Monthly Sea-Surface Temperature, Sea Ice, and Sea-Level Pressure over 1850–2023 from Coupled Data Assimilation

Vincent T. Cooper,[a] Gregory J. Hakim,[a] Kyle C. Armour,[a,b]

[a] *Department of Atmospheric and Climate Science, University of Washington, Seattle, WA, USA*

[b] *School of Oceanography, University of Washington, Seattle, WA, USA*

Corresponding author: Vincent T. Cooper, vcooper@uw.edu

Early Online Release: This preliminary version has been accepted for publication in *Journal of Climate*, may be fully cited, and has been assigned DOI 10.1175/JCLI-D-25-0021.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

ABSTRACT: Historical observations of Earth's climate underpin our knowledge and predictions of climate variability and change. However, the observations are incomplete and uncertain, and existing datasets based on these observations typically do not assimilate observations simultaneously across different components of the climate system, yielding inconsistencies that limit understanding of coupled climate dynamics. Here we use coupled data assimilation, which synthesizes observational and dynamical constraints across all climate fields simultaneously, to reconstruct globally resolved sea-surface temperature (SST), near-surface air temperature (T), sea-level pressure (SLP), and sea-ice concentration (SIC), over 1850–2023. We use a Kalman filter and forecasts from an efficient emulator (Linear Inverse Model; LIM) to assimilate observations of SST, land T, marine SLP, and satellite-era SIC. We account for model error by training LIMs on eight CMIP6 models, and we use the LIMs to generate eight independent reanalyses with 200 ensemble members, yielding 1600 total members. Key findings in the Tropics include post-1980 trends in the Walker circulation that are consistent with past variability, whereas the tropical SST contrast (the difference between warmer and colder SSTs) shows a distinct strengthening since 1975. ENSO amplitude exhibits substantial low-frequency variability and a local maximum in variance over 1875–1910. In polar regions, we find a muted cooling trend in the Southern Ocean post-1980 and substantial uncertainty. Changes in Antarctic sea ice are relatively small between 1850 and 2000, while Arctic sea ice declines by 0.5 ± 0.1 (1σ) million km^2 during the 1920s.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

SIGNIFICANCE STATEMENT: The key advance in our reconstruction is that the ocean, atmosphere, and sea ice are dynamically consistent with each other and with observations across all components, thus forming a true climate reanalysis. Existing climate datasets are typically derived separately for each component (e.g., atmosphere, ocean, and sea ice), leading to spurious trends and inconsistencies in coupled climate variability. We use coupled data assimilation to unify observations and coupled dynamics across components. We combine forecasts from climate models with observations from ocean vessels and weather stations to produce monthly state estimates spanning 1850–2023 and a novel quantification of globally resolved uncertainty. This reconstruction provides insights into historical variability and trends while motivating future efforts to reduce uncertainties in the climate record.

1. Introduction

The historical record (c. 1850–present) is central to our understanding of climate variability and Earth's response to anthropogenic forcings, but we have yet to fully extract the available information from instrumental data. Observations of sea-surface temperature (SST), near-surface air temperature (T), and sea-level pressure (SLP) from ships of opportunity and weather stations are noisy, sparse, and vary over time, which adds an incomplete-data problem (Schneider 2001) to analyses of climate variability and change that cannot be avoided and should not be ignored.

To prepare observations for climate analysis, data sources must first be homogenized (e.g., Kent and Kennedy 2021; Chan and Huybers 2019; Chan et al. 2023; Karl et al. 2015; Hausfather et al. 2017; Cowtan et al. 2018), and then the missing values must be imputed. Imputation is typically statistical, employing pattern-based methods (including empirical orthogonal functions; EOFs) derived from recent decades or kriging, and does not involve dynamical constraints (e.g., Kaplan et al. 1998; Rohde et al. 2013; Cowtan and Way 2014; Hirahara et al. 2014; Huang et al. 2017; Kadow et al. 2020; Vaccaro et al. 2021). Furthermore, when values are imputed for different climate fields, e.g., SST and SLP, there are no dynamical constraints ensuring that the coupled fields are physically consistent. Imputation and homogenization can have pronounced impacts on assessments of the climate sensitivity to increasing greenhouse gases (e.g., Sherwood et al. 2020; Forster et al. 2021; Modak and Mauritsen 2023), efforts to distinguish internal variability from forced climate change (e.g., Hegerl et al. 2019; Wills et al. 2020), understanding of atmosphere–

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

ocean variability (e.g., Battisti et al. 2019), and evaluation of climate models (e.g., Wills et al. 2022; Simpson et al. 2025). Here we apply a different approach to solve the incomplete-data problem: we use coupled data assimilation to impose observational and dynamical constraints across all climate fields simultaneously, ensuring that the full state estimate is internally consistent.

SST patterns play a ubiquitous role in regulating climate variability (e.g., Bjerknes 1969; Barsugli and Battisti 1998; Alexander et al. 2002; Deser et al. 2010a; Newman et al. 2016; Czaja et al. 2019; Capotondi et al. 2023), radiative feedbacks (e.g., Armour et al. 2013; Andrews et al. 2015; Zhou et al. 2016; Ceppi and Gregory 2017; Andrews and Webb 2018; Dong et al. 2019; Fueglistaler 2019; Andrews et al. 2022; Salvi et al. 2023; Cooper et al. 2024), and the hydrologic cycle (Hastenrath and Greischar 1993; Xie et al. 2010; Hoerling et al. 2010; Chadwick et al. 2014; Lehner et al. 2018; Siler et al. 2019; Cook et al. 2022; Kuo et al. 2023; Seager et al. 2023). There are a variety of recent (c. 1980–present) climate phenomena tied to SSTs that seem either unprecedented or unremarkable depending on what we deem to be natural variability, and this interpretation of recent trends relies on the incomplete and brief instrumental record (e.g., Wunsch 1999). In the tropical Pacific, the zonal SST gradient has strengthened (Solomon and Newman 2012; Coats and Karnauskas 2017; Lee et al. 2022; Watanabe et al. 2024), with cooling in the East Pacific and warming in the West Pacific that has coincided with an apparent strengthening of the Walker circulation post-1980 (e.g., L'Heureux et al. 2013; McGregor et al. 2014; Watanabe et al. 2023; Heede and Fedorov 2023) and distinct tropospheric temperature trends (Flannaghan et al. 2014; Fueglistaler 2019; Po-Chedley et al. 2021). At the poles, the Arctic has warmed rapidly since 1980 with substantial loss of sea ice (Dörr et al. 2023; England et al. 2021; Notz and SIMIP Community 2020), while the Southern Ocean has cooled with an overall expansion of sea ice—until 2015, after which the Southern Ocean has shown rapid warming and sea ice loss (Fan et al. 2014; Stuecker et al. 2017; Fogt et al. 2022; Espinosa et al. 2024; Roach and Meier 2024; Zhang and Li 2023; Turner et al. 2022; Dong et al. 2023; Suryawanshi et al. 2023; Bonan et al. 2024).

A major challenge for the climate dynamics community is understanding the causes of these observed changes as well as the apparent yet debated inability of our state-of-the-art coupled climate models to replicate them (e.g., Wills et al. 2022; Dong et al. 2021; Rugenstein et al. 2023; Seager et al. 2022; Olonscheck et al. 2020; Chung et al. 2019; Watanabe et al. 2021; Roach et al. 2020; Notz and SIMIP Community 2020; Chemke et al. 2022; J. Kang et al. 2024). Progress on

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

this endeavor requires robustly quantifying observational uncertainties and placing recent changes in historical context with reliable reconstructions of past climate changes. For example, are the post-1980 trends in tropical SST gradients, the Walker circulation, and polar climates unique over the historical record, or have such changes occurred often due to internal climate variability?

Existing SST datasets designed for climate analysis use a variety of statistical interpolation methods. These methods have been recently summarized in Modak and Mauritsen (2023) and Lewis and Mauritsen (2021) and described in detail in a review by Kent and Kennedy (2021), which also explains the extensive efforts to homogenize time-varying sources of in situ data. To assess the atmospheric response to SST and sea-ice concentration (SIC) over the historical record in atmospheric general circulation models (i.e., in AMIP-type simulations; Eyring et al. 2016; Webb et al. 2017), complete coverage and monthly resolution of SST/SIC is required. Combined SST/SIC datasets for this purpose include the 1870–2022 PCMDI/AMIP-II boundary condition (Hurrell et al. 2008) used as the standard for CMIP6, 1854–present NOAA ERSSTv5 (Huang et al. 2017), Met Office Hadley Centre's 1870–present HadISST1 (Rayner et al. 2003) and 1850–2010 HadISST2.1 (no longer maintained; Titchner and Rayner 2014), and the Japanese Meteorological Agency's 1850–present COBE-SST2 (Hirahara et al. 2014). Kaplan et al. (1998) developed a landmark SST analysis using optimal interpolation, and since then the incomplete-data problem has been addressed using kriging (Cowtan and Way 2014), Markov random graphs (Vaccaro et al. 2021), and machine learning (Kadow et al. 2020) to impute hybrid air-sea surface temperatures over land and ocean.

Figure 1 depicts the time-evolving observing network of in situ SST measurements in HadSST4 (Kennedy et al. 2019). As motivation for this study, we illustrate the spread (1σ) across existing datasets (HadISST1, HadISST2.1, ERSSTv5, COBE-SST2, and AMIP-II) in their preindustrial-baseline SST (mean anomaly over years 1870–1899) and the spread in their SST trends from 1900–1979 and 1980–2010. We separate the satellite era (c. 1980–present) from the earlier warming because of the variety of studies highlighting and questioning the peculiarity of recent SST trends (e.g., Fueglistaler and Silvers 2021; Andrews et al. 2022; Lewis and Mauritsen 2021). The spatial pattern of uncertainty is influenced by varying methods of imputation, homogenization of data sources, and representativeness error in using point observations as estimates of grid-scale means. Even after 1980, the data coverage over the Southern Ocean and southeast Pacific is notably

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

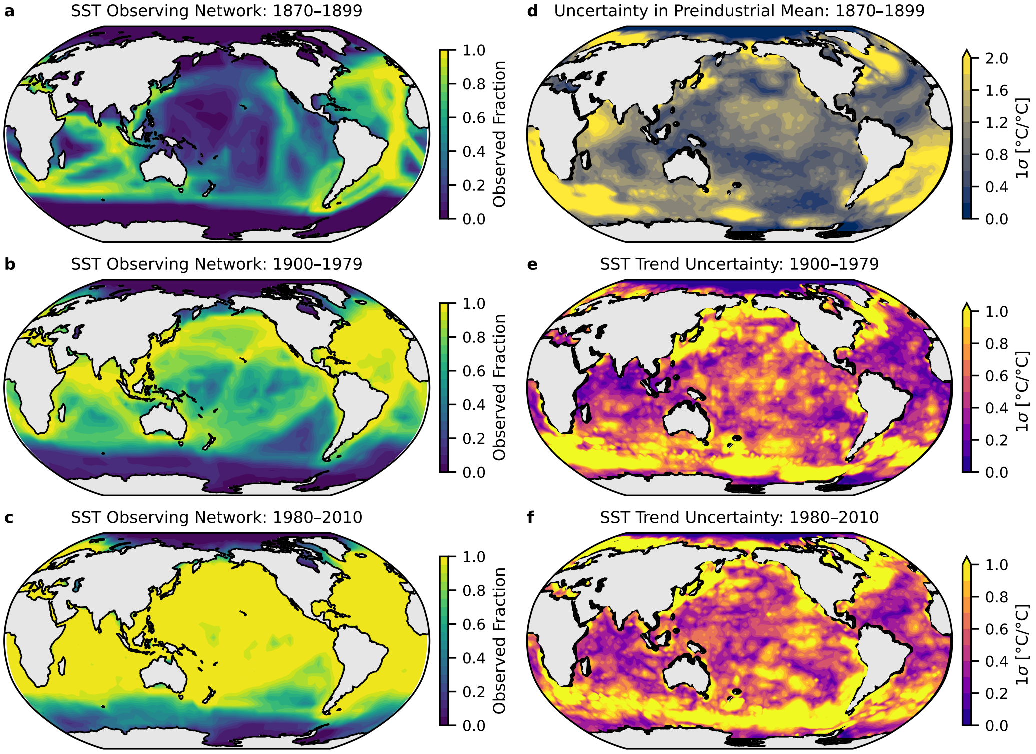

FIG. 1. **Historical observing network and SST uncertainty in pre-existing infilled datasets.** **(a–c)** Fraction of months with in situ data for SST over three time periods in HadSST4, where 1.0 indicates data in every month during the period. **(d)** Illustration of systematic uncertainty in normalized pattern of preindustrial-mean SST anomalies across existing infilled datasets, calculated as the sample standard deviation (1σ) of the 1870–1899 mean anomalies across HadISST1, HadISST2.1, ERSSTv5, PCMDI/AMIP-II, and COBE-SST2, relative to their 1961–1990 climatologies; local anomalies are divided by global-mean anomalies (60°S–60°N) to highlight uncertainty in spatial patterns. **(e–f)** Illustration of systematic uncertainty in patterns of SST trends, calculated as the 1σ of local trends across the same datasets in panel **d**; local SST trends are first divided by the global-mean SST trends (60°S–60°N) to highlight uncertainty in the patterns, and local values greater than 1.0 indicate that the local 1σ is greater than the global-mean trend. Note different colorbars in panels **d–f**.

far from complete, and the inter-dataset differences in those regions are substantial in recent decades (Figure 1c,f).

Atmospheric reanalyses address the incomplete-data problem with data assimilation, which uses a weather model's dynamics to constrain the atmospheric state. Data assimilation (DA) broadly describes the collection of methods that synthesize model forecasts with sparse and noisy observations, producing posterior analyses and uncertainties that are subject to the dynamical constraints of the model. DA is computationally intensive, hence existing reanalyses only assimilate atmospheric observations and only apply dynamical constraints to the atmospheric component, meaning that the SST and SIC boundary conditions are prescribed a priori in, for example, ERA5 (Hersbach et al. 2020; Soci et al. 2024), JRA-55 and JRA-3Q (Kobayashi et al. 2015; Kosaka et al. 2024), NOAA-CIRES-DOE's 20th Century Reanalysis (Compo et al. 2011; Slivinski et al. 2019), NCEP/NCAR Reanalysis (Kalnay et al. 1996), MERRA-2 (Gelaro et al. 2017), and ModE-RA (Franke et al. 2017; Valler et al. 2024).

Coupled atmosphere–ocean reanalysis remains a frontier and formidable challenge in climate research. ECMWF's coupled DA program, CERA-20C (Laloyaux et al. 2018), is now inactive, and ECMWF no longer hosts its output. NCEP's CFSR made a major advance (Saha et al. 2010), which assimilated observations into atmosphere and ocean components separately, and incorporated coupling by using a coupled model during the forecast step—this process is known as "weakly coupled" DA. The UFS-Replay dataset (NOAA 2024), employs a weakly coupled "replay" approach (Orbe et al. 2017), in which the coupled UFS model is nudged toward the existing ERA5 atmospheric reanalysis and ORAS5 ocean reanalysis (Zuo et al. 2019). In this study, we will use "strongly coupled" DA, which (i) ensures that the coupled atmosphere-ocean-ice state is internally consistent and (ii) synthesizes observational and dynamical constraints across each component simultaneously.

To circumvent the computational obstacles associated with DA in fully coupled models, lightweight DA methods have been developed primarily for paleoclimate reconstruction, as reviewed by Tierney et al. (2025a). The "offline" DA method uses a static, uninformed prior from pre-existing model output (e.g., Hakim et al. 2016; Franke et al. 2017; Steiger et al. 2014, 2018; Samakinwa et al. 2021; Tierney et al. 2020; Osman et al. 2021; Annan et al. 2022; Smerdon et al. 2023; Valler et al. 2024; Tierney et al. 2025b). "Online" methods use a time-evolving prior that is informed by the previous initial conditions produced by data assimilation, thus retaining memory of past observations. Online DA requires integrating a forecast model after each assimilation step, which is the main computational bottleneck.

Data-driven approaches that emulate climate models can overcome the computational bottleneck. The linear inverse model (LIM) has been tested in annual-mean DA with proxies over the last millennium (Perkins and Hakim 2021) and for subseasonal forecasting (Hakim et al. 2022). LIMs have been applied to study dynamics and predictability of the El Niño–Southern Oscillation (ENSO) (e.g., Penland and Sardeshmukh 1995; Shin et al. 2021; Vimont et al. 2014; Lou et al. 2020; Kido et al. 2023), meridional modes (Vimont 2012), global surface temperatures (Newman 2013), SSTs in the North Atlantic (e.g., Zanna 2012) and North Pacific (e.g., Newman 2007; Newman et al. 2016; Zhao et al. 2024), the Pacific-North American pattern (Henderson et al. 2020), hydroclimate (Coats et al. 2020; Tseng et al. 2021), and sea ice (Brennan et al. 2023). LIMs are computationally efficient, enabling coupled assimilation of observations across Earth system components, e.g., pressure observations in the atmosphere and SST observations in the ocean can each inform both SST and SLP in coupled DA. Combining LIMs with data assimilation presents an opportunity to constrain and quantify uncertainty in the historical climate record.

Here we use coupled data assimilation to reconstruct monthly and globally resolved SST, near-surface air temperature (T), SLP, and SIC over 1850–2023. The novelty of our approach compared to past reanalyses is that we constrain all climate fields simultaneously with (i) coupled dynamics and (ii) observations across climate components. Our DA method is made computationally tractable by efficient emulators (LIMs), which are trained on eight CMIP6 models and capture the essential dynamics for monthly reanalysis. We combine forecasts from LIMs with a Kalman filter to produce a coupled reconstruction with time-varying uncertainty quantification. Section 2 describes methods and data, including LIMs, data assimilation, validation with an out-of-sample pseudo-reconstruction, observations, and comparison datasets. Section 3 presents the historical reconstruction. Section 4 discusses the implications of the results for interpreting climate variability and change and the caveats of the method. Section 5 presents the conclusions.

2. Methods and data

In this section, we describe the reconstruction method, validation, and data sources. The reconstruction of monthly means consists of (i) a monthly forecast, for which we use LIMs that emulate eight CMIP6 models, and (ii) data assimilation in every month, for which we use the classic Kalman filter (Kalman 1960; Kalnay 2003). We validate the method with a pseudo-reconstruction of a climate model's 1850–2014 historical simulation (MPI-ESM1-2-HR), from which we draw observations that mimic the true observing network.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

a. Linear inverse models

Anomalies around an equilibrium state in the nonlinear climate system can be approximated as a stochastically forced, linear dynamical system (e.g., Hasselmann 1976; Penland and Sardeshmukh 1995; Penland 1996):

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}\mathbf{x} + \mathbf{S}\eta, \tag{1}$$

where $\mathbf{x}$ is a state vector of N principal components of SST, T, SLP, and SIC, $\mathbf{L}$ is an $N \times N$ linear operator representing the deterministic dynamics, and $\mathbf{S}\eta$ approximates the unresolvable nonlinear dynamics as stochastic forcing with an $N \times N$ noise-amplitude matrix, $\mathbf{S}$, and a vector, η, of independent, Gaussian white noise with unit variance and length N.

LIMs typically assume stationary statistics, but Shin et al. (2021) extend the LIM framework to include monthly variations in the dynamics. The monthly, or "cyclostationary" LIM, has been applied to ENSO (Shin et al. 2021; Vimont et al. 2022; Kido et al. 2023). We build on this recent work and use cyclostationary LIMs to model global SST, T, SLP, and SIC. We use the fixed-phase approach (Shin et al. 2021; OrtizBeviá 1997) to train the 12 $\mathbf{L}_j$ operators in the cyclostationary LIM, where j indicates the month:

$$\mathbf{L}_j = \tau^{-1}\log[\mathbf{C}_j(\tau)\mathbf{C}_j(0)^{-1}], \quad \text{for } j = 1, 2, \ldots, 12. \tag{2}$$

$\mathbf{C}_j(\tau)$ and $\mathbf{C}_j(0)$ are the τ-lag and zero-lag covariance matrices of $\mathbf{x}$ for month j, and $\tau = 1$ month in all of the following equations. The stochastic amplitude matrices, $\mathbf{S}_j$, are estimated from the fluctuation-dissipation relation of Equation (1) (Penland and Matrosova 1994),

$$\frac{d\mathbf{C}_j(0)}{dt} = \mathbf{L}_j\mathbf{C}_j(0) + \mathbf{C}_j(0)\mathbf{L}_j^T + \mathbf{Q}_j, \tag{3}$$

where $\mathbf{Q}_j = \mathbf{S}_j\mathbf{S}_j^T$. We follow Shin et al. (2021) in estimating the cyclostationary $\mathbf{Q}_j$ as

$$\mathbf{Q}_j = \frac{\mathbf{C}_{j+1}(0) + \mathbf{C}_{j-1}(0)}{2\Delta t} - [\mathbf{L}_j\mathbf{C}_j(0) + \mathbf{C}_j(0)\mathbf{L}_j^T], \tag{4}$$

with $\Delta t = 1$ month. Before computing $\mathbf{L}_j$ and $\mathbf{Q}_j$, we take the 3-month running means of $\mathbf{C}_j(\tau)$ and $\mathbf{C}_j(0)$, e.g., we estimate $\mathbf{C}_j(\tau) \approx \langle \mathbf{C}_{j-1}(\tau), \mathbf{C}_j(\tau), \mathbf{C}_{j+1}(\tau) \rangle$, where angle brackets denote an

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

equal-weighted average (Shin et al. 2021). As in previous LIM studies (e.g., Penland 1996), we remove any negative eigenvalues in $\mathbf{Q}_j$ and rescale remaining eigenvalues to conserve the original variance.

The LIM produces forecasts at lead $\tau = 1$ month from integrating (1) in time as

$$\mathbf{x}(t+\tau) = \mathbf{G}_j\mathbf{x}(t) + \mathbf{n}, \tag{5}$$

where $\mathbf{G}_j = \exp(\mathbf{L}_j\tau) = \mathbf{C}_j(\tau)\mathbf{C}_j(0)^{-1}$. The integrated stochastic term, $\mathbf{n}$, equals zero in a deterministic forecast, such as the prior-mean forecast in the Kalman filter as described below.

The forecast equation for the error covariance, assuming no correlation between error and state, is

$$\mathbf{P}(t+\tau) = \mathbf{G}_j\mathbf{P}(t)\mathbf{G}_j^T + \mathbf{N}_j(\tau). \tag{6}$$

We forecast the full covariance matrix with the LIM, instead of estimating it from ensemble members, because this approach is exact for a given LIM. It is equivalent to using an infinite ensemble. To solve for $\mathbf{N}_j(\tau)$, we extend the logic that applies to the stationary LIM (Hakim et al. 2022; Penland 1989) for the cyclostationary case. Equation (6) must be valid for any month's initial condition, including $\mathbf{C}_j(0)$, from which the monthly forecast must arrive at $\mathbf{C}_{j+1}(0)$ because the statistics are cyclostationary, therefore:

$$\mathbf{N}_j(\tau) = \mathbf{C}_{j+1}(0) - \mathbf{G}_j\mathbf{C}_j(0)\mathbf{G}_j^T. \tag{7}$$

We train separate LIMs to emulate the following eight CMIP6 models: CESM2, GFDL-ESM4, HadGEM3-GC3.1-LL, SAM0-UNICON, UKESM1.0-LL, NorESM2-LM, EC-Earth3, and E3SM-2-0. Our selection of models is informed by Lou et al. (2023), which found that this subgroup performs best in an analog method for ENSO forecasting, although we make two changes: we remove HadGEM3-GC3.1-MM to prevent having two versions of HadGEM3, and we substitute E3SMv2.0 (Qin et al. 2024) for CIESM because of issues simulating sea ice in CIESM (Lin et al. 2020). For training data, we use preindustrial-control simulations with historical (1850–2014) simulations appended (summary in Appendix A). LIMs are trained separately for each model using monthly mean anomalies, and each LIM has a minimum of 665 years of training data (500+

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

preindustrial and 165 historical years). While approximately 100 years of training data is sufficient for a Tropics-only cyclostationary LIM (Shin et al. 2021), global LIMs require a longer record. Thus long preindustrial simulations are essential for training, and we find that appending the historical simulations expands the footprint of regions with nonzero SIC variability in the training data, which improves reconstruction of SIC.

