Arctic temperature and precipitation extremes in present-day and future storyline-based variable resolution Community Earth System Model simulations

René R. Wijngaard[1,4], Willem Jan van de Berg[1], Christiaan T. van Dalum[1,5], Adam R. Herrington[2], and Xavier J. Levine[3,6]

[1]Institute for Marine and Atmosphere Research Utrecht, Utrecht University, Utrecht, the Netherlands
[2]Climate and Global Dynamics Laboratory, NSF National Center for Atmospheric Research, Boulder, CO, USA
[3]NORCE Norwegian Research Centre, Bjerknes Centre for Climate Research, Bergen, Norway
[4]Now at: Department of Physical Geography, Utrecht University, Utrecht, the Netherlands
[5]Now at: Royal Netherlands Meteorological Institute, De Bilt, the Netherlands
[6]Now at: Columbia University, New York, USA

Correspondence: René R. Wijngaard (r.r.wijngaard@uu.nl)

Abstract. Over the last few decades, the Arctic region has warmed up at a greater rate than elsewhere on the globe, partly resulting from the on-going loss of sea ice and seasonal snow over land. It is projected that the amplified warming of the surface will continue in the future. In addition, the intensity and frequency of temperature and precipitation means and extremes are projected to change, which may pose serious threats for human infrastructure and livelihoods. To assess (future) climate extremes, advanced modelling approaches with (regionally) refined resolution could be helpful.

In this study, we use the variable-resolution Community Earth System Model version 2.2 (VR-CESM) to evaluate and assess present-day and future ~~climate~~ temperature and precipitation extremes, such as heat waves and heavy precipitation, over the Arctic. Applying a globally uniform 1° grid and a VR grid with regional grid refinements to 28 km over the Arctic and Antarctica, we run 30-year present-day (1985–2014), 10-year present-day (2005–2014), and future (2090–2099) simulations with interactive atmosphere and land surface models, and prescribed sea ice and sea surface temperatures. We use the 30-year simulation to evaluate the ability of the VR grid to simulate climate extremes by comparison with gridded outputs of the globally uniform 1° grid, reanalysis-based datasets, and a regional climate model. The 10-year simulations follow two storylines of Arctic climate change representing a combination of strong/weak Arctic tropospheric warming and ~~strong/weakSST~~ weak/strong sea surface warming in the Barents-Kara Seas and are used to assess future climate extremes by focussing on temperature and precipitation extremes. The outcomes show that the VR grid generally performs better in simulating precipitation extremes, while the globally uniform 1° grid generally performs better in simulating temperature extremes, which is mainly related to larger negative temperature differences in the VR grid. Future projections suggest that high temperature extremes will generally increase both in intensity and duration, whereas low temperature extremes will decrease in intensity and duration, especially over regions dominated by ~~SST~~ sea surface warming and large sea ice loss. Further, wet precipitation extremes are projected to increase in intensity and frequency. The outcomes of this study may contribute to an improved understanding on future climate extremes and its implications.

Copyright statement. TEXT

1 Introduction

In recent decades, the Arctic region has warmed up at a greater rate than elsewhere on the globe, a phenomenon known as "Arctic amplification" (Rantanen et al., 2022). Although the exact causes are still under debate, numerous studies have shown that Arctic amplification is caused by several feedback mechanisms involving interactions between the atmosphere, ocean, sea ice, and land surface, such as enhanced ocean-atmosphere coupling and changing surface albedos resulting from the ongoing loss of sea ice and snow over land (Screen and Simmonds, 2010; Dai et al., 2019; Park et al., 2019; Previdi et al., 2021; Rantanen et al., 2022). Under increasingly warmer climate conditions, high temperature extremes, such as warm days/nights and heat waves, have become more frequent in recent decades, while low temperature extremes, such as cold days/nights and cold spells, have become less frequent (Matthes et al., 2015; Walsh et al., 2020). ~~Only in winter have cold extremes become more frequent over parts of North America and Asia, most likely due to the warming-induced weakening of the polar vortex, leading to more frequent cold outbreaks (Kim et al., 2014; Kug et al., 2015; Matthes et al., 2015; Johnson et al., 2018).~~ In contrast to temperature extremes, systematic trends for precipitation extremes have not been observed and are more regional in nature due to the greater spatial variability of precipitation, the lack of measurements at high latitudes and altitudes, the accuracy of measurements, and uncertainties associated with atmospheric reanalyses (Walsh et al., 2020). According to Walsh et al. (2020), increasing trends in precipitation extremes have been observed over much of the Arctic land area, although regionally decreasing trends or no trends have also been observed.

In the future, Arctic warming is projected to continue, altering the intensity and frequency of temperature extremes (Sillmann et al., 2013b; Screen et al., 2015; Landrum and Holland, 2020; Walsh et al., 2020). As warming continues, precipitation is expected to increase, generally in line with the Clausius-Clapeyron relationship, which describes an increase in atmospheric moisture content under warmer climate conditions (Pfahl et al., 2017). In addition, precipitation is expected to increasingly shift from snow to rain, and the intensity and frequency of wet ~~(dry)~~ precipitation extremes are projected to increase~~(decrease)~~, while those of dry precipitation extremes are projected to decrease (Sillmann et al., 2013b; Screen et al., 2015; Landrum and Holland, 2020; Walsh et al., 2020; Paik et al., 2023). The changing intensity and frequency of temperature and precipitation extremes could ultimately lead to more frequent floods, wildfires, and reduced agricultural production, with profound impacts on ecosystems, human infrastructure and livelihoods (Hirabayashi et al., 2013; Masrur et al., 2018; Walsh et al., 2020; Overland, 2022). Given the potential future impacts of temperature and precipitation extremes in the Arctic, there is a need to develop adaptation and mitigation strategies that can help reduce potential adverse impacts on vulnerable Arctic communities and ecosystems.

Developing future adaptation and mitigation strategies is a challenging task due to the wide range of climate change projections resulting from uncertainties in possible future greenhouse gas emission scenarios, incomplete understanding of physical processes and their representation in climate models, and natural variability within the climate system (Hawkins and Sutton, 2009; Overland et al., 2019; McCrystall et al., 2021; Levine et al., 2024). To strengthen decision-making processes related to the

development of adaptation and mitigation strategies, a number of possible climate outcomes – storylines – can be investigated. Storylines can be described as a physically self-consistent unfolding of past events or a plausible future pathway representative for regional climate change (Shepherd et al., 2018). Storylines have been generated using two distinct methodologies according to the goal they want to achieve. Event-based storylines use an extreme synoptic event and apply changes to the mean state of the atmosphere to quantify the resulting change in impacts (e.g., Sillmann et al., 2021; Chan et al., 2022). Dynamical storylines use a multi-variate linear regression to generate climate states based upon the dependence of those climate state to predetermined climate indices; those climate indices generally represent a well-known change in the atmospheric, oceanic or sea ice state (e.g., Zappa and Shepherd, 2017; Zappa, 2019; Mindlin et al., 2020; Williams et al., 2024). Recently, dynamical storylines for the Arctic have been developed by Levine et al. (2024) that describe future pathways for Arctic summer climate change. These storylines are based on two different drivers that explain substantial fractions of the surface climate response to global warming in the Arctic, namely: 1) warming of the Arctic lower troposphere and 2) warming of the sea surface in the Barents-Kara Sea. As both drivers are expected to lead to changes in the occurrence and intensity of temperature and precipitation extremes, the Arctic storylines could contribute to a better understanding of the possible range of impacts of regional climate change on the intensity and frequency of temperature and precipitation extremes over the Arctic.

To investigate present-day and future climate extremes in the Arctic, a variety of advanced modelling approaches have been used. Global climate models (GCMs) have often been used to assess the present-day state and future changes in temperature and precipitation extremes on a global scale (e.g., Sillmann et al., 2013b, a; Seneviratne and Hauser, 2020; Kim et al., 2020; Seneviratne et al., 2021). Although GCMs simulate temperature extremes and large-scale precipitation extremes reasonably well, high-resolution models ~~($\Delta x = 0.25°$ or higher)~~ with horizontal grid spacings of $0.25°$ or higher have shown the ability to simulate precipitation extremes better than coarser-gridded models (Wehner et al., 2014; O'Brien et al., 2016; Seneviratne et al., 2021). However, the use of high-resolution GCMs has been limited due to the large computational resources that are required to run these GCMs. In this context, regional climate models (RCMs) could be considered as a more suitable alternative. RCMs, such as those used in the Arctic-CORDEX experiment (https://climate-cryosphere.org/arctic-cordex/), can be run at a higher spatio-temporal resolution with horizontal grid spacings as fine as $\sim$11 km, and are therefore able to better capture spatio-temporal variability in temperature and precipitation. In addition, RCMs are often more specialized than GCMs in the simulation of (polar) processes (e.g. the treatment of snow and ice) and more optimized for specific regions of interest. Nonetheless, RCMs need to be forced with GCMs or reanalysis products, which disables two-way interactions between the region of interest and the global domain and introduces inconsistencies in terms of model physics and dynamics between RCMs and GCMs.

To overcome the limitations associated with RCMs and the computational constraints associated with high-resolution GCMs, variable-resolution GCMs or Earth System Models (ESMs), such as the variable-resolution Community Earth System Model (VR-CESM), have been developed. VR-CESM (Zarzycki et al., 2014; Lauritzen et al., 2018) is a hybrid between regional and global climate models as it applies a regional grid refinement over a region of interest within a coarse-gridded global domain (Rhoades et al., 2016). VR-CESM has been used to study various processes in several regions. For example, with regional refinements up to 7 km, VR-CESM has been used to study regional climate, atmospheric rivers,

wind extremes, glacier surface mass balance, and/or snowpack characteristics in the western USA (Rhoades et al., 2018), Chilean Andes (Bambach et al., 2021), Mediterranean (Boza et al., 2025), East Asia (Zhu et al., 2023), High Mountain Asia (Wijngaard et al., 2023), Greenland (van Kampenhout et al., 2019; Herrington et al., 2022; Loeb et al., 2024; Waling et al., 2024), and Antarctica (Datta et al., 2023). Furthermore, present-day and future climatic means and/or extremes have been assessed with VR-CESM over western US, Canada, eastern China, and Greenland ~~(Xu et al., 2022; Yin et al., 2024)~~ (Xu et al., 2022; Yin et al., 2024; Morris et al., 2023, 2024; Morris and Kushner, 2025). However, assessments focussing on temperature and precipitation extremes over the entire Arctic region have not been conducted thus far.

The main objective of this study is to evaluate and assess present-day and future temperature and precipitation extremes over the Arctic, using VR-CESM. To this end, we apply a globally uniform 1° (∼111 km) grid and a dual polar VR grid with horizontally refined grid spacings of 28 km (0.25°) over the Arctic and Antarctic with interactively coupled atmosphere and land surface models, and prescribed sea ice and sea surface temperatures. We run three different types of simulations with both model grids, namely: 1) present-day simulations covering a 30-year period (1985–2014), 2) present-day simulations covering a 10-year period (2005–2014), and 3) future simulations (2090–2099) following the high-emission SSP5-8.5 scenario. The 30-year present-day simulations are used to evaluate the performance of the VR-CESM grids through comparisons with gridded outputs from reanalysis datasets and a regional climate model. The 10-year present-day and future simulations are used to assess future changes in temperature and precipitation extremes by following two storylines of Arctic summer climate change representing a combination of strong/weak Arctic tropospheric warming and ~~strong/weakSST~~ weak/strong sea surface warming in the Barents-Kara Seas. Compared to previous VR-CESM studies focusing on climate extremes and/or the Arctic, this study has several novelties. First, this study is the first VR-CESM application that evaluates and assesses present-day and future intensity and occurrence of temperature and precipitation extremes over the entire Arctic. Second, we use a storyline approach to assess future extremes over the Arctic, which can potentially improve decision-making processes related to the development of adaptation and mitigation strategies.

This paper is organized as follows. Section 2 briefly describes the model and highlights the methods and data. Section 3 presents and discusses the main outcomes of this study. Finally, Section 4 provides further discussion and the conclusions.

2 Data and Methods

2.1 Modelling setup

We used the Community Earth System Model version 2.2 (CESM2; Danabasoglu et al., 2020; Herrington et al., 2022), a state-of-the-art global Earth system model consisting of multiple model components (including atmosphere, oceans, land surface, rivers, sea ice and land ice) that can be run in a partially or fully coupled mode. In this study, we applied CESM in a partially coupled mode by interactively coupling prognostic atmosphere and land surface components, and using prescribed daily sea ice and sea surface temperatures to replace the active ocean and sea ice components. This model configuration follows the Atmospheric Model Intercomparison Project (AMIP) protocol (Gates et al., 1999).

The atmosphere component of CESM2, the Community Atmosphere Model version 6 (CAM6), is applied with a hydrostatic spectral element dynamical core that supports unstructured grids that eliminates polar singularities and enables VR capabilities (CAM6-SE; Zarzycki et al., 2014; Lauritzen et al., 2018; Gettelman et al., 2019). Physics parameterization schemes applied in CAM6 include the Cloud Layers Unified by Binormals (CLUBB) scheme simulating shallow convection, boundary layer turbulence, and cloud macrophysics (Bogenschutz et al., 2013); a deep convection scheme (Zhang and McFarlane, 1995); a two-moment cloud microphysics scheme with prognostic treatment of precipitation (MG2; Gettelman and Morrison, 2015); the modal aerosol module (MAM4; Liu et al., 2016); a ~~cloud-aerosol~~ radiation scheme (Rapid Radiative Transfer Method for GCMs - RRTMG; Iacono et al., 2008); and anisotropic orographic gravity wave (Weimer et al., 2023) and form drag parameterization schemes (Beljaars et al., 2004). Our atmospheric model configuration uses a dry mass low-top vertical coordinate with a model top at ~40 km and 58 hybrid sigma-pressure levels in the vertical instead of the standard 32 levels of CESM2 (Lauritzen et al., 2018). The new enhanced vertical grid is planned for the CESM3 model and has a higher resolution in the planetary boundary layer and in the middle and upper troposphere, and is intended to improve the representation of moisture, temperature, and cloud profiles in the boundary layer and to reduce noise and spurious flow features (Skamarock et al., 2019; Huang et al., 2022).

CAM6 is coupled to the land surface component of CESM2, the Community Land Model version 5 (CLM5; Lawrence et al., 2019), which is applied with satellite vegetation phenology (CLM5-SP). CLM5 simulates the surface energy balance, hydrological processes, biogeochemical cycles, and their interactions with the atmosphere (Oleson et al., 2013). It includes several new and updated processes and parameterizations, such as for snow and surface discretization (Lawrence et al., 2019). ~~Snow albedo and two-way radiative transfer in the snowpack are simulated with a snowpack radiative heating model (Snow, Ice, and Aerosol Radiative model – SNICAR; Flanner and Zender, 2005; Flanner et al., 2007). A multi-layer snow model with up to 12 layers of snow and a maximum allowed depth of 10 m water equivalent (w.e.) simulates important snow processes, such as melting and refreezing, and includes modifications for wind- and temperature-dependent fresh snow density and snow compaction (van Kampenhout et al., 2017). To represent the heterogeneity of the land surface, CLM5 uses a subdivision scheme that heterogeneously subdivides CLM5 grid cells into five different land units (natural vegetation, urban, lakes, crops, and glaciers), which can be further subdivided into soil/snow columns and patches representing plant functional types (PFTs) or crop functional types (CFTs) (Lawrence et al., 2019).~~

2.1.1 Grids~~and performance~~

The VR-CESM simulations performed in this study are run with two different spectral element grids: 1) a globally uniform 1° (~111 km) grid, hereafter referred to as NE30, and 2) a dual polar VR grid (Fig. 1), hereafter referred to as POLARRES. The POLARRES grid was generated by the SQuadGen software package (Ullrich, 2014). Partly based on the existing ARCTIC and ANTSI VR grids (Herrington et al., 2022; Datta et al., 2023), POLARRES has horizontal grid spacings of 0.25° (~28 km) over the Arctic and Antarctic (poleward of 60°N/S). Further, a transition grid with horizontal grid spacings of 0.5° (~55 km; poleward of 45°N/S) serves as a buffer between the Arctic and Antarctic domains and the global 1° (~111 km) domain. ~~The~~

~~total number of horizontal grid points on the POLARRES grid sums up to 186 194, which is about 3.8 times more than on the NE30 grid (48 600).~~

~~The topography of the VR-CESM grids was interpolated from an updated 30-arcsec global topography dataset comprising of the Global Multi-resolution Terrain Elevation Data (GMTED2010; Danielson and Gesch, 2011) of the United States Geological Survey (USGS) and the BedMachine topography (Morlighem et al., 2017, 2020), which includes terrain elevations of Greenland and Antarctica (Wijngaard et al., 2023). To interpolate the topography, we used the NSF-NCAR topography generation software package (Lauritzen et al., 2015), which includes ridge finding (for the anisotropic orographic gravity wave scheme) and internal smoothing algorithms to improve the accuracy of high-resolution topography.~~

To ensure numerical stability, CAM physics (dynamics) time steps of 1800 s (300 s) and 450 s (75 s) were used for NE30 and POLARRES, respectively. ~~Furthermore, hyperviscosity coefficients preventing numerical artefacts and instability are scaled according to the dimensions of a grid element, configured such that for each halving of the grid spacing the hyperviscosity coefficients decrease by about an order of magnitude (Guba et al., 2014; Zarzycki et al., 2014).~~ The VR-CESM simulations were performed on the Cheyenne supercomputing facility at the NSF National Center for Atmospheric Research (NSF-NCAR; Computational and Laboratory, 2019). The computational cost of the NE30 and POLARRES grids amount to about 5 300 and 50 000 core hours per simulated year~~(CHPSY), which is about 2 times more expensive than a globally uniform 1° SE grid applied with 32 vertical levels (2 500 CHPSY). The POLARRES grid has a computational cost of about 50 000 CHPSY, which is about 10 times more expensive than the NE30 grid. Using 1080-2520 processors (30-70 nodes) , the throughput reached about 7 simulated years per wall-clock day for NE30 and about 0.6-1.2 simulated years per wall-clock day for POLARRES.~~ , respectively.

The topography of the VR-CESM grids was interpolated from an updated 30-arcsec global topography dataset comprising of the Global Multi-resolution Terrain Elevation Data (GMTED2010; Danielson and Gesch, 2011) of the United States Geological Survey (USGS) and the BedMachine topography (Morlighem et al., 2017, 2020), which includes terrain elevations of Greenland and Antarctica (Wijngaard et al., 2023). To interpolate the topography, we used the NSF-NCAR topography generation software package (Lauritzen et al., 2015).

2.1.2 Parameter tunings

The regionally refined resolution associated with the POLARRES simulations most likely has an impact on the representation of clouds as found in previous studies performed with VR-CESM (Wijngaard et al., 2023; Boza et al., 2025). For this reason, we implemented additional parameter tunings (largely following Wijngaard et al. (2023)) for the POLARRES simulations to tune the low-level cloudiness, and the shortwave radiation and albedo over sea ice. First, we increased the strength of the damping for the 3rd moment of the vertical velocity in the large skewness regime ($clubb\_c11b$) from 0.35 to 0.375. Increasing the damping strength reduces the vertical velocity skewness of the cloud distribution and increases the fraction of low cloud cover (Guo et al., 2015; Boza et al., 2025). Second, to tune shortwave radiation and albedo over sea ice, we ~~increased the snow grain radius (r_snw~~decreased the snowmelt onset temperature (dt_melt) over sea ice from ~~1.25 to~~ 1.5 ~~standard deviations~~ to

1.0 °C, and decreased the ~~snowmelt onset temperature (~~dt_melt~~~~snow grain radius by increasing the parameter (r_snw) over sea ice from ~~1.5 to 1.0 °C~~1.25 to 1.5 standard deviations, resulting in a higher snow albedo.

