

<https://doi.org/10.1038/s41612-024-00801-3>

Revisiting the reanalysis-model discrepancy in Southern Hemisphere winter storm track trends

Joonsuk M. Kang¹✉, Tiffany A. Shaw^{1,2}, Sarah M. Kang², Isla R. Simpson³ & Yue Yu⁴

Southern Hemisphere (SH) storminess has increased in the satellite era and recent work suggests comprehensive climate models significantly underestimate the trend. Here, we revisit this reanalysis-model trend discrepancy to better understand the mechanisms underlie it. A comprehensive like-for-like analysis shows reanalysis trends exhibit large uncertainty, and coupled climate model simulations exhibit weaker trends than most but not all reanalyses. However, simulations with prescribed sea surface temperature (SST) exhibit significantly greater storminess trends, particularly in the South Pacific, implying SST trend discrepancies in coupled simulations impact storminess trends. Using pacemaker simulations that correct Southern Ocean and tropical east Pacific SST trend discrepancies, we show that storminess trends in coupled simulations are underestimated because they do not capture the enhanced storminess resulting from Southern Ocean cooling and La-Nina-like teleconnection trends. Our findings emphasize large reanalysis uncertainty in SH circulation trends and the impact of regional SST trend discrepancies on circulation trends.

The extratropical circulation in the Southern Hemisphere (SH) is characterized by a strong storm track related to tracks of cyclones and anticyclones^{1,2}. The intensity of the storm track, hereafter the storminess, is tightly connected to surface weather in the SH^{3,4}. Understanding how storminess will change under anthropogenic forcing is important for mitigating the impacts of climate change². Climate models project that the SH storm track will intensify by the end of the 21st century under climate change^{2,5–9}, bringing increased precipitation¹⁰ and stronger surface winds¹¹.

Recent work has shown that SH storminess has increased significantly in the satellite era (1979 to present) in reanalysis data^{9,12}. However, the trends in reanalysis data were 2–3 times larger than the multi-model mean trends from models participating in the Coupled Model Intercomparison Project Phase 5 and Phase 6 (CMIP5 and CMIP6)^{13,14}. Thus, recent work concluded that climate models significantly underestimate the storminess trend in the reanalysis datasets in the satellite era. This reanalysis-model trend discrepancy in the SH winter storm track calls into question the ability of climate models to predict future weather in the SH.

A discrepancy between the climate model and observed trends can have multiple causes that can be categorized into three factors^{15,16}: (I) The observations are in error, (II) The observation-model comparison is flawed, (III) The models are deficient.

For (I), the trends can differ substantially across observational datasets¹⁷ and lead to observational uncertainty. The use of up-to-date observational data can be important in reconciling observation-model trend discrepancies^{18,19}. For storminess, the observed trend is quantified using reanalysis datasets, which involve uncertainties arising from data assimilation techniques, physical parameterizations, and evolution of observational systems^{20,21}. The reanalysis uncertainty can be particularly important in the SH where ground-based observations are limited^{22,23}. For example, recent work showed that trends in SH winter weather-scale temperature variance differ by a factor of 5 across reanalysis datasets²⁴.

For (II), there are two important aspects to consider. First, reanalysis trends involve a single realization of internal variability whereas model simulations reflect a distribution of realizations. Thus, it is important to properly sample the internal variability by using a large number of model simulations^{25,26}. Second, a like-for-like comparison whereby the observations and climate models are compared with the same temporal and spatial sampling has been important for reconciling previous discrepancies^{27,28}. A like-for-like comparison is especially important for storm tracks, which sample specific time and spatial scales²⁹.

¹Department of the Geophysical Sciences, The University of Chicago, Chicago, IL, USA. ²Max Planck Institute for Meteorology, Hamburg, Germany. ³Climate and Global Dynamics Laboratory, NSF National Center of Atmospheric Research, Boulder, CO, USA. ⁴State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, China. ✉e-mail: jmkang@uchicago.edu

For (III), the models can be deficient in either the forced response or the internal variability because they are incapable of simulating the physical mechanism responsible for the observed trend. For example, CMIP6 models fail to simulate recent sea surface temperature (SST) trends in the tropical Pacific and Southern Ocean^{30–34}. The tropical SST trends in CMIP6 models are characterized by an El Niño-like trend in the tropical Pacific, as opposed to a La Niña-like trend in the observations, with the observed trends lying at the edge or outside of the model trend distribution^{30–32}. The observed cooling trend in the Southern Ocean is also not well captured by CMIP6 models^{30,33} and it has been suggested that this is also related to the SST trend difference in the tropical Pacific^{33,35}. Previous work concluded that coupled models exhibit a systematic bias in the representation of SST trends and that differences between observed and modeled trends are very unlikely to occur due to internal variability³⁰.

The reanalysis-model discrepancy in the SH winter storm track trends should be revisited following the three factors described above. (I) It is currently unclear how increasing the number of reanalysis datasets to address reanalysis uncertainty would affect the discrepancy. (II) It is also unknown how expanding the model ensemble size and addressing like-for-like comparison (calculating storminess using the same time and spatial grids) would affect the discrepancy. (III) Lastly, the impact of observation-coupled model SST trend discrepancy on the SH winter storminess trends has not been quantified, although SST trends are related to other large-scale circulation trends in the SH^{30,36,37}.

Here we revisit the reanalysis-model discrepancy in the Southern Hemisphere winter storm track trends with additional reanalyses, expanded model ensembles, and like-for-like calculations. We also quantify the impact of SST trend discrepancy on the storm track trends, including the mechanisms connecting them.

Results

Revisiting the SH winter storminess trend discrepancy

We revisit the SH winter storminess trends in reanalyses and CMIP6 and AMIP6 multi-model ensembles (see Methods and Table S1), quantified using vertically integrated bandpass-filtered eddy kinetic energy (hereafter EKE, see Methods) as in previous work⁹. To address (I) and (II), we calculate EKE using a larger number of reanalysis data (expanding from 3 to 8) and model ensemble members (expanding from 16 to 26 for CMIP6 and from 13 to 32 for AMIP6, see Methods). We also ensure that EKE is calculated using the same time frequency and spatial grids: daily-mean, 8 pressure levels, and 1.5° by 1.5° horizontal grids (see Methods).

Expanding the number of reanalysis datasets shows the zonal-mean SH storminess trend exhibits large observational uncertainty (Fig. 1a and see Fig. S1 for time series). The reanalysis trend spread is larger than the multi-model ensemble spread, and not all trends are statistically significant. The largest and the smallest trends come from older generation reanalyses (NCEP2 and MERRA1), whereas the newer generation reanalyses are closer together (ERA5 and JRA3Q). The large spread in reanalysis trends is not significantly affected by start and end dates (Fig. S2), and a large spread is also found using a different metric for storminess (e.g., sea-level pressure variance¹¹, Fig. S3). Note that storminess climatology does not exhibit similarly large uncertainty, except for NCEP2 which has about 28% weaker climatology than other reanalyses (Fig. S1).

Six out of eight reanalysis trends, including the newest reanalysis trends, fall outside the 10–90% range of the trend distribution of coupled CMIP6 ensemble which predicts SST (Fig. 1a, Table S2). For five reanalysis trends, no more than one CMIP6 ensemble member simulates a trend as high. Thus, while there is large observational uncertainty, CMIP6 ensemble seems to be underestimating the storminess trends. Although weak, the CMIP6 trend is dominated by greenhouse gas forcing according to the

Fig. 1 | Revisiting the reanalysis-model storminess trend discrepancy for CMIP6 and AMIP6 ensembles.

a Linear trends of SH JJA EKE (40–75°S) in 8 reanalysis datasets (blue colors, 1979–2018) and 26 CMIP6 (1979–2018) and 32 AMIP6 (1979–2014) model simulations (diamonds). Statistically significant trends at the 95% confidence level are filled. The box represents the full spread of reanalysis trends and the 10–90% range of model ensemble trends. The horizontal line inside the box shows the median trend in the model ensemble. **b** Similar results to (a), but for the South Pacific (40–75°S, 170°E–60°W). Spatial pattern of SH JJA EKE trend for (c) reanalysis mean, (d) CMIP6 and (e) AMIP6 multi-model ensemble mean. Stipples indicate where reanalysis-mean or multi-model ensemble mean trends are significant at the 95% level. The green dashed lines indicate the South Pacific sector (40–75°S, 170°E–60°W).

