ARTICLE OPEN

Response of winter climate and extreme weather to projected Arctic sea-ice loss in very large-ensemble climate model simulations

Kunhui Ye [1✉], Tim Woollings[1], Sarah N. Sparrow [2], Peter A. G. Watson[3] and James A. Screen [4]

Very large (~2000 members) initial-condition ensemble simulations have been performed to advance understanding of mean climate and extreme weather responses to projected Arctic sea-ice loss under 2 °C global warming above preindustrial levels. These simulations better sample internal atmospheric variability and extremes for each model compared to those from the Polar Amplification Model Intercomparison Project (PAMIP). The mean climate response is mostly consistent with that from the PAMIP multi-model ensemble, including tropospheric warming, reduced midlatitude westerlies and storm track activity, an equatorward shift of the eddy-driven jet and increased mid-to-high latitude blocking. Two resolutions of the same model exhibit significant differences in the stratospheric circulation response; however, these differences only weakly modulate the tropospheric response. The response of temperature and precipitation extremes largely follows the seasonal-mean response. Sub-sampling confirms that large ensembles (e.g. ≥400) are needed to robustly estimate the seasonal-mean large-scale circulation response, and very large ensembles (e.g. ≥1000) for regional climate and extremes.

INTRODUCTION

Arctic amplification (AA)[1–3] is a striking aspect of climate change and Arctic sea-ice cover has decreased by approximately one half over the satellite era, now at its lowest in at least the last 1450 years[4].

The mechanisms and consequences of AA have been extensively studied through observational and modelling analyses[2,5–19]. Numerous mechanisms have been proposed linking Arctic sea-ice loss and AA to altered midlatitude weather and climate, among which are changes in the North Atlantic Oscillation (NAO)[20], jet waviness[21], storm tracks[22], planetary waves[23,24] and the stratospheric polar vortex[25]. These mechanisms for AA impacting weather and climate in both observations and modelling, and relative importance have been extensively debated[26–48].

The Polar Amplification Model Intercomparison Project (PAMIP)[9,49] has coordinated an unprecedented set of multi-model ensemble experiments to advance understanding of the causes and consequences of AA. The PAMIP multimodel ensemble suggests robust but weak influences of Arctic sea-ice loss on winter climate[9]. Whilst the PAMIP provides a very large multi-model ensemble sample, the ensemble size for each model (mostly 100 years, but in some cases, up to 500) may not be sufficiently large to fully separate an individual model's forced response from internal variability[50]. Many of the theories listed above focus on the response of extreme events and persistence of flow regimes like jet latitude and atmospheric blocking, further motivating the need for large ensemble experiments to robustly quantify changes in extremes. To complement PAMIP, we therefore in this study present very large initial-condition ensembles (~2000 members) from two climate models to robustly quantify the change in dynamical regimes and weather extremes. We also answer two additional questions: (1) what ensemble sizes are needed for robust detection of extremes, as well as seasonal-

mean responses to projected Arctic sea-ice loss? and (2) is the response dependent on resolution? Regarding the latter, a previous study[51] compared the response to sea-ice loss in a model at 3 resolutions and found no significant differences between them. However, their work was limited by the relatively small ensemble sizes. Here, we examine if resolution dependence of the response emerges in very large-ensembles.

To answer the two questions, we present analyses of a ~2200 member ensemble of 13-month simulations run with the lower-resolution version (~90 km, spatial resolution, N144) of Met Office Hadley Centre global atmospheric model Version 4 (HadAM4) and a ~1500 member of 5-month simulations run with the higher-resolution version (~60 km, spatial resolution, N216). We focus on the winter (December, January and February) weather and climate response to projected future Arctic sea-ice loss under 2 °C global warming above preindustrial levels, following the protocol introduced by PAMIP. The response is constructed by comparing an experiment (pdSST-futArcSIC) with prescribed future Arctic sea-ice concentration (SIC) and present-day sea surface temperature (SST), and a control experiment (pdSST-pdSIC) with prescribed present-day SIC and SST. All these SIC/SST boundary conditions were taken from the PAMIP project. These very large-ensemble climate model simulations were run with HadAM4 on the University of Oxford's innovative distributed computing project (Climateprediction.net; CPDN) which allows larger ensemble sizes than in many other experiments. Details of the model setup and experiments are provided in the 'Methods' section. Very large-ensemble climate model simulations based on CPDN have been highly successful in extreme event attribution (e.g., heat waves and floodings)[52,53]. The very large ensembles in our model experiments can sample the diverse atmospheric dynamics and extreme events, and can potentially capture atmospheric flow structure and extreme events that do not exist in relatively small

[1]Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, UK. [2]Oxford e-Research Centre, Engineering Science, University of Oxford, Oxford, UK. [3]School of Geographical Sciences, University of Bristol, Bristol, UK. [4]Department of Mathematics and Statistics, University of Exeter, Exeter, UK. ✉email: kunhui.ye@physics.ox.ac.uk
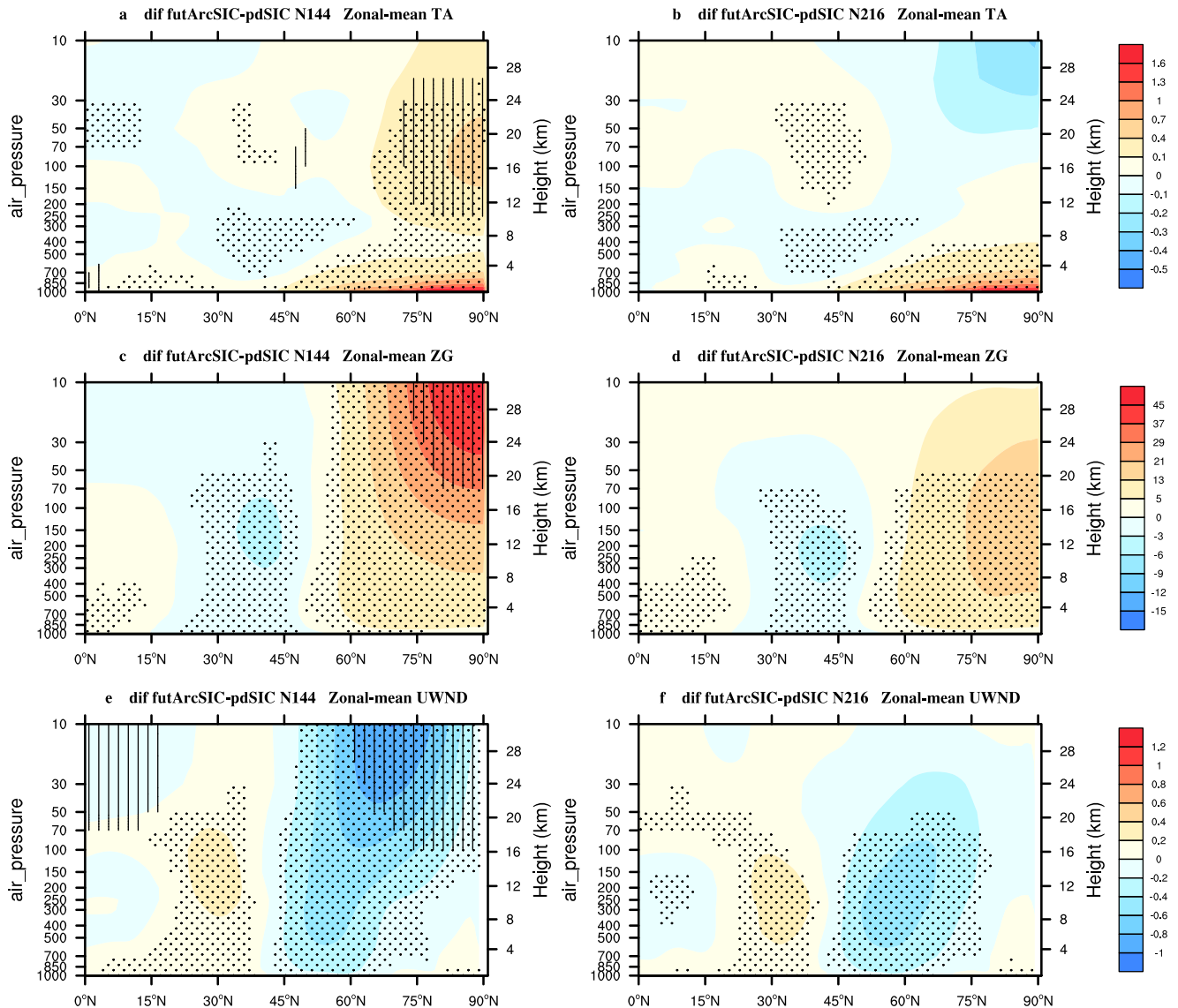
Fig. 1 **Response of latitude-pressure cross sections of zonal mean variables to projected Arctic sea-ice loss.** Response of **a**, **b** atmospheric temperature, **c**, **d** geopotential height, and **e**, **f** zonal-mean zonal wind to projected Arctic sea-ice loss. Left (right) panels are for N144 (N216) resolution. Units: °C, m s^{-1}, and gpm for temperature, zonal wind and geopotential height, respectively. Stippling indicates significance at the 5% level for the differences between the future and present day. Hatching (vertical lines) indicates significance at the 5% level for the differences in the response between the two resolutions.

ensembles. Therefore, these very large ensembles offer considerably better sampling of internal atmospheric variability and extreme events than existing single-model ensembles. In addition, we compare model output for two different resolutions (N144 and N216) to evaluate possible resolution-dependence of response to Arctic sea-ice loss.

RESULTS
Seasonal-mean weather and climate response, resolution dependence and comparison to PAMIP

We begin by examining the zonal-mean atmospheric response. The warming is unsurprisingly largest in the Arctic region and in the lowermost atmosphere; robust warming (≥1 °C), is seen extending equatorward to around 60°N and upward to around 850 hPa for N144 resolution (Fig. 1a). Another prominent region of warming is located around the lower stratosphere north of 60°N, albeit with weaker magnitude than its surface counterpart. The

geopotential height response shows a reduced meridional gradient (Fig. 1c). The Arctic warming response induces a negative zonal wind response between 45°N and 80°N and a positive response centered around 30°N (Fig. 1e), suggesting an equatorward shift of the jet and a weakening of the stratospheric polar vortex. Compared to N144 resolution, the lack of robust Arctic stratospheric responses for N216 resolution is the most important difference (cf. Fig. 1a, e, b, f). The differences between the two resolutions are statistically significant (shown by vertical lines) over the Arctic stratosphere for temperature (cf. Fig. 1a, b) and geopotential height (cf. Fig. 1c, d), and over both the Arctic and tropical stratosphere for zonal wind (cf. Fig. 1e, f). The tropospheric responses are not significantly different and therefore, do not seem to strongly depend on the stratospheric response, at least for these zonal mean features.

The tropospheric responses are similar to those from the PAMIP multimodel ensemble (see their Figs. 1–3)[9], as seen in surface air temperature (SAT), vertical profile of zonal-mean temperature,

sea-level pressure (SLP) and zonal wind. Our results reinforce the finding that the tropospheric response to projected Arctic sea-ice loss is robust across models whilst the stratospheric response is not[9].

