ARTICLE

Check for updates

Robust and perfectible constraints on human-induced Arctic amplification

Hervé Douville [1]✉

The Arctic near-surface warming is much faster than its global counterpart. Yet, this Arctic amplification occurs a rate that is season, model and forcing-dependent. The present study aims at using temperature observations and reanalyses to constrain the projections of Arctic climate during the November-to-March season. Results show that the recently observed four-fold warming ratio is not entirely due to a human influence, and will decrease with increasing radiative forcings. Global versus regional temperature observations lead to complementary constraints on the projections. When Arctic amplification is defined as the additional polar warming relative to global warming, model uncertainties are narrowed by 30% after constraint. Similar results are obtained for projected changes in the Arctic sea ice extent (40%) and when using sea ice concentration and polar temperature observations to constrain the projected polar warming (37%), thereby confirming the key role of sea ice as a positive but model-dependent surface feedback.

[1] Centre National de Recherches Météorologiques, Université de Toulouse, Météo-France, CNRS, 42 Avenue Gaspard Coriolis, 31057 Toulouse, France.
✉email: herve.douville@meteo.fr

The Arctic amplification (AA) of global warming is a well-known feature of climate projections, documented by several generations of models taking part in the Coupled Model Intercomparison Project (CMIP)[1–3]. Early climate projections[1] highlighted some diversity in the magnitude, spatial distribution, and seasonality of the near-surface polar warming in the Northern Hemisphere (NH). Yet, they showed a consistent amplification of temperature anomalies ranging typically from 1.5 to 4.5 times the global mean warming. This phenomenon was first attributed to a positive but model-dependent sea-ice albedo feedback. Projected changes in poleward ocean heat transport and in cloud cover at high latitudes were also identified as potential contributors to the enhanced Arctic warming[1]. These early findings have been corroborated by both CMIP5 and CMIP6 models. Yet, the underlying mechanisms, their seasonality, and the quantification of their relative contributions to AA have been the subject of on-going debate, as their better understanding has not translated into much more reliable projections from one generation of CMIP models to the next.

Starting with CMIP5 projections, the central role of diminishing sea-ice[2,3], has been challenged by the identification of other relevant mechanisms, such as large contributions from longwave feedbacks[4,5] and changes in atmospheric and oceanic heat transport[6,7]. While most feedbacks were reported to contribute to surface warming both over land and ocean, the positive sea-ice albedo feedback is by definition only active over the ocean and is maximum in summer[2]. Yet, the sea ice retreat is not a purely radiative feedback and is also important in winter when it is associated with a positive lapse rate feedback (LRF) and enhanced surface evaporation over the Arctic Ocean. The positive LRF is a noteworthy regional exception (given the overall negative LRF at the global scale) which is dominated by local rather than remote drivers[8]. Unlike polar warming, the total uncertainty in the CMIP5 projections of AA was shown to be dominated by model differences (rather than scenario uncertainty or internal variability), with a gradual shift of the maximum inter-model spread from autumn to winter across the 21st century[9].

In comparison with CMIP5, the latest-generation CMIP6 models show similar performance in their representation of Arctic sea ice[10]. They also lead to similar qualitative conclusions about the winter-dominated AA and the main associated feedbacks[3,11]. Yet, they still show a substantial inter-model spread in the projected Arctic warming[12]. For comparable emissions scenarios, a stronger ensemble-mean polar warming was first attributed to a larger surface albedo feedback, combined with less-negative cloud feedbacks. However, scaling the Arctic warming by the concomitant global warming yields a similar degree of AA in CMIP5 and CMIP6[3]. A radiative kernel method highlights that AA does not primarily arise from the surface albedo feedback but rather from the LRF in winter[13]. This positive LRF is a seasonal phenomenon, triggered by sea surface temperature changes after sea-ice loss, and not by the degree of atmospheric stratification. The key role of the declining sea ice cover may also explain why the inter-model spread of AA was found to decrease with increasing radiative forcings[14]. The maximum AA is found in winter, but is sensitive to the model-dependent sea ice melting in early summer, which gets amplified in cold months by oceanic heat storage/release[15,16]. In contrast, the transition from sea ice to sea water during the melting season is associated with a sharp increase in thermal inertia, which slows the Arctic warming in spring and summer[17].

To sum up, there are still inconsistencies and uncertainties regarding the drivers and magnitude of AA in global climate models. This can be related to multiple processes but also to possible methodological issues: AA definition, single versus multi-model studies, idealized increased-CO_2 versus scenario experiments, choice of the emission scenario and of the radiative kernel for feedback decomposition, focus on either annual or seasonal means. Moreover, polar feedbacks are tightly coupled to changes in oceanic and atmospheric energy transport, so that their contributions to AA should not be considered in isolation[18]. Clearly, our assessment of these complex and elusive feedbacks remains limited and other, top-down rather than bottom-up, approaches may be needed to constrain the projections in this highly sensitive region, on the front line of global climate change. It should be stressed here that the present study is not the first attempt to constrain future changes in Arctic climate (see the relevant section below), but is unique in that it aims to condition recent and future changes on observed changes in a consistent way and using a Bayesian statistical method that has already been tested successfully through the use of pseudo-observations (see Methods). Before applying this method to assess the human contribution to the recent observed AA and to future global projections, the CMIP6 model uncertainties in a high-emission scenario are first quantified and briefly discussed in the light of previously discussed regional feedbacks.

Quantifying uncertainties in CMIP6 projections of the Arctic climate

The inter-model spread in AA has not been reduced in the latest CMIP6 global climate projections[3,18] (Fig. 1). According to the latest assessment report delivered by the first working group of the Intergovernmental Panel on Climate Change (IPCC AR6 WG1), "*there remains substantial uncertainty in the magnitude of projected AA, with the Arctic warming ranging from two to four times the global average in models*". This finding is supported by our own analysis of 36 CMIP6 models. Our baseline period for present-day climate is the 1995–2014 period also chosen by the AR6 WG1. The focus is on the SSP5-8.5 high-emission scenario and on an extended winter season (ONDJFM), when the Arctic warming is the highest in both models and observations. This choice allows us to maximize the signal to noise ratio and, thus, to use a single realization for each model. Similar results are obtained with the previous-generation CMIP5 models despite their overall lower climate sensitivity (Supplementary Fig. 1). In line with former studies[11,14], a higher degree of amplification over the Arctic Ocean is found in the mid-21st century (Supplementary Figure 2) or in a weaker emission scenario (Supplementary Fig. 3). This result is consistent with a previous study[14] suggesting that, unlike over Antarctica, polar amplification becomes weaker with increasing radiative forcings over the Arctic. It is also consistent with a key role of the sea ice feedback given the limited ice volume in the Arctic compared to Antarctica. This feedback dependency to the base state may indeed explain why AA is decreasing with further warming and may suggest the potential interest of using sea ice volume rather than sea ice concentration as an observational constraint in future Arctic studies.