We regrid all training data bilinearly to 2° resolution (96×144 latitude-longitude grid). For consistency with observations, which are expressed as anomalies relative to a 1961–1990 climatology, we remove the mean and climatology in each gridcell calculated over 1961–1990 for each model. Separately for each model and state variable, we compute EOFs area-weighted by the square-root of the cosine of latitude for SST, T, SLP, Northern Hemisphere (NH) SIC, and Southern Hemisphere (SH) SIC. We retain approximately 85% of each field's variance in the truncated state. We form each model's standardized state vector from its principal components, $\mathbf{x}_k$, as:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_{SST}/\sigma_{SST} \\ \mathbf{x}_{T}/\sigma_{T} \\ \mathbf{x}_{SLP}/\sigma_{SLP} \\ \mathbf{x}_{SIC_{NH}}/\sigma_{SIC_{NH}} \\ \mathbf{x}_{SIC_{SH}}/\sigma_{SIC_{SH}} \end{bmatrix},$$

where σ_k^2 is the retained variance after EOF truncation of field k. We use the standardized state vectors $\mathbf{x}$ to compute covariance matrices for each model, and we project into and out of the LIM basis by storing the EOFs and scale factors, σ_k, for each field. Each LIM is run independently in parallel through the data assimilation framework.

b. Data assimilation

Given a prior forecast of the state's monthly mean $\mathbf{x}_f$ and error covariance $\mathbf{P}_f$, we assimilate observations to produce the posterior analysis $\mathbf{x}_a$ and $\mathbf{P}_a$ using the Kalman filter:

$$\mathbf{x}_a = \mathbf{x}_f + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_f), \tag{8}$$

$$\mathbf{P}_a = [\mathbf{I} - \mathbf{K}\mathbf{H}]\mathbf{P}_f, \tag{9}$$

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

$$\mathbf{K} = \mathbf{P}_f \mathbf{H}^T [\mathbf{H}\mathbf{P}_f \mathbf{H}^T + \mathbf{R}]^{-1}, \tag{10}$$

where $\mathbf{K}$ is the Kalman gain, $\mathbf{y}$ is the vector of observations, $\mathbf{H}$ is the linear observation operator, and $\mathbf{R}$ is the observation error covariance. After solving (8–10) for a given month, we forecast the next month from (5), with $\mathbf{n} = 0$, and (6).

Our method is "strongly coupled online DA," where "strongly coupled" means that we assimilate observations concurrently across the atmosphere–ocean–ice system, and all fields influence each other through cross-component covariances. "Online" means that we use a forecast model with the previous assimilation step's initial conditions to inform the prior. Because this method uses the classic Kalman filter and propagates $\mathbf{P}_f$ exactly, we avoid the sample error and localization issues that arise when estimating $\mathbf{P}_f$ in an ensemble Kalman filter (Evensen 1994; Houtekamer and Zhang 2016). However, ensemble-member trajectories are needed to analyze statistics of temporal variability, and this variability must be constrained by dynamics rather than sampled independently (Emile-Geay et al. 2024). We solve this problem with a modified version of the ensemble Kalman filter, described subsequently, that has no impact on the mean or covariance, (8–9), but rather simply provides sample estimates from the posterior distribution.

We generate ensemble members using the perturbed-observations version of the ensemble Kalman filter (Houtekamer and Mitchell 1998; Burgers et al. 1998), except we use the exact prior covariance forecast from the classic Kalman filter (6, 9). For each LIM, we initialize 200 ensemble members in January 1850 with random draws from a multivariate-normal distribution with covariance $\mathbf{C}_1(0)$. Each ensemble member is updated using (8), with $\mathbf{x}_f^n$ corresponding to ensemble member n in place of the ensemble mean, and $\mathbf{y}^n$ is a multivariate-normal random draw of the observations with mean $\mathbf{y}$ and covariance $\mathbf{R}$. After the assimilation, each $\mathbf{x}_a^n$ is advanced to the next month using (5). The noise term in (5), $\mathbf{n}$, becomes a random draw from $\mathbf{N}_j(\tau)$ in (7) for each ensemble member. Because our LIMs are built to forecast monthly means, we can draw from the distributions of the monthly statistics rather than stochastically integrating (Penland and Matrosova 1994) each ensemble member.

An additional benefit of the ensemble is that we can propagate temporally correlated observation errors that are associated with uncertainties in bias corrections. For example, HadSST4 (described below) provides a 200-member ensemble of monthly SST observations to represent temporally correlated errors (Kennedy et al. 2019). To incorporate these errors, we let **y** vary across the ensemble members, but each of our 200 ensemble members $\mathbf{x}^n$ is paired at every timestep with the corresponding ensemble member n from the HadSST4 ensemble.

c. Observations

We use four sources of observations corresponding to each of the four state variables (SST, T, SLP, SIC). All observations are anomalies relative to a 1961–1990 climatology, which is the period chosen by Kennedy et al. (2019) and Osborn et al. (2021).

SST observations are from HadSST4 version 4.0.1.0 (Kennedy et al. 2019), provided by the Met Office Hadley Centre on a 5° × 5° grid. HadSST4 quality controls and corrects biases in the in situ measurements from ICOADS 3.0.0 (1850–2014) and ICOADS 3.0.1 (2015–present), the central database of ship records (Freeman et al. 2017). HadSST4 provides non-infilled data as monthly means spanning 1850–present, and ship coverage varies substantially over time (Figure 1). Measurement and sampling errors are provided for every gridcell and month, and error covariance matrices are provided that estimate the spatially correlated errors. We include these sources of error in **R**. Temporally correlated errors from uncertain bias corrections are estimated with a 200-member ensemble of observations, and we account for these errors with our ensemble DA method, described in Section 2b.

Observations of near-surface air temperature (T) over land are from CRUTEM5 version 5.0.2.0 (Osborn et al. 2021). The weather-station data is quality controlled, bias-corrected, and provided as non-infilled monthly means with error estimates on a 5° × 5° grid. We include CRUTEM5's time-varying measurement and sampling errors in **R**.

SLP observations are from ICOADS Enhanced Release 3.1 for 1850–2014 and Release 3.0.2, for 2015–2023 (Freeman et al. 2017), which only includes marine data. Ideally, we would assimilate terrestrial SLP observations, but the ISPD dataset (Cram et al. 2015) of surface pressures does not have a homogenized product available that combines data at various elevations into a gridded dataset of monthly means. The lack of direct constraints on terrestrial SLP is a limitation of our reconstruction; hence, we focus our analysis on marine SLP. ICOADS marine SLP data are provided as monthly means on a 2° × 2° grid, along with the number of observations, n_{obs}, in each month and the intra-month standard deviation, s, of the observations in each gridcell. The

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

baseline climatology for anomalies is from Hersbach et al. (2020). There are a large number of SLP observations due to the finer grid of ICOADS compared to HadSST4. We eliminate observations with $n_{\mathrm{obs}} < 5$, which are expected to have a low signal-to-noise ratio. For months that have data in more than 3000 gridcells, we mask up to 40% of the values between 25°S and 60°N using random sampling. These limits increase computational efficiency of the assimilation and maintain a reasonable balance between the number of SLP and SST observations, otherwise there would be approximately five times as many SLP as SST observations. Valler et al. (2024) also reduce the number of ICOADS observations of SLP in their atmospheric reanalysis and set a similar threshold of $n_{\mathrm{obs}} = 10$ per gridcell. Past studies identified a bias in ICOADS SLP data before 1870, which is discussed in Slivinski et al. (2019), Freeman et al. (2017), and Allan and Ansell (2006). NOAA 20CRv3 performed a bias correction of the pre-1870 SLP observations, so we substitute the 1850–1870 SLP from ICOADS with the collocated values from NOAA 20CRv3. ICOADS does not provide an estimate of measurement and sampling errors which comprise the diagonal terms in $\mathbf{R}$. As described in Appendix B, we estimate $\mathbf{R}$ from the intramonth spread in individual observations and the variance across neighboring observations.

Sea ice observations are provided by the NOAA/NSIDC Climate Data Record (CDR) of Passive Microwave Sea Ice Concentration, Version 4 from 11/1978–09/2023 and Near-Real-Time, Version 2, for 10/2023–12/2023 (Meier et al. 2021b,a). We coarsen the observations from 25 km to 2° resolution. At each timestep in the assimilation with satellite data, we use a subset of the available data, which has nearly complete coverage of the polar regions. We retain all observations with SIC ranging from 0.01 to 0.98 and 40% of the remaining observations using random sampling. For measurement and sampling errors that form the diagonal terms in $\mathbf{R}$, we use the provided standard deviations of daily values, but we set the minimum error to 0.01. As described in Appendix B for SLP, these intramonth standard deviations approximate the monthly mean error. For SIC, they are calculated across both the NASA Team and Bootstrap algorithms, sampling the systematic error across data-processing methods. Errors are small in open water and pack ice but are often between 0.3 and 0.5 in partial ice cover. We do not have satellite data for sea ice from 1961–1978, but we require a full climatology from 1961–1990 to calculate the SIC anomalies relative to a baseline that is consistent with the HadSST4 anomalies. The mean of the eight models used for LIM training agrees well with observations over the satellite era (SI of Roach et al. 2020; Notz and SIMIP

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

Community 2020), so we combine the multi-model mean of the eight historical simulations from 1/1961–11/1978 with the satellite data from 12/1978–12/1990, and we use the merged climatology from 1961–1990 as the reference for SIC anomalies. Because solutions to (8) are not restricted to SIC between zero and one, we use the climatology in postprocessing to ensure that SIC is between zero and one.

d. Validation: Pseudo-reconstruction of an out-of-sample model

To test our method, we mimic the real reconstruction problem and attempt to reconstruct the 1850–2014 historical simulation from a climate model. Our target model is MPI-ESM1-2-HR, ensemble member r1i1p1f1 (Mauritsen et al. 2019). We have chosen MPI-ESM1-2-HR because it is a difficult test of the method given that, unlike nearly all other models, it has cooling in the Southern Ocean from 1980–2014. It also has a low-bias in Antarctic sea ice (Roach et al. 2020) and substantially different ENSO statistics and radiative feedbacks (Bloch-Johnson et al. 2024) compared to the models used for LIMs and priors in the data assimilation. The pseudo-reconstruction's target is out-of-sample because MPI-ESM1-2-HR is not used for LIM training; the dynamics of the target model are unknown to our eight forecast models.

We draw pseudo-observations from the target simulation at the same times and locations where real observations are available for SST, T, SLP, and SIC. Random errors are added to the pseudo-observations by sampling from the real observation errors in $\mathbf{R}$. Note that real observations also have biases and unknown, unquantified errors which make the real reconstruction more challenging than this test. On the other hand, the LIMs used as model priors are selected based on their ability to collectively emulate reality rather than the target model of the pseudo-reconstruction.

Figure 2 shows timeseries representing climate variability from the pseudo-reconstruction. The ensemble mean is calculated as the grand mean across all 1600 ensemble members (8 LIMs × 200 members), and the ensemble shading spans the 17^{th}–83^{rd} percentiles. We note that any one of the eight LIM-DA systems may not have a posterior distribution that spans the true state by itself. However, the grand ensemble of posterior distributions from all eight LIM-DA systems, which includes the spread from model error, generally spans the target (Figure 2). Because the grand ensemble represents eight separate DA systems, its distribution is non-Gaussian.

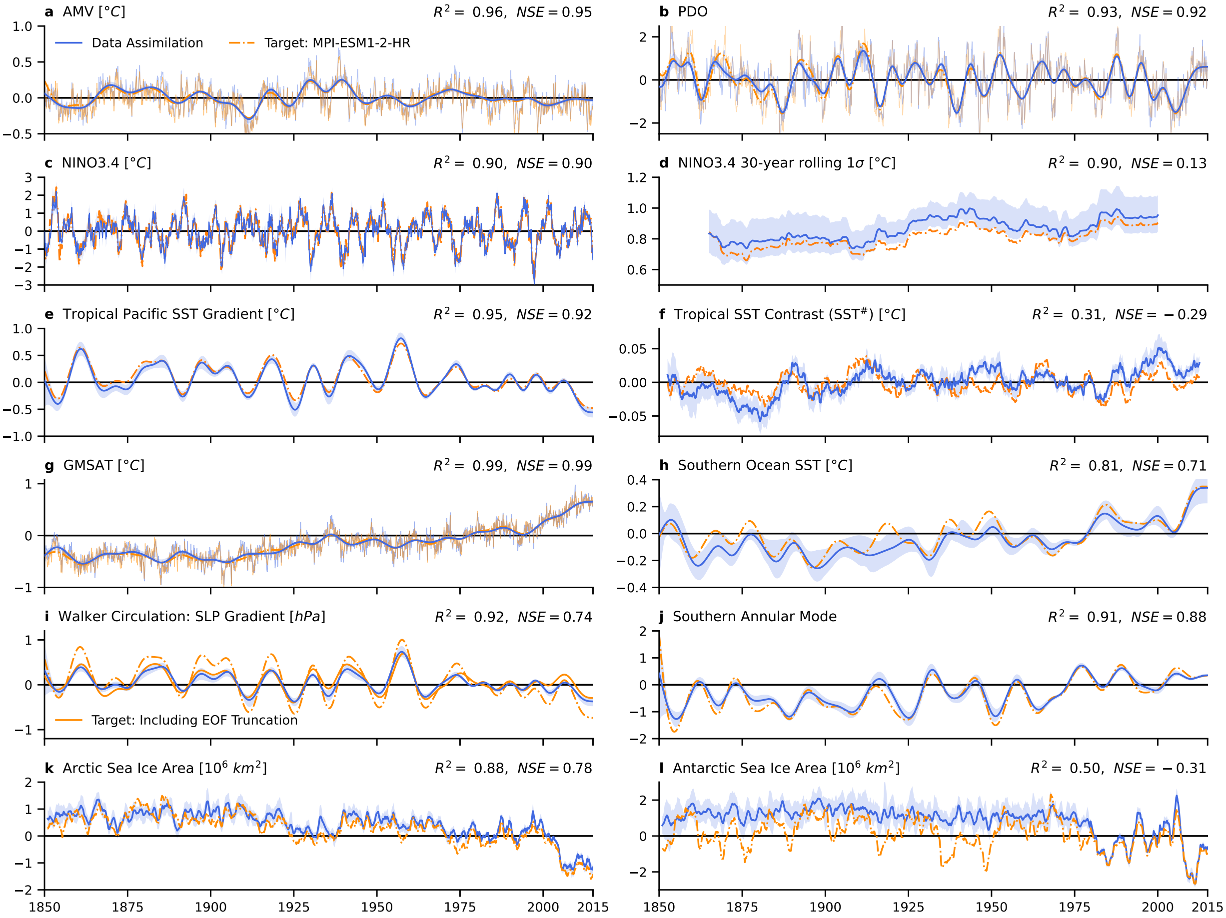

FIG. 2. **Validation by pseudo-reconstruction: timeseries.** (Orange) Values from the target model, the 1850–2014 historical simulation from MPI-ESM1-2-HR. (Blue) Result from data assimilation, showing mean of 1600 ensemble members; shading denotes ensemble 17th and 83rd percentiles, i.e., *likely* range. **(a)** Atlantic Multidecadal Variability with 10-yr low-pass filter and monthly values as thin lines. **(b)** Pacific Decadal Oscillation with 6-yr low-pass filter and monthly values as thin lines. **(c)** Monthly Nino3.4 with 30-yr running mean removed. **(d)** Rolling 30-yr standard deviation of Nino3.4 in panel **c**. **(e)** Zonal gradient of tropical Pacific SST with 10-yr low-pass filter. **(f)** Tropical SST contrast, SST$^\#$, 5-yr running mean. **(g)** Global-mean near-surface air temperature (GMSAT) with 10-yr low-pass filter and monthly values in thin lines. **(h)** Zonal mean of Southern Ocean SST (50°–70°S) with 10-yr low-pass filter. **(i)** Walker circulation, i.e., zonal SLP gradient across tropical Pacific, with 10-yr low-pass filter. **(j)** Southern Annular Mode with 10-yr low-pass filter. **(k)** Total area of Arctic and **(l)** Antarctic sea ice, with 12-month running mean applied. R^2 and Nash-Sutcliffe Efficiency (NSE) are based on the filtered metrics shown (see Fig. S5 for unfiltered results from monthly data). Calculation of metrics is described in Methods Section 2d.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

The metrics in Figure 2 are calculated as follows, with anomalies representing the departures from the 1961–1990 climatological annual cycle unless stated otherwise:

- Atlantic multidecadal variability (AMV) is the monthly mean SST anomaly in the North Atlantic (0°–60°N, 80°W–0°W) minus the global mean; the mean of the index from 1900–1970 is removed before plotting (Trenberth and Shea 2006).
- The Pacific Decadal Oscillation (PDO) is the leading EOF of the monthly mean SST anomaly in the North Pacific (20°–70°N) after removing the global mean (Newman et al. 2016).
- Nino3.4 is the monthly mean SST anomaly over 170°W to 120°W and 5°S–5°N, with the 30-yr running mean removed.
- The zonal SST gradient in the tropical Pacific is the mean SST anomaly in the west (80°E–150°E) minus the east (160°W–80°W), spanning 5°S–5°N (e.g., Heede and Fedorov 2023).
- SST$^{\#}$, which denotes the tropical SST contrast, is the mean of the warmest 30% of all tropical SSTs (30°S–30°N) minus the mean tropical SST, and the 1961–1990 mean is removed (Fueglistaler 2019). We note that SST$^{\#}$ requires actual SSTs, not just anomalies. To estimate SST$^{\#}$ for the pseudo-reconstruction, we add reconstructed anomalies to the target model's 1961–1990 climatology, which assumes outside knowledge of the target's climatology. For the actual reconstruction of SST$^{\#}$, we add the reconstructed anomalies to the 1961–1990 climatology from HadISST2.1's ensemble mean. We show the 5-year running mean of SST$^{\#}$ for consistency with Fueglistaler and Silvers (2021).
- Southern Ocean SST is the zonal-mean SST anomaly from 50°–70°S (Doddridge and Marshall 2017).
- Global-mean near-surface air temperature (GMSAT) is the global-mean T anomaly.
- The Walker circulation, measured by the zonal SLP gradient, is the mean SLP anomaly in the west Pacific (130°E–150°E) minus the central-east Pacific (160°W–120°W), spanning 5°S–5°N (e.g., Heede and Fedorov 2023).
- The Southern Annular Mode (SAM) is the standardized zonal-mean SLP anomaly at 40°S ± 2° minus the standardized zonal-mean SLP anomaly at 65°S ± 2° (Gong and Wang 1999); the reference period for standardization is 1961–1990, and each month is standardized separately.

- Sea ice area is the sum of the products of SIC and gridcell area; a common land mask is used when comparing ice area across various SIC datasets.

Most large-scale metrics are reconstructed with accuracy. We assess performance by the Pearson correlation (R), the fraction of variance explained (R^2), and the Nash-Sutcliffe Efficiency (NSE),

$$\mathrm{NSE} = 1 - \frac{\sum (x_i - \hat{x}_i)^2}{\sum (x_i - \bar{x})^2},$$

which accounts for the relative phasing of the target timeseries (x_i) versus the reconstructed timeseries ($\hat{x}_i$), the signal amplitude, and bias. The NSE has an upper bound equal to one and can become negative from biases in the mean or amplitude of variability (Nash and Sutcliffe 1970). We find $R^2 > 0.80$ for the AMV, PDO, Nino3.4, the 30-year rolling 1σ of Nino3.4, the zonal SST gradient in the tropical Pacific, GMSAT, Southern Ocean SST, the Walker circulation (zonal SLP gradient), the SAM, and Arctic ice area. The tropical SST contrast, $SST^{\#}$, has the lowest R^2 at 0.31.

The reconstruction of the Walker circulation has a damped amplitude compared to the target, which is due to the EOF truncation of SLP in the LIM training. We show an additional version of the target model's Walker circulation, which is calculated after truncating the target's SLP into the leading 30 EOFs. Truncation is expected to affect tropical SLP because the variance in tropical SLP is low compared to the variance at higher latitudes, but truncation does not appear to have a substantial influence on other metrics.

Antarctic sea ice has $R^2 = 0.50$ and is biased high in the reconstruction before 1979. The reason for this bias is that the target model is biased low relative to the multi-model mean of the LIMs and relative to the satellite record (Roach et al. 2020). There are decadal periods of abrupt ice loss in the target model which are not captured in the reconstruction. These ice-loss events are associated with brief warming episodes in Southern Ocean SST (Figure 2h), which are also not detected in the reconstruction. While we do not know whether such Antarctic ice-loss events happen in nature, we note that our pseudo-reconstruction of MPI-ESM1-2-HR does not capture its ice-loss events when observations are very sparse. Reasons for this deficiency could be (i) the LIMs used as model priors are too different from the target model, and sparse observations cannot overcome those differences, and/or (ii) the ice-loss events do not covary with available observations, and even a perfect model would be unable to reconstruct them from the data. Despite missing these decadal warmings, the lower-frequency variability in Southern Ocean SST and the SAM is captured by the reconstruction.

Figure 3 shows the pattern of trends in annual-mean SST for 1900–1979 and 1980–2014. Local trends are divided by the global-mean trend to emphasize the patterns. We also show the reconstruction's ensemble spread (1σ) in trend patterns, which highlights regions of elevated uncertainty. Note that SST is defined in all ocean gridcells at all times, even when SIC is 100%, so there are no missing values in the SST field. It is important to recall that observations in the Southern Ocean and southeast Pacific are sparse even after 1980 (Figure 1c), which is evident in our uncertainty quantification.

To further illustrate the uncertainty, we show trends from individual ensemble members (Figure 3c,g). These ensemble members show more cooling in the Southern Ocean than is seen in the ensemble mean. The key point, which is relevant to the next section on the real reconstruction, is that our DA framework is capable of reconstructing cooling over the Southern Ocean, even though the models used to train the LIMs do not show post-1980 cooling over the Southern Ocean in their historical simulations. Model biases can often be overcome if there are enough observations, and the LIM dynamics allow for cooling trends in the Southern Ocean. However, due to poor data coverage and quality in the Southern Ocean, SST trends in this region should be interpreted with caution.

Figure 4 shows trends in annual-mean SLP for 1900–1979 and 1980–2014. We only assimilate marine SLP observations, hence terrestrial SLP is expected to deviate from the target model. Large-scale patterns are consistent, but the errors in the magnitude of trends are substantial, especially over the Southern Ocean. Sparse observations and the unique physics of the target model compared to the forecast models results in considerable uncertainty. The uncertainty indicates that many ensemble members have local trends that differ substantially from the target model, and therefore accurately capturing the trend pattern requires considering the mean across the ensemble.

