LETTER • OPEN ACCESS

Steady but model dependent Arctic amplification of the forced temperature response in 21st century CMIP6 projections

ENVIRONMENTAL RESEARCH
CLIMATE

LETTER

Steady but model dependent Arctic amplification of the forced temperature response in 21st century CMIP6 projections

Stephanie Hay*, James A Screen and Jennifer L Catto

Department of Mathematics and Statistics, University of Exeter, 23 N Park Road, Exeter, Devon EX4 4QE, United Kingdom
* Author to whom any correspondence should be addressed.

E-mail: s.e.hay@exeter.ac.uk

Keywords: Arctic amplification, structural uncertainty, scenario uncertainy, internal variability

Supplementary material for this article is available online

OPEN ACCESS

RECEIVED
6 November 2023

REVISED
16 April 2024

ACCEPTED FOR PUBLICATION
23 April 2024

PUBLISHED
15 May 2024

Abstract

We examine sources of uncertainty in projections of Arctic amplification (AA) using the CMIP6 multi-model (MM) ensemble and single model initial-condition large ensembles of historical and future scenario simulations. In the CMIP6 MM mean, the annual mean AA ratio is steady at approximately 2.5, both in time and across scenarios, resulting in negligibly small scenario uncertainty in the magnitude of AA. Deviations from the steady value can be found at the low and high emission scenarios due to different root causes, with the latter being mostly evident in the summer and autumn seasons. Best estimates of model uncertainty are at least an order of magnitude larger than scenario uncertainty in CMIP6. The large ensembles reveal that irreducible internal variability has a similar magnitude to model uncertainty for most of the 21st century, except in the lowest emission scenario at the end of the 21st century when it could be twice as large.

1. Introduction

Arctic amplification (AA) of temperature change is an intrinsic part of a warming planet [1, 2]. It refers to how the Arctic surface warms at a higher rate than the global surface, and is usually defined as the ratio of Arctic mean surface warming to global mean surface warming [3], and can be sensitive to the exact definition of the Arctic [4]. AA has been found in paleoclimate observations and modelling [5–7], simplified models of the atmosphere [8, 9], satellite observations and reanalysis data [10–12], and in projections of future climate [13, 14]. The causes of AA are generally understood to arise from a set of positive feedbacks, including those related to sea-ice and snow cover [11, 15, 16], temperature [17], water vapour and clouds [18, 19], and heat transport [20–23]. Each of these feedbacks interact with each other in complicated ways, and the relative importance of each is debated [24].

Understanding and quantifying the factors that drive AA is important not only for explaining local change in the Arctic, but also, as it strongly influences the equator-to-pole temperature gradient, for predicting how the atmospheric general circulation of the planet responds to greenhouse warming. For example, model spread in projections of the North Atlantic eddy-driven jet can be linked to the ratio of upper-tropical warming to surface-based AA [25], and other jet metrics have been linked to lower-tropospheric AA [26], whilst storm tracks have been shown to respond to changes in the equator-to-pole temperature differences at upper and lower levels [27]. The impact of AA on lower latitudes remains a topic of debate [28–30], with some arguing that addressing model deficiencies, such as a too-weak eddy feedback, could help reconcile some of the differences between the observed link between AA and the mid-latitudes and the lack of link in models [31–33].

Inter-model spread of AA remains large even in state-of-the-art climate models [14]. Here we use CMIP6 models and seven single model initial condition large ensembles to systematically examine the sources of uncertainty in projections of AA: scenario uncertainty, model uncertainty, and internal variability [34], and whether these vary seasonally. Scenario-based, or radiative forcing, uncertainty, arises from our lack of knowledge of future greenhouse emissions. Model, or structural, uncertainty, arises from the various ways

individual modelling centres construct and tune climate models. Finally, internal variability is the unforced variability that arises due to natural chaotic variations in the earth system. Generally, internal variability is understood to be irreducible, whereas inter-model spread can be reduced by gaining insight in to, and improving the representation of, physical processes.

2. Data and methods

Using the complete set of CMIP6 models, listed in table 1, with historical and any of the four future emission scenarios (shared socioeconomic pathways (ssp) 126, 245, 370, and 585), we compute global mean and Arctic mean (defined as the region poleward of 66.5° N, consistent with recent work [4]) area-averaged surface temperatures for each model and scenario. The AA ratio is defined as the ratio of Arctic to global mean warming, relative to the reference period, which we choose here to be the preindustrial period and define it as the first 30 years of the historical run (1850–1879)

$$\mathrm{AA} = \frac{T_{\mathrm{Arc}}(t) - T_{\mathrm{Arc}}(t_{\mathrm{ref}})}{T_{\mathrm{Glb}}(t) - T_{\mathrm{Glb}}(t_{\mathrm{ref}})}. \tag{1}$$

For models with multiple ensemble members, we choose the first member only to compute the CMIP6 multi-model (MM) mean so as not to bias the results toward any particular model. All models are regridded to a 1×1 degree global grid before plotting, but area-averaged values are calculated on the model grid.

Seven of the models also provide large ensembles (table 1), where individual members differ from each other by only their initial state and are otherwise subject to identical forcing. We levy these ensembles alongside the CMIP6 MM ensemble to estimate model internal variability. With these, we also compute the AA ratio of trends over 30 year periods in all ensemble members:

$$\mathrm{AA} = \frac{\dot{T}_{\mathrm{Arc}}}{\dot{T}_{\mathrm{Glb}}}. \tag{2}$$

This AA metric does not require choice of reference period and can also be readily compared to reanalysis products. We choose three reanalyses: ERA5 [35], MERRA2 [36], and JRA55 [37] for which we also calculate the AA ratio of trends. For JRA55, we compute the trends over the periods 1960–1990, 1970–2000, 1980–2010, and 1990–2020, whilst for MERRA2 and ERA5 we compute the trends only over the last two periods due to data availability.

3. Results

3.1. Scenario uncertainty

In the interest of minimizing internal variability and model differences when determining scenario uncertainty, we first choose two long, well-separated time periods and compute the CMIP6 MM mean Arctic and global mean warming at the end of the 21st century, 2071–2100, relative to the reference period. Figure 1 provides the MM-mean pattern of the local amplification of warming at this time. The warming at each grid point scaled by the global surface warming reveals that each projection scenario produces consistent patterns and magnitudes of local warming amplification. Northern Hemisphere land masses and the North Pacific warm by between 1–2 times the global mean, while higher latitude land and Arctic ocean regions warm by between two and more than four times faster than the global mean, with the greatest amplification of warming found over the eastern Arctic ocean in the vicinity of the Barents Sea. The highest values spread toward the central Arctic in the higher emission scenarios, likely due to more pan-Arctic sea-ice loss. No regions are found in the MM mean that exhibit a negative local amplification in the Northern Hemisphere, though some individual models do in the lower emission scenarios, generally found in the North Atlantic warming hole region (supplemental figure 1). Figure 1 also shows more amplified warming of the Northern Hemisphere, particularly over the land masses, in the larger forcing scenarios.

As AA is usually defined as a ratio, there can be issues, particularly when the denominator approaches zero [38]. Here we note another consequence of this definition: pattern scaling. By normalizing local temperature change by global mean surface temperature change, we leverage an empirical relationship previously observed in climate models whereby the pattern of forced response in climate models is proportional to the magnitude of surface warming, independent of the details of the forcing [39, 40]. A consequence of this is that the magnitude of MM-mean AA is effectively independent of projection scenario, or in other words, scenario uncertainty is negligible over the timescales and scenarios considered here, to the extent that Arctic warming scales linearly with global warming. This is also true to a lesser extent for individual models (table 1), but as these are given by a single ensemble member it is not clear whether

Table 1. CMIP6 models, scenarios, and large initial condition ensembles used in this study. Emissions scenarios correspond to 1: 2.6 Wm^{-2}, 2: 4.5 Wm^{-2}, 3: 7.0 Wm^{-2}, and 5: 8.5 Wm^{-2} radiative forcing at the end of the 21st century. Values of AA given here represent the mean across the 21st century, relative to the reference period 1850–1879.

