa to 1 hPa and in the 12 months from January through December. The increase in the overall
519 S/N in Figure 4 with increasing trend length suggests that the observed overall month-height
520 ozone change pattern is unlikely to be explained by internal variability alone.
521
522 **Vortex coverage and vortex averaged ozone calculation**
523 Vortex coverage (shown as the dots in Extended Data Figure 7) and vortex averaged ozone (red
524 lines in Extended Data Figure 8) are obtained from MLS daily level 2 and the derived
525 meteorological products[53–55] (DMPs). The DMPs are calculated from the NASA GMAO (Global
526 Modeling and Assimilation Office) using meteorology from MERRA-2 (Modern-Era
527 Retrospective analysis for Research and Applications, Version 2) and are interpolated to the same

time and location as MLS level 2 products. Vortex edge in MLS is determined by a height-dependent sPV (scaled potential vorticity) threshold[56] on potential temperature (theta) surfaces. For consistency with our other results shown in pressure levels, we vertically interpolated this sPV threshold from theta level to pressure level (linear interpolation from log theta to log pressure) and applied it to the daily level 2 ozone from MLS and daily level 2 sPV from DMP. The monthly mean polar vortex coverage at each grid point on the MLS level 3 grid (with 4°×5° horizontal resolution) is averaged using daily vortex coverage, defined as the fraction of sPV in each grid box that meets the vortex threshold[56] against the total number of MLS overpasses in that grid box every day. Similarly, for vortex averaged ozone, we only average MLS level 2 observed ozone (weighted by cosine latitude) when its associated sPV (from the DMP) meets the vortex threshold[56] on a daily-basis in each grid box for every month.

Data availability

MLS and OMI satellite data are publicly available at https://disc.gsfc.nasa.gov. CCMI model outputs are available at https://archive.ceda.ac.uk, and the CESM model outputs are available at https://www.earthsystemgrid.org. All the pre-processed model data (e.g., monthly mean ozone averaged over 66°-82°S from CCMI and WACCM and interpolated onto MLS vertical coordinates) are available at Zenodo (https://doi.org/10.5281/zenodo.14497873)[57].

Code availability

The code used to generate all the figures in this analysis is available at Zenodo (https://doi.org/10.5281/zenodo.14497873)[57].

Additional references for the Methods section

46. Hubert, D. *et al.* Ground-based assessment of the bias and long-term stability of 14 limb and occultation ozone profile data records. *Atmospheric Measurement Techniques* **9**, 2497–2534 (2016).

47. Froidevaux, L. *et al.* Validation of Aura Microwave Limb Sounder stratospheric ozone measurements. *Journal of Geophysical Research: Atmospheres* **113**, (2008).

48. World Meteorological Organization (WMO). *Scientific Assessment of Ozone Depletion: 2010*. 516 (2011).

49. Meinshausen, M. *et al.* The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change* **109**, 213 (2011).

50. Stone, K. A., Solomon, S., Thompson, D. W. J., Kinnison, D. E. & Fyfe, J. C. On the Southern Hemisphere Stratospheric Response to ENSO and Its Impacts on Tropospheric Circulation. *Journal of Climate* **35**, 1963–1981 (2022).

51. Kay, J. E. *et al.* The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bulletin of the American Meteorological Society* **96**, 1333–1349 (2015).

52. Solomon, A. *et al.* Distinguishing the Roles of Natural and Anthropogenically Forced Decadal Climate Variability: Implications for Prediction. *Bulletin of the American Meteorological Society* **92**, 141–156 (2011).

53. Manney, G. L. *et al.* Solar occultation satellite data and derived meteorological products: Sampling issues and comparisons with Aura Microwave Limb Sounder. *Journal of Geophysical Research: Atmospheres* **112**, (2007).