For additional validation, we show the spatial distribution of correlation and Nash-Sutcliffe efficiency for multiple time periods in Supplemental Figures S1–S2. In Supplemental Figures S3–S4, we also show the correlation and Nash-Sutcliffe efficiency when using only one LIM instead of the multi-model mean of eight LIMs, which illustrates the major improvements from using multiple

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

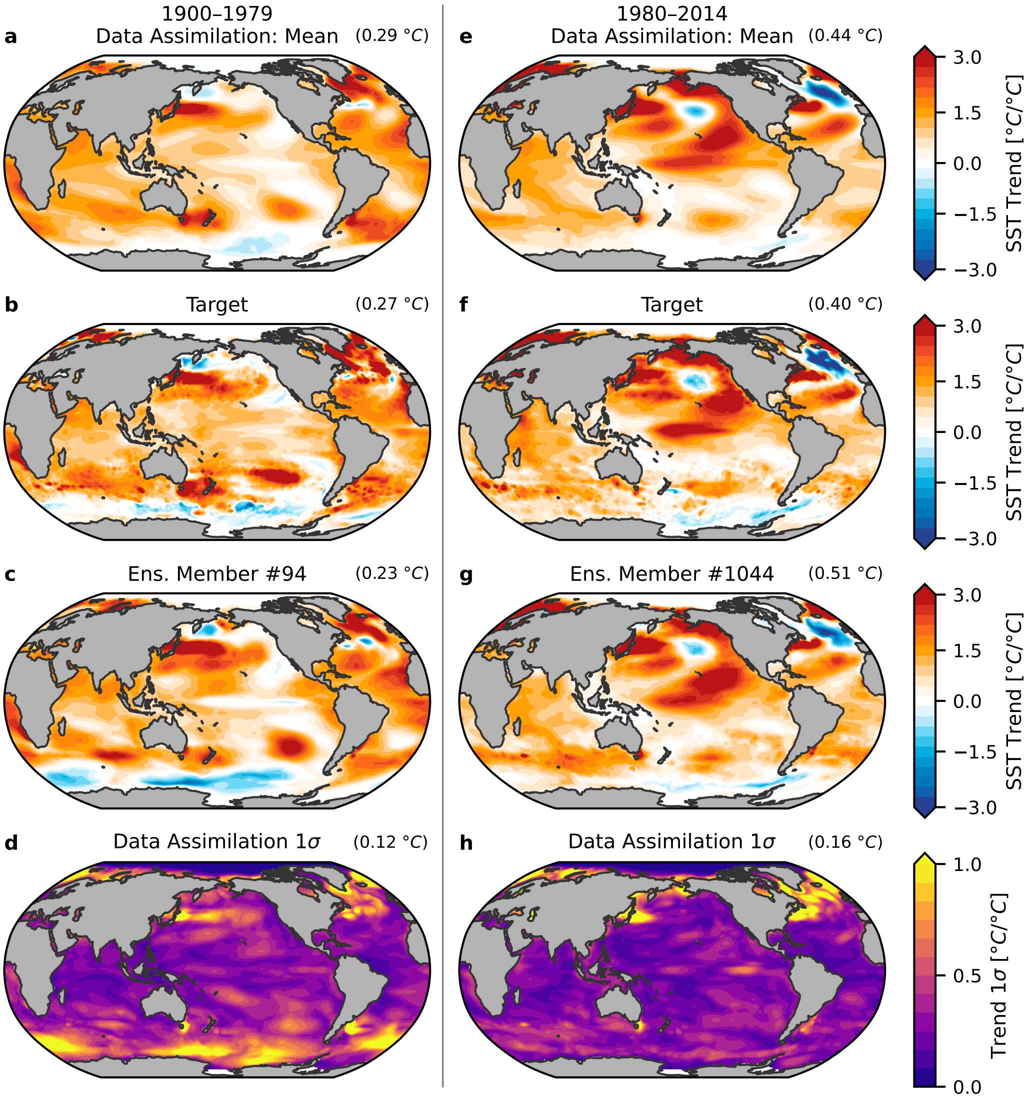

FIG. 3. **Validation by pseudo-reconstruction: SST trends.** **(a)** Normalized 1900–1979 ensemble mean of trends in the annual mean from data assimilation; local trends are divided by the global-mean trend to show SST patterns; upper-right indicates the global-mean trend before normalization, scaled by the number of years to show trend in °C per 80 years. **(b)** Repeats panel **a** but showing trends in the pseudo-reconstruction's target model, MPI-ESM1-2-HR's historical simulation. **(c)** Repeats panel **a** but shows an individual member from ensemble data assimilation. **(d)** Uncertainty in results from data assimilation, calculated as the sample standard deviation (1σ) across 1600 ensemble members' normalized trends; values greater than 1.0 indicate that local 1σ is greater than the global-mean trend; upper-right shows the global-mean of the 1σ in local trends before normalization. **(e–f)** Repeats panels **a–d** for 1980–2014.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

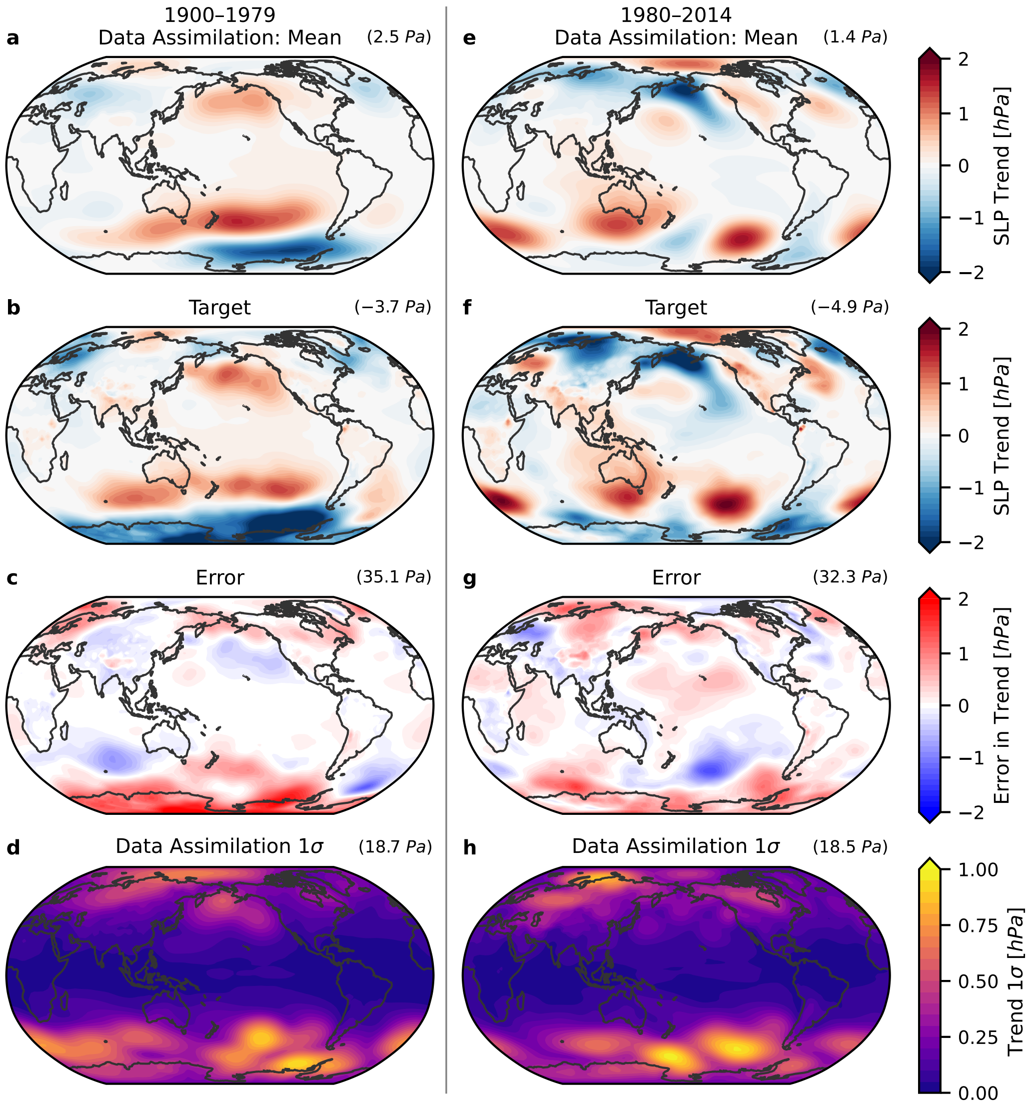

FIG. 4. **Validation by pseudo-reconstruction: trends in sea-level pressure (SLP). (a)** 1900–1979 ensemble mean of trends in the annual mean from data assimilation, scaled by the number of years to show trends in hPa per 80 years; upper-right indicates the global-mean trend in Pa per 80 years. **(b)** Repeats panel **a** but showing trends in the pseudo-reconstruction's target model, MPI-ESM1-2-HR's historical simulation. **(c)** Error, shown as mean reconstruction minus target; RMSE shown in upper right. **(d)** Uncertainty in results from data assimilation, calculated as the sample standard deviation (1σ) across trends from 1600 ensemble members; upper-right shows the global mean of the 1σ in local trends. **(e–f)** Repeats panels **a–d** for 1980–2014.

models in the reconstruction (Amrhein et al. 2020; Parsons et al. 2021). Additionally, we show monthly breakdowns of R^2 and Nash-Sutcliffe efficiency calculated without low-pass filtering the monthly resolved results for each of the metrics in Figure 2 (Supplemental Figure S5). The results show some seasonal variation in skill, which depends on the metric considered.

e. Observation validation from Desroziers statistics and HadSST4 comparison

For the reconstruction using real observations (Section 3), we also evaluate performance using the Desroziers statistics of the DA system (Desroziers et al. 2005) as described in Slivinski et al. (2021),

$$\mathrm{RMSE}_{\mathrm{actual}} = \left(\frac{1}{N_{\mathrm{obs}}} \sum_{j=1}^{N_{\mathrm{obs}}} \left(\mathbf{y}_j - [\mathbf{H}\mathbf{x}_f]_j \right)^2 \right)^{1/2}, \tag{11}$$

$$\mathrm{RMSE}_{\mathrm{exp}} = \left(\frac{1}{N_{\mathrm{obs}}} \sum_{j=1}^{N_{\mathrm{obs}}} \left(\mathbf{R}_j + [\mathbf{H}\mathbf{P}_f\mathbf{H}^T]_j \right) \right)^{1/2}, \tag{12}$$

where j is the observation index for each observation in a given month, $\mathbf{y}_j$ is observation j, $[\mathbf{H}\mathbf{x}_f]_j$ is the forecast prior mean of observation j, and the corresponding $\mathbf{R}_j$ and $[\mathbf{H}\mathbf{P}_f\mathbf{H}^T]_j$ in Equation (12) are the observation and forecast errors associated with observation j. $\mathrm{RMSE}_{\mathrm{actual}}$ is related to the innovations from Equation (8) and compares the forecasts with observations that have not yet been assimilated, while $\mathrm{RMSE}_{\mathrm{exp}}$ is related to the innovation covariance in Equation (9). The results described below are shown in Figure S6 and illustrate the calibration of the DA system.

If the calibration ratio $\mathrm{RMSE}_{\mathrm{actual}}/\mathrm{RMSE}_{\mathrm{exp}} \approx 1$, the system is well calibrated (Slivinski et al. 2021; Houtekamer and Mitchell 1998). We group the data into 20°N – 90°N (NH), 20°S – 20°N (Tropics), and 20°S – 90°S (SH), and we compute the calibration ratio using 30-year running means of the RMSE values, then we take the mean of the ratio over 1850–2023. For SST, we find calibration ratios of 1.2 (NH), 1.1 (Tropics), and 1.2 (SH). These ratios are close to 1 and confirm that the DA system is performing well for SST. Calibration ratios for SLP are 0.9 (NH), 0.7 (Tropics), and 0.8 (SH), indicating that the expected errors are larger than the actual errors due to excessive ensemble spread. For T, calibration ratios are 1.1 (NH), 1.6 (Tropics), and 1.3 (SH). Although tropical SST is very well calibrated with a ratio of 1.1, the expected errors are larger than $\mathrm{RMSE}_{\mathrm{actual}}$ for tropical SLP (ratio 0.7) and smaller than $\mathrm{RMSE}_{\mathrm{actual}}$ for tropical air temperatures over land (ratio 1.6). Overall, the Desroziers statistics suggest the DA system is well calibrated, especially for SST.

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

To illustrate observation validation at specific locations, we show sample timeseries comparing the assimilated HadSST4 observations with the real DA results (from Section 3) at seven ocean locations (Figs. S7–S14). Overall, the results show that errors in the reconstruction relative to the observations are in good agreement with observation error. When outliers appear in the data, their influence is limited by the DA prior and the other observations that are simultaneously assimilated. Figures S8–S14 also illustrate the time-varying observation density and uncertainty at various locations, reinforcing the summary calibration results in that the reconstructed ensemble mean and spread are consistent with the assimilated observations and their errors.

f. Comparison data

We include a variety of datasets for comparison with our reconstruction. For SST, we focus on datasets which are globally complete and have monthly resolution. We include PCMDI/AMIP-II (Hurrell et al. 2008), which was used for CMIP6's AMIP simulations, NOAA ERSSTv5 (Huang et al. 2017), HadISST1 (Rayner et al. 2003), HadISST2.1 (no longer maintained; Titchner and Rayner 2014), and COBE-SST2 (Hirahara et al. 2014). The statistical infilling in these products is briefly described by Modak and Mauritsen (2023) and Lewis and Mauritsen (2021), with more detail in Kent and Kennedy (2021). All products are regridded to the 2° resolution of our reconstruction.

For SLP, we show gridded reanalyses from ERA5 (1950–present) (Hersbach et al. 2020), NOAA/CIRES/DOE 20CRv3 (1836–2015) from Slivinski et al. (2019), and NCEP/NCAR (1948–present) from Kalnay et al. (1996), all regridded to 2° and monthly resolution. We also include an older product, HadSLP2 infilled (Allan and Ansell 2006). HadSLP2 is no longer maintained, but it provides monthly means of SLP and its non-infilled product would be a companion to HadSST4 if updated. We include an offline-DA reconstruction of the Walker circulation using proxy data, labeled F23 (Falster et al. 2023). We include the SAM from multiple reconstructions using offline DA (O'Connor et al. 2021; Dalaiden et al. 2021; King et al. 2023) and regression (Fogt et al. 2009), labeled as O21, D21, K23, and F09.

For SIC, we show HadISST2.2 (Titchner and Rayner 2014), HadISST1 (Rayner et al. 2003), and AMIP-II (Hurrell et al. 2008), which is largely based on HadISST1. The satellite record from NOAA/NSIDC CDR (Meier et al. 2021b,a) is shown from 11/1978–12/2023. We include the proxy-based reconstruction of Arctic SIC from Brennan and Hakim (2022), labeled BH22,

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

which has annual rather than monthly resolution. We regrid all SIC data to 2° resolution. When comparing total anomalies in sea-ice area, we restrict the comparison to only include gridcells that have SIC data in every dataset. Otherwise, one dataset may have large anomalies where another dataset has missing values from different land masks, skewing the comparison.

For global-mean T (GMSAT), we compare with HadCRUT5 (Morice et al. 2021) and BEST (Rohde et al. 2013). Note that our reconstruction is of the near-surface air temperature, while the comparison datasets are hybrids of air temperature over land and SST over ocean.

Notably, various datasets can impact one another. The lower boundary condition in ERA5 is the SST from HadISST2 until 2007 and sea ice from HadISST2 until 1979 (Hersbach et al. 2020). NOAA 20CRv3 also uses HadISST2 sea ice over 1836–2015, HadISST2 SST after 1981, and SODAsi.3 SST adjusted to HadISST2 climatology before 1981 (Slivinski et al. 2019; Giese et al. 2016). An SST dataset, ERSSTv5, also uses HadISST2 sea ice to adjust its SST values in the Southern Ocean (Huang et al. 2017). These are examples of how uncertainty in one dataset can affect others.

3. Historical Reconstruction

In this section, we share the results of our reconstruction of SST, T, SLP, and SIC from coupled atmosphere–ocean data assimilation with linear inverse models. We show timeseries and spatial trends of SST, SLP, and SIC, and the El Niño of 1877/1878.

a. Variability over 1850–2023

Figure 5 shows timeseries of the real reconstruction, as for the pseudo-reconstruction (Figure 2). The AMV and PDO are similar across datasets for most of the historical record, as described for the PDO in Newman et al. (2016), but PDO uncertainty is notably larger from 1850–1900.

Nino3.4 shows substantial inter-dataset spread before 1875, but the most interesting ENSO feature is the low-frequency evolution of ENSO variance in Figure 5d, measured by the 30-year rolling 1σ of Nino3.4. Recent studies have argued for increased ENSO variance with global warming (e.g., Cai et al. 2021, 2023), although other work suggests that ENSO variance could decrease with long-term warming (Callahan et al. 2021), and uncertainties in future ENSO variance have substantial implications for global-scale climate predictability (Amaya et al. 2025). In our results,

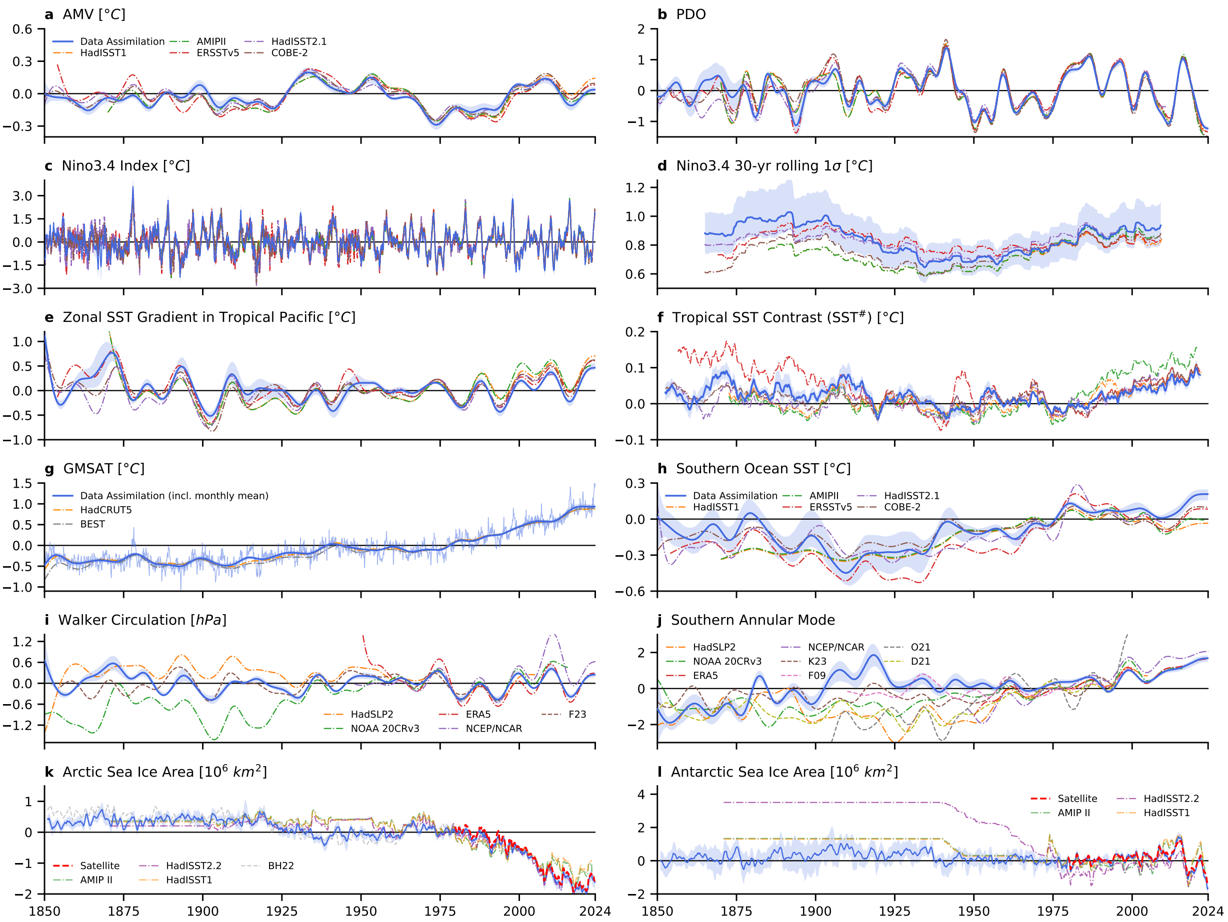

FIG. 5. **Climate variability over 1850–2023.** (Blue) Results from data assimilation, showing mean of 1600 ensemble members; shading denotes ensemble 17th and 83rd percentiles, i.e., *likely* range. Note that legend for SST datasets in panel **a** applies to panels **a–f**, and re-used line colors in SLP, T, and SIC panels do not necessarily indicate consistency with the SST datasets. **(a)** Atlantic Multidecadal Variability (SST) with 10-yr low-pass filter. **(b)** Pacific Decadal Oscillation (SST) with 6-yr low-pass filter. **(c)** Monthly SST in Nino3.4 region with 30-yr running mean removed. **(d)** Rolling 30-yr standard deviation of Nino3.4 in panel **c**. **(e)** Zonal gradient of tropical Pacific SST with 10-yr low-pass filter. **(f)** Tropical SST contrast, SST$^{\#}$, 5-yr running mean. **(g)** Global-mean near-surface air temperature (GMSAT) with 10-yr low-pass filter and monthly values from data assimilation as thin line. **(h)** Zonal mean of Southern Ocean SST (50°–70°S) with 10-yr low-pass filter. **(i)** Walker circulation, i.e., zonal SLP gradient across tropical Pacific, with 10-yr low-pass filter. **(j)** Southern Annular Mode (SLP) with 10-yr low-pass filter. **(k)** Total area of Arctic and **(l)** Antarctic sea ice, with 12-month running mean applied, and showing comparison satellite data from NOAA/NSIDC CDR. Calculation of metrics is described in Section 2d, and comparison data is summarized in Section 2f.

ENSO variance was at local maximum over 1875–1900, decreased to a local minimum over 1930–1960, and subsequently trended higher to the present. Overall, Figure 5d suggests considerable centennial-scale power in ENSO variance.

Tropical SST gradients are diagnosed using two measures. The Pacific zonal SST gradient (Figure 5e), shows that the magnitude of the strengthening trend from 1980–2023 is not clearly distinguishable from past variability, such as the weakening from 1875–1905. The long-term strengthening trend since 1900 has also been a focus of many studies (e.g., Cane et al. 1997; Karnauskas et al. 2009; Deser et al. 2010b; Solomon and Newman 2012; Coats and Karnauskas 2017; Seager et al. 2022; Lee et al. 2022), but 1900–1905 has the weakest zonal gradient during the historical record, and the gradient in 1890 is comparable to 2023. However, the $\mathrm{SST}^{\#}$ metric (Figure 5f), representing the contrast between the warmest tropical SSTs and the tropical mean (Fueglistaler 2019; Fueglistaler and Silvers 2021), shows a consistent strengthening from 1975–present. The persistent 1975–2023 trend in $\mathrm{SST}^{\#}$ may indeed be distinct compared to the variability before 1975, but further investigation is needed.

The Pacific Walker circulation (zonal SLP gradient) appears to be dominated by stationary decadal variability over the full historical record (Figure 5i). Our reconstruction does not show a trend toward weakening of the Walker circulation over the 20th century (Vecchi et al. 2006; Tokinaga et al. 2012), and the strengthening from c. 1979–2014 (e.g., Chung et al. 2019; L'Heureux et al. 2013; Watanabe et al. 2023, 2024) appears within the range of variability prior to 1975. Heede and Fedorov (2023) found large recent changes in the zonal SLP gradient in the NCEP/NCAR Reanalysis, but that product may be an outlier over 2005–2020 (Figure 5i).

Our reconstruction of the Southern Annular Mode (SAM) has relatively small ensemble spread relative to the spread across other products (Figure 5j). Notably, the pre-1980 disagreement across reanalyses and other reconstructions is larger than the decadal variability in any one product. Spurious trends in Southern Hemisphere SLP have been identified in reanalyses poleward of 60°S during the early twentieth century and c. 1950 due to the general paucity of data over much of the Southern Hemisphere (Schneider and Fogt 2018; Fogt and Connolly 2021; Laloyaux et al. 2018). Local observations in the SAM region are sparse throughout most of the historical record. Consequently, our SAM reconstruction is primarily constrained by remote observations of SLP, SST, and T, with the dynamics of the LIMs acting to connect those remote observations to the

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

SAM region's SLP. Many studies have highlighted the positive trend in the SAM from c. 1980–present (e.g., Thompson and Solomon 2002; Marshall 2003; Polvani et al. 2011; Swart et al. 2015; Banerjee et al. 2020; Fogt and Marshall 2020), but some datasets in Figure 5j show longer-term positive trends, possibly spanning the entire 20th century (O'Connor et al. 2021; Dalaiden et al. 2021; Slivinski et al. 2019; Allan and Ansell 2006). Our results indicate that the recent trend only extends from approximately 1970–present, and the trends are most notable in DJF (Fig. S16). There appears to be another prolonged positive trend from 1850–1920 in our reconstruction but not in any of the comparison data, and that the SAM trend aligns with SST cooling in the Southern Ocean over the same period. Brönnimann et al. (2024) analyzed newly digitized ship records from 1903–1916 and also find a positive SAM index and pronounced surface cooling over the Southern Ocean during the early 1900s.

Sea ice from data assimilation (Figure 5k,l) exhibits major differences compared to the HadISST and AMIP-II datasets, which have been used to assess the atmospheric response to SIC changes over the historical record. Over much of the historical record, these datasets have constant values at inferred climatologies. There are also differences in the satellite era due to uncertainties in data processing and discontinuities in satellite sources (e.g., Eisenman et al. 2014; Buckley et al. 2024), which are responsible for the spurious high values in Antarctic sea ice from 2009–2011 in HadISST1 and AMIP-II (Screen 2011), evident in Figure 5l.

For Arctic sea ice, the main difference across datasets relates to the early 20th-century warming (Brönnimann 2009; Hegerl et al. 2018). HadISST1 and AMIP-II do not have any signal of the early 20th-century warming in sea-ice area. Our reconstruction shows a loss of 0.5 ± 0.1 (1σ) million km^2 during the 1920s, measured by comparing the decadal means of the 1930s and 1910s. Note that this value should not be compared directly with other datasets unless land masks are consistently applied. The Brennan and Hakim (2022) reconstruction of annual means, using only proxy data with offline DA, agrees closely with our results.