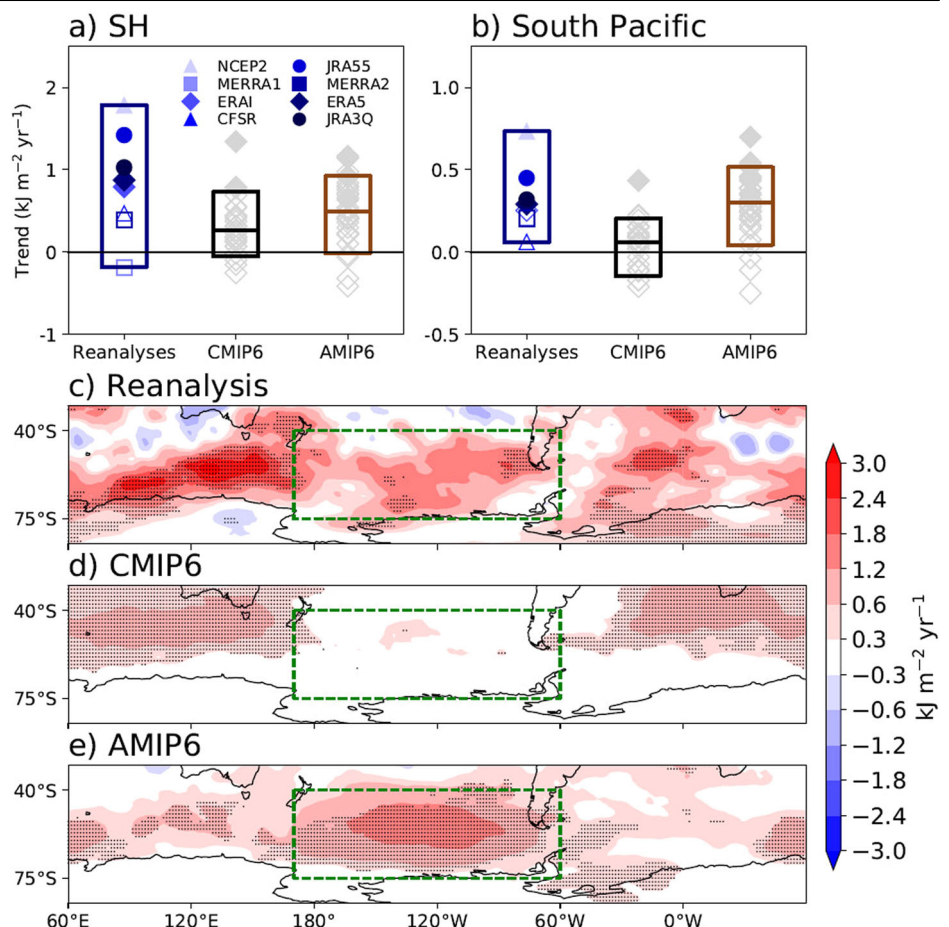
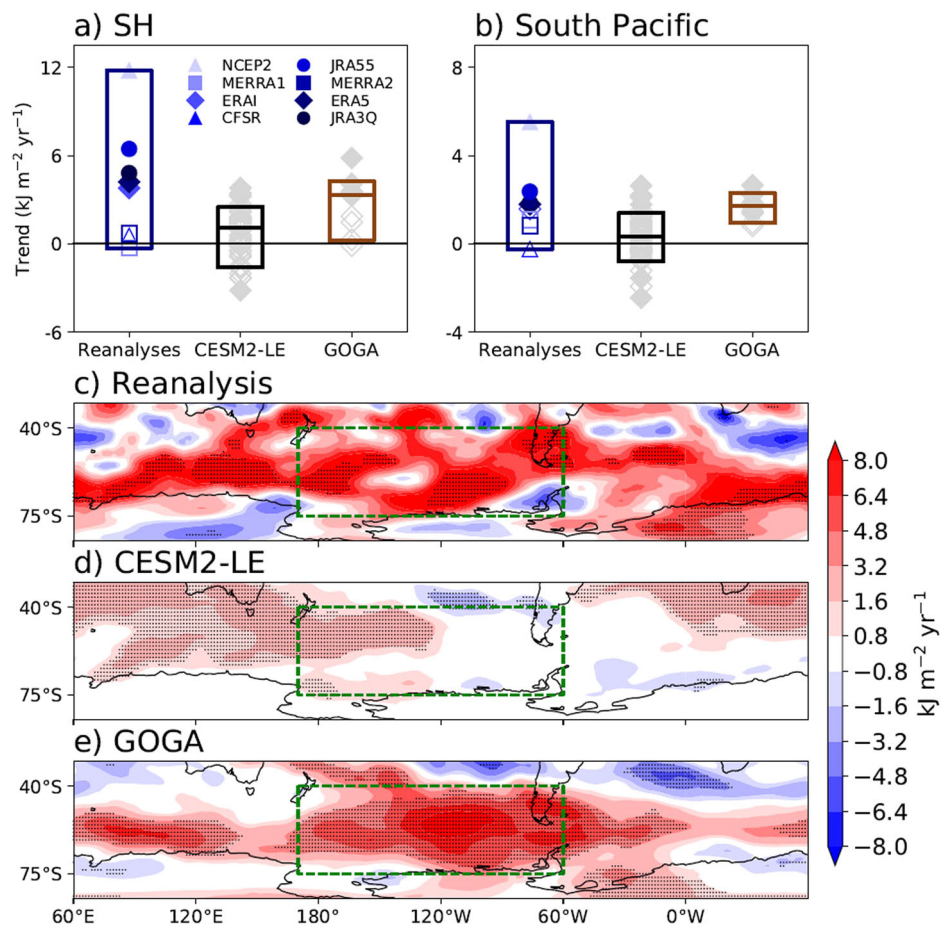


Fig. 2 | Revisiting the reanalysis-model storminess trend discrepancy using CESM2-LE and GOGA simulations. a–e Similar results to Fig. 1, but using CESM2-LE and GOGA simulations. Note that EKE is defined differently from Fig. 1 and trends are calculated from 1979 to 2013.



Detection and Attribution Model Intercomparison Project simulations (DAMIP³⁸, see Fig. S4).

Four reanalysis trends fall within the 10–90% range of the trend distribution of prescribed-SST AMIP6 ensemble (Fig. 1a, Table S2). Consistently, comparing the distributions of storminess trends between the two ensembles, the trends in the AMIP6 are significantly greater than those in CMIP6 (p value = 0.01 see Methods), indicating that prescribing observed SST trends significantly increases the storminess trends in the models. Similar results were found in previous work³⁷, where annual-mean atmospheric energy transport trends from transient eddies were greater in the AMIP6 than the CMIP6 ensemble.

The spatial pattern of the storminess trends provides insights into understanding the different trends in CMIP6 and AMIP6 ensembles (Fig. 1c–e). The multi-reanalysis-mean storminess trend is significant across most SH regions including high latitudes of the Indian Ocean, South Pacific, and South Atlantic (Fig. 1c and Fig. S5 for individual reanalysis). However, the CMIP6 ensemble-mean storminess trend in the South Pacific (170°E–60°W) is negligible (Fig. 1d). The AMIP6 ensemble-mean trend better captures the reanalysis trend, especially in the South Pacific, where CMIP6 ensemble shows a near-zero ensemble-mean trend (compare Fig. 1d, e). For more quantitative comparison, storminess trend distributions in the South Pacific are examined for the CMIP6 and AMIP6 ensembles (Fig. 1b). Five reanalysis trends fall outside the 10–90% range of CMIP6 ensemble trend distribution in the South Pacific (Fig. 1b). In contrast, seven reanalysis trends fall within the 10–90% range of AMIP6 ensemble trend distribution (Fig. 1b). Comparing the storminess trend distributions between the two ensembles, the trends in AMIP6 are significantly greater than those in the CMIP6 in the South Pacific (p value < 0.01). Overall, this demonstrates that prescribing observed SST trends significantly increases the South Pacific storminess trends in the models. These results are not sensitive to different

years used in calculating trends in CMIP6 and AMIP6, and consistent results are obtained over 1979–2013 (Fig. S6).

