

Forced and Internal Components of Winter Air Temperature Trends over North America during the past 50 Years: Mechanisms and Implications*

CLARA DESER

Climate and Global Dynamics Division, National Center for Atmospheric Research,⁺ Boulder, Colorado

LAURENT TERRAY

Sciences de L'Univers au CERFACS, URA-1875, CERFACS/CNRS, Toulouse, France

ADAM S. PHILLIPS

Climate and Global Dynamics Division, National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 24 April 2015, in final form 29 December 2015)

ABSTRACT

This study elucidates the physical mechanisms underlying internal and forced components of winter surface air temperature (SAT) trends over North America during the past 50 years (1963–2012) using a combined observational and modeling framework. The modeling framework consists of 30 simulations with the Community Earth System Model (CESM) at 1° latitude–longitude resolution, each of which is subject to an identical scenario of historical radiative forcing but starts from a slightly different atmospheric state. Hence, any spread within the ensemble results from unpredictable internal variability superimposed upon the forced climate change signal. Constructed atmospheric circulation analogs are used to estimate the dynamical contribution to forced and internal components of SAT trends: thermodynamic contributions are obtained as a residual. Internal circulation trends are estimated to account for approximately one-third of the observed wintertime warming trend over North America and more than half locally over parts of Canada and the United States. Removing the effects of internal atmospheric circulation variability narrows the spread of SAT trends within the CESM ensemble and brings the observed trends closer to the model's radiatively forced response. In addition, removing internal dynamics approximately doubles the signal-to-noise ratio of the simulated SAT trends and substantially advances the “time of emergence” of the forced component of SAT anomalies. The methodological framework proposed here provides a general template for improving physical understanding and interpretation of observed and simulated climate trends worldwide and may help to reconcile the diversity of SAT trends across the models from phase 5 of the Coupled Model Intercomparison Project (CMIP5).

1. Introduction

With emerging public awareness of human-induced climate change, a major challenge is to understand and communicate the causes of recent observed climate

trends, particularly at local and regional scales. Such trends are often interpreted in the context of rising anthropogenic emissions of greenhouse gases (GHGs) and sulfate aerosols associated with the burning of fossil fuels. However, internally generated variability may also contribute to regional climate changes over periods of several decades and longer (e.g., Van Oldenborgh et al. 2009; Hoerling et al. 2010; Kelley et al. 2012; Ting et al. 2011; Meehl et al. 2013; Wallace et al. 2014; Johnstone and Mantua 2014; Abatzoglou et al. 2014; Srivier et al. 2015; Monier et al. 2015). Distinguishing between anthropogenic and internal influences on time scales of less than 50 years and spatial scales smaller than continental remains an outstanding issue (IPCC 2007, 2013).

* Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JCLI-D-15-0304.s1>.

⁺ The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Clara Deser, Climate and Global Dynamics Division, NCAR, 1850 Table Mesa Drive, Boulder, CO 80307.

E-mail: cdeser@ucar.edu

Global coupled climate model (GCM) simulations provide estimates of the climatic impacts of anthropogenic (e.g., GHGs and sulfate aerosols) and natural (volcanic and solar) radiative forcings (IPCC 2013). However, comparison of observed and simulated trends is complicated by the fact that the chronologies of internal variability need not match (e.g., Van Oldenborgh et al. 2009; Deser et al. 2012b; Wallace et al. 2014). Systematic model biases, incomplete radiative forcing specifications, and observational uncertainty further complicate direct comparison between models and observations (e.g., Räisänen 2007; Hegerl et al. 2007; Van Oldenborgh et al. 2009).

Isolating the effects of anthropogenic climate change from those of internal multidecadal variability is relatively straightforward in climate models, provided there are enough simulations to define the forced response (Deser et al. 2012a). That is, by averaging across ensemble members from a particular model, the random sequences of internally generated variability in the individual realizations can be sufficiently muted to reveal the model's response to external forcing. Once the externally forced response is obtained, it can be subtracted from each run to find the contribution from internal variability. Most models however, including those participating in phases 3 and 5 of the Coupled Model Intercomparison Project (CMIP3 and CMIP5; Taylor et al. 2012), contain too few realizations to adequately estimate the forced response on local/regional scales (Deser et al. 2012b). Note that, while it is common practice to average single runs from multiple models to obtain a robust estimate of anthropogenic climate change [e.g., the multimodel mean (IPCC 2013)], this approach does not allow the forced and unforced components of the response to be isolated in any given model.