Antarctic sea ice is a unique result compared to existing estimates. In stark contrast to the datasets used for CMIP6/DECK/AMIP/CFMIP (Eyring et al. 2016; Webb et al. 2017) and as boundary conditions in reanalyses (e.g., Slivinski et al. 2019; Hersbach et al. 2020), our reconstruction shows much less ice loss from the preindustrial to present conditions. AMIP-II, HadISST1, and HadISST2 are at the edge or outside of our likely range for the entire pre-1980 period. Note that HadISST2 is the ice boundary condition in ERA5 and NOAA 20CRv3 before 1979, and it is used to adjust SST in NOAA ERSSTv5.

In the early 20th century, we find a wide envelope of uncertainty in Antarctic ice area that spans the range over the satellite record until 2022. Our results show a local maximum c. 1910, consistent with the SH cooling reported by Brönnimann et al. (2024). We find greater Antarctic ice cover in the early 1960s compared to the 1980s (Fan et al. 2014), consistent with Goosse et al. (2024). However, our reconstruction shows a decrease throughout the 1970s (Fig. S15) in contrast to the sharp drop in ice extent at the end of the 1970s reported by Goosse et al. (2024). Early single-channel satellite retrievals from ESMR suggest Antarctic ice cover may have been more extensive in the 1970s (Goosse et al. 2024; Kolbe et al. 2024), though the reliability of ESMR is debated (Titchner and Rayner 2014; Kolbe et al. 2024). As evident in the ensemble spreads (Fig. 5l; Fig. S15), the uncertainty before 1980 is substantial, and more work is needed to constrain Antarctic SIC. The preindustrial-mean ice area (1850–1900) does not appear clearly different from the satellite-era range until the ice loss of 2022–2023 (Roach and Meier 2024; Espinosa et al. 2024; Zhang and Li 2023; Fogt et al. 2022; Turner et al. 2022). Our results for preindustrial ice area are consistent with Edinburgh and Day (2016)'s analysis of ship records from the Heroic Age (1897–1917), who found ice expansion in the Weddell Sea but comparable conditions to 1989–2014 in the other sectors.

Finally, we consider variability in Southern Ocean SST (zonal mean 50°–70°S). We find a large spread in our ensemble before 1950 and a larger disagreement across SST datasets, which persists from 1850 to 2023. We note two interesting results in Figure 5h. First, we find a long-term warming trend from 1910–2023, which is approximately half as large as the 1910–present warming trend in GMSAT. This is consistent with expectations, since Southern Ocean warming is muted by upwelling of deep water that has not yet experienced the global warming signal (Armour et al. 2016).

Second, we find a muted cooling of the Southern Ocean from 1980–2013, and slight warming from 1980–2023. The comparison datasets show 1980–2013 cooling that is mostly outside of our likely range. In situ observations are still sparse from 1980–2023 (Figure 1; Figure S14) and the data sources change dramatically over that period, possibly introducing spurious trends from homogenizing different data sources (Kennedy et al. 2019; Huang et al. 2019; Kent and Kennedy

2021; Hausfather et al. 2017; Karl et al. 2015). We elaborate on Southern Ocean trends below and in the Discussion section.

The Southern Ocean cooling over recent decades is not unprecedented given that we find stronger cooling from 1880–1910. Brönnimann et al. (2024) report that this cooling is a real climatic phenomenon, not a data artifact. However, Sippel et al. (2024) suggest that biases in the bucket measurements of SST are responsible for a cold bias from 1910–1930. If SST-bucket biases are indeed responsible for this cooling trend, an explanation is still required for why the night-time marine air temperatures (Cornes et al. 2020) also show this 1910–1930 cooling trajectory (Figure 1 of Sippel et al. 2024).

b. Patterns of SST, SLP, and SIC trends

Figure 6 shows spatial patterns of SST trends separately for the gradual warming from 1900–1979 and the recent period of 1980–2023. We show our reconstruction and its uncertainty alongside comparison trends from NOAA ERSSTv5 and COBE-SST2. Despite similar global-mean trends from 1900–1979, there are substantial disagreements in the pattern of trends especially over the Southern Ocean and tropical Pacific. The post-1980 period is often viewed as having small uncertainty due to observation density (Figure 1), but the inter-dataset disagreements in Figure 6e–g suggest there are nontrivial uncertainties in large-scale SST gradients. The southeast Pacific and Southern Ocean regions, which have strong impacts on global climate variability and radiative feedbacks (e.g., Dong et al. 2022a; S. Kang et al. 2023c,a; Espinosa and Zelinka 2024), have the worst observation coverage (Figure 1).

Figure 7 shows spatial patterns of SLP trends for 1900–1979 and 1980–2023 from our reconstruction and comparison datasets. Note that our reconstruction only assimilates marine SLP observations, so we expect it to differ from reanalyses over land regions. From 1900–1979, there are many large-scale differences between our reconstruction, HadSLP2, and NOAA 20CRv3. The comparison datasets show strong negative trends in SLP over Antarctica and most of the Southern Ocean during both time periods, whereas we find positive trends over 1900-1979. In this region, regression-based reconstructions find positive trends in the early 20th century, also in contrast to the negative trends in existing reanalyses (Fogt et al. 2019; Fogt and Connolly 2021; Fogt et al. 2024). Schneider and Fogt (2018) and Laloyaux et al. (2018) highlight problems with the atmospheric

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

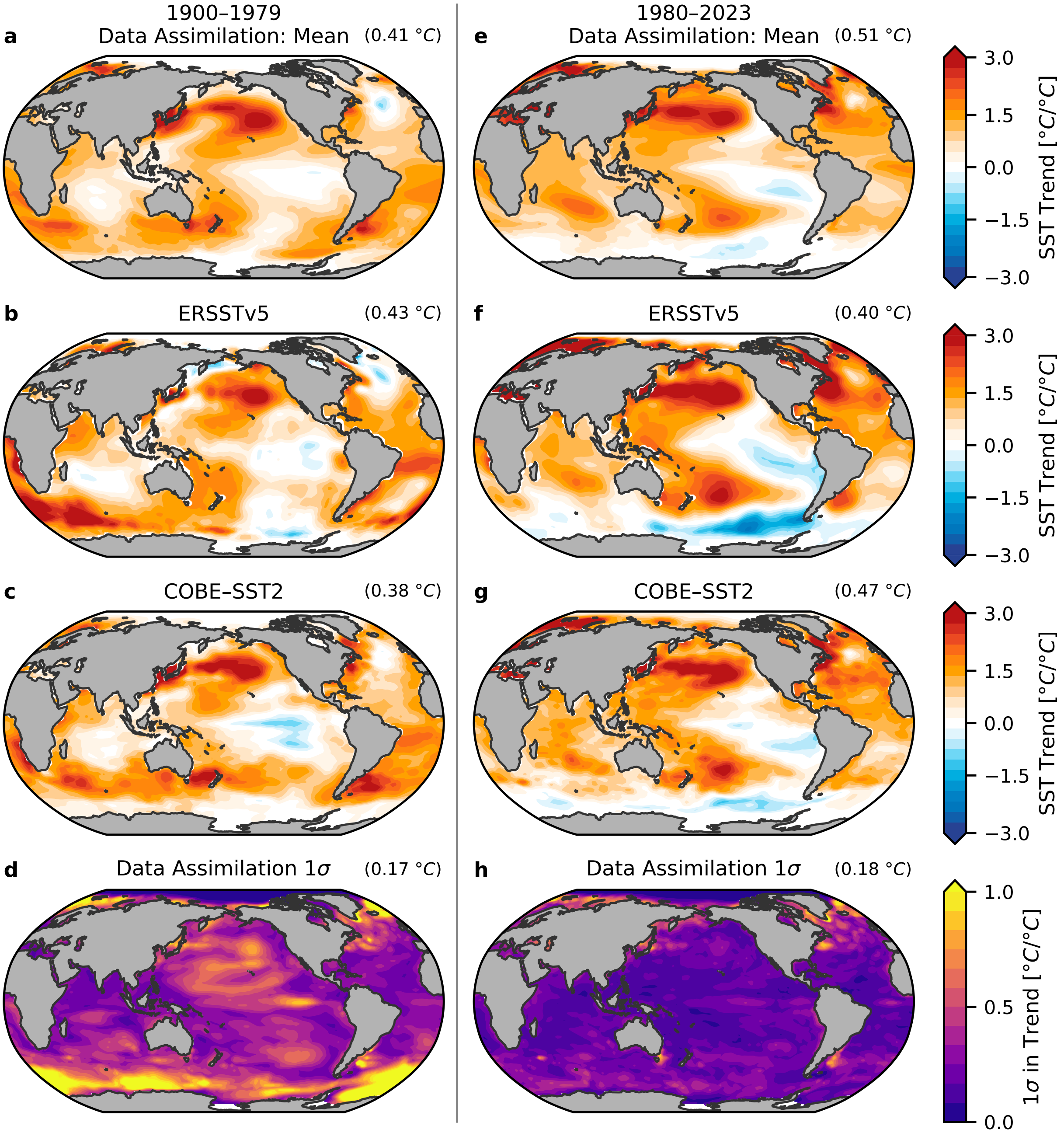

FIG. 6. **Historical patterns of SST trends.** **(a)** Normalized 1900–1979 ensemble mean of trends in the annual mean from data assimilation; local trends are divided by the global-mean trend to show SST patterns; upper-right value is the global-mean trend before normalization, scaled by the number of years to show trend in °C per 80 years. **(b)** Repeats panel **a** but showing comparison data from NOAA ERSSTv5 and **(c)** COBE-SST2. **(d)** Uncertainty in results from data assimilation, calculated as the sample standard deviation (1σ) across 1600 ensemble members' normalized trends; values greater than 1.0 indicate that local 1σ is greater than the global-mean trend; upper-right value is the global-mean of the 1σ in local trends before normalization. **(e–f)** Repeats panels **a–d** for 1980–2023.

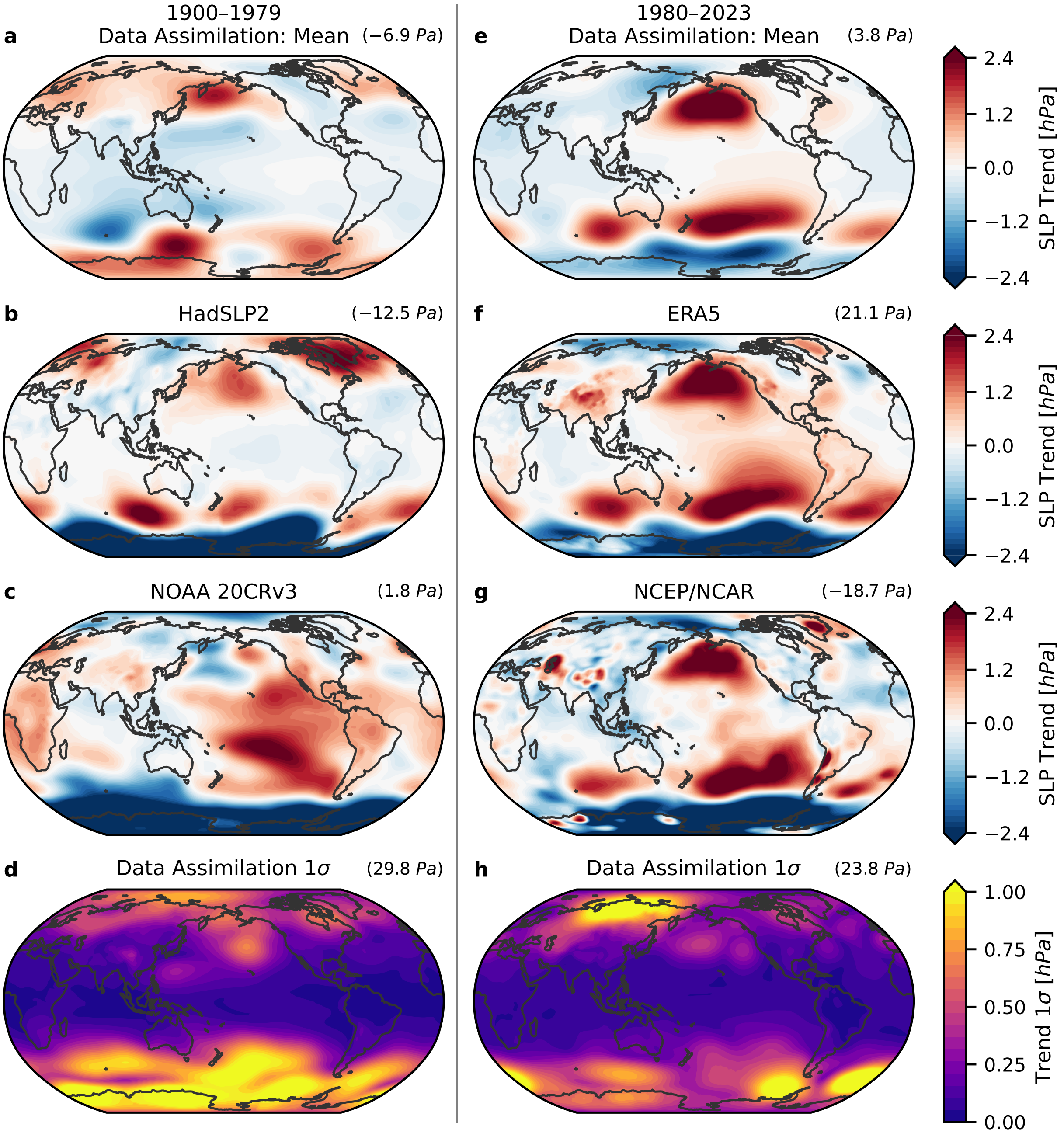

FIG. 7. **Historical trends in sea-level pressure (SLP). (a)** 1900–1979 ensemble mean of trends in the annual mean from data assimilation, scaled by the number of years to show trends in hPa per 80 years; upper-right indicates the global-mean trend in Pa per 80 years. **(b)** Repeats panel **a** but showing comparison datasets HadSLP2 and **(c)** NOAA 20CRv3. **(d)** Uncertainty in results from data assimilation, calculated as the sample standard deviation (1σ) across local trends from 1600 ensemble members; upper-right shows the global mean of the 1σ in local trends. **(e–f)** Repeats panels **a–d** for 1980–2023, with comparison reanalyses from **(f)** ERA5 and **(g)** NCEP/NCAR.

circulation in the Southern Hemisphere in multiple reanalyses and how those problems create spurious climate signals. The key problem identified in ERA-20C is that the assumed error is too small for pressure observations. This is one reason why we ensure our SLP observation error is not too small, as described in Appendix B.

Over 1980–2023, our SLP trends over the global oceans largely align with ERA5, albeit with weaker positive trends in the central and eastern Pacific (Figure 7e,f). ERA5 has a substantial trend in global-mean SLP, which increases by 21.1 Pa (44 yr)$^{-1}$ from 1980–2023, and removing this trend would improve agreement with our reconstruction in many regions. NCEP/NCAR has a substantial and opposite trend of -18.7 Pa (44 yr)$^{-1}$. Our reconstruction has a much smaller 1980–2023 trend in global-mean SLP of 3.8 Pa (44 yr)$^{-1}$ (Figure 7e) and similarly small trends from 1900–1979 and also in the pseudo-reconstruction experiment (Figure 4). Once again, our reconstruction highlights uncertainty over the Southern Ocean, especially the Amundsen Sea Low and the Atlantic sector.

Figure 8 shows trends in Arctic SIC over 1900–1979, during the early 20th-century warming from 1920–1935, and for the recent loss from 1980–2023. We compare with HadISST2, which is the pre-satellite boundary condition used in ERA5 and NOAA 20CRv3, and with the NOAA/NSIDC satellite data that we assimilate. From 1900–1979, we find ice loss in the Barents Sea between Svalbard and Russia. From 1920–1935, we find ice loss around most of the Arctic, partially offset by gains poleward of the Bering Strait. HadISST2 does not have this 1920–1935 ice loss. From 1980–2023, our ice loss looks very similar to the satellite record, but it does not match exactly because of uncertainty in the satellite data, the influence of non-SIC observations, and the particularities of our LIM and DA methods.

Figure 9 shows trends in Antarctic SIC from 1900–1979, during the 1960–1979 period of ice loss hypothesized by Fan et al. (2014), and from 1980–2023, a period with steady but small growth and then recent rapid loss (e.g., Stuecker et al. 2017). Our reconstruction of 1900–1979 shows some ice loss alongside the Southern Ocean SST warming, but we find a lesser magnitude and a different pattern compared to HadISST2. If sea ice has a relationship with the atmospheric circulation (e.g., Kohyama and Hartmann 2016), the HadISST2 boundary condition may impact the circulation in ERA5 and NOAA 20CRv3. From 1960–1979, we find ice loss in the Atlantic sector, which mostly

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

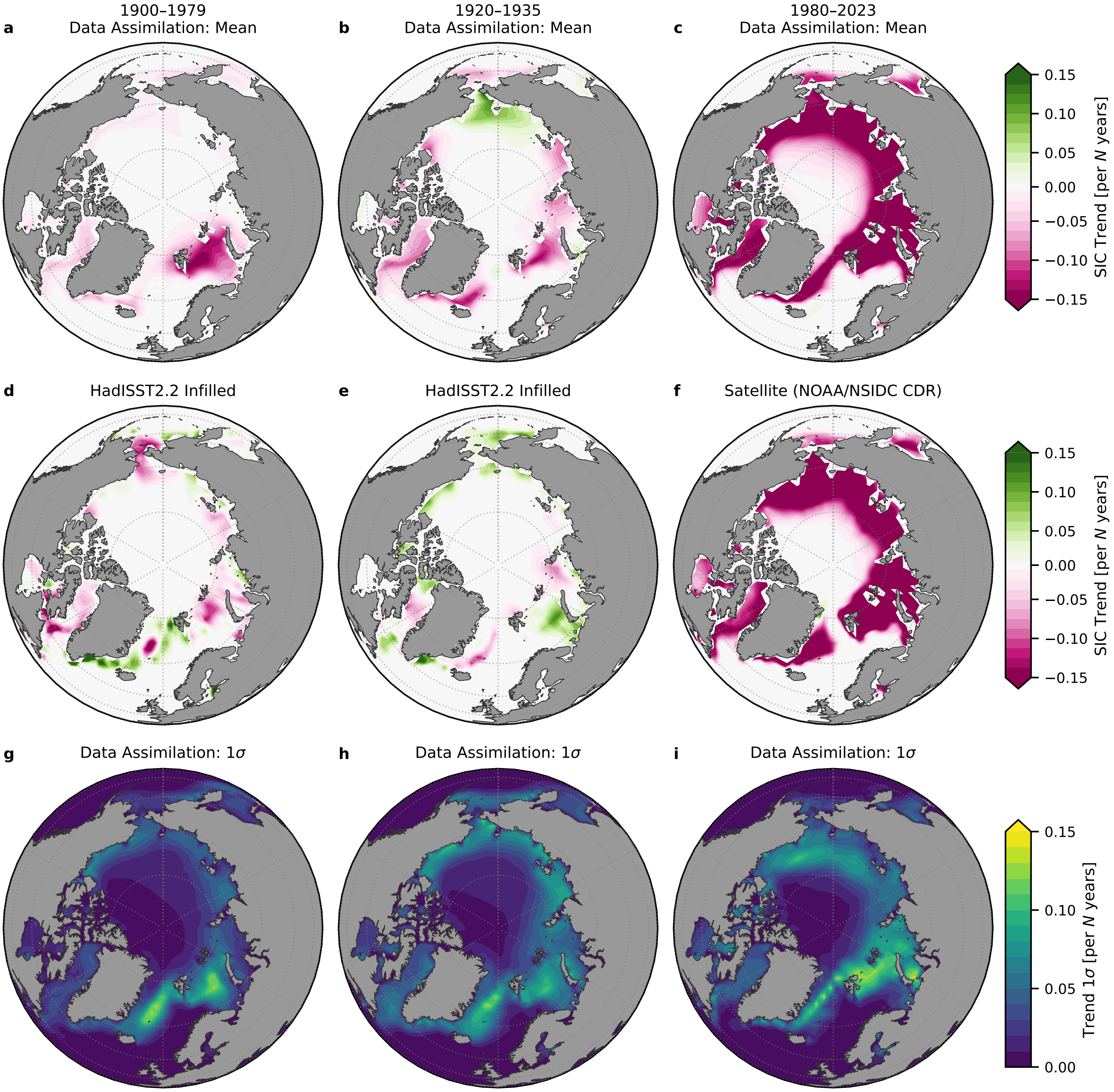

FIG. 8. **Historical trends in Arctic sea-ice concentration (SIC). (a–c)** Ensemble mean of trends from data assimilation, scaled by the number of years in each period to show trends in SIC per *N* years. **(d–f)** Repeats panels **a–c** but showing comparison datasets, with infilled HadISST2.2 in panels **d–e** and satellite data from NOAA/NSIDC CDR in panel **f**. **(g–i)** Uncertainty in results from data assimilation, calculated as local standard deviation (1σ) across 1600 ensemble members, corresponding to time periods in panels **a–c**. Note that SIC is bounded from 0 to 1.

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

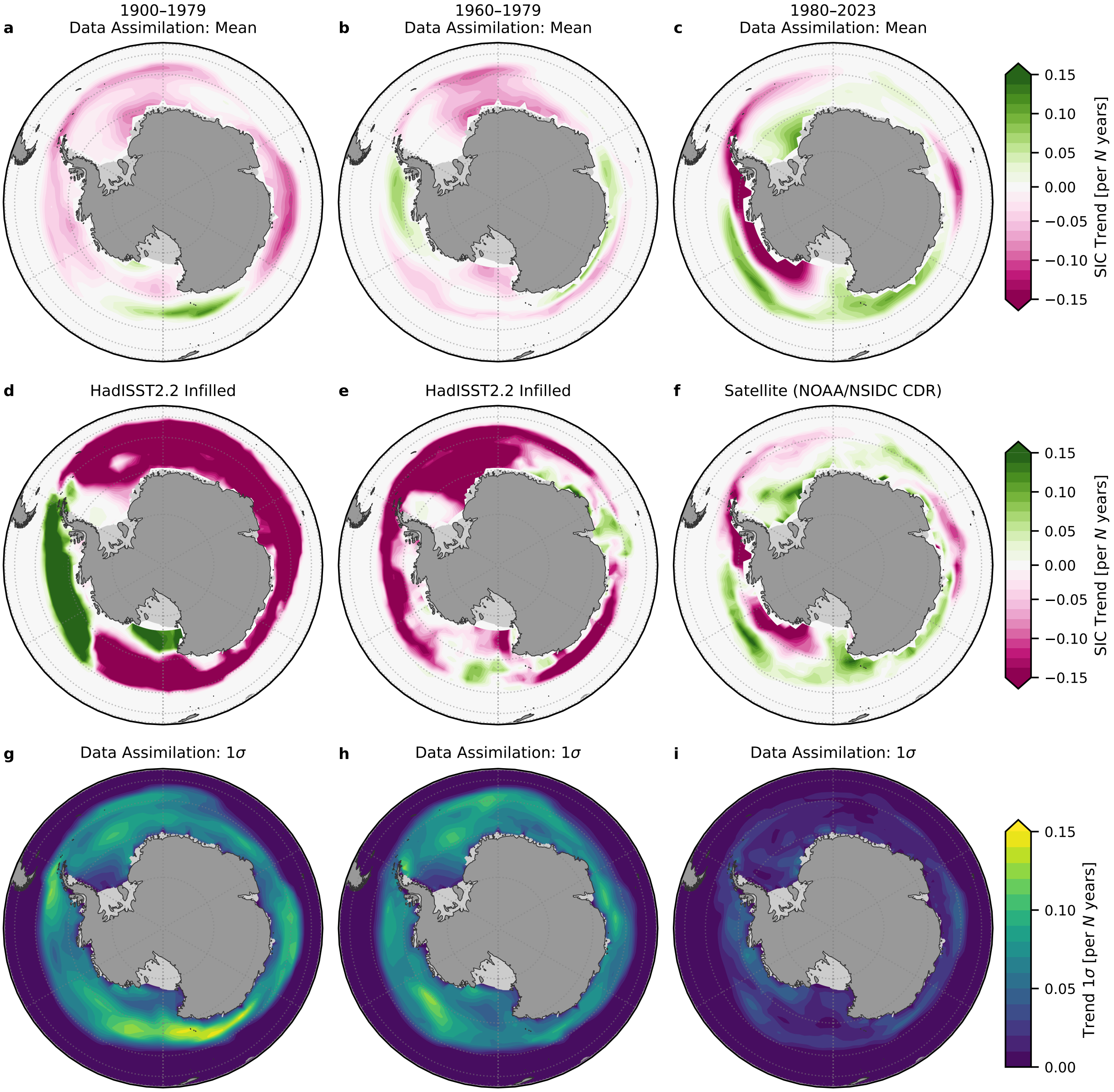

FIG. 9. **Historical trends in Antarctic sea-ice concentration (SIC). (a–c)** Ensemble mean of trends from data assimilation, scaled by the number of years in each period to show trends in SIC per *N* years. **(d–f)** Repeats panels **a–c** but showing comparison datasets, with infilled HadISST2.2 in panels **d–e** and satellite data from NOAA/NSIDC CDR in panel **f**. **(g–i)** Uncertainty in results from data assimilation, calculated as local standard deviation (1σ) across 1600 ensemble members, corresponding to time periods in panels **a–c**. Note that SIC is bounded from 0 to 1.

aligns with the pattern in HadISST2 but with a substantially different magnitude. We see a minor gain of ice in the Bellingshausen Sea, where HadISST2 shows large loss.