In addition, we implemented modifications to fix a model bug in the cloud microphysics (Shaw et al., 2022; Zhu et al., 2022). These modifications are largely based on the paleoclimate-calibrated CESM2 configuration of Zhu et al. (2022), which include a reduction of the microphysical ~~timestep~~ time-step size by increasing the number of microphysical substeps, and the removal of a limiter on the cloud ice number concentrations. For the NE30 simulations, we have adopted the number of microphysical substeps ($micro\_mg\_num\_steps$) from Zhu et al. (2022), which increases from 1 (default value) to 8~~and corresponds to a microphysical timestep of 75s~~, resulting in a microphysical time-step size of 75 s. For the POLARRES simulations, we maintained the microphysical ~~timestep of 75s, corresponding to 3 microphysical substeps .~~ time-step size of 75 s, which is derived by increasing the number of microphysical substeps from 1 to 3.

2.2 Experimental Design

To evaluate and assess present-day and future climate extremes, three different model experiments are run with and without regional grid refinement over the Arctic. First, we run present-day experiments for a 30-year period (1985–2014) while forcing VR-CESM with daily sea ice cover (SIC) and sea surface temperatures (SST) retrieved from the ERA5 reanalysis dataset (Hersbach et al., 2020). These experiments (hereafter referred to as present-day A experiments) were used to evaluate the performance of the model grids in simulating present-day climate extremes and to understand the impact of regional grid refinement on the representation of climate extremes. Second, we run future storyline experiments for the period 2090–2099 using SST and SIC fields from CMIP6 models representing selected Arctic storylines that are physically plausible outcomes of Arctic summer climate change. The selected storylines follow the high-emission SSP5-8.5 scenario and were derived by Levine et al. (2024) based on two different climate drivers: 1) Arctic atmospheric warming at the 850 hPa level and 2) ~~SST~~ sea surface warming in the Barents-Kara Seas. In this study, we used two different Arctic storylines that show opposite paths of Arctic climate change, namely:

1. Storyline ST1 (corresponding to storyline ~~B2~~ D in Levine et al. (2024)): Strong Arctic tropospheric warming combined with weak ~~SST~~ sea surface warming in the Barents-Kara Seas (~~PolAmpl~~ArcAmp~~+BKSSTWarm~~-BKWarm-). Relative to the multi-model mean (MMM) climate change, this storyline is characterized by warmer and drier continents, weaker ~~SST~~ sea surface warming over the Arctic and North Atlantic Ocean (Fig. 2c), and reduced sea ice loss (Fig. 2g).

2. Storyline ST2 (corresponding to storyline ~~B1~~ A in Levine et al. (2024)): Weak Arctic tropospheric warming combined with strong ~~SST~~ sea surface warming in the Barents-Kara Seas (~~PolAmpl-BKSSTWarm~~ArcAmp-BKWarm+). Relative to the multi-model mean (MMM) climate change, this storyline is characterized by cooler and wetter continents, stronger ~~SST~~ sea surface warming (Fig. 2d), and stronger sea ice loss, especially in the Barents-Kara Seas (Fig. 2h).

Each storyline is represented by a CMIP6 model selected for its similarity to the storylines and its skill to simulate the historical climate. Here, the r1i1p1f1 member of the Norwegian Earth System Model version 2 (NorESM2-MM; Seland et al., 2020) and the r1i1p1f2 member of the CNRM-CERFACS Earth System Model version 2 (CNRM-ESM2-1; Séférian et al., 2019)

model realizations represent storylines ST1 and ST2, respectively. Finally, we run another set of present-day experiments for a 10-year period (2005–2014) using SST and SIC fields from the same CMIP6 model realizations used for the future storyline experiments. These baseline experiments (hereafter referred to as present-day B experiments) are used (in combination with the future storyline experiments) to assess future changes in climate extremes over the Arctic. The CMIP6 model realizations representing ST1 and ST2 project a global mean near-surface warming of 3.3 K for ST1 and 4.6 K for ST2 between the future and present-day periods. All model experiments were spun-up for a period of 1 year.

Greenhouse gas concentrations, tropospheric aerosols, and ozone concentrations are time-varying and prescribed in accordance with CMIP6 historical and SSP5-8.5 forcings for the present-day and future storyline experiments, respectively. Stratospheric aerosols were not prescribed for the model experiments in this study. The land surface distributions for the VR-CESM grids are partly transient and partly constant. The plant ~~functional type (PFT)~~ and crop functional type ~~(CFT)~~ distributions are transient and are interpolated from the Land Use Harmonization (LUH2) time series, which has been developed for CMIP6 (Hurtt et al., 2020). The ~~LUH2 time series covers the period 1850-2100 and describes annual land changes based on the History Database of the Global Environment (HYDE version 3.2; Klein-Goldewijk et al., 2017) for present-day (1850–2015) and SSP-RCP-based Integrated Assessment Model (IAM) projections for the future (2015–2100). The~~ distributions of other land surface classes, such as glaciers, lakes, and urban areas, are assumed to be constant in time and set to the distributions of the year 2000.

2.3 Model evaluation

To evaluate the performance of the NE30 and POLARRES grids in simulating the climatology and extremes of temperature and precipitation, we compared gridded outputs from the present-day A experiments (1985–2014) with gridded outputs from two different reanalysis products and an RCM. To this end, we used daily precipitation sums and daily maximum, minimum, and mean 2-meter temperature from ERA5 (reanalysis), JRA-3Q (reanalysis), and RACMO (RCM):

- The ERA5 reanalysis data (Hersbach et al., 2020) were retrieved from the KNMI Climate Explorer (https://climexp.knmi.nl/start.cgi; last access: 2 July 2024). Originally, the ERA5 reanalysis data are available on a ∼31 km grid and monthly or hourly intervals. However, the data retrieved from the KNMI Climate Explorer have been regridded to a spatial resolution of 0.5°x0.5° and are available at daily intervals.

- The Japanese Reanalysis for Three Quarters of a Century (JRA-3Q) reanalysis (Kosaka et al., 2024) is the successor of the JRA-55 reanalysis, which has recently been released by the Japanese Meteorological Agency. JRA-3Q reanalysis data have been retrieved from the Research Data Archive (RDA) of NSF-NCAR (https://rda.ucar.edu/datasets/ds640.0; last access: 12 August 2024) and are available on a TL479 grid (∼40 km).

- Regional Atmosphere Climate Model (RACMO) data are based on RACMO version 2.4p1, a recently updated RCM that is specially developed to simulate polar climate processes over Greenland, the Arctic, and Antarctica (van Dalum et al., 2024). RACMO is forced with SST, SIC, and multi-level ERA5 data of wind speed, temperature, humidity, and pressure, and is available on an 11 km grid.

All CESM Although reanalysis data are commonly used to evaluate model outcomes, there may be some limitations to its use. This is because reanalysis fields can be strongly influenced by the underlying forecast model in regions with little or no observations. To allow comparison between the CESM grids and reference data, all CESM output and reference data are were regridded to a 1-degree finite volume grid (0.9°x1.25°), unless noted otherwise, to allow comparison between the CESM grids and reference data. Daily temperature fields were regridded using bilinear interpolation, and daily precipitation fields were regridded using conservative interpolation. The regridding was applied prior to the analysis of temperature and precipitation climatology and extremes. Finally, we have not used the high-resolution CARRA (Copernicus Arctic Regional Reanalysis) reanalysis due to its limited domain coverage. The release of a new version of CARRA, covering the entire Arctic domain, is planned for 2025–2026 (CARRA, 2025).

2.4 Analysis of temperature and precipitation extremes

To evaluate and assess present-day and future temperature and precipitation extremes in the Arctic, we selected eight extreme metrics (Table 1), including four temperature extreme metrics and four precipitation extreme metrics. These metrics are analyzed were analyzed on a year-by-year basis using the Climate Data Operators (CDO) functions based on the European Climate Assessment (ECA) climate indices, which are consistent with the definitions of the Expert Team on Climate Change Detection and Indices (ETCCDI).

To analyze temperature extremes, we selected extreme metrics representing the intensity and duration of temperature extremes, where the . The annual maximum of daily maximum temperature (TXx) and the annual minimum of daily minimum temperature (TNn) are were used to analyze changes in the intensity of temperature extremes. To analyze changes in the duration of temperature extremes, we used the Warm Spell Duration Index (WSDI) and Cold Spell Duration Index (CSDI) as proxies for heat waves and cold spells, respectively. Here, the warm (cold) The warm spell duration index is defined as the annual number of days in intervals of at least 6 consecutive days on which TX (TN) is higher (lower) than the daily maximum temperature (TX) exceeds the 90^{th} (10^{th}) percentile of TX. The cold spell duration index refers to intervals during which the daily minimum temperature (TN) falls below the 10^{th} percentile of TN. The 90^{th} (and 10^{th}) percentile of TX (TN) is percentiles of daily maximum and minimum temperature, respectively, are calculated for each calendar day using a 15-day running window for the base periods (1985–2014 for present-day A and 2005–2014 for present-day B). To derive the future WSDI and CSDI warm and cold spell duration metrics, we used the historical 90^{th} (and 10^{th}) percentile percentiles, calculated for the present-day B base period, as a threshold.

We assessed the intensity of precipitation extremes by analyzing the 99^{th} percentile of daily precipitation sums (P99) and the highest 5-day precipitation sums (RX5day). We also analyzed the number of heavy precipitation days (R10mm; defined as daily precipitation equal to or greater than 10 $\mathrm{mm\,d^{-1}}$) and the greatest number of consecutive dry days with daily precipitation less than 1 $\mathrm{mm\,d^{-1}}$ (CDD) to assess the frequency of heavy precipitation days throughout the year and the longest duration of dry spells, respectively.

The climatology and extremes of temperature and precipitation are analyzed by using probability density functions (PDFs) and regional averages that are calculated for four different Arctic land regions, one ocean region and one sea ice region covering

areas poleward of 60°N (Fig. 1). Arctic land regions are based on regional domains defined by Seneviratne et al. (2012), namely Alaska~~and Western~~ /western Canada (AWC; 105-168°W; 60-75°N), ~~Eastern Canadaand~~ eastern Canada/Greenland (ECG; 10-105°W; 60-85°N), Scandinavia (SCA; 0-40°E; 60-85°N) and Siberia (SIB; 40-180°E; 60-85°N). The sea ice region (SIC) is based on the maximum extent of sea ice cover during the present-day A (1985–2014) or present-day B (2005–2014) periods, taking into account all grid cells with sea ice cover greater than 15%. The ocean region (OCN) includes all ocean grid cells poleward of 60°N that are not classified as sea ice. The PDFs are constructed by sampling temperature or precipitation for all grid points in a specific region, while regional averages are based on metrics that have been calculated for each grid point within the respective regions. To avoid the effects of polar singularities on PDFs, CESM output and reference data are regridded to the NE30 grid (to allow a uniform comparison between the CESM grids and reference data), prior to the construction of the PDFs.

To evaluate statistical significance, we performed two-tailed T-tests, where outcomes are marked as statistically significant when p-values are lower than or equal to 0.05. As the statistical testing is applied to a large number of grid points, there is an increased likelihood of false positives emerging from multiple statistical testing (Wilks, 2016). To control for these false positives, we used the False Discovery Rate (FDR) approach (Benjamini and Hochberg, 1995; Wilks, 2016) to correct the p-values. Here, we set the FDR control level (α_{FDR}) to 0.1, which corresponds to a global significance level of 0.05 (Wilks, 2016).

3 Results

3.1 Evaluation

3.1.1 Temperature Extremes

Before assessing future changes in temperature and precipitation extremes in the Arctic, we evaluate the performance of NE30 and POLARRES in representing temperature and precipitation climatology and extremes. Figure 3 shows the present-day A (1985–2014) annual mean near-surface temperature differences between the NE30 and POLARRES grids, and the gridded outputs of ERA5, JRA-3Q, and RACMO. ~~Both~~ NE30 ~~and POLARRES show warm temperature biases over most of the land surface and cold biases over the~~ shows higher temperatures primarily over Europe and parts of Greenland, while these differences are less significant in POLARRES. Furthermore, both NE30 and POLARRES show lower temperatures over the central Arctic, ~~Greenland, and~~ northeastern Canada, and parts of Greenland, where POLARRES is ~~colder on average~~ on average colder than NE30. ~~Here, the cold biases relative to ERA5 are largest , with a mean bias of -1.06 °C and -1.51 °C~~ The cold temperature differences are largest compared to ERA5, with average differences of -1.32 °C for NE30 and -1.72 °C for POLARRES (Figs. 3a-b)~~, respectively (i. e. calculated over a domain poleward of 60°N), followed by RACMO(-0.98 °C /-1.44 °C~~. The differences are smaller compared to RACMO, with average differences of -1.24 °C for NE30 ~~/POLARRES ;~~ and -1.64 °C for POLARRES (Figs. 3e-f)~~and~~ , and compared to JRA-3Q~~(-0.27 °C /-0.73 °Cfor NE30/POLARRES; ~~, with average differences of -0.54 °C and -0.95 °C, respectively (Figs. 3c-d).

~~The cold biases~~ Lower temperatures over the central Arctic, ~~Greenland, and northeastern Canada~~northeastern Canada, and Greenland are also visible in the present-day A ~~annual~~ daily temperature probability density functions (PDFs), as shown in Figure 4. ~~The PDFs show~~ These PDFs demonstrate that CESM is colder than ERA5 and RACMO over ~~ECG and SIC and~~ eastern Canada/Greenland (ECG) and the sea ice region (SIC), as well as being colder than JRA-3Q over ~~SIC~~ the sea ice region (Figs. 4b,f). In particular, in low temperature regimes (i.e., the lower tail) it is 2-5 °C colder in NE30/POLARRES than in ERA5, JRA-3Q, or RACMO, with temperatures below -30 °C occurring more frequently in CESM and temperatures between -30 °C and -10 °C occurring more frequently in ERA5, JRA-3Q, or RACMO. In other regions, the PDFs of CESM and the reference data are mostly similar, except for temperature regimes around 0°C, which tend to occur more frequently in CESM over the land regions (Figs. 4a-d).

To evaluate the performance of CESM in simulating temperature extremes~~and to better understand~~ , and to gain a better understanding of the patterns of the ~~cold bias~~ lower temperatures over the Arctic, we compare CESM output with ERA5. Figure 5 shows the differences between CESM and ERA5 for the present-day A temperature extreme metrics~~(TXx, WSDI, TNn, and CSDI)~~. For the annual maximum temperature~~(TXx)~~, NE30 and POLARRES show ~~cold biases~~ lower temperatures over northeastern Canada and ~~to a lesser extent over~~ the central Arctic (Figs. 5b-c), and ~~warm biases over~~ higher temperatures over parts of the continental mid-latitudes. In general, POLARRES shows small improvements in ~~TXx (i.e., the bias decreases by~~ the annual maximum temperature with a small reduction in the temperature difference of 0.2 °C~~on average), especially~~ . Although POLARRES is generally warmer over Siberia, Alaska, and Scandinavia and colder over northeastern Canada, the differences between POLARRES and NE30 are mostly insignificant (Fig. 5d). ~~However, the cold bias is enhanced over northeastern Canada.~~ For the warm spell duration index~~(WSDI), the biases~~ , the differences between the CESM grids and ERA5 are mostly negative but insignificant (Figs. 5f-g), with the ~~biases~~ differences becoming slightly more negative in POLARRES (Fig. 5h). ~~Only a few positive biases (especially in NE30) are found over Scandinavia, Greenland, and the western Canada.~~ The annual minimum temperature ~~(TNn) is associated with larger biases than TXx, with warm biases~~ shows larger differences than the annual maximum temperature, with higher temperatures over land and, to a lesser extent, over the oceans, ~~and large cold biases~~ as well as large cold temperature differences over the central Arctic (Figs 5j-k). Here, the ~~TNn bias~~ annual minimum temperature difference is slightly more negative~~in POLARRES with decreasing TNn~~ , though still insignificant, in POLARRES, with a decrease in the annual minimum temperature over most of the Arctic, especially over Greenland, Iceland, Scandinavia, Svalbard and the Barents-Kara Sea region (Fig. 5l). The decrease in ~~TNn~~ the annual minimum temperature may be partly due to better resolved topography. Cold spell duration index ~~(CSDI) biases~~ differences are generally negative but insignificant over much of the central Arctic and Greenland in NE30 and over parts of Scandinavia in POLARRES (Figs. 5n-p), indicating a ~~decrease in the~~ smaller number of cold spell days ~~relative to~~ than in ERA5. ~~Considering that both~~ Although the annual mean temperature and the daily minimum temperature ~~(TN) show a cold bias~~ both show negative temperature differences relative to ERA5 ~~(especially~~ particularly over the central Arctic and ~~Greenland;~~ parts of Greenland (Figs. 3 and S1f-h) ~~, the negative CSDI biases are rather counter-intuitive. A~~ the negative differences are not unexpected. One possible explanation for ~~the opposite patterns~~ these differences is that the variability of the daily minimum temperature differs between ERA5 and CESM. For example, the variability of the daily minimum temperature over sea ice is lower in ERA5 than in CESM (not shown),

which may be partly due to the fixed sea ice thickness in ERA5 (i.e., 1.5 m; Batrak and Müller, 2019). In CESM, however, ~~the~~ sea ice thickness ~~is variable, which leads to a higher variability of~~ varies, leading to higher variability in the daily minimum temperature and thus a lower probability of meeting the cold spell criteria (Table 1)~~, ultimately resulting~~ . This ultimately results in a lower number of cold spell days.

Comparisons between CESM output and other reference data (i.e., JRA-3Q and RACMO) (Figs. S2 and S3) show ~~mostly similar patternsincluding cold TXx biases~~ broadly similar patterns, including lower annual maximum temperatures over north-eastern Canada, ~~cold TNn biases~~ lower annual minimum temperatures over the sea-ice-dominated regions, and ~~warm TNn (TXx) biases over most of the land (continental mid-latitudes)~~ (higher annual minimum temperatures over most land areas (Figs. S2b-c,j-k and S3b-c,j-k). ~~Spatially, the cold and warm TNn (TXx) biases~~ The annual maximum and minimum temperature differences generally correspond well to the ~~patterns of the climatological biases~~ climatological differences (Fig. 3). However, the magnitude of the extreme temperature ~~biases~~ differences is much larger. One ~~of the possible explanations for the cold biases are~~ possible explanation for the lower temperatures is the large warm near-surface temperature biases of 5-10°C over sea ice ~~that can be found~~ in reanalysis products, such as ERA5 and JRA-55 (the predecessor of JRA-3Q), during ~~winter~~ clear sky conditions in winter (Batrak and Müller, 2019; Zampieri et al., 2023). These ~~warm~~ biases are mainly ~~due to~~ caused by the poor representation of the snow layer on top of the sea ice and the thickness of the sea ice itself (Batrak and Müller, 2019; Zampieri et al., 2023). The warm biases in reanalysis products can also explain, to some extent~~explain the cold biases~~ , the cold temperature differences between CESM and RACMO~~as~~ , since RACMO is forced with sea ice, SST, and multi-level data (temperature, wind, humidity, and pressure) from ERA5, and thus inherits ERA5 uncertainties. ~~The uncertainties related to~~ Uncertainties in reanalysis products, such as ERA5, ~~indicate that the cold biases in CESM could~~ could suggest that the lower temperatures in CESM may be exaggerated. To better understand the magnitude of ~~the cold bias~~these differences, validation of near-surface temperature against observations from meteorological stations in the Arctic is recommended in future research.

