

<https://doi.org/10.1038/s43247-024-01276-z>

Cooler Arctic surface temperatures simulated by climate models are closer to satellite-based data than the ERA5 reanalysis

Check for updates

Tian Tian¹ ✉, Shuting Yang¹, Jacob Lorentsen Høyer¹, Pia Nielsen-Englyst^{1,2} & Suman Singha¹

Many climate models simulate near-surface air temperatures that are too low in the Arctic compared to the observation-based ERA5 reanalysis data, a bias that was noted in the Inter-governmental Panel on Climate Change Sixth Assessment Report (IPCC AR6). Here we present a high-resolution, satellite-derived dataset of near-surface air temperatures for the Arctic sea-ice region (1982–2020). We use it as a benchmark to reevaluate climate reanalyses and model simulations in CMIP6 (Coupled Model Intercomparison Project 6). We find that the CMIP6 simulations in the central Arctic, with generally thicker ice and snow, align well with satellite observations, with an annual mean bias of less than ± 1 °C over sea ice. By contrast, climate reanalyses like ERA5 exhibit widespread warm biases exceeding 2 °C in the same region. We conclude that reliance on ERA5 reanalysis as a reference may have led to an underestimation of climate model reliability in the Arctic region.

Several recent studies including the IPCC AR6¹ show that climate models generally simulate a too cold Arctic, indicated as a negative bias of near-surface air temperature (often measured at 2 m height, hereafter referred to as T2m). It is known as one of the long-standing biases in climate models in the past phases of the Coupled Model Intercomparison Project (CMIP)^{1–3}. Figure 3.3 in the IPCC AR6 further confirms the robust signal of Arctic T2m bias in recent decades (1995–2014) in CMIP5, CMIP6, and HighResMIP, suggesting limited improvement with advanced physics and increased horizontal resolution in climate models.

The documented climate model biases in these studies are often derived from comparisons with the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis version 5 (ERA5), one of the most advanced atmospheric reanalysis datasets⁴. ERA5 is widely applied for the Arctic assessment⁵ rather than the use of observational data sets, because the latter are typically available as anomalies relative to a reference period^{6,7}. Since the Arctic, particularly over sea ice, is data sparse, global reanalyses in this region are only weakly constrained by observations and heavily rely on model formulation with simplified physical processes tied to the radiation budget, resulting in considerable uncertainty^{6–8}. However, increasing evidence from in situ observations gathered during recent campaigns reveals that both the sea ice surface temperature (IST) and T2m in reanalyses have substantial warm biases of 5 °C or more under cold clear-sky conditions^{9–12}.

This discrepancy becomes apparent when the observed thickness of the sea ice and the overlying snow layer together exceeds the prescribed values (usually 1.5 or 2 m) in reanalysis and forecast models^{10,13,14}. These biases are known to be mainly attributed to the insufficient insulating effect of thick snow on the ice surface, resulting in an overestimated conductive heat flux from the warm ocean underneath the ice and snow layer. This issue in the Arctic has long been acknowledged in the numerical weather prediction community, and thus mitigation strategies have been explored in weather forecasts and regional reanalysis models. These strategies include a machine learning post-processing model¹³, improved representation of sea ice and snow physics over sea ice¹⁵, and even integration of atmosphere-ocean-sea ice fully coupled systems^{16,17}. Despite showing improvements in IST estimates, these strategies have not yet been implemented in existing reanalysis products covering the pan-Arctic region and recent decades^{13,15}. Additionally, the satellite IST product is currently not being assimilated by existing reanalysis products. Consequently, T2m in these reanalysis products still lacks an accurate representation of the Arctic surface state, despite its importance as an essential climate variable for Earth's climate characterisation¹. Furthermore, the seasonal and decadal variations of the T2m bias within the existing global atmospheric reanalysis and the resulting implications for climate assessment remain unclear¹³. When using these global atmospheric reanalysis products (too warm over sea ice) for model

¹National Center for Climate Research, Danish Meteorological Institute, Copenhagen 2100, Denmark. ²DTU-Space, Technical University of Denmark, Lyngby 2800, Denmark. ✉e-mail: tian@dmi.dk

validation, this bias can lead to inaccurate conclusions, such as presuming that climate models have a ‘cold temperature’ bias^{1–3}.

Hence, accurate reference data with full spatial and temporal coverage are urgently needed to benchmark climate models on the Arctic surface. It is also crucial for assessing the Arctic climate state, as many essential elements in the Arctic (e.g., sea ice and permafrost melting points, ecosystems, and their possible tipping points, etc.) respond to specific temperature thresholds. We recently developed a new high spatial resolution observational dataset of T2m (hereafter referred to as $T_{2m}^{(SAT)}$), converted from satellite-derived ISTs covering the period 1982–2021¹⁸ using the existing empirical relationships¹⁹. The dataset offers a new possibility of benchmarking climate models in the Arctic, to the best of our knowledge. In this study, we re-evaluate the ERA5 and CMIP6 model ensemble over the Arctic sea ice using this new satellite T2m product as a reference. Our findings demonstrate that relying on ERA5 leads to the erroneous conclusion of persistent cold bias in the Arctic in climate models. On the contrary, we show that the performance of the CMIP6 models in the central Arctic (with generally thicker ice and snow) aligns well with satellite observations. This study stands as a crucial reference, complementing existing massive amount of publications that rely on global reanalysis datasets, including ERA5, to benchmark CMIP5 and CMIP6 models in the Arctic.