To better understand the role of internal variability related to (II) for SH winter storminess trends, we perform similar analyses using the 50-member Community Earth System Model version 2 Large Ensemble (CESM2-LE) simulations (Fig. 2)^{39,40}. We also use the 10-member prescribed (observed) SST CESM2 simulations, namely Global Ocean Global Atmosphere (GOGA) simulations (Fig. 2). Both CESM2-LE and GOGA simulations are forced with the same radiative forcing as CMIP6 and AMIP6 ensembles from 1979 to 2014. Here we focus on trends from 1979 to 2013 (see Fig. S6 for CMIP6 and AMIP6). The CESM2-LE simulations also do not capture the observed SST trends (compare Fig. 3a, b).

For CESM2, EKE is defined from monthly anomalies due to data availability (see Methods). The zonal-mean storminess trend in reanalyses quantified using this definition of EKE also exhibits large observational uncertainty (Fig. 2a). It is also notable that the reanalysis trends are greater (especially the statistically significant ones) compared to Fig. 1a by roughly 4–5 times. This is related to using monthly anomalies instead of 2.5–6 day bandpass-filtered anomalies, which includes a broader spectrum of kinetic energy. Nevertheless, changing the EKE definition does not impact the result that storminess trends are greater in the AMIP6 than CMIP6 ensemble (Fig. S7).

Five reanalysis trends, including the newest reanalyses (ERA5 and JRA3Q), fall outside the 10–90% range of the 50-member CESM2-LE trend distribution (Fig. 2a), similar to CMIP6 ensemble in the zonal mean. The CESM2-LE simulations also show a smaller ensemble-mean storminess trend in the South Pacific compared to reanalyses (compare Fig. 2c, d). Similarly, five reanalysis trends fall outside the 10–90% of the CESM2-LE trend distribution in the South Pacific (Fig. 2b). Thus, CESM2-LE simulations underestimate reanalysis trends similar to CMIP6 ensemble, and

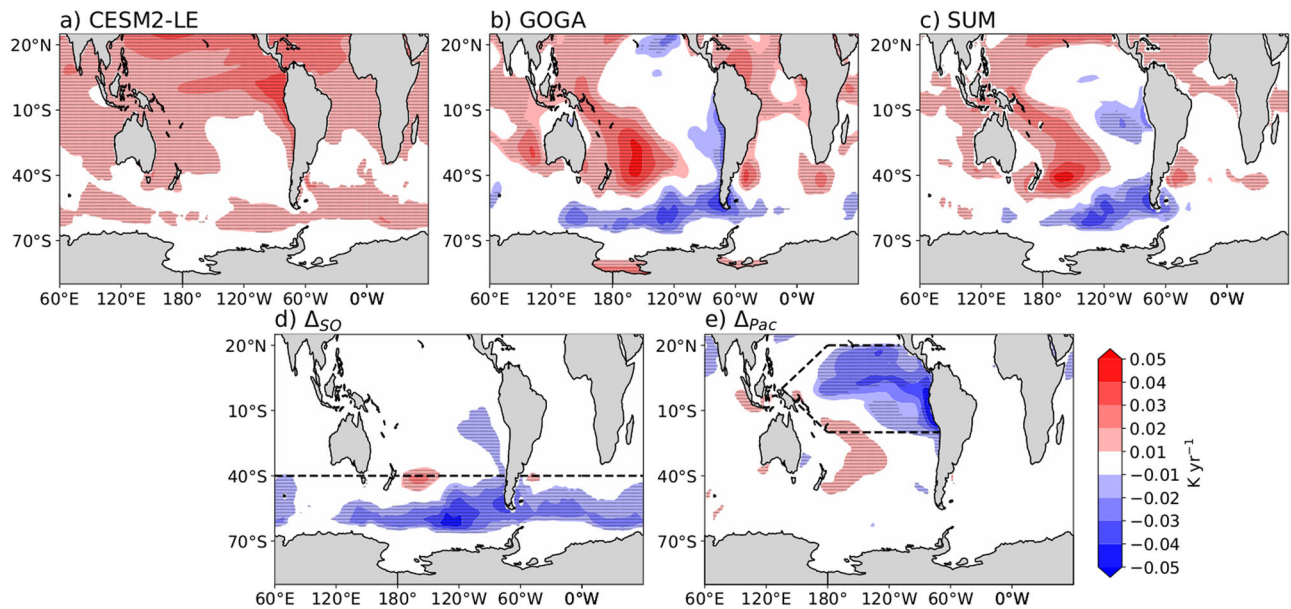


Fig. 3 | Sea surface temperature trends in CESM2-LE and the impact of pace-making. a–e Spatial pattern of ensemble-mean JJA SST trends from 1979 to 2013 for (a) CESM2-LE, (b) GOGA, (c) SUM = CESM2-LE + Δ_{Pac} + Δ_{SO} , (d) Δ_{SO} = [SOPACE] – [CESM2-LE], and (e) Δ_{Pac} = [PacPACE] – [CESM2-LE] simulations.

Stipples indicate where ensemble-mean trends are significant at the 95% level. In (d, e), the dashed black lines represent where the SST anomalies are nudged to observation.

internal variability cannot fully account for the reanalysis-model mismatch. We also find similar results for CESM2-LE and GOGA comparison (Fig. 2) as the CMIP6 and AMIP6 comparison (Fig. 1, Figs. S6, S7), where the storminess trends in prescribed SST simulations are significantly greater than those in coupled simulations, particularly in the South Pacific.

The analysis above which accounts for (I) and (II) reveals large uncertainty in reanalysis storminess trends in SH. However, most of the coupled simulations (CMIP6 and CESM2-LE) simulate trends that are smaller than most of the reanalyses, including the newest ones, thus they could be underestimating the reanalysis storminess trends. The simulations with prescribed observed SST (AMIP6 and GOGA), on the other hand, show statistically significantly greater storminess trends than the coupled simulations in the zonal mean and the South Pacific. This indicates that (III) coupled simulation's deficiency in simulating observed SST trends impacts the prediction of SH storminess trends. In what follows, we show that the important factor impacting the storminess trend underestimation in the coupled simulations is their misrepresentation of the SST trends, particularly in the Southern Ocean and tropical Pacific.

Impact of SST trend discrepancies on storminess trends

SST trend discrepancies can impact the storminess through different mechanisms. Southern Ocean SST trends reflect changes in surface fluxes and equatorward ocean energy transport⁴¹. Recent work suggested CMIP6 models underestimate annual-mean SH storminess trends because they do not capture the surface energy flux trends connected to Southern Ocean cooling¹². Furthermore, tropical Pacific SST trends likely impact the SH through Rossby wave teleconnections. More specifically, interannually La Nina leads to a stronger South Pacific storminess^{42–44}, thus a La-Nina-like SST trend would be expected to strengthen the South Pacific storminess trend.

In order to test the hypothesis that SST trend discrepancies contribute to the underestimation of SH winter storminess trends in coupled simulations, we utilize the Southern Ocean (SOPACE) and tropical Pacific (PacPACE) CESM2 pacemaker simulations (see Methods). The pacemaker simulations nudge the Southern Ocean and tropical east Pacific SST anomalies to observations (Fig. 3d, e). Thus, comparing pacemakers to the free-running CESM2-LE simulations allows us to quantify how SH

storminess trends in the coupled simulations would change if they correctly simulated the observed Southern Ocean cooling and La-Nina-like SST trend. Since the SOPACE and PacPACE simulations are forced with the same radiative forcing as CESM2-LE simulations, their ensemble-mean differences with the CESM2-LE simulations (Δ_{SO} and Δ_{Pac}) quantify the impact of capturing SST trends in the Southern Ocean and the tropical Pacific, respectively (see Methods).

Impact of Southern Ocean SST trend discrepancy on storminess trends. When CESM2 simulations are forced with historical radiative forcings and SST anomalies are nudged to observations in the Southern Ocean (Fig. 3d), zonal-mean storminess trends increase compared to the CESM2-LE simulations (Fig. 4a). Most of the members in SOPACE simulations exhibit larger trends than the CESM2-LE median trend (compare green box and black dashed line in Fig. 4a). More quantitatively, the SOPACE zonal-mean storminess trends are significantly greater than those from the CESM2-LE simulations (p value = 0.02). The strengthening is noticeable across all longitudes in SH, especially at higher latitudes (compare Figs. 2d, 4c).