Another possible explanation for the ~~cold biases~~ lower temperatures in CESM is a bias in the cloud forcing. For ~~example, Herrington et al. (2022) found~~ instance, Herrington et al. (2022) identified negative biases in the summer shortwave cloud forcing ~~over~~ across much of the Arctic, suggesting excessive reflection and cooling. Similar negative biases were also ~~found~~ identified in NE30 and POLARRES over the Arctic land regions (Figs. S4a-c), ~~especially~~ particularly over northeastern Canada. However, ~~the biases in the summer shortwave cloud forcing~~ these biases cannot fully explain the enhanced cold ~~bias~~ temperature differences over northeastern Canada in POLARRES. Although the summer shortwave cloud forcing bias ~~becomes~~ is more positive in POLARRES, the near-surface temperature still shows ~~a cooling response~~ cooling relative to NE30~~, which~~ . This is possibly due to ~~the increase in~~ increased summer surface albedo over northeastern Canada and the central Arctic (Figs. S4d-f). ~~Here, the~~ The increased summer surface albedo ~~could~~ may be related to the increased snow depth over northeastern Canada during ~~that~~ the summer season (not shown). Finally, the ~~warm bias~~ higher temperatures over land in CESM ~~is~~ are most likely related to a precipitation deficit, ~~which leads~~ leading to a soil moisture deficit~~and ultimately~~ , and ultimately resulting in less evaporative cooling and higher temperatures (Boza et al., 2025; Lin et al., 2017; Lorenz et al., 2016).

3.1.2 Precipitation Extremes

Figure 6 shows the present-day A annual mean precipitation differences between the NE30 and POLARRES grids, and the gridded outputs from ERA5, JRA-3Q, and RACMO. ~~Compared to~~ CESM generally tends to be wetter than ERA5, JRA-3Q, or RACMO ~~, CESM generally tends to be wetter~~ over the ocean and most of the Arctic land regions (i.e., poleward of 60°), and drier over most of the continental ~~midlatitudes, which~~mid-latitudes, which, to some extent, supports the existence of a precipitation deficit that contributes to the ~~warm~~ higher near-surface ~~temperature biases~~ temperatures over the continental mid-latitudes (Figs. 3 and 5). On average, CESM is wetter than ERA5, JRA-3Q, or RACMO with ~~mean biases ranging from 0.02-0.07~~ average differences ranging from 0.01-0.06 $\mathrm{mm\,d^{-1}}$ in NE30 (Figs. 6a,c,e) to ~~0.01-0.06~~ 0.02-0.07 $\mathrm{mm\,d^{-1}}$ in POLARRES (Figs. 6b,d,e), with the smallest ~~biases~~ differences relative to RACMO (Figs. 6e-f), followed by JRA-3Q (Figs. 6c-d) and ERA5 (Figs. 6a-b). ~~Although POLARRES generally tends to be slightly drier than NE30, the precipitation rates are higher~~ The wetter conditions in POLARRES are primarily related to the higher precipitation rates over elevated terrain, such as the mountainous terrains of Scandinavia, southern Iceland, and southern Alaska~~,~~. This is mainly due to the better resolved topography.

Figure 7 shows the present-day A ~~annual~~ daily precipitation PDFs, which have been constructed on a NE30 grid to ~~allow~~ enable a uniform comparison between the CESM grids and the reference data (also see Section 2.4). ~~Extreme~~ Generally, extreme precipitation rates are ~~generally~~ lowest in NE30 and highest in RACMO, particularly over ~~AWC, ECG, SCA, and SIC~~ Alaska/western Canada, eastern Canada/Greenland, the ocean and the sea ice region (Figs. 7a-c,f). POLARRES, JRA-3Q and ERA5 often ~~show~~ exhibit similar extreme precipitation rates. Only over ~~SIB~~Siberia, ERA5 and JRA-3Q, and, to lesser extent, RACMO do show lower extreme precipitation rates than NE30 ~~, POLARRES and RACMO~~ and POLARRES (Fig. 7d). ~~A possible explanation for the lower~~ It is possible that extreme precipitation rates are underestimated in ERA5 and JRA-3Q~~is that extreme precipitation rates are underestimated as~~, as, for example is found for ERA5 in Siberia (Clelland et al., 2024). ~~Over OCN~~The lower precipitation rates in RACMO could be attributed to uncertainties in the ERA5 data used to force RACMO. Over the ocean, the spread between CESM output and reference data is relatively small compared to the spread in other regions (Fig. 7e). The differences ~~between~~ in extreme precipitation rates ~~found in~~ between the CESM output and the reference data can ~~primarily~~ generally be explained by differences in horizontal resolution~~, which~~. This is lowest for NE30 (∼111 km), similar for POLARRES (∼28 km), JRA-3Q (∼40 km) and ERA5 (∼31 km), and highest for RACMO (11 km). The increase in the frequency of extreme precipitation rates with increased horizontal resolution may, for instance, be related to 1) a better representation of topography, leading to stronger orographic uplifts and orographic precipitation, 2) stronger updrafts, resulting in higher resolved-scale precipitation rates, and 3) an increase in the frequency and intensity of storm systems (e.g., Zarzycki and Jablonowski, 2014; O'Brien et al., 2016; Xu et al., 2022).

To better understand the patterns of precipitation extremes over the Arctic, we compare the CESM output with ERA5. Figure 8 shows the differences between CESM and ERA5 for the present-day A precipitation extreme metrics~~(P99, RX5day, R10mm, and CDD)~~. Compared to ERA5, the 99^{th} percentile of ~~precipitation (P99)~~ daily precipitation over the continents is lower in NE30, ~~especially over the mountainous regions~~particularly in mountainous regions, such as those along the eastern Pacific, Scandinavia, and southeastern Greenland and Iceland (Fig. 8b). In POLARRES, ~~P99~~ the 99^{th} percentile of

daily precipitation increases over the same regions, but ~~especially~~ particularly on the windward side of the mountainous regions, while ~~P99~~ it decreases on the more inland or leeward side, which is ~~especially visible along the eastern Pacific and particularly evident~~ over southeastern Greenland (Figs. 8c-d). The highest 5-day precipitation sum ~~(RX5day) biases are mostly positive~~ differences are positive over parts of eastern Asia, Greenland, and North America in NE30 (Fig. 8f). Dry ~~biases~~ precipitation differences are present over only a few regions, including parts of western Canada, ~~the northern Atlantic,~~ Europe, and western Russia. In POLARRES, ~~dry biases decrease, and wet biases increase, especially at lower latitudes over East Asia and North America~~ precipitation differences increase, particularly over parts of the Eurasian continent (Figs. 8g-h). The larger ~~biases in POLARRES could be the result of~~ differences in POLARRES may be partly due to a higher frequency and intensity of ~~(extra-tropical)~~ storm systems, which bring significant amounts of precipitation over a period of several days ~~(Zarzycki and Jablonowski, 2014; Xu et al., 2022; Zhu et al., 2023)~~(Zarzycki and Jablonowski, 2014). The number of heavy precipitation days ~~(R10mm) shows positive biases over the oceans and negative biases~~ shows positive differences over the northern Atlantic and negative differences over land in NE30, ~~especially~~ particularly over mountainous regions along the eastern Pacific, southeastern Greenland and Iceland, Scandinavia, and Scotland (Fig. 8j). In POLARRES, positive ~~biases~~ differences decrease over the North Atlantic (Figs. 8k-l). Furthermore, there is a ~~strong~~ significant increase in heavy precipitation days over southeastern Iceland, Scandinavia, ~~Scotland,~~ and the eastern Pacific, and a decrease over the more inland or leeward sides of mountainous regions in southeastern Greenland, Scandinavia, and the eastern Pacific. Finally, the greatest number of consecutive dry days ~~(CDD) shows a negative bias over~~ shows negative differences over parts of the central Arctic in NE30 and POLARRES, which could indicate a more variable precipitation climatology in CESM, reducing the length of dry periods (Figs. 8n-o). Relative to NE30, ~~CDD decreases~~ consecutive dry days increase slightly in POLARRES, ~~except for a few areas in the Canadian Arctic and East Asia that show an increase in CDD~~ however, differences between POLARRES and NE30 are mostly insignificant (Fig. 8p).

Comparisons between CESM output and other reference data (JRA-3Q and RACMO; Figs. S5 and S6) show that the spatial patterns are mostly similar compared to the differences between NE30/POLARRES and ERA5. Compared to JRA-3Q, CESM grids show a similar performance compared to ERA5, although ~~CESM tends to be slightly drier~~the dry precipitation differences for CESM appears slightly larger. Compared to RACMO, the ~~biases~~ differences are generally much more negative, indicating that precipitation in RACMO is more extreme in intensity and frequency, which is consistent with the larger extreme precipitation rates shown in Figure 7. The differences in extreme precipitation rates and metrics ~~are primarily~~ could be related to the different horizontal resolutions, with higher resolution grids generally simulating larger extreme precipitation rates and metrics, which is consistent with other studies showing the effect of horizontal resolution on (extreme) precipitation (Herrington et al., 2022; Bacmeister et al., 2014; Huang et al., 2022; Xu et al., 2022), ~~and to some extent to~~ as well as to different physics and parameterizations. The higher agreement between POLARRES and JRA-3Q and ERA5 suggests that POLARRES can better represent extreme precipitation than NE30. However, the discrepancies between POLARRES and RACMO also suggest to some extent that higher resolution grids are needed to further improve the representation of extreme precipitation, which is also demonstrated by CESM simulations on (storm-resolving) grids with a horizontal resolution of 14 km or higher (Huang et al., 2022; Xu et al., 2022).

3.2 Future projections

3.2.1 Temperature Extremes

Towards the end of the 21st century, both storylines project an increase in the magnitude of mean temperature and temperature extremes~~are projected to increase in magnitude in both storylines~~. Figure 9 shows the present-day B (2005–2014) annual mean near-surface temperature and the projected future (2090–2099) changes in annual mean near-surface temperature for two different storylines and CESM grids (NE30 and POLARRES), and the differences between the grids. The present-day panels (Figs. 9a,c,e,g) show that the annual mean temperature is relatively similar for ~~storyline~~ storylines ST1 (associated with strong Arctic tropospheric warming) and ~~storyline~~ ST2 (associated with strong ~~SST~~ sea surface warming in the Barents-Kara Seas), with ST1 being slightly colder than ST2 over the central Arctic. ~~In addition~~Additionally, the present-day B annual mean temperature is slightly lower than the present-day A annual mean temperature over the central Arctic (not shown), due to the lower SST originating from the CMIP6 models ~~that represent~~ representing the storylines (NorESM2-MM for ST1 and CNRM-ESM2-1 for ST2)~~and~~ , which force CESM.

Present-day differences between POLARRES and NE30 (Figs. 9i,k) indicate that POLARRES is generally colder than NE30, ~~except over Greenland and parts of~~ especially in ST1. Only over parts of Greenland, Alaska and eastern Siberia, ~~where~~ POLARRES is warmer than NE30. These differences are partly ~~related~~ due to the better resolved topography, ~~especially~~ particularly over Alaska, Greenland, Iceland, Scandinavia and parts of Siberia. Towards the end of the century, both storylines show ~~an~~ a significant increase in temperature (Figs. 9b,d,f,h). ~~Here,~~ ST2 ~~can be~~ is characterized by a warmer Arctic in the future, with temperature increases of up to about 19 °C over the Barents-Kara Seas (Figs. 9d,h), which can be linked to ~~the strong SST warming over the respective~~ strong sea surface warming in the region (Fig. 2d). ST1 shows ~~locally a~~ a local temperature decrease (up to about 2-3 °C) over the northern Atlantic (Figs. 9b,f), due to SST ~~cooling anomalies over the respective~~ cold anomalies in the region (Fig. 2c). These ~~SST cooling~~ cold anomalies are most likely related to increased freshwater inflow and ~~associated~~ the subsequent weakening of the Atlantic Meridional Overturning Circulation (AMOC) (Liu et al., 2020; Weijer et al., 2020). Future differences between POLARRES and NE30 (Figs. 9j,l) are relatively small and insignificant for both storylines~~with mixed signals in ST1 and a slightly colder trend in POLARRES in ST2, especially over the North American continent~~.

Stronger ST2 temperature responses are also evident from the projected changes in regional temperature extremes as shown in Figure 10. In all regions, ~~TXx and TNn~~ the annual maximum and minimum temperatures are projected to increase~~, with the~~ . The projected changes in high ~~(low) temperature extremes (TXx (TNn)) being smaller (larger)~~ temperature extremes are smaller than the projected changes in annual mean near-surface temperature, while the projected changes in low temperature extremes are larger (Figs. 10a-b). ~~In addition, the~~ Additionally, ST2 increases are ~~much larger than the~~ considerably larger than ST1 increases, with average differences of up to ~~11°C for TNn~~ about 9°C for annual minimum temperature between ST1 and ST2 (Figs. 10a-b). Specifically over ~~SIC~~the sea ice region, NE30 ~~(POLARRES) projects~~ and POLARRES project temperature increases of up to ~~11 °C (11 °C) and 21 °C (~~about 9 °C and 9 °C for annual maximum temperature, and 19 °C ~~) for TXx and TNn, respectively, suggesting that SST~~ and 20 °C for annual minimum temperature, respectively. This suggests that sea surface

warming and associated sea ice loss have the strongest ~~response~~ effect on temperature over the sea ice itself. Temperature increases over land are generally smaller with increases in the range ~~3-9°C (7-10 °C) and 6-12 °C (11-15 °C) for TXx and TNn respectively~~ of 3-8°C and 6-9 °C for the annual maximum temperature, and in the range of 7-11 °C and 12-14 °C for the annual minimum temperature in ST1 ~~(~~and ST2~~)~~, respectively. Here, POLARRES ~~predicts larger (smaller) temperature increases~~ generally predicts similar or larger increases in annual minimum temperature than NE30 ~~for TXx (TNn), with more pronounced TXx differences over ECG in ST1 and over SCA in ST1 and ST2~~and smaller annual maximum temperature increases over Alaska/western Canada and Siberia, as well as larger annual maximum temperature increases over eastern Canada/Greenland and Scandinavia. The differences over ~~ECG~~ eastern Canada/Greenland are associated with significantly stronger warming over northeastern Canada in POLARRES (Fig. S7b), mostly due to large decreases in snow depth ~~over the respective~~ in this region (not shown). As the present-day snow depth is overestimated, this warming trend is likely to be spurious. The differences over ~~SCA~~ other land regions, such as Scandinavia, are mostly related to ~~the stronger warming over the Eurasian continent~~ either stronger or weaker, though insignificant, warming over land (Fig. S7d). Temperature increases over the ocean are relatively small compared to the changes over land and sea ice, especially for ST1~~, which~~ . This is most likely ~~influenced by~~ due to the projected cooling over the northern Atlantic (Figs. S8b,j and S9b,j).

Consistent with the expected temperature increases, ~~the WSDI~~ warm spell duration also increases, ~~especially~~ particularly under ST2 and over ~~SIC, and the CSDI~~ the sea ice region. Conversely, cold spell duration decreases in all regions (Note that we have derived the future ~~WSDI and CSDI~~ warm and cold spell duration metrics based on the present-day B percentile thresholds; see Section 2.4). The ~~smaller~~ increases in warm spell duration and decreases in cold spell duration under ST1 ~~WSDI (CSDI) increase (decrease) over OCN and~~are smaller over the ocean and, to a lesser extent, over Scandinavia (for warm spell duration only) than in other regions. These smaller changes can be attributed to ~~some extentover SCA (for WSDI only) is related to~~ the projected cooling-induced ~~decrease~~ changes in warm spell (decrease) and cold spell (increase) ~~in WSDI (CSDI) in~~ duration over the northern Atlantic (Figs. S8f,n and S9f,n). ~~The differences~~

Although the differences in warm spell duration between POLARRES and NE30 are relatively small~~for the WSDI, although locally the differences are larger~~ , they are larger locally, e.g. over northeastern Canada or western Russia (Figs. S7f,h), where POLARRES projects a higher ~~WSDI. For the CSDI, the~~ warm spell duration. The cold spell duration differences are generally less negative~~, consistent with the smaller TNn increases in POLARRES (Fig. 10b). The projected increases (decreases) in WSDI (CSDI)~~ . The projected changes in warm spell (increase) and cold spell (decrease) duration are consistent with the findings of Screen et al. (2015), who assessed future changes in regional climate extremes arising from Arctic sea ice loss, using regions similar to those used in this study. Although the magnitude of change in ~~CSDI~~ cold spell duration is similar to that found in Screen et al. (2015), the projected changes in ~~WSDI~~ warm spell duration are much larger in this study. For example, in ~~ECG~~ eastern Canada/Greenland, the projected increase in ~~WSDI~~ warm spell duration in the Screen et al. (2015) study is up to about 35 $\mathrm{d\,yr^{-1}}$, while in this study the projected increases are up to about 220 $\mathrm{d\,yr^{-1}}$. A possible explanation for these differences is that the future experiments performed in Screen et al. (2015) were based on "idealized" model experiments using prescribed present-day SST and future sea ice concentrations, whereas our experiments use both prescribed future SST

and sea ice concentrations. This indicates that not only sea ice loss but also ~~SST~~ sea surface warming is an important driver of future changes in temperature extremes.

3.2.2 Precipitation Extremes

Figure 11 shows the present-day B (2005–2014) annual mean precipitation and projected future (2090–2099) changes in annual mean precipitation for two different storylines and CESM grids, and the differences between the grids. The present-day annual precipitation is similar for both ST1 and ST2 (Figs. 11a,c,e,g). However, regionally the annual mean precipitation is larger ~~(smaller)~~ in POLARRES, ~~especially~~ particularly on the windward ~~(leeward or more inland)~~ sides of mountainous ~~areas~~ regions in the eastern Pacific/southern Alaska, southeastern Greenland, Iceland, and Scandinavia. In contrast, precipitation tends to be smaller on the leeward or more inland sides of these mountainous regions (Figs. 11i,k). In the future, both NE30 and POLARRES project an increase in annual precipitation over most of the Arctic in both storylines, with the exception of the north(east)ern Atlantic where precipitation is projected to decrease (Figs. 11b,f,d,h). These decreases are larger in ST1 and are associated with decreases in SST and near-surface temperature over the region (Figs. 2c-d and 9b,f,d,h). As the projected cooling is greater in ST1, the amount of atmospheric moisture available for precipitation (as depicted by the Clausius-Clapeyron relationship) is less and can therefore explain the greater precipitation decreases in ST1. Future differences between POLARRES and NE30 (Figs. 11j,l) show consistent regional patterns in both storylines with wetter but insignificant trends in POLARRES over the northern Atlantic, the Scandinavian mountain ranges, and the eastern Pacific coast, and drier but insignificant trends over Greenland and some parts of the land (e.g. eastern Siberia in ST1 and Alaska in ST2).

Figure 12 shows the projected changes in regional precipitation extremes for two different storylines and CESM grids. In ST1~~(ST2), P99, RX5day, and R10mm increaseon average by ~3.0 (~4.0) $\mathrm{mm\,d^{-1}}$, ~10-12 (~15-18)~~ , the 99^{th} percentile of daily precipitation, the highest 5-day precipitation sum, and the number of heavy precipitation days increase, on average, by about 3 $\mathrm{mm\,d^{-1}}$, 6-7 mm, and ~~~4.0 (~5.0) d, respectively(Figs.12a-c).~~ 3 $\mathrm{d\,yr^{-1}}$, respectively, in NE30 and POLARRES (i.e. calculated for the domain 60-90°N). In ST2, the projected increases are slightly larger with average increases of about 4 $\mathrm{mm\,d^{-1}}$, 8-9 mm, and 4 $\mathrm{d\,yr^{-1}}$ for the 99^{th} percentile of daily precipitation, the highest 5-day precipitation sum, and heavy precipitation days, respectively. These increases correspond to mean relative increases of ~~26-27%(30-32%), 32-35%(44-45%), and 49-52% (56-58%) for P99, RX5day, and R10mm~~about 25-26%, 29-30%, and 36-40% in ST1 and 30-31%, 38%, and 43-50% in ST2, respectively (~~Figs~~ Fig. S10b-d), indicating that the ~~R10mm changes are relatively~~ relative changes in heavy precipitation days are the largest. As the annual mean precipitation increases by ~0.3 $\mathrm{mm\,d^{-1}}$ (~~21-22~~20-21%) in ST1 and ~0.4 $\mathrm{mm\,d^{-1}}$ (~~24-28~~25-26%) in ST2 (Figs. 11b,f,d,h; Fig. S10a), the projected changes in ~~the extremes are larger than~~ extremes exceed the changes in the means. Regionally, the ST2 increases are often larger than the ST1 increases ~~. Only over AWC~~ (Figs. 12a-c). Only over Alaska/western Canada the ST2 increases are slightly smaller than the ST1 increases for ~~P99 and R10mm.~~ the 99^{th} percentile of daily precipitation and the heavy precipitation days (Figs. 12a,c). The larger increases in ST2 are most likely related to the larger temperature increases in ST2, which lead to an increase in atmospheric moisture and thus more precipitation.