If one does not scale the projected climate change by the corresponding global warming in each CMIP6 model, the inter-model spread also includes the contrasted climate sensitivity across the CMIP6 models which adds to the previously shown AA uncertainty (Supplementary Fig. 4). The latitudinal distribution of this spread is consistent with the pattern of uncertainties in the response of sea-ice concentration (Supplementary Fig. 5), snow cover (Supplementary Fig. 6) and total precipitable water (Supplementary Fig. 7), among other variables that may contribute to the high-latitude positive feedbacks on polar warming. Not surprisingly, the projected response of total cloud cover (Supplementary Fig. 8) ranges from slightly negative to clearly positive values over the Arctic across the CMIP6 multi-model ensemble. According to these models and in line with our
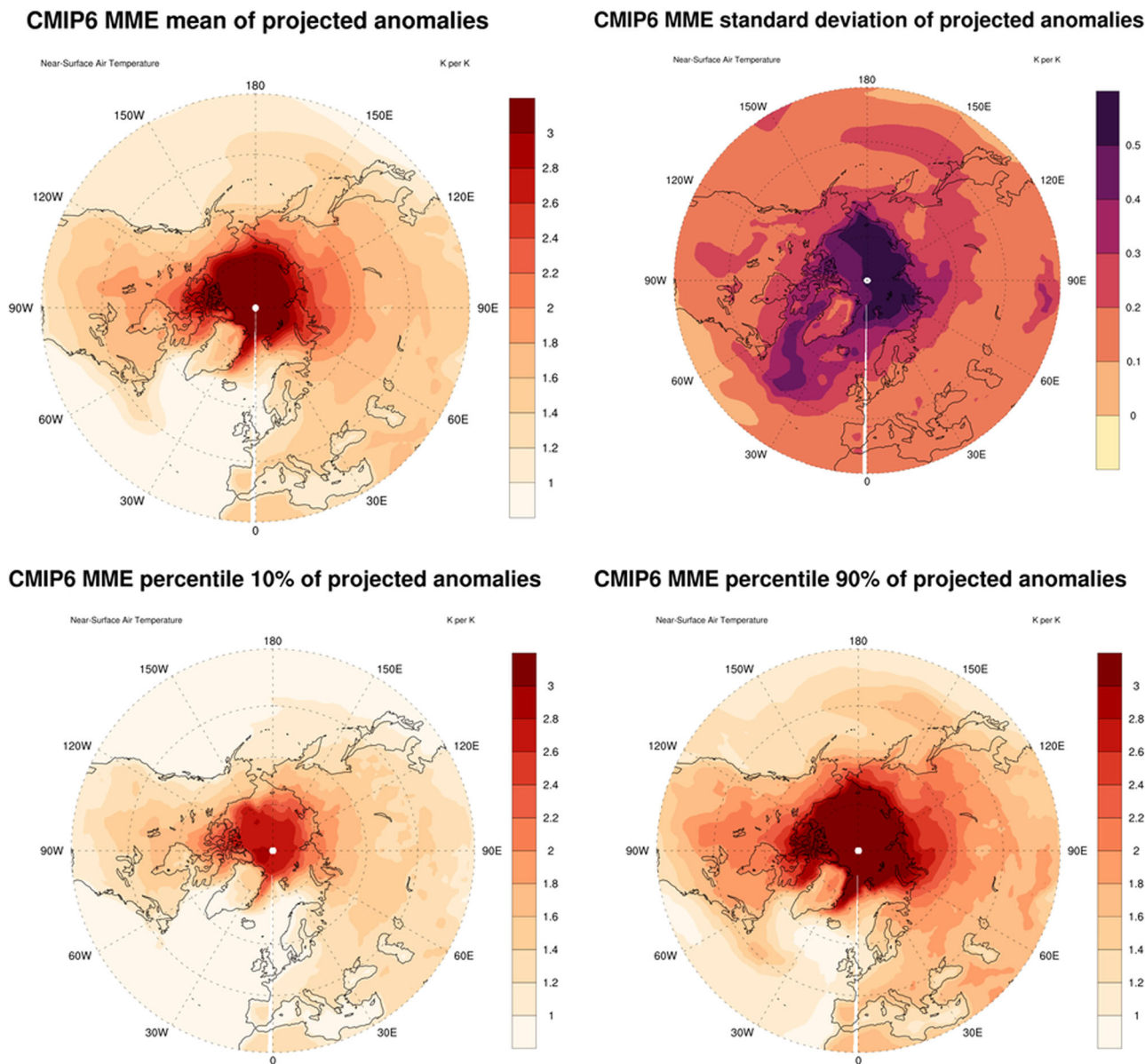
Fig. 1 Ratio of ONDJFM near-surface temperature anomalies to global warming (K per K) in the SSP5-8.5 high-emission scenario. Four multi-model ensemble statistics are shown: the ensemble mean anomalies (top left panel), the standard deviation as a measure of inter-model spread (top right panel), as well as the 10% and 90% local percentiles (bottom panels). All stereopolar maps are based on a set of thirty-six CMIP6 models with available monthly mean tas outputs. All anomalies are estimated as the differences between the 2081–2100 and 1995–2014 climatologies and scaled by the corresponding global warming.

results, the Arctic will become cloudier in a warmer climate, especially in winter when the associated positive radiative feedback occurs primarily in the longwave portion of the spectrum.

Several metrics[19] have been used to quantify the degree of AA in observations. Some are based upon the ratio of linear trends or of interannual variability between polar (here north of 60°N) and global mean near-surface air temperatures respectively (hereafter PSAT and GSAT). Another relies on the regression between polar and global mean warmings. Such definitions are sensitive to the selected baseline period and cannot be easily used to provide a continuous monitoring given the weak signal to noise ratio at the beginning of the instrumental record or even in the mid-20th century, thereby leading to undefined or very unstable values. This is the reason why another simple definition[20] will be also used in the present study, where AA is estimated as the difference between concomitant changes in PSAT and GSAT. In the

following, two regional AA metrics will be thus employed: a "fractional" AA based on the ratio between ONDJFM changes in PSAT and the annual mean changes in GSAT against the same reference period, and a "differential" AA computed as the difference between ONDJFM changes in both PSAT and GSAT.