54. Millán, L. F. *et al.* Multi-parameter dynamical diagnostics for upper tropospheric and lower stratospheric studies. *Atmospheric Measurement Techniques* **16**, 2957–2988 (2023).
55. Manney, G. L. *et al.* Jet characterization in the upper troposphere/lower stratosphere (UTLS): applications to climatology and transport studies. *Atmospheric Chemistry and Physics* **11**, 6115–6137 (2011).
56. Lawrence, Z. D., Manney, G. L. & Wargan, K. Reanalysis intercomparisons of stratospheric polar processing diagnostics. *Atmospheric Chemistry and Physics* **18**, 13547–13579 (2018).
57. Wang, P. *et al.* Data and code for "Fingerprinting the Recovery of Antarctic Ozone". *Zenodo* https://doi.org/10.5281/zenodo.14497873 (2024).

Acknowledgments

We thank Clara Deser and Pu Lin for helpful discussions. We also thank Larry Horowitz and Meiyun Lin for providing GFDL model data for this analysis. S.S. and P.W. gratefully acknowledge support from the atmospheric chemistry division of the National Science Foundation under grant 2316980 and 2128617. B.D.S. was supported by the Francis E. Fowler IV Center for Ocean and Climate at Woods Hole Oceanographic Institution (WHOI). D.E.K. was financed in part by NASA grant 80NSSC19K0952. Q.F. was in part supported by NSF Grant AGS-2202812. The Community Earth System Model (CESM) project is supported by the National Science Foundation and the Office of Science of the U.S. Department of Energy. We gratefully acknowledge high-performance computing support from Cheyenne (https://doi.org/10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory (CISL), sponsored by the National Science Foundation. Work at the Jet

595 Propulsion Laboratory, California Institute of Technology, was carried out under a contract with
596 the National Aeronautics and Space Administration (80NM0018D0004).
597

598 **Author Contributions**

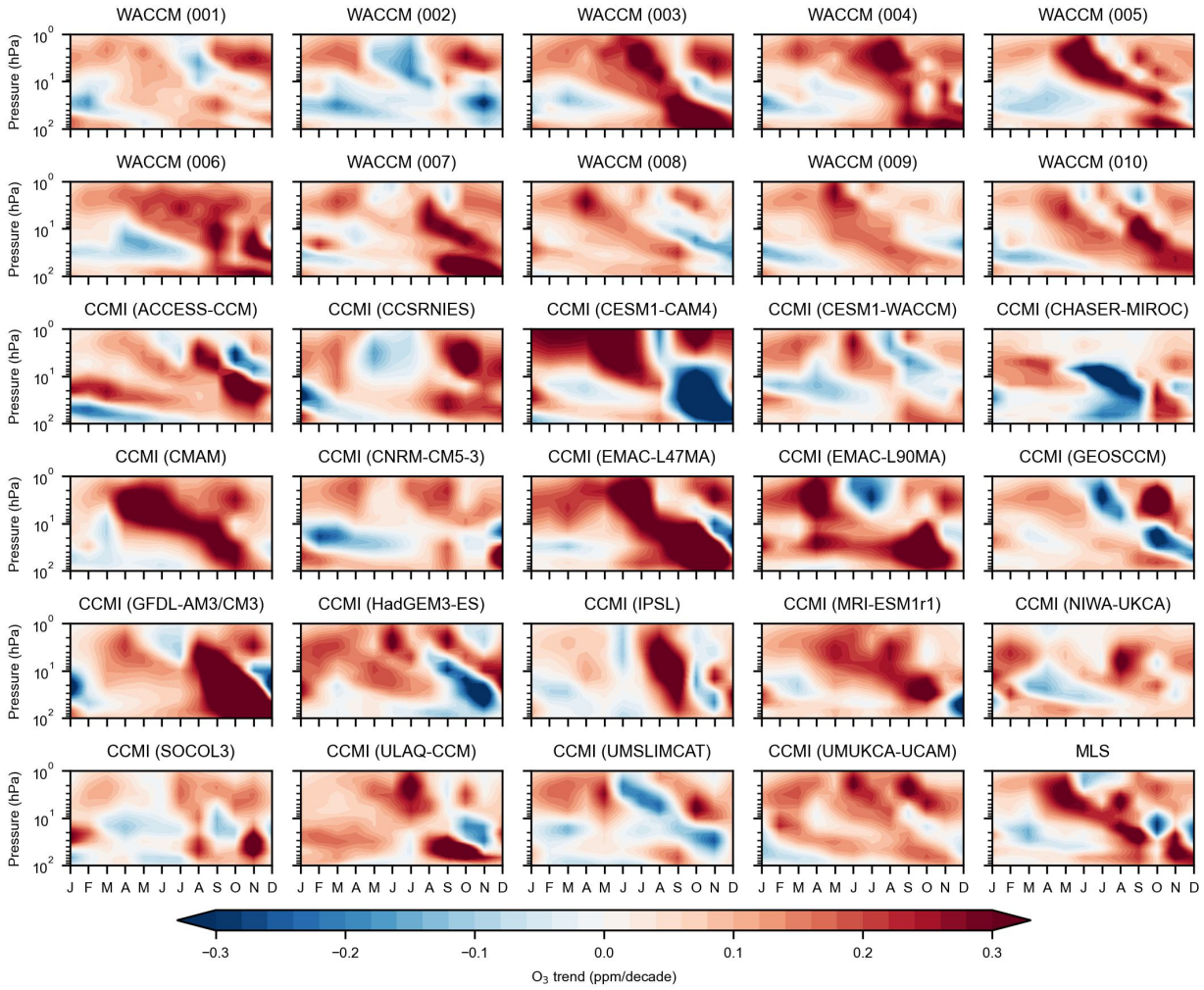
599 P.W., S.S., B.D.S. designed the study. D.E.K. designed and performed the WACCM simulations.
600 P.W. analyzed the data and produced the figures. P.W. and S.S. drafted the initial text. B.D.S.,
601 Q.F, K.A.S., J.Z., G.L.M., and L.F.M. contributed significantly to the interpretation of findings.
602