Consistent with the increase in mean and extreme precipitation, the ~~CDD shows a~~ number of consecutive dry days is projected to decrease in most regions, except for ~~SCA and AWC~~ Scandinavia where it is expected to increase (Fig. 12d). ~~Over AWC, the POLARRES grid projects an increase in CDD in ST2, while over SCA the increase in CDD is more pronounced in ST2 (both NE30 and POLARRES) and to a lesser extent in ST1 (POLARRES only). Similar CDD increases over SCA~~ Similar increases were also found ~~by Screen et al. (2015). The regions where CDD is projected to increase~~ in Scandinavia in Screen et al. (2015). Regions where consecutive dry days are projected to decrease are characterized by a ~~lower~~ high present-day ~~CDD, which is found over most of the Arctic. In regions with high present-day CDD~~ number of consecutive dry days. These regions (central Arctic, north(east)ern Greenland and Canada, and eastern Siberia; Figs. S11m-p, S12m-p) ~~, CDD is projected to decrease. This means relatively dry regions will become wetter, while wetter regions will become wetterand~~ are located at high latitudes and experience cold winters in which daily precipitation above 1 $\mathrm{mm\,d^{-1}}$ is relatively rare. With increased warming, precipitation increases and consequently consecutive dry days reduce, suggesting that these regions become wetter. Regions with a lower present-day number of consecutive dry days are generally located in the warmer, lower-latitude parts of the Arctic, where more precipitation occurs. In a warmer climate, these regions are expected to become wetter, as mean precipitation and precipitation extremes are projected to increase in intensity and frequency. However, these regions may also experience longer periods of drought ~~.~~ in summer in a warmer climate, resulting in an increased number of consecutive dry days. This results in a higher precipitation variability, particularly over regions such as Scandinavia.

Regarding the differences between NE30 and POLARRES for future precipitation (extremes), the differences often show mixed signals depending on the extreme metric and the region, although for some regions and extreme metrics the differences are more pronounced. For example, over ~~SCA,~~ Scandinavia POLARRES predicts larger absolute increases in ~~P99, RX5day, and R10mm~~ the 99^{th} percentile of daily precipitation, the highest 5-day precipitation sum, and heavy precipitation days for both storylines (Figs. 12 and S13). Furthermore, the ~~RX5day~~ highest 5-day precipitation sum differences are more pronounced over ~~AWC, SIB, and SIC~~Alaska/western Canada, Siberia, and the sea ice region, where POLARRES predicts larger ~~absolute~~ increases. The larger ~~RX5day increases could be related~~ increases in the highest 5-day precipitation sums may be partly due to the higher frequency and intensity of storm systems associated with higher resolution grids, as discussed above (Fig. 8). The larger increases over ~~SCA can~~ Scandinavia may partly be explained by increased moisture fluxes and enhanced moisture convergence associated with grid refinement, as seen in Figure 13. ~~This figure~~

Figure 13 shows the future changes in the annual mean vertically integrated moisture flux convergence, and the 850 hPa zonal wind and moisture fluxes for ST1/ST2 and NE30/POLARRES, and the differences between NE30 and POLARRES. Both NE30 and POLARRES predict a strengthening of westerly moisture transport over most of the higher mid-latitudes (i.e. roughly between 50°N and 70°N) in ST1 and over North America, northern Atlantic, and northern Europe in ST2 (Figs. 13a-c,g-i and S14). The NE30 moisture flux anomalies are associated with an eastward shift of the low-level North Atlantic jet in ST1 (Fig. 13d). The eastward shift of the North Atlantic jet also occurs in POLARRES but is accompanied by enhanced zonal wind near the climatological maximum of the jet (Figs. 13e-f). Here, the zonal wind changes are mainly due to changes occurring during winter (Figs. S15d-f and S16d-f). In ST2, the projected changes in zonal wind are much weaker compared to ~~the~~ those in ST1~~changes~~, where NE30 does not show significant changes in ~~the~~ zonal wind over the northern Atlantic region

(Fig. 13j). However, in ST2, POLARRES predicts a poleward shift of the jet and a significant intensification of zonal wind between Iceland and Scandinavia~~in ST2~~, though the area of significance is limited (Fig. 13k). These changes can be associated with zonal wind changes that mainly occur during winter and to a lesser extent during summer (Figs. S15j-l and S16j-l). The zonal wind changes over the northern Atlantic are most likely driven by SST changes and to a lesser extent by sea ice loss (Smirnov et al., 2015; Köhler et al., 2024; Wills et al., 2024), although the mechanisms for the differences between ST1 and ST2 as well as the relative contributions of SST and sea ice changes are unknown and therefore require more research in the future. Finally, the enhanced westerly moisture transport in ST1 and ST2 is accompanied by ~~enhanced~~ increased moisture convergence along the Scandinavian mountain ranges and is indicative of an intensification of precipitation over the respective region (Fig. 13b-c, ~~e-f), which can therefore explain why precipitation differences~~ h-i). However, the differences between POLARRES and NE30 are largely insignificant, which could be partly due to the relatively short time period (10 years) over which CESM model data are available. The short time period reduces the power of statistical analyses (i.e. due to the lower degrees of freedom), which may partly explain the absence of statistical significance. Therefore, the enhanced westerly moisture transport and convergence can only partially explain why the differences in precipitation between NE30 and POLARRES are more pronounced over ~~SCA and less pronounced~~ Scandinavia and less so in other regions. ~~However, it~~ It is likely that other drivers (e.g. orographic-induced updrafts and convection) also ~~play a role in the precipitation differences~~ influence the differences in precipitation between NE30 and POLARRES. ~~To understand the changes in the~~ However, to understand how the drivers of precipitation ~~means and extremes~~ change with grid refinement under different SST and sea ice regimes ~~more researchis needed as well~~requires further research.

4 Discussion and Conclusions

We have evaluated and assessed present-day and future temperature and precipitation extremes over the Arctic, using the variable-resolution Community Earth System Model version 2.2 (VR-CESM). We applied a globally uniform 1° CESM grid (NE30) and a dual-polar VR grid with regional grid refinements to 28 km over the Arctic and Antarctica (POLARRES). Using both model grids, we run three different types of simulations with interactively coupled atmosphere and land surface models, and prescribed sea ice and sea surface temperatures, namely: a 30-year simulation (1985–2014) used to evaluate the model grids, and 10-year present-day (2005–2014) and future (2090–2099) simulations used to assess future changes in temperature and precipitation extremes. Here, the 10-year present-day and future simulations follow two storylines of Arctic summer climate change representing a combination of strong/weak Arctic atmospheric warming at the 850 hPa level and ~~strong/weakSST~~ weak/strong sea surface warming in the Barents-Kara Seas. To evaluate and assess present-day and future temperature and precipitation extremes, we analyzed eight extreme metrics, including four extreme metrics each for temperature and precipitation.

The evaluations show that both NE30 and POLARRES have climatological ~~warm temperature biases over most of the land surface and climatological cold temperature biases~~ higher temperatures over Europe and parts of Greenland, and climatological lower temperatures over the central Arctic, ~~Greenland, and~~ northeastern Canada, ~~with cold biases being larger~~ and parts of Greenland, with lower temperatures in POLARRES than in NE30. For temperature extreme metrics, POLARRES performs

slightly better in simulating ~~TXx~~annual maximum temperature, while NE30 performs better for other temperature extreme metrics (~~TNn, WSDI, and CSDI~~annual mimimum temperature and warm and cold spell duration). Spatially, the cold and warm ~~biases in TNn and TXx~~ temperature differences in annual minimum and maximum temperature correspond well to the spatial patterns of the climatological biases, but the magnitude of the extreme temperature ~~biases~~ differences is much larger. The ~~cold biases~~ lower temperatures can be partly explained by large warm near-surface temperature biases in the reanalysis products, caused by poor representation of snow over sea ice and the thickness of the sea ice itself. ~~In addition, the cold biases~~ Additionally, the lower temperatures may be related to ~~increased~~ overestimated cloud cover (relative to the reference data) leading to excessive reflection and cooling, and increased summer surface albedo and snow depth over northeastern Canada and the central Arctic in POLARRES. However, to understand the magnitude of the ~~cold bias~~ lower temperatures in CESM, validation of near-surface temperature against meteorological observations in the Arctic is needed and thus recommended for future research.

NE30 and POLARRES have climatological dry ~~biases over the continent~~ precipitation differences over most of continental mid-latitudes and climatological wet ~~biases~~ precipitation differences over the oceans and most of the Arctic land regions. The dry ~~biases~~ precipitation differences indicate a precipitation deficit over land, which, to some extent, may contribute to the ~~warm biases~~ higher temperatures over the land surface via a soil moisture deficit and associated reduced evaporative cooling. POLARRES performs slightly better than NE30 in simulating the precipitation climatology, with higher precipitation rates over mountainous terrain in Scandinavia, southeastern Greenland, and Alaska, which is mainly due to the higher grid resolution and better resolved topography. With respect to precipitation extremes, the performance of POLARRES and NE30 depends on the performance criteria (i.e. ~~RMSE or bias~~RMSD or AVGD) and the data used as reference (ERA5, JRA-3Q or RACMO) for the evaluation. Using the ~~RSME~~ RSMD as the performance criterion, POLARRES generally performs better than NE30 in simulating ~~P99~~the 99^{th} percentile of daily precipitation and the number of heavy precipitation and consecutive dry days, ~~R10mm, and CDD,~~ while NE30 performs better in simulating ~~RX5day~~highest 5-day precipitation sums. The larger ~~RX5day biases in POLARRES could~~ highest 5-day precipitation sum differences in POLARRES may partly be related to a higher frequency and intensity of storm systems bringing significant amounts of precipitation over a period of several days. However, more research is needed to understand the impact of grid refinement on the frequency and intensity of storm systems in the Arctic as well as its effect on precipitation extremes in CESM.

Future projections suggest that ~~annual mean temperature, TXx, TNn, and WSDI (CSDI)~~ the annual mean, maximum and minimum temperature, and the warm spell duration will mostly increase ~~(decrease)~~ in magnitude or duration in all regions of the Arctic, ~~with the~~ while the cold spell duration is projected to decrease. Here, the rate of change ~~varying~~ varies by region and the applied storyline. The projected increases are larger for ST2, with the largest increases being projected over the currently sea ice-dominated central Arctic. The larger increases in ST2 suggest that ~~SST~~ sea surface warming and associated sea ice loss over the Arctic have a stronger response than Arctic low-level tropospheric warming, especially over the sea ice itself. In ST1, increases in temperature extremes are especially smaller over the ocean and Scandinavia, and are accompanied by a local decrease in annual mean temperature due to projected cooling anomalies over the North Atlantic. Regarding the differences in projected changes by NE30 and POLARRES, the differences are more pronounced for ~~TXx~~ annual maximum temperature

with larger increases in POLARRES over northeastern Canada~~and Greenland in ST1~~, which is mainly due to reduced snow depth~~over northeastern Canada, and over Scandinavia in ST1 and ST2, which is mainly due to stronger~~ . The annual maximum temperature differences over other land regions, such as the larger increases in Scandinavia or the smaller increases in Siberia, could be related to either stronger or weaker warming over the Eurasian continent.

Precipitation means and extremes are mostly projected to increase in intensity and frequency, with the projected changes in extremes being greater than the projected changes in mean precipitation. Only over the northern Atlantic, annual mean precipitation is projected to decrease due to local cooling-induced decreases in atmospheric moisture and precipitation. The projected changes are mostly stronger for ST2, which can be associated with the stronger warming-induced increases in atmospheric moisture and precipitation. The projected changes in ~~CDD suggest that regions with~~ consecutive dry days suggest that relatively dry regions with a high present-day ~~CDD (e.g.,~~ number of consecutive dry days (central Arctic, north(east)ern Greenland and Canada, and eastern Siberia) will become wetter ~~(i.e., CDD will decrease), and regions with~~ in a warming climate, resulting in a decreased number of consecutive dry days. Relatively wet regions with a low present-day ~~CDD~~ number of consecutive dry days (e.g., Scandinavia) ~~will~~ are expected to become wetter in a warming climate but may also experience longer periods of droughts ~~(i.e., CDD will increase).~~ in summer, resulting in an increased number of consecutive dry days.

Regarding the differences in regional projected changes by NE30 and POLARRES, the differences often show mixed signals depending on the extreme metrics and region but are especially more pronounced over Scandinavia as well as for ~~RX5day~~the highest 5-day precipitation sums. Here, POLARRES projects larger ~~RX5day increases , which is~~ increases in highest 5-day precipitation sums, which may be partly due to the higher frequency and intensity of storm systems associated with higher resolution grids. Further, POLARRES projects larger increases in ~~P99, RX5day, and R10mm, which is most likely~~ the 99^{th} percentile of daily precipitation, the highest 5-day precipitation sums, and heavy precipitation days over Scandinavia, which may be partly due to enhanced westerly moisture transport and convergence over the respective region, and possibly due to other drivers, such as orographic-induced updrafts and convection.

This study shows that, relative to the ~~used reference data~~ reference data used, VR-CESM generally outperforms ~~a~~ the coarse-resolution CESM grid in simulating precipitation climatology and extremes~~, while a~~ . In contrast, the coarse-resolution CESM grid often ~~outperforms VR-CESM~~ performs better in simulating temperature climatology and extremes. This ~~implies~~ suggests that VR-CESM can be a ~~useful modelling~~ valuable tool for simulating present-day and future extremes in the Arctic~~, but also that~~ . However, model improvements are needed to~~reduce temperature biases. Additionally, more~~ , for instance, reduce temperature differences relative to reference data. It is important to note that the evaluation of model outcomes primarily relied on reanalysis fields, which can be strongly influenced by the underlying forecast model in regions with little or no observations. Therefore, incorporating meteorological observations and remote sensing data may help to better understand the differences between CESM model output and observational references. Additionally, further research is needed to ~~improve our understanding on the~~ better understand the physical drivers and mechanisms ~~for warming-induced or grid refinement-induced~~ of changes in Arctic temperature and precipitation means and extremes, whether induced by warming or by the use of grid refinement.

Finally, this study demonstrates ~~the large range that can exist between~~ that the projected changes in Arctic temperature and precipitation extremes ~~of two different storylines. Nonetheless, the projected range in model outcomes is most likely smaller than the projected range of model outcomes that stem from CMIP6 scenario-based multi-model approaches. For example, the 5%-95% range of surface temperature changes as projected by CMIP6 models in the Arctic is about 9 °C under SSP5-8.5 (Lee et al., 2021), while the average range of~~ can vary substantially depending on the storyline applied. In this study, the average difference in near-surface temperature ~~changes~~ projections between ST1 and ST2 is about 2.5-3 °C ~~in the Arctic at the~~ by the end of the 21^{st} century. ~~This narrows down the uncertainty range , which means that the outcomes of this study, besides contributing to a better~~ However, this is considerably narrower than the 5^{th}-95^{th} percentile range of approximately 9 °C in Arctic surface temperature change projected by CMIP6 models under SSP5-8.5 (Lee et al., 2021). The narrower range of projected outcomes suggests that the storyline approach used here may support a more targeted understanding of future Arctic temperature and precipitation extremes~~in the Arctic, could also be more helpful in the development of~~ , and could help inform adaptation and mitigation strategies ~~that cope with the adverse impacts of temperature and precipitation extremes~~ in response to climate extremes in the region.

Data availability. Publicly available data will be stored in a data archive on Zenodo (https://doi.org/10.5281/zenodo.14961708) upon publication. The data archive contains the daily near-surface temperature and precipitation output and calculated temperature and precipitation extreme metrics.

Author contributions. RRW and WJB designed the study. RRW ran the model with technical support of ARH. XJL developed the Arctic storylines. CTD provided the RACMO data used for evaluation of the model output. RRW analyzed the model results and prepared the text and figures. WJB supervised RRW. All authors contributed to the final text.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. TEXT

Acknowledgements. We acknowledge the support of PolarRES (grant no. 101003590), a project of the European Union's Horizon 2020 research and innovation programme. Adam R. Herrington has been supported by the NSF National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation (grant no. 1852977). Computing and data storage resources, including the Cheyenne supercomputer (https://doi.org/10.5065/D6RX99HX, Computational and Information Systems Laboratory, 2019), were provided by the Computational and Information Systems Laboratory (CISL) at NSF NCAR.

References

Bacmeister, J. T., Wehner, M. F., Neale, R. B., Gettelman, A., Hannay, C., Lauritzen, P. H., Caron, J. M., and Truesdale, J. E.: Exploratory high-resolution climate simulations using the community atmosphere model (CAM), Journal of Climate, 27, https://doi.org/10.1175/JCLI-D-13-00387.1, 2014.

Bambach, N. E., Rhoades, A. M., Hatchett, B. J., Jones, A. D., Ullrich, P. A., and Zarzycki, C. M.: Projecting climate change in South America using variable-resolution Community Earth System Model: An application to Chile, International Journal of Climatology, https://doi.org/10.1002/joc.7379, 2021.

Batrak, Y. and Müller, M.: On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice, Nature Communications, 10, https://doi.org/10.1038/s41467-019-11975-3, 2019.

Beljaars, A. C., Brown, A. R., and Wood, N.: A new parametrization of turbulent orographic form drag, Quarterly Journal of the Royal Meteorological Society, 130, https://doi.org/10.1256/qj.03.73, 2004.

Benjamini, Y. and Hochberg, Y.: Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, Journal of the Royal Statistical Society Series B: Statistical Methodology, 57, https://doi.org/10.1111/j.2517-6161.1995.tb02031.x, 1995.

Bogenschutz, P. A., Gettelman, A., Morrison, H., Larson, V. E., Craig, C., and Schanen, D. P.: Higher-order turbulence closure and its impact on climate simulations in the community atmosphere model, Journal of Climate, 26, https://doi.org/10.1175/JCLI-D-13-00075.1, 2013.

Boza, B., Herrington, A., Ilicak, M., Danabasoglu, G., and Sen, O. L.: Exploring Climate of the Euro-Mediterranean Using a Variable-Resolution Configuration of the Global Community Earth System Model (VR-CESM), Journal of Geophysical Research: Atmospheres, 130, https://doi.org/10.1029/2024JD041300, 2025.

CARRA: Copernicus Arctic Regional Reanalysis, https://climate.copernicus.eu/copernicus-arctic-regional-reanalysis-service, 2025.

Chan, W. C., Shepherd, T. G., Facer-Childs, K., Darch, G., and Arnell, N. W.: Storylines of UK drought based on the 2010-2012 event, Hydrology and Earth System Sciences, 26, https://doi.org/10.5194/hess-26-1755-2022, 2022.

Clelland, A. A., Marshall, G. J., and Baxter, R.: Evaluating the performance of key ERA-Interim, ERA5 and ERA5-Land climate variables across Siberia, International Journal of Climatology, 44, 2318–2342, https://doi.org/10.1002/joc.8456, 2024.

Computational and Laboratory, I. S.: Cheyenne supercomputer, https://doi.org/10.5065/D6RX99HX, 2019.