Figure 2a shows a scatterplot of the differential AA in individual CMIP6 models against the corresponding projected annual mean global warming at the end of the 21st century relative to the 1995–2014 baseline period. Global warming accounts for only two-thirds of the total spread in this AA metric among the CMIP6 ensemble. This result highlights a notable contribution of regional rather than global feedbacks. This is further supported by Fig. 2b, c, showing a stronger correlation between AA and ONDJFM anomalies in NH sea ice extent and NH total precipitable water respectively, than with global warming. In contrast, there is no obvious link between AA and projected polar
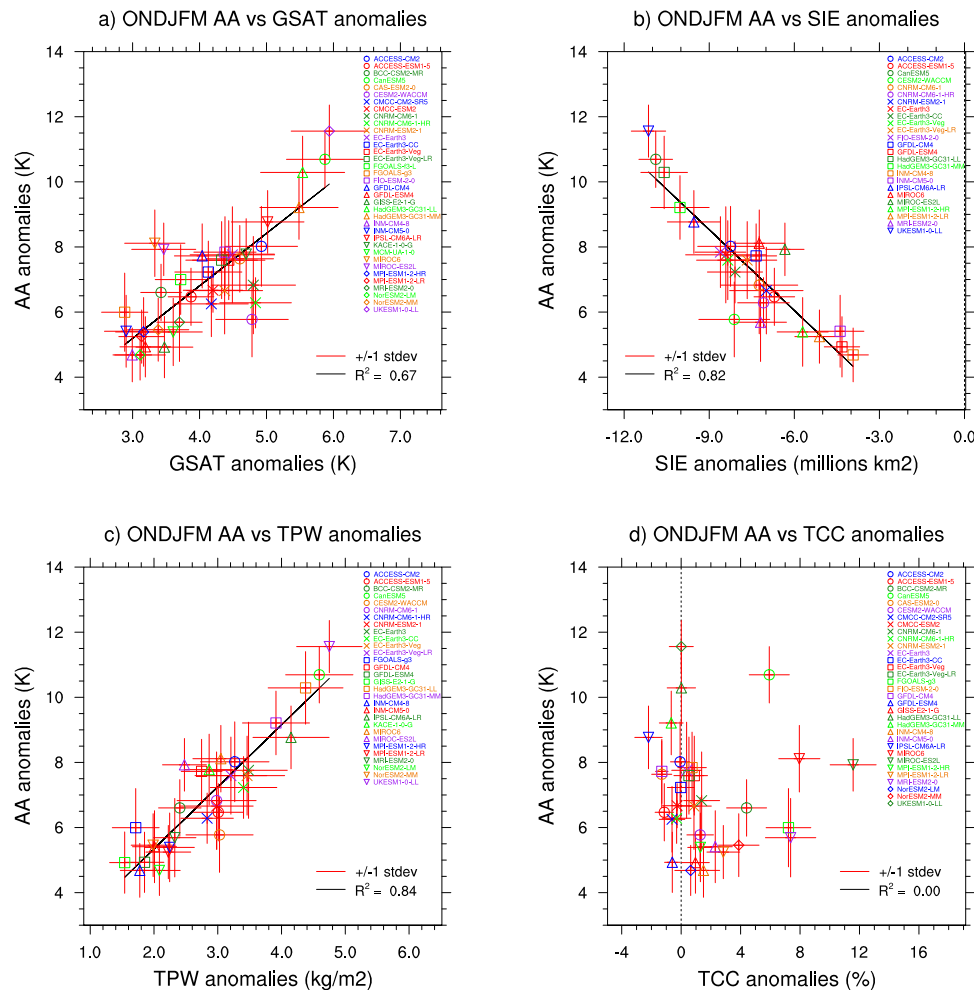
Fig. 2 Scatterplots of the differential AA (i.e., ONDJFM PSAT minus GSAT anomalies) versus other metrics across CMIP6 models. a Global warming (GSAT anomalies in K), **b** ONDJFM sea ice extent anomalies (in millions of km²), **c** ONDJFM Total Precipitable Water anomalies (in %) north of 60°N, and **d** ONDJFM Total Cloud Cover anomalies (in %) north of 60°N. All anomalies are estimated over 2080–2099 versus the 1995–2014 baseline period. All models are not available for all metrics. Red error bars denote ±1 standard deviation of interannual variability, while R^2 denotes the squared correlation between the two variables shown in each panel.

changes in total cloud cover (Fig. 2d). This finding is not surprising given the other more important regional feedbacks assessed in former studies and the fact that cloud radiative feedbacks do not only depend on changes in TCC, but also on possible changes in cloud elevation, vertical extent and radiative properties which are beyond the scope of the present study.

Supplementary Fig. 9 repeats the same analysis as in Fig. 2 but for polar warming rather than AA (i.e., without substracting the GSAT from the PSAT anomalies). The linear fit between PSAT and GSAT anomalies shows a best guess of 2.6 K per 1 K of global warming at the end of the 21st century, which is much lower than the recent amplification reported in the instrumental record. The inter-model spread in PSAT anomalies is accounted for at 84% by uncertainties in the projected global warming. Yet, and in line with our previous results, the correlation is even higher when considering NH sea ice extent or total precipitable water anomalies. This again highlights that AA is modulated by regional feedbacks and is generally stronger in models that project a large reduction in sea ice extent and a large percentage increase in tropospheric humidity. One key question is therefore to compare the effect of different observational constraints, not only on different manifestations of Arctic climate change (to check whether they lead to consistent results about the regional sensitivity of CMIP6 models), but also on polar warming (to

check whether some variables may be more efficient than GSAT to constrain this specific metric).

Constraining and attributing recent changes in Arctic warming

Beyond climate models, AA was also unveiled by paleoclimate reconstructions[21,22], instrumental and satellite records[23–25], as well as state-of-the-art atmospheric reanalyses[25,26]. An early observational study[23] based on a 125-year instrumental record did not support the simulated AA of global warming, but highlighted that the Arctic climate variability is dominated by multi-decadal fluctuations which may obscure long-term changes. The key role of sea-ice internal multi-decadal variability could not be conclusively identified, but was highlighted by several subsequent studies[27,28].

In the early 21st century, a study[25] based on the ERA-Interim reanalysis found a two-fold ratio between the Arctic and global near-surface warming over the 1989–2008 period, leading to further investigation about the role of changes in sea ice, atmospheric and oceanic circulation, cloud cover and atmospheric water vapor[25]. As in models, the observed Arctic warming was shown to be strongest at the surface and primarily consistent with the concomitant sea-ice retreat. No strong evidence was found for

a substantial radiative impact of changes in cloud cover (despite a recent increase in Arctic cloudiness[29]), but an increase in atmospheric water vapor content was reported and may have contributed to enhance the Arctic warming during summer and early autumn. More recently[26], a much stronger four-fold AA was reported over 1979–2021 from multiple temperature datasets including the state-of-the-art ERA5 reanalysis[30,31]. Such a warming ratio is extremely rare in the CMIP5 and CMIP6 simulations, thereby suggesting that it is either an extremely unlikely event or that AA is systematically underestimated by climate models[26]. This result advocates for a better understanding and consideration of internal climate variability when constraining the Arctic climate projections with the observed trends[32].