Model	Scenarios	AA	# LE members	LE AA
ACCESS-CM2	1, 2, 3, 5	2.19, 2.20, 2.26, 2.13		
ACCESS-ESM1-5	1, 2, 3, 5	2.43, 2.49, 2.39, 2.47	39, 40, 40, 40	2.53, 2.46, 2.42, 2.47
AWI-CM-1-1-MR	1, 2, 3, 5	2.53, 2.51, 2.43, 2.46		
BCC-CSM2-MR	1, 2, 3, 5	2.28, 2.31, 2.36, 2.40		
CAMS-CSM1-0	1, 2, 3, 5	1.94, 1.76, 1.88, 1.87		
CanESM5	1, 2, 3, 5	2.53, 2.44, 2.34, 2.40	50, 50, 50, 50	2.71, 2.64, 2.58, 2.59
CanESM5-CanOE	1, 2, 3, 5	2.68, 2.69, 2.68, 2.69		
CAS-ESM2-0	1, 2, 3, 5	2.24, 2.19, 2.01, 2.04		
CESM2	1, 2, 3, 5	2.64, 2.65, 2.55, 2.62	/, /, 100, /	/, /, 2.84, /
CESM2-WACCM	1, 2, 3, 5	2.31, 2.41, 2.44, 2.32		
CMCC-CM2-SR5	1, 2, 3, 5	2.93, 2.86, 2.69, 2.77		
CMCC-ESM2	1, 2, 3, 5	2.77, 2.78, 2.57, 2.59		
CNRM-CM6-1	1, 2, 3, 5	2.24, 2.19, 2.19, 2.24		
CNRM-CM6-1-HR	1, 2, 3, 5	2.59, 2.51, 2.47, 2.54		
CNRM-ESM2-0	1, 2, 3, 5	1.73, 1.88, 1.95, 2.11		
E3SM1-1	5	2.74		
EC-Earth3	1, 2, 3, 5	3.03, 2.94, 2.96, 2.91	59, 73, 58, 59	3.26, 3.13, 3.03, 3.02
EC-Earth3-AerChem	3	2.87		
EC-Earth3-CC	2, 5	3.39, 3.36		
EC-Earth3-Veg	1, 2, 3, 5	3.35, 3.32, 3.23, 3.32		
EC-Earth3-Veg-LR	1, 2, 3, 5	3.19, 3.02, 3.05, 3.00		
FGOALS-f3-L	1, 2, 3, 5	2.71, 2.68, 2.64, 2.66		
FGOALS-g3	1, 2, 3, 5	2.35, 2.43, 2.54, 2.49		
FIO-ESM2-0	1, 2, 5	2.74, 2.61, 2.71		
GFDL-CM4	2, 5	2.45, 2.40		
GFDL-ESM4	1, 2, 3, 5	1.66, 1.77, 1.61, 1.66		
GISS-E2-1-G	1, 2, 3, 5	2.52, 2.43, 2.36, 2.43	16, 25, 26, 17	2.50, 2.45, 2.40, 2.36
HadGEM3-GC31-LL	1, 2, 5	2.39, 2.38, 2.37		
HadGEM3-GC31-MM	1, 5	2.79, 2.61		
INM-CM4-8	1, 2, 3, 5	1.74, 1.87, 1.83, 1.93		
INM-CM5-0	1, 2, 3, 5	2.57, 2.31, 2.47, 2.45		
IPSL-CM5A2-INCA	1, 3	1.97, 1.87		
IPSL-CM6A-LR	1, 2, 3, 5	2.78, 2.78, 2.66, 2.65		
KACE-1-0-G	1, 2, 3, 5	2.60, 2.63, 2.52, 2.51		
KIOST-ESM	1, 2, 5	2.53, 2.51, 2.33		
MIROC6	1, 2, 3, 5	3.36, 3.21, 3.25, 3.07		
MIROC-ES2L	1, 2, 3, 5	2.36, 2.35, 2.35, 2.37	10, 30, 10, 10	2.77, 2.81, 2.72, 2.70
MPI-ESM1-2-HR	1, 2, 3, 5	2.61, 2.53, 2.63, 2.41		
MPI-ESM1-2-LR	1, 2, 3, 5	2.63, 2.52, 2.45, 2.53	30, 30, 30, 30	2.69, 2.66, 2.62, 2.63
MRI-ESM2-0	1, 2, 3, 5	2.98, 2.93, 2.79, 2.95		
NESM3	1, 2, 5	3.26, 3.22, 3.24		
NorESM2-LM	1, 2, 3, 5	2.77, 2.93, 3.04, 2.83		
NorESM2-MM	1, 2, 3, 5	2.28, 2.17, 2.07, 2.14		
TaiESM1	1, 2, 3, 5	2.50, 2.39, 2.34, 2.42		
UKESM1-0-LL	1, 2, 3, 5	2.47, 2.41, 2.34, 2.37		

differences are due to internal variability or scenario uncertainty. Averaged over the Arctic as defined above, we find values of AA that vary from only 2.46 to 2.63 between greenhouse warming scenarios. Inter-model standard deviation is reduced with increasing emission forcing, from 0.41 in ssp126 to 0.27 in ssp585. The MM-mean AA scenario uncertainty can be further reduced if we include only the models for which data from all four scenarios are available. This sampling of the CMIP6 models gives AA values of 2.53 ± 0.40 , 2.50 ± 0.37, 2.47 ± 0.37, and 2.48 ± 0.34 at the end of the century. We investigate the small decrease in MM-mean AA in the greater emissions scenarios in further detail in the next sections.

A reduction in AA from the lowest to highest forcing scenario is not simulated at the individual model level when considering only a single realization. In table 1, the average of AA across the 21st century for each model and scenario shows that no individual model for which all four scenarios were available exhibits a monotonic decrease in AA with forcing, whilst six with either two or three of the scenarios do. On the other hand, the majority of models have their lowest AA ratio in ssp370, as reflected by the MM mean. Two models monotonically increase AA with forcing and many exhibit no clear tendency. On average, the across-scenario

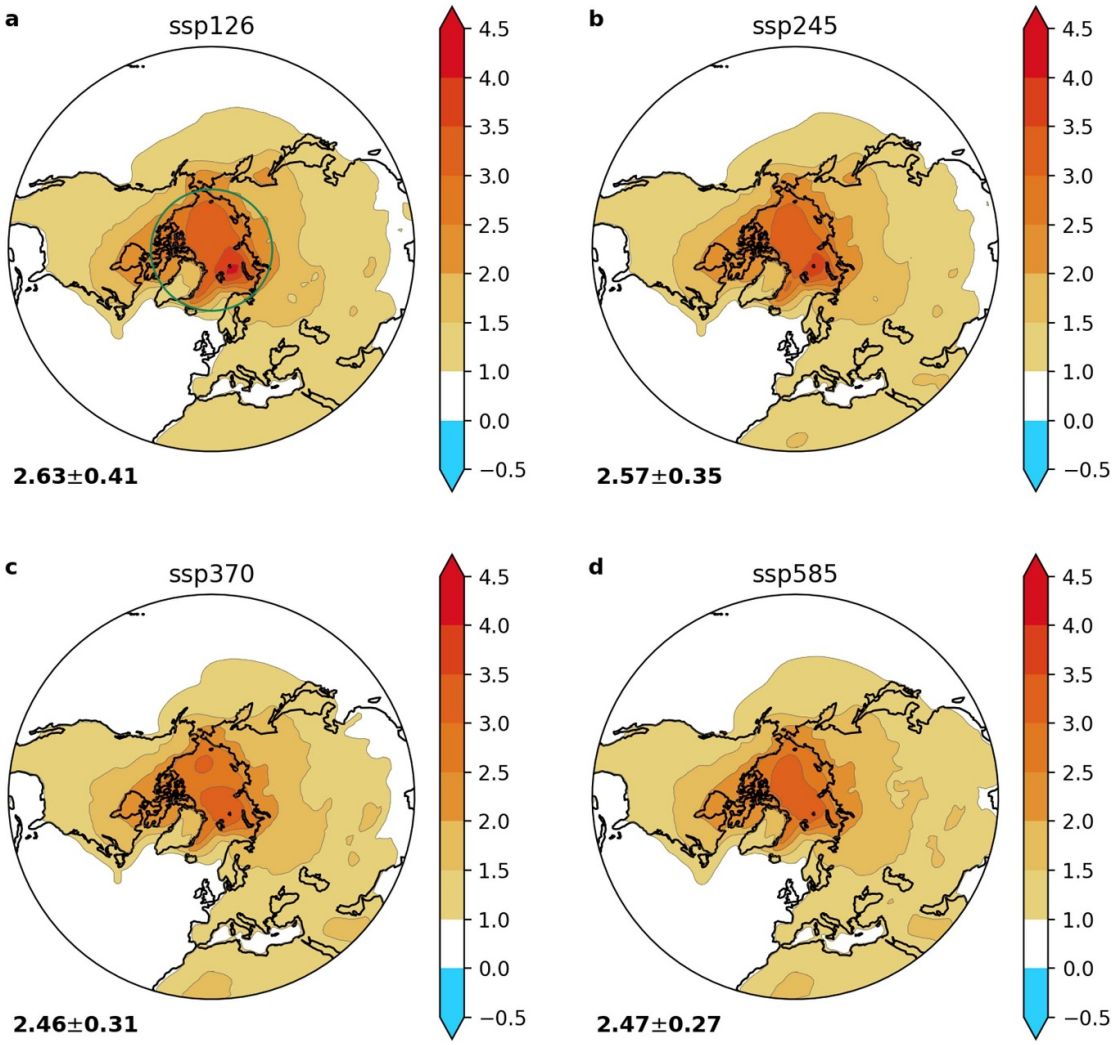

Figure 1. Annual-mean CMIP6 MM-mean local amplification of warming relative to global mean surface warming between the end of the 21st century (2071-2100) and the preindustrial period (1850-1879) for (a) ssp126, (b) ssp245, (c) ssp370, (d) ssp585. Area-weighted average of local amplification poleward of 66.5° N (region indicated by the green circle in (a)) and inter-model standard deviation is indicated in the lower-left of each panel.