Results and discussion

Improved T2m representation over the Arctic sea ice

To establish the superiority of the new $T_{2m}^{(SAT)}$ over ERA5 ($T_{2m}^{(ERA5)}$) on Arctic sea ice, we first validate both T2m datasets against various independent in situ T2m from pointwise ground measurements ($T_{2m}^{(insitu)}$), with positions shown in Supplementary Fig. S1. Table 1 shows the validation statistics including the 95% confidence intervals. The mean differences between $T_{2m}^{(SAT)}$ and $T_{2m}^{(insitu)}$ range from -0.45 °C to 0.65 °C, significantly smaller than those of $T_{2m}^{(ERA5)}$ ranging from 1.73 to 3.73 °C. Similar to ref. 18, the validation results against the North Pole drifting ice stations (which span the longest time period) meet the requirements of the Global Climate Observing System (GCOS) in terms of stability²⁰ with a trend of -0.09 °C per decade. The long-term stability and superior validation results compared to ERA5 suggest that satellite-derived T2m can be used to evaluate climate models in the Arctic. This dataset, capable of providing continuous coverage and a spatial distribution of long-term mean and variability, proves particularly useful for supplementing the sparse in situ network.

Better mean-state in CMIP6 models

Figure 1a shows the $T_{2m}^{(SAT)}$ climatology over the regions with sea ice concentrations (SIC) above 15%. Using $T_{2m}^{(SAT)}$ as a reference, ERA5 exhibits a wide-spread warm bias of 2 °C or more (Fig. 1b) for areas where the SIC is typically above 70% (compacted by the red line in Fig. 1a, and hereafter denoted as SIC70). The bias is markedly greater in winter when it can reach 6–10 °C (Supplementary Fig. S2), in agreement with previous assessments^{9–13} (and other reanalyses in Supplementary Fig. S3). In contrast,

the CMIP6 historical ensemble of 47 models (Supplementary Table S1 and Supplementary Fig. S4) simulates surface temperature ($T_{2m}^{(CMIP6)}$) remarkably well in this region, with a mean difference within ± 1 °C (Fig. 1c), which falls within the range of observational uncertainties (see Methods). This contradicts previous assessments that the Arctic is too cold in CMIP6 historical simulations^{1,3}. It’s worth noting that the results for a longer period (1982–2014) than the IPCC AR6 assessment period of 1995–2014 in Fig. 1 do not alter the conclusion (Supplementary Fig. S5).

For the marginal ice zone (MIZ, defined as outside of SIC70), prominent cold biases are observed for both ERA5 and CMIP6 (Fig. 1b, c). Relative to CMIP6, ERA5 generally shows better agreement with satellite-derived T2m, consistent with the previous assessment⁵. The cold bias at the edge of sea ice can be attributed to the following factors: 1) in the MIZ, the satellite-derived surface temperatures are a mixture of sea surface temperature (SST) and IST, thus with larger uncertainties¹⁸; 2) differences in the ice edge locations between the SIC field in the satellite-derived T2m data set and ERA5; 3) too low modelled conductive heat fluxes from the warm ocean over very thin snow-covered ice due to a prescribed sea ice depth of 1.5 m applied to all grid cells with sea ice in ERA5. CMIP6 exhibits a large cold bias across the entire North Atlantic MIZ, influenced by regional disparities between models^{3,21}. This bias is intensified by considering only models with SIC above 15% in the calculation of the mean difference (Eqs. (1)–(3) in Methods). Furthermore, the models’ sharp transition between ice and the open ocean, particularly in winter²², may also contribute to the cold bias. As suggested in the IPCC AR6, improving the resolution of ocean models may reduce this persistent systematic bias in the North Atlantic^{1–3}.

Seasonal and decadal variations of temperature bias

The CMIP6 ensemble mean $T_{2m}^{(CMIP6)}$ and satellite observations $T_{2m}^{(SAT)}$ show comparable annual mean temperatures over the SIC70 area compared to ERA5 $T_{2m}^{(ERA5)}$ (Fig. 2c, d). ERA5 consistently shows a positive mean difference of 1.89 °C against satellite observations for the entire 1982–2020 period, while the average difference for the CMIP6 ensemble mean is close to 0 °C over 1982–2014 (Fig. 2a, b). The shaded area around the ensemble mean represents the model spread, quantified by one standard deviation, mostly below ERA5’s positive difference. CMIP6 models exhibit a cold bias with respect to ERA5, but this is only an artefact due to the warm bias of ERA5 with respect to an independent dataset. This highlights the limitations of using global reanalysis like ERA5 for the evaluation of surface variables over Arctic sea ice.

