Comparison of Arctic and Antarctic stratospheric dynamics in chemistry versus no-chemistry climate models

Olaf Morgenstern[1], Douglas E. Kinnison[2], Michael Mills[2], Martine Michou[3], Larry W. Horowitz[4], Pu Lin[5], Makoto Deushi[6], Kohei Yoshida[6], Fiona M. O'Connor[7], Yongming Tang[7], Nathan Luke Abraham[8,9], James Keeble[8,9], Fraser Dennison[10], Eugene Rozanov[11,12], Tatiana Egorova[11], Timofei Sukhodolov[11,13], Guang Zeng[1]

[1]National Institute of Water and Atmospheric Research, Wellington, New Zealand
[2]National Center for Atmospheric Research, Boulder, CO, USA
[3]Centre National des Recherches Météorologiques (CNRM), MétéoFrance, Toulouse, France
[4]Geophysical Fluid Dynamics Laboratory (GFDL), National Oceanic and Atmospheric Administration, Princeton, NJ, USA
[5]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA
[6]Meteorological Research Institute (MRI), Japan Meteorological Agency, Tsukuba, Japan
[7]Hadley Centre, MetOffice, Exeter, UK
[8]Dept. of Chemistry, University of Cambridge, UK
[9]National Centre for Atmospheric Science, University of Cambridge, UK
[10]Commonwealth Scientific and Industrial Research Organization, Aspendale, Victoria, Australia
[11]Physikalisch-Meteorologisches Observatorium, World Radiation Centre, Davos, Switzerland
[12]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[13]Institute of Meteorology and Climatology, University of Natural Resources and Life Sciences, Vienna, Austria

Key Points:

- Coupling in ozone chemistry causes an increase in persistence of low temperature anomalies over both poles.
- In the Antarctic, coupling in chemistry amplifies pre-existing stratospheric cold biases.
- These effects can be masked by other dynamical differences present in some models.

Corresponding author: Olaf Morgenstern, olaf.morgenstern@niwa.co.nz

Abstract

Using nine chemistry-climate and eight associated no-chemistry models, we investigate the persistence and timing of cold episodes occurring in the Arctic and Antarctic stratosphere during the period 1980-2014. We find systematic differences in behavior between members of these model pairs. In a first group of chemistry models whose dynamical configurations mirror their no-chemistry counterparts, we find an increased persistence of such cold polar vortices, such that these cold episodes both start earlier and last longer, relative to the times of occurrence of the lowest temperatures. Also the date of occurrence of the lowest temperatures, both in the Arctic and the Antarctic, is delayed by 1-3 weeks in chemistry models, versus their no-chemistry counterparts. This behavior exacerbates a widespread problem occurring in most or all models, a delayed occurrence, in the median, of the most anomalously cold day during such cold winters. In a second group of model pairs there are differences beyond just ozone chemistry. In particular, here the chemistry models feature more levels in the stratosphere, a raised model top, and differences in non-orographic gravity wave drag versus their no-chemistry counterparts. Such additional dynamical differences can completely mask the above influence of ozone chemistry. The results point towards a need to retune chemistry-climate models versus their no-chemistry counterparts.

Plain Language Summary

Ozone is a chemical constituent of the atmosphere acting as an absorber of both solar ultraviolet light and infrared radiation emitted by the Earth. It therefore needs to be considered in climate models. Explicit ozone chemistry is a computationally challenging addition to a climate model; hence in most cases ozone is simply prescribed. Especially during relatively cold stratospheric winter/spring seasons, Antarctic and Arctic ozone depletion can be considerable. Such anomalous ozone loss is not reflected in the imposed ozone field, and hence differences in behavior are expected for such situations between chemistry- and no-chemistry models. Indeed for such cold winters/springs, we find an enhanced persistence of such cold spells in a set of chemistry-climate models, versus their no-chemistry counterparts; such enhanced persistence generally makes the chemistry model less realistic than its no-chemistry counterpart. However, if there are substantial further differences between the members of these model pairs, such as regarding their grid configuration or physical processes beyond chemistry, these can more than compensate for the effect of ozone chemistry. We thus claim that adding stratospheric ozone chemistry to a climate model necessitates retuning to counteract a deterioration of dynamics that can otherwise occur.

1 Introduction

Climate feedbacks involving ozone have long been known to be important in large-scale climate change. Most notably, stratospheric ozone depletion has been linked to a strengthening of the Southern Annular Mode (SAM) since roughly the 1970s (Son et al., 2010; Fogt & Marshall, 2020, and references therein). Ozone depletion of the Antarctic polar vortex in spring drives a cooling of this airmass, stabilizing the vortex, delaying the transition to summertime easterlies, and via deep coupling causing a strengthening of the Southern Annular Mode (SAM) during southern summer (Thompson et al., 2011; Morgenstern, 2021). In the Arctic, ozone depletion is usually less pronounced than in the Antarctic (although recent years have seen two Arctic "ozone holes"; Kuttippurath et al., 2021), residual ozone is larger, and consequently ozone depletion has not been implicated in a long-term strengthening of the Northern Annular Mode (NAM; Eyring et al., 2021). However, large ozone depletion does tend to be followed by anomalous tropospheric weather, i.e. an anomalously strong NAM (Ivy et al., 2017; Friedel et al., 2022).

The pertinent observed long-term strengthening of the NAM however remains unexplained (Eyring et al., 2021).

Climate models regularly simulate a delayed breakdown of the polar vortex. This behavior leads to too-strong stratospheric cooling following ozone depletion, driven by biases in the dynamical responses to ozone depletion (Lin et al., 2017). Also in some individual-model studies, ozone chemistry has been found to impact timescales of variability of the polar vortices (Haase & Matthes, 2019; Rieder et al., 2019; Oehrlein et al., 2020). We will investigate whether these findings apply to present-generation climate models as a group, and any learnings as these models transition from almost all excluding to in the future increasing including explicit ozone chemistry. At the time of writing, the portal of the 6$^{\text{th}}$ Coupled Model Intercomparison Project (CMIP6) lists 114 models and model variants. Morgenstern (2021) uses 29 different models in his assessment of the SAM in CMIP6, essentially sidelining many model variants to reduce redundancy. Of these 29 models, only six have explicit interactive ozone chemistry. A feature of the CMIP6 dataset is that pairs of models have participated with interactive ozone chemistry constituting the main or only point of difference between them. Simulations performed by these model pairs thus offer an opportunity to assess what the impact is of interactive chemistry versus the alternative approaches, i.e. usually prescribing the pre-computed CMIP6 ozone climatology (Checa-Garcia et al., 2018). A comparison of such model pairs will of course not only find impacts due to interactive ozone – or lack thereof – but would also be sensitive to any peculiarities of the precomputed ozone field itself, its implementation (Hardiman et al., 2019), and any differences versus the interactive ozone. For example, Morgenstern et al. (2020, 2021) have shown that the recommended CMIP6 ozone climatology (Checa-Garcia et al., 2018) greatly underestimates Northern-Hemisphere mean ozone loss over the period 1979-2000. Also in a few cases there are other differences between these pairs that extend beyond ozone chemistry, which can complicate this comparison. A recent study (Lin & Ming, 2021) finds substantially enhanced cooling in a model variant with interactive ozone versus the same model using prescribed ozone, even though the simulated and prescribed ozone are quite similar. The authors explain this as the effect of co-variance of ozone and temperature anomalies that does not exist in the no-chemistry model.

In the below we will compare simulations of pairs of CMIP6 models (supplemented with three non-CMIP6 models) with and without interactive ozone, and will assess differences between the two members of the pair regarding polar stratospheric dynamics and associated stratosphere-troposphere coupling. Where significant, such differences will be indicative of the role of climate-ozone coupling. We will assess both hemispheres, noting that Morgenstern (2021) has already made the case, using CMIP6 simulations, for why interactive ozone is important for simulating climate trends of the Southern Hemisphere. Here we will complement his analysis with a focus on timescales of variability and on anomalously cold stratospheric winters when polar ozone chemistry is particularly impactful.

2 Models and observational reference data

Models used here are listed in table 1. We use all chemistry-climate models from CMIP6 for which daily- and zonal-mean temperature and geopotential height (GPH) fields are available for "historical" simulations, and their no-chemistry CMIP6 equivalents where such models exist. Furthermore we use the SOCOL (Sukhodolov et al., 2021), ACCESS-CM2-Chem (Dennison & Woodhouse, 2022), and UKESM1-StratTrop models from the Chemistry-Climate Model Initiative Phase 2 (CCMI2) set of models (Plummer et al., 2021), and their no-chemistry CMIP6 equivalents. UKESM1-StratTrop is a further development of the UKESM1-0-LL model (Sellar et al., 2019), based on the same no-chemistry background model (Williams et al., 2018; Kuhlbrodt et al., 2018, HadGEM3-GC31-LL,) but with some updates to photolysis and other reaction rates which reduce a general over-

CCMs		No-chemistry models		Differences	References
CESM2-WACCM	3	CESM2	11	higher top, NOGWD	G19, DA20
CESM2-WACCM-FV2	3	CESM2-FV2	3	higher top, NOGWD	G19, DA20
CNRM-ESM2-1	9	CNRM-CM6-1	28	same settings	S19, V19
GFDL-ESM4	3	GFDL-CM4	1	higher top, NOGWD	D20, H19
MRI-ESM2-0	5				Y19
UKESM1-0-LL	13	HadGEM3-GC31-LL	3	same settings	SE19, K18, W18
UKESM1-StratTrop	*3*	*HadGEM3-GC31-LL*	*5*	same settings	SE19, K18, W18
ACCESS-CM2-Chem	*3*	*ACCESS-CM2*	*3*	same settings	D22, B20, BO20
SOCOL4	*3*	*MPI-ESM1-2-LR*	*3*	same settings	S21, M19