Dai, A., Luo, D., Song, M., and Liu, J.: Arctic amplification is caused by sea-ice loss under increasing CO2, Nature Communications, 10, https://doi.org/10.1038/s41467-018-07954-9, 2019.

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001916, 2020.

Danielson, J. J. and Gesch, D. B.: Global multi-resolution terrain elevation data 2010 (GMTED2010), Tech. rep., https://doi.org/10.3133/ofr20111073, 2011.

Datta, R. T., Herrington, A., Lenaerts, J. T., Schneider, D. P., Trusel, L., Yin, Z., and Dunmire, D.: Evaluating the impact of enhanced horizontal resolution over the Antarctic domain using a variable-resolution Earth system model, Cryosphere, 17, https://doi.org/10.5194/tc-17-3847-2023, 2023.

Flanner, M. G. and Zender, C. S.: Snowpack radiative heating: Influence on Tibetan Plateau climate, Geophysical Research Letters, 32, L06 501, https://doi.org/10.1029/2004GL022076, 2005.

Flanner, M. G., Zender, C. S., Randerson, J. T., and Rasch, P. J.: Present-day climate forcing and response from black carbon in snow, Journal of Geophysical Research Atmospheres, 112, https://doi.org/10.1029/2006JD008003, 2007.

Gates, W. L., Boyle, J. S., Covey, C., Dease, C. G., Doutriaux, C. M., Drach, R. S., Fiorino, M., Gleckler, P. J., Hnilo, J. J., Marlais, S. M., Phillips, T. J., Potter, G. L., Santer, B. D., Sperber, K. R., Taylor, K. E., and Williams, D. N.: An Overview of the Results of the Atmospheric Model Intercomparison Project (AMIP I), Bulletin of the American Meteorological Society, 80, 29–55, https://doi.org/10.1175/1520-0477(1999)080<0029:AOOTRO>2.0.CO;2, 1999.

Gettelman, A. and Morrison, H.: Advanced two-moment bulk microphysics for global models. Part I: Off-line tests and comparison with other schemes, Journal of Climate, 28, https://doi.org/10.1175/JCLI-D-14-00102.1, 2015.

Gettelman, A., Morrison, H., Thayer-Calder, K., and Zarzycki, C. M.: The Impact of Rimed Ice Hydrometeors on Global and Regional Climate, Journal of Advances in Modeling Earth Systems, 11, https://doi.org/10.1029/2018MS001488, 2019.

Guba, O., Taylor, M. A., Ullrich, P. A., Overfelt, J. R., and Levy, M. N.: The spectral element method (SEM) on variable-resolution grids: evaluating grid sensitivity and resolution-aware numerical viscosity, Geoscientific Model Development, 7, 2803–2816, https://doi.org/10.5194/gmd-7-2803-2014, 2014.

Guo, Z., Wang, M., Qian, Y., Larson, V. E., Ghan, S., Ovchinnikov, M., Bogenschutz, P. A., Gettelman, A., and Zhou, T.: Parametric behaviors of CLUBB in simulations of low clouds in the Community Atmosphere Model (CAM), Journal of Advances in Modeling Earth Systems, 7, https://doi.org/10.1002/2014MS000405, 2015.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, Bulletin of the American Meteorological Society, 90, https://doi.org/10.1175/2009BAMS2607.1, 2009.

Herrington, A. R., Lauritzen, P. H., Lofverstrom, M., Lipscomb, W. H., Gettelman, A., and Taylor, M. A.: Impact of Grids and Dynamical Cores in CESM2.2 on the Surface Mass Balance of the Greenland Ice Sheet, Journal of Advances in Modeling Earth Systems, 14, https://doi.org/10.1029/2022MS003192, 2022.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., and Kanae, S.: Global flood risk under climate change, Nature Climate Change, 3, 816–821, https://doi.org/10.1038/nclimate1911, nULL, 2013.

Huang, X., Gettelman, A., Skamarock, W. C., Lauritzen, P. H., Curry, M., Herrington, A., Truesdale, J. T., and Duda, M.: Advancing precipitation prediction using a new-generation storm-resolving model framework - SIMA-MPAS (V1.0): a case study over the western United States, Geoscientific Model Development, 15, https://doi.org/10.5194/gmd-15-8135-2022, 2022.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Goldewijk, K. K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren,

D. P., and Zhang, X.: Harmonization of global land use change and management for the period 850-2100 (LUH2) for CMIP6, Geoscientific Model Development, 13, https://doi.org/10.5194/gmd-13-5425-2020, 2020.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research Atmospheres, 113, https://doi.org/10.1029/2008JD009944, 2008.

Johnson, N. C., Xie, S. P., Kosaka, Y., and Li, X.: Increasing occurrence of cold and warm extremes during the recent global warming slowdown, Nature Communications, 9, https://doi.org/10.1038/s41467-018-04040-y, 2018.

Kim, B. M., Son, S. W., Min, S. K., Jeong, J. H., Kim, S. J., Zhang, X., Shim, T., and Yoon, J. H.: Weakening of the stratospheric polar vortex by Arctic sea-ice loss, Nature Communications, 5, https://doi.org/10.1038/ncomms5646, 2014.

Kim, Y. H., Min, S. K., Zhang, X., Sillmann, J., and Sandstad, M.: Evaluation of the CMIP6 multi-model ensemble for climate extreme indices, Weather and Climate Extremes, 29, https://doi.org/10.1016/j.wace.2020.100269, 2020.

Klein-Goldewijk, K., Beusen, A., Doelman, J., and Stehfest, E.: Anthropogenic land use estimates for the Holocene - HYDE 3.2, Earth System Science Data, 9, https://doi.org/10.5194/essd-9-927-2017, 2017.

Kosaka, Y., Kobayashi, S., Harada, Y., Kobayashi, C., Naoe, H., Yoshimoto, K., Harada, M., Goto, N., Chiba, J., Miyaoka, K., Sekiguchi, R., Deushi, M., Kamahori, H., Nakaegawa, T., Tanaka, T. Y., Tokuhiro, T., Sato, Y., Matsushita, Y., and Onogi, K.: The JRA-3Q reanalysis, Journal of the Meteorological Society of Japan, 102, https://doi.org/10.2151/jmsj.2024-004, 2024.

Kug, J. S., Jeong, J. H., Jang, Y. S., Kim, B. M., Folland, C. K., Min, S. K., and Son, S. W.: Two distinct influences of Arctic warming on cold winters over North America and East Asia, Nature Geoscience, 8, https://doi.org/10.1038/ngeo2517, 2015.

Köhler, D., Räisänen, P., Naakka, T., Nordling, K., and Sinclair, V. A.: The future North Atlantic jet stream and storm track: relative contributions from sea ice and sea surface temperature changes, EGUSphere, https://doi.org/10.5194/egusphere-2024-3713, 2024.

Landrum, L. and Holland, M. M.: Extremes become routine in an emerging new Arctic, Nature Climate Change, 10, https://doi.org/10.1038/s41558-020-0892-z, 2020.

Lauritzen, P. H., Bacmeister, J. T., Callaghan, P. F., and Taylor, M. A.: `NCAR_Topo` (v1.0): NCAR global model topography generation software for unstructured grids, https://doi.org/10.5194/gmd-8-3975-2015, 2015.

Lauritzen, P. H., Nair, R. D., Herrington, A. R., Callaghan, P., Goldhaber, S., Dennis, J. M., Bacmeister, J. T., Eaton, B. E., Zarzycki, C. M., Taylor, M. A., Ullrich, P. A., Dubos, T., Gettelman, A., Neale, R. B., Dobbins, B., Reed, K. A., Hannay, C., Medeiros, B., Benedict, J. J., and Tribbia, J. J.: NCAR Release of CAM-SE in CESM2.0: A Reformulation of the Spectral Element Dynamical Core in Dry-Mass Vertical Coordinates With Comprehensive Treatment of Condensates and Energy, Journal of Advances in Modeling Earth Systems, 10, https://doi.org/10.1029/2017MS001257, 2018.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Martin, M. V., and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, Journal of Advances in Modeling Earth Systems, 11, 4245–4287, https://doi.org/10.1029/2018MS001583, 2019.

Lee, J.-Y., Marotzke, J., Bala, G., Cao, L., Corti, S., Dunne, J., Engelbrecht, F., Fischer, E., Fyfe, J., Jones, C., Maycock, A., Mutemi, J., Ndiaye, O., Panickal, S., and Zhou, T.: Future Global Climate: Scenario-based Projections and Near-term Information, pp. 553–672, Cambridge University Press, https://doi.org/10.1017/9781009157896.006, 2021.

Levine, X. J., Williams, R. S., Marshall, G., Orr, A., Graff, L. S., Handorf, D., Karpechko, A., Köhler, R., Wijngaard, R. R., Johnston, N., Lee, H., Nieradzik, L., and Mooney, P. A.: Storylines of summer Arctic climate change constrained by Barents-Kara seas and Arctic tropospheric warming for climate risk assessment, Earth System Dynamics, 15, 1161–1177, https://doi.org/10.5194/esd-15-1161-2024, 2024.

Lin, Y., Dong, W., Zhang, M., Xie, Y., Xue, W., Huang, J., and Luo, Y.: Causes of model dry and warm bias over central U.S. and impact on climate projections, Nature Communications, 8, 881, https://doi.org/10.1038/s41467-017-01040-2, 2017.

Liu, W., Fedorov, A. V., Xie, S. P., and Hu, S.: Climate impacts of a weakened Atlantic meridional overturning circulation in a warming climate, Science Advances, 6, https://doi.org/10.1126/sciadv.aaz4876, 2020.

Liu, X., Ma, P. L., Wang, H., Tilmes, S., Singh, B., Easter, R. C., Ghan, S. J., and Rasch, P. J.: Description and evaluation of a new four-mode version of the Modal Aerosol Module (MAM4) within version 5.3 of the Community Atmosphere Model, Geoscientific Model Development, 9, https://doi.org/10.5194/gmd-9-505-2016, 2016.

Loeb, N. A., Crawford, A., Herrington, A., McCrystall, M., Stroeve, J., and Hanesiak, J.: Projections and Physical Drivers of Extreme Precipitation in Greenland Baffin Bay, Journal of Geophysical Research: Atmospheres, 129, https://doi.org/10.1029/2024JD041375, 2024.

Lorenz, R., Argüeso, D., Donat, M. G., Pitman, A. J., Hurk, B. V. D., Berg, A., Lawrence, D. M., Chéruy, F., Ducharne, A., Hagemann, S., Meier, A., Milly, P. C., and Seneviratne, S. I.: Influence of land-atmosphere feedbacks on temperature and precipitation extremes in the GLACE-CMIP5 ensemble, Journal of Geophysical Research, 121, https://doi.org/10.1002/2015JD024053, 2016.

Masrur, A., Petrov, A. N., and DeGroote, J.: Circumpolar spatio-temporal patterns and contributing climatic factors of wildfire activity in the Arctic tundra from 2001-2015, https://doi.org/10.1088/1748-9326/aa9a76, 2018.

Matthes, H., Rinke, A., and Dethloff, K.: Recent changes in Arctic temperature extremes: Warm and cold spells during winter and summer, Environmental Research Letters, 10, https://doi.org/10.1088/1748-9326/10/11/114020, 2015.

McCrystall, M. R., Stroeve, J., Serreze, M., Forbes, B. C., and Screen, J. A.: New climate models reveal faster and larger increases in Arctic precipitation than previously projected, Nature Communications, 12, https://doi.org/10.1038/s41467-021-27031-y, 2021.

Mindlin, J., Shepherd, T. G., Vera, C. S., Osman, M., Zappa, G., Lee, R. W., and Hodges, K. I.: Storyline description of Southern Hemisphere midlatitude circulation and precipitation response to greenhouse gas forcing, Climate Dynamics, 54, https://doi.org/10.1007/s00382-020-05234-1, 2020.

Morlighem, M., Williams, C. N., Rignot, E., An, L., Arndt, J. E., Bamber, J. L., Catania, G., Chauché, N., Dowdeswell, J. A., Dorschel, B., Fenty, I., Hogan, K., Howat, I., Hubbard, A., Jakobsson, M., Jordan, T. M., Kjeldsen, K. K., Millan, R., Mayer, L., Mouginot, J., Noël, B. P., O'Cofaigh, C., Palmer, S., Rysgaard, S., Seroussi, H., Siegert, M. J., Slabon, P., Straneo, F., van den Broeke, M. R., Weinrebe, W., Wood, M., and Zinglersen, K. B.: BedMachine v3: Complete Bed Topography and Ocean Bathymetry Mapping of Greenland From Multibeam Echo Sounding Combined With Mass Conservation, Geophysical Research Letters, 44, https://doi.org/10.1002/2017GL074954, 2017.

Morlighem, M., Rignot, E., Binder, T., Blankenship, D., Drews, R., Eagles, G., Eisen, O., Ferraccioli, F., Forsberg, R., Fretwell, P., Goel, V., Greenbaum, J. S., Gudmundsson, H., Guo, J., Helm, V., Hofstede, C., Howat, I., Humbert, A., Jokat, W., Karlsson, N. B., Lee, W. S., Matsuoka, K., Millan, R., Mouginot, J., Paden, J., Pattyn, F., Roberts, J., Rosier, S., Ruppel, A., Seroussi, H., Smith, E. C., Steinhage,

D., Sun, B., den Broeke, M. R., Ommen, T. D., van Wessem, M., and Young, D. A.: Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the Antarctic ice sheet, Nature Geoscience, 13, https://doi.org/10.1038/s41561-019-0510-8, 2020.

Morris, M. and Kushner, P. J.: Mechanisms of the Extreme Wind Speed Response to Climate Change in Variable-Resolution Climate Simulations of Western, Central, and Atlantic Canada, Journal of Climate, https://doi.org/10.1175/JCLI-D-24-0532.1, 2025.

Morris, M., Kushner, P. J., Moore, G. W., and Mercan, O. Y.: Atmospheric Circulation Patterns Associated with Extreme Wind Events in Canadian Cities, Journal of Climate, 36, https://doi.org/10.1175/JCLI-D-22-0719.1, 2023.

Morris, M., Kushner, P. J., Moore, G. W. K., and Mercan, O.: Resolution Dependence of Extreme Wind Speed Projections in the Great Lakes Region, Journal of Climate, 37, 3153–3171, https://doi.org/10.1175/JCLI-D-23-0547.1, 2024.

O'Brien, T. A., Collins, W. D., Kashinath, K., Rübel, O., Byna, S., Gu, J., Krishnan, H., and Ullrich, P. A.: Resolution dependence of precipitation statistical fidelity in hindcast simulations, Journal of Advances in Modeling Earth Systems, 8, https://doi.org/10.1002/2016MS000671, 2016.

Oleson, K., Lawrence, D., Bonan, G., Drewniak, B., Huang, M., C.D., K., Levis, S., Li, F., Riley, W., Subin, Z., Swenson, S., P.E., T., Bozbiyik, A., Fisher, R., Heald, C., Kluzek, E., Lamarque, J.-F., Leung, L. R., Lipscomb, W., Muszala, S., Ricciuto, D., Sacks, W., Tang, J., and Yang, Z.-L.: Technical description of version 4.5 of the Community Land Model (CLM). Ncar Tech. Note NCAR/TN-503+STR. National Center for Atmospheric Research, Boulder, Geophysical Research Letters, 37, 1–434, https://doi.org/10.5065/D6RR1W7M, 2013.

Overland, J., Dunlea, E., Box, J. E., Corell, R., Forsius, M., Kattsov, V., Olsen, M. S., Pawlak, J., Reiersen, L. O., and Wang, M.: The urgency of Arctic change, https://doi.org/10.1016/j.polar.2018.11.008, 2019.

Overland, J. E.: Arctic Climate Extremes, https://doi.org/10.3390/atmos13101670, 2022.

Paik, S., An, S. I., Min, S. K., King, A. D., and Kim, S. K.: Emergent constraints on future extreme precipitation intensification: from global to continental scales, Weather and Climate Extremes, 42, https://doi.org/10.1016/j.wace.2023.100613, 2023.

Park, H. S., Kim, S. J., Stewart, A. L., Son, S. W., and Seo, K. H.: Mid-Holocene Northern Hemisphere warming driven by Arctic amplification, Science Advances, 5, https://doi.org/10.1126/sciadv.aax8203, 2019.

Pfahl, S., O'Gorman, P. A., and Fischer, E. M.: Understanding the regional pattern of projected future changes in extreme precipitation, Nature Climate Change, 7, https://doi.org/10.1038/nclimate3287, 2017.

Previdi, M., Smith, K. L., and Polvani, L. M.: Arctic amplification of climate change: A review of underlying mechanisms, https://doi.org/10.1088/1748-9326/ac1c29, 2021.

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., and Laaksonen, A.: The Arctic has warmed nearly four times faster than the globe since 1979, Communications Earth and Environment, 3, https://doi.org/10.1038/s43247-022-00498-3, 2022.

Rhoades, A. M., Huang, X., Ullrich, P. A., and Zarzycki, C. M.: Characterizing Sierra Nevada Snowpack Using Variable-Resolution CESM, Journal of Applied Meteorology and Climatology, 55, 173–196, https://doi.org/10.1175/JAMC-D-15-0156.1, 2016.

Rhoades, A. M., Ullrich, P. A., Zarzycki, C. M., Johansen, H., Margulis, S. A., Morrison, H., Xu, Z., and Collins, W. D.: Sensitivity of Mountain Hydroclimate Simulations in Variable-Resolution CESM to Microphysics and Horizontal Resolution, Journal of Advances in Modeling Earth Systems, 10, 1357–1380, https://doi.org/10.1029/2018MS001326, 2018.

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

Screen, J. A., Deser, C., and Sun, L.: Projected changes in regional climate extremes arising from Arctic sea ice loss, Environmental Research Letters, 10, https://doi.org/10.1088/1748-9326/10/8/084006, 2015.

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J.-F., Moine, M.-P., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-y Mélia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., and Madec, G.: Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate, Journal of Advances in Modeling Earth Systems, 11, 4182–4227, https://doi.org/10.1029/2019MS001791, 2019.

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geoscientific Model Development, 13, 6165–6200, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

Seneviratne, S., Nicholls, N., Easterling, D., Goodess, C., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., and Zhang, X.: Changes in Climate Extremes and their Impacts on the Natural Physical Environment, pp. 109–230, Cambridge University Press, ISBN 1107025060, https://doi.org/10.1017/CBO9781139177245.006, nULL, 2012.

Seneviratne, S. I. and Hauser, M.: Regional Climate Sensitivity of Climate Extremes in CMIP6 Versus CMIP5 Multimodel Ensembles, Earth's Future, 8, https://doi.org/10.1029/2019EF001474, 2020.

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Luca, A. D., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., and Zhou, B.: Weather and Climate Extreme Events in a Changing Climate, pp. 1513–1766, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896.013, 2021.

Shaw, J., McGraw, Z., Bruno, O., Storelvmo, T., and Hofer, S.: Using Satellite Observations to Evaluate Model Microphysical Representation of Arctic Mixed-Phase Clouds, Geophysical Research Letters, 49, e2021GL096 191, https://doi.org/https://doi.org/10.1029/2021GL096191, e2021GL096191 2021GL096191, 2022.

Shepherd, T. G., Boyd, E., Calel, R. A., Chapman, S. C., Dessai, S., Dima-West, I. M., Fowler, H. J., James, R., Maraun, D., Martius, O., Senior, C. A., Sobel, A. H., Stainforth, D. A., Tett, S. F., Trenberth, K. E., van den Hurk, B. J., Watkins, N. W., Wilby, R. L., and Zenghelis, D. A.: Storylines: an alternative approach to representing uncertainty in physical aspects of climate change, Climatic Change, 151, https://doi.org/10.1007/s10584-018-2317-9, 2018.

Sillmann, J., Kharin, V., Zhang, X., Zwiers, F., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, Journal of Geophysical Research Atmospheres, 118, 1716–1733, https://doi.org/10.1002/jgrd.50203, 2013a.