We now turn to study spatial features of the responses to projected Arctic sea-ice loss. Figure 2 shows the spatial patterns of responses for SAT, precipitation, and SLP. Robust and large warming is found in the Arctic region with weaker warming found farther south (Fig. 2a), likely representing advective mixing of the warming response to neighbouring regions. Two conspicuous centers of significant but weak cooling are seen in northwest Europe and East Asia (note the asymmetry of the colour scale). The strongest increase in precipitation is found in regions with the largest sea-ice loss and warming (Fig. 2c), most likely caused by strong local warming and evaporation. There is a significant precipitation response dipole in the North Atlantic with wetting centered around 40°N and drying centered around 60°N, consistent with the PAMIP multimodel ensemble[54]. A significant drying response is also seen in Northern Eurasia and Northwest Pacific. In terms of SLP response, a significant negative SLP response is collocated with the largest sea-ice loss and resembles a 'heat low' feature (Fig. 2e). The negative NAO response is robust, and a band of positive SLP response is seen in Eurasia with indication of a strengthening Siberian High. The negative NAO response (dynamics) is balanced by the Arctic warming effect (thermodynamics) in controlling the cooling response in Southern Europe, which makes the temperature response in this region highly uncertain. Comparing the two resolutions suggests that the responses are mostly similar and statistically indistinguishable (cf. Fig. 2a, c, e, b, d, f). Robust atmospheric circulation responses in winds, storm track activity and atmospheric blocking frequency are also simulated (Supplementary Fig. 1). Again, the tropospheric responses are very similar to those in the PAMIP multimodel ensemble, except that the cooling responses in the two resolutions in Europe and Siberia are not seen in PAMIP; this suggests that these are model dependent features of the response and/or that a large ensemble is needed to detect them in response to projected Arctic sea-ice loss.

The highly similar tropospheric responses between the two resolutions studied here are notable. This provides further evidence for the minimal role of the differing stratospheric responses in modulating the tropospheric responses to projected Arctic sea-ice loss.

North Atlantic/East Asian jet stream
The influence of AA on midlatitude jet has been contentious[55]. The trimodality of the North Atlantic jet[56] is evident in both the pdSST-pdSIC (Fig. 3a, black bars) and pdSST-futArcSIC (Fig. 3a, red hatched bars) experiments for N144 resolution. The jet is more often located in central and southern latitudes and less often at northern jet latitudes in response to projected Arctic sea-ice loss, consistent with the equatorward shift of the zonal wind (Fig. 1). Similarly, the occurrence of weak jets is increased in response to projected Arctic sea-ice loss (Fig. 3c; see the method section for defining jet speed in a particular domain). Despite the significant weakening of the jet, no change in jet persistence is found; similar frequencies of persistent jet events are seen for both experiments (Fig. 3e), including for extremely rare jet persistence events occurring at the most southern and northern latitudes. This suggests that the mechanisms driving the jet latitude and speed responses to projected Arctic sea-ice loss do not alter jet persistence. Compared to N144, the trimodal distribution of jet latitudes is less prominent for N216 and a weaker overall southward shift is seen (Fig. 3b). The seasonal-mean jet latitude shift is smaller in N216 than in N144 (−0.074 versus −0.36 degree, the difference is significant at the 1% level). The change in the jet speed is similar between N216 and N144, both in the seasonal

mean (−0.21 versus −0.24 m/s) and the daily distributions (Fig. 3d). The jet persistence response is also not robust for most of the latitude bands in N216 (Fig. 3f). The different jet latitude regime structure and the different stratospheric response may play a role in driving the jet latitude response difference between the two resolutions.

The East Asian jet distribution is unimodal in both resolutions (Fig. 4a, b), albeit with a flatter shape for N216 resolution. Consistent with the North Atlantic jet response, the East Asian jet also shows a southward shift in response to projected Arctic sea-ice loss, which is again weaker in N216 than N144; the seasonal-mean shift in N216 is −0.021 degrees, roughly half of the shift in N144 (−0.053 degrees). However, more strong jet speed days are seen in response to projected Arctic sea-ice loss (Fig. 4c, d) with respective seasonal-mean jet speed response of 0.25 and 0.38 m/s in N144 and N216 respectively (the difference is significat at the 10% level), suggesting a strengthened jet for both resolutions. This is consistent with the strengthening Siberian High and cooling response in Asia, characteristic of a stronger winter East Asian monsoon[57]. Again, the response in the jet persistence is weak and mostly not significant (Fig. 4e, f). The North Atlantic and East Asian jets respond with opposite signs to the projected Arctic sea-ice loss in terms of jet speed. This highlights regional differences in response and the regional cooling over the European and East Asian regions is linked to these different regional dynamical responses.

Daily temperature and precipitation extremes
Arctic sea-ice loss and the associated warming have been linked to a decrease in temperature variability and, more contentiously, to more frequent cold extremes[1,44,58] in the Northern Hemisphere (NH). The response of daily temperature variability in the winter season is characterized by significant increases over Hudson Bay, the eastern and inner Arctic Ocean, and significant decreases in a large part of mid-high latitude oceans and land areas (Fig. 5a). The largest change in temperature variability is found over the oceanic areas where sea ice is lost. Overall, decreases in daily temperature variability are dominant in the response, alluding to the thermodynamic effect of Arctic warming[59,60]. These are consistent between the two resolutions.

To understand the extreme temperature change in response to projected Arctic sea-ice loss, temperature differences are given for the lower 5 and 1% (Fig. 5c, e) and for the lower 0.1% (Supplementary Fig. 2) of the daily temperature distributions for N144 resolution. Corresponding results for upper 5, 1 and 0.1% are displayed in Fig. 6. Interestingly, the patterns resemble those of the seasonal-mean SAT response (see Fig. 2a). The significance of the daily temperature distribution is indicated by the stippling in Supplementary Fig. 2a, b. The change is dominated by less severe cold extremes as the warming difference is ubiquitous. Cold extremes will be less severe particularly for those areas with largest seasonal-mean warming. Our simulations do not support a link between Arctic sea-ice loss and more frequent or severe North American cold extremes as suggested in previous studies[1,33]. Over these regions, the thermodynamic effects of projected sea-ice loss are dominant. In contrast, at least in our simulations, sea-ice loss (in isolation from the other effects of greenhouse warming) makes cold extremes more severe over Asia (Fig. 5c, e). The dynamical response to projected sea-ice loss, including the strengthening Siberian High and East Asian jet, appears a more important driver of changes in East Asian cold extremes than the thermodynamical response (i.e. more severe cold extremes occur despite the warmer Arctic air mass).

To characterise the change in daily temperature variability, we now consider whether the temperature distribution shows an equal shift of its cold and hot tails. Comparing Figs. 5, 6 shows that the shift of the distribution is generally not equal, with one tail
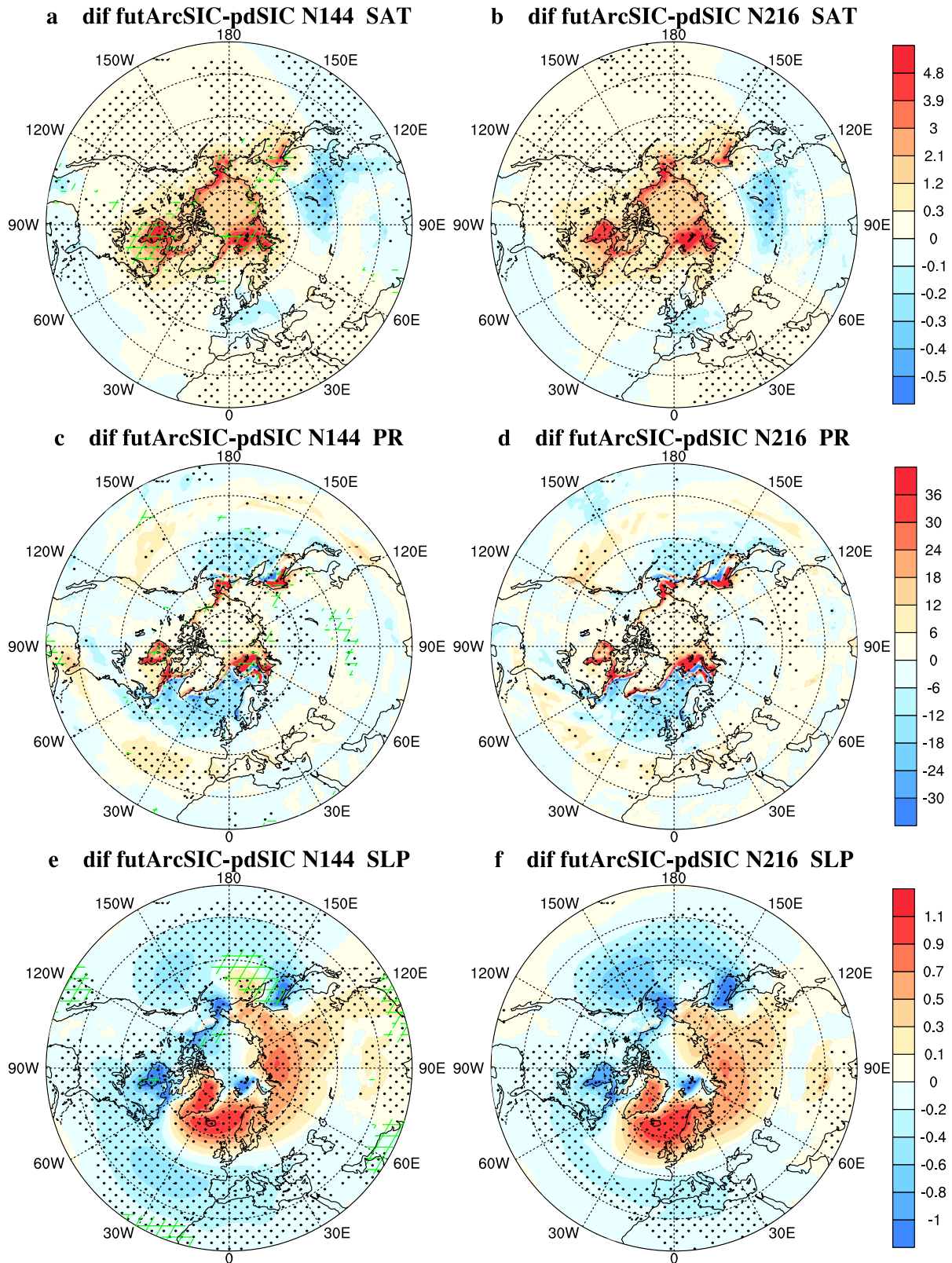
Fig. 2 Response of surface variables to projected Arctic sea-ice loss. Response of **a**, **b** surface air temperature, **c**, **d** precipitation and **e**, **f** sea level pressure. Left (right) panels are for N144 (N216) resolution. Units: °C for SAT, mm for precipitation and hPa for sea level pressure. Stippling indicates significance at the 5% level for the differences between the future and present day. Hatching (mesh) indicates significance at the 5% level for the differences in the response between the two resolutions.
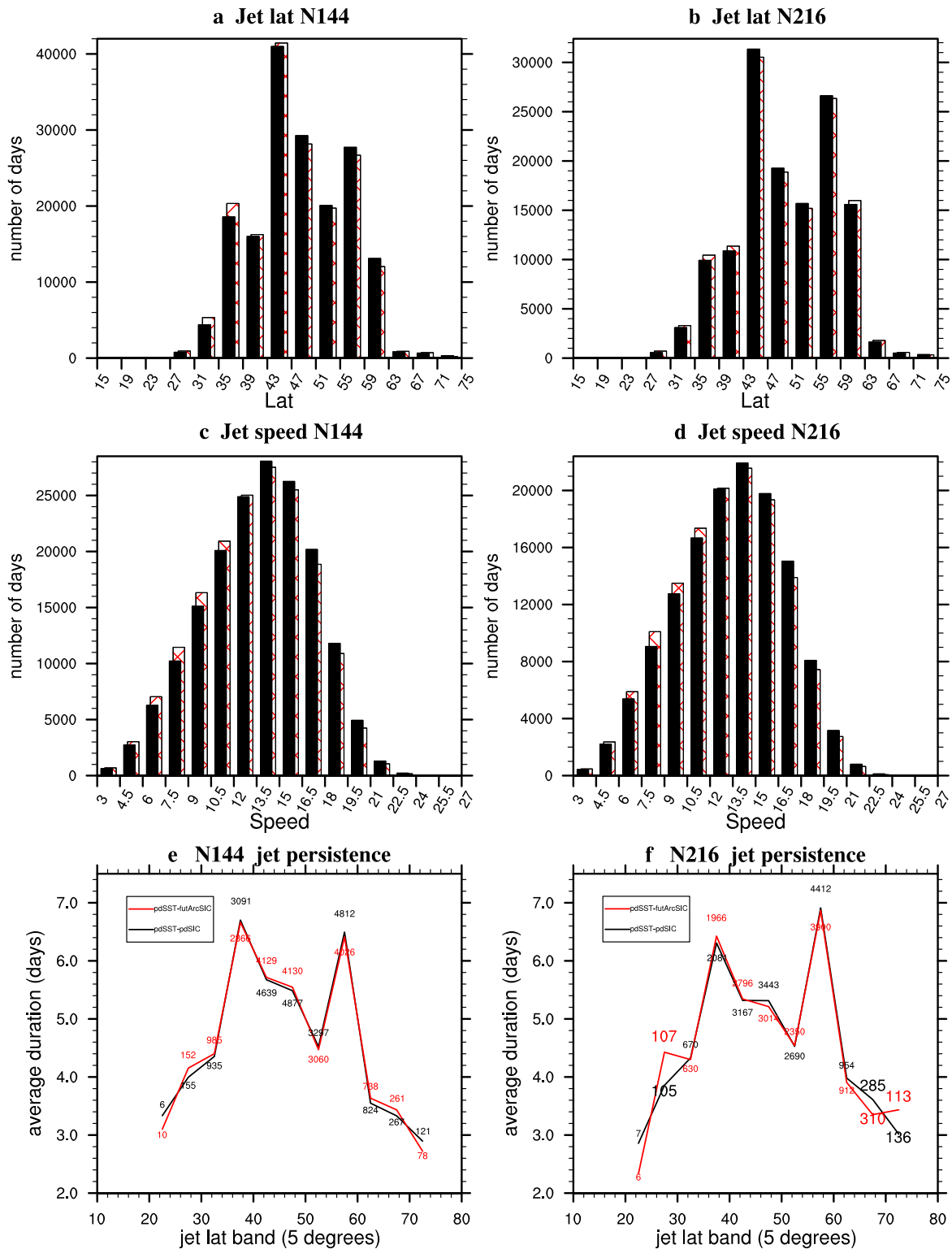
North Atlantic Jet response