variation in AA in an individual model varies between 0.01 and 0.38, with an average of 0.16. This result differs from other studies that have found scenario differences either under larger forcing than that considered here [41] or with large ensembles of a single model [42]. This points to the potentially significant influence of internal variability in a single realization of a model. Indeed, another study has found that one realization per model in CMIP6 does not show a significant difference between scenarios, and at least 10 realizations is required to detect scenario differences [42]. The values of AA from each available scenario in the large ensembles are given in the final column of table 1. Although only one of the six models (GISS-E2-1-G) simulates a monotonic decrease of AA with forcing, the other five have generally higher AA in the two lower forcing scenarios than in the two higher ones. Again, ssp370 exhibits typically the lowest AA. We discuss internal variability in more detail in section 3.3.

3.2. Model uncertainty

A second consequence of the scaling of Arctic warming with global warming is that MM mean AA has little variation in time as well as across scenarios (figure 2). We determine the AA from the temperature anomaly for eight 30 year periods throughout the 21st century, relative to the preindustrial period. In all scenarios, inter-model spread is reduced in time from its largest value of 3.3 (from 0.9 in GFDL-ESM4 to 4.2 in NorESM2-LM) in the period 2000–2029. However, the spread remains significant even at the end of the highest forcing scenario of between 1.9 (CAMS-CSM1-0) and 3.3 (MIROC6). It is worth noting that any single model may not show the constant AA in time found in the MM-mean, and whether that is due to model difference or internal variability is a point we return to in section 3.3. The inter-model spread when the forcing is largest might be understood to be the best estimate of inter-model spread that can be obtained

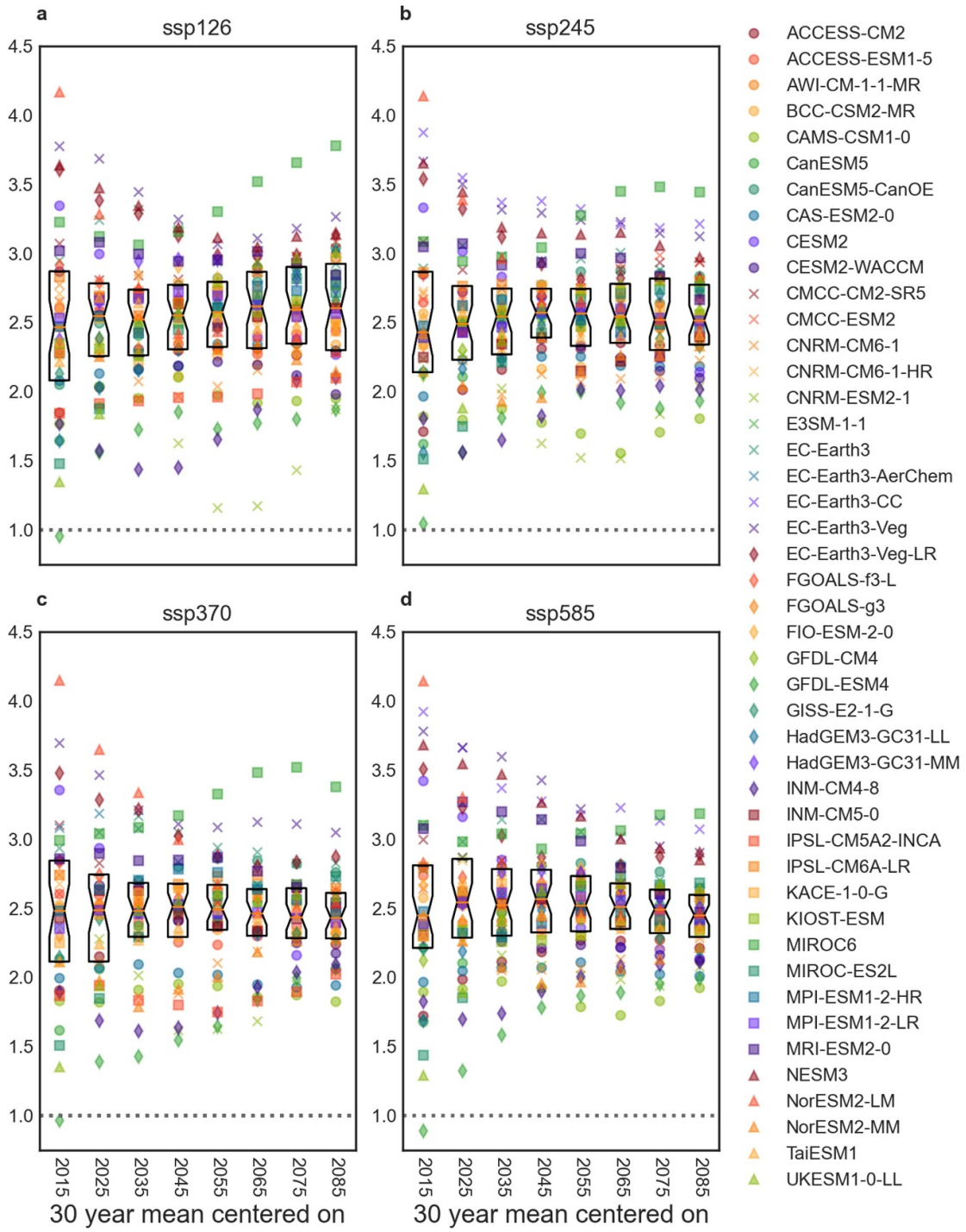

Figure 2. Annual-mean CMIP6 AA for selected 30 year time periods relative to the pre-industrial period (1850–1879) for (a) ssp126, (b) ssp245, (c) ssp370, (d) ssp585. The dashed grey lines indicate where AA = 1.0.

from the CMIP6 mean, as it maximises the signal-to-noise ratio, rather than the larger values found in weaker emissions scenarios which may include a large contribution from internal variability.

The linearity upon which the scaling depends begins to break down by the end of the 21st century when looking at the lowest and highest emission scenarios, but for different reasons. MM-mean AA begins to increase slightly in ssp126, attributable to a slowing of the increase, or even a decrease, in some models in global mean surface temperature as radiative forcing peaks mid-century and then decreases monotonically in that emissions scenario (supplemental figure 2). The reason that AA increases is that while warming ceases, local feedbacks in the Arctic, such as those related to the presence of sea ice, continue to amplify temperature change there for some time [42]. On the other hand, in ssp585, the MM-mean AA decreases after the middle of the century, which has been argued to arise from the loss of perennial sea-ice [41]. As the Arctic transitions to a seasonally ice-free state in some models, this results in a weakening of the sea-ice related feedbacks,

reducing amplification of warming over the Arctic [41]. Nevertheless, the variance in MM-mean AA across time is only 0.03 in both ssp126 and ssp585, a factor of 100 smaller than the inter-model spread. We note that one study has found a time dependence for AA [43], most evident when looking past the end of the 21st century. This suggests that, past this point, we may expect a further break down of the conditions necessary for maintaining approximately constant AA.

AA varies across seasons, reaching its maximum in wintertime and minimum in summer, generally understood to be linked by heat storage in the warm season that is released in the cold season [44]. Heat storage is increased as more open ocean is created by sea-ice loss in summer, which generates a feedback of increased latent heat transport in summer that leads to more sea-ice loss [45]. The effective heat capacity of ice, differences in lapse rate feedbacks due to the absence of summertime thermal inversions, latent and and dry heat transports, and differences in seasonal energy exchanges have all been linked to this seasonal asymmetry, and to model spread in simulated AA [45–47].