The use of large "initial condition" GCM ensembles to identify the relative roles of forced and internal variability in determining future climate trends has led to important insights regarding uncertainty in climate change projections at local/regional scales (e.g., Deser et al. 2012b; Hu and Deser 2013; Fischer et al. 2013; Deser et al. 2014; Wettstein and Deser 2014; Wallace et al. 2014; Hawkins et al. 2015; Thompson et al. 2015). Such ensembles typically include 30–40 members, each of which is subject to the identical scenario of radiative forcing but starts from a slightly different atmospheric state. The degree to which the different ensemble members diverge over time is indicative of the importance of unpredictable, internal climate variability (Deser et al. 2012a). Such ensembles are designed to sample the range of possible trend outcomes resulting from the superposition of internal climate variability and anthropogenic climate change. In addition, the large number of integrations allows for a robust determination of the forced response at local/regional scales.

Here we apply a new 30-member initial-condition ensemble covering the period 1920–2100 conducted with the National Center for Atmospheric Research (NCAR) Community Earth System Model, version 1 (CESM1; Kay et al. 2015), to the understanding of observed climate trends, particularly wintertime surface air temperature (SAT) over North America during the past 50 years. Configured at a spatial resolution of 1° latitude/longitude (approximately 85 km at 40°N), this ensemble provides important context for the interpretation of the one realization in nature: specifically, the relative importance of internal variability and external radiative forcing at local/regional scales.

In addition to apportioning trends into forced and internal contributions, we assess the physical mechanisms underlying these components: in particular, the roles of dynamics (atmospheric circulation changes in the absence of radiatively induced changes in SAT) and thermodynamics (changes in SST, sea ice, and land surface properties in the absence of atmospheric circulation changes). Previous work has highlighted the importance of dynamics to observed trends in Northern Hemisphere winter SAT (e.g., Wallace et al. 1995; Hurrell 1996; Thompson et al. 2009; Cattiaux et al. 2010; Smoliak et al. 2015). However, none distinguished between internal and forced dynamical components, a unique contribution of our study enabled by the inclusion of the CESM1 Large Ensemble. While the focus of this work is on winter SAT trends over North America during the past 50 years, our methodology is generic and can be used to improve physical understanding and interpretation of observed and simulated climate variations, including trends of any length, worldwide.

The remainder of this study is organized as follows. Section 2 contains a description of the model experiments, observational datasets, and methods for determining dynamical and thermodynamic contributions to forced and internal SAT trend components. Results are presented in section 3 and discussed in section 4. Implications for model evaluation and interpretation of the CMIP5 archive are also included in section 4. A summary is provided in section 5.

2. Data and methods

a. The CESM1 Large Ensemble

We make use of a new set of simulations conducted with the Community Earth System Model, version 1 (Community Atmosphere Model, version 5) [CESM1(CAM5)], at 1° spatial resolution, referred to here as the CESM1 Large Ensemble (CESM-LE). The CESM-LE consists of 30 simulations for the period 1920–2100, each subject to the identical external radiative forcing but beginning from

slightly different atmospheric initial conditions. Following the CMIP5 design protocol, historical natural and anthropogenic radiative forcing was applied for 1920–2005 and representative concentration pathway 8.5 (RCP8.5) radiative forcing was used for 2006–2100 (Taylor et al. 2012). The first ensemble member was initialized with 1 January conditions taken from a randomly selected year of a 1700-yr preindustrial (1850) control integration (PiCTL) of the same model. This first ensemble member was then integrated forward from 1850 to 2080. Ensemble members 2–30 were all started on 1 January 1920 with initial conditions taken from the first ensemble member. A small [$O(10^{-14})$ K] random round-off difference to the initial air temperature field was added to each ensemble member. This infinitesimal perturbation serves to create spread among the ensemble members as internally generated modes of climate variability grow over time. A full description of the CESM-LE is given in Kay et al. (2015).