In Fig. 3, the seasonal cycle of decadal mean differences in the SIC70 area further illustrates that ERA5 consistently shows larger biases compared to CMIP6. During the last three decades, the consecutive decadal mean differences are 2.24 °C, 1.81 °C, and 1.75 °C for ERA5, 0.40 °C, 0.06 °C, and -0.38 °C for CMIP6, and 1.84 °C, 1.76 °C and 2.12 °C between ERA5 and CMIP6 (Fig. 2c). The most notable decadal changes in temperature are observed during winter (December–March) in both ERA5 and CMIP6. The winter warm bias (DJFM) in ERA5 reduced from

Table 1 | Validation statistics of the daily surface air temperature datasets against in situ measurements over the Arctic sea ice

Observation type	N	$T_{2m}^{(SAT)} - T_{2m}^{(insitu)}$			N	$T_{2m}^{(ERA5)} - T_{2m}^{(insitu)}$		
		MD	STD	RMSE		MD	STD	RMSE
NP drifting ice stations	7663	0.65	3.19	3.26	7832	3.73	3.22	4.93
95% CI		± 0.07	3.15–3.24			± 0.07	3.18–3.26	
ECMWF drifting buoys	55290	-0.45	3.33	3.36	56160	1.73	3.34	3.76
95% CI		± 0.03	3.31–3.35			± 0.03	3.32–3.36	
CRREL drifting buoys	22974	-0.17	3.33	3.34	22794	1.80	3.07	3.56
95% CI		± 0.04	3.31–3.36			± 0.04	3.05–3.09	

As Table 2 in ref. 18, validation was performed for observation $T_{2m}^{(SAT)}$ and reanalysis $T_{2m}^{(ERA5)}$, using in situ measurements $T_{2m}^{(insitu)}$ from 14 Russian North Pole (NP) drifting ice stations, 116 drifting buoys from the European Centre for Medium-Range Weather Forecasts (ECMWF), and 96 drifting buoys from the U.S. Army Cold Regions Research Engineering Laboratory (CRREL). The table shows the number of observations (N), the mean difference (MD), standard deviation of the difference (STD) and the root mean squared error (RMSE). The corresponding 95% confidence intervals (CI) for the mean differences and the standard deviation of the difference are estimated²⁹ and listed. The spatial and temporal coverage of the in situ measurements used here is illustrated in Supplementary Fig. S1. Units: °C.

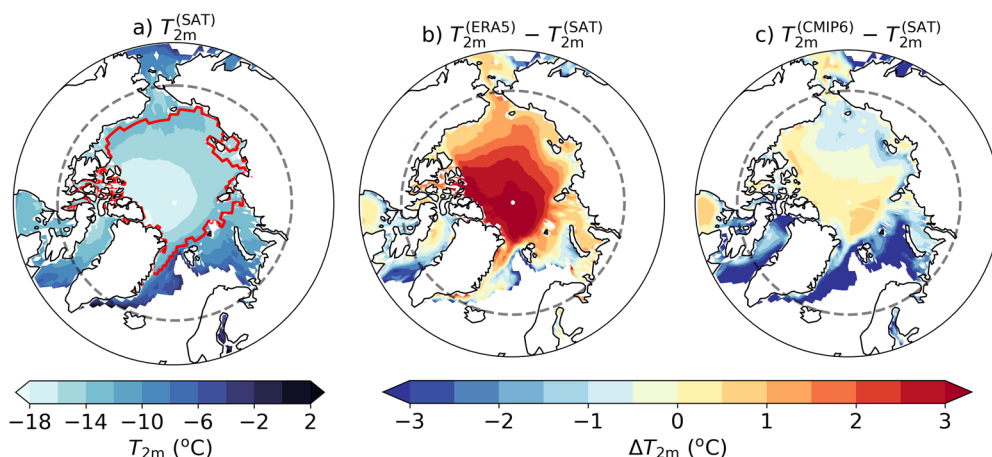


Fig. 1 | Observed climatology of surface air temperature, along with bias in ERA5 and the CMIP6 ensemble mean over the Arctic sea ice. a The 20-year mean satellite-derived $T_{2m}^{(SAT)}$ over sea ice (SIC > 15%) for the period 1995–2014 (see Methods). The climatological mean difference of surface air temperature from (b) ERA5 $T_{2m}^{(ERA5)}$ and (c) the CMIP6 ensemble mean $T_{2m}^{(CMIP6)}$ versus $T_{2m}^{(SAT)}$ for the same period. The maps are bounded at 58°N with the dashed line marking 66.5°N. The red line in a indicates the observed SIC ≥ 70% (SIC70) averaged over 1995–2014. Units: °C.