Sillmann, J., Kharin, V., Zwiers, F., Zhang, X., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections, Journal of Geophysical Research Atmospheres, 118, 2473–2493, https://doi.org/10.1002/jgrd.50188, 2013b.

Sillmann, J., Shepherd, T. G., van den Hurk, B., Hazeleger, W., Martius, O., Slingo, J., and Zscheischler, J.: Event-Based Storylines to Address Climate Risk, https://doi.org/10.1029/2020EF001783, 2021.

Skamarock, W. C., Snyder, C., Klemp, J. B., and Park, S. H.: Vertical resolution requirements in atmospheric simulation, Monthly Weather Review, 147, https://doi.org/10.1175/MWR-D-19-0043.1, 2019.

Smirnov, D., Newman, M., Alexander, M. A., Kwon, Y.-O., and Frankignoul, C.: Investigating the Local Atmospheric Response to a Realistic Shift in the Oyashio Sea Surface Temperature Front, Journal of Climate, 28, 1126–1147, https://doi.org/10.1175/JCLI-D-14-00285.1, 2015.

Ullrich, P.: SQuadGen: spherical quadrilateral grid generator, https://climate.ucdavis.edu/squadgen.php, 2014.

van Dalum, C. T., van de Berg, W. J., Gadde, S. N., van Tiggelen, M., van der Drift, T., van Meijgaard, E., van Ulft, L. H., and van den Broeke, M. R.: First results of the polar regional climate model RACMO2.4, The Cryosphere, 18, 4065–4088, https://doi.org/10.5194/tc-18-4065-2024, 2024.

van Kampenhout, L., Lenaerts, J. T. M., Lipscomb, W. H., Sacks, W. J., Lawrence, D. M., Slater, A. G., and van den Broeke, M. R.: Improving the Representation of Polar Snow and Firn in the Community Earth System Model, Journal of Advances in Modeling Earth Systems, 9, 2583–2600, https://doi.org/10.1002/2017MS000988, 2017.

van Kampenhout, L., Rhoades, A. M., Herrington, A. R., Zarzycki, C. M., Lenaerts, J. T. M., Sacks, W. J., and van den Broeke, M. R.: Regional grid refinement in an Earth system model: impacts on the simulated Greenland surface mass balance, The Cryosphere, 13, 1547–1564, https://doi.org/10.5194/tc-13-1547-2019, 2019.

Waling, A., Herrington, A., Duderstadt, K., Dibb, J., and Burakowski, E.: Using variable-resolution grids to model precipitation from atmospheric rivers around the Greenland ice sheet, Weather and Climate Dynamics, 5, 1117–1135, https://doi.org/10.5194/wcd-5-1117-2024, 2024.

Walsh, J. E., Ballinger, T. J., Euskirchen, E. S., Hanna, E., Mård, J., Overland, J. E., Tangen, H., and Vihma, T.: Extreme weather and climate events in northern areas: A review, Earth-Science Reviews, 209, https://doi.org/10.1016/j.earscirev.2020.103324, 2020.

Wehner, M. F., Reed, K. A., Li, F., Prabhat, Bacmeister, J., Chen, C. T., Paciorek, C., Gleckler, P. J., Sperber, K. R., Collins, W. D., Gettelman, A., and Jablonowski, C.: The effect of horizontal resolution on simulation quality in the Community Atmospheric Model, CAM5.1, Journal of Advances in Modeling Earth Systems, 6, https://doi.org/10.1002/2013MS000276, 2014.

Weijer, W., Cheng, W., Garuba, O. A., Hu, A., and Nadiga, B. T.: CMIP6 Models Predict Significant 21st Century Decline of the Atlantic Meridional Overturning Circulation, Geophysical Research Letters, 47, https://doi.org/10.1029/2019GL086075, 2020.

Weimer, M., Wilka, C., Kinnison, D. E., Garcia, R. R., Bacmeister, J. T., Alexander, M. J., Dörnbrack, A., and Solomon, S.: A Method for Estimating Global Subgrid-Scale Orographic Gravity-Wave Temperature Perturbations in Chemistry-Climate Models, Journal of Advances in Modeling Earth Systems, 15, https://doi.org/10.1029/2022MS003505, 2023.

Wijngaard, R. R., Herrington, A. R., Lipscomb, W. H., Leguy, G. R., and An, S. I.: Exploring the ability of the variable-resolution Community Earth System Model to simulate cryospheric-hydrological variables in High Mountain Asia, Cryosphere, 17, https://doi.org/10.5194/tc-17-3803-2023, 2023.

Wilks, D. S.: "The stippling shows statistically significant grid points": How research results are routinely overstated and overinterpreted, and what to do about it, Bulletin of the American Meteorological Society, 97, https://doi.org/10.1175/BAMS-D-15-00267.1, 2016.

Williams, R. S., Marshall, G. J., Levine, X., Graff, L. S., Handorf, D., Johnston, N. M., Karpechko, A. Y., Andrew, O. R., van de Berg, W. J., Wijngaard, R. R., and Mooney, P. A.: Future Antarctic Climate: Storylines of Midlatitude Jet Strengthening and Shift Emergent from CMIP6, Journal of Climate, 37, https://doi.org/10.1175/JCLI-D-23-0122.1, 2024.

Wills, R. C. J., Herrington, A. R., Simpson, I. R., and Battisti, D. S.: Resolving Weather Fronts Increases the Large-Scale Circulation Response to Gulf Stream SST Anomalies in Variable-Resolution CESM2 Simulations, Journal of Advances in Modeling Earth Systems, 16, https://doi.org/10.1029/2023MS004123, 2024.

Xu, Z., Chang, A., and Vittorio, A. D.: Evaluating and projecting of climate extremes using a variable-resolution global climate model (VR-CESM), Weather and Climate Extremes, 38, https://doi.org/10.1016/j.wace.2022.100496, 2022.

Yin, Z., Herrington, A. R., Datta, R., Subramanian, A. C., Lenaerts, J. T. M., and Gettelman, A.: Improved Understanding of Multicentury Greenland Ice Sheet Response to Strong Warming in the Coupled CESM2-CISM2 with Regional Grid Refinement, Journal of Advances in Modeling Earth Systems, https://doi.org/10.22541/au.170967825.56554188/v1, 2024.

Zampieri, L., Arduini, G., Holland, M., Keeley, S. P., Mogensen, K., Shupe, M. D., and Tietsche, S.: A Machine Learning Correction Model of the Winter Clear-Sky Temperature Bias over the Arctic Sea Ice in Atmospheric Reanalyses, Monthly Weather Review, 151, https://doi.org/10.1175/MWR-D-22-0130.1, 2023.

Zappa, G.: Regional Climate Impacts of Future Changes in the Mid–Latitude Atmospheric Circulation: a Storyline View, Current Climate Change Reports, 5, https://doi.org/10.1007/s40641-019-00146-7, 2019.

Zappa, G. and Shepherd, T. G.: Storylines of atmospheric circulation change for European regional climate impact assessment, Journal of Climate, 30, https://doi.org/10.1175/JCLI-D-16-0807.1, 2017.

Zarzycki, C. M. and Jablonowski, C.: A multidecadal simulation of Atlantic tropical cyclones using a variable-resolution global atmospheric general circulation model, Journal of Advances in Modeling Earth Systems, 6, 805–828, https://doi.org/10.1002/2014MS000352, 2014.

Zarzycki, C. M., Levy, M. N., Jablonowski, C., Overfelt, J. R., Taylor, M. A., and Ullrich, P. A.: Aquaplanet Experiments Using CAM's Variable-Resolution Dynamical Core, Journal of Climate, 27, 5481–5503, https://doi.org/10.1175/JCLI-D-14-00004.1, 2014.

Zhang, G. J. and McFarlane, N. A.: Sensitivity of climate simulations to the parameterization of cumulus convection in the canadian climate centre general circulation model, Atmosphere - Ocean, 33, https://doi.org/10.1080/07055900.1995.9649539, 1995.

Zhu, H., Zhang, J., Xu, Z., Vittorio, A. V. D., Xin, X., Xiao, C., and Li, Y.: Advantages of a variable-resolution global climate model in reproducing the seasonal evolution of East Asian summer monsoon, International Journal of Climatology, 43, https://doi.org/10.1002/joc.7796, 2023.

Zhu, J., Otto-Bliesner, B. L., Brady, E. C., Gettelman, A., Bacmeister, J. T., Neale, R. B., Poulsen, C. J., Shaw, J. K., McGraw, Z. S., and Kay, J. E.: LGM Paleoclimate Constraints Inform Cloud Parameterizations and Equilibrium Climate Sensitivity in CESM2, Journal of Advances in Modeling Earth Systems, 14, https://doi.org/10.1029/2021MS002776, 2022.

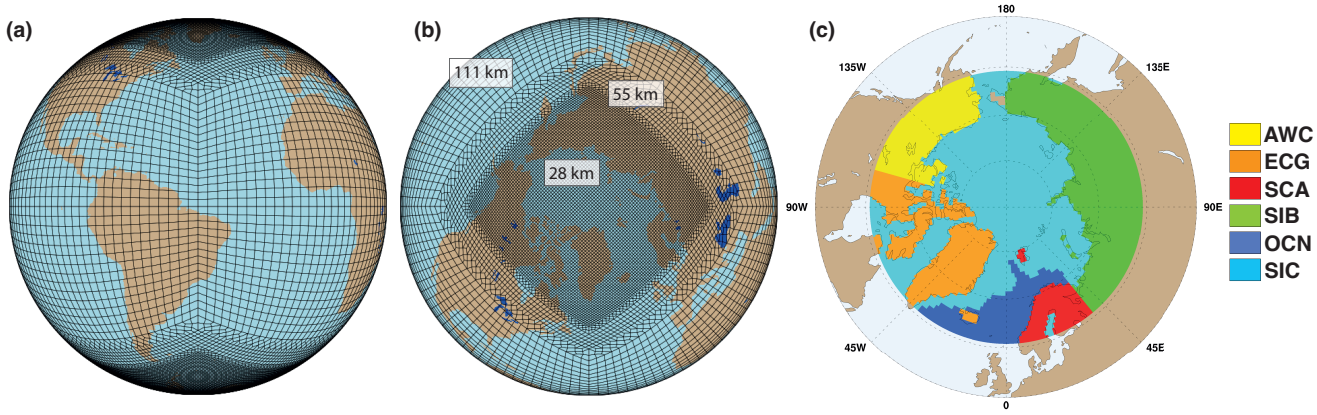

Figure 1. Schematic representation of the variable resolution (VR) spectral element grid (**a**) with regional grid refinements over the Arctic (**b**), and an overview of regions used for analysis throughout this study (**c**), which include four land regions AWC (Alaska and Western Canada; yellow), ECG (Eastern Canada and Greenland; orange), SCA (Scandinavia; red), and SIB (Siberia; green), a sea ice region (SIC; ~~diagonal pattern~~blue), and an ocean region (OCN; ~~crossed pattern~~cyan).

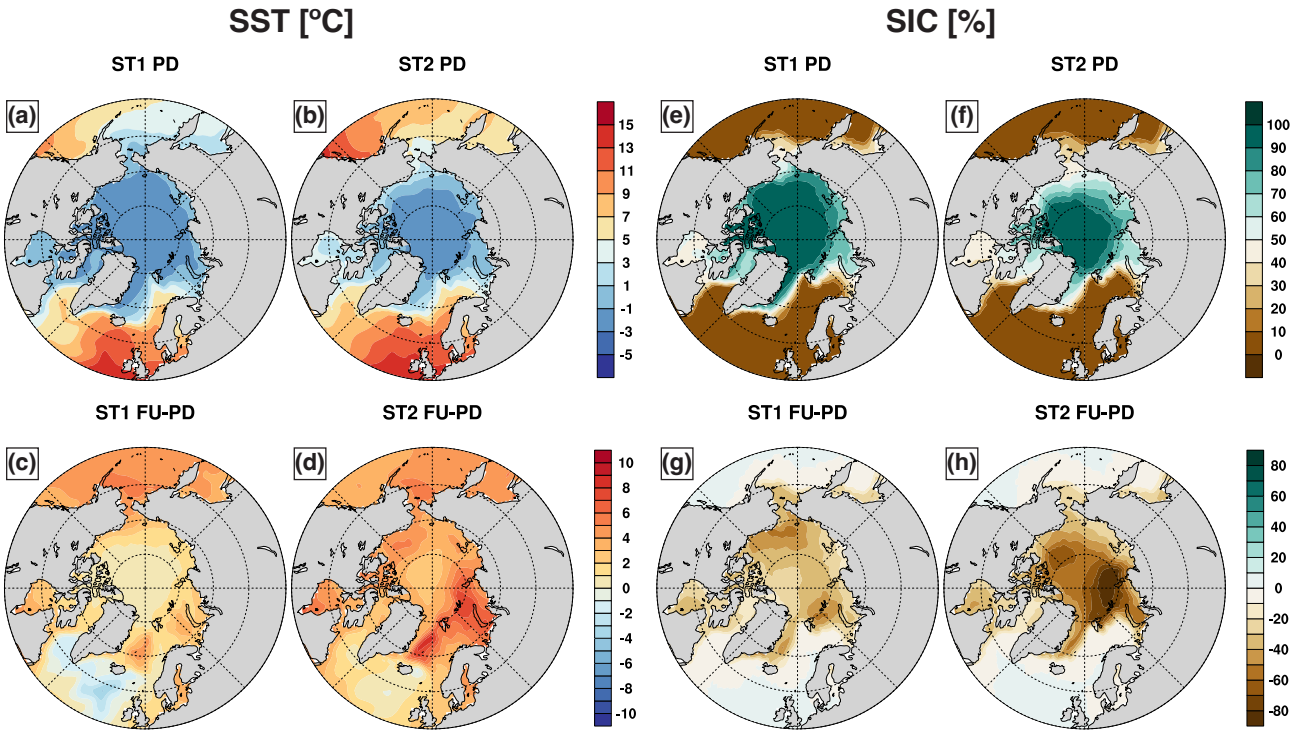

Figure 2. Present-day B (2005–2014) state (**a-b,e-f**) and future (2090–2099) changes (**c-d,g-h**) in northern hemisphere (poleward of 50°N) sea surface temperature (SST (°C); **a-d**) and sea ice cover (SIC (%); **e-h**), as ~~outputted~~ output by the NE30 grid for two different storylines: ST1 (**a,c,e,g**) and ST2 (**b,d,f,h**). The future changes are expressed as absolute differences relative to present-day.

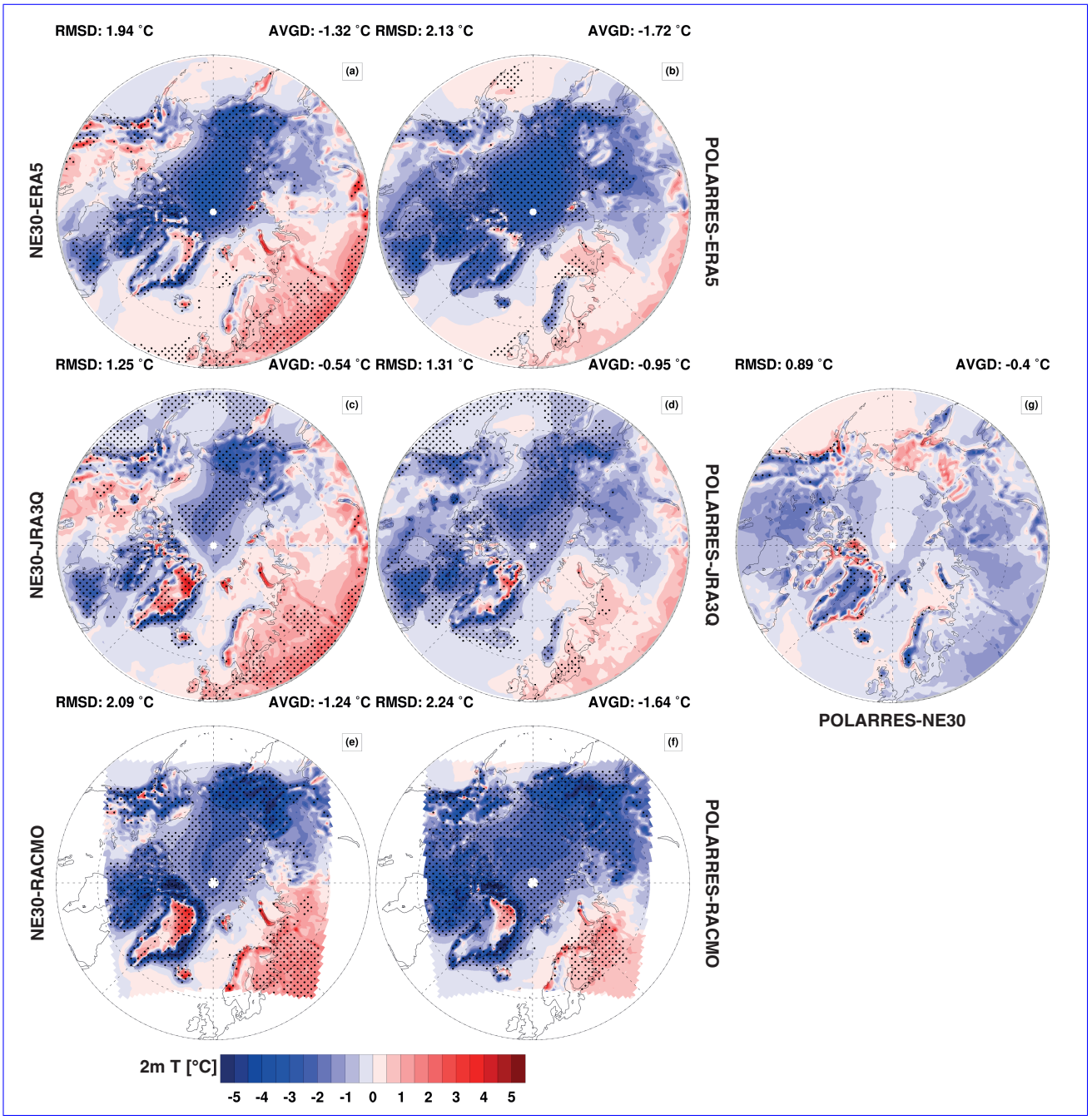

Figure 3. Present-day A (1985–2014) northern hemisphere (poleward of 50°N) annual mean 2m temperature (°C) differences between NE30 and ERA5 (**a**), JRA-3Q (**c**) or RACMO (**e**), and between POLARRES and ERA5 (**b**), JRA-3Q (**d**), or RACMO (**f**). The area-weighted ~~root-mean-square-error~~ root-mean-square-difference (~~RMSE~~RMSD) and ~~mean bias~~ average difference (AVGD) listed above the panels are calculated for the domain 60-90°N. The ~~dots represent~~ stippling represents the significance of temperature differences at the 95% confidence level.