Table 1. CMIP6/CCMI2 chemistry and corresponding CMIP6 no-chemistry models considered here. The 2nd and 4th columns denote the number of "historical", REF-D1, or AMIP simulations used in the analysis. For the purposes of this paper, models listed in italics are atmosphere-only; we use their CCMI2 REF-D1 and CMIP6 AMIP simulations, respectively. References: B20 = Bi et al. (2020), BO20 = Bodman et al. (2020), D20 = Dunne et al. (2020), D22 = Dennison and Woodhouse (2022), DA20 = Danabasoglu et al. (2020), G19 = Gettelman et al. (2019), H19 = Held et al. (2019), K18 = Kuhlbrodt et al. (2018), M19 = Mauritsen et al. (2019), S19 = Séférian et al. (2019), S21 = Sukhodolov et al. (2021), SE19 = Sellar et al. (2019), V19 = Voldoire et al. (2019), W18= Williams et al. (2018), Y19 = Yukimoto, Kawai, et al. (2019).

estimation of ozone in the extrapolar stratosphere. (Other CCMI2 models are not used here because they do not have a no-chemistry equivalent in the CMIP6 group of models.) References in table 1 are for the chemistry models (Morgenstern, 2021). In the CCMI2 "REF-D1" simulations used here the three CCMI2 models are not coupled to an interactive ocean; rather they use prescribe observational (HadISST) sea-surface conditions (Rayner et al., 2003). The simulations are therefore more comparable to the Atmosphere Model Intercomparison Project (AMIP) simulations of CMIP6 (although these use a different observational climatology for sea surface conditions than the REF-D1 simulations; Taylor et al., 2015). CESM2-FV2 and CESM2-WACCM-FV2 are identical to CESM2 and CESM2-WACCM but with the atmospheric resolution degraded from about $\sim 1^\circ$ to $\sim 2^\circ$. ACCESS-CM2-Chem and ACCESS-CM2 share an atmosphere model with UKESM1-0-LL and HadGEM3-GC31-LL but use different land models.

Previous evaluations have shown that the UKESM1-0-LL and CNRM-ESM2-1 models well simulate 1979-2000 Arctic ozone trends, GFDL-ESM4, CESM2-WACCM, and MRI-ESM2-0 underestimate Arctic ozone depletion (Morgenstern et al., 2020), and SOCOL quite faithfully reproduces extrapolar ozone (Sukhodolov et al., 2021). The ozone field used to drive the no-chemistry models HadGEM3-GC31-LL, MPI-ESM1-2-LR, GFDL-CM4, CESM1, CESM2-FV2, and ACCESS-CM2 however much underestimates these Northern-Hemisphere and especially Arctic ozone trends (Morgenstern et al., 2020), with hemispheric- and annual-mean TCO trends for 1979-2000 in the CMIP6 climatology (Checa-Garcia et al., 2018) only reaching approximately a third of observed trends (Morgenstern, 2021). In the Antarctic, UKESM1-0-LL and MRI-ESM2-0 under- and overestimate, respectively, Antarctic ozone during spring, whereas the other CMIP6 chemistry-climate models simulate more realistic Antarctic ozone depletion (Morgenstern et al., 2020).

The results will be compared to version 2 of the National Center for Environmental Prediction (NCEP)/Department of Energy (DOE) / NCEP-DOE2 reanalysis (Kanamitsu et al., 2002) and the Multi-Sensor Reanalysis 2 total-column ozone climatology (van der A et al., 2015a). Replacing NCEP-DOE2 with ERA5 (Hersbach et al., 2020) leads to no discernible difference in figure 3, suggesting that model biases and shortcomings are much bigger factors in our analysis than any observational uncertainty.

3 Method

In a seminal paper Baldwin and Dunkerton (2001) showed how stratospheric circulation anomalies in the Arctic propagate to low altitudes and affect tropospheric circulation for the approximately two months that such features may last. For example, impacts include anomalous states of the NAM, the positions of the northern storm tracks, and mid-latitude storms. Equivalent influences of the stratosphere on the weather of the Southern Hemisphere have also been demonstrated (Thompson et al., 2005). Baldwin and Dunkerton (2001)'s method also lends itself to a comparison of chemistry versus equivalent no-chemistry models presented here. While Baldwin and Dunkerton (2001) present a composite of a stratospheric NAM index for composites of several strong and weak-NAM events, here we modify their method to using polar-cap mean stratospheric temperature as our key metric. The reason for this is that (a) this diagnostic is available for both chemistry- and no-chemistry models, unlike e. g. ozone, and (b) wintertime low temperatures are associated with heterogeneous chlorine activation on polar stratospheric clouds followed by ozone depletion in models with interactive chemistry. Much of the rest of our analysis is inspired by Baldwin and Dunkerton (2001), namely:

1. We use available "historical", REF-D1, or AMIP daily- and zonal-mean temperature and GPH fields on pressure levels for 1980-2014, for the models listed in table 1.
2. We calculate the polar-cap (75°N-90°N and 90°S-75°S, respectively) average temperatures and GPH fields.
3. We smoothen both fields using 15-day boxcar filters, to reduce the impact of outliers, and subtract off the mean annual cycles of polar-cap temperature and GPH, creating temperature and GPH anomaly timeseries.
4. We determine, for every year starting on 1 September (for the Arctic) and 1 May (for the Antarctic) and for every ensemble member, the lowest value at 70 hPa of the polar-cap average temperature anomaly, and the day of its occurrence. This temperature is then used to rank the years by stratospheric temperature.
5. We form the averages for all 231-day periods, from 130 days before the coldest day to 100 days after the coldest day, for the 20% coldest years.

4 Results

4.1 General model performance for monthly-mean ozone and temperature

Arctic total-column ozone, in the decades before ∼1995, experienced a decline of nearly 100 DU in March since 1979 but recovered slightly thereafter, see the Multi-Sensor Reanalysis 2 (MSR-2; van der A et al., 2015a, 2015b) panel of figure 1. Losses in other seasons were much smaller. The loss was mainly driven by increasing halogens in a well-understood mechanism involving chlorine activation on polar stratospheric clouds (WMO, 2018). In the nine chemistry-climate models and the CMIP6 climatology (itself derived from model results, Checa-Garcia et al., 2018), this springtime loss is captured but with varying degrees of realism. March trends come close to MSR-2 in UKESM1-0-LL, UKESM1-StratTrop, and ACCESS-CM2-Chem, but in these models, unrealistically, the ozone loss is bigger in April than in March. SOCOL and CNRM-ESM2-1 also both simulate substantial though underestimated ozone loss. CESM2-WACCM, CESM2-WACCM-FV2, GFDL-ESM4, and MRI-ESM2-0 all substantially underestimate the amount of ozone loss, as does the CMIP6 ozone climatology used to force no-chemistry CMIP6 models. A failure to simulate a realistic impact of halogen increases on Arctic ozone can indicate that chlorine activation in these models is not realistic, for example because of a stratospheric warm bias reducing the occurence of polar stratospheric clouds, or for other reasons.

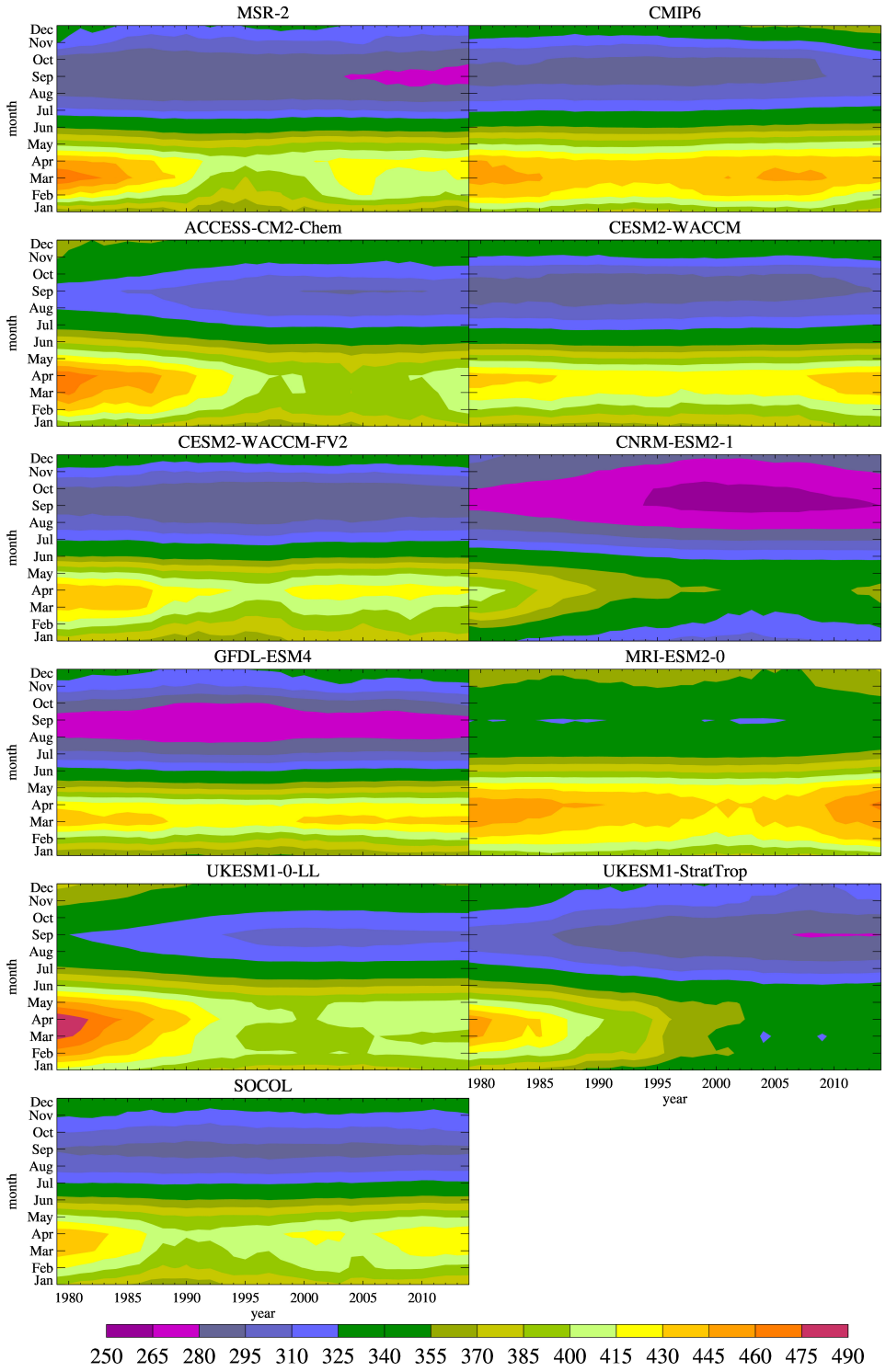

Figure 1. 1979-2014 monthly-mean TCO (DU) averaged over the Arctic polar cap (north of 75°N), expressed as functions of the year and month of the year and smoothed with an 11-year boxcar filter, for the MSR-2 observational reference (van der A et al., 2015a), the CMIP6 ozone forcing dataset (Checa-Garcia et al., 2018), and the single-model ensemble-means of the "historical" and REF-D1 simulations, respectively, by the nine chemistry-climate models.