a Jet lat N144

b Jet lat N216

c Jet speed N144

d Jet speed N216

e N144 jet persistence

f N216 jet persistence

Fig. 3 Histograms of North Atlantic jet parameters for the two experiments. Histograms of North Atlantic **a**, **b** jet latitude, **c**, **d** speed and **e**, **f** persistence event duration as a function of 5 degrees latitude band. Black (red) bars represent pdSST-pdSIC (pdSST-futArcSIC) experiment. Left (right) panels are for N144 (N216) resolution. In **a**–**d**, the distributions are significantly different at the 5% level according to the Kolmogorov-Smirnov two-sample test. In **e** and **f**, numbers in black (red) denote the number of persistence events for pdSST-pdSIC (pdSST-futArcSIC) experiment; Numbers with larger font size indicate significance at the 5% level. Units: latitude north, m s^{-1}, and days for jet latitude, jet speed and jet persistence duration, respectively.
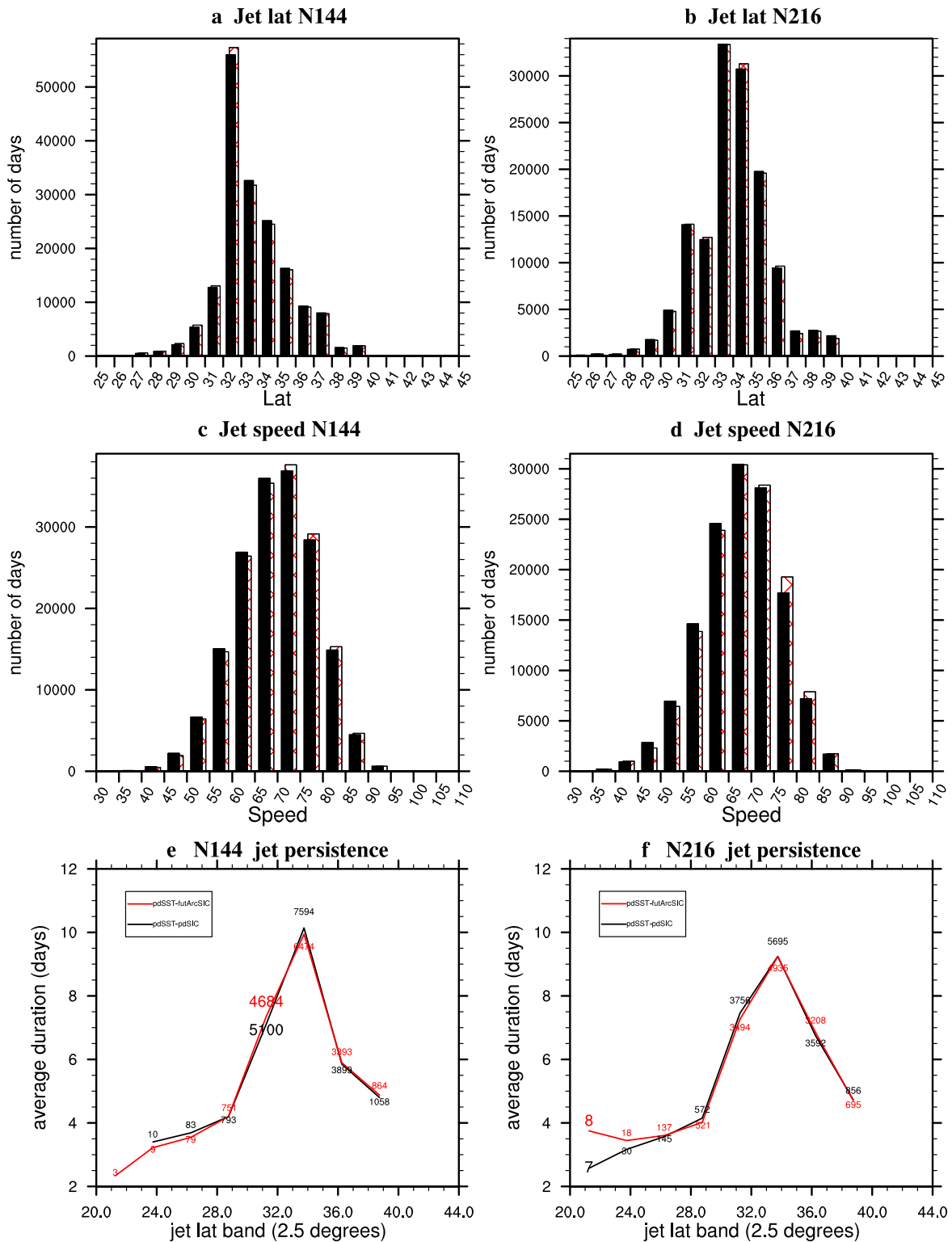
East Asian Jet response

Fig. 4 Histograms of East Asian jet parameters for the two experiments. Histograms of East Asian **a**, **b** jet latitude, **c**, **d** speed and **e**, **f** persistence event duration as a function of 2.5 degrees latitude band. Black (red) bars represent pdSST-pdSIC (pdSST-futArcSIC) experiment. Left (right) panels are for N144 (N216) resolution. In **a–d**, the distributions are significantly different at the 5% level according to the Kolmogorov-Smirnov two-sample test. In **e** and **f**, numbers in black (red) denote the number of persistence events for pdSST-pdSIC (pdSST-futArcSIC) experiment; Numbers with larger font size indicate significance at the 5% level. Units: latitude north, m s^{-1}, and days for jet latitude, jet speed and jet persistence duration, respectively.