From an energetic perspective, storminess trends are affected by surface energy flux, top-of-the-atmosphere radiation, and stationary circulation trends (see Eq. 1 in Shaw et al.¹²). Among these factors, Shaw et al.¹² hypothesized the underestimation of annual-mean SH storminess in the CMIP6 models was due to an underestimated surface energy flux trends in CMIP6 models in the SH, which is connected to Southern Ocean cooling trends. The surface energy flux trends reflect ocean heat transport divergence and storage trends⁴⁵. Thus, if the surface energy flux implies equatorward ocean heat transport, it also strengthens the SH storminess because it steepens the atmospheric equator-to-pole energy gradient (i.e., energy is moved away from the South Pole to the equator by the ocean). This influence of surface energy flux trends on the storminess reflecting ocean heat transport trends can be quantified through the moist static energy budget^{7,46} with the poleward atmospheric energy transport implied from surface energy flux (F_{SFC} ; see Methods).

The poleward atmospheric energy transport (F_{SFC}) induced from surface energy flux shows positive ensemble-mean trends across SH in SOPACE simulations (green, Fig. 5a), which is consistent with strengthening SH storminess, similar to ERA5 (blue, Fig. 5a). In contrast, the

Fig. 4 | Storminess trends when SST trends are corrected in coupled simulations. a–e Similar results to Fig. 2, but using SOPACE, PacPACE, and SUM simulations. In (a, b), the median trend from the CESM2-LE and GOGA simulations are shown in dashed lines. The reanalysis trends are repeated from Fig. 2.

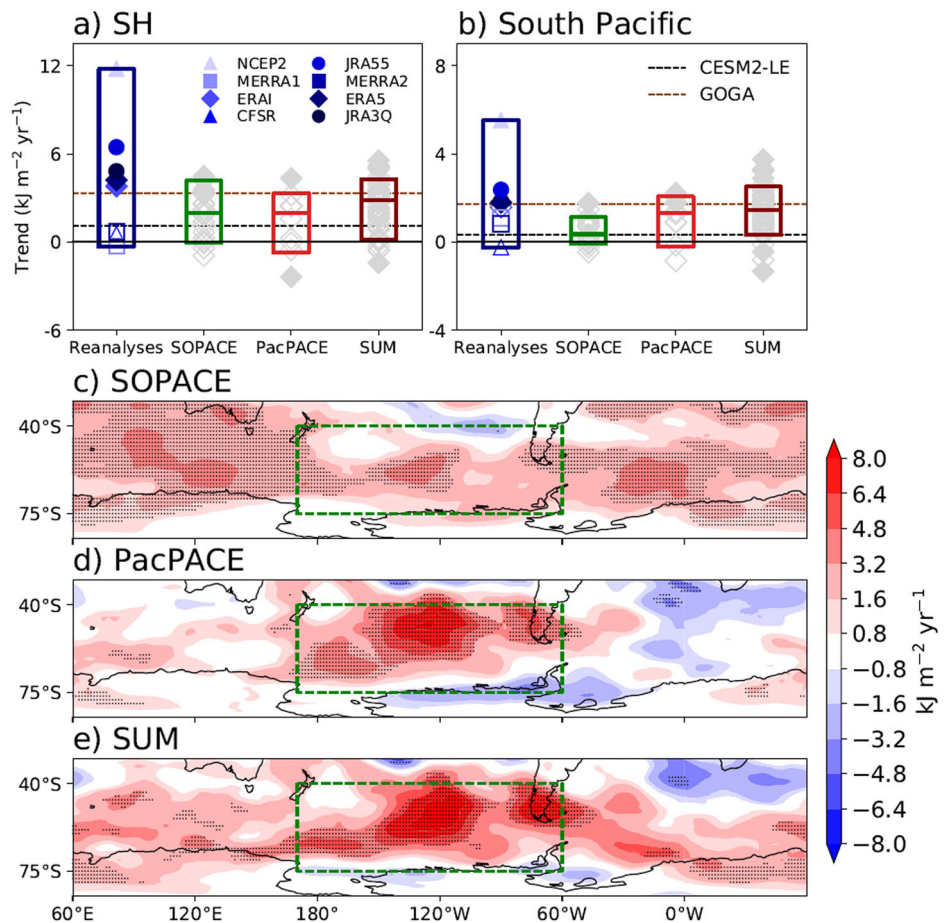
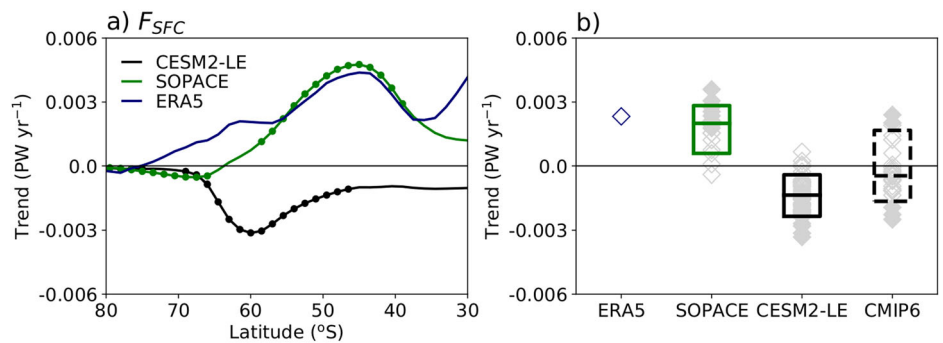


Fig. 5 | Southern Ocean pacemaker improves surface energy flux induced atmospheric energy transport trends. a Linear trends in JJA zonal-mean F_{SFC} from 1979 to 2013 for ERA5 (dark blue), ensemble-mean SOPACE (green) and CESM2-LE (black) simulations. Statistically significant trends at the 95% level are dotted. b Linear trends of SH JJA F_{SFC} (40–75°S) from 1979 to 2013 for ERA5, SOPACE, CESM2-LE, and CMIP6 ensemble (diamonds). Statistically significant trends at the 95% confidence level are filled. The box represents the 10–90% range of model ensemble trends. The horizontal line inside the box shows the median trend in the model ensemble.



ensemble-mean trends in CESM2-LE simulations are negative across SH indicating storminess weakening (black, Fig. 5a). Most of the ensemble members in the SOPACE simulations and ERA5 show positive F_{SFC} trends, whereas most of the members in the CESM2-LE have negative trends (Fig. 5b). Thus, the F_{SFC} trends from the SOPACE simulations are significantly greater than those from CESM2-LE simulations (p value < 0.01 , compare green and black in Fig. 5b). Consistent results are found from the CMIP6 ensemble where most of the models simulate negative F_{SFC} trends (black dashed, Fig. 5b).

Thus, the SOPACE simulations simulate surface energy flux trends consistent with reanalysis trends, which the CESM2-LE simulations struggled to simulate. Consequently, storminess trends are larger in SOPACE than CESM2-LE due to surface energy flux trends that better capture reanalysis trends. The causality is less likely to be the other way around since temperature at the surface is being nudged in the SOPACE simulations.

Note that while surface energy flux and F_{SFC} trends indicate storminess weakening in CESM2-LE and CMIP6, other factors such as top-of-the-atmosphere radiation trends strengthen the storminess (Fig. S8), leading to weak but positive ensemble-mean storminess trends (e.g., Figs. 1, 2).

While nudging SST anomalies in the Southern Ocean significantly strengthens storminess trends in the zonal mean, its impact in the South Pacific is weak. Quantitatively, the SOPACE and CESM2-LE trend distributions in the South Pacific (Fig. 4b) are not significantly different (p value = 0.29). This suggests that other processes are important for storminess trends in the South Pacific.

