nature geoscience

ARTICLES

https://doi.org/10.1038/s41561-021-00803-3

Emergence of Southern Hemisphere stratospheric circulation changes in response to ozone recovery

Brian Zambri[1]✉, Susan Solomon[1], David W. J. Thompson[2] and Qiang Fu[3]

Depletion of stratospheric ozone in the Southern Hemisphere (SH) during the late twentieth century cooled local air temperature, which resulted in stronger stratospheric westerly winds near 60° S and altered SH surface climate. However, Antarctic ozone has been recovering since around 2001 thanks to the implementation of the Montreal Protocol, which banned production and consumption of ozone-depleting substances. Here we show that the post-2001 increase in ozone has resulted in significant changes to trends in SH temperature and circulation. The trends are generally of opposite sign to those that resulted from stratospheric ozone losses, including a warming of the SH polar lower stratosphere and a weakening of the SH stratospheric polar vortex. Observed post-2001 trends of temperature and circulation in the stratosphere are about 50–75% smaller in magnitude than the trends during the ozone depletion era. The response is broadly consistent with expectations based on modelled depletion-era trends and variability of both ozone and reactive chlorine. The differences in observed stratospheric trends between the recovery and depletion periods are statistically significant ($P < 0.05$), providing evidence for the emergence of dynamical impacts of the healing of the Antarctic ozone hole.

Over the last ~30 years of the twentieth century, the atmosphere over Antarctica experienced profound changes. Total column ozone (O_3) losses exceeded 50% of pre-O_3-hole values during October throughout the 1990s[1]. Because of the decreased abundance of O_3, the Southern Hemisphere (SH) polar lower stratosphere cooled by more than 10 K during November–December from 1979 to 2001[2] (Fig. 1). The temperature (T) changes in the polar lower stratosphere led to circulation changes, including strengthened zonal winds (U) near 60° S throughout the stratosphere, extending into the troposphere and affecting SH surface climate through a strengthening of the Southern Annular Mode (SAM)[3,4]. Reported tropospheric impacts of circulation changes and stratospheric O_3 depletion include a strengthening and poleward shift of the SH extratropical westerly jet[5], increased summer precipitation over much of subtropical Australia[6], and cooler- and warmer-than-average summer Ts over subtropical Australia[6] and New Zealand[7], respectively. With the phaseout of O_3-depleting substances (ODS) under the Montreal Protocol, and despite an unexpected increase in chlorofluorocarbon-11 emissions since 2012[8], the Antarctic O_3 'hole' stabilized around 2001, and O_3 itself has now begun to show signs of healing according to a range of metrics[9–12]. As a result, trends in SH climate very different from those experienced with O_3 depletion should occur[5,13–16]. A negative phase of the SAM, such as was experienced in 2019–2020[17,18], has been associated with exceptionally hot and dry conditions in Australia[19]; a pause or reversal of recent positive SAM trends could act to accelerate warming trends in this region. Identifying changes in climate trends that signal the onset of changes in circulation as the O_3 hole begins to recover is the focus of this paper.

In this study, we use data from the Total Ozone Mapping Spectrometer/Ozone Monitoring Instrument (TOMS/OMI) merged ozone dataset[20,21] and the European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5)[22] to compare trends in the SH stratospheric circulation in the late twentieth and early twenty-first centuries. The natural variability of the climate system plays a large role in the observed trends in stratospheric ozone[9,23]. The real world represents only a single realization, or 'ensemble member'; for this reason, model ensembles are useful tools to help distinguish forced trends from those due to variability[24], provided that the model or models are capable of simulating a reasonably realistic response[25]. Therefore, we also make use of a ten-member ensemble of opportunity composed of simulations conducted with the Community Earth System Model 1 (CESM1) Whole Atmosphere Community Climate Model (WACCM)[26,27] to estimate significance of trends compared with variability. WACCM is an interactive climate–chemistry model, and its ability to represent polar ozone chemistry and climate has been previously documented[28,29]. We analyse trends in O_3, total inorganic chlorine (Cl_y), T and geopotential height (Z) over the Antarctic polar cap (65–90° S), along with U over the SH polar vortex edge region (55–65° S). Following the phasing out of ODS under the Montreal Protocol, the concentrations of these gases stopped rising in the polar stratosphere in 2001, as illustrated by the decrease in equivalent effective stratospheric chlorine (EESC[30]; Fig. 1a). Therefore, we calculate trends for ERA5 for the O_3 'depletion' era (1979–2001 for ERA5 and 1975–2001 for WACCM; see Methods for details on calculation of trends) and the 'recovery' era (2001–2018). Although still useful as an indicator, EESC has been shown to lead to quantitative errors in trends in O_3 if used for regression over the full period of O_3 loss and recovery[31]; we therefore calculate piecewise linear trends separately over the depletion and recovery eras and examine whether the changes in trends between the two eras are statistically significant[32]. A change from strong and systematic cooling to, for example, near zero or warming would be consistent with a turnaround in O_3 abundances.

Climate response to O_3 recovery

TOMS/OMI November total column O_3 shows the extensively studied deepening of the O_3 hole from 1979 to 2001, with November

[1]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. [2]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA. [3]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA.
✉e-mail: bzambri@mit.edu

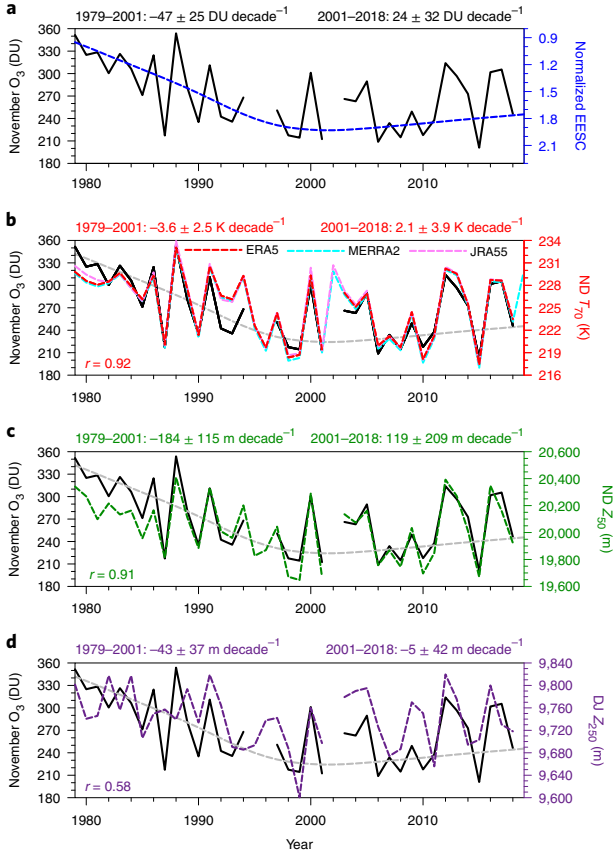


Fig. 1 | Time series of Antarctic O_3, indicators of Southern Hemisphere stratospheric climate and EESC. a–d, TOMS/OMI November total column O_3 (DU) and normalized EESC (**a**), November–December (ND) T_{70} (K) (**b**), ND Z_{50} (m) (**c**) and December–January (DJ) Z_{250} (m) (**d**). Above each panel are the linear trends and 95% confidence intervals for the indicated periods. The grey lines in **b**–**d** are the normalized EESC; *r*-values represent the Pearson correlation coefficients between the time series displayed in each panel. The O_3, *T* and *Z* are 65–90° S averages.

total column O_3 decreasing from 351 DU in 1979 to 212 DU in 2001, a decrease of about 40% (Fig. 1a). Stratospheric O_3 absorbs incoming solar radiation, providing a significant source of heat for the atmosphere there. The large decrease in Antarctic O_3 abundances resulted in less heating (net cooling) of the lower polar stratosphere and is accompanied by large concurrent negative trends in November–December 70 hPa *T* (T_{70}) and 50 hPa *Z* (Z_{50}) (Fig. 1b–c). The tight coupling between O_3 and *T* on the interannual time scale ($r = 0.92$ over the 41-year period; $P < 0.00001$) probably reflects both the radiative response of *T* to variations in O_3 and the response of both O_3 and *T* to dynamic variability associated with the large-scale stratospheric circulation. After the levelling off of concentrations of ODS in 2001 (consistent with the normalized EESC, Fig. 1a), Fig. 1a indicates that O_3 concentrations have slowly increased over 2001–2018, opposing the trend observed over the preceding decades. The subsequent cessation of SH stratospheric temperature and circulation trends in response to O_3 healing in the mid-twenty-first century has been predicted on numerous occasions using climate models[13–16]; Fig. 1b,c suggests that lower polar stratospheric *T*s have indeed begun to respond to the observed reversal in O_3 trends, with *T* and *Z* trends exhibiting similar changes between the two periods, increasing by about 2.1 K decade^{-1} and 119 m decade^{-1} over 2001–2018, respectively. Temperatures from the JRA55[33] and MERRA2[34] reanalyses show excellent agreement with the ERA5 *T* time series (Fig. 1b). While a complete reversal of the O_3 depletion-induced large-scale circulation trends in the troposphere is not yet evident given the large interannual variability[35], the 250 hPa upper tropospheric level suggests that the trend of decreasing *Z* observed during the depletion era has also flattened.

Differences in depletion and recovery trends

The post-2001 trends in Antarctic stratospheric O_3, *T* and *Z* are considerably smaller in magnitude than their counterparts during the depletion era (magnitudes of the recovery-period trends are 25–50% those of the depletion-era trends). This is certainly due in large part to the very long atmospheric lifetime of the ODS, which causes the EESC to decrease much more slowly than it originally increased (Fig. 1a). However, there are other factors that may contribute to the stronger trends over 1979–2001 and weaker trends over 2001–2018 in Fig. 1b–d. For example, continued increases in greenhouse gas (GHG) concentrations contribute to the cooling of the lower stratosphere, which works in the direction of the stratospheric cooling during the depletion era and against the warming trend observed after 2001[13].

The weaker dynamical trends over the past 18 years are expected, given the relatively small amplitudes of O_3 increases since ~2001 compared with the very large amplitude of losses during the preceding decades. However, the critical point is not that the trends during the recovery era are different from zero. Rather, the critical point is that the trends during the recovery era are different from *t*he trends during the depletion era, which is suggestive of a change in forcing. In Fig. 2, the observed trends are compared. The 95% confidence intervals are constructed using the adjusted standard error[32] to account for autocorrelation (Methods). Figure 2 shows that, in the lower stratosphere (Fig. 2a–c), the observed 2001–2018 trends in O_3, *T* and *Z* are significantly different ($P < 0.05$) from the 1979–2001 trends, further supporting that the altered trends over the past two decades do indeed indicate a change in forcing consistent with the cessation of Antarctic O_3 depletion. The change in trends is also generally insensitive to the choice of turnaround year (Extended Data Fig. 1): the trend differences between the recovery and depletion periods for O_3, T_{70} and Z_{50} are all statistically significant at the $P < 0.05$ level regardless of the four turnaround years used (the only exceptions are for T_{70} and Z_{50} with the turnaround year 2000 ($P < 0.072$ and $P < 0.068$, respectively) and Z_{50} with turnaround year 2002 ($P < 0.052$)). Trend differences in the upper troposphere (Fig. 2d) are of the expected sign but are generally not significantly different from zero at the $P < 0.05$ level except when considering 1999 as the turnaround year (Extended Data Fig. 1). The relatively weak differences at 250 hPa indicate that the signal of ozone recovery-induced climate change is not yet statistically significant in the troposphere.