603 **Competing Interests**

604 The authors declare no competing interests.
605

606 **Correspondence and requests for materials**

607 Peidong Wang (pdwang@mit.edu) and Susan Solomon (solos@mit.edu)
608



609

610 **Extended Data Figure 1. Variability superimposed on external forcing.** Ozone trends from

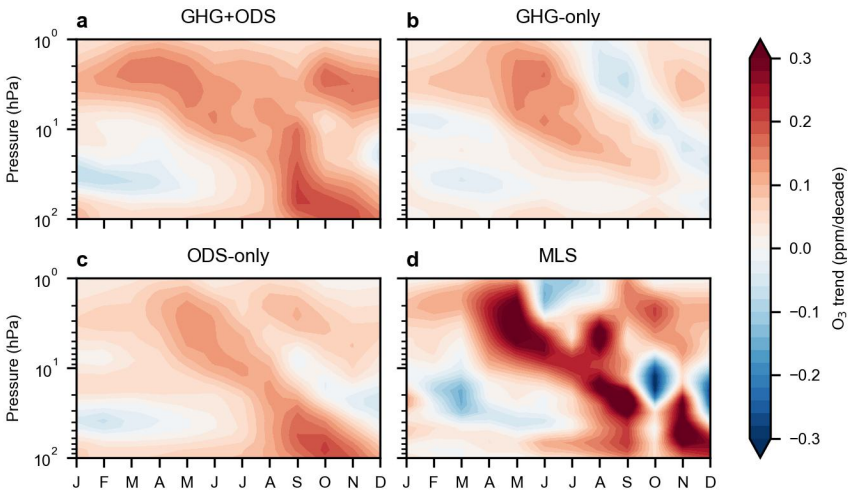
611 2005-2018 in individual WACCM realizations (top two rows, with the realization number

612 indicated in each panel's title) and in individual models from CCMI-1 (last four rows, with