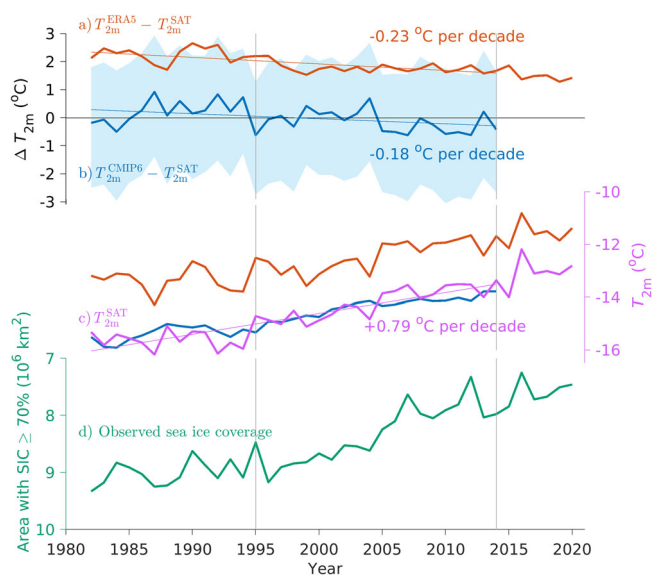


Fig. 2 | Evolution of annual mean surface air temperatures and their differences relative to satellite-derived observation over the central Arctic sea ice. a Surface air temperature difference for ERA5 reanalysis [$T_{2m}^{(ERA5)} - T_{2m}^{(SAT)}$] in red. b Surface air temperature difference for the CMIP6 ensemble mean [$T_{2m}^{(CMIP6)} - T_{2m}^{(SAT)}$] in thick blue line, with model spread (shaded area) calculated as one standard deviation from the mean. c Satellite-derived observations of surface air temperature $T_{2m}^{(SAT)}$ in purple, together with $T_{2m}^{(ERA5)}$ and $T_{2m}^{(CMIP6)}$ in red and blue, respectively. All temperatures (a–c) are averaged over the sea ice area with observed SIC ≥ 70% (SIC70) in units of °C. We also show (d) annual mean total sea ice area for SIC70 (green, on an inverted y-axis) in units of km². Thin colour lines in a–c represent the respective linear trends (°C per decade) calculated for the common period 1982–2014 (see Methods). More details of the temperature time series are provided in Supplementary Fig. S6a.

2.82 °C (1985–1994) to 1.98–1.99 °C (1995–2014). These multi-decadal variations align with previous studies that large warm biases in reanalysis surface temperature on sea ice often correspond to cold temperatures in winter and in regions with thick ice, thick snow, or a combination of both states (as seen in ERA5 in Fig. 2 of ref. 13). As for CMIP6, there is an increase in winter cold bias (DJFM) from −0.50 °C and −0.78 °C over the first two decades (1985–2004) to −1.81 °C (2005–2014).

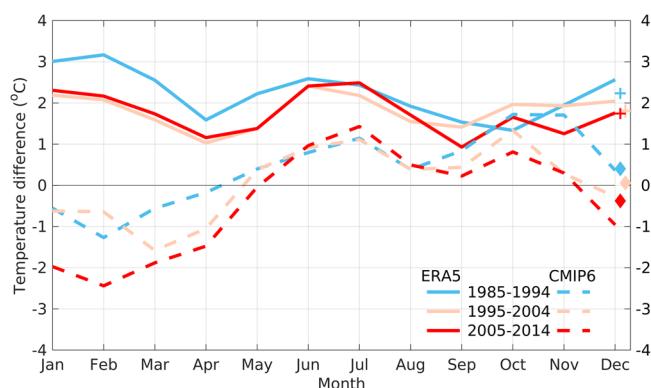


Fig. 3 | Annual cycle of decadal mean surface air temperature differences relative to satellite-derived data over the central Arctic sea ice. It shows the monthly temperature difference averaged over sea ice areas with observed SIC ≥ 70% (SIC70, north of 66.5°N) for the three recent decades in ERA5 (solid lines) and the CMIP6 ensemble mean (dashed lines). Symbols next to the right y-axis indicate the annual mean for each period. The time series of annual and seasonal mean temperatures are provided in Supplementary Fig. S6a–c and the maps of mean difference in winter are provided in Supplementary Fig. S2. Units: °C.

The recent increase of the winter cold bias seems a result from the underestimated temperature rise over sea ice (2005–2014 in Supplementary Figs. S2 and S6b). This is probably due to the overestimated Arctic sea ice mass in winter by climate models, as highlighted in ref. 23 (see their Fig. 2). Consequently, most models fail to simulate the steeper decline in sea ice area since the mid-2000s (Fig. 2d) and a plausible evolution of Arctic warming at the same time^{22,24,25} (Supplementary Fig. S6). Recent studies suggest the need for climate models to consider not only changes in sea ice^{3,21} but also the thinning of snow cover on Arctic sea ice to accurately represent and predict Arctic surface warming because even small changes in snow thickness can lead to significant changes in the ice-atmosphere heat exchange²⁶. In a warming climate, the modelled decline in snow thickness on Arctic sea ice is primarily attributed to factors such as later sea ice formation in autumn, an increasing ratio of liquid-to-solid precipitation, and a transition from perennial to seasonal sea-ice cover (in terms of different mean sea-ice states)^{26,27}.

During the summer and fall months, no noticeable trends in biases between decades are evident for both ERA5 and CMIP6, indicating a relatively limited influence of the state of sea ice or snow during this period.

Similarly, satellite-derived temperatures also exhibit low temporal variability during summer¹⁹. Consequently, the winter biases in both ERA5 and CMIP6 play a predominant role in shaping their annual mean biases as well as contributing to the multi-decadal variations.