SAT
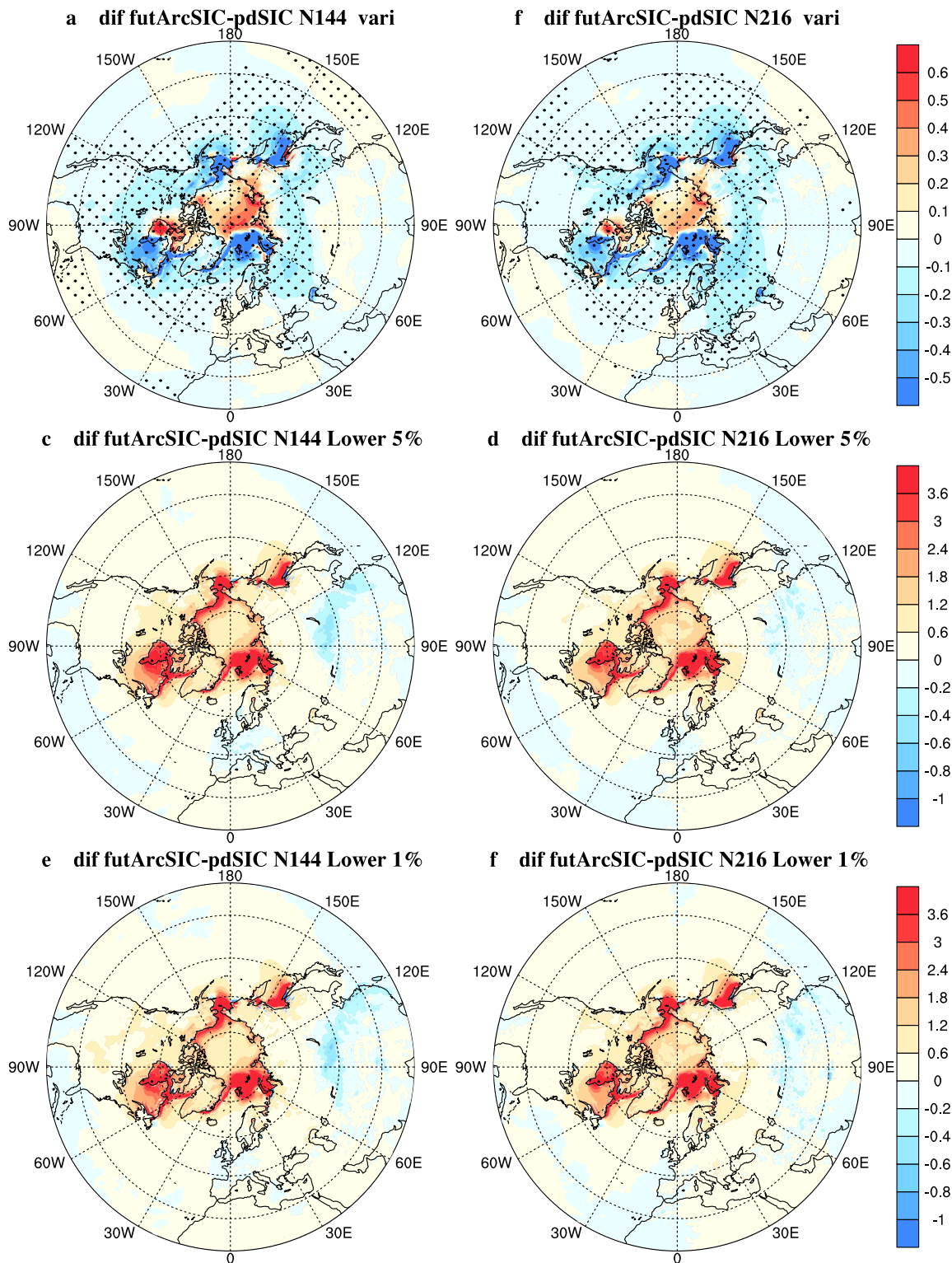
a   dif futArcSIC-pdSIC N144 vari

f   dif futArcSIC-pdSIC N216 vari

c   dif futArcSIC-pdSIC N144 Lower 5%

d   dif futArcSIC-pdSIC N216 Lower 5%

e   dif futArcSIC-pdSIC N144 Lower 1%

f   dif futArcSIC-pdSIC N216 Lower 1%

Fig. 5   Response of temperature variability and cold temperature extremes to projected Arctic sea-ice loss. Response of **a**, **b** daily surface air temperature variability, **c**, **d** lower 5% and **e**, **f** lower 1% in surface air temperature to projected Arctic sea-ice loss. Right (left) panels are for N144 (N216) resolution. Units: °C.
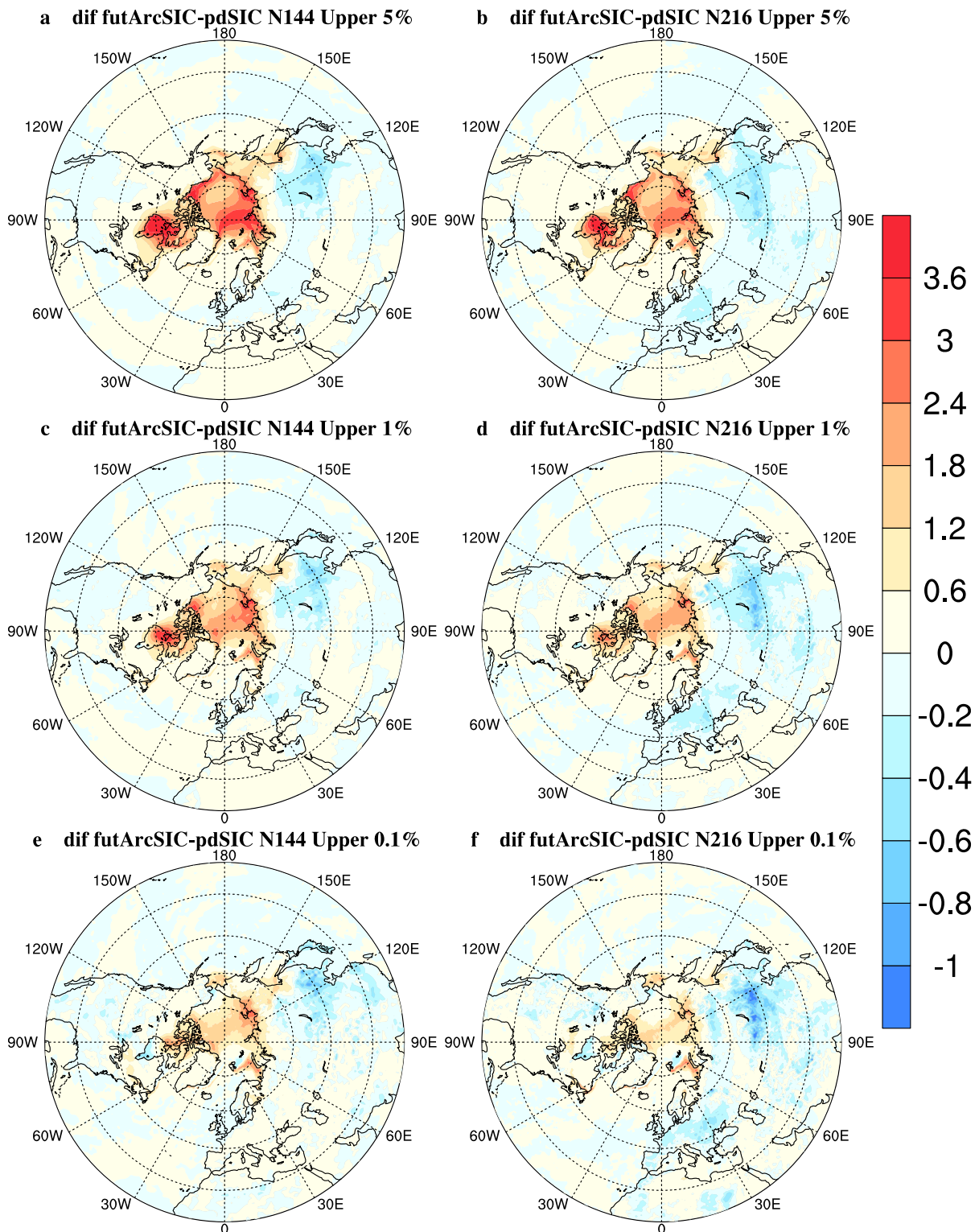
SAT

Fig. 6 Response of warm temperature extremes to projected Arctic sea-ice loss. Response of **a**, **b** upper 5%, **c**, **d** upper 1% and **e**, **f** upper 0.1% in surface air temperature to projected Arctic sea-ice loss. Right (left) panels are for N144 (N216) resolution. Units: °C.

affected more than the other for most of the regions except for southern Europe and particularly Asia. Over the Arctic Ocean, there is stronger warming of the warm tail of the distribution than the cold tail and hence, a widening of the distribution and increase in variability; likely due to the closer proximity of open ocean from where an influence is felt during warm events. The situation is reversed for east Canada, the Barents/Kara Seas, Bering Sea and Sea of Okhotsk, where loss of local sea ice strongly warms

the cold events, leading to a narrowing of the temperature distribution and decrease in variability. Similarly, warming of the cold tail and weak cooling of the warm tail leads to narrowing of the temperature distribution and reduction of temperature variability over North America and Northern Eurasia. For Asia, the cooling of the warm tail is comparable to that of the cold tail. Hence, there is small change in the daily temperature variability but also suggests that the dynamic cooling effects is relatively stronger in this region than others. Similar features are found for N216 (Figs. 5d, f, 6b, d) except that the change in cold extremes is weaker than in N144 but the change in warm extremes is stronger than in N144. Therefore, the response of extreme temperature to projected Arctic sea-ice loss is generally robust across different criteria and between different resolutions.

To further understand the influences of projected Arctic sea-ice loss on the daily SAT distribution, we study the latitude-weighted area-mean SAT distributions for Northern Europe, Asia, and East Asia (Fig. 7). The change in the number of days in percentage relative to the control experiment is highlighted in red. Over Northern Europe, Arctic sea-ice loss clearly reduces the number of extreme and moderate cold days and increases the number of extreme and moderate warm days, for both resolutions (Fig. 7a, b). This is consistent with the warming response over the region being dominated by thermodynamic effects. The most consistent shift towards increasing cold extremes is found for Asia (Fig. 7c–f) and particularly East Asia (Fig. 7e, f) for both resolutions. This is also highlighted in the percentile difference maps in Fig. 5. The sign of this response suggests that dynamical effects emerge as the dominant driver for changing cold extremes in these regions, at least in the models used here.

The link of Arctic sea-ice loss and Arctic warming to Eurasia temperature cooling/extremes has been controversial with the notion that either causality is not clear, or model ensembles are too limited to infer robust response. Our very large-ensemble simulations support the physical link, and Asia and particularly East Asia are the two notable regions where sea-ice loss favours more cold extremes. This is linked to strengthening of the Siberian High and faster East Asian jet, which usually accompanies a stronger East Asian winter monsoon. However, the results strongly suggest that extreme temperature changes (and the associated seasonal-mean changes) are relatively weak and likely hard to detect in the real world given other factors including the global warming.

Another question concerns the influence of projected Arctic sea-ice loss on precipitation extremes (Fig. 8a, b). This has been less explored although Arctic sea-ice loss and Arctic warming are linked to extreme European snowfall[61] and increase in precipitation in Eurasia[62]. The patterns of extreme precipitation response are reminiscent of the seasonal-mean precipitation response and the response is consistent between the extreme and seasonal-mean precipitation. The daily precipitation distribution response is only statistically significant in limited areas for both resolutions (stippling), mostly in the North Atlantic, Arctic and Northwest Pacific. As expected, regions with the largest warming and sea-ice loss will have more extreme precipitation days and this mainly represents the thermodynamic effects. Dynamical effects likely play a leading role in the North Atlantic, North Pacific and Northern Eurasia as in the seasonal-mean precipitation response. Our very large-ensemble simulations suggest that outside the Arctic regions, the precipitation extremes are influenced minimally by the expected increase in evaporation due to Arctic sea-ice loss. Precipitation response in these simulations outside the Arctic regions is instead constrained mostly by the dynamical response, for example the negative NAO response and equatorward storm track shift.

Ensemble-size dependence in weather and climate response to projected Arctic sea-ice loss

To test the importance of ensemble size we systematically form sub-samples of varying size from the large pool of all available members (see 'Methods' for details). The sub-sampling randomly draws sub-samples of desired size (e.g., 100) from the large ensembles with replacement being allowed to construct a large number of samples. The analyses are performed for both the seasonal-mean (Fig. 9 and Supplementary Figs. 3, 6) and extremes responses (Fig. 10 and Supplementary Fig. 4) for a set of critical variables, considering the standard deviation and the 95% confidence range of the 100,000 samples for the seasonal-mean and 5000 samples for extremes responses.

We first discuss the uncertainty for the seasonal-mean response as a function of ensemble size from 100 to 1900 members for N144 (Fig. 9a, c) and from 100 to 1400 members for N216 (Fig. 9b, d). All variables exhibit large uncertainties for 100 members, which is typical for many PAMIP models and other model experiments. In particular, the NAO response can range from −3 hPa to 1 hPa, showing both sign and magnitude uncertainty. Internal atmospheric variability may therefore explain much of the disagreement on NAO response to Arctic sea-ice loss in existing studies. The uncertainty range decreases considerably with increasing ensemble size. The sign of the NAO response is considered to be reasonably constrained when the 95% range excludes zero, and this is seen to require at least 200 members at N144 and 400 members at N216. An ensemble size larger than 100 would already narrow the uncertainty and well constrain the size of the tropospheric zonal-mean zonal wind response. The warming response in the lower Arctic atmosphere, Northern Europe and Western North America are already robust for an ensemble size of 200. The temperature response in Southern Europe could have either sign even for large ensemble size but that in Asia (Fig. 9) and East Asia (Supplementary Fig. 3) is more confident to be a cooling response for very large ensembles such as 1000 members. The temperature response in Southern Europe cannot be separated from internal variability even for a large ensemble size. This might reflect the lack of any response or potentially conflicting dynamical and thermodynamical influences from Arctic warming and the equatorward jet shift.