Impact of tropical Pacific SST trend discrepancy on storminess trends. When CESM2 simulations are forced with historical radiative forcings and SST anomalies are nudged to observations in the tropical east Pacific (Fig. 3e), South Pacific storminess trends increase compared

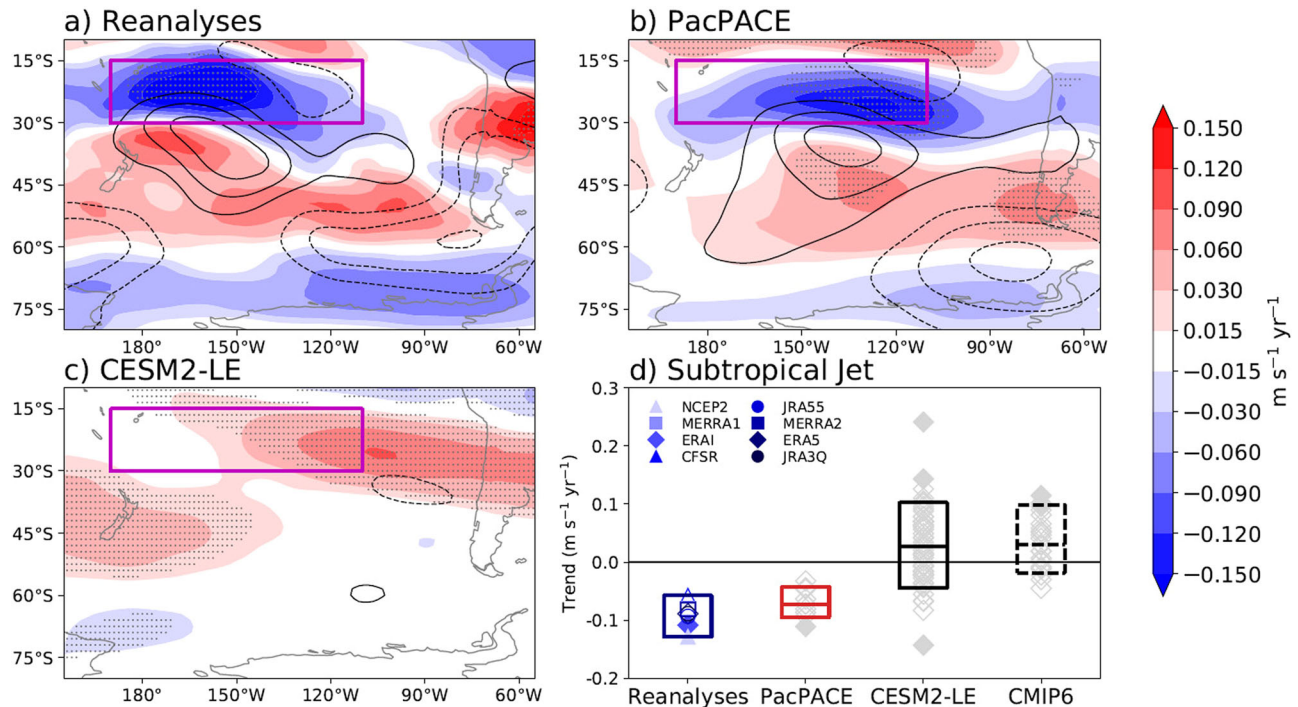


Fig. 6 | Pacific pacemaker improves La-Nina-like teleconnection trends in the South Pacific. Spatial pattern of 250-hPa zonal wind (colors) and eddy geopotential height (contours) trends for (a) reanalysis mean, (b) PacPACE and (c) CESM2-LE simulation ensemble mean. The positive and negative eddy geopotential height trends are respectively depicted in solid and dashed contours in 0.3 m yr⁻¹ intervals (zero contour is suppressed). Stipples indicate where reanalysis-mean or ensemble-mean trends are significant at the 95% level. The magenta box

indicates the South Pacific subtropical jet sector (15–30°S, 170°E–110°W).

d Linear trends of South Pacific subtropical jet from 1979 to 2013 for reanalyses (blue colors), PacPACE, CESM2-LE, and CMIP6 ensemble (diamonds). Statistically significant trends at the 95% confidence level are filled. The box represents the full spread of reanalysis trends and the 10–90% range of model ensemble trends. The horizontal line inside the box shows the median trend in the model ensemble.

to the CESM2-LE simulations (Fig. 4b, d). Most of the ensemble members in the PacPACE simulations show larger trends than the CESM2-LE simulation median trend (compare red box and black dashed line in 4b). Thus, the South Pacific storminess trends from the PacPACE simulations are significantly greater than those from the CESM2-LE simulations (p value = 0.02). Note that the zonal-mean trend distribution from the PacPACE simulations is not statistically different (p value = 0.35) from the trend distribution from the CESM2-LE simulations (Fig. 4a).

We hypothesized that the La Nina-like SST trend in the tropical Pacific induces a Rossby wave teleconnection trend to the South Pacific, characterized by weaker subtropical jet and strengthened storminess in the South Pacific, similar to what is seen interannually^{42–44}. Specifically, during La Nina winters, the colder tropical Pacific temperature leads to weaker subtropical jet directly via thermal wind balance. Weaker subtropical jet then leads to upper tropospheric eddy momentum flux convergence, which induces adiabatic descent to warm the midlatitudes (see Eq. 7 in Seager et al.⁴²). The poleward eddy heat flux in the midlatitudes increases to alleviate the adiabatic warming, strengthening the storminess in the South Pacific⁴². As such, the subtropical jet weakening can lead to South Pacific storminess strengthening.

The 250-hPa zonal wind and eddy geopotential (deviation from the zonal mean) trends in the reanalyses and PacPACE simulations show a clear La Nina-like teleconnection pattern that is absent in the CESM2-LE simulations (Fig. 6a–c). In particular, all reanalyses and all members in PacPACE simulations have weakening subtropical jet trends across the South Pacific. The subtropical jet trends in PacPACE simulations are significantly different (p value < 0.01) from those in the CESM2-LE simulations where most of the members have strengthening subtropical jet similar to CMIP6 ensemble (dashed box, Fig. 6d). This confirms that PacPACE simulations exhibit stronger South Pacific storminess by capturing La Nina-like teleconnection trends in reanalysis.

Combined impact of Southern Ocean and tropical Pacific SST trend discrepancies on storminess trends. To investigate the combined impact of both pacemakers ($\Delta_{Pac} + \Delta_{SO}$) on the coupled simulations, we create a synthetic large ensemble named SUM with 50 members, which is defined as:

$$\text{SUM} = \text{CESM2-LE} + \Delta_{Pac} + \Delta_{SO} \quad (1)$$

Note that we are adding ensemble-mean impacts ($\Delta_{Pac} + \Delta_{SO}$) to individual ensemble members of CESM2-LE. This synthetic large ensemble is meant to estimate the results for ensemble simulations that nudge the Southern Ocean and tropical Pacific SST simultaneously. It assumes the ensemble-mean impacts of pacemaker simulations are combined with the forced response and internal variability in the CESM2-LE simulations and that the impacts of pacemakers can be added linearly. A similar approach was taken in Kang et al.³³ to create synthetic ensemble simulations using SOPACE simulations.

The SUM ensemble provides valuable insights since it captures the observed SST trend in the ensemble mean (compare Fig. 3b, c). This is due to the remote impacts of the pacemaker simulations on SST trends outside the nudged area (see dashed lines in Fig. 3d, e). The SOPACE simulations affect the SST trend in the Southeast Pacific and around Antarctica³³ while the PacPACE simulations reverse the trend in the tropical Pacific and enhance the warming in the Southwest Pacific (Fig. 3e).

The SUM ensemble shows storminess trends are larger than individual pacemaker simulations both in the zonal mean and South Pacific (Fig. 4a, b) and exhibit significant storminess trends across the SH (Fig. 4e). More importantly, the trends in the SUM ensemble are not statistically different from the prescribed-SST GOGA simulations in both zonal mean (p value = 0.41) and South Pacific (p value = 0.26).

Discussion

Our study revisits the reanalysis-model SH winter storminess trend discrepancy and examines factors underlying the discrepancy (observational uncertainty, flawed observation-model comparison, and model deficiency). Expanding the number of reanalysis datasets compared to previous work⁹ shows that there is significant observational (reanalysis) uncertainty. The results show that simulations with SST prescribed from observations (AMIP) have significantly greater storminess trends, particularly in the South Pacific, than the coupled simulations (CMIP), which do not capture the observed SST trends. Our results further imply that AMIP6 model trends are not significantly different from reanalysis trends, but CMIP6 model trends are still likely discrepant due to SST trend discrepancies. Our conclusion that AMIP6 models capture storm track trends in reanalysis data differs from previous work⁹ because we used additional reanalysis and model datasets and ensured a like-for-like comparison. This also demonstrates that, among many factors, SST trend discrepancies are an important factor underlying the reanalysis-model SH winter storminess trend discrepancy.