Estimated warming trends over the Arctic sea ice

The present multi-decadal variations in sea ice related T2m bias in ERA5, together with the declining sea ice state (Fig. 2d) not only present a challenge for benchmarking climate models (as demonstrated above) but also pose a risk of inadequately presenting the warming trend over the Arctic sea ice. When focusing on regions covered by sea ice (with SIC \geq 70%, Fig. 2), the warming rate is estimated to be 0.61 (0.56) °C per decade for CMIP6 (ERA5), compared to the warming rate of 0.79 °C per decade derived from $T_{2m}^{(SAT)}$ for the 1982–2014 period. When considering the Arctic region as north of 66.5 °N^{8,25} including areas of all sea ice cover, open water and land, CMIP6 simulates a warming trend that aligns slightly better with the estimate derived from the combined observational SST and near-surface air temperature than ERA5 (Supplementary Table S2).

Conclusions

Our analysis of surface temperature representations offers a new perspective on climate model performance in the Arctic. Using the satellite-derived T2m dataset as an alternative reference, we show a considerable discrepancy of a profound warm bias in ERA5. This bias prevails particularly in winter in the central Arctic region characterised by thicker sea ice (SIC70), raising concerns on the use of current global reanalysis datasets for model evaluation of near-surface air temperatures. Our analysis shows that the CMIP6 models exhibit reasonable performance in these areas, displaying minor deviations within the range of observational uncertainties (see methods). Re-evaluating warming trends further refines our understanding by revealing that, supplementing previous assessments³, CMIP6 slightly outperforms ERA5 in capturing the warming trend over SIC70. Outside SIC70, it is evident that ERA5 aligns well with observations, while the cold bias in the North Atlantic MIZ in CMIP6 remains consistent with the associated well-documented systematic model bias in the North Atlantic^{1–3}. These findings highlight the imperative role of integrating new observational data for benchmarking climate model.

Methods

Reference datasets

For the near-surface air temperature on the Arctic sea ice, we employed two sets (T2m and SIC) of datasets: one from satellite-observations and one from global reanalysis. The observational T2m dataset is derived from the satellite-based DMI/CMEMS daily gap-free (so called L4) sea surface temperature (SST) and sea ice surface temperature (IST) climate data record spanning from 1 January 1982 to 31 May 2021, covering the Arctic region (>58 °N)¹⁸. The derivation is based on the empirical model developed in¹⁹, which converts clear-sky satellite-observed ISTs to all-sky T2m. Clear-sky ISTs are estimated by excluding the clear-sky bias correction from the daily all-sky L4 DMI/CMEMS ISTs dataset¹⁸. The model is applied to these clear-sky ISTs for grid cells with SIC > 15% (using the DMI-SIC available in the DMI/CMEMS L4 SST/IST dataset¹⁸), resulting in daily gap-free all-sky T2m fields.

For further evaluation of climate models, re-gridding was performed from a 0.05-degree regular lat/lon grid to a coarser 1-degree grid using the nearest neighbour (NN) method for the $T_{2m}^{(SAT)}$ dataset. The NN method using the Climate Data Operators (CDO²⁸ with function -remapnn) is used to choose every 20th grid point along both longitude and latitude dimensions without any interpolation. This choice is made to preserve the original data as much as possible and to ensure that the analysis is representative of the available information, also in the case where there is a mismatch between the coastlines or in the sea ice edge (SIC \leq 15%).

For the ERA5 reanalysis⁴, monthly mean outputs of T2m and SIC were transformed from a 0.25-degree regular lat/lon grid to a coarser 1-degree grid using the nearest neighbour method in CDO (that is, selecting every 4th

grid point along both dimensions). Unlike the NN method, bilinear interpolation (BL) that calculates an output cell value as a weighted average of the four nearest cell centres is commonly used for continuous data, such as the global T2m datasets. Remarkably, for ERA5 T2m with 1 degree resolution, both the NN and BL interpolation methods produce same scientific results. This means that the choice of regridding methods for reference products does not alter the conclusions of this study. To specifically identify $T_{2m}^{(ERA5)}$ on sea ice, a mask file based on ERA5's monthly mean SIC (>15%) was applied for each respective month for the common period 1982–2020.

Validation of reference products

The validation process used for DMI/CMEMS L4 IST, is outlined in Table 2¹⁸, was repeated for the daily T2m datasets of observational $T_{2m}^{(SAT)}$ and reanalysis $T_{2m}^{(ERA5)}$ (on their original grids), respectively. Following ref. 18, we used in situ T2m measurements from three different sources, including 14 Russian North Pole (NP) drifting ice stations, 116 drifting buoys distributed through ECMWF, and 96 drifting buoys from the U.S. Army Cold Regions Research Engineering Laboratory (CRREL). Only matchups with SIC above 15% are considered, and differences deviating more than three times the standard deviation (i.e. lying outside the 99.7% of normally distributed data) from the mean temperature difference have been excluded. This is done to avoid outliers (i.e. from erroneous in situ observations) affecting the results and gives a slightly different number of matchups for $T_{2m}^{(SAT)}$ and $T_{2m}^{(ERA5)}$. As indicated in Supplementary Fig. S1, the in situ matchups are found in all seasons and almost all regions and are thus assumed to be representative of the varying conditions in the Arctic. The statistical analysis against in situ T2m is summarised in Table 1. Note that the true uncertainty of the T2m products is expected to be lower than the standard deviations reported here, since the comparisons also include sampling errors and uncertainties in the in situ observations, assumed to be on the order of 0.5 °C and 0.1 °C, respectively¹⁹. To illustrate the robustness of the validation results, the confidence intervals²⁹ are provided in Table 1.