b. The CMIP5 ensemble

We also make use of the CMIP5. Specifically, we use a single run from each of the 38 models that submitted both historical and RCP8.5 simulations to the CMIP5 archive [see Table 9.A.1 in IPCC (2013)]. If multiple runs exist for a given model, we use the first one.

c. Observational datasets

We make use of three monthly mean SAT datasets: 1) Merged Land–Ocean Surface Temperature analysis (MLOST), version 3.5, (Vose et al. 2012) on a 5° latitude \times 5° longitude grid; 2) GISS Surface Temperature Analysis (GISTEMP) on a 2° latitude \times 2° longitude grid and smoothed with a 250-km spatial filter (Hansen et al. 2010); and 3) Climatic Research Unit Temperature, version 4, (CRUTEM4) on a 5° latitude \times 5° longitude grid (Osborn and Jones 2014). Unlike MLOST and GISTEMP, CRUTEM4 is an uninterpolated, land-only dataset. We also make use of monthly mean sea level pressure (SLP) from the Twentieth Century Reanalysis, version 2, (Compo et al. 2011) on a 2° latitude \times 2° longitude grid. Similar results are obtained for the different SAT datasets; for conciseness, we focus mainly on results obtained with MLOST.

d. Methods

For each dataset and model simulation, we compute monthly anomalies by subtracting the long-term (1963–2012) monthly means from the corresponding month of each year. We then form 3-month winter [December–February (DJF)] averages from the monthly anomalies. Finally, we compute linear trends over the 50-yr period 1963–2012 using least squares regression analysis. This

time period was chosen to match the length of record used in Deser et al. (2014) in their study of projected North American climate trends over the next 50 years (note that 2012 was the most recent year available for the observational datasets at the time this study was carried out). When computing pattern correlations and RMS differences between model simulations and observations, we regrid the model output to the observational grid.

e. Dynamical adjustment using the constructed circulation analog approach

The objective of dynamical adjustment is to empirically determine the component of SAT variability due solely to atmospheric circulation changes, all other factors being equal (e.g., no changes in ocean or land surface conditions). Here we use a variation of the constructed circulation analog method (Van den Dool et al. 2003) to find the dynamical contribution to SAT anomalies in both the CESM-LE and observations. Circulation analogs were pioneered by Lorenz (1969), Van den Dool (1994), and Van den Dool et al. (2003) as a statistical approach to weather prediction and have been applied more recently to downscale climate projections from coarse-resolution models (e.g., Zorita et al. 1995) as well as to infer the contribution of dynamics to observed European SAT trends (Cattiaux et al. 2010). A detailed description of our methodology is provided in the appendix; a summary is given here.

1) APPLICATION TO THE CESM-LE

For each month and year of each of the 30 CESM-LE simulations, we find a set of N_a closest SLP analogs within the 1700-yr CESM PiCTL, ranked according to their Euclidean distance from the target CESM-LE SLP field. For example, the SLP analogs for January 1920 of run 1 are found by searching all of the January SLP fields in the PiCTL. The advantage of using the model's control run to obtain the circulation analogs is twofold: one, there is no effect of forced climate change on either the analogs or the associated SAT anomalies; two, the number of samples from which to draw analogs is very large. We then randomly subsample N_s of the N_a analogs and compute their optimal linear combination that best fits the target CESM-LE SLP field. The dynamically induced SAT anomaly field is then defined as the corresponding optimal linear combination of PiCTL SAT anomalies associated with the N_s SLP analogs. We then repeat this random sampling procedure N_r times. Finally, we average the N_r optimal sets of SLP analogs and associated SAT anomalies to obtain a “best estimate” of the circulation-induced component of SAT anomalies in the absence of climate change. The repeated subsampling of optimal linear combinations of analogs ensures robustness

of the results (see [appendix](#)). We use the domain 20° – 90° N, 180° – 350° E for the SLP analogs following [Wallace et al. \(2012\)](#) and [Deser et al. \(2014\)](#), but there is little sensitivity to the precise region used (e.g., within $\pm 20^{\circ}$ of longitude and $\pm 10^{\circ}$ of latitude; not shown). The results shown here are based on the parameter values $N_a = 150$, $N_s = 100$, and $N_r = 50$. However, there is little sensitivity to the precise choice of these values as long as they lie within the parameter space that has converged (see the [appendix](#)).