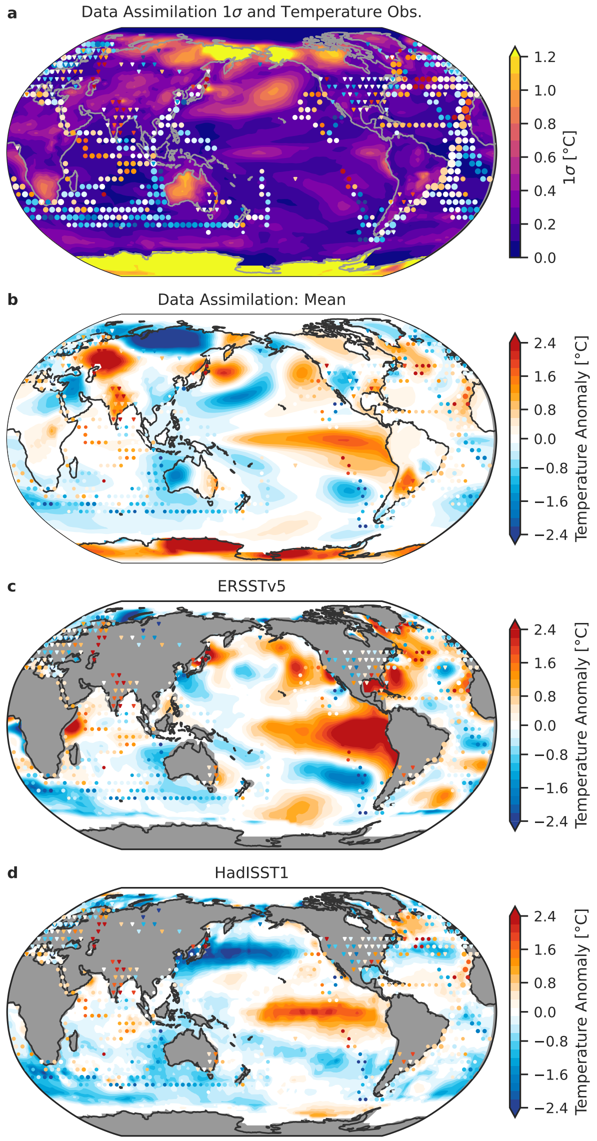

FIG. 10. **El Niño in July 1877: reconstruction, observations, and uncertainty.** **(a)** Contours show uncertainty in the data assimilation, calculated as the sample standard deviation (1σ) across the local anomalies in SST and near-surface air temperatures (T) over land for the 1600 ensemble members; scattered dots show anomalies in SST from HadSST4, with size inversely proportional to error, while triangles show land T from CRUTEM5; T and SST points use colorbar from panels **b–d**. **(b)** Contours show ensemble mean of SST anomalies and land T from data assimilation, with HadSST4 and CRUTEM5 observations. **(c–d)** Repeats panel **b** but with comparison SST datasets, NOAA ERSSTv5 and HadISST1.

c. *El Niño in 1877*

The extreme El Niño that began in 1877, which is the largest event in the historical record, is an instructive comparison case for infilled datasets. Observations are sparse but the signal is large. Recent reconstructions of hybrid air/sea-surface temperature also focused on this event (Vaccaro et al. 2021; Kadow et al. 2020) to illustrate how different the imputed values can be for different datasets.

Figure 10 shows the onset of El Niño in July 1877. We show the ensemble spread in our reconstructed SST and land T, the observations of SST and station temperatures, and two comparison datasets. ERSSTv5 depicts the center of action in the coastal-eastern Pacific, whereas the central Pacific is most notable in HadISST1. Our ensemble mean displays some commonalities with each dataset, but we find higher confidence in the east Pacific El Niño (Fig 10b), and we note that our method leverages teleconnections with observations of SLP and land temperatures to constrain the SSTs. Our results also show large uncertainties in the central Pacific and the coastal-eastern Pacific (Fig 10a); i.e., uncertainty in the type of ENSO (e.g., Newman et al. 2011; Karnauskas 2013; Capotondi et al. 2015). There are also large differences across datasets in the North Pacific. Around the Southern Ocean in ERSSTv5, the influence of the HadISST2 sea ice is evident as a ring of cold anomalies. This results from the expansion of Antarctic sea ice in HadISST2 (Figure 5l).

4. Discussion

With coupled DA, we provide a dynamically and observationally constrained perspective on coupled variability and trends over the historical record. These results suggest it may be worth revisiting assessments of forced versus internal variability and climate-model biases using this internally consistent reconstruction. Many studies have characterized post-1980 trends, but placing those changes in the context of the longer record may help disentangle the mechanisms and causes of both variability and trends. Several large-scale model biases, including those in the Southern Ocean and the Tropics, now appear less drastic than previously estimated, suggesting climate models may perform better than indicated by comparison with earlier datasets (e.g., Wills et al. 2022; Simpson et al. 2025).

a. Tropical trends

The zonal SST gradient and Walker circulation in the tropical Pacific has been a focus of many discussions of forced versus internal variability (Vecchi and Soden 2007; DiNezio et al. 2009; Coats and Karnauskas 2017; Kohyama et al. 2017; Seager et al. 2019; Lee et al. 2022; S. Kang et al. 2023b; Watanabe et al. 2024; Jiang et al. 2024). In our results, the 1979–2014 strengthening trend in the Walker circulation (Pacific zonal SLP gradient) does not appear distinct from variability over the historical record. The Pacific zonal SST gradient has a more notable trend from 1980–present, but it is difficult to convincingly say that the trend is outside of the range of natural variability.

Over the full twentieth century, we do not find a long-term weakening of the Walker circulation (Tokinaga et al. 2012; Vecchi et al. 2006) nor a clear strengthening of the zonal SST gradient (Coats and Karnauskas 2017; Seager et al. 2022) that is distinct from past variability. If the recent trend is a forced response to global warming from CO_2 (e.g, Clement et al. 1996; Seager et al. 2019), that trend is not yet distinct from past variability in the reconstruction.

However, our results indicate that there is a peculiar trend from c. 1975–present in the strengthening of the SST contrast between the warmest SSTs and the mean SST over the entire Tropics ($SST^{\#}$; Fueglistaler and Silvers 2021; Zhang and Fueglistaler 2020; Fueglistaler 2019). Fueglistaler and Silvers (2021) questioned whether the recent trend in $SST^{\#}$ could be due to data artifacts in the SST record or purely coincidence, i.e., a rare occurrence of variability during the satellite record. Data artifacts are still a possible influence, but the dynamical constraints in our method reduce the likelihood of that explanation, especially considering that the SST is also informed by SLP observations and station temperatures. Further analysis of paleoclimate proxy data in the Tropics (e.g., Deutsch et al. 2014; Sanchez et al. 2020, 2021) could help assess the role of possible data artifacts and the range of natural variability in SST contrasts.

b. Southern annular mode

The positive trend in the SAM (c. 1980–present) has been associated with stratospheric ozone depletion, CO_2 forcing, natural variability, and other factors (Doddridge and Marshall 2017; Polvani et al. 2021; Bitz and Polvani 2012; Seviour et al. 2016; Thomas et al. 2015; Thompson et al. 2011; England et al. 2016; Fogt and Marshall 2020; Banerjee et al. 2020). Efforts to determine what has caused the SAM trend have been complicated by recent results, included in Figure 5j, depicting a positive trend over the entire twentieth century (O'Connor et al. 2021; Dalaiden et al. 2021; King et al. 2023; Slivinski et al. 2019). Our findings, which show no trend from 1925–1970, then a prolonged positive trend from 1970–present, are consistent with a trend onset that is associated with stratospheric ozone depletion (Thompson and Solomon 2002; Fogt et al. 2009; Polvani et al. 2011; Thompson et al. 2011). The regression-based reconstruction of F09 is in general agreement with our results over 1920–1970, showing no significant SAM trend until a positive trend emerges in DJF around 1970 (Fogt et al. 2009). We also find that DJF has the strongest SAM trend over 1970–present (Fig. S16). Another large positive trend over 1850–1920 warrants further investigation into possible drivers and the role of data quality, particularly given the sparse and imperfect SLP observations in the early record. We note that Brönnimann et al. (2024) also report positive SAM in the early 1900s in newly digitized ship data, supporting our results.

c. Southern Ocean cooling

Studies of the post-1980 cooling in the Southern Ocean typically use SSTs from NOAA ERSST, the latest of which is Version 5 (Huang et al. 2017). Even when nudging a climate model (CESM1) to ERA reanalysis winds, the model does not reproduce the Southern Ocean SST cooling from ERSST (Blanchard-Wrigglesworth et al. 2021; Dong et al. 2022a). Therefore, it seems that the winds alone cannot explain the SST cooling over the Southern Ocean (Dong et al. 2023), and other explanations have been proposed (e.g., Zhang et al. 2019; Haumann et al. 2020; Dong et al. 2022b; Swart et al. 2023; Schmidt et al. 2023).

Pacemaker experiments, which nudge a coupled climate model's SST in the Southern Ocean to match an infilled SST dataset (typically NOAA ERSST), have been used to investigate how SST cooling of the Southern Ocean affects global climate, radiative feedbacks, and the atmospheric circulation (Zhang et al. 2021; S. Kang et al. 2023c,a). The Southern Ocean cooling has also been proposed as a driver of cooling in the tropical east Pacific (Dong et al. 2022a), possibly forced by the ozone hole (Hartmann 2022) or other means (Watanabe et al. 2024). J. Kang et al. (2024) leverage the pacemaker experiments, but they also highlight the importance of regional-scale discrepancies in SST trends for the atmospheric circulation and uncertainty in post-1979 trends across reanalyses in the Southern Hemisphere.

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

In our results, we find much less cooling over the Southern Ocean compared to NOAA ERSSTv5. While more work is needed before definitive conclusions can be made about which reconstruction is more accurate, we compare the non-infilled SST dataset that we use to inform our data assimilation, HadSST4, with the non-infilled SST data from ERSSTv5 and from a recent product that has undergone extensive bias corrections (DCENT, Chan et al. 2024). Then we also compare our results with 1980–2023 trends in other infilled SST datasets.

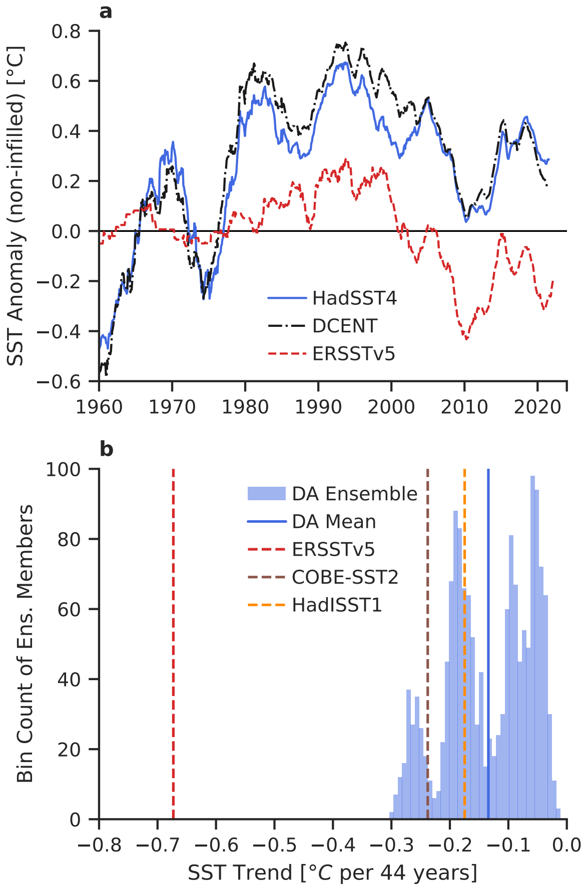

FIG. 11. **Recent evolution of Southern Ocean SST: comparing in situ data and infilled trends in the southeast-Pacific sector.** Both panels analyze the mean of values in the region west of the Drake Passage, spanning latitudes 50°S to 70°S and longitudes 70°W to 140°W. **(a)** Comparison of non-infilled SST anomalies, illustrating differences from the homogenization of time-varying in situ sources; for visual clarity, 5-yr running mean is applied and the 1961–1979 mean is removed. **(b)** Infilled SST trends for 1980–2023 from data assimilation (DA), with 1600 ensemble members shown as histogram; the distribution is shaped by the eight distinct model priors. Vertical lines indicate the mean trend and comparison datasets.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

Figure 11a compares the non-infilled anomalies in the southeast-Pacific sector of the Southern Ocean (latitudes 50°S–70°S and longitudes 70°W–140°W). HadSST4 and DCENT show similar trajectories, but they have a substantial offsets relative to ERSSTv5. This suggests that not only the infilling but also the homogenization of time-varying data sources affects trends in this region. Kennedy et al. (2019) show the major transition from bucket measurements to drifting buoys between 1980 and 2005, and Huang et al. (2019) find substantial differences in SST analyses from 2000–2016 when including drifting buoys and/or ARGO floats in NOAA ERSSTv5. ERSSTv5 has a detailed bias-correction procedure and consequently could provide the best estimate in this region. A key point is that the processing of time-varying data sources could have a spurious influence on what appear to be climate trends.

Figure 11b shows the distribution of 1980–2023 SST trends in the southeast-Pacific sector of the Southern Ocean (latitudes 50°S–70°S and longitudes 70°W–140°W). Our reconstruction shows a wide range of uncertainty, with possible trends ranging from −0.3°C to 0.0°C (44 yr)$^{-1}$. Our distribution is shaped by the uncertainty in bias corrections from HadSST4 and by the eight LIMs used as priors in the assimilation. COBE-SST2 and HadISST1 are within our uncertainty range, but ERSSTv5 has a much larger trend of −0.7°C (44 yr)$^{-1}$. Determining which of these trends is correct may be important for advancing understanding of the mechanisms driving Southern Ocean cooling. For example, nudging a climate model's winds to reanalysis may not explain the magnitude of cooling in ERSSTv5, but wind-nudging might be sufficient to explain all of the cooling in our reconstruction. Thus the result of weaker cooling in our reconstruction supports the notion that ozone depletion, through its influence on the SAM and surface winds, may be a key driver of the observed SST trends (Hartmann 2022). The possibility that our reconstruction is closer to the true (but unknown) trend motivates revisiting investigations of Southern Ocean cooling, as well as its impacts on the tropical Pacific and global climate (e.g., S. Kang et al. 2023a,c).

d. Radiative feedbacks and historical pattern effects

The pattern effect on climate sensitivity, i.e., the dependence of radiative feedbacks on spatial patterns of SST and SIC anomalies (Armour et al. 2013; Andrews et al. 2015; Zhou et al. 2016; Ceppi and Gregory 2017; Andrews and Webb 2018; Fueglistaler 2019; Dong et al. 2019, 2020; Cooper et al. 2024), has strong ties to the incomplete-data problem. The pattern effect over the

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

historical record (Andrews et al. 2018, 2022; Marvel et al. 2018; Salvi et al. 2023; Armour et al. 2024) depends on what the SST patterns were in the past, and recent studies have revealed that differences across infilled SST datasets lead to disparate interpretations of the historical pattern effect (Fueglistaler and Silvers 2021; Lewis and Mauritsen 2021), or possibly no pattern effect at all (Modak and Mauritsen 2023).

Uncertainty in sea ice is typically omitted from studies of the pattern effect, but Andrews et al. (2018, SI) found that differences in sea ice between AMIP-II and HadISST2 change the shortwave clear-sky feedback by approximately 0.6 W m^{-2} K^{-1}. This change from sea ice alone is approximately the same magnitude as the total pattern effect over the historical record, as the mean pattern effect is 0.48 W m^{-2} K^{-1} using HadISST1 (Andrews et al. 2022). Constraining uncertainty in Antarctic sea ice is important for quantifying historical pattern effects.

We find many differences in the spatial patterns of SST and SIC anomalies relative to AMIP-II and HadISST1, which have been used to account for historical pattern effects and quantify variability in feedbacks over the historical record (Zhou et al. 2016; Andrews et al. 2018; Marvel et al. 2018; Dong et al. 2019; Gregory et al. 2020; Sherwood et al. 2020; Lewis and Mauritsen 2021; Andrews et al. 2022; Salvi et al. 2023; Modak and Mauritsen 2023). Our reconstruction of monthly SST and SIC can be used as boundary conditions in atmospheric general circulation models to examine the implications for historical feedbacks, pattern effects, and climate sensitivity.

e. Future opportunities and caveats of the method

Future efforts to reconstruct the historical record could improve on our results in a variety of ways, and we list a few of them here:

- LIMs and DA: Future investigations could elaborate on optimizing the LIMs, their training data, and possibly consider machine-learning methods (e.g., Meng and Hakim 2024). Our method uses climate models to train the LIMs, and therefore inherits some of the problems in climate models. We mitigate this effect by using eight different CMIP6 models to sample the range of systematic uncertainty and through DA. There are many varieties of DA that could improve on our results, including 4D-Var, quantile-conserving filtering, or multi-model Kalman filtering with a large ensemble generated by various LIMs (Kalnay 2003; Houtekamer and Zhang 2016; Anderson 2022; Bach and Ghil 2023). Our method assumes state variables can be approximated with Gaussian distributions, which appears to work reasonably well for SIC but could likely be improved in future studies.

- Pressure data: An update of non-infilled HadSLP2 (Allan and Ansell 2006) would be helpful, as no quality-controlled dataset of gridded monthly mean SLP with error estimates is currently available. ICOADS provides only marine data (Freeman et al. 2017) and does not include observation errors. Including terrestrial pressure data (Cram et al. 2015) could improve our reconstruction, but no gridded product exists, and elevation differences are a considerable source of error.

- Sea ice: There are many observations available before the satellite era (e.g., Walsh et al. 2019; Edinburgh and Day 2016; Titchner and Rayner 2014), but we do not have a current compilation of this data in a format that can be used in reconstructions. A dataset structured like HadSST4 or DCENT but with historical SIC observations would be helpful.

- SST: Ongoing efforts to digitize new data, quantify error, and correct the biases of existing data will continue to be critical (e.g., Brönnimann et al. 2019, 2024; Chan et al. 2019, 2023; Kent and Kennedy 2021; Kennedy et al. 2019). For SST anomalies (also T and SLP), it would be helpful to use a climatological period that overlaps with satellite observations of SIC (i.e., post-1979).

5. Conclusions

The historical record is essential to our understanding of coupled climate dynamics and variability, but instrumental observations are sparse and noisy. Moreover, existing observational datasets are typically derived separately for each component of the climate system, leading to inconsistencies in coupled variability when they are combined.

In this study, we develop a method for climate reanalysis using strongly coupled data assimilation. The key advance of our method compared to past work is that we (i) ensure that the coupled atmosphere-ocean-ice state is internally consistent and (ii) synthesize observational and dynamical constraints across all components simultaneously. Using a Kalman filter, we combine monthly forecasts from linear inverse models (LIMs), which are trained on eight CMIP6 models to account

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

for model error, with observations of SST, land temperature, marine sea-level pressure, and satellite-era sea ice.

We first validate the method through pseudo-reconstruction of an out-of-sample climate model, then we present the actual reconstruction on a global 2° × 2° grid with monthly resolution of SST, near-surface air temperature, sea-level pressure, and sea-ice concentration over 1850–2023. We also provide a novel quantification of the time-varying uncertainty in all fields and its spatial fingerprints.

In many ways, our results differ from comparison datasets regarding how recent trends (c. 1980–present) compare to past variability. The recent evolution of the Walker circulation appears consistent with past variability, but the SST contrast ($\mathrm{SST}^{\#}$; warmest regions versus the tropical mean) exhibits a prolonged strengthening from 1975–present that appears distinct from past variability.

In the Southern Ocean, we find a weaker SST cooling post-1980 compared to the strong cooling in other estimates (namely ERSSTv5), which climate models have been unable to replicate. We emphasize the observational uncertainty over the Southern Ocean, which merits more attention due to sparse and problematic data even after 1980. The Southern Annular Mode appears well constrained but differs substantially from existing estimates before 1980. Antarctic sea ice also follows a different trajectory in our reconstruction compared to other estimates over the majority of the record (1850–1980). Our constraints on Antarctic sea ice are a key result, as we find much less ice loss over 1900–1980 compared to existing datasets, but with large uncertainty.

Our historical reconstruction is designed for climate analysis and is publicly available. We provide the grand mean of all 1600 ensemble members, the separate ensemble means for each of the eight model priors, and a subset of 200 fully gridded ensemble members. Our monthly SST and sea ice can also be used as boundary conditions in atmospheric general circulation models (i.e., in AMIP-type simulations). Through coupled data assimilation, this reconstruction improves constraints on coupled climate dynamics and variability, highlights key uncertainties in the historical record, and guides future investigations into coupled atmosphere–ocean–ice interactions.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

Acknowledgments. The authors acknowledge funding from: a National Defense Science and Engineering Graduate (NDSEG) Fellowship from the Department of Defense (V.T.C.); National Science Foundation (NSF) awards OCE-2002276 (V.T.C., K.C.A., and G.J.H.), AGS-1752796 (V.T.C. and K.C.A.), AGS-2203543 (V.T.C. and K.C.A.), AGS-2402475 (G.J.H), OPP-2213988 (G.J.H), AGS-2202526 (G.J.H); Heising-Simons Foundation award 2023-4715 (G.J.H); and a Calvin Professorship in Oceanography (K.C.A.). The authors acknowledge computing support from the Casper system (`https://ncar.pub/casper`) and the Derecho system (`https://doi.org/10.5065/qx9a-pg09`) provided by the NSF National Center for Atmospheric Research (NCAR), sponsored by the National Science Foundation. The authors acknowledge helpful feedback from Dominik Stiller, early discussions with Gil Compo and Laura Slivinski, and expert descriptions of datasets from the Climate Data Guide (Schneider et al. 2013) and `reanalyses.org`, which is maintained by CIRES and supported by NOAA/ESRL Physical Sciences Division, the NOAA Climate Program Office, and the US Department of Energy's (DOE) Office of Science. The authors declare no conflicts of interest.

Data availability statement. The reconstruction will be publicly available in a Zenodo repository upon acceptance.

Observation data is available as follows: HadSST4.0.1.0 at `https://www.metoffice.gov.uk/hadobs/hadsst4/index.html` at CRUTEM.5.0.2.0, `https://www.metoffice.gov.uk/hadobs/crutem5/data/CRUTEM.5.0.2.0/download.html`; ICOADS SLP at `https://downloads.psl.noaa.gov/Datasets/icoads/2degree/enh/`; NOAA/NSIDC CDRv4 at `https://doi.org/10.7265/efmz-2t65` and NRTv2 at `https://doi.org/10.7265/tgam-yv28`.

SST and SIC comparison data is available as follows: PCMDI/AMIP-II at `https://aims2.llnl.gov/` with specifier input4MIPs.CMIP6Plus.CMIP.PCMDI.PCMDI-AMIP-1-1-9; HadISST1 at `https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html`, HadISST2.1.0.0 SST at `https://www.metoffice.gov.uk/hadobs/hadisst2/data/HadISST.2.1.0.0/index.html` and HadISST2.2.0.0 SIC at `https://www.metoffice.gov.uk/hadobs/hadisst2/data/download.html`; COBE-SST2 at `https://downloads.psl.noaa.gov/Datasets/COBE2/`; ERSSTv5 from `https://www.ncei.noaa.gov/products/`

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

`extended-reconstructed-sst` or `https://doi.org/10.5065/JZ08-3W17`. Brennan and Hakim (2022) is available at `https://doi.org/10.5281/zenodo.5809703`

GMSAT comparison data is available as follows: BEST at `https://climate.metoffice.cloud/temperature.html`; HadCRUT5 at `https://www.metoffice.gov.uk/hadobs/hadcrut5/data/HadCRUT.5.0.2.0/download.html`.