Climate model data

We analysed monthly mean outputs of the climate models involved in the CMIP6 project³⁰. Modelled SIC (percentage of grid cell covered by sea ice) and T2m (in K) were extracted from in total 47 models at various horizontal resolutions. This was done for the historical period (1982–2014), considering only the first realisation (r1i1p1) of each model used in our analysis. Among these models, 21 models offer SIC on the atmospheric model grid (variable name: siconca), and 26 have SIC on the ocean model grid (variable name: siconc). All data were transformed onto a globally 1-degree regular lat/lon grid (dimensions 180 × 360) using BL in CDO (cdo -remapbil). The models and variables used in our study are provided in Supplementary Table S1. References of the model simulations used in our study are provided in Supplementary Table S1. The process of model selection is illustrated in Supplementary Fig. S4.

Choice of analysis periods

The satellite-derived T2m dataset covers the period from 1982 to 2021 (May), but it's important to note that the CMIP6 historical simulations only extend up to 2014. In Fig. 1, the period 1995–2014 aligns with the IPCC AR6 assessment period, allowing us to assess the impact of the warm bias in ERA5 T2m during this specific time frame.

Figure 2 is designed to further illustrate the evolving nature of biased T2m in both ERA5 (1982–2020) and CMIP6 (1982–2014) throughout the period of available data sets. This extended period reveals that the T2m in ERA5 is consistently warmer than the satellite-derived T2m, although the extent and magnitude varies slightly over time with the changing Arctic conditions. The choice of 1982–2014 for CMIP6 aligns with the available model data for a more meaningful comparison.

Benchmarking the mean climate over the Arctic sea ice

Following the approach outlined in the CMIP6 model evaluation, as detailed in IPCC AR6^{1–3}, we compared the monthly mean air temperature data of

CMIP6 and ERA5 with the corresponding satellite-derived data in Fig. 1. To isolate T2m over Arctic sea ice in the CMIP6 models, we applied a mask of SIC (>15 %, >58 °N) for each respective month m for each respective individual model n . The ensemble mean for month m at grid point (x, y) is then the average for models with the corresponding SIC within the mask limit, i.e.,

$$\mu(x, y)_m^{(CMIP6)} = \frac{1}{N_m} \sum_{n=1}^{N_m} T(x, y)_{m,n}^{(CMIP6)}, \text{ if SIC}(x, y)_{m,n} > 15\% \quad (1)$$

where N_m is the number of models at month m and grid point (x, y) which SIC is within the SIC mask. Similarly, the monthly mean temperatures $\mu(x, y)_m^{(SAT|ERA5)}$ were calculated using a mask defined by the monthly mean SIC (>15%) from DMI-SIC and ERA5, respectively. Subsequently, the monthly mean differences were calculated and averaged over time to obtain the bias:

$$\Delta T(x, y)_m^{(ERA5|CMIP6)} = \mu(x, y)_m^{(ERA5|CMIP6)} - \mu(x, y)_m^{(SAT)}, \text{ if } \mu(x, y)_m \neq \text{nan} \quad (2)$$

$$T_{\text{bias}}(x, y)^{(ERA5|CMIP6)} = \frac{1}{M} \sum_{m=1}^M \Delta T(x, y)_m^{(ERA5|CMIP6)}, \text{ if } \Delta T(x, y)_m \neq \text{nan} \quad (3)$$

where M is the number of months at grid point (x, y) which SIC is within the SIC mask during the analysis period.

It is worth noting that there is little difference in sea ice concentration between ERA5 and DMI-SIC, and the inconsistency is found mainly along coasts and sea ice edges. In Supplementary Fig. S4, which evaluates the seasonal cycle of the total sea ice area north of 58 °N, the data derived from both DMI-SIC and ERA5 demonstrate the closest agreement among the five observational SIC data sets. The averaged temperature and its biases for ERA5 may be slightly different between that calculated using the monthly mean temperature and SIC mask and that using the daily temperature and sea ice mask. However, the conclusions regarding the warm bias in ERA5 over Arctic sea ice remain the same.