We further provide quantitative uncertainty analysis for these variables to show the standard deviation relative to that for a sample of 100 members. This removes the impact of different units for the response in different valuables in Fig. 9a, b, and shows that these uncertainty lines all collapse onto the relation (10/√N), where N is ensemble size. Hence the standard error σ/√N provides an excellent prediction of the uncertainty associated with a sample of size N. It is obvious that the uncertainty reduction is less significant as ensemble size increases. The uncertainty is roughly reduced by 50% (upper horizontal line) and 70% (lower horizontal line), respectively, for a member size of 400 and 1100. We select these two ensemble sizes (400 and 1100) for reference when quantifying the expected reduction in uncertainty in response to future Arctic sea-ice loss. As will be seen below, these relationships hold for both seasonal-mean variables and extremes. In particular, an ensemble size of 400 is necessary to constrain the sign of the NAO response in the high-resolution (Fig. 9b).

Very large-ensemble simulations are particularly important for sampling extreme events. Here, we focus on the lower 1% SAT (Fig. 10) and upper 5% precipitation (Supplementary Fig. 4) responses to quantify the impacts of ensemble size on the uncertainty in the response of extremes. For a member size of 100 the extreme temperature response shows large uncertainty in both sign and magnitude. As in the seasonal-mean response, the uncertainty range narrows as ensemble size increases but the sign uncertainty is still large even for relatively large ensembles. For Northern Europe, the sign of the warming response is well constrained for an ensemble size of 400, For southern Europe, a response with sign constrained needs
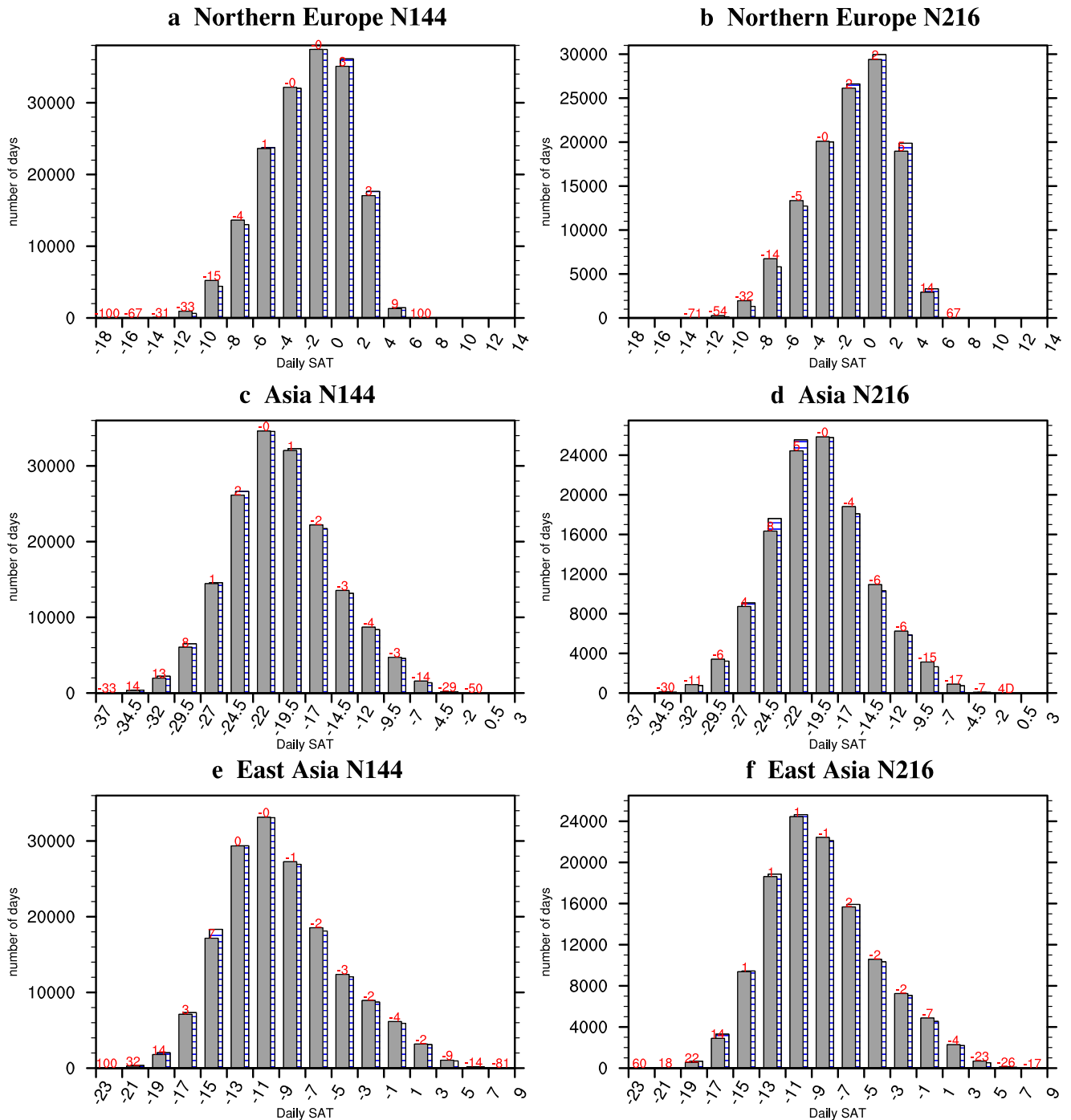
Grey: pdSIC   Blue: futArcSIC

Fig. 7 **Histograms of area-mean surface air temperature for different regions. a**, **b** Northern Europe; **c**, **d** Asia; **e**, **f** East Asia. Black (blue) bars represent pdSST-pdSIC (pdSST-futArcSIC) experiment. Left (right) panels are for N144 (N216) resolution. Red numbers denote the change in percentage (%; "–" for negative change) relative to pdSST-pdSIC experiment. Numbers with 'D' indicate the absolute change of days if there is no data value in pdSST-pdSIC experiment. Units: °C.

an ensemble size of around 1000 and then less extreme temperature is expected. This means that the true response of extreme temperature in southern Europe to projected Arctic sea-ice loss is likely thermodynamically controlled. Similarly, it also needs a large ensemble size to have a response with sign constrained for both Asia and East Asia in N144 but in this case colder extreme temperatures are expected. The extreme temperature response in Asia is still highly

uncertain in terms of sign even for large ensembles in N216 as the response is weaker than in N144 (cf. Fig. 5d, f, c, e). In contrast, colder extreme temperature in East Asia is more confident, in particular in N144 in both resolutions with N216 response slightly straddling the '0' line. In terms of quantitative analysis of uncertainty reduction (Fig. 10c, d), the standard error is again found to give a very good prediction of the reduction in uncertainty expected from a given
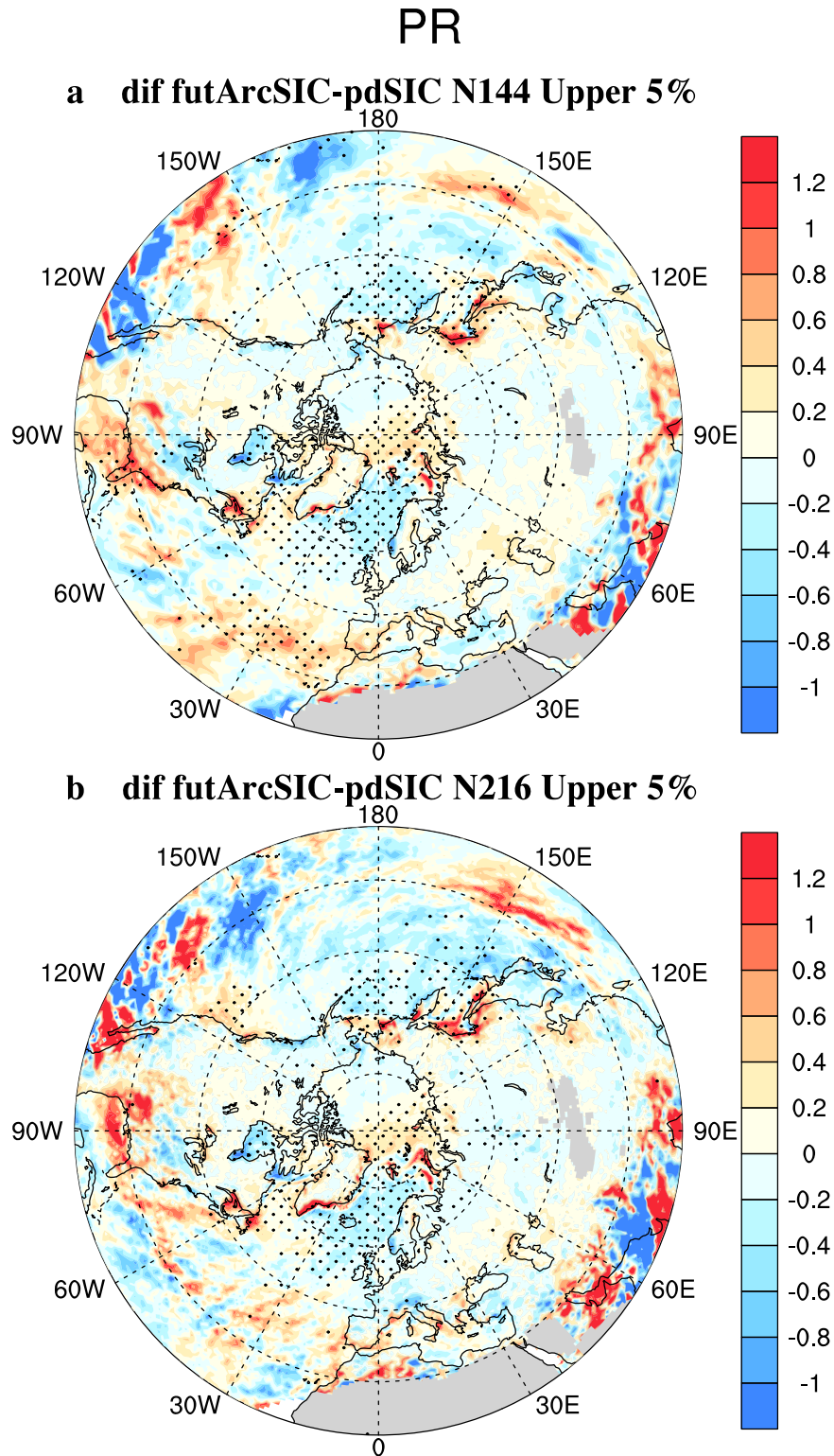
PR

a dif futArcSIC-pdSIC N144 Upper 5%

b dif futArcSIC-pdSIC N216 Upper 5%

Fig. 8 Response of upper 5% in daily precipitation based on wet days only to projected Arctic sea-ice loss. a N144 resolution; **b** N216 resolution. Stippling indicates that distributions are significantly different between the future and present day at the 5% level according to the Kolmogorov–Smirnov two-sample test. Units: mm day^{-1}. Wet days refer to daily precipitation ≥1 mm.

increase in ensemble size. However, much larger ensembles are necessary to detect robust extremes response given their rarity and complex drivers.