Similar differences in behaviour of trends are evident throughout the SH polar lower stratosphere (Fig. 3). Before 2001, large cooling trends (up to 6 K decade^{-1}) are observed from about 250 hPa to 40 hPa, extending from the pole to about 65° S (Fig. 3a–b). The stratospheric cooling is in contrast to the warming trends over the 2001–2018 period (Fig. 3d,e); although the warming is weaker, the trend differences over most of the polar lower stratosphere are significant at the 95% level. The WACCM ensemble-mean trends are also shown. Compared with the ERA5 and Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2) trends, WACCM generally simulates less cooling in the depletion era (Fig. 3c); the method used to calculate the WACCM depletion-era trends probably underestimates the trend (Methods), although there is also a contribution from averaging over ensemble members, which smooths out some of the internal variability.

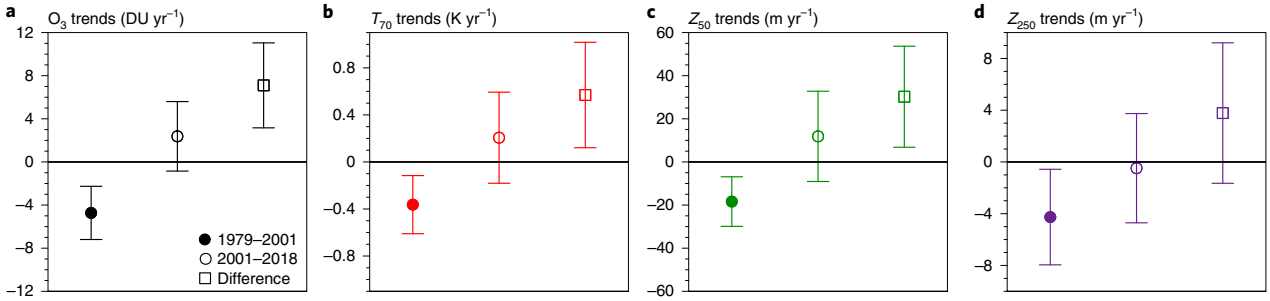


Fig. 2 | SH trend differences. a–d, Trends in November O_3 (DU yr^{-1}) (**a**), ND T_{70} (K yr^{-1}) (**b**), ND Z_{50} (m yr^{-1}) (**c**) and DJ Z_{250} (m yr^{-1}) (**d**) for 1979–2001 (filled circles), 2001–2018 (open circles) and the difference (squares). Error bars represent the adjusted 95% confidence intervals of the trends (Methods).

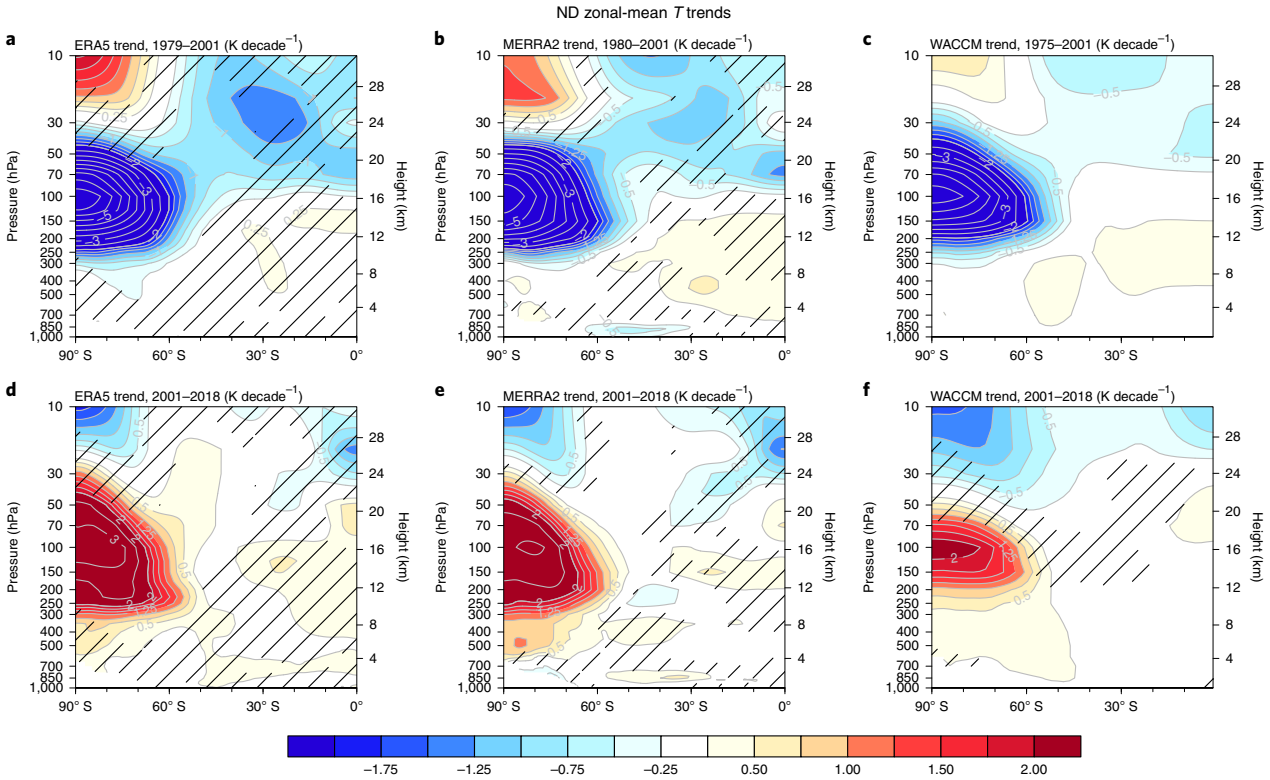


Fig. 3 | SH ND zonal-mean *T* trends. a–f, SH ND zonal-mean *T* trends (K decade^{-1}) for ERA5 1979–2001 (**a**), MERRA2 1979–2001 (**b**), WACCM ensemble mean, 1975–2001 (**c**), and ERA5 (**d**), MERRA2 (**e**) and WACCM ensemble mean (**f**) 2001–2018. Contours above 2 K decade^{-1} are drawn at intervals of 0.5 K decade^{-1}. Hatching for reanalysis indicates regions where the trend differences are insignificant at the 95% level; hatching for the model indicates regions where the trend is insignificant at the 95% level using a two-tailed Student's *t*-test based on the variability in the ensemble.

The latter effect is also seen in Fig. 3f, which shows generally lower magnitudes in the WACCM ensemble-mean trends for 2001–2018 compared with the reanalysis.

Figure 4 further illustrates the latter effect by showing trends over the depletion and recovery eras for each of the ten ensemble members. There is noticeably less spread between ensemble members in the depletion-era cooling (Fig. 4a) due to the relatively large forcing then. While there is substantial variability among ensemble members in the recovery era (Fig. 4b), as would be expected due to the weaker forcing, 9 of the 10 ensemble members do indicate warming trends in the SH polar lower stratosphere. The WACCM ensemble mean can be interpreted as representing the 'forced' response, while reanalysis represents the single realization available in the real world, which includes substantial contributions from natural variability. The variability among the ensemble members highlights the need for and utility of large numbers of simulations in a model ensemble aimed at distinguishing forced from natural responses.

The SH circulation trends of the late twentieth century were shown to be highly seasonal[3]. This is illustrated in Fig. 5, which shows that the largest *Z* trends in the depletion era occurred from about September to December in the stratosphere, with maximum magnitude trend near 30 hPa in November. In addition, *Z* trends in the troposphere peak in December–January. ERA5 *Z* trends again

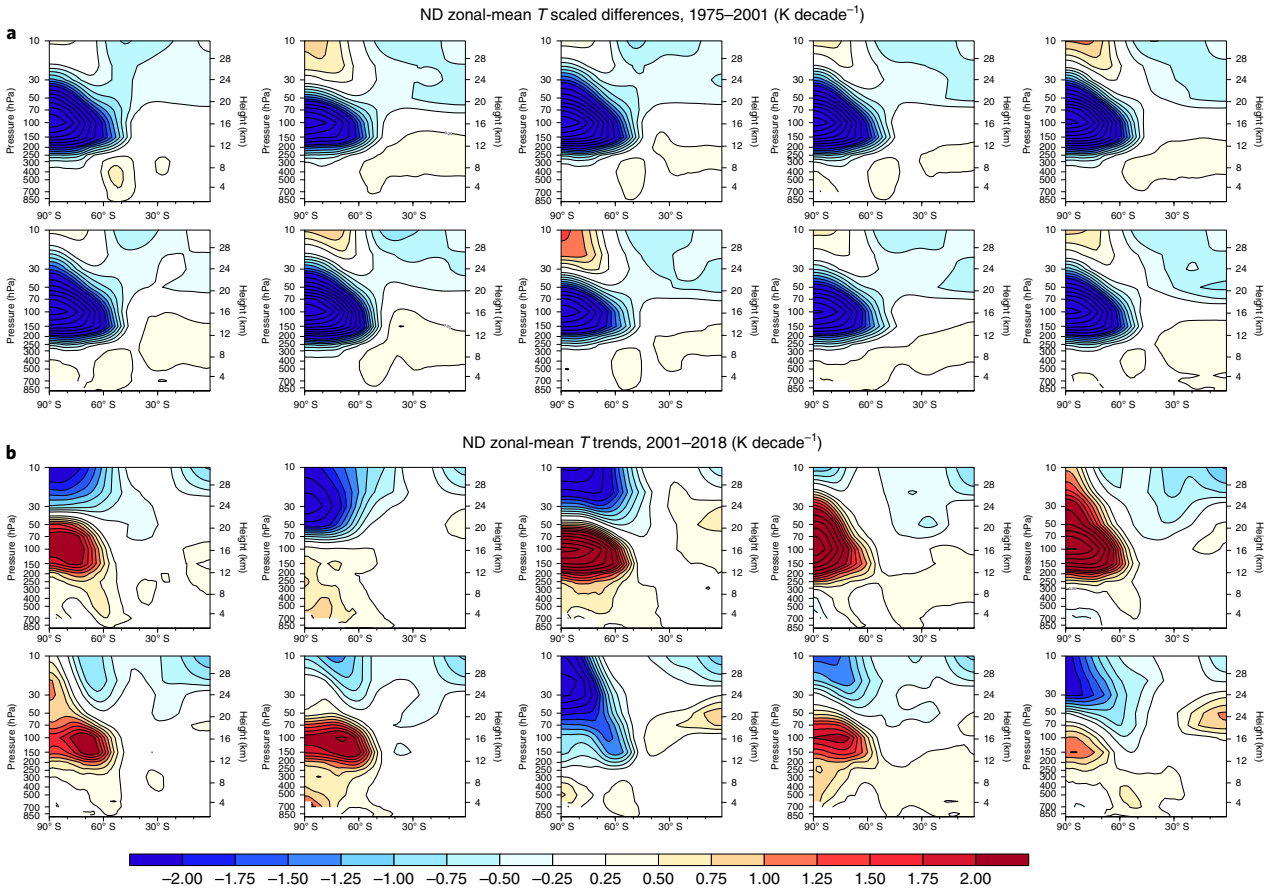


Fig. 4 | Variability in SH *T* trends. a,**b**, SH ND zonal-mean *T* trends (K decade^{-1}) for individual WACCM ensemble members for 1975–2001 (**a**) and 2001–2018 (**b**). Contours above 2 K decade^{-1} in the absolute value are drawn at intervals of 0.5 K decade^{-1}.

show a clear change from 2001 to 2018, with generally increasing *Z*s but of smaller magnitude, over this period. However, in contrast to the depletion-era trends, the maximum trend for the recovery period is shifted to December, while still occurring near 30 hPa. The temporal shift in Fig. 5 is due to the influence of GHGs in the WACCM model (Extended Data Fig. 2). Consistent with Fig. 1 and the comparably small upward trend in O_3 from 2001 to 2018 shown there, magnitudes of *Z* trends for 2001–2018 are about 25–50% of those over 1979–2001. As a result, the reversals of the lower tropospheric trends in *Z* and the SAM are not yet evident given variability.