Area averaged with SIC at or above 70%

The climatological mean for a 20-year period (1995–2014) was calculated for the observed SIC. A reference line was inserted in Fig. 1a to compass the area with the average SIC $\geq 70\%$ (referred to as SIC70) in the central Arctic ($\geq 66.5^\circ\text{N}$). The total sea ice area for SIC70 is defined as the observed sea ice coverage (Fig. 2d). This area was computed using the monthly mean data with a mask file of SIC $\geq 70\%$ for the period 1982–2020 as

$$A_{\text{SIC70},m} = \sum_x \sum_y \text{gridarea}(x, y), \text{ if SIC}(x, y)_m^{(SAT)} \geq 70\% \quad (4)$$

$$\bar{T}_{\text{SIC70},m} = \frac{\sum_x \sum_y T(x, y)_m \cdot \text{gridarea}(x, y)}{A_{\text{SIC70},m}}, \text{ if SIC}(x, y)_m^{(SAT)} \geq 70\% \quad (5)$$

Correspondingly, the monthly T2m data from the observations, ERA5 and the CMIP6 ensemble mean on the SIC70 grid cells were averaged (Eq. (5)) and then further compared in annual mean (Fig. 2a–c) and decadal monthly mean (Fig. 3). The linear trend of the area-averaged T2m for the common period (1982–2014) was determined using the matlab polyfit function. Individual trends in $\bar{T}_{\text{SIC70}}^{(SAT|ERA5|CMIP6)}$ are statistically significant ($p < 0.05$) based on t -tests.

Data availability

The satellite-derived IST and sea ice data used for the T2m calculation are available from <https://doi.org/10.48670/moi-00123>¹⁸. The resulting daily T2m dataset with a 0.05-degree resolution is archived on the DMI data server, and the monthly data used in this analysis are archived in Zenodo (10.5281/zenodo.10555223³¹). ERA-5 reanalysis data are freely available in the Copernicus Climate Change Service Climate Data Store (<https://cds.climate.copernicus.eu/>). The CMIP6 model ensemble is obtained using the first member of each model that is available freely and publicly from the Earth System Grid Federation (ESGF, <https://esgf.llnl.gov>). The in situ data from ECMWF covering 1993–2015 can be accessed through the World Meteorological Organization’s Global Telecommunication System (GTS) and is available for members in the ECMWF Meteorological Archival and Retrieval System (MARS). The U.S. Army Cold Regions Research Engineering Laboratory (CRREL) mass balance buoys for the period 2001–2017 are available from <http://imb-crrel-dartmouth.org>^{32,33} and the Russian North Pole (NP) drifting ice stations are available from NCAR RDA Dataset ds474.0 (ucar.edu)³⁴.

Code availability

The scripts and source data used to produce the figures and tables in this paper are archived in Zenodo (<https://doi.org/10.5281/zenodo.10208080>³⁵).

Received: 16 July 2023; Accepted: 19 February 2024;

Published online: 01 March 2024

References

- Eyring, V. et al. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, chap. 3. Human Influence on the Climate System, 423–552 (Cambridge University Press, 2021).
- Bock, L. et al. Quantifying progress across different CMIP phases with the ESMValTool. *J. Geophys. Res.* **125**, e2019JD032321 (2020).
- Davy, R. & Outten, S. The Arctic surface climate in CMIP6: status and developments since CMIP5. *J. Clim.* **33**, 8047–8068 (2020).
- Hersbach, H. et al. The ERA5 global reanalysis. *Quart. J. Roy. Meteorol. Soc.* **146**, 1999–2049 (2020).
- Graham, R. M., Hudson, S. R. & Maturilli, M. Improved performance of ERA5 in Arctic gateway relative to four global atmospheric reanalyses. *Geophys. Res. Lett.* **46**, 6138–6147 (2019).
- Simmons, A. et al. A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets. *Quart. J. Roy. Meteorol. Soc.* **143**, 101–119 (2017).
- Benestad, R., Erlandsen, H., Mezghani, A. & Parding, K. Geographical distribution of thermometers gives the appearance of lower historical global warming. *Geophys. Res. Lett.* **46**, 7654–7662 (2019).
- AMAP. *Arctic Climate Change Update 2021: Key Trends and Impacts. Summary for Policy-Makers* (Arctic Monitoring and Assessment Programme (AMAP), 2021).
- Wang, C., Graham, R. M., Wang, K., Gerland, S. & Granskog, M. A. Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic sea ice: effects on sea ice thermodynamics and evolution. *Cryosphere* **13**, 1661–1679 (2019).
- Batrak, Y. & Müller, M. On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice. *Nat. Commun.* **10**, 4170 (2019).
- Zhang, Z.-L. et al. On the turbulent heat fluxes: A comparison among satellite-based estimates, atmospheric reanalyses, and in-situ observations during the winter climate over Arctic sea ice. *Adv. Clim. Change Res.* **14**, 347–362 (2023).
- Heinemann, G., Schefczyk, L., Willmes, S. & Shupe, M. D. Evaluation of simulations of near-surface variables using the regional climate model CCLM for the MOSAiC winter period. *Elem. Sci. Anth.* **10**, 00033 (2022).