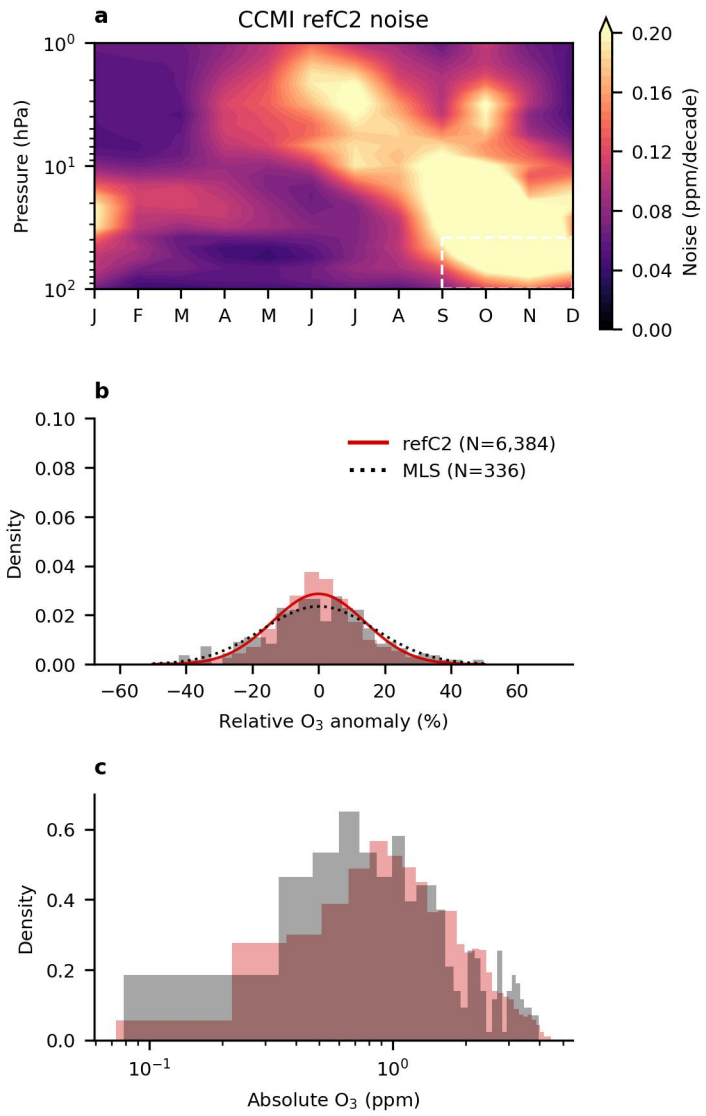
613 model names indicated in each panel's title) under the refC2 scenario. The MLS observed trend

614 is shown in the bottom right panel.