We use the Southern Ocean and tropical Pacific pacemaker simulations to demonstrate the underestimated SH storminess trends in the coupled simulations are connected to SST trend discrepancies through well-understood mechanisms. When the coupled simulation SST trends are corrected across the Southern Ocean, the zonal-mean storminess trends strengthen significantly. The improvement of zonal-mean storminess trends involves simulating surface energy flux trends closer to reanalysis that lead to equatorward ocean energy transport and increased storminess. When the coupled simulation SST trends are corrected across the tropical Pacific, the South Pacific storminess trends strengthen significantly. The improvement of South Pacific storminess trends is a result of capturing a teleconnection trend connecting the La Nina-like SST trend to South Pacific storminess trend, consistent with previous work on interannual timescales^{42–44}. Thus, the pacemaker simulations show, when the coupled simulation SST trend discrepancies are corrected, the storminess trends strengthen significantly across the SH and are in better agreement with trends in most reanalyses and the newest reanalyses in particular. This highlights the importance of SST trend discrepancies on the SH winter circulation trends.

Our results show that observational uncertainty is large in the SH circulation trends. Since the SH exhibits a significant observation uncertainty (large spread in reanalyses trends), it is important to use all available reanalysis data as in previous work^{19,24} until there is good reason to rule out a certain product. The large spread in reanalyses trends, which is comparable to that in the large ensemble simulations, also poses a challenge for reanalysis-model comparison in the SH. Reducing the spread is important and may be helped by looking at direct observations rather than reanalysis products.

The proposed mechanisms connecting SST trends and storminess enhance our understanding of the SH winter storminess trends. Southern Ocean cooling involves equatorward ocean heat transport and storage, which strengthens the SH zonal-mean storminess by changing the equator-to-pole energy gradient. Thus, if the coupled simulations correctly represented Southern Ocean cooling and equatorward ocean heat transport and storage, then storminess trends are stronger¹². When the coupled simulations properly represent the tropical SST trends, the subtropical jet weakens, which can affect the barotropic growth rate of storminess making them stronger⁹.

The pacemaker simulations show that when the SST trend discrepancy is corrected the coupled simulation circulation trend becomes significantly stronger. Thus, it is important to understand the SST trend discrepancy and its underlying mechanisms. For the tropical Pacific, many mechanisms have been proposed^{31,32,34}. For the Southern Ocean, proposed mechanisms involving Antarctic meltwater seem to be important^{47,48} as well as multi-decadal variability⁴⁹. Understanding the mechanisms underlying the emergent responses is important for having confidence in climate model projections⁵⁰.

Methods

Storminess trends

We quantify storminess in the SH winter (June–August) using vertically integrated eddy kinetic energy (hereafter EKE), which is defined as

$$EKE = \frac{1}{g} \int_{p_t}^{p_s} u'^2 + v'^2 dp, \quad (2)$$

where g is the gravitational acceleration, p_s is the surface pressure, p_t is the pressure at the highest vertical level (Table S1), and u and v are daily-mean zonal and meridional winds, respectively. Here, the primes denote 2.5–6 day bandpass-filtered anomalies. To produce the bandpass-filtered anomalies, timeseries of u and v with 92 days of SH winter padded with 10 days at both ends are first created. We then apply a first-order Butterworth filter to the time series to obtain 2.5–6 day bandpass-filtered anomalies. We use p_s data that has the same time frequency as u and v for most datasets, but monthly-mean p_s when high-frequency data is not available. To ensure like-for-like comparison, both reanalysis and model u and v are linearly interpolated onto a common $1.5^\circ \times 1.5^\circ$ grid. Vertical integration is performed using the eight vertical levels available to the CMIP6 models (1000, 850, 700, 500, 250, 100, 50, and $p_t = 10$ hPa). While the standard is to multiply 1/2 to the right-hand side of Eq. (2), the factor is not multiplied following Chemke et al.⁹.

After quantifying storminess each year, the long-term trends are calculated using the least-squares linear regression. The statistical significance of the trend is evaluated as the 95% confidence level using the Mann-Kendall test.

When the like-for-like calculation is not performed and EKE is calculated in unprocessed time and spatial grids in Table S1, the zonal-mean SH winter storminess trends are greater in most reanalyses by about 24% (Fig. S9). This exemplifies the importance of like-for-like calculation.

Atmospheric energy transport from surface energy fluxes

To connect surface energy flux and storminess, we calculate the poleward (in SH) atmospheric energy transport (F_{SFC}) induced from surface energy flux (S , positive upward) trends. They are related as follows^{7,12,46}:

$$\nabla \cdot f_{SFC} = S, \quad (3)$$

where S is the zonal-mean surface energy flux (in $W m^{-2}$) with the global average removed (defined as positive upward), and $F_{SFC} = -2\pi a \cos \phi f_{SFC}$ (in PW), represents the poleward atmospheric energy flux induced by surface energy flux gradient at latitude ϕ where a is the Earth's radius. In F_{SFC} , $2\pi a \cos \phi$ factor represents zonal integration and a minus sign is multiplied to represent poleward transport in SH. If the ocean heat storage trends are neglected, F_{SFC} also represents equatorward ocean heat transport. The surface energy flux in ERA5 is obtained by subtracting mass-consistent atmospheric total energy flux divergence and energy tendency from the top-of-the-atmosphere radiation⁵¹. Note that other reanalyses do not have mass-consistent energy flux datasets available for this calculation.

Similarly, top-of-the-atmosphere radiation and storminess trends can be connected. In Fig. S8, $F_{TOA} = -2\pi a \cos \phi f_{TOA}$ is the poleward (in SH) atmospheric energy induced from top-of-atmosphere radiation, where $\nabla \cdot f_{TOA} = Ra_{TOA}$ and Ra_{TOA} is the net top-of-atmosphere radiation (positive downward).

Reanalysis datasets

Storminess trends are quantified in 8 reanalysis datasets (observation-based products). The 8 reanalysis datasets consist of the two latest reanalysis datasets from four different reanalysis centers. They are NCEP2⁵², MERRA1⁵³, ERA-Interim⁵⁴, CFSR/CFSv2^{55,56}, JRA-55⁵⁷, MERRA2⁵⁸, ERA5⁵⁹, and JRA3Q⁶⁰. We focus on the 40 years from 1979 to 2018 as in previous work. MERRA2 starts from 1980 so its trend is calculated from 1980. MERRA1 ends in 2015 (for JJA) so its trend is

calculated to 2015. To create daily-mean variables, we use six-hourly instantaneous data, which is the highest frequency common to all reanalysis datasets, although ERA5 and MERRA2 data are available at higher frequencies. The CFSR trend is obtained by merging CFSR (1979–2010) and CFSv2 (2011–2018) datasets.

CMIP6 and AMIP6 model simulations

Storminess trends are quantified in 26 CMIP6 model simulations¹⁴ using the historical (1979 to 2014) and SSP5-8.5 (2015 to 2018) scenarios (Table S1). We use the SSP5-8.5 scenario following previous work^{9,12}. Scenario uncertainty is a small contributor since the scenarios begin in 2015. In addition, we quantify storminess in AMIP6 simulations (1979 to 2014) in 32 models (Table S1) with observed SSTs prescribed from 1979 to 2014. We refer to the CMIP6 and AMIP6 model simulations as multi-model ensembles. The difference between the CMIP6 and AMIP6 multi-model ensembles quantifies the impact of discrepancies in SST trends in the CMIP6 models^{31,32} on storminess trends. We quantify the statistical significance of the difference between trend distributions in the multi-model ensembles using the Mann–Whitney U test⁶¹ (hereafter MW test) at the 95% level (p value < 0.05), which is a non-parametric statistical test (similar results are found using Student's t test). The number of models used in each ensemble is based on the availability of daily-mean zonal and meridional wind data on pressure levels. We use the 'r1i1p1f1' ensemble member for all models to equally weight the structural uncertainty across different models.