Models as validation

WACCM geopotential height trends show broad agreement in spatial and temporal evolution with the trends in the reanalysis, although again the magnitudes of the trends are generally smaller in the ensemble average, which removes some of the internal variability (Fig. 5). Figure 5c,f shows the WACCM signal-to-noise ratios (SNRs[23,36]; Methods), which identify where one should expect to see robust forced trends, that is, to distinguish significant trends against the variability in the reanalysis. As expected, SNRs are much higher for the depletion era (Fig. 5c) than for the recovery era (Fig. 5f). Although the magnitude of the depletion-era trends is largest in November near 30 hPa, the SNR for the model ensemble is greatest in January due to the reduced variability in the SH polar stratosphere in summer. During the recovery era, SNRs reach a maximum of 2 in the lowermost stratosphere in December and March. They are between 1 and 2 in the uppermost troposphere near 250 hPa, where the changes in *Z* trends can be expected to affect eddy activity and storm tracks[37]. By contrast, the low SNR in the recovery period in April–September indicates that the trends in ERA5 in those months are likely unforced and are instead due to natural variability, although contributions from changes in the Brewer–Dobson circulation are possible in September[38].

Analysis of the pre- and post-2001 trends in O_3 in the WACCM ensemble mean can be viewed as estimates of the forced response to the trends in O_3-depleting Cl_y (Fig. 6). Similarly, WACCM ensemble-mean trends in *T* and circulation follow from the forced O_3 trends, although O_3 is not the lone contributor to the changes in these simulations. Figure 6 indicates that modelled stratospheric chlorine decreased since 2001 at a rate of about one-half to one-third the rate at which it increased over 1975–2001. In the WACCM ensemble mean, stratospheric levels of Cl_y decrease overall by as much as 32% decade^{-1} over 2001–2018, with the maximum occurring in January (Fig. 6e). The WACCM ensemble mean indicates a maximum O_3 recovery of 12–14% decade^{-1} at 150 hPa in October through January (Fig. 6f), broadly consistent with what would be expected on the basis of the Cl_y–O_3 trend relationship seen in the depletion era. The O_3 change leads to maximum warming trends in the stratosphere of more than 2 K decade^{-1} and a decrease in the stratospheric *U*s of about 2 m s^{-1} decade^{-1}, both occurring in December (Fig. 6g,h), again consistent with what would be expected on the basis of the ratios of each to the O_3 trends seen in the depletion era.

Stratospheric *T* and *U* trends exhibit the expected temporal sequencing in both periods: lagged cooling and strengthened *U*s

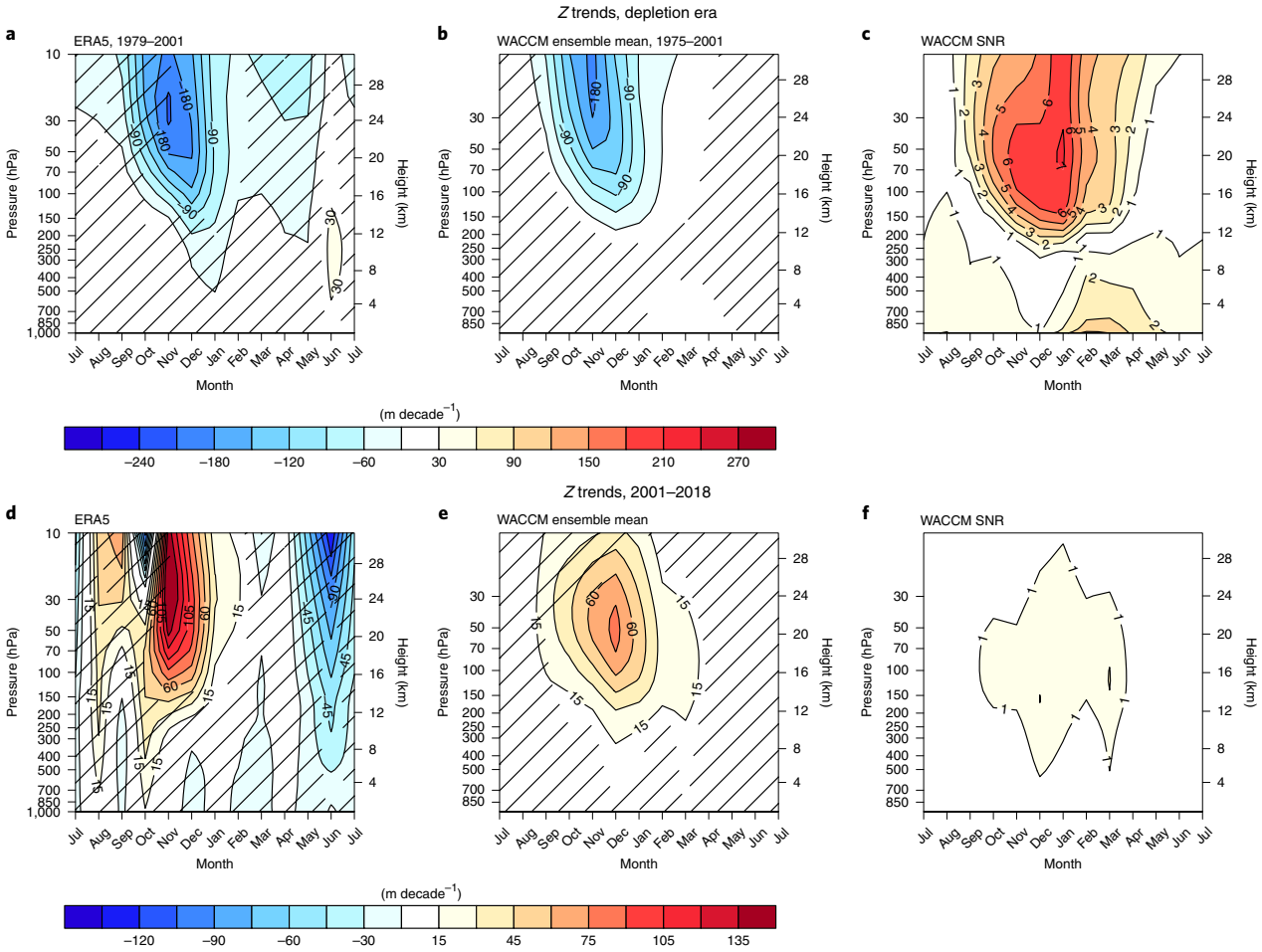


Fig. 5 | Seasonality of *Z* trends. a,**b**,**d**,**e**, Monthly polar-cap averaged (65–90°S) *Z* trends (m decade^{-1}) for ERA5 1979–2001 (**a**), WACCM ensemble mean, 1975–2001 (**b**), and ERA5 (**d**) and WACCM ensemble mean (**e**) 2001–2018. **c**,**f**, WACCM trend SNR for the 1975–2001 (**c**) and 2001–2018 (**f**) periods. Hatching for reanalysis indicates regions where the trend differences are insignificant at the 95% level; hatching for the model indicates regions where the trend is insignificant at the 95% level using a two-tailed Student's *t*-test based on the variability in the ensemble.

following O_3 depletion, and similarly lagged warming and weakened *U*s following the positive trends in O_3 (Fig. 6). Similar to the *Z* trends in Fig. 5, the maximum magnitudes of trends in *T* and *U* in the recovery era are shifted downwards and occur later relative to the depletion-era trends; this is once again due to the contribution of increasing GHG concentrations in the model[13] (Extended Data Fig. 3).

Implications for future O_3-based climate changes

Fingerprints of O_3 healing have been identified through different characterizations of the Antarctic O_3 hole, such as decreases in its areal extent[9]. Here we have identified and detected the associated impacts of O_3 healing on the thermal and dynamical structure of the atmosphere. The magnitudes of the *T* and circulation trends since 2001 are smaller than their O_3-depletion-era counterparts, as expected. However, the differences in the trends between the O_3-depletion and O_3-healing eras are statistically significant, indicating a change in ODS-based climate forcing. The differences in observed trends provide further evidence that O_3 healing is indeed under way, while also validating past predictions made about the impacts of future O_3 recovery[14–16].

The observed trends are convolved with interannual and other sources of natural variability. The significance of the ensemble-mean model trends in both periods highlights the importance and utility of ensemble modelling for disentangling a relatively small forcing from various sources of natural variability. We made use of ensemble simulations from a single coupled climate–chemistry model. But it would be interesting to explore the responses across multiple climate models to probe model-to-model uncertainty in the forced response[16], which can provide complementary predictive skill[39]. We emphasize that the results presented here are robust to the methods used for ensemble initialization, trend calculation, and between two different reanalyses (Extended Data Figs. 4–9).