Given the scalability of AA previously shown, we pool the scenarios and each of the time periods considered in figure 2 and plot as a function of global warming (figure 3), for each season. The CMIP6 MM-mean AA in DJF is, on average, 3.33–3.42 for all four scenarios and little variation is apparent in the mean with increasing global warming or time. At small amounts of global warming, the models simulate between 1 and 6 times more warming in the Arctic than globally, and for the models which warm the most (5 °C–6 °C globally), there is a convergence of AA toward the MM mean. Whilst the spread in MAM is slightly smaller at low amounts of warming, there is a similar convergence toward the mean of 2.24–2.29 for the greatest amounts of warming. Neither of these seasons exhibit a decreasing AA in time for ssp585 that was seen in the global mean. Convergence toward the MM mean at higher amounts of warming indicates that the models converge toward a forced value of AA as the influence of internal variability decreases.

It is SON that accounts for the decrease in AA with time in the higher emission scenarios. The decrease is also evident in ssp370, which was not found in the annual mean. Inter-model spread at low values of global warming is largest in SON (0.5–6.5), and while AA remains approximately steady at 3.11 in ssp126, it begins to decrease slightly in the other emissions scenarios in time. There is not a clear convergence toward the MM mean as in DJF and MAM. On the other hand, AA increases in JJA with scenario and time. AA barely exceeds 1.0 at low values of warming, but increases up to 1.4 at higher values. A reduction in the seasonal amplitude of AA with increased CO_2 forcing was also found in a study using a single model [41]. That, and another study [48], also found that the peak in AA shifted from late autumn to early winter over the 21st century, with our results here suggesting that is caused by a decrease in autumn AA, while winter AA remains approximately constant.

We briefly explore the role of changing mean states in these two seasons in figure 4, first focusing on the importance of the presence of sea ice in summer. Looking only at the end of the century, the role of sea-ice on AA in JJA is clear. By separating the models into two groups, those that remain partially ice-covered (in purple) and those which are 'ice-free' ($<$1 million km^2 remaining ice on average in JJA) (in gold) there is a clear separation of the distribution of AA for the ssp370 and ssp585 scenarios (figure 4(a).) We hypothesise that as sea-ice decreases, incoming solar radiation at the surface can go to warming the surface air as it no longer used largely as input to melt the sea ice and warm the ocean surface, resulting in increased AA in JJA. Indeed, turbulent heat fluxes averaged poleward of 66.5° N in summer increase in the CMIP6 mean by 0.05 W m^{-2} per year over the 21st century in ssp585, a potential physical mechanism for increasing AA in JJA.

In SON, somewhat surprisingly, there is only a small effect on the distribution of AA when separating models into those that are ice-free or partially ice-covered, however only nine models remain partially ice-covered in ssp370, and six models in ssp585 therefore a large influence of ice-free models is present when considering all models. A smaller AA in higher emissions scenarios than in lower emissions scenarios has generally been assumed to be driven by a reduced ice-albedo feedback [41], but here we find only a modestly larger AA ratio in ssp370 and ssp585 at the end of the century for the partially ice-covered compared to the ice-free models. In fact, in contrast to other seasons, model spread in the AA ratio in SON is uncorrelated with sea-ice area, even in scenarios or during periods where all models remain partially ice-covered (figure 4(c)). The increasing correlation between AA ratio and sea-ice area in JJA with forcing in figure 4(c) is consistent with the findings in figure 4(a). Interestingly, model spread in sea-ice loss and AA ratio become less correlated in time, with variance explained dropping from approximately 50% to 20% (figure 4(d)).

A second feedback that depends on the mean state is that of the Planck feedback. This is a negative feedback that arises due to the fact that a warmer planet radiates energy back to space more effectively. Whilst this feedback negative everywhere, it is less negative in the relatively colder poles, and hence contributes to AA [17]. As the equator-to-pole temperature difference is reduced, its contribution to AA is reduced [49]. In all seasons but SON, model spread in AA at the end of the century correlates with model spread in the temperature difference between the pole and the tropics in each of the scenarios separately, explaining between 19% and 66% of the variance, depending on scenario and season (figure 4(e)). In SON, on the other

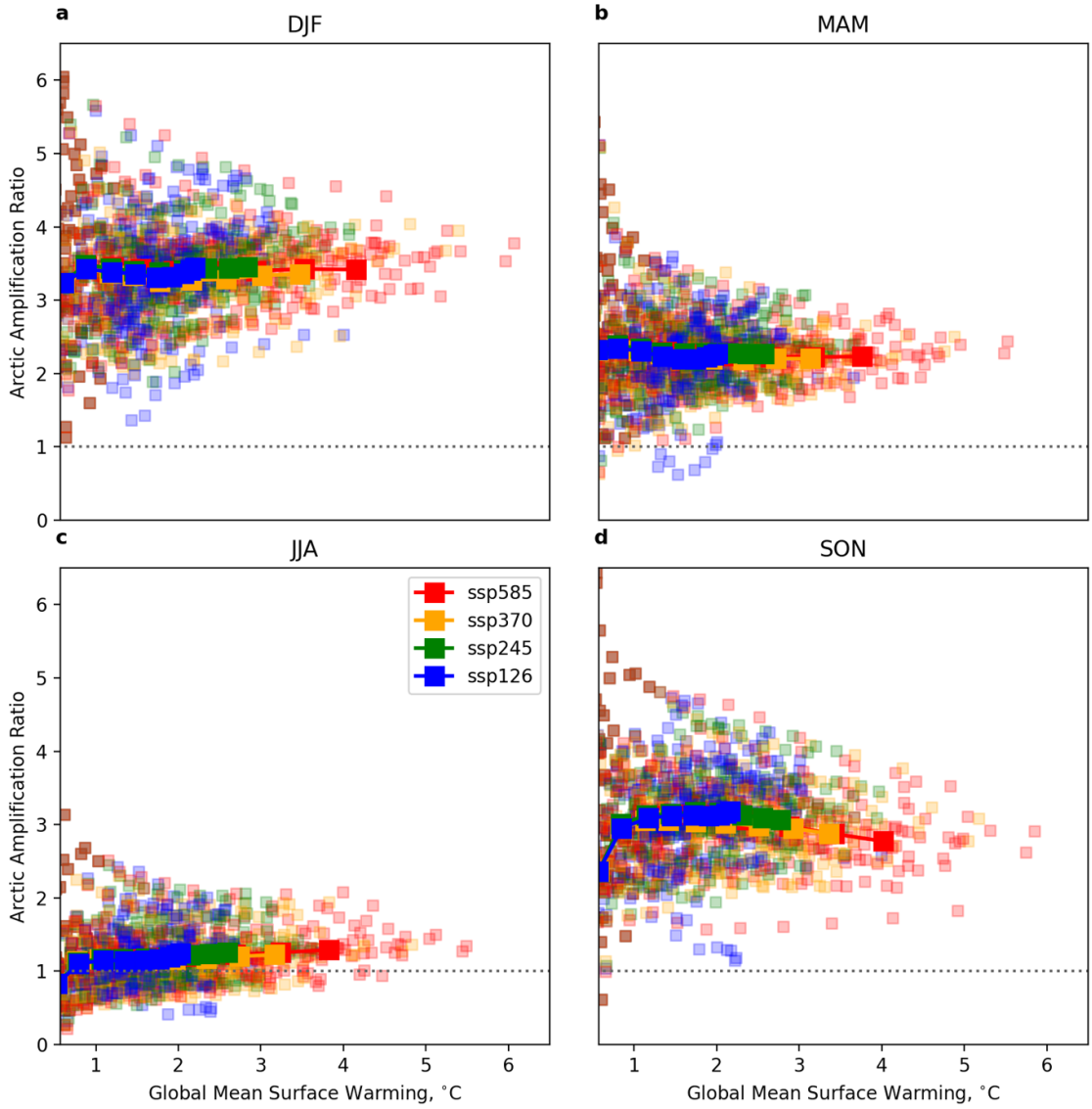

Figure 3. Global mean surface warming vs AA (relative to preindustrial) for all the models and 30 year time periods in figure 2 across the seasons (a) DJF, (b) MAM, (c) JJA, (d) SON. The dashed grey lines indicate where AA = 1.0.

hand, there is no statistically significant correlation across models in any scenario, and we see the correlation switch sign in ssp585. In other words, in SON only, the smaller the magnitude of equator-to-pole temperature difference, the smaller the AA. The equator-to-pole temperature difference is not correlated with sea-ice loss at the end of the 21st century in SON.

The time-dependence of the relationship between the AA ratio and the change in the equator-to-pole temperature difference (figure 4(f)) shows that the scenarios begin to separate from each another in the middle of the 21st century. Although in both JJA and SON the correlations decrease over time, the drop in variance explained is particularly clear in SON in the highest forcing scenario.