Our analysis is based on the KCC statistical package[33,34] and on the combined use of GSAT and PSAT reconstructions (see Methods). This Bayesian technique allows us to derive a posterior distribution of the projected anomalies from the prior distribution derived from the raw model outputs. Doing so, it takes account of both internal variability and observational uncertainties to constrain recent and future simulated climate changes in a consistent way. Besides the HadCRUT5 gridded dataset at a relatively coarse spatial resolution (5° by 5°), the ERA5 reanalysis has been also used to provide alternative regional constraints, but with no estimate of observational errors. This may be partly justified by the improved quality of ERA5 in the Arctic compared to previous global atmospheric reanalyses[31]. Similarly, our reconstructions of total precipitable water and total cloud cover north of 60°N have been derived from ERA5 and should be considered cautiously given the evolution of the global observing system since 1959. Finally, the Arctic sea ice extent was also estimated from the ERA5 sea ice concentration outputs but was prescribed in ERA5 and originates from observations[30].

Figure 3a shows a scatterplot of recent changes in ONDJFM PSAT versus recent changes in annual mean GSAT across ten CMIP6 models and in HadCRUT5 observations. This subset of models has been selected because it also provides historical

simulations driven by individual radiative forcings (Fig. 3b). Despite a wider definition of the Arctic domain (north of 60°N compared to 66.5°N in a previous study[26]), our results support an observed four-fold ratio between the Arctic and the globe respectively (cf. the black cross in Fig. 3a). Yet, such a high value is outside the range of simulated values found in CMIP6 models and of the related 90% confidence interval of the prior joint distribution (blue ellipse fitted on raw model outputs). All CMIP6 models seem to underestimate the observed Arctic amplification factor, but part of this underestimation may be due to internal variability which contributes to the overlap between the blue and black ellipses in Fig. 3a. In the end, the uncertainty in the forced response of the CMIP6 models (red ellipse) is much less than in the prior distribution (blue ellipse). Yet, our constrained estimate of the polar warming amplification factor (around 3) is not much different from the unconstrained estimate. This value is higher than the 2.6 ratio found at the end of the 21st century (Supplementary Fig. 9a), which is fully consistent with the fact that the Arctic feedbacks are time-evolving and forcing-dependent[18]. Some feedbacks may be for instance sensitive to the base state and thus to potential biases in climate models. Such a link can be obscured by a few outliers, but is clear in the case of the cloud response (Supplementary Fig. 10). Some CMIP6 models show a stronger than observed Arctic cloudiness[29] which is then so close to 100% that it cannot increase across the 21st century. This result highlights the need to further improve climate models and/or to constrain their projections with reliable observations, but over a long enough period to account for a possible time-dependence of the Arctic feedbacks.

The KCC method also allows us to constrain and attribute simulated changes in PSAT since 1850 (Fig. 3b). Not surprisingly, the results show that the simulated changes cannot be explained by natural forcings (cf. the timeseries shown in blue) and are thus mostly due to human activities (cf. other colors). Yet, the polar warming induced by the greenhouse gas (GHG) emissions has been parly offset by anthropogenic aerosols across the 20th century. Both unconstrained effects however appear to be
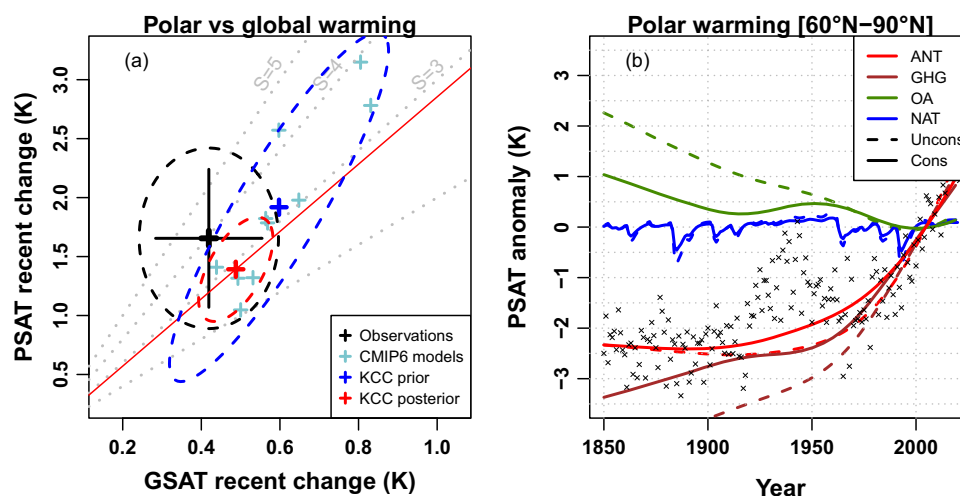

Fig. 3 GSAT-dependence and attribution of recent observed changes in ONDJFM PSAT. a Scatterplot of recent changes in ONDJFM PSAT (K) versus annual mean GSAT (K). Recent changes are simply estimated as the difference between the 2001–2020 and 1981–2000 climatologies, respectively. Individual CMIP6 models are shown as cyan crosses while the ensemble mean and ensemble spread (5–95% confidence interval) of their *prior* and *posterior* joint distributions are shown as crosses and ellipses in blue and red respectively. The red solid line shows the linear regression fit of the *posterior* joint distribution. Observed changes and related uncertainties are shown in black. The gray dotted lines denote four illustrative amplification rates ranging from 2 to 5 K per 1 K of global warming; **b** Constrained (solid lines) and unconstrained (dashed lines) timeseries of the ensemble-mean PSAT response to natural (NAT, with no significant change between the constrained and unconstrained timeseries), greenhouse-gas (GHG), other-anthropogenic (OA), and anthropogenic (ANT = GHG + OA) forcings over the period 1850–2020. Black crosses denote the median HadCRUT5 anomalies across the 200 available members.

overestimated compared to the KCC results. This may arise from several forcing and/or model deficiencies, such as a missing warming effect due to the deposition of black carbon aerosols on sea-ice. Interestingly, the constrained ensemble mean polar warming due to all anthropogenic forcings (red solid line) is more gradual than in the raw model outputs (red dashed line). This seems realistic given the steady increase in GHG concentrations and their overall dominant effect on near-surface temperatures. The observed increase in PSAT (black crosses) is mostly attributable to the GHG emissions but is only partly captured by the ensemble mean CMIP6 historical simulations. The apparent mismatch between the observed and forced component of the polar warming, especially in the mid-twentieth century, suggests a potential strong influence of internal climate variability, in line with former studies highlighting the key influence of the ocean multi-decadal variability in the Pacific and/or North Atlantic oceans[27,28].

Constraining future changes in Arctic climate
As discussed earlier, AA was found in both paleo- and 21st century climate simulations, thereby suggesting that ice-core-based reconstructions may provide quantitative insights on future climate changes[21,22]. Yet, it may be inappropriate to simply scale an observational estimate of past temperature changes to predict the future climate sensitivity[35]. Moreover, the documented dependence of both Arctic and tropical feedbacks on control climate[18] may challenge the feasibility of constraining climate projections with paleo data.