For precipitation extremes, the sign of the response is relatively well constrained for a very large ensemble size (~1000) except for the southern part of the North Atlantic for both resolutions. In response to projected Arctic sea-ice loss, less extreme precipitation is expected for a very large ensemble size for these regions. This suggests the dominant role of dynamics in driving the extreme precipitation response. The quantitative analysis of
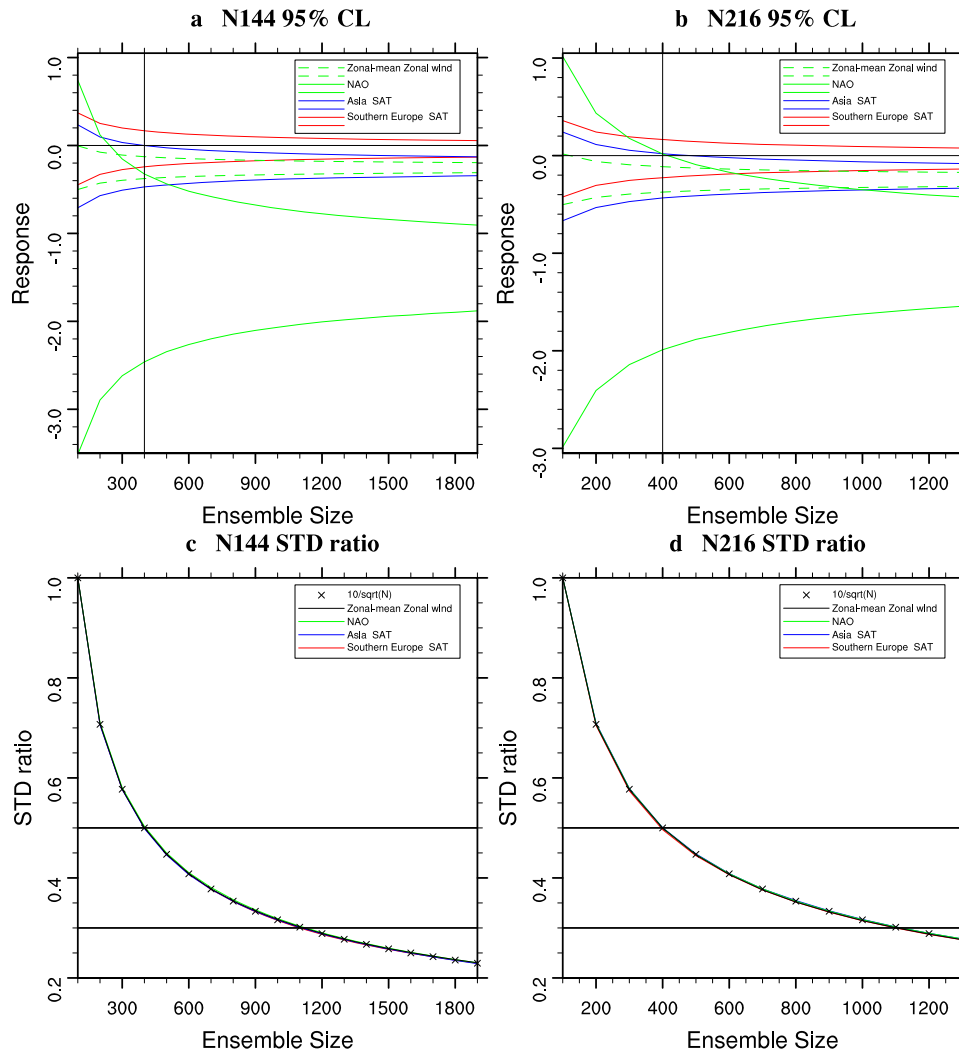
DJF-mean response

Fig. 9 (Top) Sub-sampled 95% range of the ensemble-mean response to projected Arctic sea-ice loss as a function of ensemble size and **(bottom)** the standard deviation ratio against 100-member size as a function of ensemble size. **a, c** N144 resolution; **b, d** N216 resolution. The vertical line at the top panels indicates 400 members. The upper (lower) horizontal line at the bottom panels indicates 0.5 and 0.3, respectively. Units: m s−1, hPa and °C for tropospheric zonal-mean zonal wind, NAO, and temperature response, respectively.

uncertainty reduction for extreme precipitation (Supplementary Fig. 4c, d) is again similar to the seasonal-mean response.

The results for both extreme temperature and precipitation suggest that an ensemble size of 1000 is necessary to constrain the sign of the response to this forcing, while reducing roughly 70% uncertainty relative to an ensemble size of 100.

DISCUSSION

Our very large-ensemble climate simulations are based on an atmosphere-only model, and coupled climate model simulations would provide further insights into the response to projected Arctic sea-ice loss[50,63]. We note that coupled climate model simulations may introduce more complicated coupled internal variability like El Niño–Southern Oscillation that will obfuscate and modulate true signals[50]. Projected Arctic sea-ice loss is undoubtedly only one of the aspects of climate change that influences weather and climate projections[1]. Further comparisons with other factors such as global SSTs are useful to understand the broader picture, including the tug

of war between Arctic warming and tropical warming in driving the NH climate and weather variability and change.

The marked differences in stratospheric response between the two resolutions and the weaker jet latitude response in N216 suggest some possible modulating effects of stratospheric circulation. The climatological stratospheric circulation is much stronger in N216 than in N144 that may be associated with weaker wave-driving effects (Supplementary Fig. 5). A stronger mean-state of the polar vortex can inhibit tropospheric forcing of the stratosphere[64] and sudden stratospheric warmings[65]. The downward impacts of the polar vortex on the North Atlantic jet are well known, and the lack of significant polar vortex response may translate into uncertainty in the North Atlantic jet shift in N216[66]. Even for very large ensemble size (≥1000), the sign of the polar vortex, upper-level (50–200 hPa) Arctic temperature and the North Atlantic jet latitude responses in N216 (Supplementary Fig. 6) are still uncertain. This is very different from the situation in N144. This seems to support the role of modulating effects from stratospheric circulation. We have also performed sub-sampling for N216 to compare the responses of North Atlantic jet latitude and the NAO
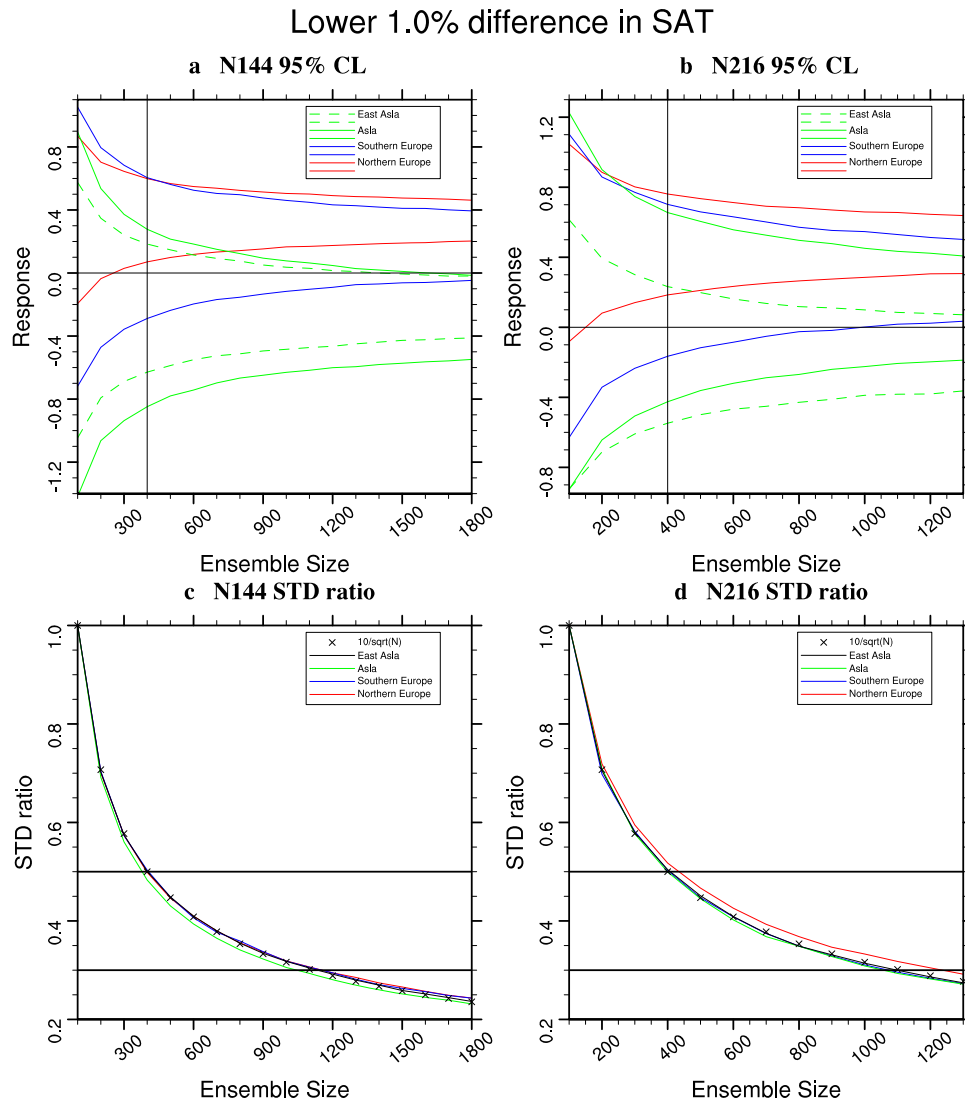
Lower 1.0% difference in SAT

Fig. 10 (Top) Sub-sampled 95% range of the ensemble-mean response of extreme temperature measured by the lower 1% temperature difference to projected Arctic sea-ice loss as a function of ensemble size and (bottom) the standard deviation ratio against 100-member size as a function of ensemble size. a, c N144 resolution; b, d N216 resolution. The vertical line at the top panels indicates 400 members. The upper (lower) horizontal line at the bottom panels indicates 0.5 and 0.3, respectively. Units: °C.

between sub-samples having similar stratospheric polar vortex response to N144 and those having negligible response (Supplementary Fig. 7). The results suggest that the stratospheric response may weakly modulate the tropospheric response in our experiments and further model experiments, involving switching on/off stratospheric pathway, are needed to clarify the issue.

Our simulations agree with PAMIP results that the winter tropospheric circulation response to projected Arctic sea-ice loss is robust but weak compared to interannual variability. This includes a robust equatorward shift of jets, weakening of midlatitude westerlies, increase in mid-high latitude atmospheric blocking frequency, weakening of mid-high latitude storm track activities, and robust climate response. The influences on the extreme daily temperature and precipitation are also robust in many regions across the NH. These are largely consistent with the seasonal-mean climate response, with the exception of well-understood changes in temperature variability linked to thermal advection[67,68]. Over the land areas in the NH, a decrease in daily temperature variability is hence the dominant response as warming effects play a major role. East Asia is a notable exception in showing an increase in severe

cold temperature extremes. This is likely due to the dominant dynamic effects in these regions and the very large ensembles which allow this very weak signal to be extracted. The changes in precipitation extremes are mainly located over the North Atlantic and Northwest Pacific, shaped predominantly by dynamic effects.

Uncertainty analysis using a sub-sampling method supports the finding that a large ensemble is necessary to extract robust responses in both seasonal-means and extremes for projected Arctic sea-ice loss. For this forcing, large ensembles (≥400) are needed to robustly estimate the seasonal-mean large-scale circulation response, and very large ensembles (≥1000) are needed to simulate regional climate and extremes, although it might be possible to statistically approximate changes in extremes from an ensemble of several hundred members. Increasing ensemble size will be particularly important for small-ensemble (e.g., <400 members) model simulations to have more confident estimates of projected Arctic sea-ice loss impacts. Our large ensembles have allowed a deeper understanding of this weak signal and a robust quantification of associated changes in extremes. Extreme events often involve complicated and

persistent dynamical factors, hence providing the motivation for our large ensemble simulations at relatively high resolution. We confirm that no dynamical changes distinct to extremes have been found, for example, in the persistence of jet shifts. Even for extreme events, the reduction in uncertainty with ensemble size is very well predicted by standard error analysis, providing guidance for the design of future large ensembles.