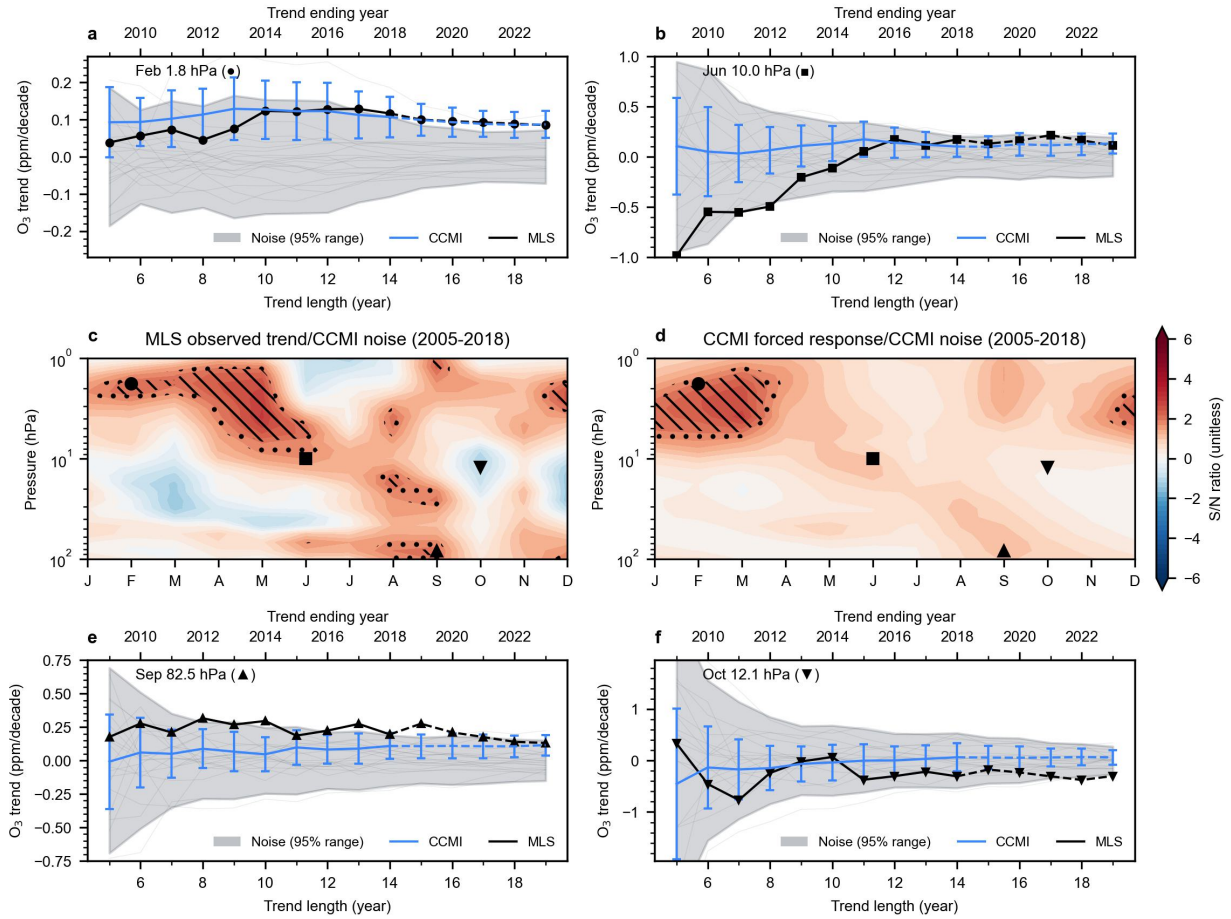
615



Extended Data Figure 2. Ozone trends due to different forcings. Ensemble-mean ozone trends (2005-2018) averaged over 10 WACCM members for each scenario. Results indicate the forced responses in ozone due to: **a** combined time-evolving GHG and ODS forcing (refC2), **b** evolving GHG forcing only (fODS); and **c** evolving ODS forcing only (fGHG). A detailed description of each scenario is given in the Methods section. The observed ozone trend from MLS in 2005-2018 is also shown in panel **d** for visual comparison with forced ozone trends due to different forcings.



Extended Data Figure 3. Ozone variability modulated by external forcing. Similar to Figure 2 but for the CCMI models. Note that the spread in ozone trends in CCMI arises not only from internal variability, but also from cross-model differences and errors (discussed in detail in the Methods section). This convolving of internal variability with model differences and errors contributes to the larger noise in panel **a** compared to the noise derived from the WACCM single-model refC2 ensemble in Figure 2b.