A possible alternative to narrow model uncertainties in climate projections is to use the so-called "emergent constraint" (EC) technique[36,37]. This empirical statistical method consists of linking future climate changes to observable metrics that can be more or less accurately simulated by CMIP-class models. The relevance of these metrics is usually supported by the existence of a correlation with the projected future climate response across an ensemble of models. ECs are generally applied in a simple regression framework, where the ensemble is used to define a predictive relationship that can be combined with observations to produce an estimate of constrained projections. While this method has drawn a growing interest and has been applied to multiple variables, there are so far only few studies related to the Arctic[32,38,39]. They all rely on observed trends (in Arctic sea ice or near-surface temperature) rather than on model biases. This strategy is consistent with our results that show no apparent link between sea ice sensitivity and sea ice mean state across the CMIP6 models (Supplementary Fig. 11).

Interestingly, a recent EC study[40] was aimed at constraining the large CMIP6 scatter in AA with a broad set of recent observations co-located to model data. The results suggested that the lower thermodynamic structure of the atmosphere is more realistically depicted in climate models with limited AA (weakly positive polar LRF) in the recent past. In contrast, remote influences that can shape the warming structure in the free troposphere are more realistically captured by models with a strong AA (strongly positive Arctic LRF). The two contrasted findings highlight the difficulty to define and combine relevant ECs based on present-day climate. A former CMIP6 study[39] found that the projected Arctic warming is positively correlated with the simulated global warming trend from 1981 to 2011 across the multi-model ensemble. Given this simple EC, the concomitant observed global warming suggests a weaker Arctic warming compared to the CMIP6 median projection. This study supports our choice to focus on recent temperature changes rather than on mean climate, but the KCC method will take advantage of the full historical record to constrain the projections.

Moreover, there is increasing evidence that most ECs that have been proposed to constrain CMIP5 projections are generally less efficient when applied to CMIP6 models[37]. This spurious behavior can arise from model interdependency[37,38] through common structural model assumptions and can lead to overconfident constrained projections. Our KCC technique does not build on empirical linear regression schemes and has been tested successfully in a perfect model framework[33]. It has been already applied at both global[33] and local scales[34]. Beyond temperature, KCC has been also used to constrain other variables, such as global total precipitable water[41] or global land surface relative humidity[42], leading to consistent results for both CMIP6 and CMIP5 models.

Here, the method is first applied to 36 CMIP6 models under the SSP5-8.5 high-emission scenario. In line with the limited GSAT influence on our differential AA index (Fig. 2a), the KCC results show the added value of using both global and regional HadCRUT5 observations for constraining this metric (Fig. 4) and the corresponding Arctic warming (Supplementary Fig. 12). Using first only observations of global mean surface temperature (GMST, combination of SAT over land and sea ice and of sea surface temperature over the ocean), KCC leads to a narrowing and downward shift of the AA projections (Fig. 4b), in line with the well-known[33] overestimation of global warming by some CMIP6 models (Fig. 4a). In contrast, the regional constraint of PSAT does not much change the ensemble mean projection but mainly narrows the plausible range of projected AA. The obtained reduction of the 90% confidence interval for the AA index projected at the end of the century is however slightly weaker (~23%) than when applying the GMST constraint (~26%). Finally, and not surprisingly, the combination of both, global and regional, observational constraints within KCC leads to an even stronger narrowing of model uncertainty (30%) and suggests that the most extreme CMIP6 responses are not compatible with the HadCRUT5 observations.

Not surprisingly given our differential AA definition, similar results are obtained when focusing on PSAT projections (Supplementary Fig. 12) since GSAT and PSAT are not independent metrics. Moreover, our conclusions are qualitatively similar when using ERA5 rather than HadCRUT5 as an observational constraint of projected changes in PSAT (Supplementary Fig. 13), although the narrowing of model uncertainty is here slightly stronger given the assumption of no observational error (since the ERA5 reanalysis only provides one reconstruction of historical temperatures). The results are also not much sensitive to the choice of the prior distribution (CMIP5 instead of CMIP6), although the narrowing of model uncertainty is then less noticeable given the lower unconstrained inter-model spread across the CMIP5 models (Supplementary Fig. 14). Whatever the prior distribution is, KCC suggests that the constrained ensemble-mean PSAT warming relative to 1995–2014 is in 2100 slightly above 10 K.

Discussion and conclusions
KCC and ERA5 can be also used to constrain other projected changes in the Arctic climate. It is for instance possible to narrow uncertainties in the projected ONDJFM sea ice extent with an overall reduction of 5–95% confidence interval by 40% at the end of the century (Supplementary Fig. 15). Note that the ERA5 sea ice concentration, that has been here used as an observational constraint, was derived from satellite data since ERA5 is a global atmospheric reanalysis driven by observed oceanic boundary conditions[30]. Similarly, ERA5 can be used to constrain the average total precipitation water (Supplementary Fig. 16) or total cloud cover (Supplementary Fig. 17) projected north of 60°N. The
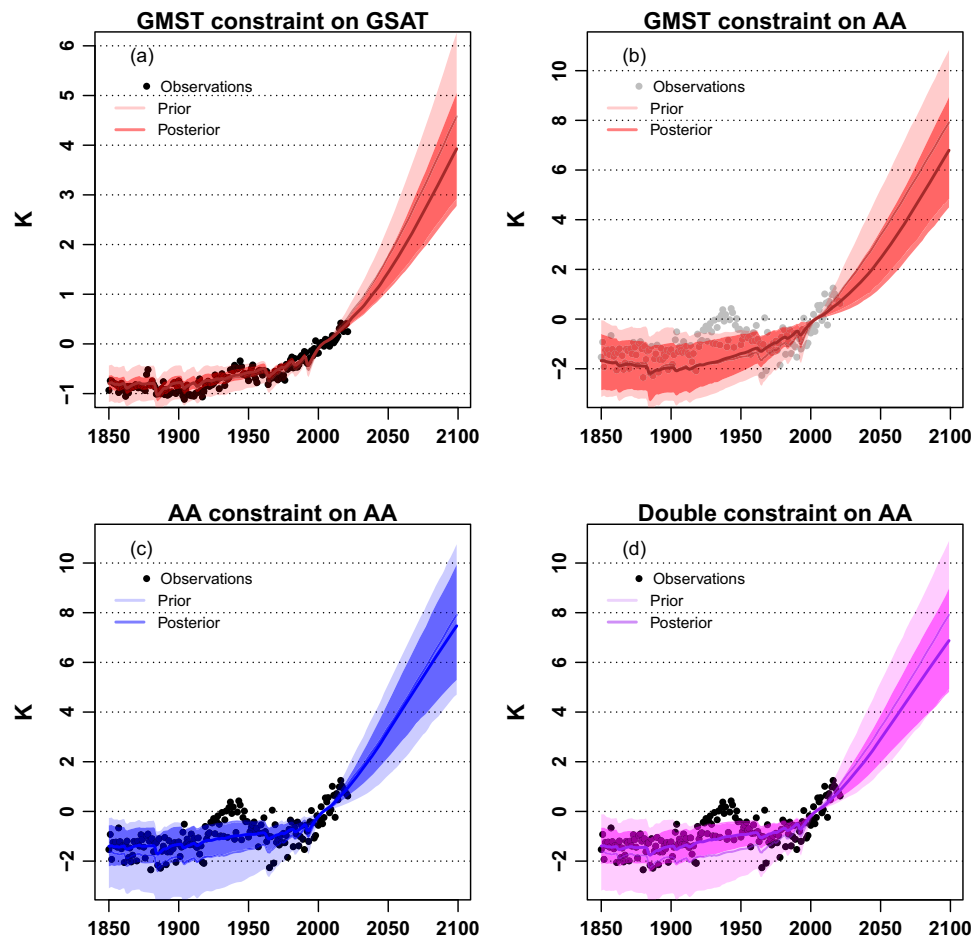
Fig. 4 Constrained versus unconstrained changes in GSAT and ONDJFM AA. Mean (solid lines) and 5–95% confidence interval (shading) of the *prior* (unconstrained) and *posterior* (constrained) distributions of the annual mean GSAT (K) and differential ONDJFM AA (K) forced response to both natural and anthropogenic radiative forcings in historical simulations and SSP5-8.5 projections from 36 CMIP6 models: **a** HadCRUT5 GMST constraint on GSAT; **b** HadCRUT5 GMST constraint on AA; **c** HadCRUT5 AA constraint on AA; **d** HadCRUT5 GMST and AA constraints on AA. After the constraint, the 5–95% confidence interval at the end of the 21st century is reduced by 32%, 26%, 23%, and 30% in (**a**, **b**, **c**, and **d**), respectively. Note that these percentages are minimum values since KCC leads to even more confident projections during the early to mid-21st century. The median values of the observed anomalies are show as black (gray) filled circles when they are (not) used for constraining the models' response.