METHODS

CPDN very large-ensemble initial-condition climate model simulations

HadAM4 is the latest version of a climate configuration of the Hadley Centre Unified forecast and climate model[69]. It has some significant improvements over its predecessor, HadAM3, including enhanced vertical resolution, introduction of a cloud area parameterization, a new mixed-phase precipitation scheme and others[70]. The CPDN distributed computing platform enables large-ensemble, high-resolution climate simulations by using donated computing time from a massive number of computers around the world. HadAM4 has been well supported and successfully run on the CPDN distributed computing platform. In our CPDN climate simulations, HadAM4 was configured with 38 vertical levels with a lid of around 4.6 hPa. Two resolutions, N144 resolution (1.25×0.83 degrees) and N216 resolution (0.83×0.56 degrees), were considered for studying resolution dependence of response to projected Arctic sea-ice loss. Two experiments, pdSST-pdSIC and pdSST-futArcSIC, similar to the exp 1.1 and exp 1.6 in PAMIP[9], were run for both resolutions and they differ only in the boundary conditions as mentioned below. For both experiments, the SST/SIC boundary conditions were taken from the PAMIP archive to ensure fair comparisons with PAMIP. Specifically, pdSST-pdSIC used present-day climatology of SST and SIC. For pdSST-futArcSIC, it used present-day climatology of SST over the globe and SIC outside the Arctic while future SIC was used in the Arctic region with future SST also specified if SIC loss is over 10%. All other forcings including greenhouse gases, ozone, volcano, and solar forcings are specified as present-day climatology and the same for both experiments for both resolutions. The model was run from October 1st (November 1st) until October 31st (March 31st) next year for the N144 (N216) resolution. The shorter run duration for the N216 resolution was intended to shorten computing times as our analysis focuses on winter season. Initial conditions for 2500 ensembles were generated by perturbing the potential temperature field using a stochastic method in model output taken from a long HadAM4 model run. Although we have aimed for the maximum number of ensembles for our analysis, the eventual ensembles available for analysis are smaller given the nature of the distributing computing and other issues contributing to loss of members. There are around 2200 members successfully returned and extracted for the N144 resolution and around 1500 for the N216 resolution for the pdSST-pdSIC experiment for most of the valuables and around 200 members fewer for the pdSST-futArcSIC experiment. Owning to limited storage space, monthly variables were outputted at 18 vertical levels (1000 to 10 hPa) while daily variables were outputted at 3 vertical levels (850, 250 and 100 hPa) or 4 vertical levels (850, 250, 100 and 50 hPa).

NAO, Siberian High, polar vortex and tropospheric zonal-mean zonal wind

The NAO is defined as the difference between latitude-weighted area-mean SLP over the Azores (28–20°W, 36–40°N) and Iceland (25–16°W, 63–70°N). The square-root of the cosine of the latitude is used for weighting. The Siberian High is defined as the latitude-weighted area-mean SLP over the domain of 80–120°E, 40–65°N, and the polar vortex is defined as zonal-mean zonal wind averaged over 54–70°N at 10 hPa. Tropospheric zonal-mean zonal wind is defined as zonal-mean zonal wind averaged over 45–60°N and 1000-500 hPa.

North Atlantic/East Asian Jet indices and persistence

For the North Atlantic Jet, daily zonal wind at 850 hPa was used to compute the two jet indices—latitude and speed—following established procedures[56]. A jet persistence event over a latitude band of 5 degrees is identified when the jet latitude falls into the latitude band for at least two consecutive days. The jet persistence duration for an event is simply the number of consecutive days when the jet latitude falls into the latitude band. For the East Asian jet, similar procedures are followed except that the domain is changed to 120°E–180°E, 20°N–40°N and the latitude window for jet persistence is 2.5 degrees using daily zonal wind at 250 hPa. Note that the results are insensitive to the latitude window.

Storm track activity and atmospheric blocking frequency

Storm track activity is measured by the variance of 2.5–6-day band-pass filtered meridional wind. A two-dimensional blocking index is defined following a previous study[71] with some modifications that uses the geopotential height gradient at 500 hPa to identify blocking. Before the computation of the blocking index, the latitude grid was regridded onto 1 (0.75) degree for N144 (N216) resolution for easier computation. Similar to a previous study[72], the blocking index was computed for every grid latitude between 35°N and 75°N. This has allowed the detection of blocking over an extended latitude range.

Daily surface air temperature variability

The daily temperature variability is defined as the standard deviation for the winter season for each ensemble member.

Percentile computation of distribution of daily variables and domains for area-mean

The percentile was computed for all ensembles over the winter season for every grid point. Note that the same ensemble size was used for both the experiments, and this is defined as the smaller ensemble size between the two experiments. The regions used for latitude-weighted area-mean computation are defined as follows. The square-root of the cosine of the latitude is used for weighting. Northern (southern) Europe is bounded by 0°E–20°E, 60°N–70°N (40°N–60°N) for daily area-mean temperature; Northern (southern) Europe is bounded by 0°E–30°E, 60°N–70°N (45°N–60°N) for seasonal area-mean temperature; Siberia is bounded by 90°E–130°E, 45°N–60°N; Asia is bounded by 72°E–120°E, 42°N–58°N for daily area-mean temperature; Asia is bounded by 90°E–120°E, 45°N–60°N for seasonal area-mean temperature; East Asia is bounded by 100°E–120°E, 30°N–45°N; Southern (northern) part of North Atlantic is bounded by 310°E–360°E, 30°N–45°N (45°N–65°N); North Pacific is bounded by 150°E–200°E, 45°N–60°N; Western North America is bounded by 230°E–270°E, 40°N–60°N.

Sub-sampling method for estimating uncertainty in response

This sub-sampling method is similar to a bootstrapping method. We define N as the ensemble size starting from 100 with an interval of 100. The sub-sampling method sub-samples N members from all the available ensembles with replacement. Firstly, the indices (0, 1, 2, etc.) of all the available ensembles are generated. Secondly, a new set of indices with replacement (e.g., same index number can appear twice or more) is randomly drawn from the original set of indices each time and the N members corresponding to the first N index numbers in the new set of indices are selected as one N-member subsample. This is performed for both experiments and the difference between the ensemble-means of the two experiments is defined as a response. For extremes response, the difference between the percentiles of

daily temperature and precipitation in the two experiments is computed. This process is repeated 100,000 times for seasonal-mean variables and, due to demanding computing time, 5000 times for percentile differences of daily temperature and precipitation. Therefore, we obtain 100,000 samples of the response to Arctic sea-ice loss for seasonal-mean variables, and 5000 samples of the response for percentile differences of daily temperature and precipitation. The 95% range and the standard deviation across samples of response can then be computed.

Sub-sampling method for estimating the role of stratospheric response in modulating tropospheric response between the two resolutions

Two groups of sub-sampled responses using a sub-sampling method are obtained respectively for polar vortex response in N216 resembling that in N144 (their absolute difference < 0.01 m s^{-1}) and negligible polar vortex response in N216 (absolute response <0.001 m s^{-1}). The sub-sampling method is similar to the sub-sampling method for estimating uncertainty in response described above but differs in two ways. First, it does not allow replacement and is done for N216 only. Second, each pair of sub-samples (500 members for each set of sub-samples) should have similar polar vortex response to N144 (first group) or negligible polar vortex response (second group). A total of 5000 pairs of sub-samples are obtained for each group to construct the response to projected Arctic sea-ice loss. By comparing the distribution of the response between the two groups and those responses using all available ensembles in both N144 and N216, we can infer the contribution of stratospheric response to the difference in the tropospheric response. As seen from Supplementary Fig. 7, including significant stratospheric response does shift the distribution of both the North Atlantic jet latitude and, particularly the NAO response in N216 toward the N144 response. Comparing this with the difference in the response using all available ensembles between N144 and N216 suggests that the stratospheric response may partly explain the resolution difference in the North Atlantic jet latitude and the NAO response. However, further model experiments are needed to more clearly reveal the causality on this issue.

Significance test

The Student's t test is used for testing the significance of the response to projected Arctic sea-ice loss and the differences in the response between the two resolutions. Kolmogorov-Smirnov two-sample test (KS test) is non-parametric and distribution free and is used to test whether two samples are drawn from the same distribution. The KS test is applied to the distributions of jet latitude/speed, daily SAT and precipitation.

DATA AVAILABILITY

The datasets generated and analysed during the current study are not publicly available due to ongoing work to create a permanent data deposition at Centre for Environmental Data Analysis (https://www.ceda.ac.uk/), but are available from the corresponding author on reasonable request.

CODE AVAILABILITY

The source codes for analyzing the model output in this study are available from the corresponding author upon reasonable request.