CESM2 large ensemble and pacemaker simulations

We use the Community Earth System Model version 2 Large Ensemble (CESM2-LE) simulations^{39,40}. The CESM2-LE simulations are an initial condition ensemble with a nominal 1-degree spatial resolution in both atmosphere and ocean. We use the first 50 simulations from this ensemble that are forced with historical radiative forcing and standard biomass burning from 1850 to 2014 consistent with CMIP6 simulations. The SST trends in the CESM2-LE simulations during this period fail to capture the observed SST trends in the tropical Pacific and Southern Ocean^{30,33} consistent with the CMIP6 multi-model ensemble.

We also use the Southern Ocean pacemaker simulations³³ (hereafter called SOPACE) and Pacific pacemaker simulations (hereafter called PacPACE, see <https://www.cesm.ucar.edu/working-groups/climate/simulations/cesm2-pacific-pacemaker> for details). The SOPACE and PacPACE simulations have 21 and 10 ensemble members, respectively. The same CMIP6 historical forcing is used for SOPACE (1979–2013) and PacPACE simulations (1880–2014). They have the same horizontal resolution as CESM2-LE. They are fully coupled except in the regions where SST anomalies (relative to observed 1880–2019 climatology) are nudged to observed SST anomalies from ERSSTv5⁶². More specifically, in SOPACE, SST anomalies are nudged to observations poleward of 40°S. In PacPACE, SST anomalies are nudged to observation within a wedge-shaped area of 20°S–20°N from the American coast to the western Pacific. We quantify the impact of pacemaking on the simulated trends as

$$\begin{aligned}\Delta_{SO} &= [\text{SOPACE}] - [\text{CESM2} - \text{LE}], \\ \Delta_{Pac} &= [\text{PacPACE}] - [\text{CESM2} - \text{LE}],\end{aligned}\quad (4)$$

where the squared brackets denote the ensemble mean³³.

Finally, we utilize AMIP-style CESM2 simulations, namely Global Ocean Global Atmosphere (GOGA) simulations, with 10 members. The GOGA simulations are forced with the same CMIP6 historical forcing from 1880 to 2014 and take observed SSTs from ERSSTv5 as boundary conditions. We quantify the impact of SST trend discrepancy by comparing the trend distributions in CESM2-LE and GOGA simulations using the MW test.

For the CESM2 simulations, EKE is calculated using the monthly-mean kinetic energy output following Kang et al.⁶³ due to data availability:

$$EKE = \frac{1}{g} \int_{p_t}^{p_s} (\overline{u^2} + \overline{v^2} - \overline{u}^2 - \overline{v}^2) dp, \quad (5)$$

where the $\overline{u^2}$ and $\overline{v^2}$ are the monthly averages of the square of u and v at each model time step (every 30 min). Since $\overline{u^2} = \overline{u^2} - \overline{u}^2$ and $\overline{v^2} = \overline{v^2} - \overline{v}^2$, this method is identical to defining primes in Eq. (2) as deviation from the monthly mean and using model time step (every 30 min) data. As such, this EKE represents the kinetic energy due to sub-monthly variations. For most reanalysis datasets except ERA5, $\overline{u^2}$ and $\overline{v^2}$ at model time step is not provided and it has to be calculated from six-hourly data. However, for ERA5, we find that the difference between calculating $\overline{u^2} + \overline{v^2}$ at model time step versus six-hourly time step is negligible (about 0.1%, Fig. S10). We extract the 8 CMIP6 pressure levels from all reanalysis datasets. The CESM2 data are interpolated from model levels to the 8 pressure levels. Then, both reanalysis and CESM2 data are linearly interpolated onto a common $1.5^\circ \times 1.5^\circ$ grid and vertically integrated over the 8 pressure levels ($p_t = 10$ hPa). Here we focus on the trend from 1979 to 2013, which is the common period for the CESM2 simulations.

Data availability

The reanalysis data used here are available online (CFSR: <https://rda.ucar.edu/datasets/ds093.0/> and <https://rda.ucar.edu/datasets/ds094.0/>, ERA5: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=form>, ERA-Interim: <https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-interim>, JRA3Q: <https://rda.ucar.edu/datasets/ds640-1/>, JRA55: <https://rda.ucar.edu/datasets/ds628.1/>, MERRA1: <https://goldsmr3.gesdisc.eosdis.nasa.gov/opensap/>, MERRA2: <https://disc.gsfc.nasa.gov/datasets?project=MERRA-2>, NCEP2: <https://psl.noaa.gov/data/gridded/data.ncep.reanalysis2.html>). The CMIP6 and AMIP6 model data are downloadable from the CMIP6 data search interface <https://esgf-node.llnl.gov/search/cmip6/>. The CESM2-LE simulations are accessible online at <https://www.cesm.ucar.edu/community-projects/lens2>. The GOGA and PacPACE simulations are available at <https://www.cesm.ucar.edu/working-groups/climate>. The SOPACE simulation data are archived at https://github.com/yuyuyaoyao/CESM2_SOPACE.

Code availability

The codes to reproduce the main results of the manuscript are accessible at <https://doi.org/10.5281/zenodo.12617000>.

Received: 2 July 2024; Accepted: 7 October 2024;

Published online: 19 October 2024

References

- Hoskins, B. J. & Hodges, K. I. A new perspective on Southern Hemisphere storm tracks. *J. Clim.* **18**, 4108–4129 (2005).
- Shaw, T. et al. Storm track processes and the opposing influences of climate change. *Nat. Geosci.* **9**, 656–664 (2016).
- Pfahl, S. & Wernli, H. Quantifying the relevance of cyclones for precipitation extremes. *J. Clim.* **25**, 6770–6780 (2012).
- Pepler, A. Record lack of cyclones in southern Australia during 2019. *Geophys. Res. Lett.* **47**, e2020GL088488 (2020).
- O’Gorman, P. A. Understanding the varied response of the extratropical storm tracks to climate change. *Proc. Natl. Acad. Sci.* **107**, 19176–19180 (2010).
- Chang, E. K., Guo, Y. & Xia, X. CMIP5 multimodel ensemble projection of storm track change under global warming. *J. Geophys. Res.* **117**, D231118 (2012).
- Shaw, T. A., Barpanda, P. & Donohoe, A. A moist static energy framework for zonal-mean storm-track intensity. *J. Atmos. Sci.* **75**, 1979–1994 (2018).