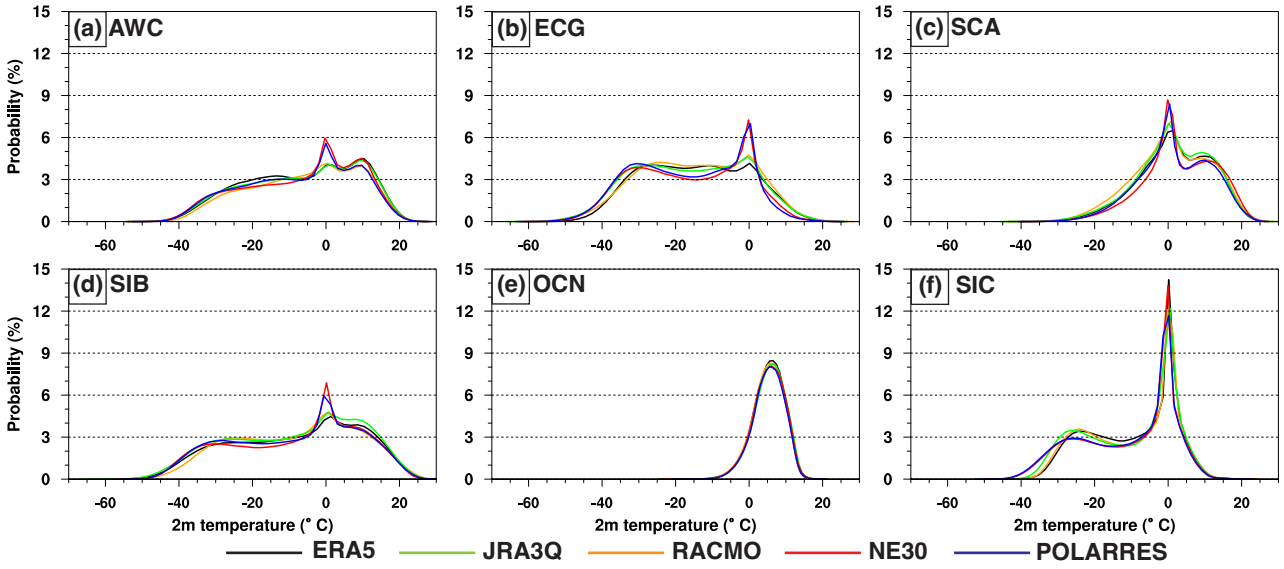

Figure 4. Present-day A (1985–2014) ~~annual~~ daily 2m temperature (°C) probability density functions (PDFs) for NE30 (red), POLARRES (blue), ERA5 (black), JRA-3Q (green), and RACMO (orange). The PDFs are calculated on a NE30 grid for the six different regions as shown in Fig. 1.

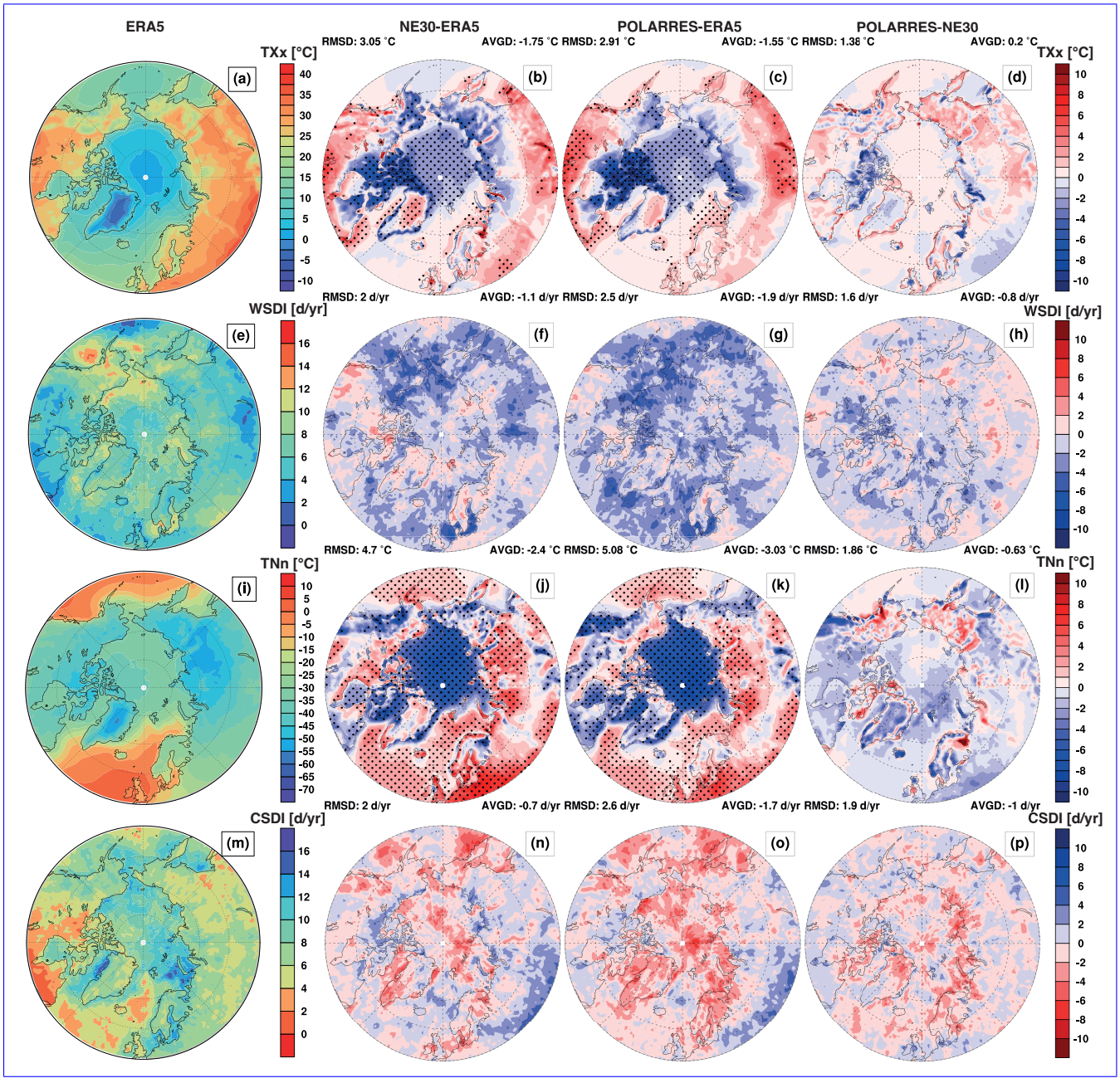

Figure 5. Present-day A (1985–2014) northern hemisphere (poleward of 50°N) annual maximum ~~of TX~~ temperature (TXx (°C); **a-d**), warm spell duration index (WSDI ($\mathrm{d\,yr^{-1}}$); **e-h**), annual minimum ~~of TN~~ temperature (TNn (°C); **j-l**), and cold spell duration index (CSDI ($\mathrm{d\,yr^{-1}}$); **m-p**) state (1^{st} column) and differences between NE30 and ERA5 (2^{nd} column), POLARRES and ERA5 (3^{rd} column), and POLARRES and NE30 (4^{th} column). The area-weighted root-mean-square-difference (~~RMSE~~RMSD) and ~~mean bias~~ average difference (AVGD) listed above the 2^{nd}-4^{th} column panels are calculated for the domain 60-90°N. The stippling represents the significance of differences at the 95% confidence level.

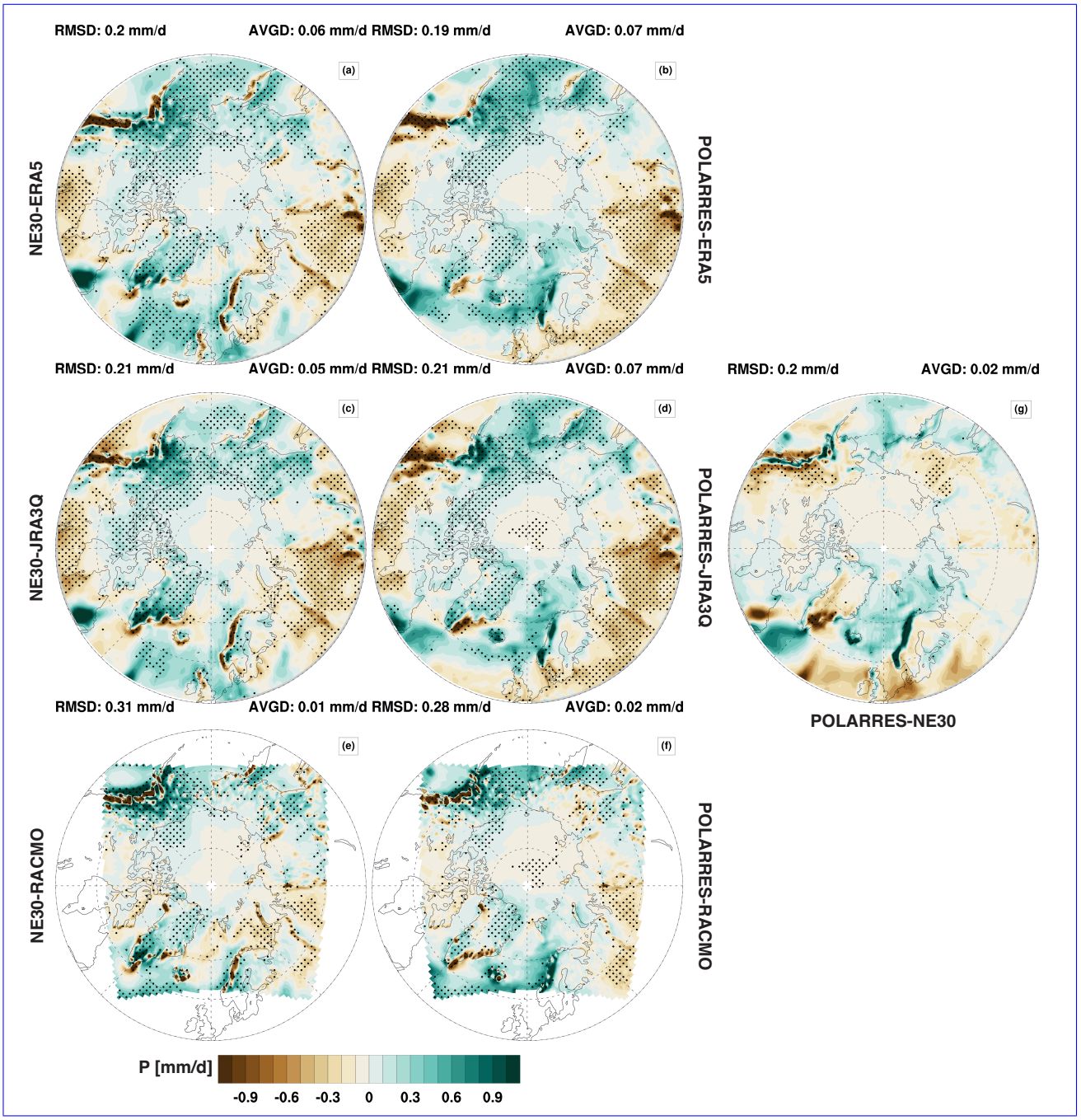

Figure 6. Same as Fig. 3, but for annual mean precipitation ($\mathrm{mm\,d^{-1}}$)

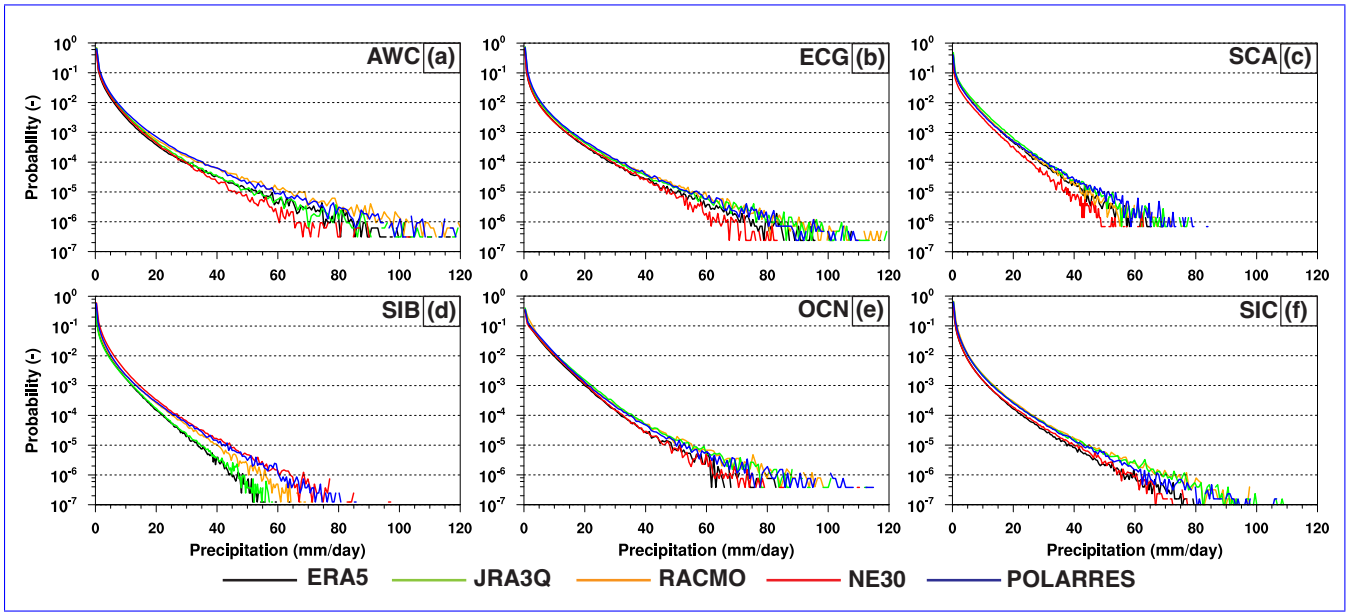

Figure 7. Same as Fig. 4, but for daily precipitation rates ($\mathrm{mm\,d^{-1}}$)

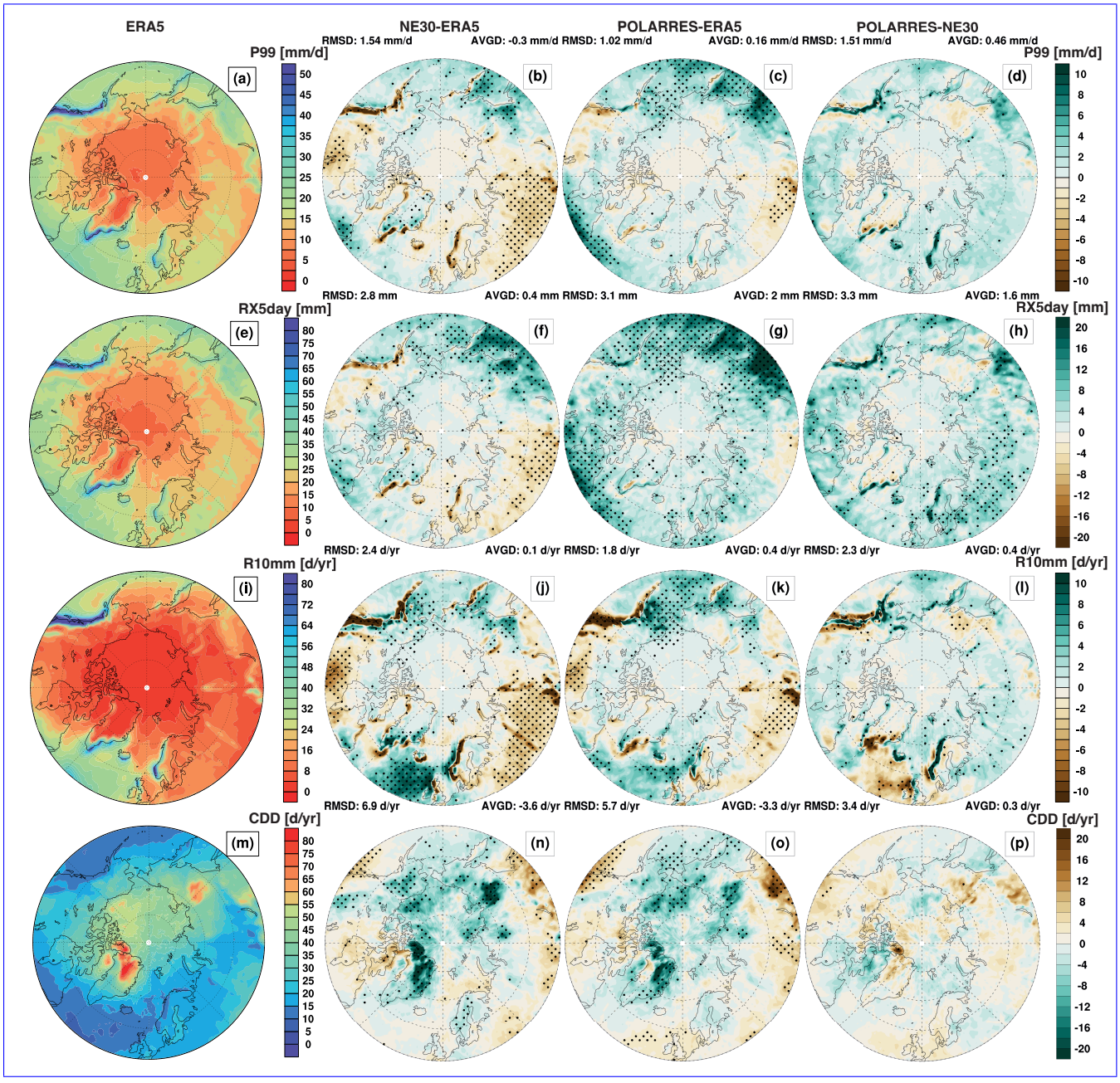

Figure 8. Same as Fig.5, but for the 99^{th} percentile of daily precipitation (P99 ($\mathrm{mm\,d^{-1}}$); **a-d**), highest 5-day precipitation (RX5day (mm); **e-h**), heavy precipitation days (R10mm ($\mathrm{d\,yr^{-1}}$); **i-l**), and consecutive dry days (CDD (~~d~~$\mathrm{d\,yr^{-1}}$); **m-p**)

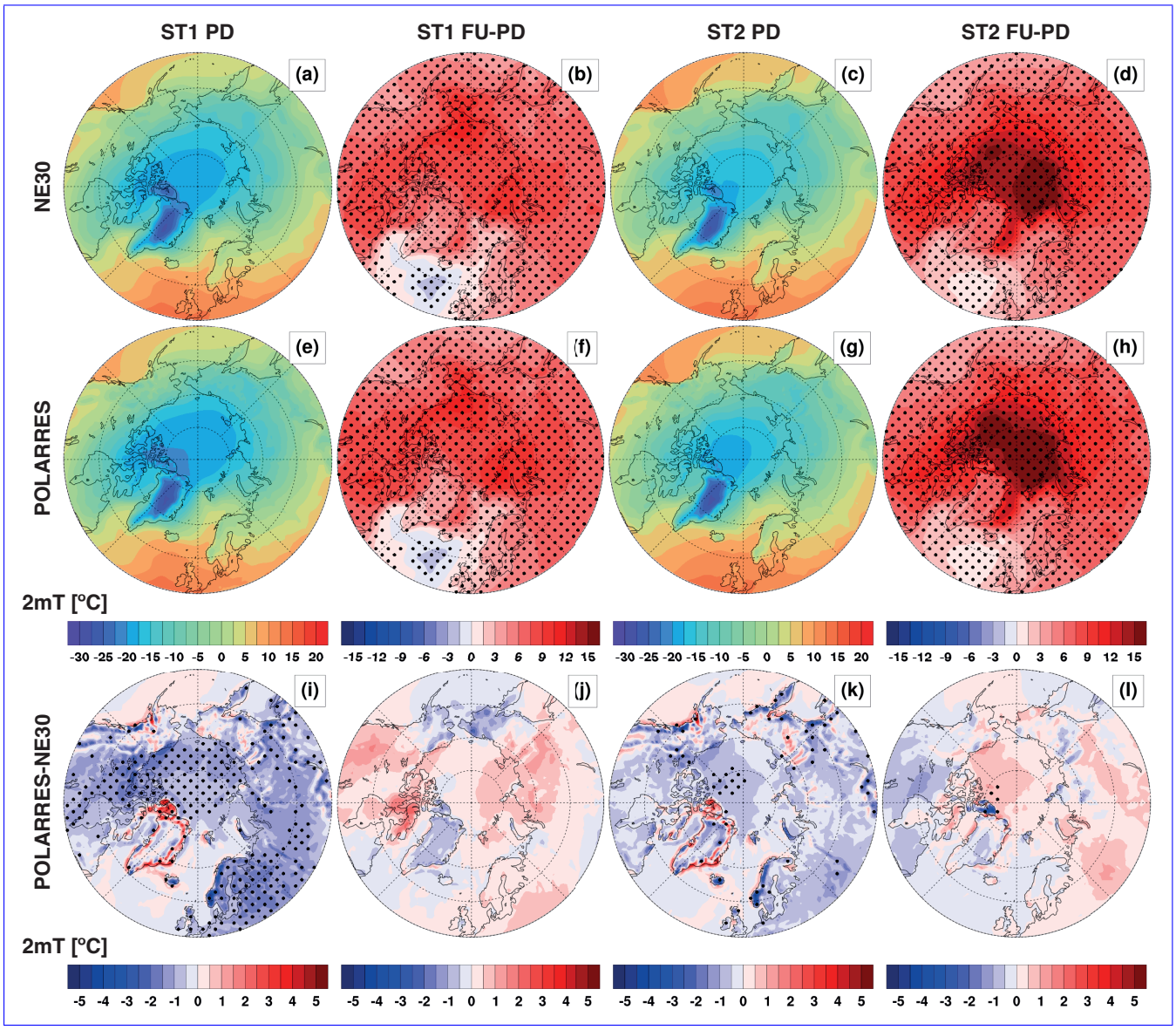

Figure 9. Northern hemisphere (poleward of 50°N) present-day B (2005–2014) annual 2m temperature and projected future (2090–2099) changes in annual 2m temperature (°C) for NE30 (**a-d**) and POLARRES (**e-h**) in ST1 (1^{st} and 2^{nd} columns) and ST2 (3^{rd} and 4^{th} columns). The future changes are expressed as absolute differences relative to present-day. The differences between NE30 and POLARRES are shown in the bottom row (**i-l**). The stippling represents the significance of differences at the 95% confidence level.