The steadiness of AA in DJF and MAM extends throughout the troposphere. The time evolution of AA at the model pressure levels up to 400 hPa in JJA and SON is shown in figure 5. AA is defined at each level as the ratio of Arctic to global warming at that level. AA has been found to extend in to the midtroposphere [50], and it has been argued that AA away from the surface is more likely to have an influence on the circulation of the lower latitudes (e.g. [26]). In JJA, larger values of AA first appear in the mid-troposphere in the early-to-mid 21st century and then reach the surface only near the end of the century. The mid-tropospheric warming in JJA is generally understood to arise from remote influences [50, 51]. The downwelling longwave radiation from this warmer air aloft in conjunction with incoming solar radiation and lower sea-ice is what drives the increase in surface AA with time in JJA.

The bottom-heavy structure of AA in SON is evident. In all scenarios, the depth of AA increases in time; warming is not amplified above 550 hPa initially, but the amplified warming spreads up to 400 hPa by the

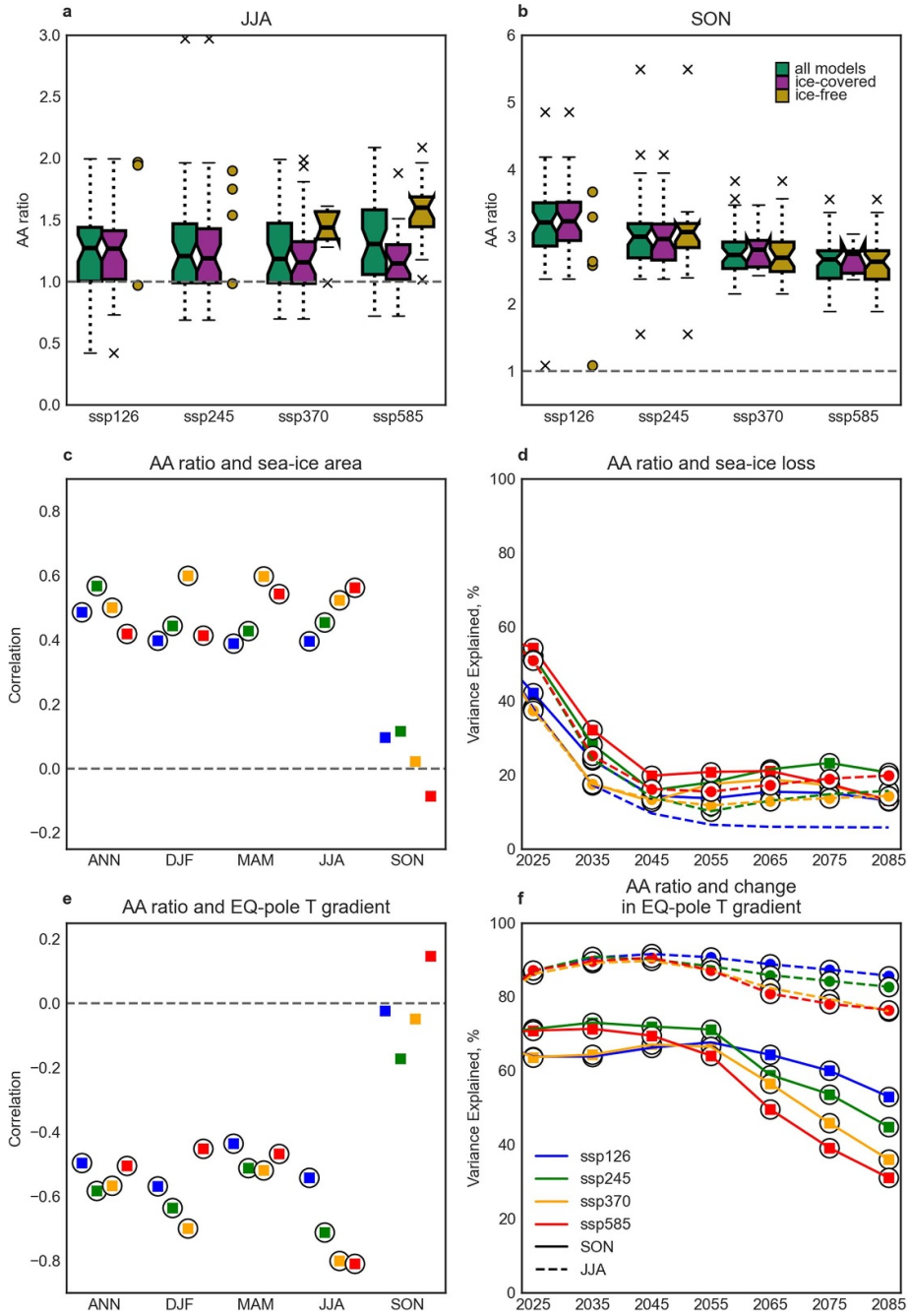

Figure 4. AA in 2071–2100 relative to 1850–1879 in each scenario for all models in green, for models with a partially ice-covered Arctic in purple, and those with <1 million km^{-2} sea ice in gold for (a) JJA, and (b) SON. Dashed grey lines indicate where AA $= 1.0$. (c) The inter-model correlation between the sea-ice area and the AA ratio in the four scenarios across seasons, where statistically significant correlations are denoted with a black circle. The thin dashed grey line indicates zero correlation. (d) JJA (dashed coloured lines) and SON (solid coloured lines) variance explained over time by model spread in the loss of sea-ice area and the AA ratio for the four scenarios. (e) and (f) as in (c) and (d), but for the equator-to-pole temperature difference.

end of the 21st century. This could be an important factor in the impact of AA on lower latitudes, as deeper AA in models has been shown to be related to a larger thermodynamic response in the midlatitudes [52, 53]. Surface AA is relatively constant in ssp126 and ssp245, but notably shallower in ssp370, and becoming shallower in ssp585 with time. However, at upper levels, AA continues to increase with time in all scenarios.

3.3. Internal variability

Next, we examine the internal variability in AA in large ensembles (figure 6). To do so, we calculate 30 year linear trends in Arctic and global mean temperature and take their ratio only where the trend in each is statistically significant at the 95% level via a Student's t-test so as to prevent spurious values of AA due to a small numerator or denominator. We use both the CMIP6 ensemble and seven large initial condition

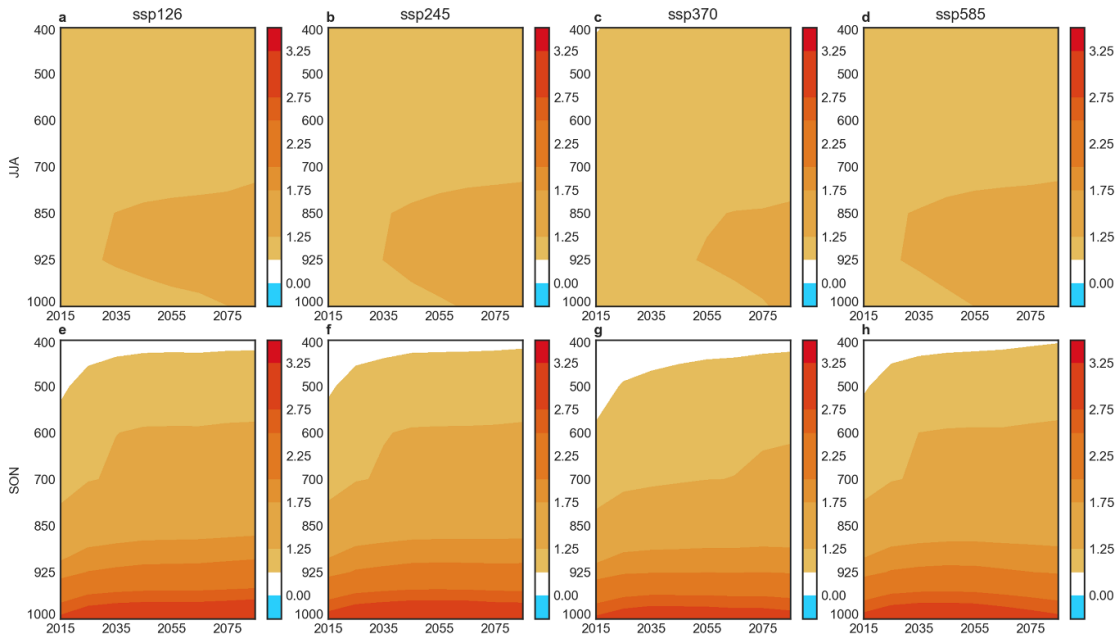

Figure 5. The CMIP6 MM mean vertical distribution of AA relative to preindustrial in JJA for (a) ssp126, (b) ssp245, (c) ssp370, and (d) ssp585. (e)–(h) as in (a)–(d) but for SON.

ensembles of varying size (table 1) and plot the linear trend in global temperature against the magnitude of AA.