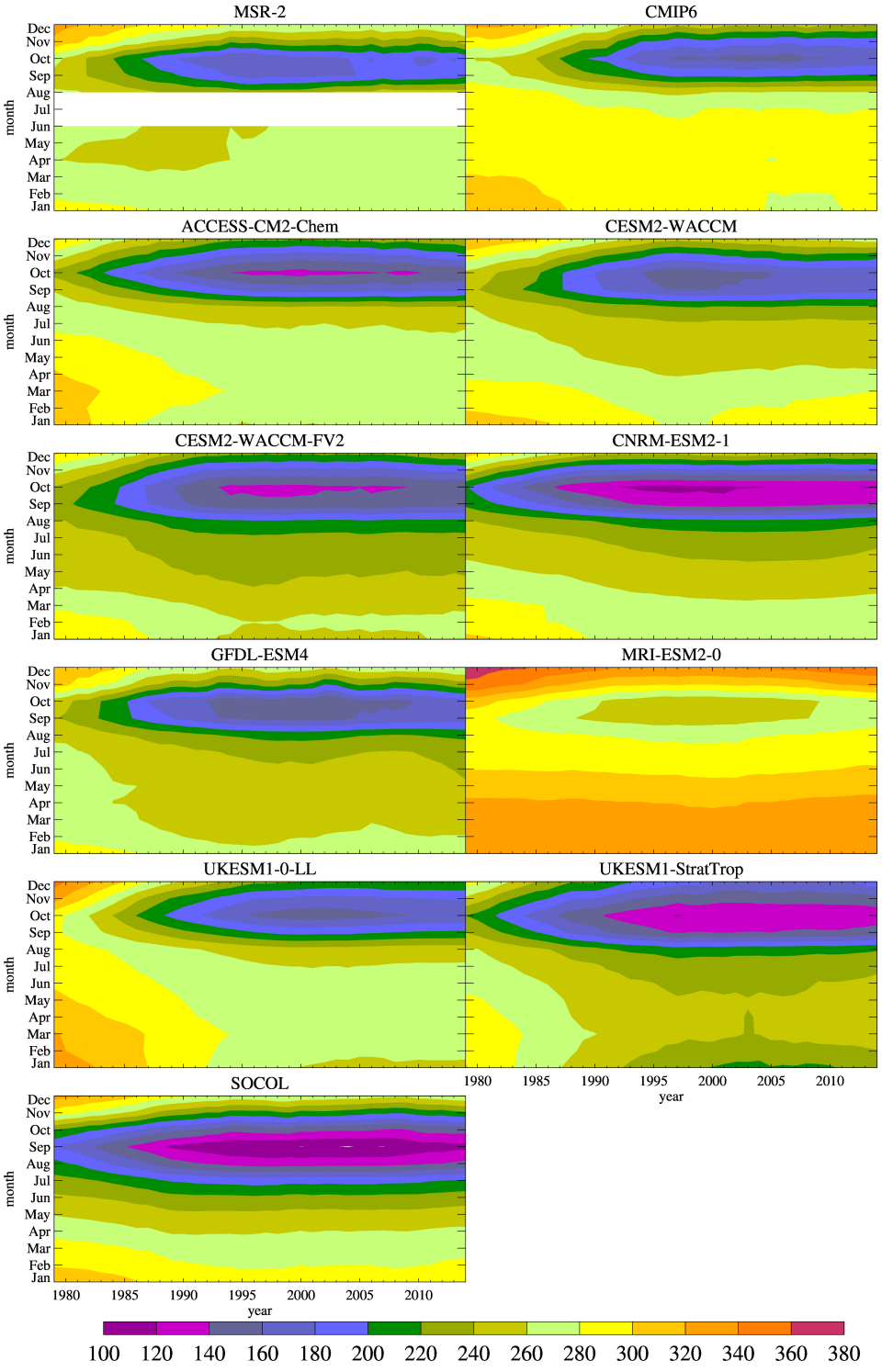

Figure 2. Same as figure 1, but for the Antarctic polar cap (south of 75°S).

Similarly, the Antarctic has experienced substantial ozone loss in spring, manifesting as the Antarctic "ozone hole" (figure 2). The models capture this, but again with various biases. Several models have severe ozone loss persisting for too long into summer (the UKESM1 models, ACCESS-CM2-Chem, the CESM2-WACCM versions, and CNRM-ESM2-1). The MRI-ESM2-0 model substantially underestimates ozone loss. The SOCOL model simulates an early onset of the ozone hole, with lowest polar ozone occurring in September not October. The GFDL-ESM4 model overall has the most realistic timing and small biases of Antarctic ozone – we note however the much underestimated ozone loss in the Arctic in this model.

Next we assess the simulation of temperature in these models.

An inspection of the mean 1980-2014 temperature bias and standard deviation for the 70 hPa polar-cap mean temperature (figure 3)indicates that for both polar regions, there is excellent agreement between the NCEP-DOE2 (Kanamitsu et al., 2002) and the newer ERA5 reanalyses (Hersbach et al., 2020), with essentially identical standard deviations and absolute biases between the two reanalyses of mostly less than 1K, much smaller than typical model biases. We therefore use NCEP-DOE2 in the following for ease of handling. A majority of models (chemistry and no-chemistry alike) exhibits cold biases during spring. In the Antarctic, the cold bias reaches -15 to -20K in November in ACCESS-CM2-Chem, CNRM-ESM2-1, UKESM1-0-LL, and UKESM1-StratTrop. These biases are all worsened versus their no-chemistry counterparts. The cold biases are reflected in an increase in stratospheric variability during December and January, indicating an extension of the lifetime of the Antarctic polar vortex versus their no-chemistry counterparts. The CESM2-WACCM models exhibit largely unchanged biases and variability in the Antarctic versus the no-chemistry equivalents, but a decreased cold bias in the Arctic in the chemistry versions. The GFDL-ESM4 model exhibits only a small warm bias in the Antarctic but a substantial warm bias ($\sim$ 5K) in the Arctic in spring, explaining its extremely small Arctic ozone loss. SOCOL simulates relatively small biases in both polar regions but exaggerated variability in the Antarctic in September and October, reflecting the early onset of ozone depletion in this model noted above.

4.2 Temperature variablity and cold episodes in chemistry- and no-chemistry models

Figures 4 and 5 confirm that practically all variability in polar-cap 70 hPa temperature occurs during the cold season – during summer this variability is no more than a few K but in the daily polar-cap average can reach and exceed $\pm$20 K during winter and spring. For both polar regions there is an asymmetry between cold and warm winters: For warm winters, the anomalies occur nearly symmetrically around the middle of the cold season (in the Arctic, approximately day 30, i.e. 31 January; in the Antarctic, approximately day -40, i.e. 22 November), whereas during extremely cold winters the temperature anomaly builds until the wintertime circulation collapses and temperatures rapidly return to the average, with the largest cold anomalies occurring in spring or even summer. Also models with larger ensembles (e.g. CNRM-CM6-1, MPI-ESM1-2-LR) show that for warm anomalies there is no sharp upper bound for the largest warm anomalies that can occur, whereas the cold anomalies, until well into spring, are sharply bounded by a lower envelope function which decreases during the course of the winter. During spring some rare extremely cold events occur, i.e. long-lasting cold polar vortices, e.g. in the UKESM1, CNRM-ESM2-1, and ACCESS-CM2-Chem models. This assymmetric nature of variability reflects coupling with mid-latitudes, or lack thereof. During warm winters, the Arctic and Antarctic receive their heat from mid-latitudes in dynamical disturbances. This mechanism is different from the radiative cooling that dominates during cold, dynamically relatively unperturbed winter seasons and causes temperatures to gradually drop throughout the season, until the final warming marks the end of the polar vortex.