The changes in stratospheric composition and circulation due to Antarctic O_3 depletion are juxtaposed upon numerous other changes in SH surface climate[3,4,6,7,40], albeit whether the various changes are causal or coincidental has not been determined. The low *Z*s over the Antarctic during the O_3-depletion period are consistent with anomalously strong westerly flow along ~60° S and the positive polarity of the SAM, and vice versa for the recovery period. Recent *U* trends near 60° S not only are no longer positive but suggest a small negative trend since 2001. A weakened polar vortex and trend towards the negative polarity of the SAM have been associated with exceptionally hot and dry conditions in Australia[19], and it follows that a continued weakening of the SH stratospheric

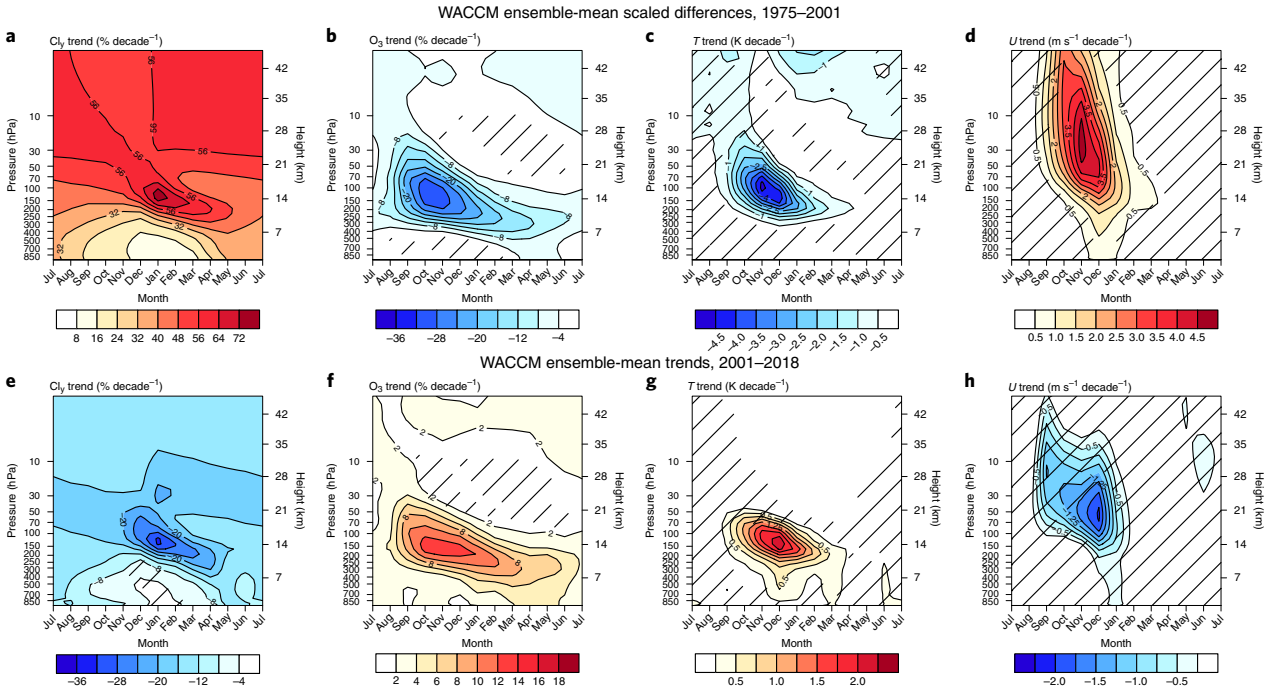


Fig. 6 | SH model trends in Cl$_y$, O$_3$, *T* and U. a–h, WACCM ensemble-mean trends in Cl$_y$ (% decade^{-1}) (**a**,**e**), O$_3$ (% decade^{-1}) (**b**,**f**), *T* (K decade^{-1}) (**c**,**g**) and *U* (m s^{-1} decade^{-1}) (**d**,**h**) for 1975–2001 (**a**–**d**) and 2001–2018 (**e**–**h**). Note the contour intervals in the bottom row are half those of the top row. Cl$_y$, O$_3$ and *T* are averaged over the polar cap (65–90° S); *U* is averaged over 55–65° S. Hatching indicates regions where the trend is insignificant at the 95% level using a two-tailed Student's *t*-test based on the variability in the ensemble.

summer circulation could act to accelerate warming trends in this region. Further O$_3$ healing can be expected to begin to affect storm tracks and potentially surface climate throughout much of the SH as recovery proceeds.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41561-021-00803-3.

Received: 20 December 2019; Accepted: 23 June 2021;
Published online: 23 August 2021

References

1. *Scientific Assessment of Ozone Depletion: 1998* Global Ozone Research and Monitoring Project Report 44 (WMO, 1999).
2. Randel, W. J. & Wu, F. Cooling of the Arctic and Antarctic polar stratospheres due to ozone depletion. *J. Clim.* **12**, 1467–1479 (1999).
3. Thompson, D. W. J. & Solomon, S. Interpretation of recent Southern Hemisphere climate change. *Science* **296**, 895–899 (2002).
4. Gillett, N. P. & Thompson, D. W. J. Simulation of recent Southern Hemisphere climate change. *Science* **302**, 273–275 (2003).
5. Son, S. -W., Tandon, N. F., Polvani, L. M. & Waugh, D. W. Ozone hole and Southern Hemisphere climate change. *Geophys. Res. Lett.* **36**, L15705 (2009).
6. Hendon, H. H., Thompson, D. W. J. & Wheeler, M. C. Australian rainfall and surface temperature variations associated with the Southern Hemisphere annular mode. *J. Clim.* **20**, 2452–2467 (2007).
7. Renwick, J. A. & Thompson, D. W. J. The Southern Annular Mode and New Zealand climate. *Water Atmos.* **14**, 24–25 (2006).
8. Rigby, M. et al. Increase in CFC-11 emissions from eastern China based on atmospheric observations. *Nature* **569**, 546–560 (2019).
9. Solomon, S. et al. Emergence of healing in the Antarctic ozone layer. *Science* **353**, 269–274 (2016).
10. Solomon, S. et al. Mirrored changes in Antarctic ozone and stratospheric temperature in the late 20th versus early 21st centuries. *J. Geophys. Res. Atmos.* **122**, 8940–8950 (2017).
11. *Scientific Assessment of Ozone Depletion: 2018* (WMO, 2018); https://www.esrl.noaa.gov/csd/assessments/ozone/2018/downloads/2018OzoneAssessment.pdf
12. Strahan, S. E. & Douglass, A. R. Decline in Antarctic ozone depletion and lower stratospheric chlorine determined from Aura Microwave Limb Sounder observations. *Geophys. Res. Lett.* **45**, 382–390 (2018).
13. Shindell, D. T., Rind, D. & Lonergan, P. Increased polar stratospheric ozone losses and delayed eventual recovery owing to increasing greenhouse-gas concentrations. *Nature* **392**, 589–592 (1998).
14. Son, S.-W. et al. The impact of stratospheric ozone recovery on the Southern Hemisphere westerly jet. *Science* **320**, 1486–1489 (2008).
15. Son, S.-W. et al. Impact of stratospheric ozone on Southern Hemisphere circulation change: a multimodel assessment. *J. Geophys. Res.* **115**, D00M07 (2010).
16. Banerjee, A. et al. A pause in Southern Hemisphere circulation trends due to the Montreal protocol. *Nature* **579**, 544–548 (2020).
17. Lim, E.–P. et al. The 2019 Antarctic sudden stratospheric warming. *SPARC Newsl.* **54**, 10–13 (2020).
18. Van Oldenborgh, G. J. et al. Attribution of the Australian bushfire risk to anthropogenic climate change. *Nat. Hazards Earth Syst. Sci.* **21**, 941-960 https://doi.org/10.5194/nhess-21-941-2021 (2021).
19. Lim, E.–P. et al. Australian hot and dry extremes induced by weakenings of the stratospheric polar vortex. *Nat. Geosci.* **12**, 896–901 (2019).
20. McPeters, R. D. et al. *Earth Probe Total Ozone Mapping Spectrometer (TOMS) Data Products User's Guide* NASA/TP-1998–206895 (NASA, 1998).
21. McPeters, R. et al. Validation of the Aura Ozone Monitoring Instrument total column ozone product. *J. Geophys. Res.* **113**, D15S14 (2008).
22. *ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate* (Copernicus Climate Change Service Climate Data Store, accessed 3 October 2019); https://cds.climate.copernicus.eu/cdsapp#!/home
23. Stone, K. A., Solomon, S. & Kinnison, D. E. On the identification of ozone recovery. *Geophys. Res. Lett.* **45**, 5158–5165 (2018).
24. 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. *Bull. Am. Met. Soc.* **96**, 1333–1349 (2015).
25. Seviour, W. J. M. Weakening and shift of the Arctic stratospheric polar vortex: internal variability or forced response? *Geophys. Res. Lett.* https://doi.org/10.1002/2017GL073071 (2017).

26. Marsh, D. R. et al. Climate change from 1850 to 2005 simulated in CESM1(WACCM). *J. Clim.* **26**, 7372–7391 (2013).
27. Garcia, R. R., Smith, A. K., Kinnison, D. E., de la Cámara, Á. & Murphy, D. J. Modification of the gravity wave parameterization in the Whole Atmosphere Community Climate Model: motivation and results. *J. Atmos. Sci.* **74**, 275–291 (2017).
28. Solomon, S., Kinnison, D., Bandoro, J. & Garcia, R. Simulation of polar ozone depletion: an update. *J. Geophys. Res. Atmos.* **120**, 7958–7974 (2015).
29. Calvo, N., Garcia, R. R. & Kinnison, D. E. Revisiting Southern Hemisphere polar stratospheric temperature trends in WACCM: the role of dynamical forcing. *Geophys. Res. Lett.* **44**, 3402–3410 (2017).
30. Velders, G. J. M. & Daniel, J. S. Uncertainty analysis of projections of ozone-depleting substances: mixing ratios, EESC, ODPs, and GWPs. *Atmos. Chem. Phys.* **14**, 2757–2776 (2014).
31. Kuttippurath, J., Bodeker, G. E., Roscoe, H. K. & Nair, P. J. A cautionary note on the use of EESC-based regression analysis for ozone trend studies. *Geophys. Res. Lett.* **42**, 162–168 (2015).
32. Santer, B. D. et al. Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *J. Geophys. Res.* **105**, 7337–7356 (2000).
33. Kobayashi, S. et al. The JRA-55 reanalysis: general specifications and basic characteristics. *J. Meteor. Soc. Japan* **93**, 5–48 (2016).
34. Gelaro, R. et al. The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454 (2017).
35. Seviour, W. J., Waugh, D. W., Polvani, L. M., Correa, G. J. P. & Garfinkel, C. L. Robustness of the simulated tropospheric response to ozone depletion. *J. Clim.* **30**, 2577–2585 (2017).
36. Zambri, B. et al. Modeled and observed volcanic aerosol control on stratospheric NO_y and Cl_y. *J. Geophys. Res. Atmos.* https://doi.org/10.1029/2019JD031111 (2019).
37. Yuval, J. & Kaspi, Y. Eddy activity sensitivity to changes in the vertical structure of baroclinicity. *J. Atmos. Sci.* **73**, 1709–1726 (2016).
38. Fu, Q. S., Solomon, S., Pahlavan, H. A. & Lin, P. Observed changes in the Brewer–Dobson circulation for 1980–2018. *Envir. Res. Lett.* **14**, 114026 (2019).
39. Smith, D. M. et al. Robust skill of decadal climate predictions. *NPJ Clim. Atmos. Sci.* **2**, 13 (2019).
40. Bandoro, J., Solomon, S., Donohoe, A., Thompson, D. W. J. & Santer, B. Influences of the Antarctic ozone hole on Southern Hemisphere summer climate change. *J. Clim.* **27**, 6245–6264 (2014).

Methods

ERA5 stratospheric *T* bias. A stratospheric *T* bias spanning the period 2000–2006 was discovered for the ERA5 reanalysis product[41]. The updated ERA5.1 was not available at the time this manuscript was submitted, so we instead included MERRA2 and Japanese 55-year reanalysis (JRA55) temperature in Fig. 1 and MERRA2 temperature trends in Fig. 3. In both cases, ERA5 is nearly identical to the other reanalysis products; therefore, the apparent temperature bias does not affect the analysis carried out herein.