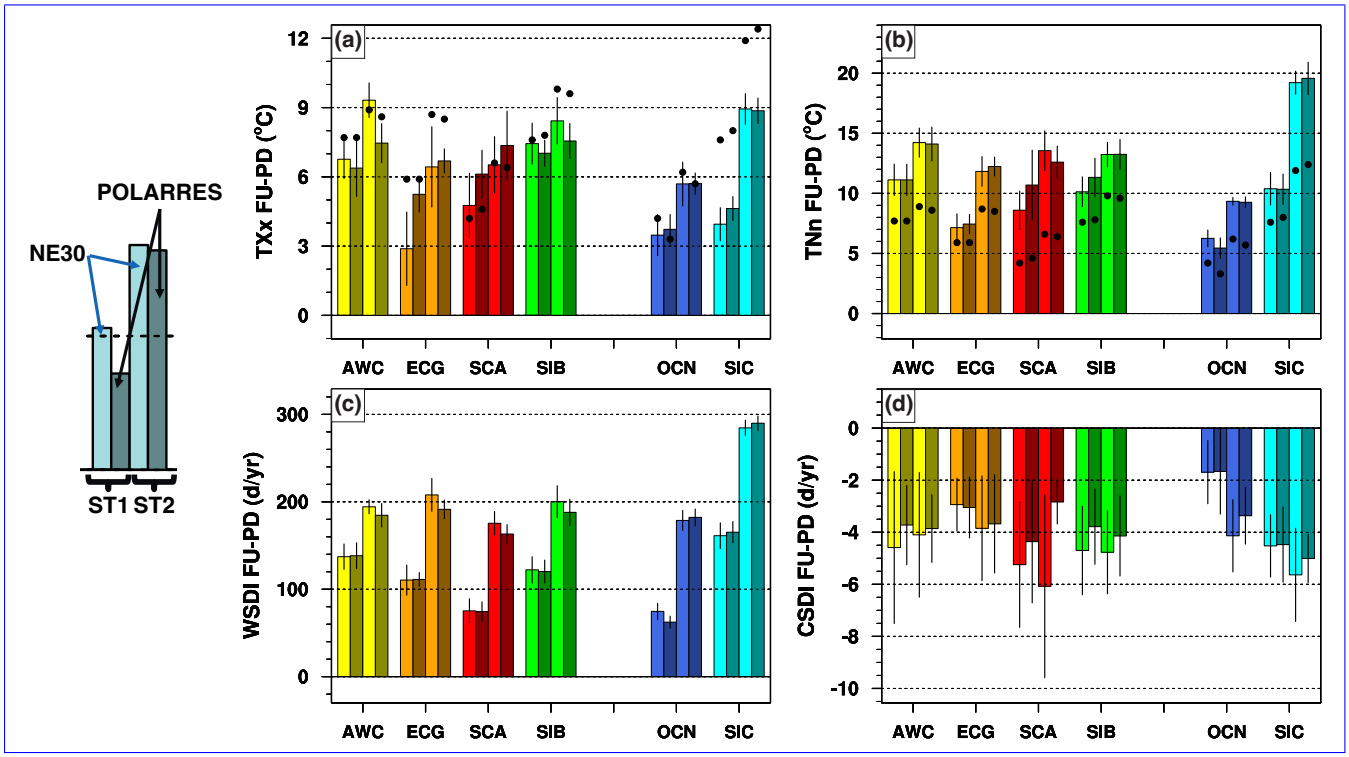

Figure 10. Projected future (2090–2099) changes in annual maximum temperature (TXx (°C); **a**), annual minimum temperature (TNn (°C; **b**), warm spell duration index (WSDI ($\mathrm{d\,yr^{-1}}$); **c**), and cold spell duration index (CSDI ($\mathrm{d\,yr^{-1}}$); **d**) for six different regions as shown in Fig. 1. The bars show the projected changes in NE30 (1^{st} and 3^{rd} bars) and POLARRES (2^{nd} and 4^{the} bars) for ST1 (1^{st} and 2^{nd} bars) and ST2 (3^{rd} and 4^{th} bars).The black dots denote the projected changes in annual mean near-surface temperature (°C). The error bars (black) represent the 95% confidence interval.

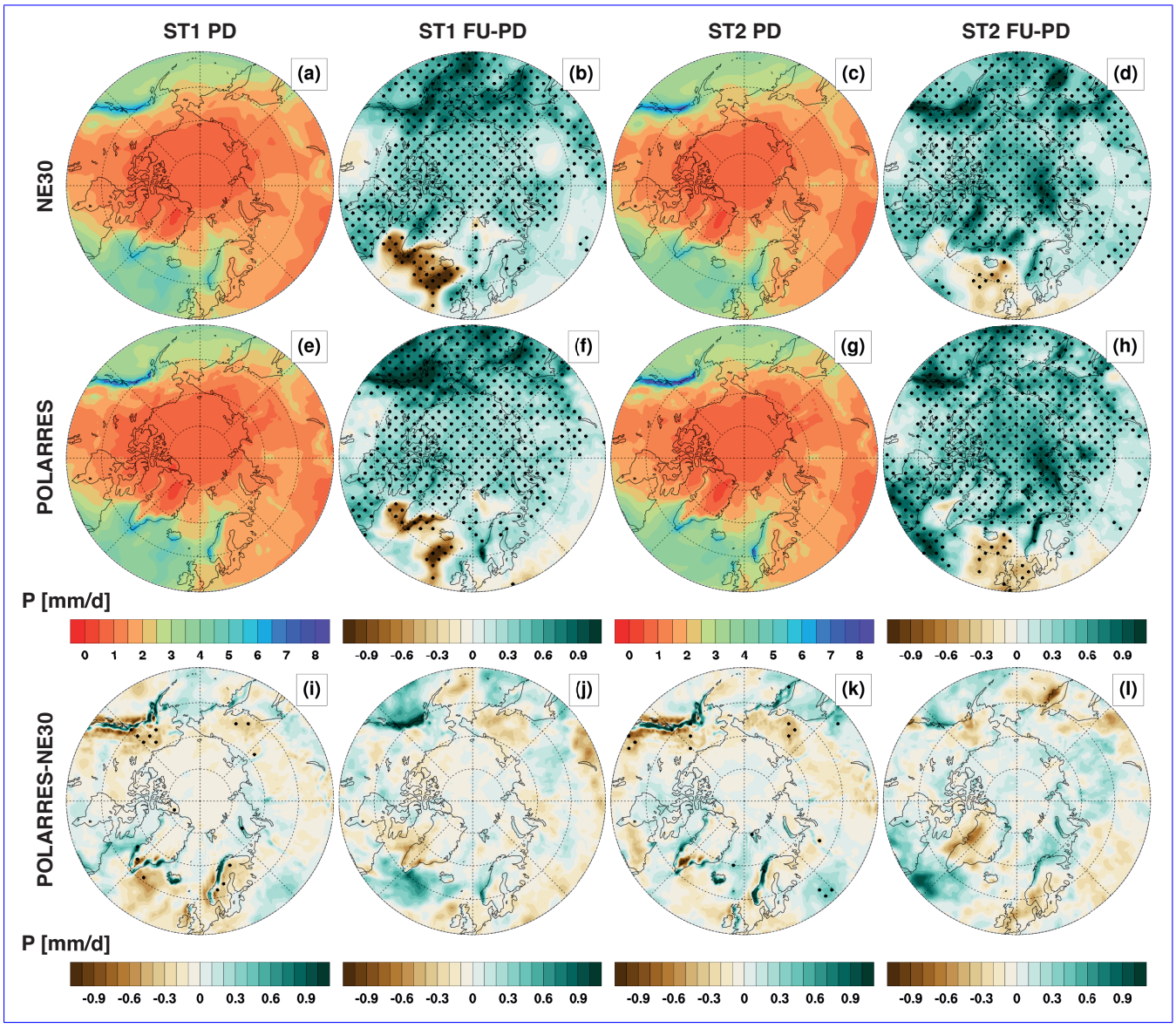

Figure 11. Same as Fig. 9, but for annual mean precipitation ($\mathrm{mm\,d^{-1}}$)

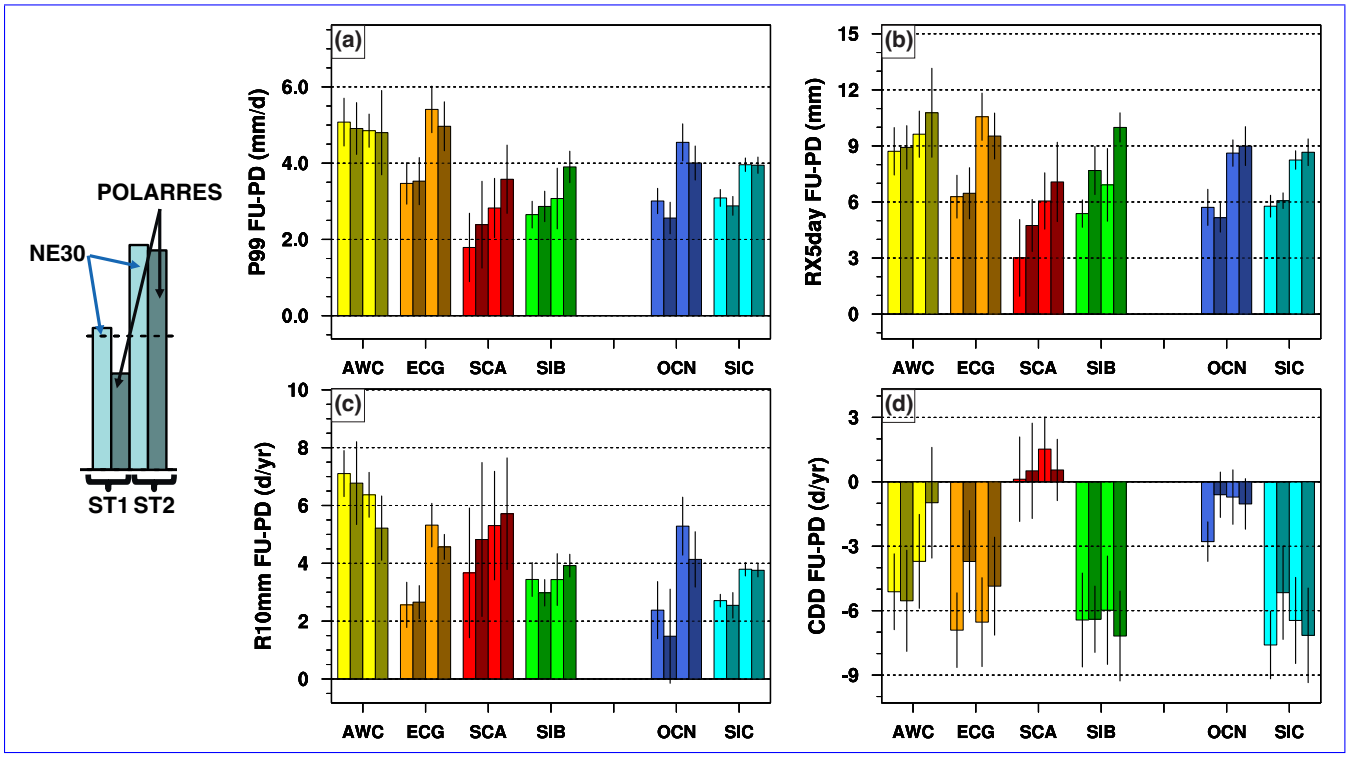

Figure 12. Same as Fig. 10, but for the 99^{th} percentile of daily precipitation (P99 ($\mathrm{mm\,d^{-1}}$); **a**), highest 5-day precipitation (RX5day (mm); **b**), heavy precipitation days (R10mm ($\mathrm{d\,yr^{-1}}$); **c**), and consecutive dry days (CDD (~~d~~$\mathrm{d\,yr^{-1}}$); **d**)

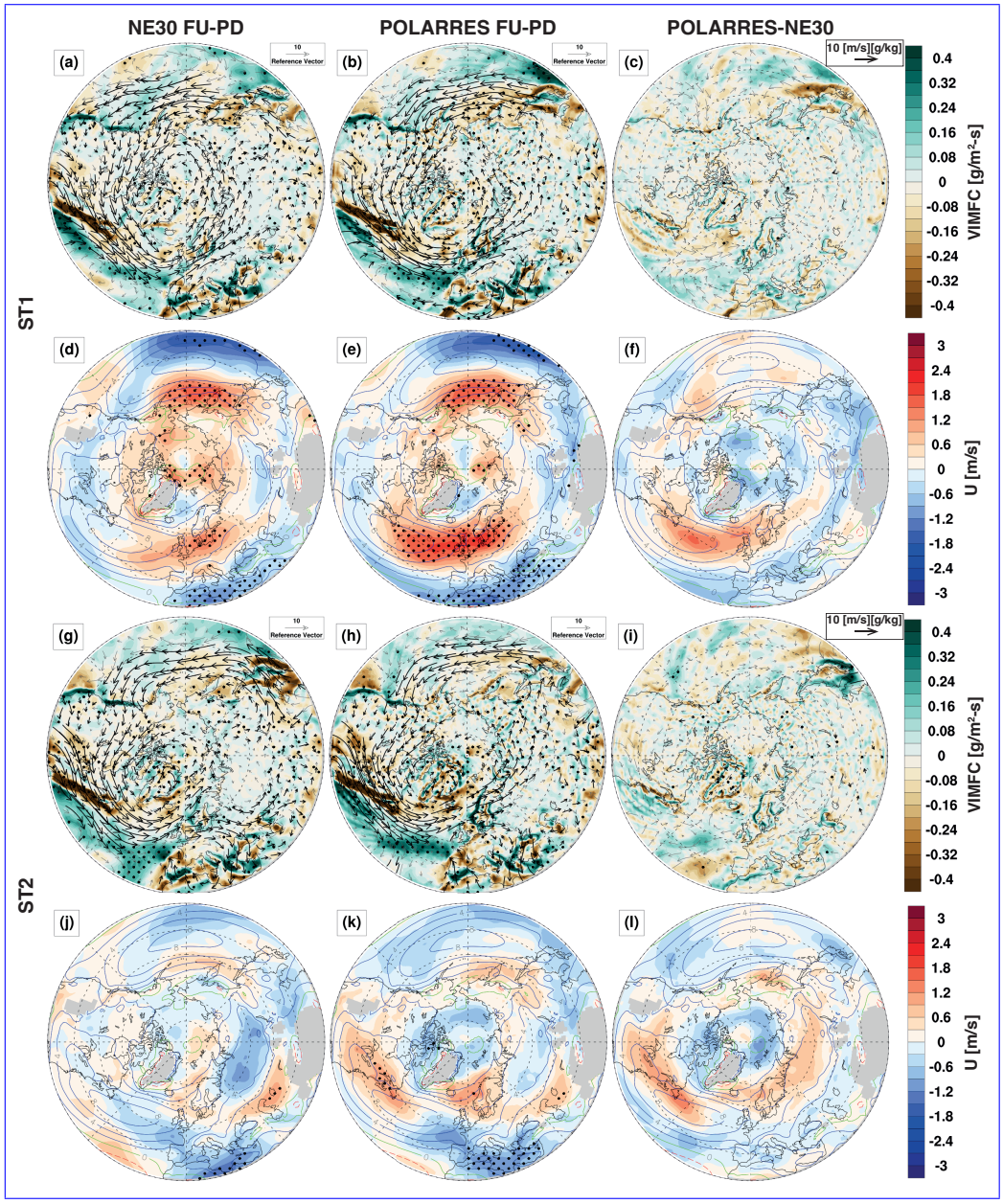

Figure 13. Projected future (2090–2099) changes in northern hemisphere (poleward of 30°N) vertically-integrated moisture flux convergence (VIMFC $(\mathrm{g/m2-s})$; shading) and 850 hPa moisture flux $((\mathrm{m/s})/\ (\mathrm{g/kg})$; vectors) (**a-c,g-i**), and 850 hPa zonal wind (**d-f,j-l**; $\mathrm{m/s}$; shading), as ~~outputted~~ output by NE30 (1^{st} column) and POLARRES (2^{nd} column) for ST1 (**a-f**) and ST2 (**g-l**). The differences between NE30 and POLARRES are shown in the 3^{rd} column. The zonal wind contours denote the climatological mean of the present-day zonal wind at 850 hPa, where red stands for positive (eastward) and blue stands for negative (westward). The ~~dots~~ stippling and the vectors (black) represent the significance of future changes at the 95% confidence level. The grey vectors represent the changes in moisture flux that are insignificant.

Table 1. Metrics of daily temperature and precipitation extremes used in this study

Extreme Metric	Index	Description
Temperature extremes		
Annual maximum temperature	TXx (°C)	Annual maximum of daily maximum temperature (TX)
Annual minimum temperature	TNx (°C)	Annual minimum of daily minimum temperature (TN)
Warm spell duration	WSDI ($\mathrm{d\,yr^{-1}}$)	Warm spell duration index, defined as the annual number of days in intervals of at least 6 consecutive days where TX > TX90, where TX90 is the 90^{th} percentile of daily maximum temperature, calculated for each calendar day using a running window of 15 days
Cold spell duration	CSDI ($\mathrm{d\,yr^{-1}}$)	Cold spell duration index, defined as the annual number of days in intervals of at least 6 consecutive days where TN < TN10, where TN10 is the 10^{th} percentile of daily minimum temperature, calculated for each calendar day using a running window of 15 days
Precipitation extremes		
99^{th} percentile of daily precipitation	P99 ($\mathrm{mm\,d^{-1}}$)	99^{th} percentile of daily precipitation sums
Highest 5-day precipitation sum	RX5day (mm)	Highest amount of precipitation over an interval of 5 days
Number of heavy precipitation days	R10mm ($\mathrm{d\,yr^{-1}}$)	Number of days with heavy precipitation, defined as daily precipitation that is equal to or higher than $10\,\mathrm{mm\,d^{-1}}$
~~CDD (d~~Number of consecutive dry days	CDD ($\mathrm{d\,yr^{-1}}$)	Greatest number of consecutive days with daily precipitation less than $1\,\mathrm{mm\,d^{-1}}$