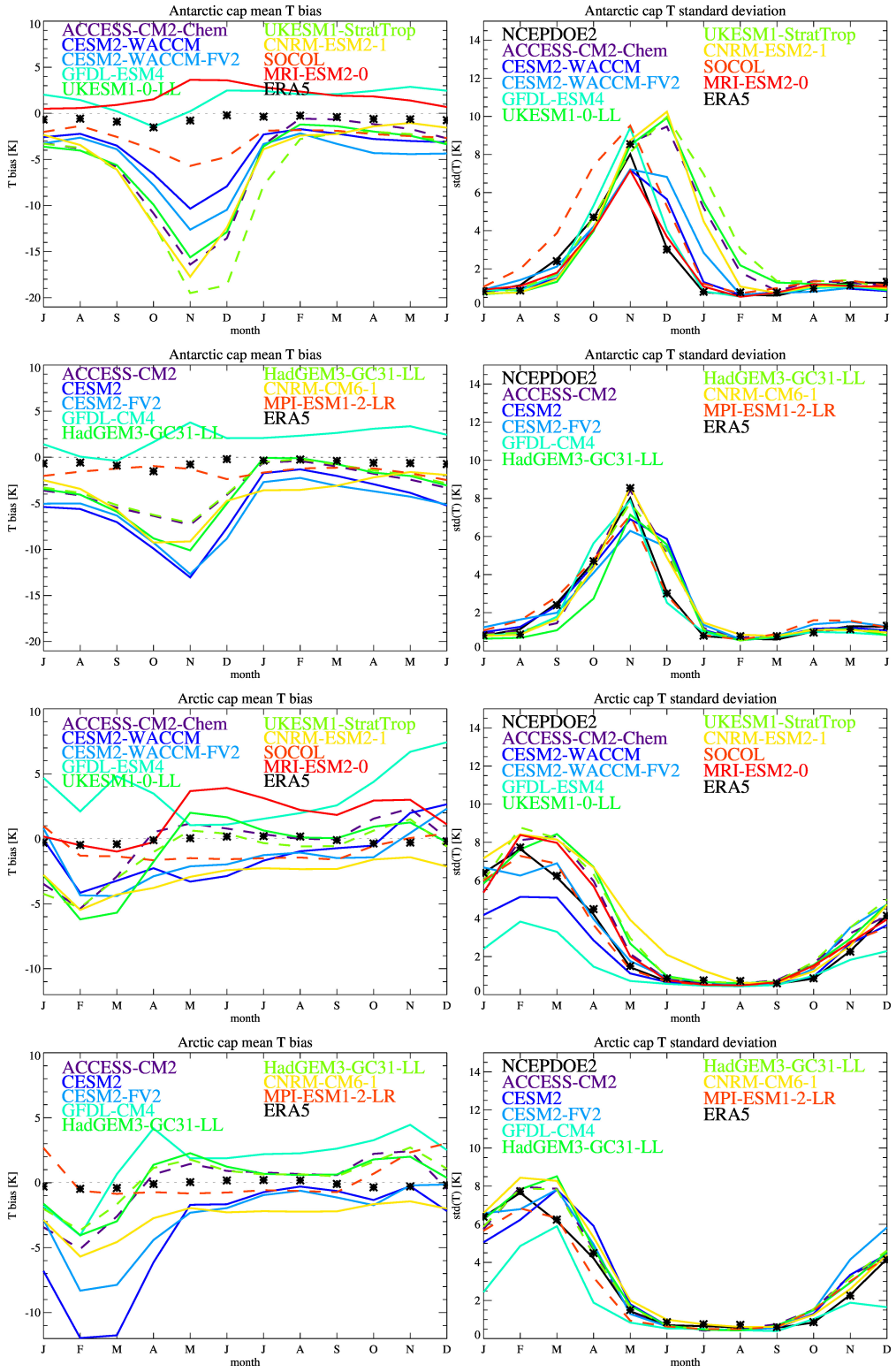

Figure 3. (left) Monthly-mean 70 hPa temperature bias (K) relative to NCEP-DOE2 for the period 1980-2014. (right) Standard deviation of monthly-mean 70 hPa temperature (K). (1st row) Chemistry-climate models, Antarctic polar cap mean. (2nd row) Same for the associated no-chemistry models. (3rd row) Chemistry-climate models, Arctic polar cap mean. (4th row) Same for the associated no-chemistry models. Solid lines represent CMIP6 "historical" ensembles, dashed lines are CMIP6 AMIP (ACCESS-CM2, HadGEM3-GC31-LL, MPI-ESM1-2-LR) and CCMI2 REF-D1 (ACCESS-CM2-Chem, UKESM1-StratTrop, SOCOL) ensembles. respectively. Black '*' symbols denote the ERA5 reanalysis.

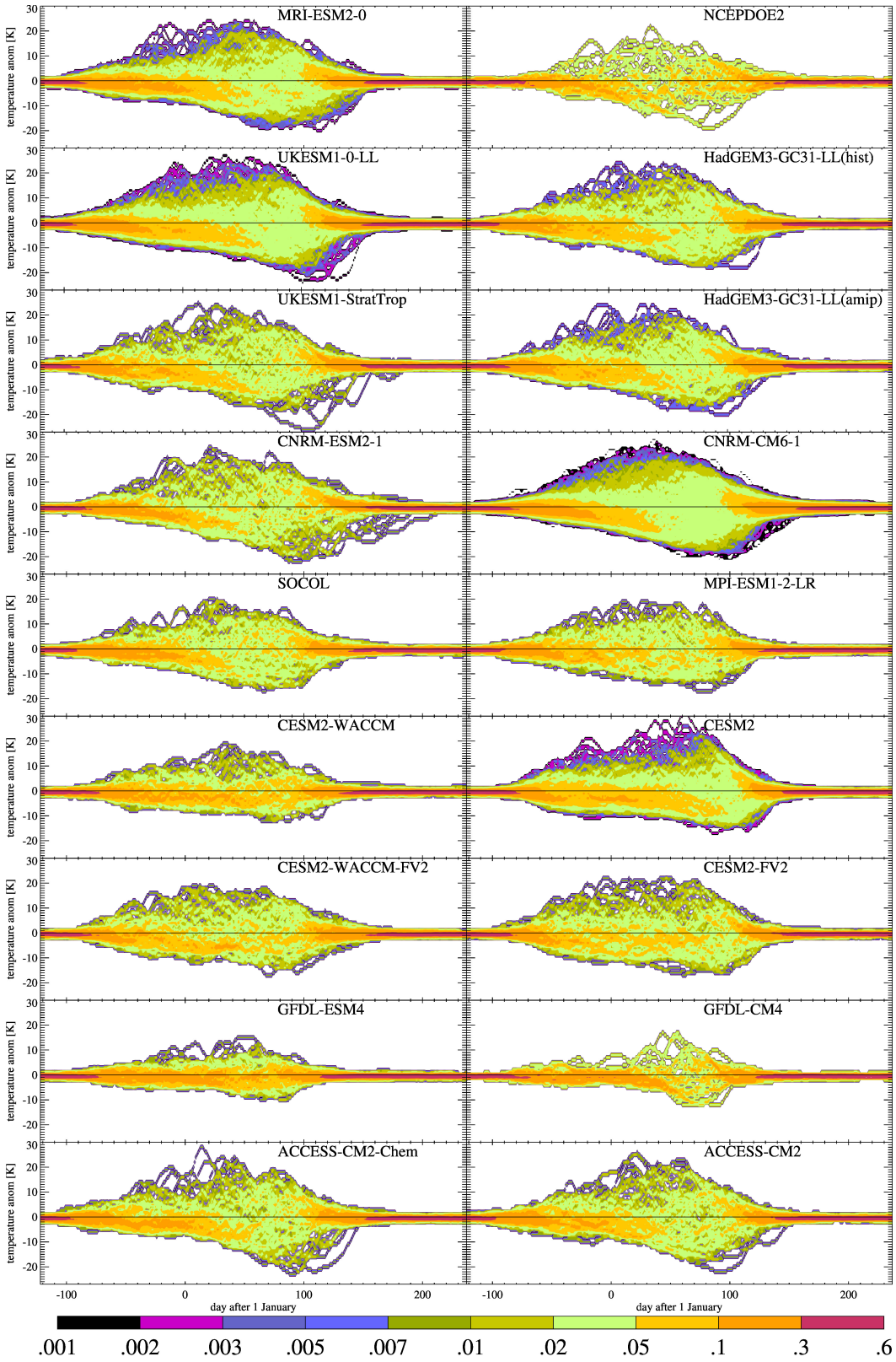

Figure 4. Probability density plot (K^{-1}) of the Arctic-mean (75°N-90°N) temperature anomaly relative to the 1980-2014 mean seasonal cycle at 70 hPa, in the NCEP-DOE2 reanalysis and the climate models as a function of the day of the year, for September 1980 to August 2014. Models with larger ensembles allow for better sampling of low-probability temperature anomalies (colored in blue and violet); these colors are therefore absent for small-ensemble models and the reanalysis.

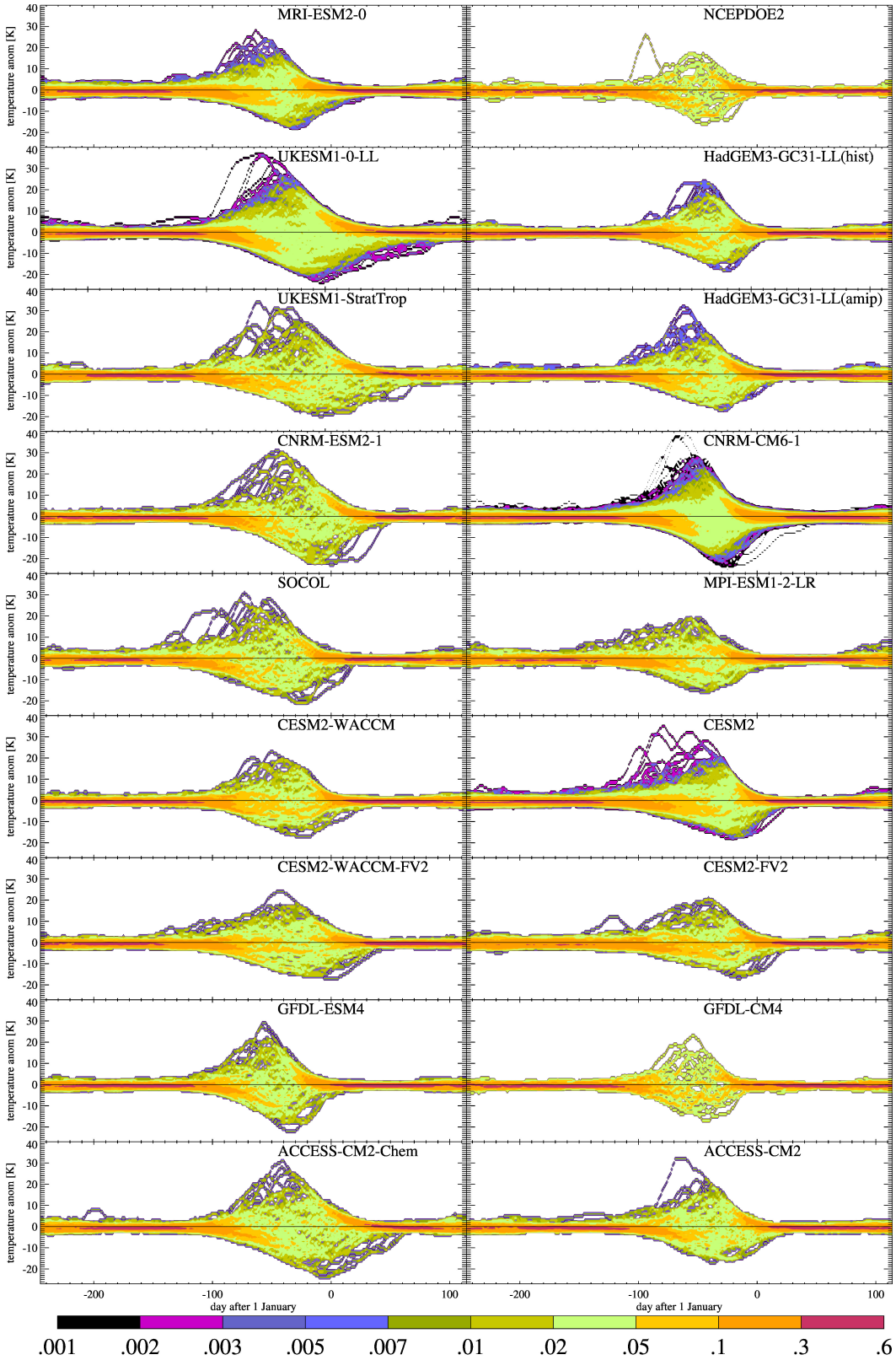

Figure 5. Same as figure 4, but for Antarctic 70 hPa polar-cap mean temperatures.