As in figure 4, we note that, in some ensembles more clearly than in others, AA converges toward a value at high magnitude of warming. The value to which the different ensembles converge varies from 2.0–2.5. For 30 year periods with small global warming trends, variability can be quite large in some ensembles, particularly in the ssp126 scenario, in which mitigation reduces global warming to near zero. In all cases, these 30 year periods are at the end of the century and the large values of AA arise from the slowing of global warming while local feedbacks continue to amplify the Arctic warming. Generally, the convergence of AA values at higher global warming trend suggests that imprint of internal variability depends on how quickly greenhouse warming is proceeding, and that the spread in magnitude of AA can be quite large at low warming rates. Furthermore, the converging AA at faster warming seems to imply a preferred ratio of Arctic to global warming in each model and it is not far off from the 2.5 found for all CMIP6 models. At moderate values of global warming, internal variability is quite similar across models, and we find values of AA of just below 1 to approximately 4, similar to the inter-model spread. It is interesting that even during periods of moderate warming, the AA ratio is not always greater than 1. Variability when considering shorter time periods for trend calculations is even larger.

To better assess how much of the model spread in figure 2 can be attributed to internal variability, we compare the ensemble-mean AA ratio for each of the large ensembles and the CMIP6 MM mean relative to preindustrial in ssp370, for which all large ensembles have members, in figure 7(a). We note the approximate flatness of the AA curves in time in CMIP6 is replicated by most large ensembles, except for CanESM5 which increases until 2040 before flattening, and CESM2, which decreases over the same time period before becoming flat. The reason for the divergent behaviour in these two models is unclear but a future study to understand it may reveal important differences in processes leading to AA. Internal variability in each ensemble is given by the standard deviation across ensemble members and is represented by the shaded region surrounding each model's curve. We find that in all large ensembles, internal variability decreases in time. At the end of the 21st century, the large ensembles' means are spread approximately equally around the CMIP6 MM mean, from 2.2 in GISS-E2-1-G to in 2.8 in EC-Earth3.

If instead we define AA as in figure 6 from 30 year trends, we can plot the AA ratio, again for ssp370, as a distribution over the 21st century for each ensemble (figure 7(b). This gives a very similar estimate for the CMIP6 mean (2.58), as well as for the model uncertainty (2.11–2.94). Internal variability is provided by the width of the distribution here, and we can see how over individual 30 year periods, this is larger than model uncertainty, defined as the spread in ensemble means.

We use the large ensembles to additionally estimate the relative contributions of scenario and model uncertainty, and internal variability to our estimates of AA. Mean scenario uncertainty is given by the MM

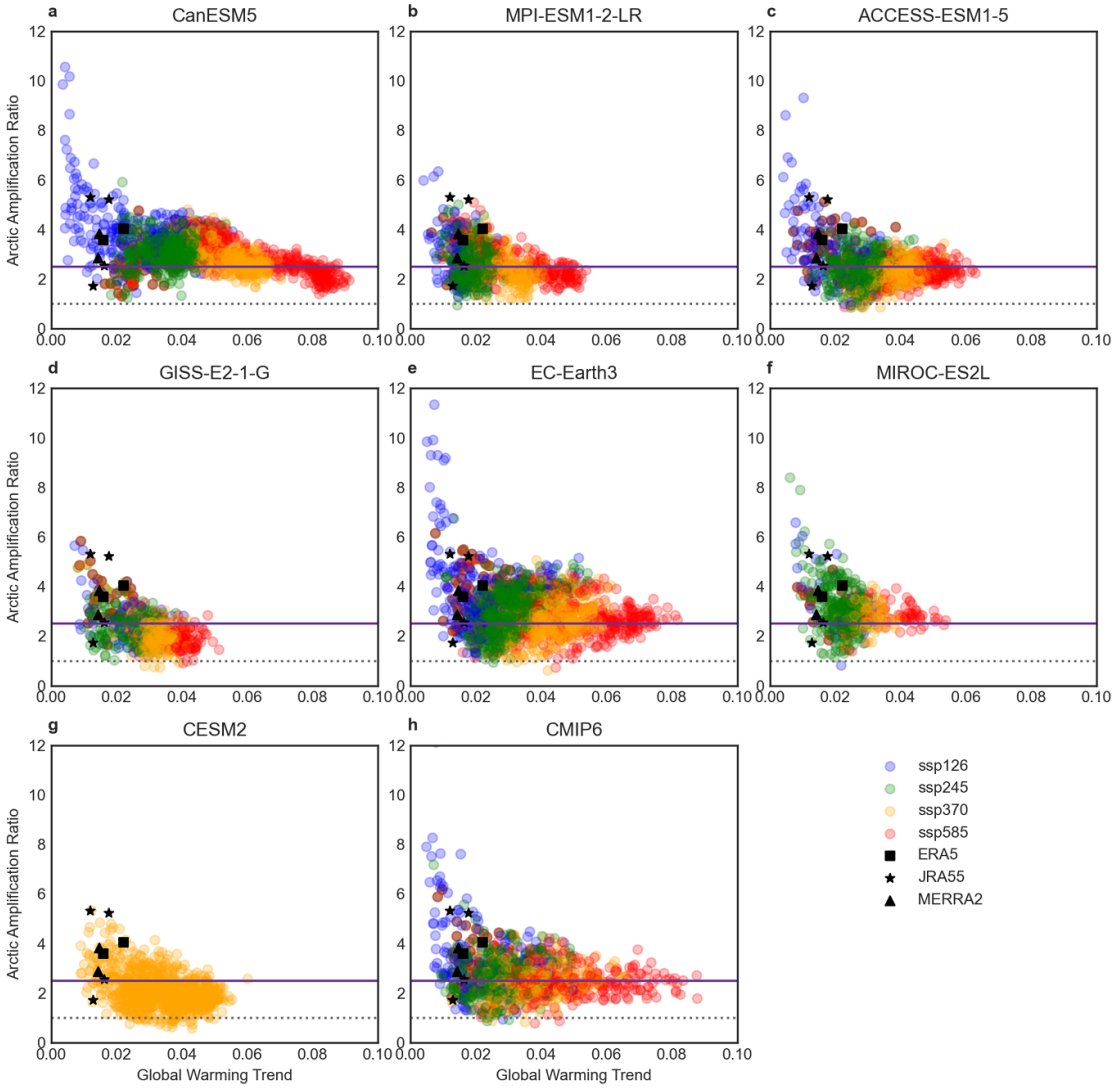

Figure 6. Global mean surface warming trend vs AA determined as the ratio of 30-year trends in Arctic and global warming for the periods considered in figure 2 for each ensemble member and scenario of (a) CanESM5 (b) MPI-ESM1-2-LR , (c) ACCESS-ESM1-5, (d) GISS-E2-1-G, (e) EC-Earth3, (f) MIROC-ES2L, (g) CESM2, and (h) CMIP6. AA over the same periods from three reanalysis products are indicated with black markers. The dashed line indicates where AA = 1.0, and the purple line indicates where AA = 2.5.

ensemble-mean standard deviation of AA across the scenarios, and the different single model ensemble means provide the spread about this estimate. Model uncertainty is defined as the standard deviation of the ensemble- and scenario-mean AA across models, and the different scenarios provide the spread about this estimate. Finally, internal variability is estimated by the standard deviation across members of an ensemble, given a model and scenario, and different models and scenarios provide the spread about this estimate.

Using the large ensembles to compare sources of uncertainty in projections of the surface-based AA ratio either defined relative to preindustrial or by 30 year trends leads to different conclusions as to the largest sources of uncertainty in time (figure 7(c) compared to (d)). Defining AA relative to to preindustrial (figure 7(c)), scenario uncertainty is small, but grows over time, as suggested by our earlier analysis and in other studies (e.g. [41, 42]). Internal variability is initially the largest source of uncertainty, although an estimate of internal variability varies depending on the ensemble considered and this gives a large spread around our estimate, particularly in the early 21st century. Nevertheless, all ensembles agree that this source of uncertainty is reduced in time, eventually reaching a value just larger than the uncertainty arising from scenario uncertainty. Model uncertainty is approximately constant in time, becoming the largest source of uncertainty by the end of the 21st century.