Note that by using so many circulation analogs and by weighting them according to their degree of resemblance with the target CESM-LE SLP field, our technique can be viewed in some sense as a type of linear regression rather than a strict circulation analog approach. An advantage of our method compared to that of [Wallace et al. \(2012\)](#) and [Smoliak et al. \(2015\)](#) is that the analog weights can be applied to any field in order to estimate the dynamical contribution, including the SLP field itself. In this way, we are able to evaluate how much of the SLP anomaly field in question is reproduced by our optimal linear combination of circulation analogs. Finally, despite the different algorithms, our results are very similar to those of [Wallace et al. \(2012\)](#) and [Smoliak et al. \(2015\)](#), lending confidence to both approaches (see Fig. S1 in the supplemental material).

To obtain the forced dynamical contribution, we average the dynamical contributions for all 30 ensemble members. To obtain the internal dynamical contribution for each ensemble member, we subtract the forced dynamical contribution from the total dynamical contribution. Thermodynamic contributions are obtained as residuals (e.g., total minus dynamical) for both forced and internal components.

2) APPLICATION TO OBSERVATIONS

To dynamically adjust the observed SAT, we use SLP analogs from the Twentieth Century Reanalysis (20CR) over the period 1899–2012, leaving out the targeted month/year, and then construct their optimal linear set following the procedures described above. For the results shown here, we use $N_a = 80$, $N_s = 50$, and $N_r = 100$. The results are not sensitive to the precise choices of these parameters (see the [appendix](#)). Since we do not have an observational “control run” from which to draw the SLP analogs and their associated effects on SAT, we high-pass filter the SAT data before computing their dynamically induced component to minimize any potential effects from forced, thermodynamically induced SAT; a similar strategy was employed in [Smoliak et al. \(2015\)](#). We use the framework of the CESM-LE to test how well this approach works and to determine the sensitivity to our choice of high-pass filter. We find that subtracting a quadratic trend or a spline trend over the

full period 1920–2012 produces the closest estimates of the true unforced dynamically induced SAT (on average, within 3.7% of the true RMS amplitude using six ensemble members as test cases); 10- and 20-yr high-pass filters perform less well (within 16% and 23% of the true RMS amplitude in a smaller sample of test cases). Further information on the sensitivity to choice of high-pass filter is provided in the [appendix](#). For simplicity, the results shown here are based on quadratic trend removal.

To obtain the internal dynamical contribution to the observed SAT anomalies, we have performed a separate dynamical adjustment based on the internal component of the observed SLP anomalies, obtained by subtracting the CESM-LE ensemble-mean SLP anomaly from the observed SLP anomaly at each time step. The forced dynamical contribution to observed SAT anomalies is then obtained by subtracting the internal dynamical contribution from the total dynamical contribution. Thermodynamic contributions are obtained as residuals (e.g., total minus dynamical) for both forced and internal components.

3. Results

This section is organized as follows. We begin by showing the diversity of SAT and SLP trends simulated across the 30 CESM-LE ensemble members and placing the observed trends within this context ([section 3a](#)). We then demonstrate the key role of internal dynamics in producing much of the trend spread within the CESM-LE and its influence on observed SAT trends ([section 3b](#)). We show next that removing the influence of internal dynamics brings the observed SAT trends into closer alignment with the radiatively forced component simulated by the CESM-LE ([section 3c](#)). Additional benefits of removing internal dynamics from the SAT trends are given in [section 3d](#), including reduction of ensemble spread, enhancement of signal to noise and advancement of time of emergence. We conclude with a complete decomposition of the simulated and observed SAT trends into internal and forced dynamical and thermodynamic components, both for observations and an illustrative member of the CESM-LE ([section 3e](#)). In the remainder of this section, we use the term “dynamically adjusted SAT trends” to denote that the internal component of circulation-induced SAT trends has been removed, leaving the total (dynamic plus thermodynamic) forced component plus the internal thermodynamic component.