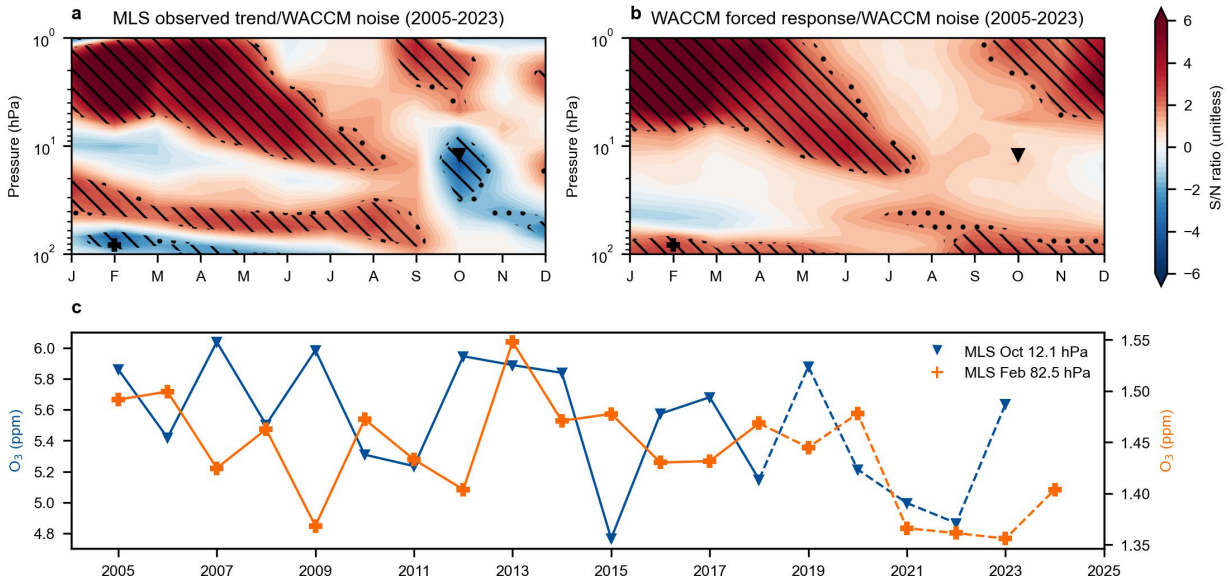


632

633 **Extended Data Figure 4. The local signal-to-noise characteristics of ozone changes.** Similar

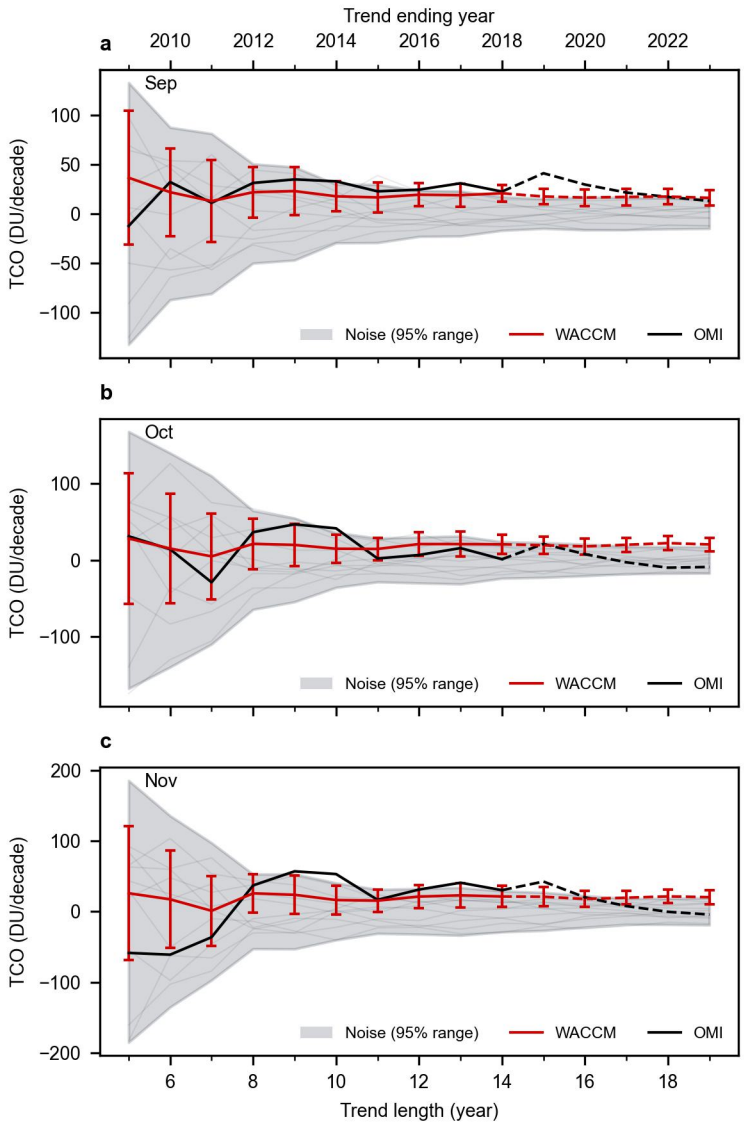
634 to Figure 3 but with results for the CCMI models.