8. Priestley, M. D. & Catto, J. L. Future changes in the extratropical storm tracks and cyclone intensity, wind speed, and structure. *Weather Clim. Dyn.* **3**, 337–360 (2022).
9. Chemke, R., Ming, Y. & Yuval, J. The intensification of winter mid-latitude storm tracks in the Southern Hemisphere. *Nat. Clim. Change* **12**, 553–557 (2022).
10. Yettella, V. & Kay, J. E. How will precipitation change in extratropical cyclones as the planet warms? Insights from a large initial condition climate model ensemble. *Clim. Dyn.* **49**, 1765–1781 (2017).
11. Chang, E. K. Projected significant increase in the number of extreme extratropical cyclones in the Southern Hemisphere. *J. Clim.* **30**, 4915–4935 (2017).
12. Shaw, T. A., Miyawaki, O. & Donohoe, A. Stormier Southern Hemisphere induced by topography and ocean circulation. *Proc. Natl. Acad. Sci.* **119**, e2123512119 (2022).
13. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteor. Soc.* **93**, 485–498 (2012).
14. 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).
15. Schmidt, G. On mismatches between models and observations. <https://www.realclimate.org/index.php/archives/2013/09/on-mismatches-between-models-and-observations/> Accessed: 04 September 2023 (2013).
16. Shaw, T. A. et al. Regional climate change: consensus, discrepancies, and ways forward. *Front. Clim.* **6**, 1391634 (2024).
17. Deser, C., Phillips, A. S. & Alexander, M. A. Twentieth century tropical sea surface temperature trends revisited. *Geophys. Res. Lett.* **37**, L10701 (2010).
18. Santer, B. D. et al. Consistency of modelled and observed temperature trends in the tropical troposphere. *Int. J. Climatol.* **28**, 1703–1722 (2008).
19. Grise, K. M. et al. Recent tropical expansion: Natural variability or forced response? *J. Clim.* **32**, 1551–1571 (2019).
20. Bengtsson, L., Hagemann, S. & Hodges, K. I. Can climate trends be calculated from reanalysis data? *J. Geophys. Res.* **109**, D11111 (2004).
21. Fujiwara, M. et al. Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems. *Atmos. Chem. Phys.* **17**, 1417–1452 (2017).
22. Guo, Y. & Chang, E. K. Impacts of assimilation of satellite and rawinsonde observations on Southern Hemisphere baroclinic wave activity in the NCEP–NCAR reanalysis. *J. Clim.* **21**, 3290–3309 (2008).
23. Guo, Y., Chang, E. K. & Leroy, S. S. How strong are the Southern Hemisphere storm tracks? *Geophys. Res. Lett.* **36**, L22806 (2009).
24. Martineau, P., Behera, S. K., Nonaka, M., Nakamura, H. & Kosaka, Y. Seasonally dependent increases in subweekly temperature variability over Southern Hemisphere landmasses detected in multiple reanalyses. *Weather Clim. Dyn.* **5**, 1–15 (2024).
25. Deser, C. et al. Insights from Earth system model initial-condition large ensembles and future prospects. *Nat. Clim. Change* **10**, 277–286 (2020).
26. Jain, S. et al. Importance of internal variability for climate model assessment. *npj Clim. Atmos. Sci.* **6**, 68 (2023).
27. Po-Chedley, S., Thorsen, T. J. & Fu, Q. Removing diurnal cycle contamination in satellite-derived tropospheric temperatures: Understanding tropical tropospheric trend discrepancies. *J. Clim.* **28**, 2274–2290 (2015).
28. Santer, B. D. et al. Comparing tropospheric warming in climate models and satellite data. *J. Clim.* **30**, 373–392 (2017).
29. Chang, E. K., Lee, S. & Swanson, K. L. Storm track dynamics. *J. Clim.* **15**, 2163–2183 (2002).
30. Wills, R. C., Dong, Y., Proistosescu, C., Armour, K. C. & Battisti, D. S. Systematic climate model biases in the large-scale patterns of recent sea-surface temperature and sea-level pressure change. *Geophys. Res. Lett.* **49**, e2022GL100011 (2022).
31. Lee, S. et al. On the future zonal contrasts of equatorial Pacific climate: Perspectives from observations, simulations, and theories. *npj Clim. Atmos. Sci.* **5**, 82 (2022).
32. Seager, R., Henderson, N. & Cane, M. Persistent discrepancies between observed and modeled trends in the tropical Pacific Ocean. *J. Clim.* **35**, 4571–4584 (2022).
33. Kang, S. M. et al. Global impacts of recent southern ocean cooling. *Proc. Natl. Acad. Sci.* **120**, e2300881120 (2023).
34. Watanabe, M. et al. Possible shift in controls of the tropical Pacific surface warming pattern. *Nature* **630**, 315–324 (2024).
35. Dong, Y., Armour, K. C., Battisti, D. S. & Blanchard-Wrigglesworth, E. Two-way teleconnections between the Southern Ocean and the tropical Pacific via a dynamic feedback. *J. Clim.* **35**, 6267–6282 (2022).
36. Purich, A., Cai, W., England, M. H. & Cowan, T. Evidence for link between modelled trends in Antarctic sea ice and underestimated westerly wind changes. *Nat. Commun.* **7**, 10409 (2016).
37. Cox, T., Donohoe, A., Armour, K. C., Frierson, D. M. & Roe, G. H. Trends in atmospheric heat transport since 1980. *J. Clim.* **37**, 1539–1550 (2024).
38. Gillett, N. P. et al. The detection and attribution model intercomparison project (DAMIP v1. 0) contribution to CMIP6. *Geosci. Model. Dev.* **9**, 3685–3697 (2016).
39. Danabasoglu, G. et al. The community earth system model version 2 (CESM2). *J. Adv. Model. Earth Syst.* **12**, e2019MS001916 (2020).
40. Rodgers, K. B. et al. Ubiquity of human-induced changes in climate variability. *Earth Syst. Dyn.* **12**, 1393–1411 (2021).
41. Armour, K. C., Marshall, J., Scott, J. R., Donohoe, A. & Newsom, E. R. Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. *Nat. Geosci.* **9**, 549–554 (2016).
42. Seager, R., Harnik, N., Kushnir, Y., Robinson, W. & Miller, J. Mechanisms of hemispherically symmetric climate variability. *J. Clim.* **16**, 2960–2978 (2003).
43. Nakamura, H., Sampe, T., Tanimoto, Y. & Shimpo, A. Observed associations among storm tracks, jet streams and midlatitude oceanic fronts. *AGU Geophys. Monogr. Ser.* **147**, 329–345 (2004).
44. Ashok, K., Nakamura, H. & Yamagata, T. Impacts of ENSO and Indian Ocean dipole events on the Southern Hemisphere storm-track activity during austral winter. *J. Clim.* **20**, 3147–3163 (2007).
45. Barpanda, P. & Shaw, T. A. Surface fluxes modulate the seasonality of zonal-mean storm tracks. *J. Atmos. Sci.* **77**, 753–779 (2020).
46. Kang, S. M., Held, I. M., Frierson, D. M. & Zhao, M. The response of the ITCZ to extratropical thermal forcing: Idealized slab-ocean experiments with a GCM. *J. Clim.* **21**, 3521–3532 (2008).
47. Bronselaer, B. et al. Change in future climate due to Antarctic meltwater. *Nature* **564**, 53–58 (2018).
48. Roach, L. A. et al. Winds and meltwater together lead to Southern Ocean surface cooling and sea ice expansion. *Geophys. Res. Lett.* **50**, e2023GL105948 (2023).
49. Chung, E.-S. et al. Antarctic sea-ice expansion and Southern Ocean cooling linked to tropical variability. *Nat. Clim. Change* **12**, 461–468 (2022).
50. Shaw, T. A. Mechanisms of future predicted changes in the zonal mean mid-latitude circulation. *Curr. Clim. Change Rep.* **5**, 345–357 (2019).
51. Mayer, J., Mayer, M. & Haimberger, L. Consistency and homogeneity of atmospheric energy, moisture, and mass budgets in ERA5. *J. Clim.* **34**, 3955–3974 (2021).
52. Kanamitsu, M. et al. Ncep–doe amip–ii reanalysis (r-2). *Bull. Am. Meteor. Soc.* **83**, 1631–1644 (2002).
53. Rienecker, M. M. et al. MERRA: NASA’s modern-era retrospective analysis for research and applications. *J. Clim.* **24**, 3624–3648 (2011).

54. Dee, D. P. et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597 (2011).
55. Saha, S. et al. The NCEP climate forecast system reanalysis. *Bull. Am. Meteor. Soc.* **91**, 1015–1058 (2010).
56. Saha, S. et al. The NCEP climate forecast system version 2. *J. Clim.* **27**, 2185–2208 (2014).
57. Kobayashi, S. et al. The JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan* **93**, 5–48 (2015).
58. Gelaro, R. et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454 (2017).
59. Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
60. Kosaka, Y. et al. The JRA-3Q reanalysis. *J. Meteor. Soc. Japan* **102**, 49–109 (2024).
61. Mann, H. B. & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Annals Math. Stat.* **18**, 50–60 (1947).
62. Huang, B. et al. Extended reconstructed sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons. *J. Clim.* **30**, 8179–8205 (2017).
63. Kang, J. M., Shaw, T. A. & Sun, L. Arctic sea ice loss weakens Northern Hemisphere summertime storminess but not until the late 21st century. *Geophys. Res. Lett.* **50**, e2022GL102301 (2023).

Acknowledgements

J.M.K. and T.A.S. are supported by the National Oceanic and Atmospheric Administration award NA23OAR4310597. I.R.S. is supported by the National Science Foundation (NSF) National Center of Atmospheric Research, which is a major facility sponsored by the National Science Foundation under the Cooperative Agreement 1852977. We acknowledge the Climate Variability and Change Working Group at the NSF National Center of Atmospheric Research for running the Pacific pacemaker simulations and GOGA simulations and making them available. The Southern Ocean pacemaker simulations are supported by the high-performance computing cluster of State Key Laboratory of Satellite Ocean Environment Dynamics.

Author contributions

J.M.K. performed analysis and wrote the manuscript together with T.A.S., I.R.S., S.M.K., and Y.Y.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at

<https://doi.org/10.1038/s41612-024-00801-3>.

Correspondence and requests for materials should be addressed to Joonsuk M. Kang.

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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2024