CESM1 (WACCM) ensembles. CESM1 is a fully coupled climate model with atmosphere, ocean, land and sea-ice components[42]. We use the free-running version of WACCM version 4, the high-top atmospheric component of CESM1. This version includes updates to the gravity wave parameterizations as well as of the heterogeneous chemistry in the model[27]. The model has a horizontal resolution of 1.9° latitude by 2.5° longitude and 66 vertical levels with a model top near 140 km in altitude[26,27,28]. The WACCM chemical scheme used in the study includes representations of chemistry in the troposphere, stratosphere, mesosphere and lower thermosphere[43,44]. The species included within this mechanism are contained within the O_x, NO_x, HO_x, ClO_x and BrO_x chemical families, along with CH_4 and its degradation products. In addition, 20 primary nonmethane hydrocarbons and related oxygenated organic compounds are represented along with their surface emissions. There is a total of 183 different species, 341 gas-phase reactions, 114 photolytic processes and 17 heterogeneous reactions on aerosols (sulfate, nitric acid trihydrate and water ice).

Using this model, two fully coupled ensembles of simulations are generated. The ensembles span the pre-O_3-depletion era of 1955–1979 and the beginnings of the recovery era, 1995–2024. These simulations have a repeated cyclic 28-month quasi-biennial oscillation based on rocketsonde data[45] and no solar cycle. Sulfate aerosol surface-area densities include volcanic and nonvolcanic sources and are specified on the basis of calculations from Mills et al.[46], which used volcanic SO_2 injections from Neely and Schmidt[47] for the period spanning 1999–2014. Ensemble members are generated by randomly perturbing the initial *T* fields of a single Chemistry–Climate Model Initiative (CCMI[11]) REF-C2 simulation by roundoff magnitude (order 10^{-14} K). This simulation was previously initialized from a previously spun-up control run. GHG concentrations in the simulations evolve according to the representative concentration pathway 6.0[48,49], and concentrations of O_3-depleting substances are prescribed according to the CCMI[50]. The 95% significance on the ten-member ensemble-mean trends is calculated using a Student's *t*-test based on the variability in trends among ensemble members.

To analyse differences in the methods for calculating trends, we also use the three REF-C2 simulations from the WACCM contribution to the CCMI. Both the ten-member ensemble and the three-member CCMI ensemble were initialized using 'micro-perturbations', adding random perturbations of magnitude 10^{-14} to the initial atmospheric *T* field. While this method of ensemble initialization is standard for atmospheric trend attribution studies[51], we also performed the same analysis on the three-member ensemble from the Climate Model Intercomparison Project, version 6 (CMIP6[52]; Extended Data Figs. 4 and 5) to illustrate the robustness of the results to the ensemble initialization method. In particular, the WACCM–CMIP6 ensemble was initialized using 'macro-perturbations', using different years from a pre-industrial control run to sample dominant modes of climate variability. The 95% confidence intervals on the trend differences are computed by constructing the Student's *t*-distributions of trend differences for non-overlapping, consecutive 40-year periods. For each 40-year period, we calculate the difference in trends between the first 23 and last 18 years, which corresponds to the trend differences between 1979–2001 (O_3 depletion) and 2001–2018 (O_3 recovery). For WACCM–CCMI, a 1,000-year control run yielded a total of 42 trend differences; a 500-year control run for WACCM–CMIP6 yielded a total of 20 trend differences.

Hatching in Extended Data Figs. 4 and 5 indicates regions where the trend differences are not significantly different from the distributions of trend differences in the control runs ($P > 0.05$). Extended Data Fig. 4 shows that the trend differences for O_3, *T* and *Z* are significant in the lower stratosphere in austral late spring/early summer in the CMIP6 ensemble, while Extended Data Fig. 5 shows that the trend differences in the ND lower stratospheric *T* are significant in both ensembles ($P < 0.05$), demonstrating the robustness of the impact of the change in forcing to micro- and macro-perturbations in the ensembles.

Choice of turnaround year. We use 2001 as the turnaround year and calculate trends for 1979–2001 (or 1975–2001; see the following) and 2001–2018 because 2001 is the year of maximum EESC (Fig. 1) and so is the point at which O_3 and O_3-induced SH climate change recovery should begin to occur. The trends obtained by defining other years (1999, 2000, 2002) as turnaround years are reported in Extended Data Fig. 1. In addition, we exclude years 2002 and 2019 because of the breakdown of the SH stratospheric polar vortex in these years, which leads to anomalously high O_3 values.

Calculating trends. For the 1979–2001 period in ERA5 and for the 2001–2018 period for both ERA5 and WACCM, we compute the trends using simple linear regression, where for example, for monthly O_3 anomalies, we compute

$$O_3'(t) = b_0 + b_1 x(t) + R(t), \tag{1}$$

where $O_3'(t)$ represents the O_3 anomaly, b_0 is the constant term, b_1 is the linear trend coefficient corresponding to the linear regression function $x(t)$, and $R(t)$ is the residual.

The WACCM ensembles span 1955–1979 and 1995–2024, respectively. Therefore, the full 1979–2001 period analysed in ERA5 is not simulated explicitly. For the 1975–2001 period, WACCM 'trends' are actually scaled differences and are calculated by taking the difference of the means of two ten-year periods and dividing by the number of years between them. In particular, using O_3 as an example again, we take the mean of the 1970–1979 and 1996–2005 periods, difference them, and divide by the 26 years between the two interval midpoints:

$$b(O_3, 1975-2001) = \frac{\overline{O_3(1996-2005)} - \overline{O_3(1970-1979)}}{26} \tag{2}$$

Because the years 1980–1994 are not explicitly simulated for each ensemble member, the two ensembles are not coherent on a member-by-member basis, although they are coherent on an ensemble-mean basis. Therefore, we assess the depletion-era response by subtracting the ensemble-mean 1970–1979 climatology from the 1996–2005 mean of each ensemble member and then averaging:

$$b(O_3, 1975-2001) = \frac{1}{26}\frac{1}{10}\sum_{i=1}^{10}\left(\overline{O_3(1996-2005)_i} - \frac{1}{10}\sum_{j=1}^{10}\overline{O_3(1970-1979)_j}\right), \tag{3}$$

where *i* and *j* are the ensemble members. Furthermore, we assess the variance by generating the 100 possible combinations of ensemble members:

$$b_{ij}(O_3, 1975-2001) = \frac{1}{26}\left(\overline{O_{3i}(1996-2005)} - \overline{O_{3j}(1970-1979)}\right); i, j = 1, 2, \ldots, 10 \tag{4}$$

The mean in equation (3) is the same as, for example, averaging over the set defined by equation (4), subtracting the 1970–1979 climatology (the average over all ensemble members) from each 1996–2005 ensemble member, subtracting the means of the two ensembles or taking the mean difference of each of the ten ensemble members for each period:

$$\frac{1}{10}\sum_{i=1}^{10}\left(\overline{O_3(1996-2005)_i} - \frac{1}{10}\sum_{j=1}^{10}\overline{O_3(1970-1979)_j}\right) = \frac{1}{100}\sum_{i=1}^{10}\sum_{j=1}^{10}\overline{O_3(1996-2005)_i} - \overline{O_3(1970-1979)_j} \tag{5}$$

$$= \frac{1}{10}\sum_{i=1}^{10}\overline{O_3(1996-2005)_i} - \frac{1}{10}\sum_{j=1}^{10}\overline{O_3(1970-1979)_j} \tag{6}$$

$$= \frac{1}{10}\sum_{i=1}^{10}\overline{O_3(1996-2005)_i} - \frac{1}{10}\sum_{i=1}^{10}\overline{O_3(1970-1979)_i} \tag{7}$$

$$= \frac{1}{10}\sum_{i=1}^{10}\left(\overline{O_3(1996-2005)_i} - \overline{O_3(1970-1979)_i}\right) \tag{8}$$

To consider the implications of this methodological choice, the trends for 1975–2001 and 2001–2018 are calculated using equation (1) and equation (2) separately for a three-member ensemble of WACCM–CCMI simulations; results are displayed in Extended Data Fig. 5. Differences for the 1975–2001 period are up to 1.5 K decade^{-1}, with equation (2) underestimating the trends (Extended Data Fig. 5c). However, the impact of using equation (2) instead of equation (1) is probably smaller for the ten-member ensemble. Extended Data Fig. 6 shows the same comparison for the ten-member ensemble for 2001–2018. For the average of ten ensemble members, the differences in the polar lower stratosphere are everywhere below 0.5 K decade^{-1} (Extended Data Fig. 6c), as compared with differences up to 1 K decade^{-1} for the three-member ensemble (Extended Data Fig. 5f). Furthermore, Extended Data Fig. 7 shows that for $n > 6$, the trend differences in the polar lower stratosphere are below 0.5 K decade^{-1}. Therefore, it is likely that underestimates of the 1975–2001 WACCM trends as a result of using equation (2) are less than 1 K decade^{-1}.

Finally, Extended Data Fig. 8 shows the JRA55 temperature trends calculated for 1975–2001 (Extended Data Fig. 8a), 1979–2001 (Extended Data Fig. 8b) and the difference (Extended Data Fig. 8c). Differences are everywhere below 0.5 K decade^{-1}, which indicates that comparing 1979–2001 trends in ERA5 with 1975–2001 trends in WACCM is appropriate.

SNR. We also make use of the ensemble of coupled climate–chemistry model simulations by analysing the SNR[53]. We define the SNR as the ensemble-mean

trend—the forced response—divided by the standard deviation of the trends across ensemble members, which quantifies the internal variability. The SNR analysis provides a quantitative comparison of the respective magnitudes of the forced response and the internal variability; this is useful for comparing the model ensemble with the observations in that it can help to identify times and locations in the observations at which one might best be able to see a forced response that is distinguishable from the natural variability.

Significance of observed trends. We construct the 95% confidence intervals for the observed trends as in Santer et al.[32]. The adjusted standard error, s'_b, of the trend, b, is used to account for autocorrelation in the time series, and is defined as

$$s'_b = \frac{s'_e}{\sqrt{\sum_{t=1}^{n_t}(t-\bar{t})^2}}, \quad (9)$$

where s'_e is the adjusted standard deviation of the residuals about the regression line, and n_t is the sample size. The adjusted standard deviation, s'_e, is defined as

$$s'_e = \sqrt{\frac{1}{n_e-2}\sum_{t=1}^{n_t}(R(t))^2}, \quad (10)$$

where $R(t)$ are the residuals in equation (1) and

$$n_e = n_t\frac{1-r_1}{1+r_1} \quad (11)$$

is the effective sample size for r_1, the lag-1 autocorrelation coefficient of $R(t)$. The adjusted 95% confidence interval is $b \pm t_{\mathrm{inv}}s'_b$, where t_{inv} is obtained by inverting the Student's t-distribution for n_e degrees of freedom and $P=0.975$ (two-tailed test). For the model ensembles, we test whether the mean of ten individual trends lies outside the 2.5–97.5% confidence intervals for the respective periods. For the single realization of the reanalysis, we assess whether the trend differences are statistically significant; for trend differences, the adjusted 95% confidence interval is given by[54]

$$b_1 - b_2 \pm t_{\mathrm{inv}}\sqrt{s'^2_{b_1} + s'^2_{b_2}} \quad (12)$$

Influence of increasing GHG concentrations on recovery-era trends. Extended Data Fig. 2 shows the Z trends for the WACCM full-forcing ensemble mean, the GHG-only ensemble mean and the difference, which can be interpreted as the ODS-forced response. The GHG response (Extended Data Fig. 2b) is a small negative trend that partially cancels the ODS-forced trend (Extended Data Fig. 2c). Extended Data Fig. 3 shows the ODS-forced trends in CL_y, O_3, T and U, which show much more symmetric trends between the depletion and recovery eras.