13. Zampieri, L. et al. A machine learning correction model of the winter clear-sky temperature bias over the Arctic sea ice in atmospheric reanalyses. *Mon. Weather Rev.* **151**, 1443–1458 (2023).
14. Herrmannsdörfer, L., Müller, M., Shupe, M. & Rostosky, P. Surface temperature comparison of the Arctic winter MOSAiC observations, ERA5 reanalysis, and MODIS satellite retrieval. *Elem. Sci. Anth.* **11**, 00085 (2023).
15. Batrak, Y., Cheng, B. & Kallio-Myers, V. Sea ice cover in the Copernicus Arctic Regional Reanalysis. *Cryosphere Discuss.* **2023**, 1–40 (2023).
16. Arduini, G. et al. On the importance of representing snow over sea-ice for simulating the Arctic boundary layer. *J. Adv. Model. Earth Syst.* **14**, e2021MS002777 (2022).
17. Day, J. J. et al. Benefits and challenges of dynamic sea ice for weather forecasts. *Wea. Clim. Dyn.* **3**, 713–731 (2022).
18. Nielsen-Englyst, P. et al. A combined sea and sea-ice surface temperature climate dataset of the Arctic, 1982–2021. *Remote Sens. Environ.* **284**, 113331 (2023).
19. Nielsen-Englyst, P. et al. Deriving Arctic 2 m air temperatures over snow and ice from satellite surface temperature measurements. *Cryosphere* **15**, 3035–3057 (2021).
20. GCOS. The global observing system for climate: Implementation needs. *World Meteorol. Organ.* **200**, 316 (2016).
21. Shen, Z., Duan, A., Li, D. & Li, J. Assessment and ranking of climate models in Arctic Sea ice cover simulation: From CMIP5 to CMIP6. *J. Clim.* **34**, 3609–3627 (2021).
22. Horvat, C. Marginal ice zone fraction benchmarks sea ice and climate model skill. *Nat. Commun.* **12**, 2221 (2021).
23. Keen, A. et al. An inter-comparison of the mass budget of the Arctic sea ice in CMIP6 models. *Cryosphere* **15**, 951–982 (2021).
24. Notz, D. & Community, S. Arctic sea ice in CMIP6. *Geophys. Res. Lett.* **47**, e2019GL086749 (2020).
25. Rantanen, M. et al. The Arctic has warmed nearly four times faster than the globe since 1979. *Commun. Earth Environ.* **3**, 168 (2022).
26. Landrum, L. L. & Holland, M. M. Influences of changing sea ice and snow thicknesses on simulated Arctic winter heat fluxes. *Cryosphere* **16**, 1483–1495 (2022).
27. Webster, M., DuVivier, A., Holland, M. & Bailey, D. Snow on Arctic sea ice in a warming climate as simulated in CESM. *J. Geophys. Res. Oceans* **126**, e2020JC016308 (2021).
28. Schulzweida, U. CDO User Guide (1.9.9). *Zenodo* <https://doi.org/10.5281/zenodo.4246983> (2020).
29. Thomson, R. E. & Emery, W. J. *Data analysis methods in physical oceanography* (Newnes, 2014).
30. Eyring, V. et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).
31. Nielsen-Englyst, P. & Høyer, J. Monthly near-surface air temperatures over Arctic sea ice based on satellite observations. *Zenodo* <https://doi.org/10.5281/zenodo.10555223> (2024).
32. Perovich, D., Richter-Menge, J. & Polashenski, C. Observing and understanding climate change: Monitoring the mass balance, motion, and thickness of Arctic sea ice, The CRREL-Dartmouth Mass Balance Buoy Program, CRREL-Dartmouth, available at: <http://imb-crrel-dartmouth.org> (2016). Last access: 24 November 2016.
33. Richter-Menge, J. et al. Ice mass-balance buoys: a tool for measuring and attributing changes in the thickness of the Arctic sea-ice cover. *Ann. Glaciol.* **44**, 205–210 (2006).
34. RU/FSR/HME/AARI & NSIDC. AARI Russian North Polar Drifting Station Data, from NSIDC. <https://doi.org/10.5065/V47S-KW40> (1993).
35. Tian, T. & Nielsen-Englyst, P. Scripts for figures and tables in “Cooler Arctic surface temperatures simulated by climate models are closer to satellite-based data than the ERA5 reanalysis”. *Zenodo* <https://doi.org/10.5281/zenodo.10208080> (2023).

Acknowledgements

The authors express gratitude to Dr. Heike Langenberg, the editor, Dr. Lorenzo Zampieri, and the two anonymous reviewers for their valuable suggestions, which have greatly improved this paper. Special thanks are also extended to Dr. Kristian Mogensen, Dr. Sarah Keeley, and Dr. Steffen Tietsche from ECMWF for sharing information on the progress in mitigating IST bias in the reanalysis models, which inspired this work. The authors were supported by the Danish National Center for Climate Research (NCKF). SY is also partly funded by the EU Horizon Europe research and innovation project OptimESM under grant agreement No101081193. JLH and PNE are also funded by Copernicus Marine Services.

Author contributions

Conception of the work: TT, SY, JLH. Acquisition of data: TT, PNE and SSI. TT designed analysis, produced figures and wrote the manuscript. PNE performed the validation of T2m against in situ observations and produced the figure. All authors read, improved, and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43247-024-01276-z>.

Correspondence and requests for materials should be addressed to Tian Tian.

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

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024