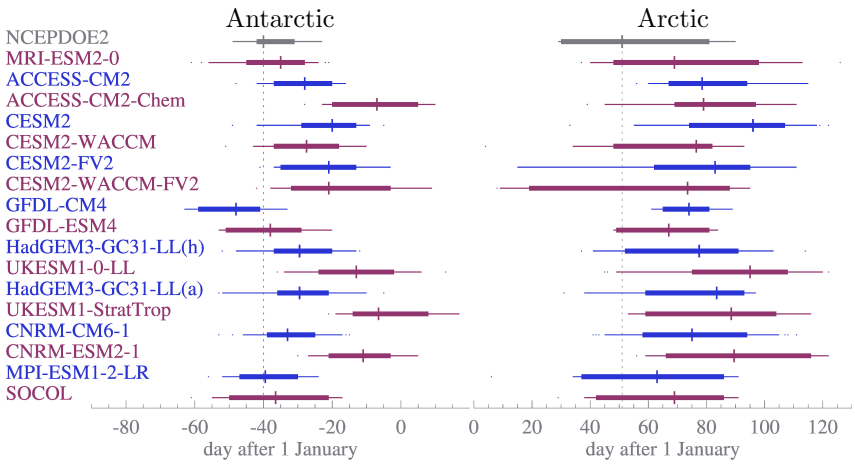

Figure 6. Date of occurrence of lowest temperatures in years with the minimum of the 15-day filtered 70 hPa polar-mean temperature anomaly in the lowest 20%. Left: Antarctica (75-90°S). Right: Arctic (75-90°N). Vertical bars: Median day. Thick lines: 16 to 84-percentile range. Thin lines: 2.5 to 97.5-percentile range. Dots: Outliers outside the 2.5 to 97.5 percentiles. Note that the NCEP-DOE2 reanalysis and the GFDL-CM4 model simulations only have 7 such cold winters each in this 20% category (out of a total of 34, for the period of September 1980 - August 2014). The long dashed vertical lines mark the medians of the coldest days in the NCEP-DOE2 reanalysis. Chemistry models are represented in red, no-chemistry models in blue. '(h)' stands for the "historical" ensemble of HadGEM3-GC31-LL, '(a)' for AMIP.

A remarkable warm outlier is seen in the NCEP-DOE2 reanalysis around day −100 (i.e. 23 September; figure 5). This is the vortex breakup and major stratospheric warming of 2002 which at the time was considered very unusual as it had never been seen before in the observational record (Newman & Nash, 2003). Both chemistry (UKESM1, SOCOL) and no-chemistry (CNRM-CM6-1, CESM2) models exhibit similar extremely warm episodes around this time of the year, meaning that some CMIP6/CCMI2 models can qualitatively simulate such events (Jucker et al., 2021).

Restricting our attention to the 20% of years with the lowest 15-day mean temperature anomalies in the Arctic and Antarctic at 70 hPa, figure 6 indicates that the median of such cold days, in the reanalysis, is around day 51 (21 February) in the Arctic and day −40 (22 November) in the Antarctic. It is noteworthy that all 17 models considered here simulate a later median date for the thus defined coldest day in the Arctic at 70 hPa by 15 days or more. Also in the Antarctic most models simulate a delay in the coldest day. Both findings may illustrate that climate models struggle with correctly capturing stratospheric dynamics in the polar regions, although it is impossible to be sure given that only the seven coldest winters are considered in the reanalysis (out of 34 in total). In the Arctic, all models have some degree of overlap of the 16 to 84-percentile interval for this coldest date with the reanalysis, whereas in the Antarctic, where the reanalysis shows relatively little variation in the date of the coldest day, the models simulating the most severe ozone depletion (ACCESS-CM2-Chem, UKESM1, and CNRM-ESM2-1) all have 16 to 84-percentile ranges for this diagnostic that do not overlap with those of the reanalysis.

For these four model pairs (ACCESS, HadGEM3 / UKESM1 – both versions, CNRM) the chemistry variants simulate a delay in the median occurrence of the coldest day by around ∼10 to 30 days versus their no-chemistry counterparts. This holds in both po-

lar regions. However, for other model pairs this is not the case: The MPI-ESM1-2-LR/SOCOL pair exhibits quite similar behavior for both polar regions, the GFDL pair simulates shifts in the coldest day of different signs in the two polar regions, and the CESM2 pairs, in most cases, produce an earlier coldest day if interactive chemistry is used. We will discuss these findings more in section 5.

4.3 Composite analysis of cold stratospheric winters

Next we produce composites for temperature and GPH, similar to Baldwin and Dunkerton (2001)'s method. We express these fields relative to the time of occurrence of the largest absolute temperature anomaly (deviation from the mean) at 70 hPa. Baldwin and Dunkerton (2001) and Thompson et al. (2005) had used NAM and SAM indices instead, respectively.

Figures 7 and 8 show that the 20% coldest winters, at the time of the lowest temperature, are generally between 10 and 20 K colder at 70 hPa than the average winter, in agreement with figures 4 and 5. Substantial cold anomalies however often start at least two months before the largest temperature anomalies occur, and last 30-50 days beyond this date. They are accompanied by corresponding negative GPH anomalies that typically extend into the troposphere, in agreement with Baldwin and Dunkerton (2001)'s and Thompson et al. (2005)'s findings.

For both polar regions, agreement between the CMIP6 models is generally remarkably good. However, in several model pairs temperature and GPH anomalies are systematically more persistent, both at the start and end, for both polar regions, and also often of larger-amplitude in the chemistry-climate models (UKESM1 – both variants, CNRM-ESM2-1, SOCOL, MRI-ESM2-0, ACCESS-CM2-Chem) than in the corresponding no-chemistry models (HadGEM3, CNRM-CM6-1, MPI-ESM1, ACCESS-CM2). In particular, these six chemistry models all maintain substantial cold anomalies in the lower stratosphere well beyond day 50 after the lowest temperatures occur in the Arctic at 70 hPa. In all six cases, the chemistry models produce more persistent cold anomalies than their no-chemistry counterparts. The cold anomalies lasting well into spring are reflected in GPH anomalies also lasting longer and spawning low-GPH anomalies in the troposphere, signaling impacts of this behavior on simulated tropospheric weather.

Exceptions to this behavior are the GFDL and the CESM2 / CESM2-FV2 families which do not exhibit substantial differences in the persistence of the Arctic cold anomalies between the chemistry and no-chemistry models. An inspection of figure 4 shows that GFDL-ESM4 simulates less polar temperature variability in the Arctic than the other CCMs. It is characterized by a substantial warm bias with practically no "cold" Arctic winters with corresponding ozone depletion (figure 3) and far too weak ozone trends for the 1979-2000 period (figure 1; Morgenstern et al., 2020). By contrast, CESM2-WACCM has a cold bias but also simulates too weak Arctic ozone trends for 1979-2000 (Morgenstern et al., 2020).

Comparing now the climate models with the NCEP-DOE2 reanalysis (Kanamitsu et al., 2002), in the cases where the chemistry models exhibit increased persistence, the no-chemistry counterparts are in better agreement with observations than the chemistry models. This means the persistence of cold anomalies long into spring seen in most CCMs is not reflected in the NCEP-DOE2 reanalysis.

5 Discussion

We have analyzed the dynamics of stratospheric cold winters in 13 CMIP6 and three CCMI2 climate and chemistry-climate models and compared them to reanalyses. The behavior of the chemistry models depends crucially on whether substantial additional