SLP comparison data is available as follows: NOAA/CIRES/DOE 20th Century Reanalysis (V3) and additional datasets listed here are provided by the NOAA PSL, Boulder, Colorado, USA, from their website at `https://psl.noaa.gov`; NCEP/NCAR Reanalysis at `https://downloads.psl.noaa.gov//Datasets/ncep.reanalysis/Monthlies/surface/`; ERA5 post-1979 at `https://doi.org/10.5065/P8GT-0R61` and 1950–1978 back extension at `https://doi.org/10.5065/JAXB-X906`; HadSLP2 at `https://www.metoffice.gov.uk/hadobs/hadslp2/data/download.html`; King et al. (2023) at `https://doi.org/10.5281/zenodo.8156908`; O'Connor et al. (2021) at `https://doi.org/10.5281/zenodo.5507607`; Dalaiden et al. (2021) at `https://doi.org/10.5281/zenodo.4770179`; Fogt et al. (2009) at `https://polarmet.osu.edu/ACD/sam/sam_recon.html`; Falster et al. (2023) at `https://doi.org/10.5281/zenodo.8280559`.

CMIP6 output is available on NCAR Glade and at `https://esgf-node.llnl.gov/search/cmip6/`.

APPENDIX A

Summary of LIM training data

Model	Total Years (piControl range)	Ens. Mem.	EOFs	Reference
CESM2	1166 (200–1200)	r1i1p1f1	408	Danabasoglu et al. (2020)
UKESM1.0	1754 (2250–3839)	r1i1p1f2	408	Sellar et al. (2019)
SAM0-UNICON	865 (1–700)	r1i1p1f1	306	Park et al. (2019)
GFDL-ESM4	665 (1–500)	r1i1p1f1	306	Dunne et al. (2020)
NorESM2-LM	666 (1600–2100)	r1i1p1f1	306	Seland et al. (2020)
EC-Earth3	1165 (2103–3102)	r2i1p1f1	408	Döscher et al. (2022)
HadGEM3-GC31-LL	2165 (1850–3849)	r1i1p1f1	408	Kuhlbrodt et al. (2018)
E3SM-2	665 (1–500)	r1i1p1f1	306	Qin et al. (2024)

Table A1. CMIP6 training data for 8 linear inverse models. All models with 408 EOFs have the following distribution across state variables: 108 SST, 108 T, 48 SLP, 72 Arctic SIC, 72 Antarctic SIC. Models with 306 EOFs have 92 SST, 84 T, 30 SLP, 50 Arctic SIC, 50 Antarctic SIC. Note that Total Years includes piControl plus 165 years of historical simulation (1850–2014).

APPENDIX B

Observation error for sea-level pressure

To estimate **R** for observations of monthly mean SLP, we apply a method similar to that in Kaplan et al. (2000). The intramonth standard deviation (s) provided by ICOADS is comprised of submonthly variability, measurement error, and representativeness error, thus providing an estimate of the observation error in the monthly mean (Leith 1973). We take the local time-average of $s^2 \frac{n_{\text{obs}}}{n_{\text{obs}}-1}$ over the well observed period 1961–2023 to estimate the climatological error variance, σ^2, in the monthly mean for each gridcell, and we restrict the estimate to gridcells with $n_{\text{obs}} > 30$ in a given month. Again using a similar approach to Kaplan et al. (2000), we then spatially smooth the resulting climatological maps of σ using a running-mean window of 12° latitude × 50° longitude equatorward of 52°N/S and a window of 18° latitude × 100° longitude poleward of 52°N/S. This results in 12 monthly 2° × 2° fields of the random measurement and sampling error, σ_{random}.

We then must assign a time-varying error, σ, to each monthly value of SLP. We start with the random error described above, then reduce the random error by the number of intramonth observations in a gridcell. To account for autocovariance and possible sampling errors even when n_{obs} is large, reduce n_{obs} to $n_{\text{adjusted}} = n_{\text{obs}}/2$, and we set the maximum of n_{adjusted} at 30 (Leith 1973; Bretherton et al. 1999). We then consider the systematic component of the total error, $\sigma^2 = \sigma^2_{\text{systematic}} + \sigma^2_{\text{random}}/n_{\text{adjusted}}$, as discussed in Kennedy (2014). We estimate $\sigma^2_{\text{systematic}}$ from the variance across neighboring observations. The idea is that if neighboring observations consistently differ, the differences are from irreducible, systematic errors. Separately for each month from 1961–2023, we calculate the spatial variance across a running-mean window of 16° latitude × 32° longitude, restricting the calculation to gridcells with $n_{\text{obs}} \geq 5$. We use the zonal mean of the climatology of this field to represent $\sigma^2_{\text{systematic}}$. We make one adjustment by setting the minimum $\sigma_{\text{systematic}}$ at 6 hPa south of 72°S, preventing the error from decreasing near the Antarctic coastline. The systematic error ranges from approximately 1 hPa on the equator to 7 hPa in polar regions, with a local maximum of 9.5 hPa over the Southern Ocean at 55°S.

References

Alexander, M. A., I. Bladé, M. Newman, J. R. Lanzante, N.-C. Lau, and J. D. Scott, 2002: The Atmospheric Bridge: The Influence of ENSO Teleconnections on Air–Sea Interaction over the Global Oceans. *J. Climate*, **15 (16)**, 2205–2231, https://doi.org/10.1175/1520-0442(2002)015⟨2205:TABTIO⟩2.0.CO;2.

Allan, R., and T. Ansell, 2006: A New Globally Complete Monthly Historical Gridded Mean Sea Level Pressure Dataset (HadSLP2): 1850–2004. *J. Climate*, **19 (22)**, 5816–5842, https://doi.org/10.1175/JCLI3937.1.

Amaya, D. J., N. Maher, C. Deser, M. G. Jacox, M. A. Alexander, M. Newman, J. Dias, and J. Lou, 2025: Linking Projected Changes in Seasonal Climate Predictability and ENSO Amplitude. *J. Climate*, **38 (3)**, 675–688, https://doi.org/10.1175/JCLI-D-23-0648.1.

Amrhein, D. E., G. J. Hakim, and L. A. Parsons, 2020: Quantifying Structural Uncertainty in Paleoclimate Data Assimilation With an Application to the Last Millennium. *Geophys. Res. Lett.*, **47 (22)**, https://doi.org/10.1029/2020GL090485.

Anderson, J. L., 2022: A Quantile-Conserving Ensemble Filter Framework. Part I: Updating an Observed Variable. *Monthly Weather Review*, **150 (5)**, 1061–1074, https://doi.org/10.1175/MWR-D-21-0229.1.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

Andrews, T., J. M. Gregory, and M. J. Webb, 2015: The Dependence of Radiative Forcing and Feedback on Evolving Patterns of Surface Temperature Change in Climate Models. *J. Climate*, **28 (4)**, 1630–1648, https://doi.org/10.1175/JCLI-D-14-00545.1.

Andrews, T., and M. J. Webb, 2018: The Dependence of Global Cloud and Lapse Rate Feedbacks on the Spatial Structure of Tropical Pacific Warming. *J. Climate*, **31 (2)**, 641–654, https://doi.org/10.1175/JCLI-D-17-0087.1.

Andrews, T., and Coauthors, 2018: Accounting for Changing Temperature Patterns Increases Historical Estimates of Climate Sensitivity. *Geophys. Res. Lett.*, **45 (16)**, 8490–8499, https://doi.org/10.1029/2018GL078887.

Andrews, T., and Coauthors, 2022: On the Effect of Historical SST Patterns on Radiative Feedback. *J. Geophys. Res.: Atmospheres*, **127 (18)**, https://doi.org/10.1029/2022JD036675.

Annan, J. D., J. C. Hargreaves, and T. Mauritsen, 2022: A new global surface temperature reconstruction for the Last Glacial Maximum. *Climate of the Past*, **18 (8)**, 1883–1896, https://doi.org/10.5194/cp-18-1883-2022.

Armour, K. C., C. M. Bitz, and G. H. Roe, 2013: Time-Varying Climate Sensitivity from Regional Feedbacks. *J. Climate*, **26 (13)**, 4518–4534, https://doi.org/10.1175/JCLI-D-12-00544.1.

Armour, K. C., J. Marshall, J. R. Scott, A. Donohoe, and E. R. Newsom, 2016: Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. *Nature Geoscience*, **9 (7)**, 549–554, https://doi.org/10.1038/ngeo2731.

Armour, K. C., and Coauthors, 2024: Sea-surface temperature pattern effects have slowed global warming and biased warming-based constraints on climate sensitivity. *Proc. Natl. Acad. Sci. (USA)*, **121 (12)**, e2312093 121, https://doi.org/10.1073/pnas.2312093121.

Bach, E., and M. Ghil, 2023: A Multi-Model Ensemble Kalman Filter for Data Assimilation and Forecasting. *Journal of Advances in Modeling Earth Systems*, **15 (1)**, https://doi.org/10.1029/2022MS003123.

Banerjee, A., J. C. Fyfe, L. M. Polvani, D. Waugh, and K.-L. Chang, 2020: A pause in Southern Hemisphere circulation trends due to the Montreal Protocol. *Nature*, **579 (7800)**, 544–548, https://doi.org/10.1038/s41586-020-2120-4.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

Barsugli, J. J., and D. S. Battisti, 1998: The basic effects of atmosphere-ocean thermal coupling on midlatitude variability. *Journal of the Atmospheric Sciences*, https://doi.org/10.1175/1520-0469(1998)055⟨0477:TBEOAO⟩2.0.CO;2.

Battisti, D. S., D. J. Vimont, and B. P. Kirtman, 2019: 100 Years of Progress in Understanding the Dynamics of Coupled Atmosphere–Ocean Variability. *Meteorological Monographs*, **59**, 1–8, https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0025.1.

Bitz, C. M., and L. M. Polvani, 2012: Antarctic climate response to stratospheric ozone depletion in a fine resolution ocean climate model. *Geophys. Res. Lett.*, **39 (20)**, https://doi.org/10.1029/2012GL053393.

Bjerknes, J., 1969: Atmospheric Teleconnections from the Equatorial Pacific. *Monthly Weather Review*, **97 (3)**, 163–172, https://doi.org/10.1175/1520-0493(1969)097⟨0163:ATFTEP⟩2.3.CO;2.

Blanchard-Wrigglesworth, E., L. A. Roach, A. Donohoe, and Q. Ding, 2021: Impact of Winds and Southern Ocean SSTs on Antarctic Sea Ice Trends and Variability. *J. Climate*, **34 (3)**, 949–965, https://doi.org/10.1175/JCLI-D-20-0386.1.

Bloch-Johnson, J., and Coauthors, 2024: The Green's Function Model Intercomparison Project (GFMIP) Protocol. *Journal of Advances in Modeling Earth Systems*, **16 (2)**, https://doi.org/10.1029/2023MS003700.

Bonan, D. B., J. Dörr, R. C. J. Wills, A. F. Thompson, and M. Årthun, 2024: Sources of low-frequency variability in observed Antarctic sea ice. *The Cryosphere*, **18 (4)**, 2141–2159, https://doi.org/10.5194/tc-18-2141-2024.

Brennan, M. K., and G. J. Hakim, 2022: Reconstructing Arctic Sea Ice over the Common Era Using Data Assimilation. *J. Climate*, **35 (4)**, 1231–1247, https://doi.org/10.1175/JCLI-D-21-0099.1.

Brennan, M. K., G. J. Hakim, and E. Blanchard-Wrigglesworth, 2023: Monthly Arctic Sea-Ice Prediction With a Linear Inverse Model. *Geophys. Res. Lett.*, **50 (7)**, e2022GL101 656, https://doi.org/10.1029/2022GL101656.

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

Bretherton, C. S., M. Widmann, V. P. Dymnikov, J. M. Wallace, and I. Bladé, 1999: The effective number of spatial degrees of freedom of a time-varying field. *J. Climate*, https://doi.org/10.1175/1520-0442(1999)012⟨1990:TENOSD⟩2.0.CO;2.

Brönnimann, S., 2009: Early twentieth-century warming. *Nature Geoscience*, **2 (11)**, 735–736, https://doi.org/10.1038/ngeo670.

Brönnimann, S., Y. Brugnara, and C. Wilkinson, 2024: Early 20th century Southern Hemisphere cooling. *Climate of the Past*, **20 (3)**, 757–767, https://doi.org/10.5194/cp-20-757-2024.

Brönnimann, S., and Coauthors, 2019: Unlocking Pre-1850 Instrumental Meteorological Records: A Global Inventory. *Bull. Amer. Meteor. Soc.*, **100 (12)**, ES389–ES413, https://doi.org/10.1175/BAMS-D-19-0040.1.

Buckley, E. M., C. Horvat, and P. Yoosiri, 2024: Sea Ice Concentration Estimates from ICESat-2 Linear Ice Fraction. Part 1: Multi-sensor Comparison of Sea Ice Concentration Products. URL https://egusphere.copernicus.org/preprints/2024/egusphere-2024-3861/, https://doi.org/10.5194/egusphere-2024-3861.

Burgers, G., P. Jan van Leeuwen, and G. Evensen, 1998: Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review*, **126 (6)**, 1719–1724, https://doi.org/10.1175/1520-0493(1998)126⟨1719:ASITEK⟩2.0.CO;2.

Cai, W., and Coauthors, 2021: Changing El Niño–Southern Oscillation in a warming climate. *Nature Reviews Earth & Environment*, **2 (9)**, 628–644, https://doi.org/10.1038/s43017-021-00199-z.

Cai, W., and Coauthors, 2023: Anthropogenic impacts on twentieth-century ENSO variability changes. *Nature Reviews Earth & Environment*, **4 (6)**, 407–418, https://doi.org/10.1038/s43017-023-00427-8.

Callahan, C. W., C. Chen, M. Rugenstein, J. Bloch-Johnson, S. Yang, and E. J. Moyer, 2021: Robust decrease in El Niño/Southern Oscillation amplitude under long-term warming. *Nature Climate Change*, **11 (9)**, 752–757, https://doi.org/10.1038/s41558-021-01099-2.

Cane, M. A., A. C. Clement, A. Kaplan, Y. Kushnir, D. Pozdnyakov, R. Seager, S. E. Zebiak, and R. Murtugudde, 1997: Twentieth-Century Sea Surface Temperature Trends. *Science*, **275 (5302)**, 957–960, https://doi.org/10.1126/science.275.5302.957.

Capotondi, A., and Coauthors, 2015: Understanding ENSO Diversity. *Bull. Amer. Meteor. Soc.*, **96 (6)**, 921–938, https://doi.org/10.1175/BAMS-D-13-00117.1.

Capotondi, A., and Coauthors, 2023: Mechanisms of tropical Pacific decadal variability. *Nature Reviews Earth & Environment*, **4 (11)**, 754–769, https://doi.org/10.1038/s43017-023-00486-x.

Ceppi, P., and J. M. Gregory, 2017: Relationship of tropospheric stability to climate sensitivity and Earth's observed radiation budget. *Proc. Natl. Acad. Sci. (USA)*, **114 (50)**, 13 126–13 131, https://doi.org/10.1073/pnas.1714308114.

Chadwick, R., P. Good, T. Andrews, and G. Martin, 2014: Surface warming patterns drive tropical rainfall pattern responses to CO ¡sub¿2¡/sub¿ forcing on all timescales. *Geophys. Res. Lett.*, **41 (2)**, 610–615, https://doi.org/10.1002/2013GL058504.

Chan, D., G. Gebbie, and P. Huybers, 2023: Global and Regional Discrepancies between Early-Twentieth-Century Coastal Air and Sea Surface Temperature Detected by a Coupled Energy-Balance Analysis. *J. Climate*, **36 (7)**, 2205–2220, https://doi.org/10.1175/JCLI-D-22-0569.1.

Chan, D., G. Gebbie, P. Huybers, and E. C. Kent, 2024: A Dynamically Consistent ENsemble of Temperature at the Earth surface since 1850 from the DCENT dataset. *Scientific Data*, **11 (1)**, 953, https://doi.org/10.1038/s41597-024-03742-x.

Chan, D., and P. Huybers, 2019: Systematic Differences in Bucket Sea Surface Temperature Measurements among Nations Identified Using a Linear-Mixed-Effect Method. *J. Climate*, **32 (9)**, 2569–2589, https://doi.org/10.1175/JCLI-D-18-0562.1.

Chan, D., E. C. Kent, D. I. Berry, and P. Huybers, 2019: Correcting datasets leads to more homogeneous early-twentieth-century sea surface warming. *Nature*, **571 (7765)**, 393–397, https://doi.org/10.1038/s41586-019-1349-2.

Chemke, R., Y. Ming, and J. Yuval, 2022: The intensification of winter mid-latitude storm tracks in the Southern Hemisphere. *Nature Climate Change*, **12 (6)**, 553–557, https://doi.org/10.1038/s41558-022-01368-8.

Chung, E.-S., A. Timmermann, B. J. Soden, K.-J. Ha, L. Shi, and V. O. John, 2019: Reconciling opposing Walker circulation trends in observations and model projections. *Nature Climate Change*, **9 (5)**, 405–412, https://doi.org/10.1038/s41558-019-0446-4.

Clement, A. C., R. Seager, M. A. Cane, and S. E. Zebiak, 1996: An Ocean Dynamical Thermostat. *J. Climate*, **9 (9)**, 2190–2196, https://doi.org/10.1175/1520-0442(1996)009⟨2190:AODT⟩2.0.CO;2.

Coats, S., and K. B. Karnauskas, 2017: Are Simulated and Observed Twentieth Century Tropical Pacific Sea Surface Temperature Trends Significant Relative to Internal Variability? *Geophys. Res. Lett.*, **44 (19)**, 9928–9937, https://doi.org/10.1002/2017GL074622.

Coats, S., J. E. Smerdon, S. Stevenson, J. T. Fasullo, B. Otto-Bliesner, and T. R. Ault, 2020: Paleoclimate Constraints on the Spatiotemporal Character of Past and Future Droughts. *J. Climate*, **33 (22)**, 9883–9903, https://doi.org/10.1175/JCLI-D-20-0004.1.

Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137 (654)**, 1–28, https://doi.org/10.1002/qj.776.

Cook, B. I., and Coauthors, 2022: Megadroughts in the Common Era and the Anthropocene. *Nature Reviews Earth & Environment*, **3 (11)**, 741–757, https://doi.org/10.1038/s43017-022-00329-1.

Cooper, V. T., and Coauthors, 2024: Last Glacial Maximum pattern effects reduce climate sensitivity estimates. *Science Advances*, **10 (16)**, 9461, https://doi.org/10.1126/sciadv.adk9461.

Cornes, R. C., E. Kent, D. Berry, and J. J. Kennedy, 2020: CLASSnmat: A global night marine air temperature data set, 1880–2019. *Geoscience Data Journal*, **7 (2)**, 170–184, https://doi.org/10.1002/gdj3.100.

Cowtan, K., R. Rohde, and Z. Hausfather, 2018: Evaluating biases in sea surface temperature records using coastal weather stations. *Quart. J. Roy. Meteor. Soc.*, **144 (712)**, 670–681, https://doi.org/10.1002/qj.3235.

Cowtan, K., and R. G. Way, 2014: Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Quart. J. Roy. Meteor. Soc.*, **140 (683)**, 1935–1944, https://doi.org/10.1002/qj.2297.

Cram, T. A., and Coauthors, 2015: The International Surface Pressure Databank version 2. *Geoscience Data Journal*, **2 (1)**, 31–46, https://doi.org/10.1002/gdj3.25.

Czaja, A., C. Frankignoul, S. Minobe, and B. Vannière, 2019: Simulating the Midlatitude Atmospheric Circulation: What Might We Gain From High-Resolution Modeling of Air-Sea Interactions? *Current Climate Change Reports*, **5 (4)**, 390–406, https://doi.org/10.1007/s40641-019-00148-5.

Dalaiden, Q., H. Goosse, J. Rezsöhazy, and E. R. Thomas, 2021: Reconstructing atmospheric circulation and sea-ice extent in the West Antarctic over the past 200 years using data assimilation. *Climate Dynamics*, **57 (11-12)**, 3479–3503, https://doi.org/10.1007/s00382-021-05879-6.

Danabasoglu, G., and Coauthors, 2020: The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, **12 (2)**, https://doi.org/10.1029/2019MS001916.

Deser, C., M. A. Alexander, S.-P. Xie, and A. S. Phillips, 2010a: Sea Surface Temperature Variability: Patterns and Mechanisms. *Annual Review of Marine Science*, **2 (1)**, 115–143, https://doi.org/10.1146/annurev-marine-120408-151453.

Deser, C., A. S. Phillips, and M. A. Alexander, 2010b: Twentieth century tropical sea surface temperature trends revisited. *Geophys. Res. Lett.*, **37 (10)**, https://doi.org/10.1029/2010GL043321.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J. Roy. Meteor. Soc.*, **131 (613)**, 3385–3396, https://doi.org/10.1256/qj.05.108.

Deutsch, C., and Coauthors, 2014: Centennial changes in North Pacific anoxia linked to tropical trade winds. *Science*, **345 (6197)**, 665–668, https://doi.org/10.1126/science.1252332.

DiNezio, P. N., A. C. Clement, G. A. Vecchi, B. J. Soden, B. P. Kirtman, and S. K. Lee, 2009: Climate Response of the Equatorial Pacific to Global Warming. *J. Climate*, **22 (18)**, 4873–4892, https://doi.org/10.1175/2009JCLI2982.1.

Doddridge, E. W., and J. Marshall, 2017: Modulation of the Seasonal Cycle of Antarctic Sea Ice Extent Related to the Southern Annular Mode. *Geophys. Res. Lett.*, **44 (19)**, 9761–9768, https://doi.org/10.1002/2017GL074319.

Dong, Y., K. C. Armour, D. S. Battisti, and E. Blanchard-Wrigglesworth, 2022a: Two-Way Teleconnections between the Southern Ocean and the Tropical Pacific via a Dynamic Feedback. *J. Climate*, **35** **(19)**, 6267–6282, https://doi.org/10.1175/JCLI-D-22-0080.1.

Dong, Y., K. C. Armour, M. D. Zelinka, C. Proistosescu, D. S. Battisti, C. Zhou, and T. Andrews, 2020: Intermodel Spread in the Pattern Effect and Its Contribution to Climate Sensitivity in CMIP5 and CMIP6 Models. *J. Climate*, **33** **(18)**, 7755–7775, https://doi.org/10.1175/JCLI-D-19-1011.1.

Dong, Y., A. G. Pauling, S. Sadai, and K. C. Armour, 2022b: Antarctic Ice-Sheet Meltwater Reduces Transient Warming and Climate Sensitivity Through the Sea-Surface Temperature Pattern Effect. *Geophys. Res. Lett.*, **49** **(24)**, https://doi.org/10.1029/2022GL101249.

Dong, Y., L. M. Polvani, and D. B. Bonan, 2023: Recent Multi-Decadal Southern Ocean Surface Cooling Unlikely Caused by Southern Annular Mode Trends. *Geophys. Res. Lett.*, **50** **(23)**, https://doi.org/10.1029/2023GL106142.

Dong, Y., C. Proistosescu, K. C. Armour, and D. S. Battisti, 2019: Attributing Historical and Future Evolution of Radiative Feedbacks to Regional Warming Patterns using a Green's Function Approach: The Preeminence of the Western Pacific. *J. Climate*, **32** **(17)**, 5471–5491, https://doi.org/10.1175/JCLI-D-18-0843.1.

Dong, Y., and Coauthors, 2021: Biased Estimates of Equilibrium Climate Sensitivity and Transient Climate Response Derived From Historical CMIP6 Simulations. *Geophys. Res. Lett.*, **48** **(24)**, https://doi.org/10.1029/2021GL095778.

Dörr, J. S., D. B. Bonan, M. Årthun, L. Svendsen, and R. C. J. Wills, 2023: Forced and internal components of observed Arctic sea-ice changes. *The Cryosphere*, **17** **(9)**, 4133–4153, https://doi.org/10.5194/tc-17-4133-2023.

Döscher, R., and Coauthors, 2022: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6. *Geoscientific Model Development*, **15** **(7)**, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022.

Dunne, J. P., and Coauthors, 2020: The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics. *Journal of Advances in Modeling Earth Systems*, **12 (11)**, https://doi.org/10.1029/2019MS002015.