635



636

637 **Extended Data Figure 5. The local signal-to-noise pattern of ozone changes.** Similar to
638 Figure 3 but for signal and noise estimates based on ozone trends over 2005-2023 (rather than
639 over 2005-2018). There is a significant decrease in MLS ozone in October and November in the
640 middle stratosphere, and in January through May in the lowermost stratosphere. This raises the
641 question of whether these two features may be linked. The bottom panel shows the timeseries of
642 ozone mixing ratios from MLS in October at 12.1 hPa (blue) and in February at 82.5 hPa
643 (orange). The decrease in February at 82.5 hPa is mainly due to continued low ozone after 2021
644 (panel **c**), which lags the behavior in October by about a season, suggesting that they may be
645 linked.



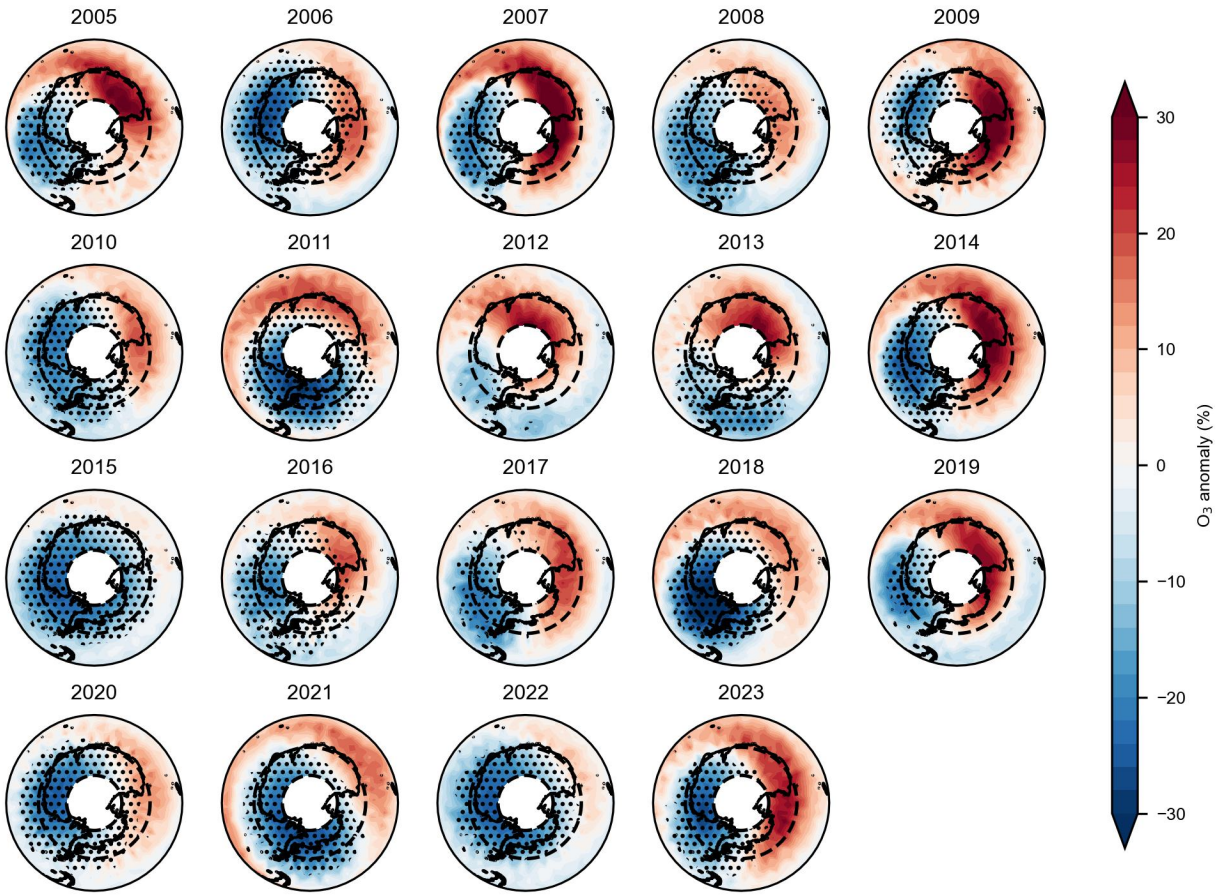
646

647 **Extended Data Figure 6. Time of emergence of springtime total column ozone recovery.**