Data availability

TOMS/OMI ozone data are available from https://ozoneaq.gsfc.nasa.gov/. ERA5 data are available from https://cds.climate.copernicus.eu/cdsapp#!/search?text=ERA5. MERRA2 data are available from https://disc.gsfc.nasa.gov/datasets?keywords=%22MERRA-2%22&page=1&source=Models%2FAnalyses%20MERRA-2. JRA55 data are available from https://rda.ucar.edu/datasets/ds628.1/. Model output from the ten-member ensembles used in the analysis presented here is available at https://doi.org/10.7910/DVN/YGWVSB. WACCM–CCMI and WACCM–CMIP6 model output are available from https://www.earthsystemgrid.org.

Code availability

Computer code is available from the corresponding author upon reasonable request.

References

41. Simmons, A. et al. *Global Stratospheric Temperature Bias and Other Stratospheric Aspects of ERA5 and ERA5.1* Technical Memorandum No. 859 (ECMWF, 2020); https://ecmwf.int/node/19362
42. Hurrell, J. W. et al. The Community Earth System Model: a framework for collaborative research. *Bull. Am. Meteorol. Soc.* **94**, 1339–1360 (2013).
43. Kinnison, D. E. et al. Sensitivity of chemical tracers to meteorological parameters in the MOZART3 chemical transport model. *J. Geophys. Res.* **112**, D20302 (2007).
44. Tilmes, S. et al. Representation of the Community Earth System Model (CESM1) CAM4-Chem within the chemistry–climate model initiative. *Geosci. Model Dev.* **9**, 1853–1890 (2016).
45. Matthes, K. et al. Role of the QBO in modulating the influence of the 11 year solar cycle on the atmosphere using constant forcings. *J. Geophys. Res.* **115**, D18110 (2010).
46. Mills, M. J. et al. Global volcanic aerosol properties derived from emissions, 1990–2014, using CESM1(WACCM). *J. Geophys. Res. Atmos.* **121**, 2332–2348 (2016).
47. Neely, R.R. III & Schmidt, A. *VolcanEESM: Global Volcanic Sulphur Dioxide (SO_2) Emissions Database from 1850 to Present—Version 1.0* (Centre for Environmental Data Analysis, 2016); https://doi.org/10.5285/76ebdc0b-0eed-4f70-b89e-55e606bcd568
48. Fujino, J., Nair, R., Kainuma, M., Masui, T. & Matsuoka, Y. Multi-gas mitigation analysis on stabilization scenarios using AIM global model. *Energy J.* https://doi.org/10.5547/ISSN0195-6574-EJ-VolSI2006-NoSI3-17 (2006).
49. Hijioka, Y., Matsuoka, Y., Nishimoto, H., Masui, M. & Kainuma, M. Global GHG emissions scenarios under GHG concentration stabilization targets. *J. Glob. Environ. Eng.* **13**, 97–108 (2008).
50. Eyring, V. et al. Overview of IGAC/SPARC Chemistry–Climate model initiative (CCMI) community simulations in support of upcoming ozone and climate assessments. *SPARC Newsl.* **40**, 48–66 (2013).
51. Deser, C. et al. Insights from Earth system model initial-condition large ensembles and future prospects. *Nat. Clim. Change* **10**, 277–286 (2020).
52. Gettelman, A. et al. The Whole Atmosphere Community Climate Model version 6 (WACCM6). *J. Geophys. Res. Atmos.* **124**, 12380–12403 (2019).
53. Deser, C., Philips, A. S., Alexander, M. A. & Smoliak, B. V. Projecting North American climate over the next 50 years: uncertainty due to internal variability. *J. Clim.* **27**, 2271–2296 (2014).
54. Andrade, J. M. & Estévez–Pérez, M. G. Statistical comparison of the slopes of two regression lines: a tutorial. *Anal. Chim. Acta* **838**, 1–12 (2014).

Acknowledgements

Helpful discussion and comments by D. Kim are gratefully acknowledged. B.Z. and S.S. were supported by grants from the National Science Foundation NSF 1539972 and 1848863, and a gift to MIT from an anonymous donor. D.W.J.T. is supported by NSF AGS-1848785. Q.F. is supported by NSF AGS-1821437. We would like to acknowledge high-performance computing support from Cheyenne (https://doi.org/10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

Author contributions

B.Z. and S.S. designed the study. B.Z., S.S., D.W.J.T. and Q.F. analysed and interpreted the results. B.Z. led the writing, and all authors contributed to the editing of the manuscript and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

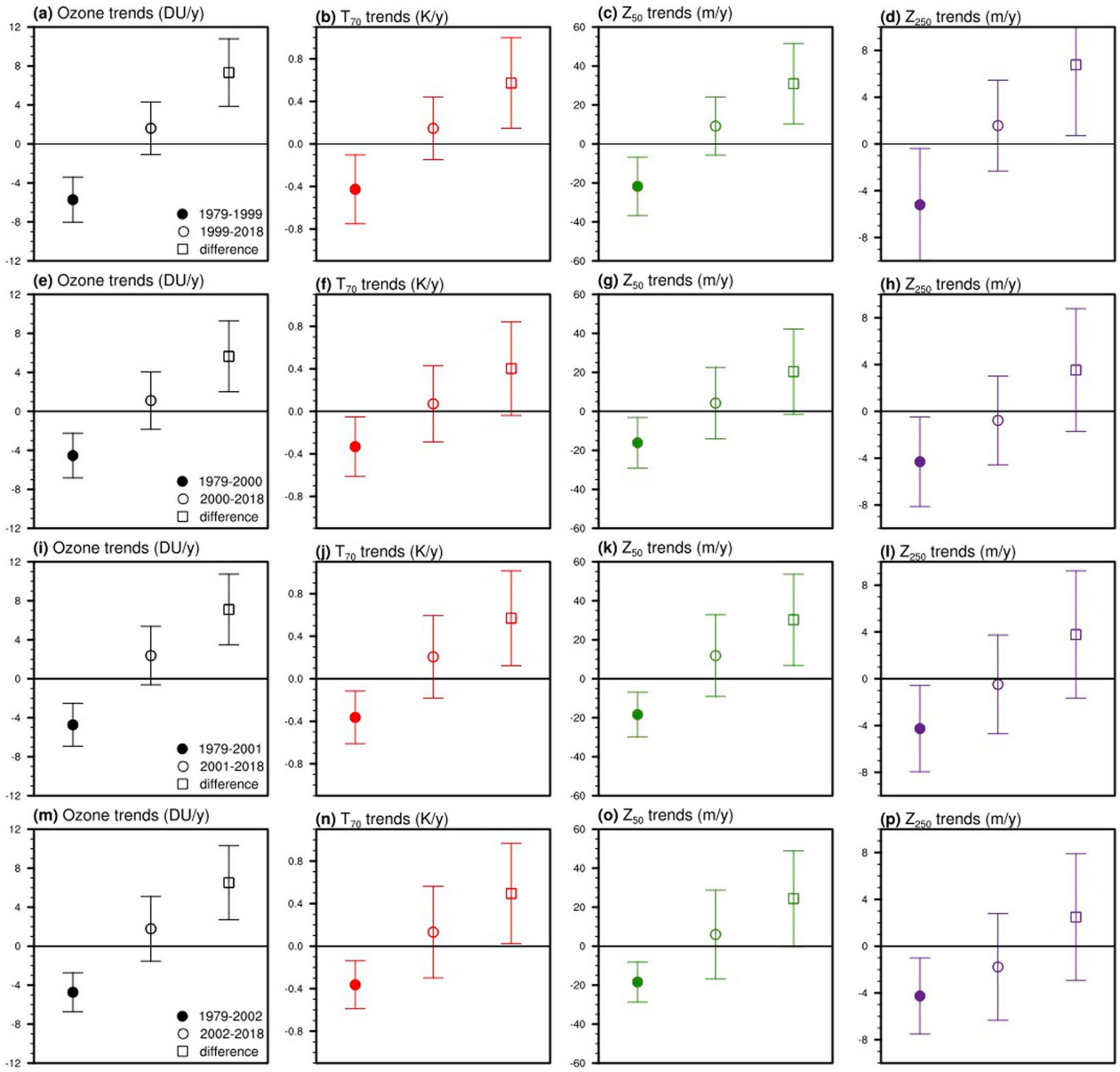
Extended data is available for this paper at https://doi.org/10.1038/s41561-021-00803-3.

Supplementary information is available for this paper at https://doi.org/10.1038/s41561-021-00803-3.

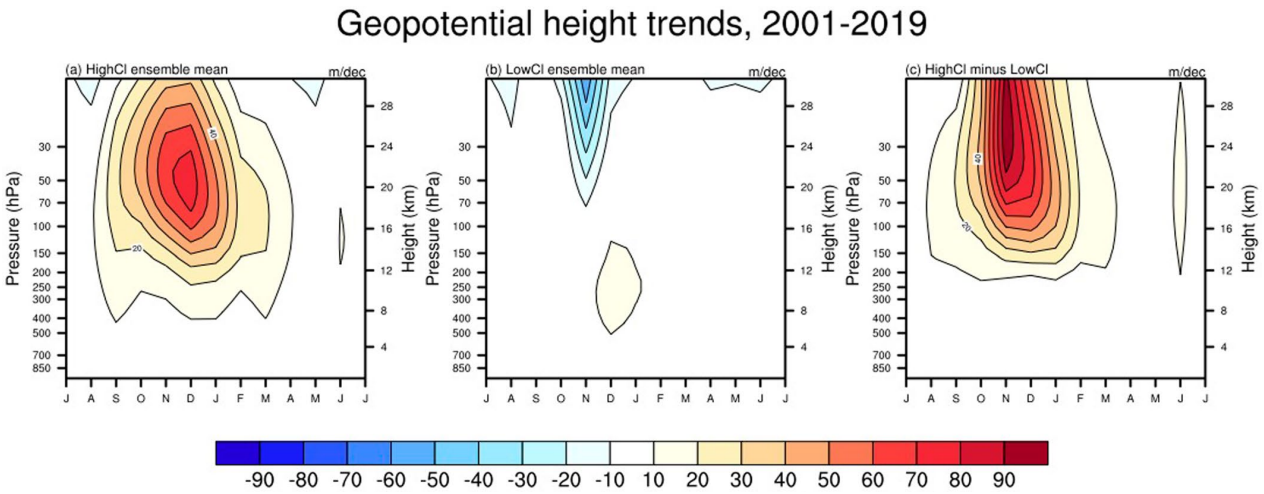
Correspondence and requests for materials should be addressed to B.Z.

Peer review information *Nature Geoscience* thanks Sabine Bischof, Olaf Morgenstern, Yousuke Yamashita and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Clare Davis, Tom Richardson.