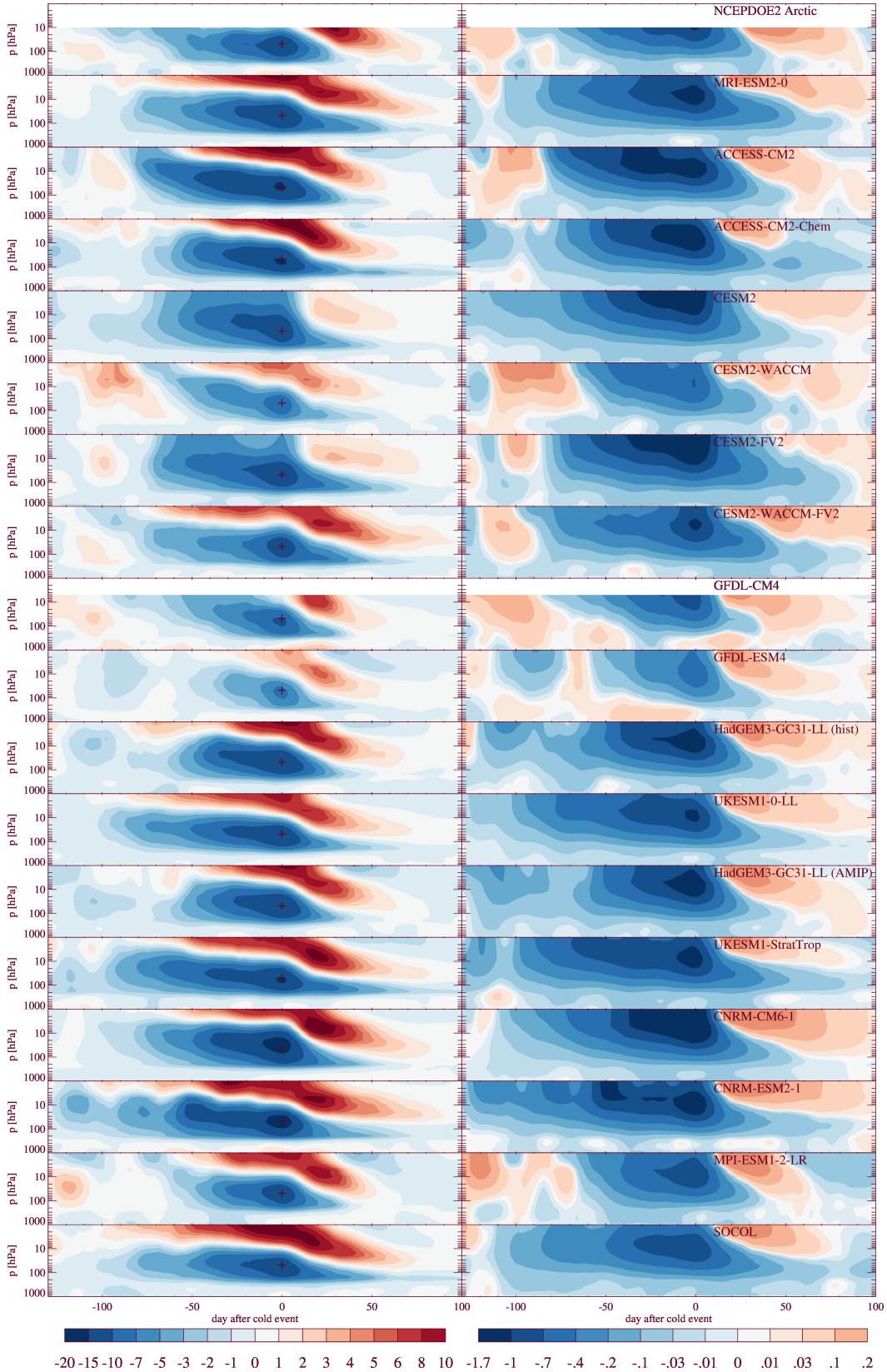

Figure 7. Arctic polar cap (75°N-90°N) mean temperature (left; in K) and GPH (right; in km) anomalies (relative to their 1980-2014 mean seasonal cycles) for the 20% coldest winters in the chemistry-climate and no-chemistry models. Time is relative to the day of occurrence of the coldest day at 70 hPa, marked by a small '+' symbol.

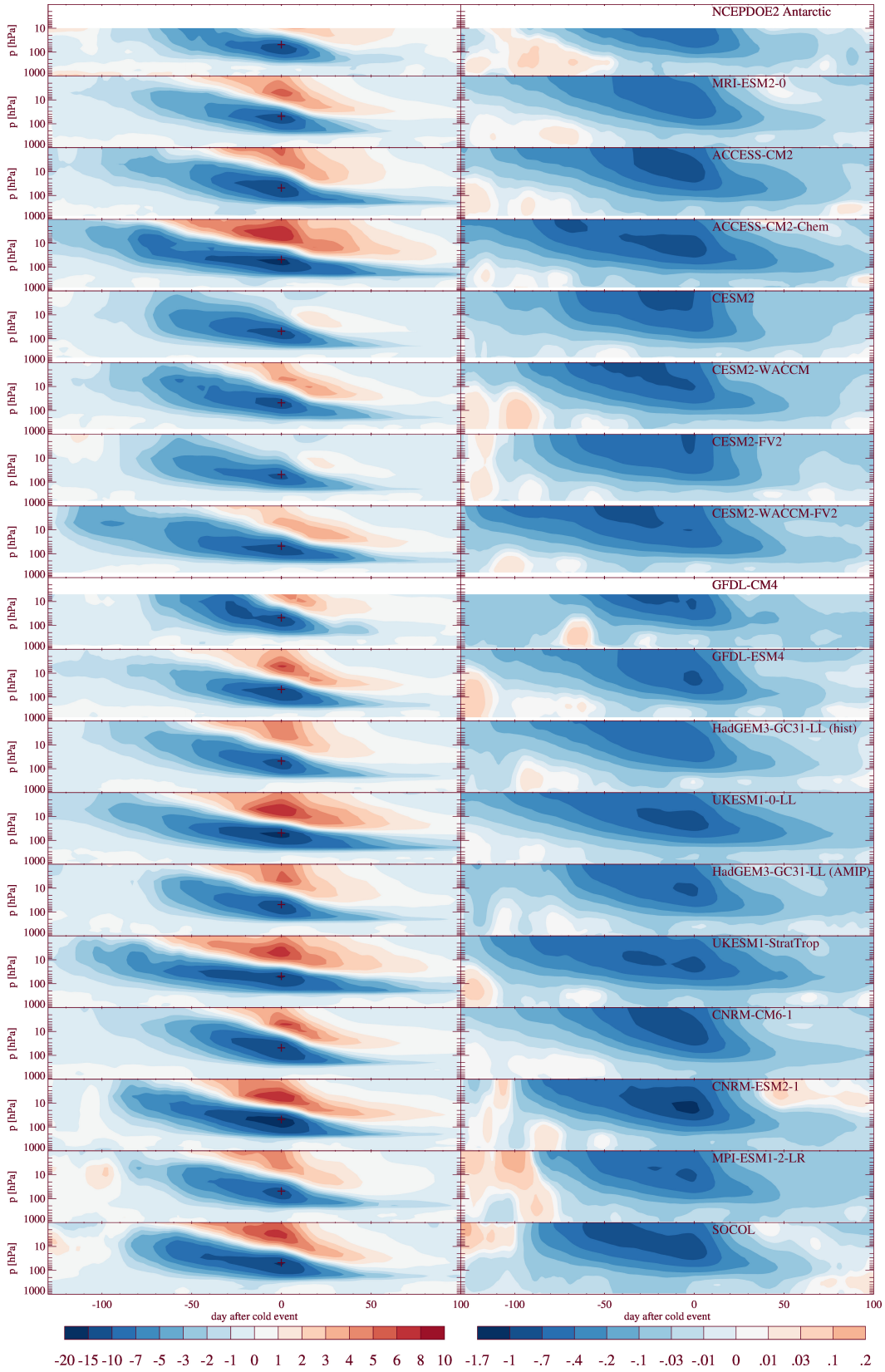

Figure 8. Same a figure 7, but for the Antarctic polar cap (75°S-90°S).

CCM	no-chemistry model	tuning	winter bias	springtime variability	coldest day	persistence
ACCESS-CM2-Chem	ACCESS-CM2	no	cold	high	late	increased
CESM2-WACCM	CESM2	yes	cold	good/low	early	similar
CESM2-WACCM-FV2	CESM2-FV2	yes	cold	good	early/unchanged	similar
CNRM-ESM2-1	CNRM-CM6-1	no	cold	high	late	increased
GFDL-ESM4	GFDL-CM4	yes	warm/small	good/low	late/early	similar
MRI-ESM2-0			small	good/high		
SOCOL	MPI-ESM1-2	no	small	good	unchanged	increased
UKESM1-0-LL	HadGEM3-GC31-LL	no	cold	high	late	increased
UKESM1-StratTrop	HadGEM3-GC31-LL	no	cold	high	late	increased

Table 2. Summary of findings. "Tuning" refers to any substantial differences in the dynamical part of the model relative to the no-chemistry base model. The "winter bias" and "variability" are for the monthly-mean 70 hPa temperatures at 75°S-90°S and 75°N-90°N, relative to ERA-Interim, for the CCMs (figure 3). The "coldest day" refers to the shift in the median occurrence of the coldest day relative to the corresponding no-chemistry model (hence it is "not applicable" to MRI-ESM2-0; figure 6). The "persistence" is qualitatively discerned from figures 7 and 8 and is relative to the corresponding no-chemistry models. Again this is "not applicable" to the MRI-ESM2-0 model which however behaves similarly to the other CCMs with "increased" persistence times.

differences, extending beyond interactive ozone chemistry, exist between the chemistry models and their no-chemistry equivalents. In four cases where the dynamics configurations are essentially unchanged versus the no-chemistry configuration (ACCESS-CM2-Chem, CNRM-ESM2-1, UKESM1-0-LL/UKESM1-StratTrop), coupling in chemistry results in a delay in the occurrence of the coldest day, both in the Arctic and Antarctic lower stratosphere. As a result of this extension of the cold season, all of these models unrealistically simulate maximum polar ozone loss during the months of April and November, respectively. The SOCOL model does not exhibit any shift in the timing of the coldest day versus its reference model MPI-ESM1-2-LR; SOCOL is also characterized by a generally good representation of ozone trends (Sukhodolov et al., 2021), albeit with an early onset of ozone loss in the Antarctic, and a good representation of Arctic temperature and variability (figure 3). All five of these chemistry models exhibit timescales of persistence of stratospheric cold anomalies over both poles that are longer than in their no-chemistry counterparts, reflecting extensions of the lifetimes of both polar vortices. Possibly in the SOCOL model, this increased persistence is counteracted by the early onset of ozone depletion in the Antarctic, resulting in no shift of the occurrence of the coldest day relative to the backgroud model, MPI-ESM1-2-LR. The MRI-ESM2-0 model also behaves similarly to these chemistry models. These extendeded lifetimes of the polar vortices compare worse to a reanalysis than the shorter lifetimes of the polar vortices characterizing the correspondiong no-chemistry models. This impact of interactive chemistry is consistent with earlier studies based on fewer models (Haase & Matthes, 2019; Oehrlein et al., 2020; Lin & Ming, 2021).