a. SAT and SLP trends in the CESM-LE and observations

[Figure 1](#) shows maps of 50-yr (1963–2012) winter SAT and SLP trends over North America from each of the

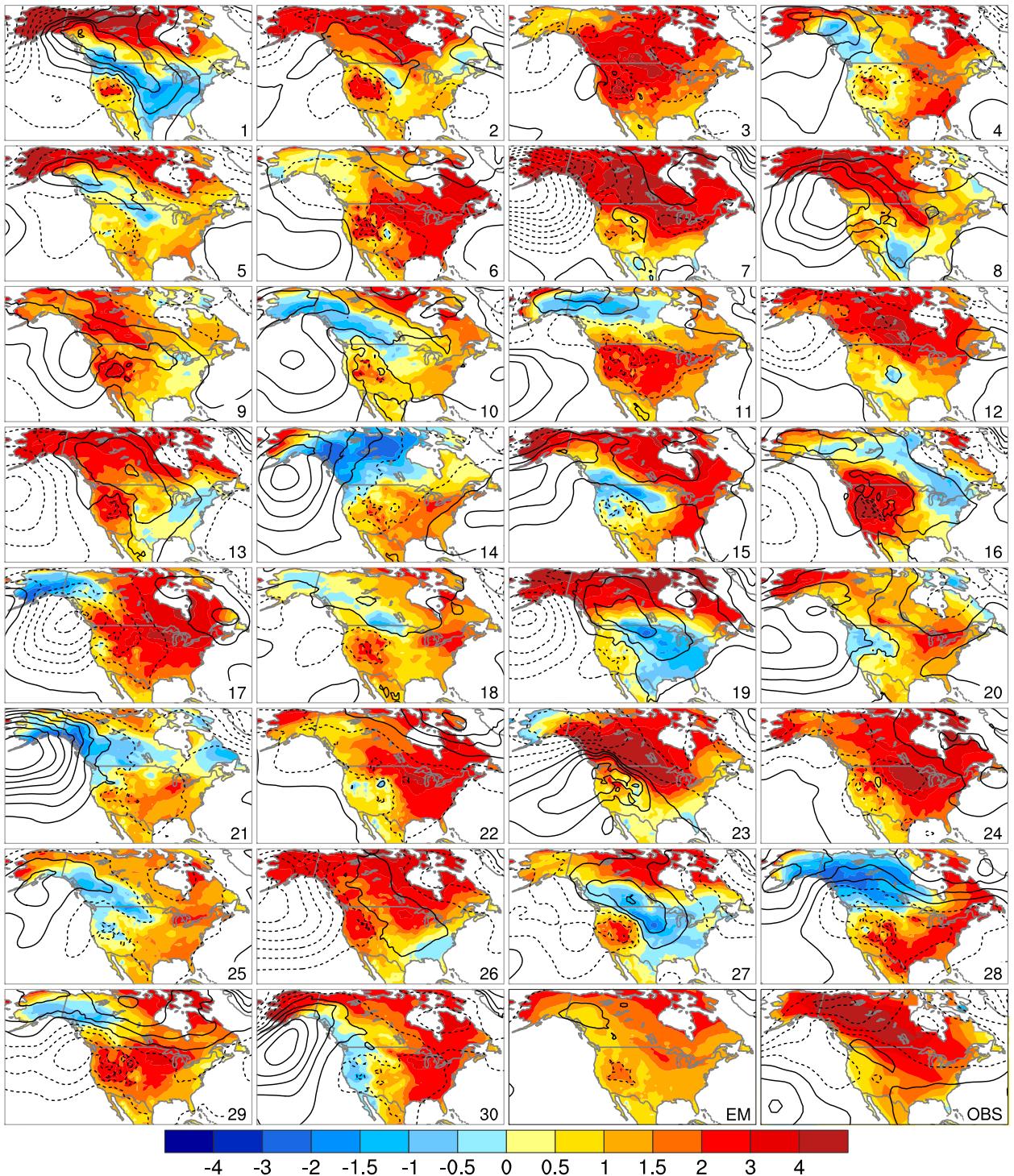


FIG. 1. DJF SAT [color shading; $^{\circ}\text{C} (50 \text{ yr})^{-1}$] and SLP [contours; $\text{hPa} (50 \text{ yr})^{-1}$] trends (1963–2012) for each member of the CSM-LE (labeled 1–30), the CSM-LE ensemble-mean trend (labeled EM), and observations (labeled OBS). SLP contour interval is $1 \text{ hPa} (50 \text{ yr})^{-1}$ starting at $\pm 0.5 \text{ hPa} (50 \text{ yr})^{-1}$, with solid (dashed) contours for positive (negative) values. SAT and SLP observation are from MLOST and 20CR, respectively.

30 members of the CESM-LE (labeled 1–30), as well as from observations and the CESM-LE ensemble mean. The SAT and SLP trends vary widely in pattern, polarity, and magnitude across the individual ensemble members, despite the fact that each simulation was conducted with the same model and subjected to the identical radiative forcing. Focusing first on SAT, it is evident that most ensemble members show warming, consistent with physical expectation because of the increase in GHG concentrations. However, despite the increased radiative forcing, it is also clear that large regions of North America can experience cooling in any single realization. Further, the pattern of SAT trends can be nearly opposite between individual runs. For example, run 28 exhibits cooling over northwestern North America and warming in the southeast, while run 19 shows opposite-signed trends. Many other examples can be found where individual ensemble members show SAT trends of opposite polarity or disparate amplitude at a given location.