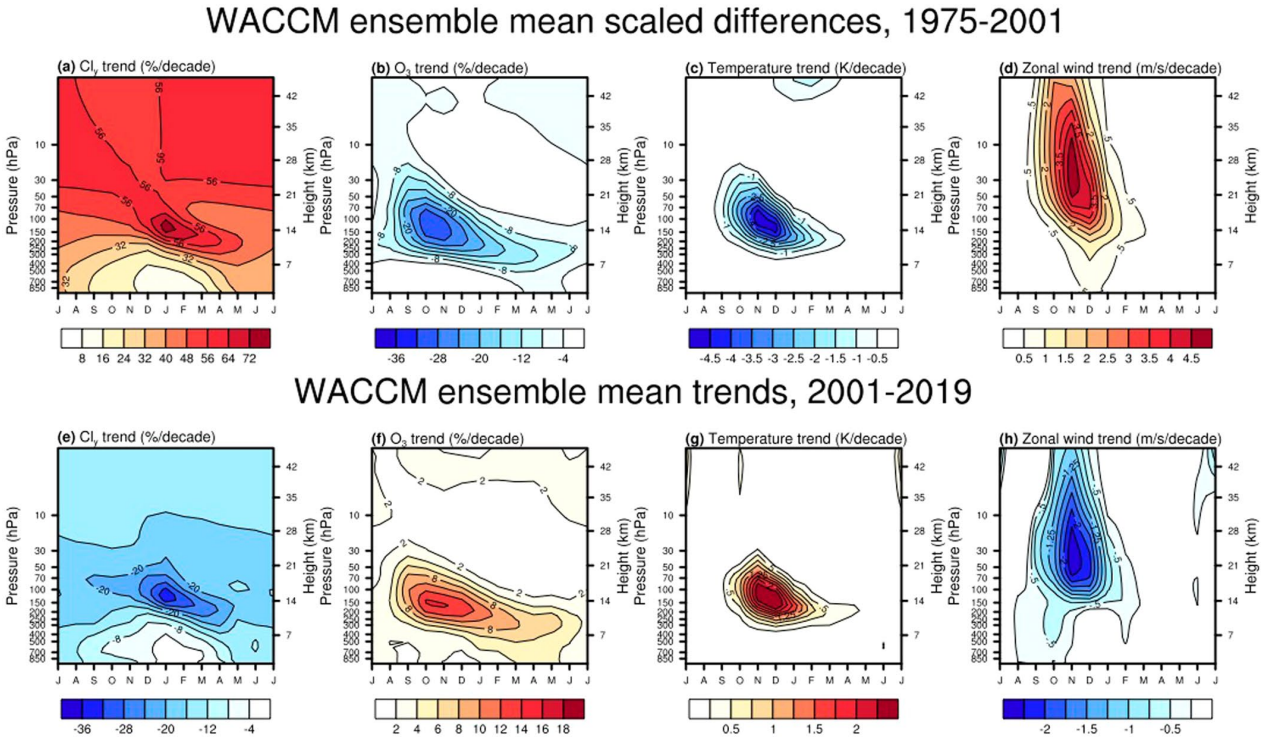
Reprints and permissions information is available at www.nature.com/reprints.



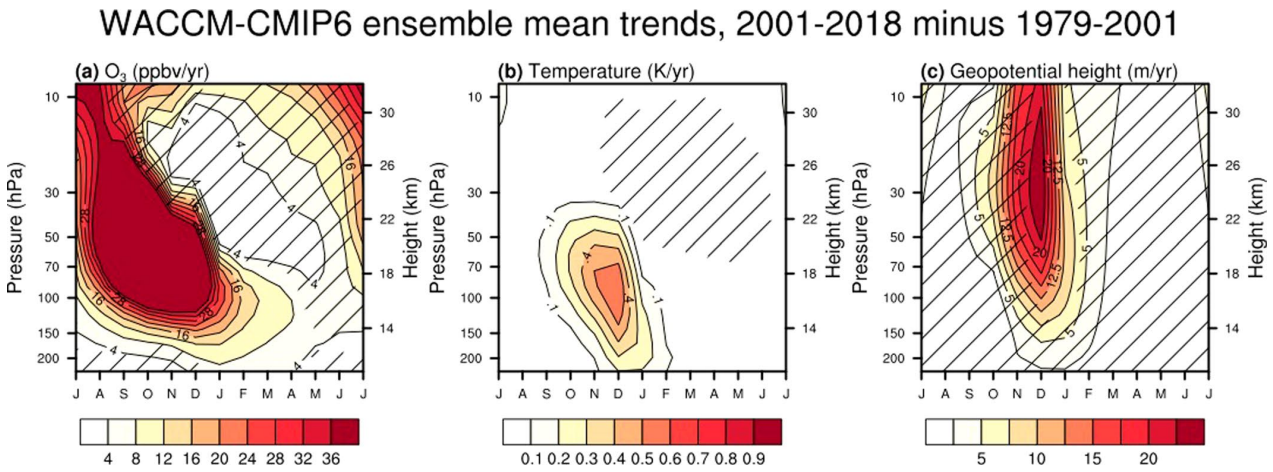
Extended Data Fig. 1 | Trends and trend differences for different turnaround years. Ozone and circulation trends for ozone depletion era (filled circles), ozone recovery era (open circles), and their differences (squares) for turnaround year defined as (**a–d**) 1999, (**e–h**) 2000, (**i–l**) 2001, and (m–p) 2002. Vertical lines indicate the 95% confidence intervals on the trends.



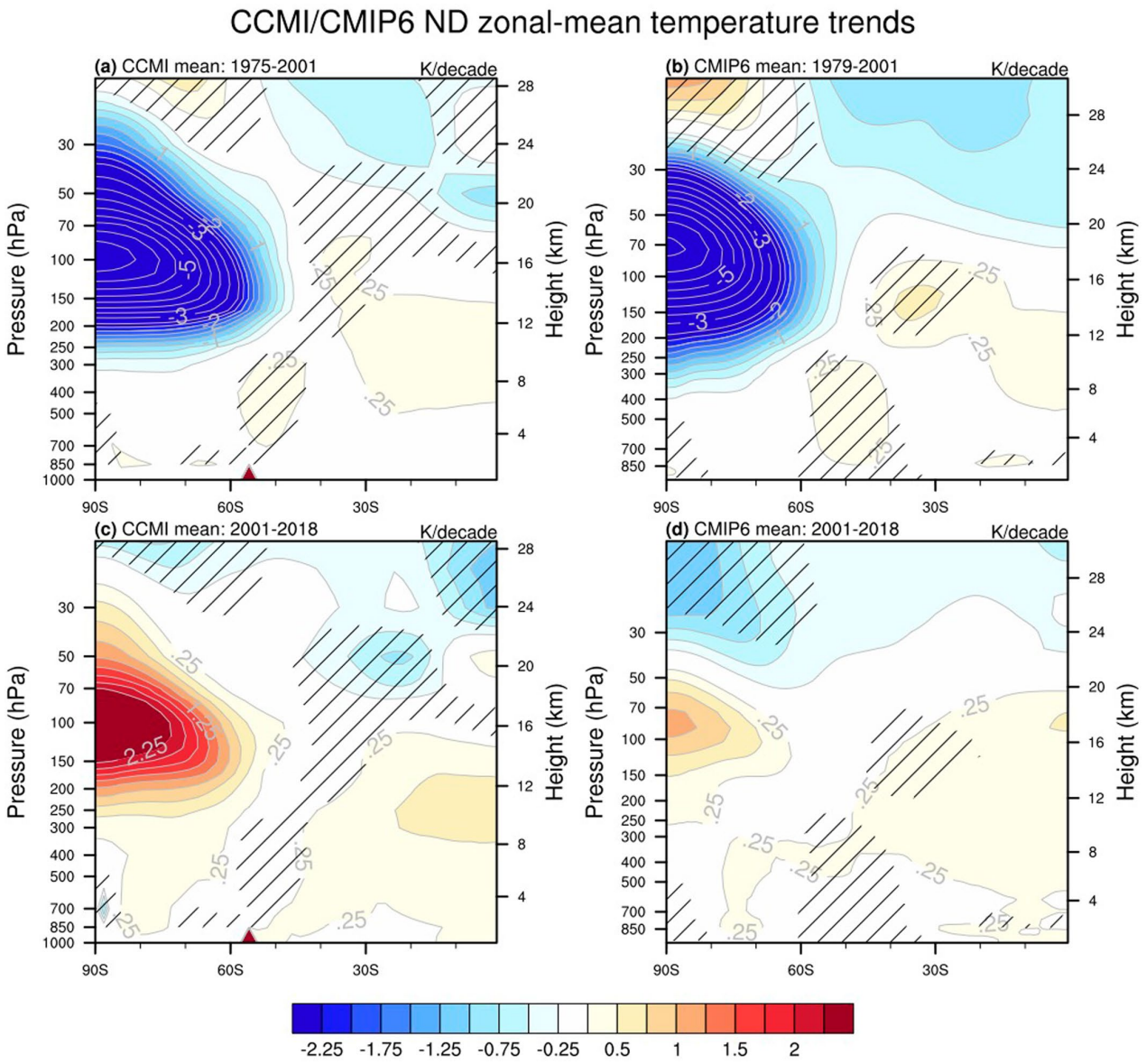
Extended Data Fig. 2 | GHG contribution to WACCM geopotential height trends. WACCM ensemble mean geopotential height trends (m/decade) for 2001–2018 for (a) ODS+GHG, (b) GHG-only, and (c) the difference (approximately the ODS-only response).



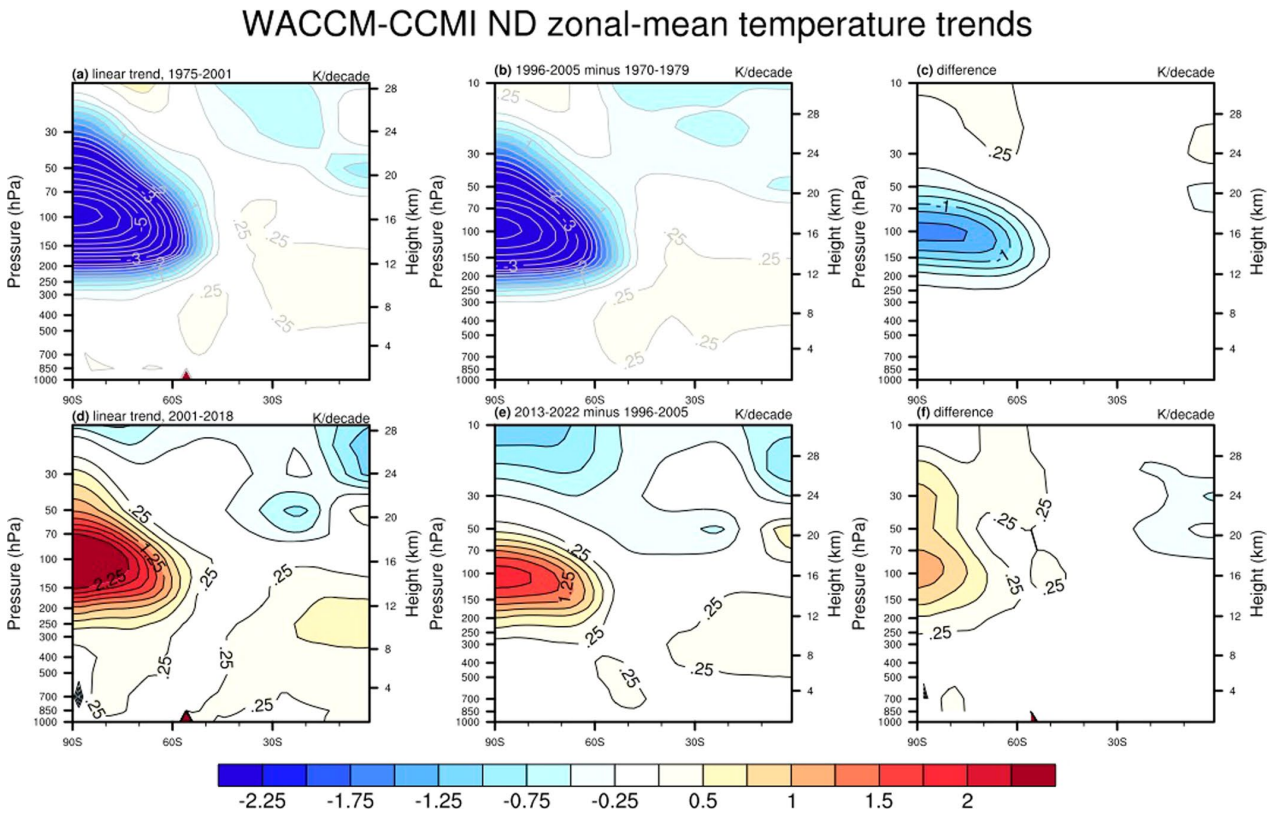
Extended Data Fig. 3 | ODS-forced trends. As Fig. 6, but for ODS+GHG minus GHG-only.