REFERENCES
1. Cohen, J. et al. Recent Arctic amplification and extreme mid-latitude weather. Nat. Geosci. 7, 627–637 (2014).
2. Cohen, J. et al. Divergent consensuses on Arctic amplification influence on midlatitude severe winter weather. Nat. Clim. Change 10, 20–29 (2020).
3. Rantanen, M. et al. The Arctic has warmed nearly four times faster than the globe since 1979. Commun. Earth Environ. 3, 168 (2022).
4. Kinnard, C. et al. Reconstructed changes in Arctic sea ice over the past 1450 years. Nature 479, 509–512 (2011).
5. Screen, J. A. & Simmonds, I. The central role of diminishing sea ice in recent arctic temperature amplification. Nature 464, 1334–1337 (2010).
6. Walsh, J. E. Intensified warming of the Arctic: causes and impacts on middle latitudes. Glob. Planet. Change 117, 52–63 (2014).
7. Gao, Y. et al. Arctic sea ice and Eurasian climate: a review. Adv. Atmos. Sci. 32, 92–114 (2015).
8. Screen, J. A., Deser, C. & Simmonds, I. Local and remote controls on observed Arctic warming. Geophys. Res. Lett. 39, 10709 (2012).
9. Smith, D. M. et al. Robust but weak winter atmospheric circulation response to future Arctic sea ice loss. Nat. Commun. 13, 1–15 (2022).
10. Deser, C., Magnusdottir, G., Saravanan, R. & Phillips, A. The effects of North Atlantic SST and sea ice anomalies on the winter circulation in CCM3. Part II: Direct and indirect components of the response. J. Clim. 17, 877–889 (2004).
11. Overland, J. et al. The Melting Arctic and Midlatitude weather patterns: are they connected? J. Clim. 28, 7917–7932 (2015).
12. Blackport, R. & Screen, J. A. Weakened evidence for mid-latitude impacts of Arctic warming. Nat. Clim. Change 10, 1065–1066 (2020).
13. Francis, J. A. & Vavrus, S. J. Evidence linking Arctic amplification to extreme weather in mid-latitudes. Geophys. Res. Lett. 39, L06801 (2012).
14. Wallace, J. M., Held, I. M., Thompson, D. W. J., Trenberth, K. E. & Walsh, J. E. Global warming and winter weather. Science 343, 729–730 (2014).
15. Vihma, T. Effects of Arctic Sea Ice decline on weather and climate: a review. Surv. Geophys. 35, 1175–1214 (2014).
16. Shepherd, T. G. Effects of a warming Arctic. Science 353, 989–990 (2016).
17. McCusker, K. E., Fyfe, J. C. & Sigmond, M. Twenty-five winters of unexpected Eurasian cooling unlikely due to Arctic sea-ice loss. Nat. Geosci. 9, 838–842 (2016).
18. Screen, J. A. et al. Consistency and discrepancy in the atmospheric response to Arctic sea-ice loss across climate models. Nat. Geosci. 11, 155–163 (2018).
19. Dai, A. & Song, M. Little influence of Arctic amplification on mid-latitude climate. Nat. Clim. Change 10, 231–237 (2020).
20. Alexander, M. A. et al. The atmospheric response to realistic Arctic sea ice anomalies in an AGCM during winter. J. Clim. 17, 890–905 (2004).
21. Francis, J. A. & Vavrus, S. J. Evidence for a wavier jet stream in response to rapid Arctic warming. Environ. Res. Lett. 10, 014005 (2015).
22. Overland, J. E., Wood, K. R. & Wang, M. Warm Arctic—cold continents: climate impacts of the newly open Arctic Sea. Polar Res 30, 15787 (2011).
23. Honda, M., Inoue, J. & Yamane, S. Influence of low Arctic sea-ice minima on anomalously cold Eurasian winters. Geophys. Res. Lett 36, https://doi.org/10.1029/2008GL037079 (2009).
24. Petoukhov, V. & Semenov, V. A. A link between reduced Barents-Kara sea ice and cold winter extremes over northern continents. J. Geophys. Res. Atmos. 115, https://doi.org/10.1029/2009JD013568 (2010).
25. Kim, B. M. et al. Weakening of the stratospheric polar vortex by Arctic sea-ice loss. Nat. Commun. 5, 1–8 (2014).
26. Barnes, E. A. Revisiting the evidence linking Arctic amplification to extreme weather in midlatitudes. Geophys. Res. Lett. 40, 4734–4739 (2013).
27. Screen, J. A. & Simmonds, I. Exploring links between Arctic amplification and mid-latitude weather. Geophys. Res. Lett. 40, 959–964 (2013).
28. Sun, L., Deser, C. & Tomas, R. A. Mechanisms of stratospheric and tropospheric circulation response to projected Arctic sea ice loss. J. Clim. 28, 7824–7845 (2015).
29. Nakamura, T. et al. The stratospheric pathway for Arctic impacts on midlatitude climate. Geophys. Res. Lett. 43, 3494–3501 (2016).
30. Peings, Y. Ural blocking as a driver of early-winter stratospheric warmings. Geophys. Res. Lett. 46, 5460–5468 (2019).
31. Zhang, P. et al. A stratospheric pathway linking a colder Siberia to Barents-Kara Sea sea ice loss. Sci. Adv. 4, eaat6025 (2018).
32. Mori, M., Watanabe, M., Shiogama, H., Inoue, J. & Kimoto, M. Robust Arctic sea-ice influence on the frequent Eurasian cold winters in past decades. Nat. Geosci. 7, 869–873 (2014).
33. Kug, J. S. et al. Two distinct influences of Arctic warming on cold winters over North America and East Asia. Nat. Geosci. 8, 759–762 (2015).
34. Wu, B., Yang, K. & Francis, J. A. A cold event in Asia during January–February 2012 and its possible association with Arctic sea ice loss. J. Clim. 30, 7971–7990 (2017).
35. Ye, K. & Messori, G. Inter-model spread in the wintertime Arctic amplification in the CMIP6 models and the important role of internal climate variability. Glob. Planet. Change 204, 103543 (2021).
36. Cohen, J., Jones, J., Furtado, J. C. & Tziperman, E. Warm Arctic, cold continents: a common pattern related to Arctic sea ice melt, snow advance, and extreme winter weather. Oceanography 26, 150–160 (2013).

37. Luo, D. et al. Impact of Ural blocking on winter warm Arctic–cold Eurasian anomalies. Part II: The link to the North Atlantic Oscillation. *J. Clim.* **29**, 3949–3971 (2016b).

38. Inoue, J., Hori, M. E. & Takaya, K. The role of Barents Sea ice in the wintertime cyclone track and emergence of a warm-Arctic cold-Siberian anomaly. *J. Clim.* **25**, 2561–2568 (2012).

39. Ogawa, F. et al. Evaluating impacts of recent Arctic sea ice loss on the northern hemisphere winter climate change. *Geophys. Res. Lett.* **45**, 3255–3263 (2018).

40. Sato, K., Inoue, J. & Watanabe, M. Influence of the Gulf Stream on the Barents Sea ice retreat and Eurasian coldness during early winter. *Environ. Res. Lett.* **9**, 084009 (2014).

41. Simmonds, I. & Govekar, P. D. What are the physical links between Arctic sea ice loss and Eurasian winter climate. *Environ. Res. Lett.* **9**, 101003 (2014).

42. Sun, L., Perlwitz, J. & Hoerling, M. What caused the recent "Warm Arctic, Cold Continents" trend pattern in winter temperatures? *Geophys. Res. Lett.* **43**, 5345–5352 (2016).

43. Sung, M. K., Kim, S. H., Kim, B. M. & Choi, Y. S. Interdecadal variability of the warm Arctic and cold Eurasia pattern and its North Atlantic origin. *J. Clim.* **31**, 5793–5810 (2018).

44. Tang, Q., Zhang, X., Yang, X. & Francis, J. A. Cold winter extremes in northern continents linked to Arctic sea ice loss. *Environ. Res. Lett.* **8**, 014036 (2013).

45. Ye, K., Jung, T. & Semmler, T. The influences of the Arctic troposphere on the midlatitude climate variability and the recent Eurasian cooling. *J. Geophys. Res.* **18**, 10,162–10,184 (2018).

46. Ye, K. & Messori, G. Two leading modes of wintertime atmospheric circulation drive the recent warm Arctic–cold Eurasia temperature pattern. *J. Clim.* **33**, 5565–5587 (2020).

47. Mori, M., Kosaka, Y., Watanabe, M., Nakamura, H. & Kimoto, M. A reconciled estimate of the influence of Arctic sea-ice loss on recent Eurasian cooling. *Nat. Clim. Change* **9**, 123–129 (2019).

48. Outten, S. et al. Reconciling conflicting evidence for the cause of the observed early 21st century Eurasian cooling. *Weather Clim. Dyn.* **4**, 95–114 (2023).

49. Smith, D. M. et al. The Polar Amplification Model Intercomparison Project (PAMIP) contribution to CMIP6: investigating the causes and consequences of polar amplification. *Geosci. Model Dev.* **12**, 1139–1164 (2019).

50. Peings, Y., Labe, Z. M. & Magnusdottir, G. Are 100 ensemble members enough to capture the remote atmospheric response to+ 2 °C Arctic sea ice loss? *J. Clim.* **34**, 3751–3769 (2021).

51. Streffing, J., Semmler, T., Zampieri, L. & Jung, T. Response of Northern Hemisphere weather and climate to Arctic Sea Ice decline: resolution Independence in polar amplification model intercomparison project (PAMIP) simulations. *J. Clim.* **34**, 8445–8457 (2021).

52. Otto, F. E., Massey, N., van Oldenborgh, G. J., Jones, R. G. & Allen, M. R. Reconciling two approaches to attribution of the 2010 Russian heat wave. *Geophys. Res. Lett.* **39** (2012).

53. Pall, P. et al. Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature* **470**, 382–385 (2011).

54. Ye, K., Woollings, T. & Screen, J. A. European Winter Climate Response to Projected Arctic Sea-Ice Loss Strongly Shaped by Change in the North Atlantic Jet. *Geophys. Res. Lett.* **50**, e2022GL102005 (2023).

55. Barnes, E. A. & Screen, J. A. The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it? *Wiley Interdiscip. Rev. Clim. Change* **6**, 277–286 (2015).

56. Woollings, T., Hannachi, A. & Hoskins, B. Variability of the North Atlantic eddy-driven jet stream. *Q. J. R. Meteorol. Soc.* **136**, 856–868 (2010).

57. Jhun, J. G. & Lee, E. J. A new East Asian winter monsoon index and associated characteristics of the winter monsoon. *J. Clim.* **17**, 711–726 (2004).

58. Lo, Y. E., Mitchell, D. M., Watson, P. A. & Screen, J. A. Changes in winter temperature extremes from future Arctic Sea-Ice Loss and Ocean Warming. *Geophys. Res. Lett.* **50**, e2022GL102542 (2023).

59. Screen, J. A. & Simmonds, I. Amplified mid-latitude planetary waves favour particular regional weather extremes. *Nat. Clim. Change* **4**, 704–709 (2014).

60. Blackport, R., Fyfe, J. C. & Screen, J. A. Decreasing subseasonal temperature variability in the northern extratropics attributed to human influence. *Nat. Geosci.* **14**, 719–723 (2021).

61. Bailey, H. et al. Arctic sea-ice loss fuels extreme European snowfall. *Nat. Geosci.* **14**, 283–288 (2021).

62. Liu, J., Curry, J. A., Wang, H., Song, M. & Horton, R. M. Impact of declining Arctic sea ice on winter snowfall. *Proc. Natl Acad. Sci. USA* **109**, 4074–4079 (2012).

63. Blackport, R. & Kushner, P. J. The role of extratropical ocean warming in the coupled climate response to Arctic sea ice loss. *J. Clim.* **31**, 9193–9206 (2018).

64. Hall, R. J., Mitchell, D. M., Seviour, W. J. & Wright, C. J. Persistent model biases in the CMIP6 representation of stratospheric polar vortex variability. *J. Geophys. Res. Atmos.* **126**, e2021JD034759 (2021).

65. Rao, J., Garfinkel, C. I., Wu, T., Lu, Y. & Chu, M. Mean state of the Northern Hemisphere stratospheric polar vortex in three generations of CMIP models. *J. Clim.* **35**, 4603–4625 (2022).

66. Sun, L., Deser, C., Simpson, I. & Sigmond, M. Uncertainty in the winter tropospheric response to Arctic Sea ice loss: the role of stratospheric polar vortex internal variability. *J. Clim.* **35**, 3109–3130 (2022).

67. Ayarzagüena, B. & Screen, J. A. Future Arctic sea ice loss reduces severity of cold air outbreaks in midlatitudes. *Geophys. Res. Lett.* **43**, 2801–2809 (2016).

68. Holmes, C. R., Woollings, T., Hawkins, E. & De Vries, H. Robust future changes in temperature variability under greenhouse gas forcing and the relationship with thermal advection. *J. Clim.* **29**, 2221–2236 (2016).

69. Cullen, M. J. The unified forecast/climate model. *Meteorol. Mag.* **122**, 81–94 (1993).

70. Webb, M., Senior, C., Bony, S. & Morcrette, J. J. Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Clim. Dyn.* **17**, 905–9922 (2001).

71. Tibaldi, S. & Molteni, F. On the operational predictability of blocking. *Tellus A* **42**, 343–365 (1990).

72. Scherrer, S. C., Croci-Maspoli, M., Schwierz, C. & Appenzeller, C. Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region. *Int. J. Climatol.* **26**, 233–249 (2006).

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

K.Y. led the conception and experiment design, performed the model experiments and analysis of model output and wrote the first draft of the paper. T.W., P.A.G.W., S.N.S. and J.A.S. participated in conceiving the ideas, experiment design and analysis of model output. S.N.S. also assisted in performing the model experiments. All authors contributed to the revision of the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41612-023-00562-5.

Correspondence and requests for materials should be addressed to Kunhui Ye.

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.