The behavior of this group of models contrasts with the GFDL and two CESM2 pairs of models. In these pairs, the chemistry models differ more substantially in their dynamics configurations from their no-chemistry counterparts, namely the chemistry versions operate on a vertically extended grid with more levels in the stratosphere, compared to their no-chemistry counterparts. CESM2-WACCM and CESM-WACCM-FV2

include an additional non-orographic gravity wave drag (NOGWD) parameterization (Gettelman et al., 2019) completely absent in CESM2 and CESM2-FV2. GFDL-ESM4 also differs in terms of NOGWD and a few other aspects (Dunne et al., 2020). NOGWD drives the Brewer-Dobson Circulation and influences the stability of the polar vortices (Eichinger et al., 2020, for a recent review see), so may well explain the differences in behavior between the CCMs and their no-chemistry equivalents. GFDL-ESM4 is the only chemistry model studied here with a substantial warm bias in the Arctic stratosphere in winter. Together with the much underestimated variability (figure 3) this indicates this model does not realistically simulate Arctic ozone depletion (Morgenstern et al., 2020), but ranks amongst the top-performing models for Antarctic ozone depletion. CESM2-WACCM, like most other models studied here, has a cold bias in the Arctic winter stratosphere. Together also with the underestimated variability this suggests that the model simulates too many "cold" polar vortices with too regular ozone depletion. Both in the CESM2 and the GFDL models, however, the timings of the coldest days, for both polar regions, are either unchanged or more realistic in the chemistry models. The timescales of persistence are not appreciably different between the chemistry and no-chemistry configurations of these models.

The findings illustrate that in the cases where ozone chemistry is the only significant difference between two model configurations, ozone chemistry introduces additional "memory" into the atmosphere. Feedbacks of ozone chemistry onto radiation, for a cold winter, enhance radiative cooling and stabilize the vortex to last longer into spring; similar results were found in earlier single-model studies (Oehrlein et al., 2020; Lin & Ming, 2021). These effects can however be counterbalanced by retuning and/or additional physics, especially the non-orographic gravity wave scheme added or modified in GFDL-ESM4 and CESM2-WACCM (both versions).

The findings illustrate that additional "physics" that, based on first principles, can be expected to better capture Earth system feedbacks, such as ozone chemistry, will only lead to a better reproduction of atmospheric dynamics and climate if other processes are tuned to account for its presence in a climate model. In particular, NOGWD schemes are often adjusted to improve the simulation of stratospheric dynamics. In the absence of such tuning, adding in interactive ozone chemistry may degrade performance, which might erroneously be understood to count against including this process in a climate model.

Data availability

CMIP6 data are available at `https://esgf-node.llnl.gov/search/cmip6/`. Specifically, the following datasets are used: Dix et al. (2019); Danabasoglu (2019a, 2019b, 2019c, 2019d); Séférian (2018); Voldoire (2018); Guo et al. (2018); Krasting et al. (2018); Yukimoto, Koshiro, et al. (2019); Wieners et al. (2019); Tang et al. (2019); Byun (2020); Ridley et al. (2019a, 2019b). CCMI2 data are downloaded from `ftp://ftp.ceda.ac.uk/badc/ccmi/data/post-cmip6/ccmi-2022`. Specifically, the following datasets have been used: Dennison and Woodhouse (2021); Rozanov et al. (2021); Abraham and Keeble (2021). NCEP-DOE2 data were provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, from their web site at `https://psl.noaa.gov/data/gridded/data.ncep.reanalysis2.pressure.html`. Hersbach et al. (2019) was downloaded from the Copernicus Climate Change Service (C3S) Climate Data Store. The results contain modified Copernicus Climate Change Service information. Neither the European Commission nor ECMWF are responsible for any use that may be made of the Copernicus information or data it contains.

MSR-2 data are available at `https://www.temis.nl/protocols/O3global.php` (van der A et al., 2015b).

Acknowledgments

We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. We also acknowledge IGAC, SPARC, and CEDA for their support of CCMI. We acknowledge NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, for providing the NCEP/DOE Reanalysis 2 data, ECMWF for providing the ERA5 data, and KNMI for providing the MSR-2 total-column ozone climatology.

OM and GZ were supported by the NZ Government's Strategic Science Investment Fund (SSIF) through the NIWA programme CACV. MD was supported by the Japan Society for the Promotion of Science KAKENHI (grant number: JP20K04070). FMO'C and YT were supported by the Met Office Hadley Centre Climate Programme funded by BEIS. PL is under award NA18OAR4320123 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce. All calculations with the SOCOLv4.0 model were supported by the Swiss National Supercomputing Centre (CSCS) under projects S-901 (ID 154), S-1029 (ID 249), and S-903. ER, TE, and TS were supported by the Swiss National Science Foundation (SNSF) project POLE (grant no. 200020-182239).

References

Abraham, N. L., & Keeble, J. (2021). *CCMI-2022: REF-D1 data produced by the UKESM1-StratTrop model at NCAS Cambridge.* NERC EDS Centre for Environmental Data Analysis. Retrieved from TBD

Baldwin, M. P., & Dunkerton, T. J. (2001). Stratospheric harbingers of anomalous weather regimes. *Science*, *294*(5542), 581-584. Retrieved from https://www.science.org/doi/abs/10.1126/science.1063315 doi: 10.1126/science.1063315

Bi, D., Dix, M., Marsland, S., O'Farrell, S., Sullivan, A., Bodman, R., ... Heerdegen, A. (2020). Configuration and spin-up of ACCESS-CM2, the new generation Australian Community Climate and Earth System Simulator Coupled Model. *Journal of Southern Hemisphere Earth Systems Science*, *70*, 225-251.

Bodman, R. W., Karoly, D. J., Dix, M. R., Harman, I. N., Srbinovsky, J., Dobrohotoff, P. B., & Mackallah, C. (2020). Evaluation of CMIP6 AMIP climate simulations with the ACCESS-AM2 model. *Journal of Southern Hemisphere Earth Systems Science*, *70*, 166-179. Retrieved from https://doi.org/10.1071/ES19033

Byun, Y.-H. (2020). *NIMS-KMA UKESM1.0-LL model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.8379 doi: 10.22033/ESGF/CMIP6.8379

Checa-Garcia, R., Hegglin, M. I., Kinnison, D., Plummer, D. A., & Shine, K. P. (2018). Historical tropospheric and stratospheric ozone radiative forcing using the CMIP6 database. *Geophysical Research Letters*, *45*(7), 3264-3273. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL076770 doi: 10.1002/2017GL076770

Danabasoglu, G. (2019a). *NCAR CESM2-FV2 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.11297 doi: 10.22033/ESGF/CMIP6.11297

Danabasoglu, G. (2019b). *NCAR CESM2 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/

10.22033/ESGF/CMIP6.7627 doi: 10.22033/ESGF/CMIP6.7627

Danabasoglu, G. (2019c). *NCAR CESM2-WACCM-FV2 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.11298 doi: 10.22033/ESGF/CMIP6.11298

Danabasoglu, G. (2019d). *NCAR CESM2-WACCM model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.10071 doi: 10.22033/ESGF/CMIP6.10071

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., ... Strand, W. G. (2020). The Community Earth System Model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001916 (e2019MS001916 2019MS001916) doi: https://doi.org/10.1029/2019MS001916

Dennison, F., & Woodhouse, M. (2021). *CCMI-2022: REF-D1 data produced by the ACCESS-CM2-Chem model at CSIRO-ARCCSS.* NERC EDS Centre for Environmental Data Analysis. Retrieved from TBD

Dennison, F., & Woodhouse, M. T. (2022). ACCESS-CM2-Chem: Evaluation of Southern Hemisphere ozone and its effect on the Southern Annular Mode. *Journal of Southern Hemisphere Earth System Science*, *in review*.

Dix, M., Bi, D., Dobrohotoff, P., Fiedler, R., Harman, I., Law, R., ... Yang, R. (2019). *CSIRO-ARCCSS ACCESS-CM2 model output prepared for CMIP6 CMIP AMIP.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.4239 doi: 10.22033/ESGF/CMIP6.4239

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., ... Zhao, M. (2020). The GFDL Earth System Model version 4.1 (GFDL-ESM4.1): Overall coupled model description and simulation characteristics. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS002015. Retrieved from https://doi.org/10.1029/2019MS002015

Eichinger, R., Garny, H., Šácha, P., & Oberländer-Hayn, S. (2020). Effects of missing gravity waves on stratospheric dynamics; part 1: climatology. *Climate Dynamics*, *54*, 3165–3183. Retrieved from https://doi.org/10.1007/s00382-020-05166-w

Eyring, V., Gillett, N., AchutaRao, K., Barimalala, R., Parrillo, M. B., Bellouin, N., ... Sun, Y. (2021). Human Influence on the Climate System. In V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (chap. 3). Cambridge University Press, Cambridge, UK and New York, NY.

Fogt, R. L., & Marshall, G. J. (2020). The Southern Annular Mode: Variability, trends, and climate impacts across the Southern Hemisphere. *WIREs Climate Change*, *11*(4), e652. Retrieved from https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcc.652 doi: https://doi.org/10.1002/wcc.652

Friedel, M., Chiodo, G., Stenke, A., Domeisen, D., Fueglistaler, S., Anet, J., & Peter, T. (2022). *Robust effects of springtime Arctic ozone depletion on surface climate.* Retrieved from https://doi.org/10.21203/rs.3.rs-496081/v1

Gettelman, A., Mills, M. J., Kinnison, D. E., Garcia, R. R., Smith, A. K., Marsh, D. R., ... Randel, W. J. (2019). The Whole Atmosphere Community Climate Model version 6 (WACCM6). *Journal of Geophysical Research: Atmospheres*, *124*(23), 12380-12403. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD030943 doi: 10.1029/2019JD030943

Guo, H., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., ... Zhang, R. (2018). *NOAA-GFDL GFDL-CM4 model output historical.* Earth

System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.8594 doi: 10.22033/ESGF/CMIP6.8594

Haase, S., & Matthes, K. (2019). The importance of interactive chemistry for stratosphere–troposphere coupling. *Atmospheric Chemistry and Physics*, *19*(5), 3417–3432. Retrieved from https://acp.copernicus.org/articles/19/3417/2019/ doi: 10.5194/acp-19-3417-2019

Hardiman, S. C., Andrews, M. B., Andrews, T., Bushell, A. C., Dunstone, N. J., Dyson, H., ... Wood, R. A. (2019). The impact of prescribed ozone in climate projections run with HadGEM3-GC3.1. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3443-3453. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001714 doi: https://doi.org/10.1029/2019MS001714

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., ... Zadeh, N. (2019). Structure and performance of GFDL's CM4.0 climate model. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3691-3727. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001829 doi: https://doi.org/10.1029/2019MS001829

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., ... Thépaut, J.-N. (2019). *ERA5 monthly averaged data on pressure levels from 1979 to present.* Copernicus Climate Change Service (C3S) Climate Data Store (CDS), accessed on 03-May-2022. doi: 10.24381/cds.6860a573

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049. Retrieved from https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803 doi: https://doi.org/10.1002/qj.3803

Ivy, D. J., Solomon, S., Calvo, N., & Thompson, D. W. J. (2017). Observed connections of Arctic stratospheric ozone extremes to Northern Hemisphere surface climate. *Environmental Research Letters*, *12*(2), 024004.