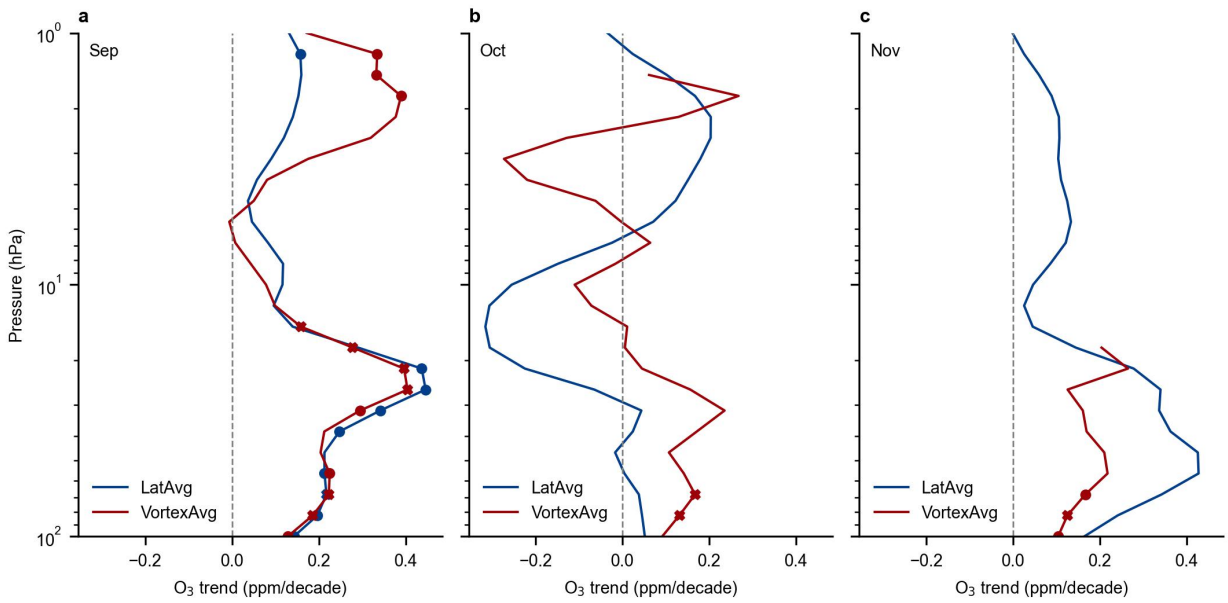
648 Similar to the top and bottom panels in Figure 3, except the trends are the TCO from WACCM

649 and OMI.

650



Extended Data Figure 7. Map of MLS ozone anomalies and polar vortex in October at 12.1 hPa. The color shadings indicate the ozone anomaly relative to the zonal mean in 2005-2018. Because the location of the polar vortex can vary considerably over time, dotted markers indicate that the polar vortex has occupied a given grid box for more than 25% of the time in that month. Black dashed lines encompass the area between latitudes 66°-82°S.



Extended Data Figure 8. MLS springtime ozone trends (2005-2018) using fixed latitude averages versus vortex averages. Monthly mean ozone trends in September, October, and November from 2005 to 2018 are shown as blue lines (where ozone is averaged across fixed latitudes between 66°-82°S) and red lines (where ozone is averaged inside the polar vortex) at different pressure levels. A detailed description of the vortex calculation is provided in the Methods section. Dots and crosses indicate trends significant at the 5% and 10% levels, respectively. Note that the statistical confidence in this figure is based solely on p-values from linear regression. It does not rely on internal variability noise generated by WACCM or the CCMI models, as shown in other figures.