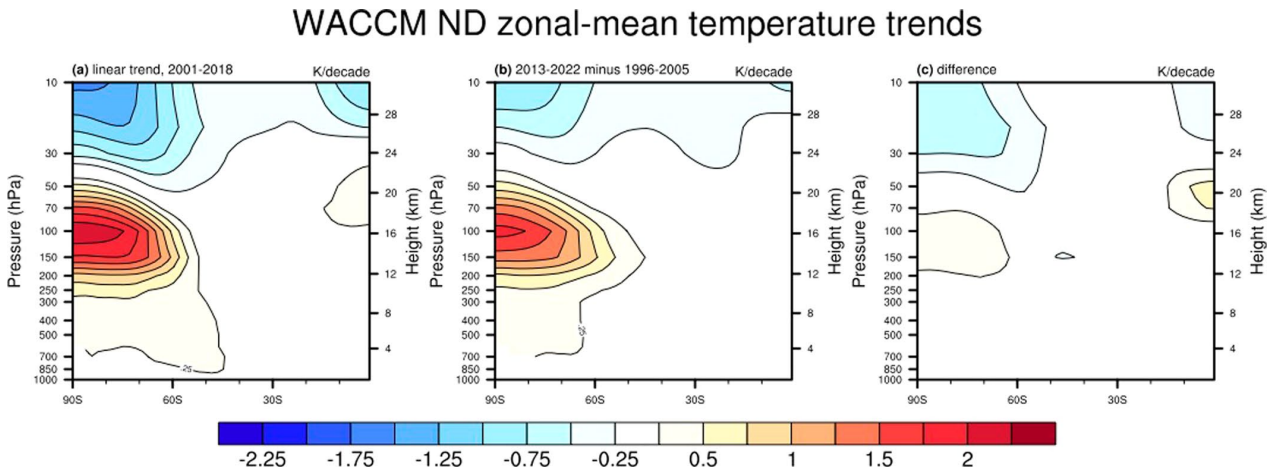
Extended Data Fig. 4. | WACCM-CMIP6 trend differences. WACCM-CMIP6 November-December Southern Hemisphere trend differences between 1979–2001 and 2001-2018 in ozone, temperature, and geopotential height. Hatching indicates regions where the trend differences are not significantly different from the distributions of trend differences in the control run ($p > 0.05$; Methods).



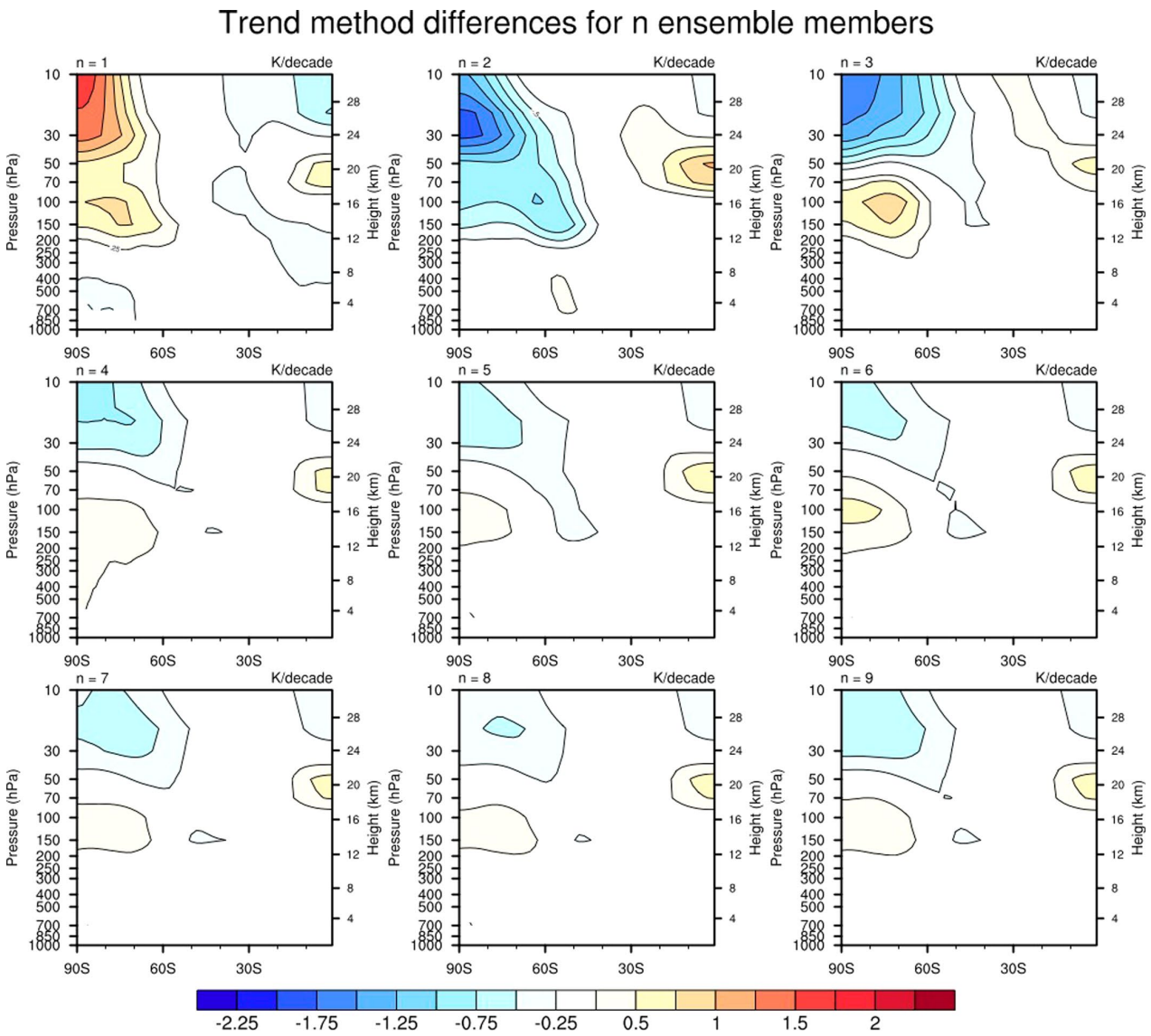
Extended Data Fig. 5 | WACCM CCMI and CMIP6 temperature trends. SH ND WACCM ensemble mean zonal-mean temperature trends for CCMI (**a, c**) and CMIP6 (**b, d**) for the ozone depletion (**a, b**) and recovery (**c, d**) periods. Hatching indicates regions where the trend differences are not significantly different from the distributions of trend differences in the control runs ($p > 0.05$; Methods).



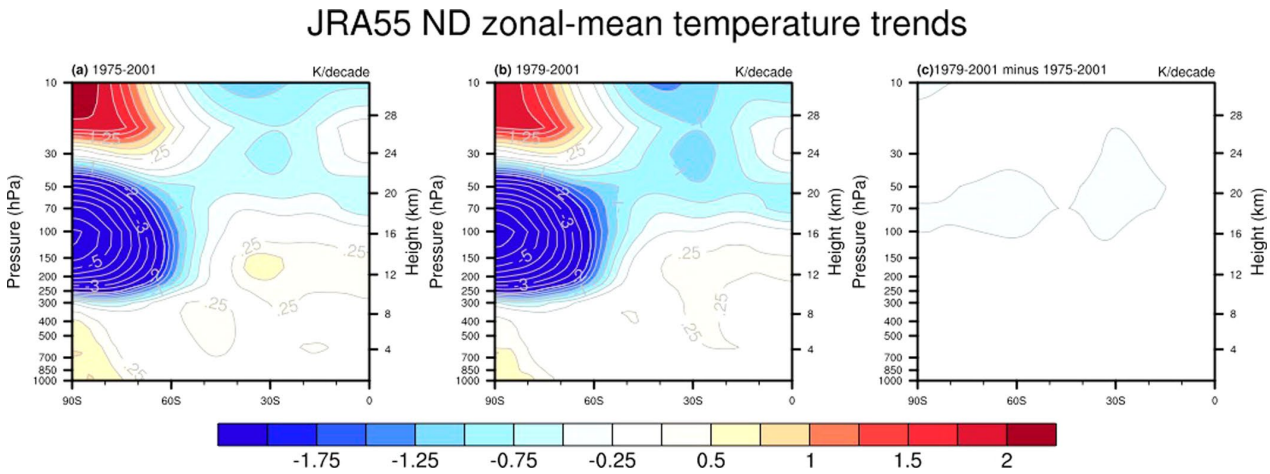
Extended Data Fig. 6 | WACCM-CCMI temperature trends. SH ND zonal-mean temperature trends for the WACCM-CCMI ensemble mean calculated using (**a,d**) linear regression, equation 1, (**b,e**) equation 2 using scaled differences over 10 year periods, and (**c,f**) the difference for the periods (**a–c**) 1975–2001 and (**d–f**) 2001–2018 using the two methods.



Extended Data Fig. 7 | WACCM temperature trends. SH ND zonal-mean temperature trends for the WACCM ensemble mean for the period 2001–2018 calculated using (**a**) linear regression (equation 1), (**b**) scaled differences (equation 2), and (**c**) the difference.



Extended Data Fig. 8 | Trend differences arising from trend calculation methods: the role of ensemble size. SH ND zonal-mean temperature trend shown in each panel is the difference between using the linear trend (for example, as in Extended Data Fig. 6a) and differencing the climatologies (for example, as in Extended Data Fig. 6b) for the average of $1 \leq n \leq 9$ ensemble members (the difference for $n = 10$ is shown in Extended Data Fig. 6c).



Extended Data Fig. 9 | JRA55 temperature trends. SH ND zonal-mean temperature linear trends for JRA55 for (**a**) 1975–2001, (**b**) 1979–2001, and (**c**) the difference.