KCC results should be here considered with more caution since ERA5 has been shown to exhibit spurious global trends in tropospheric humidity given the gradual evolution of the assimilated data[38]. Yet, they are physically consistent with those obtained for PSAT and the northern hemisphere sea ice extent. GMST observations lead to lower the upper bound of the projected Arctic warming and, consistently, have similar effects on the projected high-latitude total precipitable water and cloud cover. In contrast, our regional ERA5 constraints lead to exclude the lower bound values of the CMIP6 ensemble. This is also consistent with the opposite effects on the projected sea ice extent. This robust finding confirms that regional climate change does not scale accurately with global warming across different models[43], and that local observations are also important to constrain regional climate change[34,44]. The results also suggest that more reliable reconstructions of atmospheric humidity and cloudiness would be very useful to better constrain the projections.

Finally, and given the partial redundancy between global and regional temperature variations, we can use the ERA5 sea ice extent rather than GMST to constrain the surface polar warming (Supplementary Fig. 18). Results are then fully consistent with our previous attempt to constrain PSAT (Supplementary Fig. 12). Yet, the double constraint based on both ERA5 sea ice extent and

ERA5 PSAT leads to a slightly greater narrowing (by 37% instead of 34%) of the 5–95% confidence interval, still without a major change in the ensemble mean response. This result emphasizes the added value of reliable gridded observations for constraining climate change projections at the regional scale. Note that no ERA5 uncertainty, only internal variability, is here accounted for in KCC. Results are however not much sensitive if a 20% random mesurement error is introduced in the ERA5 timeseries. Clearly, KCC is a robust method for constraining both past and future climate change, and will provide even more tightly constrained projections as soon as we can use more reliable observations or longer timeseries. It could be thus used in a semi-operational context where climate projections are constrained on a regular basis (every year), using best quality-checked and updated datatsets.

Methods

Two simple regional AA metrics are considered in the present study. The fractional AA index simply writes as the following ratio:

$$\Delta PSAT/\Delta GSAT$$

where $\Delta PSAT$ denotes the seasonal (here ONDJFM) mean increase in near-surface warming north of 60°N (PSAT) and $\Delta GSAT$ denotes the annual mean global warming relative to the same reference period (i.e., 1979–1998 in Fig. 3).

Such a definition is not suitable for constraining AA over the full historical period (given the limited changes in GSAT in the early instrumental record), so that an alternative differential AA index is defined as:

$$\Delta PSAT - \Delta GSAT$$

where both ΔPSAT and ΔGSAT denote ONDJFM changes in PSAT and GSAT relative to a common reference period (1995–2014 in Figs. 2, 4).

Regarding the two CMIP generations of global climate models, we use all models that provide monthly mean near-surface air temperature outputs (named tas in the CMIP archives) for one realization (run 1) of both historical simulation (1850–2014 for CMIP6, 1850–2005 for CMIP5) and the corresponding high-emission scenario (2015–2100 for CMIP6, 2006–2100 for CMIP5). The list of available CMIP6 models reads as follows (36 models): ACCESS-CM2, ACCESS-ESM-1–5, BCC-CSM2-MR, CanESM5, CAS-ESM2-0, CESM2-WACCM, CMCC-CM2-SR5, CMCC-ESM2, CNRM-CM6-1, CNRM-CM6-1-HR, CNRM-ESM2-1, EC-Earth3, EC-Earth3-CC, EC-Earth3-Veg, EC-Earth3-Veg-LR, FGOALS-f3-L, FGOALS-g3, FIO-ESM-2-0, GFDL-CM4, GFDL-ESM4, GISS-E2-1-G, HadGEM3-GC31-LL, HadGEM3-GC31-MM, INM-CM4-8, INM-CM5-0, IPSL-CM6A-LR, KACE-1-0, MCM-UA-1-0, MIROC-ES2L, MIROC6, MPI-ESM1-2-HR, MPI-ESM1-2-LR, MRI-ESM2-0, NorESM2-LM, NorESM2-MM, and UKESM1-0-LL.The list of available CMIP5 models is slightly shorter (28 models): bcc-csm1-1-m, BNU-ESM, CanESM2, CCSM4, CESM1-CAM5, CMCC-CM, CMCC-CMS, CNRM-CM5, CSIRO-Mk3-6-0, FIO-ESM, GISS-E2-H, GISS-E2-H-CC, GISS-E2-R, GISS-E2-R-CC, HadGEM2-ES, inmcm4, IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR, MIROC5, MIROC-ESM, MIROC-ESM-CHEM, MPI-ESM-LR, MPI-ESM-MR, MRI-CGCM3, MRI-ESM1, NorESM1-M, NorESM1-ME.