The observed SAT trends (OBS) (Fig. 1, bottom-right panel) are positive everywhere, with the largest warming ($>4^{\circ}\text{C}$ over 50 yr) over northwestern Canada and smaller-amplitude warming ($<1^{\circ}\text{C}$ over 50 yr) over the western United States and far eastern Canada. Some members (e.g., runs 7, 12, and 24) resemble the observed trend pattern while others are nearly orthogonal (e.g., runs 14, 21, and 28). Viewed in the context of the 30-member CESM-LE, the observed SAT trend distribution lies within the range of simulated outcomes, each of which is consistent with the response to radiative forcing. As discussed further below, ensemble member 7 happens to provide a particularly close match to the observed trends, both in terms of pattern and magnitude, because of the chance superposition of forced and internal variability.

The degree of resemblance between the observed and simulated trend maps can be quantified in terms of centered pattern correlation (e.g., the area-average SAT trend is removed before computing the pattern correlation) and RMS difference (Fig. 2, gray dots). Pattern correlations range from -0.7 to $+0.7$, and RMS differences range from 0.6° to 1.8°C (50 yr^{-1}), with a close linear relationship between the two (e.g., lowest RMS corresponding to highest pattern correlation and vice versa). In other words, numerous SAT trend configurations are possible because of the superposition of forced and internally generated components, and no single CESM-LE run need match the one realization in nature (although given enough ensemble members, one would expect some to be realistic if the model is credible).

The externally forced component of the simulated SAT trends can be obtained by averaging the trends from all 30 members together [the CESM-LE ensemble-mean trend