Jucker, M., Reichler, T., & Waugh, D. W. (2021). How frequent are Antarctic sudden stratospheric warmings in present and future climate? *Geophysical Research Letters*, *48*(11), e2021GL093215. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021GL093215 (e2021GL093215 2021GL093215) doi: https://doi.org/10.1029/2021GL093215

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., & Potter, G. L. (2002). NCEP-DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society*, 1631-1643.

Krasting, J. P., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., ... Zhao, M. (2018). *NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.8597 doi: 10.22033/ESGF/CMIP6.8597

Kuhlbrodt, T., Jones, C. G., Sellar, A., Storkey, D., Blockley, E., Stringer, M., ... Walton, J. (2018). The low-resolution version of HadGEM3 GC3.1: Development and evaluation for global climate. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2865-2888. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001370 doi: https://doi.org/10.1029/2018MS001370

Kuttippurath, J., Feng, W., Müller, R., Kumar, P., Raj, S., Gopikrishnan, G. P., & Roy, R. (2021). Exceptional loss in ozone in the Arctic winter/spring of 2019/2020. *Atmospheric Chemistry and Physics*, *21*(18), 14019–14037. Retrieved from https://acp.copernicus.org/articles/21/14019/2021/ doi: 10.5194/acp-21-14019-2021

Lin, P., & Ming, Y. (2021). Enhanced climate response to ozone depletion from ozone-circulation coupling. *Journal of Geophysical Research: Atmospheres*,

126(7), e2020JD034286. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JD034286 (e2020JD034286 2020JD034286) doi: https://doi.org/10.1029/2020JD034286

Lin, P., Paynter, D., Polvani, L., Correa, G. J. P., Ming, Y., & Ramaswamy, V. (2017). Dependence of model-simulated response to ozone depletion on stratospheric polar vortex climatology. *Geophysical Research Letters*, *44*(12), 6391-6398. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL073862 doi: https://doi.org/10.1002/2017GL073862

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., ... Roeckner, E. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing co_2. *Journal of Advances in Modeling Earth Systems*, *11*(4), 998-1038. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001400 doi: https://doi.org/10.1029/2018MS001400

Morgenstern, O. (2021). The Southern Annular Mode in 6^{th} Coupled Model Intercomparison Project models. *Journal of Geophysical Research: Atmospheres*, *126*(5), e2020JD034161. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JD034161 (e2020JD034161 2020JD034161) doi: https://doi.org/10.1029/2020JD034161

Morgenstern, O., Frith, S. M., Bodeker, G. E., Fioletov, V., & van der A, R. J. (2021). Reevaluation of total-column ozone trends and of the effective radiative forcing of ozone-depleting substances. *Geophysical Research Letters*, *48*(21), e2021GL095376. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021GL095376 (e2021GL095376 2021GL095376) doi: https://doi.org/10.1029/2021GL095376

Morgenstern, O., O'Connor, F. M., Johnson, B. T., Zeng, G., Mulcahy, J. P., Williams, J., ... Kinnison, D. E. (2020). Reappraisal of the climate impacts of ozone-depleting substances. *Geophysical Research Letters*, *47*(20), e2020GL088295. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL088295 (e2020GL088295 10.1029/2020GL088295) doi: https://doi.org/10.1029/2020GL088295

Newman, P., & Nash, E. (2003). The unusual Southern Hemisphere stratosphere winter of 2002. *Journal of the Atmospheric Sciences*, *62*. doi: 10.1175/JAS-3323.1

Oehrlein, J., Chiodo, G., & Polvani, L. M. (2020). The effect of interactive ozone chemistry on weak and strong stratospheric polar vortex events. *Atmospheric Chemistry and Physics*, *20*(17), 10531–10544. Retrieved from https://acp.copernicus.org/articles/20/10531/2020/ doi: 10.5194/acp-20-10531-2020

Plummer, D., Nagashima, T., Tilmes, S., Archibald, A., Chiodo, G., Fadnavis, S., ... Peter, T. (2021). CCMI-2022: A new set of Chemistry-Climate Model Initiative (CCMI) community simulations to update the assessment of models and support upcoming Ozone Assessment activities. *SPARC Newsletter*, *57*, 22-30. Retrieved from https://www.sparc-climate.org/wp-content/uploads/sites/5/2021/07/SPARCnewsletter_Jul2021_web.pdf

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., ... Kaplan, A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, *108*(D14). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD002670 doi: https://doi.org/10.1029/2002JD002670

Ridley, J., Menary, M., Kuhlbrodt, T., Andrews, M., & Andrews, T. (2019a). *MOHC HadGEM3-GC31-LL model output prepared for CMIP6 CMIP AMIP.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.5853 doi: 10.22033/ESGF/CMIP6.5853

Ridley, J., Menary, M., Kuhlbrodt, T., Andrews, M., & Andrews, T. (2019b). *MOHC HadGEM3-GC31-LL model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.6109 doi: 10.22033/ESGF/CMIP6.6109

Rieder, H. E., Chiodo, G., Fritzer, J., Wienerroither, C., & Polvani, L. M. (2019). Is interactive ozone chemistry important to represent polar cap stratospheric temperature variability in Earth-system models? *Environmental Research Letters*, *14*(4), 044026.

Rozanov, E., Egorova, T., & Sukhodolov, T. (2021). *CCMI-2022: REF-D1 data produced by the SOCOL model at ETH-PMOD.* NERC EDS Centre for Environmental Data Analysis. Retrieved from https://catalogue.ceda.ac.uk/uuid/f088e7a33a25409197ea9b6aa3b90864

Séférian, R. (2018). *CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.4068 doi: 10.22033/ESGF/CMIP6.4068

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., ... Madec, G. (2019). Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth system processes in present-day and future climate. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4182-4227. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001791 doi: 10.1029/2019MS001791

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., ... Zerroukat, M. (2019). UKESM1: Description and evaluation of the U.K. Earth System Model. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4513-4558. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001739 doi: 10.1029/2019MS001739

Son, S.-W., Gerber, E. P., Perlwitz, J., Polvani, L. M., Gillett, N. P., Seo, K.-H., ... Yamashita, Y. (2010). Impact of stratospheric ozone on Southern Hemisphere circulation change: A multimodel assessment. *Journal of Geophysical Research: Atmospheres*, *115*(D3). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JD014271 doi: https://doi.org/10.1029/2010JD014271

Sukhodolov, T., Egorova, T., Stenke, A., Ball, W. T., Brodowsky, C., Chiodo, G., ... Rozanov, E. (2021). Atmosphere–ocean–aerosol–chemistry–climate model SOCOLv4.0: description and evaluation. *Geoscientific Model Development*, *14*(9), 5525–5560. Retrieved from https://gmd.copernicus.org/articles/14/5525/2021/ doi: 10.5194/gmd-14-5525-2021

Tang, Y., Rumbold, S., Ellis, R., Kelley, D., Mulcahy, J., Sellar, A., ... Jones, C. (2019). *MOHC UKESM1.0-LL model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.6113 doi: 10.22033/ESGF/CMIP6.6113

Taylor, K. E., Williamson, D., & Zwiers, F. (2015). *AMIP sea surface temperature and sea ice concentration boundary conditions.* Retrieved from https://pcmdi.llnl.gov/mips/amip/details/index.html

Thompson, D., Baldwin, M., & Solomon, S. (2005). Stratosphere–troposphere coupling in the Southern Hemisphere. *Journal of the Atmospheric Sciences*, *3, 62*, 708-715. Retrieved from https://journals.ametsoc.org/view/journals/atsc/62/3/jas-3321.1.xml doi: 10.1175/JAS-3321.1

Thompson, D., Solomon, S., Kushner, P., England, M. H., Grise, K. M., & Karoly, D. J. (2011). Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change. *Nature Geoscience*, *4*, 741–749. Retrieved from https://doi.org/10.1038/ngeo1296

van der A, R. J., Allaart, M. A. F., & Eskes, H. J. (2015a). Extended and refined multi sensor reanalysis of total ozone for the period 1970–2012. *Atmospheric*

Measurement Techniques, *8*, 3021–3035. Retrieved from https://doi.org/10.5194/amt-8-3021-2015

van der A, R. J., Allaart, M. A. F., & Eskes, H. J. (2015b). *Multi-Sensor Reanalysis (MSR) of total ozone, version 2 dataset.* Royal Netherlands Meteorological Institute (KNMI). doi: 10.21944/temis-ozone-msr2

Voldoire, A. (2018). *CMIP6 simulations of the CNRM-CERFACS based on CNRM-CM6-1 model for CMIP experiment historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.4066 doi: 10.22033/ESGF/CMIP6.4066

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., ... Waldman, R. (2019). Evaluation of CMIP6 DECK experiments with CNRM-CM6-1. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2177-2213. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001683 doi: https://doi.org/10.1029/2019MS001683

Wieners, K.-H., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M., Bittner, M., ... Roeckner, E. (2019). *MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP AMIP.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.6464 doi: 10.22033/ESGF/CMIP6.6464

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., ... Xavier, P. K. (2018). The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *Journal of Advances in Modeling Earth Systems*, *10*(2), 357-380. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017MS001115 doi: https://doi.org/10.1002/2017MS001115

WMO. (2018). *Scientific Assessment of Ozone Depletion: 2018.* Geneva, Switzerland: World Meteorological Organization. Retrieved from https://www.esrl.noaa.gov/csl/assessments/ozone/2018/

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., ... Ishii, M. (2019). The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and basic evaluation of the physical component. *Journal of the Meteorological Society of Japan Ser. II*. doi: 10.2151/jmsj.2019-051

Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., Yoshida, K., Urakawa, S., ... Adachi, Y. (2019). *MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.6842 doi: 10.22033/ESGF/CMIP6.6842