The observational constraint method, called Kriging for Climate Change (KCC), has been previously applied to global and local warming[31,32], and can be easily applied to other climate variables[38,39] as long as their internal variability can be fitted with a simple mix of auto-regressive processes (Supplementary Fig. 19). KCC consists of three consecutive steps. First, the forced response of each climate model is estimated over the whole simulated period (here 1850–2099), using a spline smoothing with 6 degrees of freedom, and the response to specific individual forcings is also estimated for the attribution component of the study. Second, these forced responses sampled from the available climate models (CMIP5 or CMIP6) are used to build a prior of the real-world forced response. Finally, KCC allows us to derive a posterior distribution of the past and/or future forced responses conditional on the selected observations. The method can be summarized using the following equation:

$$y = Hx + \varepsilon \tag{1}$$

where y represents the observed timeseries (a long vector including observed changes in both GSAT and PSAT, here ONDJM anomalies relative to a common 1995–2014 baseline period from year 1850 to 2020 of the corresponding early winter season), x is the forced model response (a long vector of the corresponding simulated timeseries, here smooth ONDJFM anomalies relative to the same 1995–2014 baseline period as for observations from year 1850 to 2099 of the corresponding early winter season), H is an observational operator (matrix), ε is the random noise associated with internal variability and measurement errors (again a long vector providing a time-varying error estimate for GSAT and PSAT if available, as is the case with HadCRUT5 observations), and $\varepsilon \sim N(0, \Sigma_y)$, where N stands for the multivariate Gaussian distribution. Note that there is no threoretical motivation and not enough fully-independent CMIP6 models to assume a non-Gaussian distribution. Raw model outputs are thus used to construct the prior on x as: $\pi(x) = N(\mu_x, \Sigma_x)$. Then the posterior distribution given observations y can be derived as $p(x\,|\,y) = N(\mu_p, \Sigma_p)$. More details about the KCC method can be found in the related original studies and the R scripts and related data file of the current application are available from https://doi.org/10.5281/zenodo.8004439.

Interestingly, KCC can be used exactly in the same way to constrain historical and projected changes, but also historical changes driven by all or individual radiative forcings as concatenated in the x vector (e.g., anthropogenic greenhouse gases or natural radiative forcings only). By experiment design, the responses to individual forcings are considered to be additive (without interaction between different forcings or with internal climate variability) and so are the KCC-derived constraints on the model responses to individual forcings given the use of a generalized additive model (GAM) when estimating the prior model distributions.

In the following, we assess the forced response of Arctic climate (e.g., spatially averaged near-surface temperature north of 60°N or integrated Northern Hemisphere sea-ice extent), as well as the response to specific subsets of radiative forcings (attribution to GHG versus natural or all forcings respectively). These forced responses are then constrained by the historical observed global warming (https://www.metoffice.gov.uk/hadobs/hadcrut5/) and/or by ERA5 reanalyses (https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5).

We thus consider a very long CMIP6 vector consisting of four successive components:

$$x = (G_{all}, R_{all}, R_{ghg}, RE_{nat}) \tag{2}$$

where each element is an entire 1850–2099 timeseries of the forced response, G and R stand for global mean surface temperature and a regional variable, respectively. "all", "ghg" or "nat" are the subsets of external forcings considered. Similarly, we

define an observed vector made of two components:

$$y = (G_{obs}, R_{obs}) \tag{3}$$

i.e., only observed timeseries are used in y. The length of these timeseries depends on the selected dataset: 1850–2020 for HadCRUT5 and 1950–2020 for ERA5 (we have noticed that ERA5 has been recently updated from 1940 onwards, but we feel that the 1940s will not provide a notable additional constraint within KCC given the limited GHG influence in the mid-20th century). Note the HadCRUT5 dataset merges near-surface air temperature over land and sea ice, but sea surface temperature over the free ocean. This blended global mean surface temperature (GMST) is compared with the global mean near-surface air temperature (GSAT) from climate models. This approximation was discussed in details in former studies[31,32] and does not represent as a major issue compared to the other sources of uncertainty considered in the present study. The limited number of near-surface observations over sea ice and the coarse resolution (5° by 5°) of HadCRUT5 are, however, an incentive to also estimate the observed Arctic warming or other recent changes in the Arctic climate using ERA5 at a much higher resolution (here 0.5° by 0.5°).

All attribution or projection diagnoses presented below can be derived from the posterior distribution $p(x\,|\,y)$. μ_x and Σ_x are estimated as the sample mean and covariance matrix of the forced responses. Σ_y requires a statistical modeling of internal variability (Supplementary Fig. 19) and can also account for measurement errors (if available). The intrinsic variance of the selected global and Arctic climate indices is derived from observations after subtracting the multi-model mean estimate of the forced response from HadCRUT5 and ERA5 data respectively. We also assume a dependence between the global and regional variability, by accounting for the correlation between the two residuals in Σ_y. The assessment of measurement uncertainty in near-surface temperature is based on the HadCRUT5 ensemble (200 members) for both GMST and PSAT. In contrast, no observational error is accounted for when using ERA5 as an alternative or additional observational constraint.

To sum up, there are multiple advantages of the KCC method compared to more empirical (non-bayesian) statistical methods which have been proposed so far and, in a few cases, applied to the projections of Arctic climate change. On the methodological side, KCC does not rely on empirical relationships between observable and future climate properties (based on a limited number of not fully independent models), but relies on a more straightforward use of the full instrumental record to constrain both past and future climate changes in a consistent way. So doing, it accounts for both model and observational uncertainties and only assumes a Gaussian fit for both the prior and posterior distributions of simulated changes. It is thus not much sensitive to the choice of the prior distribution (i.e., CMIP model generation), at least much less than overfitted emergent constraints based on linear regressions in which potential outliers may have a spurious effect on the constrained distribution. Although the method does not need a full understanding of the dominant physical processes (which is a prerequisite for the use of more empirical emergent constraints), the limited number of observational constraints however raises the issue of selecting the most relevant observations in order to obtain a maximum narrowing of the posterior confidence interval (cf. the final discussion on using sea ise observations since 1950 compared to global mean surface temperature since 1850 and the corresponding Supplementary Information).

Data availability

Code availability

References
1. Holland, M. M. & Bitz, C. M. Polar amplification of climate change in coupled models. *J. Clim.* 21, 221–232 (2003).
2. Laîné, A., Yoshimori, M. & Abe-Ouchi, A. Surface Arctic amplification factors in CMIP5 models: land and oceanic surfaces and seasonality. *J. Clim* 29, 3297–3316 (2016).

3. Hahn, L. C., Armour, K. C., Zelinka, M. D., Bitz, C. M. & Donohoe, A. Contributions to polar amplification in CMIP5 and CMIP6 models. *Front. Earth Sci.* **9**, 710036 (2021).

4. Pithan, F. & Mauritsen, T. Arctic amplification dominated by temperature feedbacks in contemporary climate models. *Nat. Geosci.* **7**, 181–184 (2014).

5. Goosse, H. et al. Quantifying climate feedbacks in polar regions. *Nat. Commun.* **9**, 1919 (2018).

6. Huang, Y., Xia, Y. & Tan, X. On the pattern of CO_2 radiative forcing and poleward energy transport. *J. Geophys. Res. Atmos.* **122**, 10578–10593 (2017).

7. Park, K., Kang, S. M., Kim, D., Stuecker, M. F. & Jin, F.-F. Contrasting local and remote impacts of surface heating on polar warming and amplification. *J. Clim* **31**, 3155–3166 (2018).