On the other hand, when we use shorter periods to define AA as a ratio of trends (figure 7(d)), internal variability remains the largest source of uncertainty throughout most of the 21st century. Scenario uncertainty grows substantially in the latter half of the century, exceeding that of internal variability by the end of it. This is mainly due to ssp126, where large values of AA are found when global warming slows much

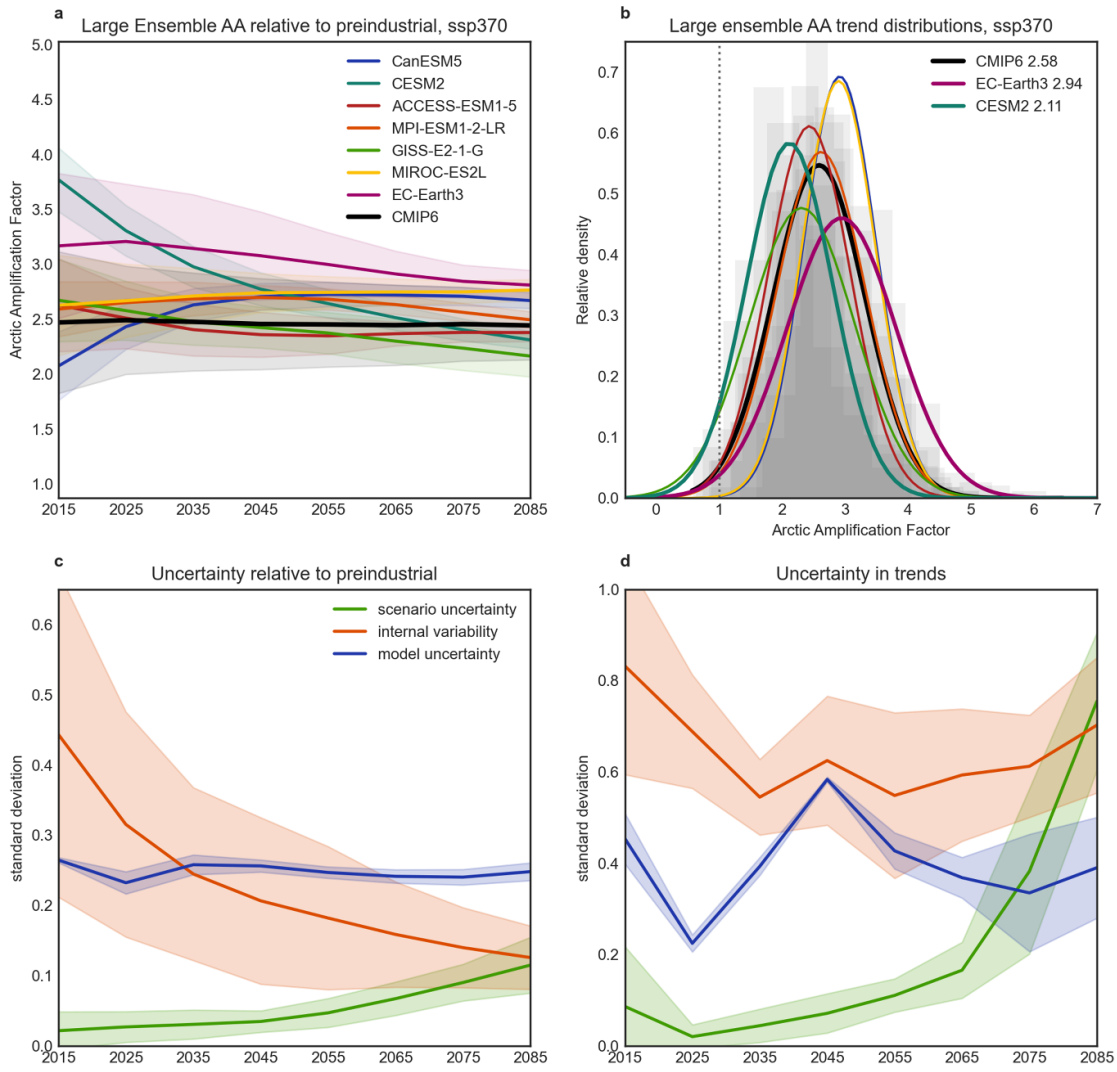

Figure 7. (a) Timeseries of mean AA relative to preindustrial in ssp370 for each of the large ensembles and the CMIP6 mean. (b) The distribution of AA determined by 30 year trends for all 30 year periods with statistically significant AA in ssp370 for all large ensembles and the CMIP6 mean. The dashed line indicates where AA = 1. (c) Shows the uncertainty in estimates of AA relative to preindustrial, partitioned in to scenario uncertainty (green), internal variability (orange), and model uncertainty (blue). (d) as in (c) but for uncertainty in AA determined by 30 year trends.

more than Arctic warming does in some models (figure 6, supplemental figure 2, and [54]). This is also the cause of the spread around estimates of model uncertainty at this time, as not all models simulate these very high values of AA.

4. Discussion

When averaging across a sufficient number of models or by looking at large enough forcing (global warming), in the annual mean, the Arctic warms by approximately 2.5 times the global mean. This number varies from the summer value of just greater than 1.0 to the winter value of approximately 3.3. Scenario uncertainty is found to be negligibly small largely due to the linear scaling of Arctic warming with global warming, but begins to grow near the end of the 21st century as mean-state dependence of the feedbacks that amplify temperatures in the Arctic are altered sufficiently. This is particularly apparent when calculating the AA ratio as trends over 30 year periods. Past this point, the steady nature of AA found here may be invalid (e.g. [43]. In particular, the sea-ice state impacts AA in the summer and potentially to a lesser degree in the autumn, and the reduced equator-to-pole temperature difference may lead to a decrease in AA in the autumn. Whilst model uncertainty of AA is an order of magnitude larger than scenario uncertainty at the start of the 21st century, it nevertheless remains approximately 3 times as large at the end of the century. Finally, uncertainty due to internal variability depends on the rate of warming, and can be quite large at low rates of warming, such as those found at the end of the 21st century in low emissions scenarios, or at present

day warming. At higher rates of warming, where the signal is maximised relative to the noise of internal variability, initial condition ensembles tend to converge toward a value of between 2.0–3.0 times more warming in the Arctic than globally.

A best estimate value of an annual mean AA of approximately 2.5 seems to emerge from this analysis repeatedly. Additionally, a similar value was found over 21 000 years of a paleoclimate simulation [5]. Another paleoclimate study based on observations [6] found a somewhat larger value of AA over four different warming periods, but one that was consistent across periods. We ask whether the value we find here of approximately 2.5 is reasonable, and if the models are accurately representing AA by comparing to three reanalysis products, ERA5, JRA55, and MERRA2. By examining the same 30 year periods for which the trends were calculated over for figure 6, AA in reanalysis varies from 1.72 to 5.32. For the same 30 year period of 1990–2020, AA is 3.82, 4.05, and 5.23 in the three reanalyses. This is quite a large spread despite small differences in global warming. The combination of uncertainty in the numerator, due to less well-constrained estimate of Arctic warming (0.056, 0.092, 0.089) in reanalysis due to observational difficulties in the harsh environment, and a small denominator (0.015, 0.018, 0.022) creates this surprisingly large spread. However, these values of AA all fit within the distribution of AA for most ensembles, suggesting that model internal variability is capturing the spread seen in observations. Due to high internal variability in AA, particularly at the relatively low values of warming so far observed, it might be very difficult to constrain AA using observations and thus understand how to reduce inter-model spread.

Finally, the inter-model correlations of AA with other metrics (e.g. sea-ice area, sea-ice loss, equator-to-pole temperature differences) shown in figure 4 are fairly low, and decrease in time. The loss of sea-ice area below some critical value needed to maintain sea-ice feedbacks could explain the decorrelation with AA over the 21st century in the larger forcing scenarios, but few models go 'ice-free' in the lower forcing scenarios. This, together with the low correlation between the equator-to-pole temperature difference and AA suggests caution in using this particular metric of AA if you are interested in understanding how Arctic change influences atmospheric circulation, particularly that which depends on local temperature gradients. Other metrics of AA that use ratios of temperature (Arctic warming relative to mid-latitude warming or tropical warming, for example) can be expected to exhibit a similar flat curve in time due to the linear scaling of temperature and would suffer similar problems when the denominator approaches zero [38], but metrics that define AA as a difference between temperature in the Arctic and elsewhere may be more suitable.

In conclusion, we have used the available CMIP6 historical and four future warming scenarios to look at the different sources of uncertainty or spread in the AA ratio: model uncertainty, scenario uncertainty, internal variability. For the first time, we take advantage of the linear scaling of Arctic temperature with global warming and show that, across the CMIP6 models, the pattern and magnitude of annual mean Arctic warming is nearly independent of scenario in the 21st century. This does not imply that the Arctic does not undergo large changes in time or across scenarios, but rather that its change relative to the global temperature remains approximately constant. This allows us to pool the scenarios to demonstrate a convergence, given large enough warming, in CMIP6 and in the large ensembles toward a best estimate value for AA. This value does not depend strongly on whether AA is defined from long-term changes or shorter-term trends, but is subject to model uncertainty.