Edinburgh, T., and J. J. Day, 2016: Estimating the extent of Antarctic summer sea ice during the Heroic Age of Antarctic Exploration. *The Cryosphere*, **10 (6)**, 2721–2730, https://doi.org/10.5194/tc-10-2721-2016.

Eisenman, I., W. N. Meier, and J. R. Norris, 2014: A spurious jump in the satellite record: has Antarctic sea ice expansion been overestimated? *The Cryosphere*, **8**, 1289–1296, https://doi.org/10.5194/tc-8-1289-2014.

Emile-Geay, J., G. J. Hakim, F. Viens, F. Zhu, and D. E. Amrhein, 2024: Temporal Comparisons Involving Paleoclimate Data Assimilation: Challenges & Remedies. *J. Climate*, **-1 (aop)**, https://doi.org/10.1175/JCLI-D-24-0101.1.

England, M. R., I. Eisenman, N. J. Lutsko, and T. J. Wagner, 2021: The Recent Emergence of Arctic Amplification. *Geophys. Res. Lett.*, **48 (15)**, https://doi.org/10.1029/2021GL094086.

England, M. R., L. M. Polvani, K. L. Smith, L. Landrum, and M. M. Holland, 2016: Robust response of the Amundsen Sea Low to stratospheric ozone depletion. *Geophys. Res. Lett.*, **43 (15)**, 8207–8213, https://doi.org/10.1002/2016GL070055.

Espinosa, Z. I., E. Blanchard-Wrigglesworth, and C. M. Bitz, 2024: Understanding the drivers and predictability of record low Antarctic sea ice in austral winter 2023. *Communications Earth & Environment*, **5 (1)**, 723, https://doi.org/10.1038/s43247-024-01772-2.

Espinosa, Z. I., and M. D. Zelinka, 2024: The Shortwave Cloud-SST Feedback Amplifies Multi-Decadal Pacific Sea Surface Temperature Trends: Implications for Observed Cooling. *Geophys. Res. Lett.*, **51 (18)**, https://doi.org/10.1029/2024GL111039.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, **99 (C5)**, https://doi.org/10.1029/94jc00572.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, **9 (5)**, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

Falster, G., B. Konecky, S. Coats, and S. Stevenson, 2023: Forced changes in the Pacific Walker circulation over the past millennium. *Nature*, https://doi.org/10.1038/s41586-023-06447-0.

Fan, T., C. Deser, and D. P. Schneider, 2014: Recent Antarctic sea ice trends in the context of Southern Ocean surface climate variations since 1950. *Geophys. Res. Lett.*, **41 (7)**, 2419–2426, https://doi.org/10.1002/2014GL059239.

Flannaghan, T. J., S. Fueglistaler, I. M. Held, S. Po-Chedley, B. Wyman, and M. Zhao, 2014: Tropical temperature trends in Atmospheric General Circulation Model simulations and the impact of uncertainties in observed SSTs. *J. Geophys. Res.: Atmospheres*, **119 (23)**, 327–340, https://doi.org/10.1002/2014JD022365.

Fogt, R. L., and C. J. Connolly, 2021: Extratropical Southern Hemisphere Synchronous Pressure Variability in the Early Twentieth Century. *J. Climate*, **34 (14)**, 5795–5811, https://doi.org/10.1175/JCLI-D-20-0498.1.

Fogt, R. L., Q. Dalaiden, and G. K. O'Connor, 2024: A comparison of South Pacific Antarctic sea ice and atmospheric circulation reconstructions since 1900. *Climate of the Past*, **20 (1)**, 53–76, https://doi.org/10.5194/cp-20-53-2024.

Fogt, R. L., and G. J. Marshall, 2020: The Southern Annular Mode: Variability, trends, and climate impacts across the Southern Hemisphere. *WIREs Climate Change*, **11 (4)**, https://doi.org/10.1002/wcc.652.

Fogt, R. L., J. Perlwitz, A. J. Monaghan, D. H. Bromwich, J. M. Jones, and G. J. Marshall, 2009: Historical SAM Variability. Part II: Twentieth-Century Variability and Trends from Reconstructions, Observations, and the IPCC AR4 Models*. *J. Climate*, **22 (20)**, 5346–5365, https://doi.org/10.1175/2009JCLI2786.1.

Fogt, R. L., D. P. Schneider, C. A. Goergens, J. M. Jones, L. N. Clark, and M. J. Garberoglio, 2019: Seasonal Antarctic pressure variability during the twentieth century from spatially complete reconstructions and CAM5 simulations. *Climate Dynamics*, **53 (3-4)**, 1435–1452, https://doi.org/10.1007/S00382-019-04674-8/METRICS.

Fogt, R. L., A. M. Sleinkofer, M. N. Raphael, and M. S. Handcock, 2022: A regime shift in seasonal total Antarctic sea ice extent in the twentieth century. *Nature Climate Change*, **12 (1)**, 54–62, https://doi.org/10.1038/s41558-021-01254-9.

Forster, P., and Coauthors, 2021: 2021: The Earth's energy budget, climate feedbacks, and climate sensitivity. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou, Eds., Cambridge Univ. Press, Cambridge, UK and New York, NY, chap. 7, https://doi.org/10.1017/9781009157896.009.

Franke, J., S. Brönnimann, J. Bhend, and Y. Brugnara, 2017: A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past climatic variations. *Scientific Data*, **4 (1)**, 170 076, https://doi.org/10.1038/sdata.2017.76.

Freeman, E., and Coauthors, 2017: ICOADS Release 3.0: a major update to the historical marine climate record. *International Journal of Climatology*, **37 (5)**, 2211–2232, https://doi.org/10.1002/joc.4775.

Fueglistaler, S., 2019: Observational Evidence for Two Modes of Coupling Between Sea Surface Temperatures, Tropospheric Temperature Profile, and Shortwave Cloud Radiative Effect in the Tropics. *Geophys. Res. Lett.*, **46 (16)**, 9890–9898, https://doi.org/10.1029/2019GL083990.

Fueglistaler, S., and L. Silvers, 2021: The Peculiar Trajectory of Global Warming. *J. Geophys. Res.: Atmospheres*, **126 (4)**, 1–15, https://doi.org/10.1029/2020JD033629.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Climate*, **30 (14)**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Giese, B. S., H. F. Seidel, G. P. Compo, and P. D. Sardeshmukh, 2016: An ensemble of ocean reanalyses for 1815–2013 with sparse observational input. *J. Geophys. Res.: Oceans*, **121 (9)**, 6891–6910, https://doi.org/10.1002/2016JC012079.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

Gong, D., and S. Wang, 1999: Definition of Antarctic Oscillation index. *Geophys. Res. Lett.*, **26 (4)**, 459–462, https://doi.org/10.1029/1999GL900003.

Goosse, H., Q. Dalaiden, F. Feba, B. Mezzina, and R. L. Fogt, 2024: A drop in Antarctic sea ice extent at the end of the 1970s. *Communications Earth & Environment*, **5 (1)**, 628, https://doi.org/10.1038/s43247-024-01793-x.

Gregory, J. M., T. Andrews, P. Ceppi, T. Mauritsen, and M. J. Webb, 2020: How accurately can the climate sensitivity to $$\hbox {CO}_{2}$$ be estimated from historical climate change? *Climate Dynamics*, **54 (1-2)**, 129–157, https://doi.org/10.1007/s00382-019-04991-y.

Hakim, G. J., J. Emile-Geay, E. J. Steig, D. Noone, D. M. Anderson, R. Tardif, N. Steiger, and W. A. Perkins, 2016: The last millennium climate reanalysis project: Framework and first results. *J. Geophys. Res.: Atmospheres*, **121 (12)**, 6745–6764, https://doi.org/10.1002/2016JD024751.

Hakim, G. J., C. Snyder, S. G. Penny, and M. Newman, 2022: Subseasonal Forecast Skill Improvement From Strongly Coupled Data Assimilation With a Linear Inverse Model. *Geophys. Res. Lett.*, **49 (11)**, https://doi.org/10.1029/2022GL097996.

Hartmann, D. L., 2022: The Antarctic ozone hole and the pattern effect on climate sensitivity. *Proc. Natl. Acad. Sci. (USA)*, **119 (35)**, https://doi.org/10.1073/pnas.2207889119.

Hasselmann, K., 1976: Stochastic climate models: Part I. Theory. *Tellus A: Dynamic Meteorology and Oceanography*, **28 (6)**, 473, https://doi.org/10.3402/tellusa.v28i6.11316.

Hastenrath, S., and L. Greischar, 1993: Circulation mechanisms related to northeast Brazil rainfall anomalies. *J. Geophys. Res.: Atmospheres*, **98 (D3)**, 5093–5102, https://doi.org/10.1029/92JD02646.

Haumann, F. A., N. Gruber, and M. Münnich, 2020: Sea-Ice Induced Southern Ocean Subsurface Warming and Surface Cooling in a Warming Climate. *AGU Advances*, **1 (2)**, https://doi.org/10.1029/2019AV000132.

Hausfather, Z., K. Cowtan, D. C. Clarke, P. Jacobs, M. Richardson, and R. Rohde, 2017: Assessing recent warming using instrumentally homogeneous sea surface temperature records. *Science*

Advances, **3 (1)**, https://doi.org/10.1126/SCIADV.1601207/SUPPL{_}FILE/1601207{_}SM.PDF.

Heede, U. K., and A. V. Fedorov, 2023: Colder Eastern Equatorial Pacific and Stronger Walker Circulation in the Early 21st Century: Separating the Forced Response to Global Warming From Natural Variability. *Geophys. Res. Lett.*, **50 (3)**, https://doi.org/10.1029/2022GL101020.

Hegerl, G. C., S. Brönnimann, A. Schurer, and T. Cowan, 2018: The early 20th century warming: Anomalies, causes, and consequences. *WIREs Climate Change*, **9 (4)**, https://doi.org/10.1002/wcc.522.

Hegerl, G. C., and Coauthors, 2019: Causes of climate change over the historical record. *Environmental Research Letters*, **14 (12)**, 123 006, https://doi.org/10.1088/1748-9326/AB4557.

Henderson, S. A., D. J. Vimont, and M. Newman, 2020: The Critical Role of Non-Normality in Partitioning Tropical and Extratropical Contributions to PNA Growth. *J. Climate*, **33 (14)**, 6273–6295, https://doi.org/10.1175/JCLI-D-19-0555.1.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146 (730)**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-Scale Sea Surface Temperature Analysis and Its Uncertainty. *J. Climate*, **27 (1)**, 57–75, https://doi.org/10.1175/JCLI-D-12-00837.1.

Hoerling, M., J. Eischeid, and J. Perlwitz, 2010: Regional Precipitation Trends: Distinguishing Natural Variability from Anthropogenic Forcing. *J. Climate*, **23 (8)**, 2131–2145, https://doi.org/10.1175/2009JCLI3420.1.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review*, **126 (3)**, 796–811, https://doi.org/10.1175/1520-0493(1998)126⟨0796:DAUAEK⟩2.0.CO;2.

Houtekamer, P. L., and F. Zhang, 2016: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, **144 (12)**, 4489–4532, https://doi.org/10.1175/MWR-D-15-0440.1.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

Huang, B., C. Liu, G. Ren, H. M. Zhang, and L. Zhang, 2019: The Role of Buoy and Argo Observations in Two SST Analyses in the Global and Tropical Pacific Oceans. *J. Climate*, **32 (9)**, 2517–2535, https://doi.org/10.1175/JCLI-D-18-0368.1.

Huang, B., and Coauthors, 2017: Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons. *J. Climate*, **30 (20)**, 8179–8205, https://doi.org/10.1175/JCLI-D-16-0836.1.

Hurrell, J. W., J. J. Hack, D. Shea, J. M. Caron, and J. Rosinski, 2008: A New Sea Surface Temperature and Sea Ice Boundary Dataset for the Community Atmosphere Model. *J. Climate*, **21 (19)**, 5145–5153, https://doi.org/10.1175/2008JCLI2292.1.

Jiang, F., R. Seager, and M. A. Cane, 2024: A climate change signal in the tropical Pacific emerges from decadal variability. *Nature Communications*, **15 (1)**, 8291, https://doi.org/10.1038/s41467-024-52731-6.

Kadow, C., D. M. Hall, and U. Ulbrich, 2020: Artificial intelligence reconstructs missing climate information. *Nature Geoscience*, **13 (6)**, 408–413, https://doi.org/10.1038/s41561-020-0582-5.

Kalman, R. E., 1960: A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, **82 (1)**, 35–45, https://doi.org/10.1115/1.3662552.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77 (3)**, https://doi.org/10.1175/1520-0477(1996)077⟨0437:TNYRP⟩2.0.CO;2.

J. Kang, T. A. Shaw, S. M. Kang, I. R. Simpson, and Y. Yu, 2024: Revisiting the reanalysis-model discrepancy in Southern Hemisphere winter storm track trends. *npj Climate and Atmospheric Science*, **7 (1)**, 252, https://doi.org/10.1038/s41612-024-00801-3.

S. Kang, P. Ceppi, Y. Yu, and I.-S. Kang, 2023a: Recent global climate feedback controlled by Southern Ocean cooling. *Nature Geoscience*, https://doi.org/10.1038/s41561-023-01256-6.

S. Kang, Y. Shin, H. Kim, S.-P. Xie, and S. Hu, 2023b: Disentangling the mechanisms of equatorial Pacific climate change. *Science Advances*, **9 (19)**, https://doi.org/10.1126/sciadv.adf5059.

S. Kang, Y. Yu, C. Deser, X. Zhang, I.-S. Kang, S.-S. Lee, K. B. Rodgers, and P. Ceppi, 2023c: Global impacts of recent Southern Ocean cooling. *Proc. Natl. Acad. Sci. (USA)*, **120 (30)**, https://doi.org/10.1073/pnas.2300881120.

Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856–1991. *J. Geophys. Res.: Oceans*, **103 (C9)**, 18 567–18 589, https://doi.org/10.1029/97JC01736.

Kaplan, A., Y. Kushnir, and M. A. Cane, 2000: Reduced Space Optimal Interpolation of Historical Marine Sea Level Pressure: 1854–1992*. *J. Climate*, **13 (16)**, 2987–3002, https://doi.org/10.1175/1520-0442(2000)013⟨2987:RSOIOH⟩2.0.CO;2.

Karl, T. R., and Coauthors, 2015: Possible artifacts of data biases in the recent global surface warming hiatus. *Science*, **348 (6242)**, 1469–1472, https://doi.org/10.1126/SCIENCE.AAA5632/SUPPL{_}FILE/PAPV2.PDF.

Karnauskas, K. B., 2013: Can we distinguish canonical El Niño from Modoki? *Geophys. Res. Lett.*, **40 (19)**, 5246–5251, https://doi.org/10.1002/grl.51007.

Karnauskas, K. B., R. Seager, A. Kaplan, Y. Kushnir, and M. A. Cane, 2009: Observed Strengthening of the Zonal Sea Surface Temperature Gradient across the Equatorial Pacific Ocean*. *J. Climate*, **22 (16)**, 4316–4321, https://doi.org/10.1175/2009JCLI2936.1.

Kennedy, J. J., 2014: A review of uncertainty in in situ measurements and data sets of sea surface temperature. *Rev. Geophys.*, **52 (1)**, 1–32, https://doi.org/10.1002/2013RG000434.

Kennedy, J. J., N. A. Rayner, C. P. Atkinson, and R. E. Killick, 2019: An Ensemble Data Set of Sea Surface Temperature Change From 1850: The Met Office Hadley Centre HadSST.4.0.0.0 Data Set. *J. Geophys. Res.: Atmospheres*, **124 (14)**, 7719–7763, https://doi.org/10.1029/2018JD029867.

Kent, E. C., and J. J. Kennedy, 2021: Historical Estimates of Surface Marine Temperatures. *Annual Review of Marine Science*, **13 (1)**, 283–311, https://doi.org/10.1146/annurev-marine-042120-111807.

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

Kido, S., I. Richter, T. Tozuka, and P. Chang, 2023: Understanding the interplay between ENSO and related tropical SST variability using linear inverse models. *Climate Dynamics*, **61 (3-4)**, 1029–1048, https://doi.org/10.1007/s00382-022-06484-x.

King, J., K. J. Anchukaitis, K. Allen, T. Vance, and A. Hessl, 2023: Trends and variability in the Southern Annular Mode over the Common Era. *Nature Communications*, **14 (1)**, 2324, https://doi.org/10.1038/s41467-023-37643-1.

Kobayashi, S., and Coauthors, 2015: The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, **93 (1)**, https://doi.org/10.2151/jmsj.2015-001.

Kohyama, T., and D. L. Hartmann, 2016: Antarctic Sea Ice Response to Weather and Climate Modes of Variability*. *J. Climate*, **29 (2)**, 721–741, https://doi.org/10.1175/JCLI-D-15-0301.1.

Kohyama, T., D. L. Hartmann, and D. S. Battisti, 2017: La Niña–like Mean-State Response to Global Warming and Potential Oceanic Roles. *J. Climate*, **30 (11)**, 4207–4225, https://doi.org/10.1175/JCLI-D-16-0441.1.

Kolbe, W. M., R. T. Tonboe, and J. Stroeve, 2024: Mapping of sea ice concentration using the NASA NIMBUS 5 Electrically Scanning Microwave Radiometer data from 1972–1977. *Earth System Science Data*, **16 (3)**, 1247–1264, https://doi.org/10.5194/essd-16-1247-2024.

Kosaka, Y., and Coauthors, 2024: The JRA-3Q Reanalysis. *Journal of the Meteorological Society of Japan*, **102 (1)**, https://doi.org/10.2151/jmsj.2024-004.

Kuhlbrodt, T., and Coauthors, 2018: The Low-Resolution Version of HadGEM3 GC3.1: Development and Evaluation for Global Climate. *Journal of Advances in Modeling Earth Systems*, **10 (11)**, 2865–2888, https://doi.org/10.1029/2018MS001370.

Kuo, Y., H. Kim, and F. Lehner, 2023: Anthropogenic Aerosols Contribute to the Recent Decline in Precipitation Over the U.S. Southwest. *Geophys. Res. Lett.*, **50 (23)**, https://doi.org/10.1029/2023GL105389.

Laloyaux, P., and Coauthors, 2018: CERA-20C: A Coupled Reanalysis of the Twentieth Century. *Journal of Advances in Modeling Earth Systems*, **10 (5)**, 1172–1195, https://doi.org/10.1029/2018MS001273.

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

Lee, S., M. L'Heureux, A. T. Wittenberg, R. Seager, P. A. O'Gorman, and N. C. Johnson, 2022: On the future zonal contrasts of equatorial Pacific climate: Perspectives from Observations, Simulations, and Theories. *npj Climate and Atmospheric Science*, **5 (1)**, 82, https://doi.org/10.1038/s41612-022-00301-2.

Lehner, F., C. Deser, I. R. Simpson, and L. Terray, 2018: Attributing the U.S. Southwest's Recent Shift Into Drier Conditions. *Geophys. Res. Lett.*, **45 (12)**, 6251–6261, https://doi.org/10.1029/2018GL078312.

Leith, C. E., 1973: The Standard Error of Time-Average Estimates of Climatic Means. *J. Appl. Meteor.*, **12 (6)**, 1066–1069, https://doi.org/10.1175/1520-0450(1973)012⟨1066:TSEOTA⟩2.0.CO;2.

Lewis, N., and T. Mauritsen, 2021: Negligible Unforced Historical Pattern Effect on Climate Feedback Strength Found in HadISST-Based AMIP Simulations. *J. Climate*, **34 (1)**, 39–55, https://doi.org/10.1175/JCLI-D-19-0941.1.

Lin, Y., and Coauthors, 2020: Community Integrated Earth System Model (CIESM): Description and Evaluation. *Journal of Advances in Modeling Earth Systems*, **12 (8)**, e2019MS002 036, https://doi.org/10.1029/2019MS002036.

Lou, J., M. Newman, and A. Hoell, 2023: Multi-decadal variation of ENSO forecast skill since the late 1800s. *npj Climate and Atmospheric Science*, **6 (1)**, 89, https://doi.org/10.1038/s41612-023-00417-z.

Lou, J., T. J. O'Kane, and N. J. Holbrook, 2020: A Linear Inverse Model of Tropical and South Pacific Seasonal Predictability. *J. Climate*, **33 (11)**, 4537–4554, https://doi.org/10.1175/JCLI-D-19-0548.1.

L'Heureux, M. L., S. Lee, and B. Lyon, 2013: Recent multidecadal strengthening of the Walker circulation across the tropical Pacific. *Nature Climate Change*, **3 (6)**, 571–576, https://doi.org/10.1038/nclimate1840.

Marshall, G. J., 2003: Trends in the Southern Annular Mode from Observations and Reanalyses. *J. Climate*, **16 (24)**, 4134–4143, https://doi.org/10.1175/1520-0442(2003)016⟨4134:TITSAM⟩2.0.CO;2.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

Marvel, K., R. Pincus, G. A. Schmidt, and R. L. Miller, 2018: Internal Variability and Disequilibrium Confound Estimates of Climate Sensitivity From Observations. *Geophys. Res. Lett.*, **45 (3)**, 1595–1601, https://doi.org/10.1002/2017GL076468.

Mauritsen, T., and Coauthors, 2019: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO¡sub¿2¡/sub¿. *Journal of Advances in Modeling Earth Systems*, **11 (4)**, 998–1038, https://doi.org/10.1029/2018MS001400.

McGregor, S., A. Timmermann, M. F. Stuecker, M. H. England, M. Merrifield, F.-F. Jin, and Y. Chikamoto, 2014: Recent Walker circulation strengthening and Pacific cooling amplified by Atlantic warming. *Nature Climate Change*, **4 (10)**, 888–892, https://doi.org/10.1038/nclimate2330.

Meier, W. N., F. Fetterer, A. K. Windnagel, and J. S. Stewart, 2021a: Near-Real-Time NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration (G10016, Version 2). Tech. rep., NSIDC: National Snow and Ice Data Center, Boulder, CO. https://doi.org/10.7265/tgam-yv28.

Meier, W. N., F. Fetterer, A. K. Windnagel, and J. S. Stewart, 2021b: NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 4. Tech. rep., NSIDC: National Snow and Ice Data Center, Boulder, Colorado USA. https://doi.org/https://doi.org/10.7265/efmz-2t65.

Meng, Z., and G. J. Hakim, 2024: Reconstructing the Tropical Pacific Upper Ocean Using Online Data Assimilation With a Deep Learning Model. *Journal of Advances in Modeling Earth Systems*, **16 (11)**, https://doi.org/10.1029/2024MS004422.

Modak, A., and T. Mauritsen, 2023: Better-constrained climate sensitivity when accounting for dataset dependency on pattern effect estimates. *Atmospheric Chemistry and Physics*, **23 (13)**, 7535–7549, https://doi.org/10.5194/acp-23-7535-2023.

Morice, C. P., and Coauthors, 2021: An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set. *J. Geophys. Res.: Atmospheres*, **126 (3)**, https://doi.org/10.1029/2019JD032361.

Nash, J. E., and J. V. Sutcliffe, 1970: River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, **10 (3)**, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6.

Newman, M., 2007: Interannual to Decadal Predictability of Tropical and North Pacific Sea Surface Temperatures. *J. Climate*, **20 (11)**, 2333–2356, https://doi.org/10.1175/JCLI4165.1.

Newman, M., 2013: An Empirical Benchmark for Decadal Forecasts of Global Surface Temperature Anomalies. *J. Climate*, **26 (14)**, 5260–5269, https://doi.org/10.1175/JCLI-D-12-00590.1.

Newman, M., S.-I. Shin, and M. A. Alexander, 2011: Natural variation in ENSO flavors. *Geophys. Res. Lett.*, **38 (14)**, n/a–n/a, https://doi.org/10.1029/2011GL047658.

Newman, M., and Coauthors, 2016: The Pacific Decadal Oscillation, Revisited. *J. Climate*, **29 (12)**, 4399–4427, https://doi.org/10.1175/JCLI-D-15-0508.1.

NOAA, 2024: The Global Ensemble Forecast System (version 13) Replay dataset. URL https://psl.noaa.gov/data/ufs_replay/.

Notz, D., and SIMIP Community, 2020: Arctic Sea Ice in CMIP6. *Geophys. Res. Lett.*, **47 (10)**, e2019GL086 749, https://doi.org/10.1029/2019GL086749.

Olonscheck, D., M. Rugenstein, and J. Marotzke, 2020: Broad Consistency Between Observed and Simulated Trends in Sea Surface Temperature Patterns. *Geophys. Res. Lett.*, **47 (10)**, e2019GL086 773, https://doi.org/10.1029/2019GL086773.