8. Feldl, N. et al. Sea ice and atmospheric circulation shape the high-latitude lapse reate feedback. *npj Clim. Atmos. Sci.* **3**, 41 (2020).

9. Wu, Y.-T. et al. Exploiting SMILEs and the CMIP5 archive to understand Arctic climate change seasonality and uncertainty. *Geophys. Res. Lett.* **50**, e2022GL100745 (2023).

10. Shu, Q. et al. Assessment of sea ice extent in CMIP6 with comparison to observations and CMIP5. *Geophys. Res. Lett.* **47**, e2020GL087965 (2020).

11. Cai, S., Hsu, P.-C. & Liu, F. Changes in polar amplification in response to increasing warming in CMIP6. *Atmos. Ocean. Sci. Lett.* **14**, 100043 (2021).

12. Cai, Z. et al. Arctic warming revealed by multiple CMIP6 models: evaluation of historical simulations and quantification of future projection uncertainties. *J. Clim.* **34**, 4871–4892 (2021).

13. Boeke, R. C., Taylor, P. C. & Sejas, S. A. On the nature of the Arctic's positive lapse-rate feedback. *Geophys. Res. Lett.* **48**, e2020GL091109 (2021).

14. Xie, A. et al. Polar amplification comparison among Earth's three poles under different socioeconomic scenarios from CMIP6 surface air temperature. *Sci. Rep.* **12**, 16548 (2022).

15. Chung, E. S. et al. Cold-season Arctic amplification driven by Arctic ocean-mediated seasonal energy transfer. *Earth's Future* **9**, 1–17 (2021).

16. Hu, X., Liu, Y., Kong, Y. & Yang, Q. A quantitative analysis of the source of inter-model spread in Arctic surface warming response to increased CO_2 concentration. *Geophys. Res. Lett.* **49**, e2022GL100034 (2022).

17. Hahn, L. C., Armour, K. C., Battisti, D. S., Eisenman, I. & Bitz, C. M. Seasonality in Arctic warming driven by sea ice effective heat capacity. *J. Clim* **35**, 1629–1642 (2022).

18. Previdi, M., Smith, K. L. & Polvani, M. L. Arctic amplification of climate change: a review of underlying mechanisms. *Environ. Res. Lett.* **16**, 093003 (2021).

19. Davy, R., Chen, L. & Hanna, E. Arctic amplification metrics. *Int. J. Climatol.* **38**, 4384–4394 (2018).

20. Francis, J. A. & Vavrus, S. J. Evidence for a wavier jet stream in response to rapid Arctic warming. *Env. Res. Lett.* **10**, 014005 (2015).

21. Masson-Delmotte, V. et al. Past and future polar amplification of climate change: climate model intercomparisons and ice-core constraints. *Clim. Dyn.* **26**, 513–529 (2006).

22. Miller, G. et al. Arctic amplification: can the past constrain the future? *Quat. Sci. Rev.* **29**, 1779–1790 (2010).

23. Polyakov, I. V. et al. Observationally based assessment of polar amplification of global warming. *Geophys. Res. Lett.* **29**, 1878 (2002).

24. Serreze, M. & Francis, J. The Arctic amplification debate. *Climatic Change* **76**, 241–264 (2006).

25. Screen, J. A. & Simmonds, I. The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature* **464**, https://doi.org/10.1038/nature09051 (2010).

26. Rantanen, M. et al. The Arctic has warmed nearly four times faster than the globe since 1979. *Commun. Earth Environ.* **3**, 168 (2022).

27. Day, J. J. et al. Sources of multi-decadal variability in Arctic sea ice extent. *Environ. Res. Lett.* **7**, 034011 (2012).

28. Liu, W. & Fedorov, A. Interaction between Arctic sea ice and the Atlantic meridional overturning circulation in a warming climate. *Clim. Dyn.* **58**, 1811–1827 (2022).

29. Wang, X. et al. Seasonal trends in clouds and radiation over the Arctic seas from satellite observations during 1982 to 2019. *Remote Sens.* **13**, 3201 (2021).

30. Hersbasch, H. et al. The ERA5 global reanalysis. *Quaterly J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).

31. Graham, R. M., Hudson, S. R. & Maturilli, M. Improved performance of ERA5 in Arctic gateway relative to four global atmospheric reanalyses. *Geophys. Res. Lett.* **46**, 6138–6147 (2019).

32. Boé, J., Hall, A. & Qu, X. September sea-ice cover in the Arctic Ocean projected to vanish by 2100. *Nat. Clim. Change*, https://doi.org/10.1038/NGEO467 (2009).

33. Ribes, A., Qasmi, S. & Gillett, N. Making climate projections conditional on historical observations. *Sci. Adv.* **7**, eabc0671 (2021).

34. Qasmi, S. & Ribes, A. Reducing uncertainty in local temperature projections. *Sci. Adv.* **8**, eabo6872 (2022).

35. Crucifix, M. Does the Last Glacial Maximum constrain climate sensitivity? *Geophys. Res. Lett.* **33**, L18701 (2006).

36. Hall, A., Cox, P., Huntingford, C. & Klein, S. Progressing emergent constraints on future climate change. *Nat. Clim. Change* **9**, 269–278 (2019).

37. Sanderson, B. et al. The potential for structural errors in emergent constraints. *Earth Syst. Dyn.* **12**, 899–918 (2021).

38. Knutti, R. et al. A climate model projection weighting scheme accounting for performance and interdependence. *Geophys. Res. Lett.* **44**, 1909–1918 (2017).

39. Hu, X. M., Ma, J. R., Ying, R., Cai, M. & Kong, Y. Q. Inferring future warming in the Arctic from the observed global warming trend and CMIP6 simulations. *Adv. Clim. Chang. Res.* **12**, 499–507 (2021).

40. Linke, O., et al. Constraints on simulated past Arctic amplification and lapse-rate feedback from observations, *Atmos. Chem. Phys. Discuss.* [preprint], https://doi.org/10.5194/acp-2022-836, in review, (2023).

41. Douville, H., Qasmi, S., Ribes, A. & Bock, O. Global warming at near-constant tropospheric relative humidity is supported by observations. *Commun. Earth Environ.* **3**, 237 (2022).

42. Douville, H. & Willett, K. A dry future revisited. *Sci. Adv.* **9**, eade6253 (2023).

43. Douville, H. et al. Water remains a blind spot in climate change policies. *PLOS Water* **1**, e0000058 (2022).

44. Ribes, A. et al. An updated assessment of past and future warming over France based on a regional observational constraint. *Earth Syst. Dyn.* esd-2022-7 (2022).

Author contributions
H.D. designed the research and is entirely responsible for the content of the manuscript.

Competing interests
The author declares no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-023-00949-5.

Correspondence and requests for materials should be addressed to Hervé. Douville.

Peer review information *Communications Earth & Environment* thanks P Jonathan and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Heike Langenberg. A peer review file is available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.