Model uncertainty emerges as the largest source of uncertainty, whilst scenario uncertainty is very small and plausible physical mechanisms can be leveraged to understand how it varies in some seasons. For example, by examining the relationship between the model spread in AA and other metrics, our results indicate that the spread in the magnitude of AA in SON is not clearly related to either sea-ice area in that season at the time, and becoming less correlated in time with the loss of sea-ice relative to the reference period. On the other hand, the presence or absence of sea-ice in JJA appears to be an important factor in determining the magnitude of AA. Significant scenario uncertainty in projections of September sea-ice area and model uncertainty have been found by a previous study [55], with only a small contribution from internal variability, which may help explain why JJA and SON seasons exhibit more scenario uncertainty. Additionally, the Planck feedback may be reduced in SON in the larger forcing scenarios, leading to lower AA. Internal variability, as measured by seven different large initial condition ensembles, is of similar magnitude to model uncertainty in the early 21st century, but reduces in time when considering AA relative to preindustrial. When calculating AA as the ratio of 30 year trends, internal variability is the largest source of uncertainty throughout most of the 21st century. Crucially, this is an irreducible uncertainty and values of the AA ratio in reanalysis products fall within the range we find for internal variability, given the relatively low warming trend thus far observed.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://esgf-node.llnl.gov/search/cmip6/ and www.cesm.ucar.edu/community-projects/lens2/data-sets.

Acknowledgments

S H, J A S and J L C are all supported by Natural Environment Research Council Grant NE/V005855/1.

ORCID iDs

Stephanie Hay https://orcid.org/0000-0003-4007-842X
James A Screen https://orcid.org/0000-0003-1728-783X

References

[1] Manabe S and Wetherald R T 1975 *J. Atmos. Sci.* **32** 3–15
[2] Manabe S and Stouffer R J 1980 *J. Geophys. Res.* **85** 5529–54
[3] Davy R, Chen L and Hanna E 2018 *Int. J. Climatol.* **38** 4384–94
[4] Rantanen M, Karpechko A Y, Lipponen A, Nordling K, Hyvärinen O, Ruosteenoja K, Vihma T and Laaksonen A 2022 *Commun. Earth Environ.* **3** 168
[5] Yan Y and Wen X 2022 *Research Square* (https://doi.org/10.21203/rs.3.rs-1111875/v1)
[6] Miller G H, Alley R B, Brigham-Grette J, Fitzpatrick J J, Polyak L, Serreze M C and White J C 2010 *Quat. Sci. Rev.* **29** 1779–90
[7] Park H-S, Kim S-J, Kim S J, Kim S J, Stewart A L, Son S W, Lee J Y and Seo K H 2019 *Sci. Adv.* **5** eaax8203
[8] Roe G H, Feldl N, Armour K C, Hwang Y T and Frierson D M W 2015 *Nat. Geosci.* **8** 135–9
[9] Feldl N and Merlis T M 2021 *Geophys. Res. Lett.* **48** e2021GL094130
[10] Serreze M C, Barrett A P, Stroeve J, Kindig D N and Holland M M 2009 The *Cryosphere* **3** 11–19
[11] Screen J A and Simmonds I 2010 *Nature* **464** 1334–7
[12] Esau I, Pettersson L, Cancet M, Chapron B, Chernokulsky A, Donlon C, Sizov O, Soromotin A and Johannessen J 2023 *Remote Sens.* **15** 1354
[13] Holland M M and Bitz C M 2003 *Clim. Dyn.* **21** 221–32
[14] Hahn L, Armour K, Zelinka M, Bitz C and Donohoe A 2021 *Front. Earth Sci.* **9** 641760
[15] Dai A, Luo D, Song M and Liu J 2019 *Nat. Commun.* **10** 121
[16] Sejas S and Taylor P 2023 *Environ. Res.* **2** 035008
[17] Pithan F and Mauritsen T 2014 *Nat. Geosci.* **7** 768–76
[18] Curry J A, Schramm J L, Rossow W B and Randall D A 1996 *J. Clim.* **9** 1731–64
[19] Taylor P C, Kato S, Xu K-M and Cai M 2015 *J. Geophys. Res.* **120** 12656–78
[20] Cai M 2005 *Geophys. Res. Lett.* **32** L22712
[21] Olascoaga M J 2006 *Geophys. Res. Lett.* **33** L22603
[22] Spielhagen R F, Werner K, Sørensen S A, Zamelczyk K, Kandiano E, Budéus G, Husum K, Marchitto T M and Hald M 2011 *Science* **331** 450–3
[23] Yoshimori M, Abe-Ouchi A and Lainé A 2017 *Clim. Dyn.* **49** 3457–72
[24] Goosse H *et al* 2018 *Nat. Commun.* **9** 1919
[25] Peings Y, Cattiaux J, Cattiaux J, Vavrus S J and Magnusdottir G 2018 *Environ. Res. Lett.* **13** 074016
[26] Barnes E A and Polvani L M 2015 *J. Clim.* **28** 5254–71
[27] Harvey B J, Shaffrey L and Woollings T 2014 *Clim. Dyn.* **43** 1973–91
[28] Cohen J *et al* 2020 *Nat. Clim. Change* **10** 20–29
[29] Blackport R, Screen J A, van der Wiel K and Bintanja R 2019 *Nat. Clim. Change* **9** 697–704
[30] Cohen J, Agel L, Barlow M and Entekhabi D 2023 *Commun. Earth Environ.* **4** 341
[31] Cohen J, Screen J A, Furtado J C, Barlow M, Whittleston D, Coumou D, Francis J A, Dethloff K, Entekhabi D and Overland J E 2014 *Nat. Geosci.* **7** 627–37
[32] Smith D *et al* 2022 *Nat. Commun.* **13** 727
[33] Screen J A, Eade R, Smith D M, Thomson S and Yu H 2022 *Geophys. Res. Lett.* **49** e2022GL100523
[34] Hawkins E and Sutton R 2009 *Bull. Am. Meteorol. Soc.* **90** 1095–108
[35] Hersbach H *et al* 2020 *Q. J. R. Meteorol. Soc.* **146** 1999–2049
[36] Gelaro R *et al* 2017 *J. Clim.* **30** 5419–54
[37] Kobayashi S *et al* 2015 *J. Meteorol. Soc. Japan* **93** 5
[38] Hind A, Zhang Q and Brattström G 2016 *Sci. Rep.* **6** 30469
[39] Santer B D, Wigley T M L, Schelsinger M E and Mitchell J F B 1990 Developing climate scenarios from equilibrium GCM results (available at: https://api.semanticscholar.org/CorpusID:127017512)
[40] Tebaldi C and Arblaster J M 2014 *Clim. Change* **122** 459–71
[41] Liang Y C, Polvani L M and Mitevski I 2022 *npj Clim. Atmos. Sci.* **5** 14
[42] Ono J, Watanabe M, Komuro Y, Tatebe H and Abe M 2022 *Commun. Earth Environ.* **3** 27
[43] Holland M M and Landrum L 2021 *Front. Earth Sci.* **9** 719024
[44] Chung E, Ha K, Timmermann A, Stuecker M F, Bodai T and Lee S K 2021 *Earth's Future* **9** e2020EF001898
[45] Hahn L, Armour K C, Battisti D S, Donohoe A and Fajber R 2023 *Geophys. Res. Lett.* **50** e2023GL105156
[46] Hahn L, Armour K, Battisti D, Eisenman I and Bitz C 2022 *J. Clim.* **35** 1629–42
[47] Boeke R C, Taylor P T, Taylor P C and Taylor P C 2018 *Nat. Commun.* **9** 5017

[48] Wu Y, Liang Y, Kuo Y, Lehner F, Previdi M, Polvani L M, Lo M and Lan C W 2023 *Geophys. Res. Lett.* **50** e2022GL100745
[49] Previdi M, Smith K and Polvani L M 2021 *Environ. Res. Lett.* **16** 093003
[50] Screen J A, Deser C and Simmonds I 2012 *Geophys. Res. Lett.* **39** L01807
[51] Laliberté F and Kushner P J 2013 *Geophys. Res. Lett.* **40** 606–11
[52] Labe Z, Peings Y and Magnusdottir G 2020 *Geophys. Res. Lett.* **47** e2020GL088583
[53] He S, Xu X, Furevik T and Gao Y 2020 *Geophys. Res. Lett.* **47** e2020GL087212
[54] Davy R and Griewank P 2023 *Environ. Res. Lett.* **18** 084003
[55] Bonan D, Lehner F and Holland M M 2020 *Environ. Res. Lett.* **16** 044002