Orbe, C., L. D. Oman, S. E. Strahan, D. W. Waugh, S. Pawson, L. L. Takacs, and A. M. Molod, 2017: Large-Scale Atmospheric Transport in GEOS-Replay Simulations. *Journal of Advances in Modeling Earth Systems*, **9 (7)**, 2545–2560, https://doi.org/10.1002/2017MS001053.

OrtizBeviá, M. J., 1997: Estimation of the cyclostationary dependence in geophysical data fields. *J. Geophys. Res.: Atmospheres*, **102 (D12)**, 13 473–13 486, https://doi.org/10.1029/97JD00243.

Osborn, T. J., P. D. Jones, D. H. Lister, C. P. Morice, I. R. Simpson, J. P. Winn, E. Hogan, and I. C. Harris, 2021: Land Surface Air Temperature Variations Across the Globe Updated to 2019: The CRUTEM5 Data Set. *J. Geophys. Res.: Atmospheres*, **126 (2)**, https://doi.org/10.1029/2019JD032352.

Osman, M. B., J. E. Tierney, J. Zhu, R. Tardif, G. J. Hakim, J. King, and C. J. Poulsen, 2021: Globally resolved surface temperatures since the Last Glacial Maximum. *Nature*, **599 (7884)**, 239–244, https://doi.org/10.1038/s41586-021-03984-4.

O'Connor, G. K., E. J. Steig, and G. J. Hakim, 2021: Strengthening Southern Hemisphere Westerlies and Amundsen Sea Low Deepening Over the 20th Century Revealed by Proxy-Data Assimilation. *Geophys. Res. Lett.*, **48 (24)**, e2021GL095 999, https://doi.org/10.1029/2021GL095999.

Park, S., J. Shin, S. Kim, E. Oh, and Y. Kim, 2019: Global Climate Simulated by the Seoul National University Atmosphere Model Version 0 with a Unified Convection Scheme (SAM0-UNICON). *J. Climate*, **32 (10)**, 2917–2949, https://doi.org/10.1175/JCLI-D-18-0796.1.

Parsons, L. A., D. E. Amrhein, S. C. Sanchez, R. Tardif, M. K. Brennan, and G. J. Hakim, 2021: Do Multi-Model Ensembles Improve Reconstruction Skill in Paleoclimate Data Assimilation? *Earth and Space Science*, **8 (4)**, e2020EA001 467, https://doi.org/10.1029/2020EA001467.

Penland, C., 1989: Random Forcing and Forecasting Using Principal Oscillation Pattern Analysis. *Monthly Weather Review*, **117 (10)**, 2165–2185, https://doi.org/10.1175/1520-0493(1989) 117⟨2165:RFAFUP⟩2.0.CO;2.

Penland, C., 1996: A stochastic model of IndoPacific sea surface temperature anomalies. *Physica D: Nonlinear Phenomena*, **98 (2-4)**, 534–558, https://doi.org/10.1016/0167-2789(96)00124-8.

Penland, C., and L. Matrosova, 1994: A Balance Condition for Stochastic Numerical Models with Application to the El Niño-Southern Oscillation. *J. Climate*, **7 (9)**, 1352–1372, https://doi.org/ https://doi.org/10.1175/1520-0442(1994)007{\%}3C1352:ABCFSN{\%}3E2.0.CO;2.

Penland, C., and P. D. Sardeshmukh, 1995: The Optimal Growth of Tropical Sea Surface Temperature Anomalies. *J. Climate*, **8 (8)**, 1999–2024, https://doi.org/10.1175/1520-0442(1995) 008⟨1999:TOGOTS⟩2.0.CO;2.

Perkins, W. A., and G. J. Hakim, 2021: Coupled Atmosphere–Ocean Reconstruction of the Last Millennium Using Online Data Assimilation. *Paleoceanography and Paleoclimatology*, **36 (5)**, https://doi.org/10.1029/2020PA003959.

Po-Chedley, S., B. D. Santer, S. Fueglistaler, M. D. Zelinka, P. J. Cameron-Smith, J. F. Painter, and Q. Fu, 2021: Natural variability contributes to model–satellite differences in tropical tropospheric warming. *Proc. Natl. Acad. Sci. (USA)*, **118 (13)**, https://doi.org/10.1073/pnas.2020962118.

Polvani, L. M., D. W. Waugh, G. J. P. Correa, and S.-W. Son, 2011: Stratospheric Ozone Depletion: The Main Driver of Twentieth-Century Atmospheric Circulation Changes in the Southern Hemisphere. *J. Climate*, **24 (3)**, 795–812, https://doi.org/10.1175/2010JCLI3772.1.

Polvani, L. M., and Coauthors, 2021: Interannual SAM Modulation of Antarctic Sea Ice Extent Does Not Account for Its Long-Term Trends, Pointing to a Limited Role for Ozone Depletion. *Geophys. Res. Lett.*, **48 (21)**, https://doi.org/10.1029/2021GL094871.

Qin, Y., X. Zheng, S. A. Klein, M. D. Zelinka, P. Ma, J. Golaz, and S. Xie, 2024: Causes of Reduced Climate Sensitivity in E3SM From Version 1 to Version 2. *Journal of Advances in Modeling Earth Systems*, **16 (1)**, https://doi.org/10.1029/2023MS003875.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.: Atmospheres*, **108 (D14)**, 4407, https://doi.org/10.1029/2002JD002670.

Roach, L. A., and W. N. Meier, 2024: Sea ice in 2023. *Nature Reviews Earth & Environment*, **5 (4)**, 235–237, https://doi.org/10.1038/s43017-024-00542-0.

Roach, L. A., and Coauthors, 2020: Antarctic Sea Ice Area in CMIP6. *Geophys. Res. Lett.*, **47 (9)**, https://doi.org/10.1029/2019GL086729.

Rohde, R., and Coauthors, 2013: An Overview Berkeley Earth Temperature Averaging Process. *Geoinformatics & Geostatistics: An Overview*, **(June)**, https://doi.org/10.4172/gigs.1000103.

Rugenstein, M., S. Dhame, D. Olonscheck, R. J. Wills, M. Watanabe, and R. Seager, 2023: Connecting the SST Pattern Problem and the Hot Model Problem. *Geophys. Res. Lett.*, **50 (22)**, https://doi.org/10.1029/2023GL105488.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91 (8)**, 1015–1058, https://doi.org/10.1175/2010BAMS3001.1.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

Salvi, P., J. M. Gregory, and P. Ceppi, 2023: Time-Evolving Radiative Feedbacks in the Historical Period. *J. Geophys. Res.: Atmospheres*, **128 (20)**, https://doi.org/10.1029/2023JD038984.

Samakinwa, E., V. Valler, R. Hand, R. Neukom, J. J. Gómez-Navarro, J. Kennedy, N. A. Rayner, and S. Brönnimann, 2021: An ensemble reconstruction of global monthly sea surface temperature and sea ice concentration 1000–1849. *Scientific Data*, **8 (1)**, 261, https://doi.org/10.1038/s41597-021-01043-1.

Sanchez, S. C., G. J. Hakim, and C. P. Saenger, 2021: Climate Model Teleconnection Patterns Govern the Niño-3.4 Response to Early Nineteenth-Century Volcanism in Coral-Based Data Assimilation Reconstructions. *J. Climate*, **34 (5)**, 1863–1880, https://doi.org/10.1175/JCLI-D-20-0549.1.

Sanchez, S. C., N. Westphal, G. H. Haug, H. Cheng, R. L. Edwards, T. Schneider, K. M. Cobb, and C. D. Charles, 2020: A Continuous Record of Central Tropical Pacific Climate Since the Midnineteenth Century Reconstructed From Fanning and Palmyra Island Corals: A Case Study in Coral Data Reanalysis. *Paleoceanography and Paleoclimatology*, **35 (8)**, 1877–1878, https://doi.org/10.1029/2020PA003848.

Schmidt, G. A., and Coauthors, 2023: Anomalous Meltwater From Ice Sheets and Ice Shelves Is a Historical Forcing. *Geophys. Res. Lett.*, **50 (24)**, https://doi.org/10.1029/2023GL106530.

Schneider, D. P., C. Deser, J. Fasullo, and K. E. Trenberth, 2013: Climate Data Guide Spurs Discovery and Understanding. *Eos, Trans. Amer. Geophys. Union*, **94 (13)**, 121–122, https://doi.org/10.1002/2013EO130001.

Schneider, D. P., and R. L. Fogt, 2018: Artifacts in Century-Length Atmospheric and Coupled Reanalyses Over Antarctica Due To Historical Data Availability. *Geophys. Res. Lett.*, **45 (2)**, 964–973, https://doi.org/10.1002/2017GL076226.

Schneider, T., 2001: Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values. *J. Climate*, **14 (5)**, 853–871, https://doi.org/10.1175/1520-0442(2001)014⟨0853:AOICDE⟩2.0.CO;2.

Screen, J. A., 2011: Sudden increase in Antarctic sea ice: Fact or artifact? *Geophys. Res. Lett.*, **38 (13)**, n/a–n/a, https://doi.org/10.1029/2011GL047553.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-25-0021.1.

Seager, R., M. Cane, N. Henderson, D.-E. Lee, R. Abernathey, and H. Zhang, 2019: Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. *Nature Climate Change*, **9 (7)**, 517–522, https://doi.org/10.1038/s41558-019-0505-x.

Seager, R., N. Henderson, and M. Cane, 2022: Persistent Discrepancies between Observed and Modeled Trends in the Tropical Pacific Ocean. *J. Climate*, **35 (14)**, 4571–4584, https://doi.org/10.1175/JCLI-D-21-0648.1.

Seager, R., M. Ting, P. Alexander, H. Liu, J. Nakamura, C. Li, and M. Newman, 2023: Ocean-forcing of cool season precipitation drives ongoing and future decadal drought in southwestern North America. *npj Climate and Atmospheric Science*, **6 (1)**, 141, https://doi.org/10.1038/s41612-023-00461-9.

Seland, Ø., and Coauthors, 2020: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, **13 (12)**, 6165–6200, https://doi.org/10.5194/gmd-13-6165-2020.

Sellar, A. A., and Coauthors, 2019: UKESM1: Description and Evaluation of the U.K. Earth System Model. *Journal of Advances in Modeling Earth Systems*, **11 (12)**, 4513–4558, https://doi.org/10.1029/2019MS001739.

Seviour, W. J. M., A. Gnanadesikan, and D. W. Waugh, 2016: The Transient Response of the Southern Ocean to Stratospheric Ozone Depletion. *J. Climate*, **29 (20)**, 7383–7396, https://doi.org/10.1175/JCLI-D-16-0198.1.

Sherwood, S. C., and Coauthors, 2020: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence. *Rev. Geophys.*, **58 (4)**, https://doi.org/10.1029/2019RG000678.

Shin, S.-I., P. D. Sardeshmukh, M. Newman, C. Penland, and M. A. Alexander, 2021: Impact of Annual Cycle on ENSO Variability and Predictability. *J. Climate*, **34 (1)**, 171–193, https://doi.org/10.1175/JCLI-D-20-0291.1.

Siler, N., C. Proistosescu, and S. Po-Chedley, 2019: Natural Variability Has Slowed the Decline in Western U.S. Snowpack Since the 1980s. *Geophys. Res. Lett.*, **46 (1)**, 346–355, https://doi.org/10.1029/2018GL081080.

Simpson, I. R., and Coauthors, 2025: Confronting Earth System Model trends with observations. *Science Advances*, **11 (11)**, 8035, https://doi.org/10.1126/sciadv.adt8035.

Sippel, S., and Coauthors, 2024: Early-twentieth-century cold bias in ocean surface temperature observations. *Nature*, **635 (8039)**, 618–624, https://doi.org/10.1038/s41586-024-08230-1.

Slivinski, L. C., and Coauthors, 2019: Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quart. J. Roy. Meteor. Soc.*, **145 (724)**, 2876–2908, https://doi.org/10.1002/qj.3598.

Slivinski, L. C., and Coauthors, 2021: An Evaluation of the Performance of the Twentieth Century Reanalysis Version 3. *J. Climate*, **34 (4)**, 1417–1438, https://doi.org/10.1175/JCLI-D-20-0505.1.

Smerdon, J. E., E. R. Cook, and N. J. Steiger, 2023: The Historical Development of Large-Scale Paleoclimate Field Reconstructions Over the Common Era. *Rev. Geophys.*, **61 (4)**, https://doi.org/10.1029/2022RG000782.

Soci, C., and Coauthors, 2024: The ERA5 global reanalysis from 1940 to 2022. *Quart. J. Roy. Meteor. Soc.*, **150 (764)**, 4014–4048, https://doi.org/10.1002/qj.4803.

Solomon, A., and M. Newman, 2012: Reconciling disparate twentieth-century Indo-Pacific ocean temperature trends in the instrumental record. *Nature Climate Change*, **2 (9)**, 691–699, https://doi.org/10.1038/nclimate1591.

Steiger, N. J., G. J. Hakim, E. J. Steig, D. S. Battisti, and G. H. Roe, 2014: Assimilation of time-averaged pseudoproxies for climate reconstruction. *J. Climate*, **27 (1)**, https://doi.org/10.1175/JCLI-D-12-00693.1.

Steiger, N. J., J. E. Smerdon, E. R. Cook, and B. I. Cook, 2018: A reconstruction of global hydroclimate and dynamical variables over the Common Era. *Scientific Data*, **5 (1)**, 180 086, https://doi.org/10.1038/sdata.2018.86.

Stuecker, M. F., C. M. Bitz, and K. C. Armour, 2017: Conditions leading to the unprecedented low Antarctic sea ice extent during the 2016 austral spring season. *Geophys. Res. Lett.*, **44 (17)**, 9008–9019, https://doi.org/10.1002/2017GL074691.

Accepted for publication in *Journal of Climate.* DOI 10.1175/JCLI-D-25-0021.1.

Suryawanshi, K., B. Jena, C. C. Bajish, and N. Anilkumar, 2023: Recent Decline in Antarctic Sea Ice Cover From 2016 to 2022: Insights From Satellite Observations, Argo Floats, and Model Reanalysis. *Tellus, Series A: Dynamic Meteorology and Oceanography*, **75 (1)**, 193–212, https://doi.org/10.16993/TELLUSA.3222.

Swart, N. C., J. C. Fyfe, N. Gillett, and G. J. Marshall, 2015: Comparing Trends in the Southern Annular Mode and Surface Westerly Jet. *J. Climate*, **28 (22)**, 8840–8859, https://doi.org/10.1175/JCLI-D-15-0334.1.

Swart, N. C., and Coauthors, 2023: The Southern Ocean Freshwater Input from Antarctica (SOFIA) Initiative: scientific objectives and experimental design. *Geoscientific Model Development*, **16 (24)**, 7289–7309, https://doi.org/10.5194/gmd-16-7289-2023.

Thomas, J. L., D. W. Waugh, and A. Gnanadesikan, 2015: Southern Hemisphere extratropical circulation: Recent trends and natural variability. *Geophys. Res. Lett.*, **42 (13)**, 5508–5515, https://doi.org/10.1002/2015GL064521.

Thompson, D. W. J., and S. Solomon, 2002: Interpretation of Recent Southern Hemisphere Climate Change. *Science*, **296 (5569)**, 895–899, https://doi.org/10.1126/science.1069270.

Thompson, D. W. J., S. Solomon, P. J. Kushner, M. H. England, K. M. Grise, and D. J. Karoly, 2011: Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change. *Nature Geoscience*, **4 (11)**, 741–749, https://doi.org/10.1038/ngeo1296.

Tierney, J. E., E. J. Judd, M. B. Osman, J. M. King, O. J. Truax, N. J. Steiger, D. E. Amrhein, and K. J. Anchukaitis, 2025a: Advances in Paleoclimate Data Assimilation. *Annu. Rev. Earth Planet. Sci.*, **28**, 55, https://doi.org/10.1146/annurev-earth-032320-064209.

Tierney, J. E., J. King, M. B. Osman, J. T. Abell, N. J. Burls, E. Erfani, V. T. Cooper, and R. Feng, 2025b: Pliocene Warmth and Patterns of Climate Change Inferred From Paleoclimate Data Assimilation. *AGU Advances*, **6 (1)**, https://doi.org/10.1029/2024AV001356.

Tierney, J. E., J. Zhu, J. King, S. B. Malevich, G. J. Hakim, and C. J. Poulsen, 2020: Glacial cooling and climate sensitivity revisited. *Nature*, **584 (7822)**, 569–573, https://doi.org/10.1038/s41586-020-2617-x.

Titchner, H. A., and N. A. Rayner, 2014: The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations. *J. Geophys. Res.: Atmospheres*, **119 (6)**, 2864–2889, https://doi.org/10.1002/2013JD020316.

Tokinaga, H., S.-P. Xie, A. Timmermann, S. McGregor, T. Ogata, H. Kubota, and Y. M. Okumura, 2012: Regional Patterns of Tropical Indo-Pacific Climate Change: Evidence of the Walker Circulation Weakening. *J. Climate*, **25 (5)**, 1689–1710, https://doi.org/10.1175/JCLI-D-11-00263.1.

Trenberth, K. E., and D. J. Shea, 2006: Atlantic hurricanes and natural variability in 2005. *Geophys. Res. Lett.*, **33 (12)**, 12 704, https://doi.org/10.1029/2006GL026894.

Tseng, K.-C., N. C. Johnson, E. D. Maloney, E. A. Barnes, and S. B. Kapnick, 2021: Mapping Large-Scale Climate Variability to Hydrological Extremes: An Application of the Linear Inverse Model to Subseasonal Prediction. *J. Climate*, **34 (11)**, 4207–4225, https://doi.org/10.1175/JCLI-D-20-0502.1.

Turner, J., and Coauthors, 2022: Record Low Antarctic Sea Ice Cover in February 2022. *Geophys. Res. Lett.*, **49 (12)**, https://doi.org/10.1029/2022GL098904.

Vaccaro, A., J. Emile-Geay, D. Guillot, R. Verna, C. Morice, J. Kennedy, and B. Rajaratnam, 2021: Climate Field Completion via Markov Random Fields: Application to the HadCRUT4.6 Temperature Dataset. *J. Climate*, **34 (10)**, 4169–4188, https://doi.org/10.1175/JCLI-D-19-0814.1.

Valler, V., and Coauthors, 2024: ModE-RA: a global monthly paleo-reanalysis of the modern era 1421 to 2008. *Scientific Data*, **11 (1)**, 36, https://doi.org/10.1038/s41597-023-02733-8.

Vecchi, G. A., and B. J. Soden, 2007: Global Warming and the Weakening of the Tropical Circulation. *J. Climate*, **20 (17)**, 4316–4340, https://doi.org/10.1175/JCLI4258.1.

Vecchi, G. A., B. J. Soden, A. T. Wittenberg, I. M. Held, A. Leetmaa, and M. J. Harrison, 2006: Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature*, **441 (7089)**, 73–76, https://doi.org/10.1038/nature04744.

Vimont, D. J., 2012: Analysis of the Atlantic Meridional Mode Using Linear Inverse Modeling: Seasonality and Regional Influences. *J. Climate*, **25 (4)**, 1194–1212, https://doi.org/10.1175/JCLI-D-11-00012.1.

Vimont, D. J., M. A. Alexander, and M. Newman, 2014: Optimal growth of Central and East Pacific ENSO events. *Geophys. Res. Lett.*, **41 (11)**, 4027–4034, https://doi.org/10.1002/2014GL059997.

Vimont, D. J., M. Newman, D. S. Battisti, and S.-I. Shin, 2022: The Role of Seasonality and the ENSO Mode in Central and East Pacific ENSO Growth and Evolution. *J. Climate*, **35 (11)**, 3195–3209, https://doi.org/10.1175/JCLI-D-21-0599.1.

Walsh, J. E., W. L. Chapman, F. Fetterer, and S. Stewart, 2019: Gridded Monthly Sea Ice Extent and Concentration, 1850 Onward, Version 2. Tech. rep., NSIDC: National Snow and Ice Data Center., Boulder, Colorado USA. https://doi.org/10.7265/jj4s-tq79.

Watanabe, M., J.-L. Dufresne, Y. Kosaka, T. Mauritsen, and H. Tatebe, 2021: Enhanced warming constrained by past trends in equatorial Pacific sea surface temperature gradient. *Nature Climate Change*, **11 (1)**, 33–37, https://doi.org/10.1038/s41558-020-00933-3.

Watanabe, M., T. Iwakiri, Y. Dong, and S. M. Kang, 2023: Two Competing Drivers of the Recent Walker Circulation Trend. *Geophys. Res. Lett.*, **50 (23)**, e2023GL105 332, https://doi.org/10.1029/2023GL105332.

Watanabe, M., S. M. Kang, M. Collins, Y.-T. Hwang, S. McGregor, and M. F. Stuecker, 2024: Possible shift in controls of the tropical Pacific surface warming pattern. *Nature*, **630 (8016)**, 315–324, https://doi.org/10.1038/s41586-024-07452-7.

Webb, M. J., and Coauthors, 2017: The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6. *Geoscientific Model Development*, **10 (1)**, 359–384, https://doi.org/10.5194/gmd-10-359-2017.

Wills, R. C., D. S. Battisti, K. C. Armour, T. Schneider, and C. Deser, 2020: Pattern Recognition Methods to Separate Forced Responses from Internal Variability in Climate Model Ensembles and Observations. *J. Climate*, **33 (20)**, 8693–8719, https://doi.org/10.1175/JCLI-D-19-0855.1.

Wills, R. C. J., Y. Dong, C. Proistosecu, K. C. Armour, and D. S. Battisti, 2022: Systematic Climate Model Biases in the Large-Scale Patterns of Recent Sea-Surface Temperature and Sea-Level Pressure Change. *Geophys. Res. Lett.*, **49 (17)**, e2022GL100 011, https://doi.org/10.1029/2022GL100011.

Wunsch, C., 1999: The Interpretation of Short Climate Records, with Comments on the North Atlantic and Southern Oscillations. *Bull. Amer. Meteor. Soc.*, **80 (2)**, 245–255, https://doi.org/10.1175/1520-0477(1999)080⟨0245:TIOSCR⟩2.0.CO;2.

Xie, S. P., C. Deser, G. A. Vecchi, J. Ma, H. Teng, and A. T. Wittenberg, 2010: Global Warming Pattern Formation: Sea Surface Temperature and Rainfall. *J. Climate*, **23 (4)**, 966–986, https://doi.org/10.1175/2009JCLI3329.1.

Zanna, L., 2012: Forecast Skill and Predictability of Observed Atlantic Sea Surface Temperatures. *J. Climate*, **25 (14)**, 5047–5056, https://doi.org/10.1175/JCLI-D-11-00539.1.

Zhang, C., and S. Li, 2023: Causes of the record-low Antarctic sea-ice in austral summer 2022. *Atmospheric and Oceanic Science Letters*, **16 (6)**, 100 353, https://doi.org/10.1016/j.aosl.2023.100353.

Zhang, L., T. L. Delworth, W. Cooke, and X. Yang, 2019: Natural variability of Southern Ocean convection as a driver of observed climate trends. *Nature Climate Change*, **9 (1)**, 59–65, https://doi.org/10.1038/s41558-018-0350-3.

Zhang, X., C. Deser, and L. Sun, 2021: Is There a Tropical Response to Recent Observed Southern Ocean Cooling? *Geophys. Res. Lett.*, **48 (5)**, https://doi.org/10.1029/2020GL091235.

Zhang, Y., and S. Fueglistaler, 2020: How Tropical Convection Couples High Moist Static Energy Over Land and Ocean. *Geophys. Res. Lett.*, **47 (2)**, https://doi.org/10.1029/2019GL086387.

Zhao, Y., D. Sun, E. Di Lorenzo, G. Liu, and S. Wu, 2024: Separate the Role of Southern and Northern Extra-Tropical Pacific in Tropical Pacific Climate Variability. *Geophys. Res. Lett.*, **51 (16)**, https://doi.org/10.1029/2024GL109466.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2016: Impact of decadal cloud variations on the Earth's energy budget. *Nature Geoscience*, **9 (12)**, 871–874, https://doi.org/10.1038/ngeo2828.

Zuo, H., M. A. Balmaseda, S. Tietsche, K. Mogensen, and M. Mayer, 2019: The ECMWF operational ensemble reanalysis-analysis system for ocean and sea ice: A description of the system and assessment. *Ocean Science*, **15 (3)**, https://doi.org/10.5194/os-15-779-2019.