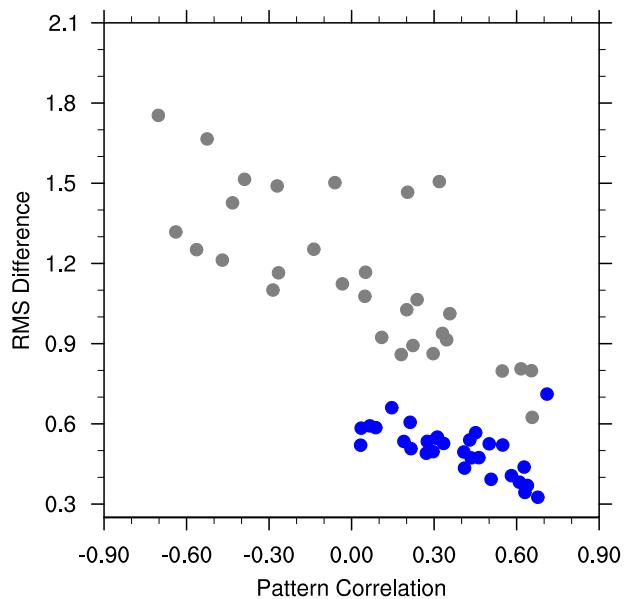


FIG. 2. Pattern correlation (x axis) vs RMS difference [y axis; $^{\circ}\text{C}$ (50 yr^{-1})] of DJF SAT trends (1963–2012) over North America for each member of the CESM-LE against observations (MLOST; 20CR). Gray dots are for total trends, and blue dots are for dynamically adjusted trends (internal dynamics removed). See text for explanation.

(EM)] (Fig. 1, bottom panel). This component shows a generally poleward-amplified warming pattern, with maximum values $\sim 2^{\circ}\text{--}3^{\circ}\text{C}$ (50 yr^{-1}) along the Arctic border compared to $\sim 0.5^{\circ}\text{--}1.5^{\circ}\text{C}$ (50 yr^{-1}) over most of the United States. In addition to poleward amplification, the forced warming trend in the CESM-LE exhibits zonal contrasts across Canada, with smaller magnitudes in the west compared to the east, and a local maximum over the U.S. Rocky Mountains. Areas of amplified warming may be due to local positive feedbacks associated with reductions in snow and sea ice.

It is instructive to compare the observed SAT trend pattern with the CESM-LE forced response, noting that agreement is not necessarily expected because of the added contribution of internal variability in observations. Like the model's forced response, the observed trend pattern shows some evidence for poleward amplification, but the largest warming occurs over western Canada [$>4^{\circ}\text{C}$ (50 yr^{-1})] instead of the Canadian archipelago and the Alaskan west coast (note, however, that the comparison is hindered by sparse data coverage over the Canadian archipelago). The pattern correlation between SAT trends in observations and the CESM-LE ensemble mean is only 0.12, with an RMS difference of 0.9°C (50 yr^{-1}) (recall that individual ensemble members exhibit much larger pattern correlations and smaller RMS differences than the ensemble mean; Fig. 2).

The diversity of SAT trends within the CESM-LE is accompanied by a variety of atmospheric circulation (SLP) trend patterns and polarities (Fig. 1). Opposite-signed trends of similar magnitude are evident over the North Pacific in individual ensemble members: for example, negative SLP trends in runs 7, 13, 17, 19, and 26 compared to positive trends in runs 8, 10, 14, 20, 21, and 30, with maximum amplitudes approximately $4\text{--}8\text{ hPa (50 yr)}^{-1}$. Notably, the ensemble-mean SLP trends during 1963–2012 are near zero everywhere, indicating that SLP trends in any given realization are almost entirely due to internal variability. This point will become especially important when we turn to the diagnosis of dynamical versus thermodynamic contributions to the forced and internal components of the SAT trends.

The observed SLP trend pattern shows negative values over the northeast Pacific extending into western Canada and weaker positive values to the south (Fig. 1, bottom right). The observed SLP trend pattern bears some resemblance to individual ensemble members, notably runs 12 and 17, and a lack of resemblance to others (e.g., runs 8, 10, 14, and 21). Consistent with this visual impression, the pattern correlations between the simulated and observed SLP trend maps range from -0.6 to $+0.8$ (Fig. 3). Interestingly, there is a close relationship between the pattern correlations of the simulated SLP and SAT trends with observations (Fig. 3). That is, ensemble members that show large positive (negative) SLP trend pattern correlations with observations also show large positive (negative) SAT trend pattern correlations with observations. This linear dependence of the SLP and SAT trend pattern correlations suggests that the circulation plays an important role in the diversity of SAT trends across the CESM-LE and, by extension, in the observed SAT trends. Next, we quantify this role through the use of dynamical adjustment.

b. Dynamically adjusted SAT trends in the CESM-LE and observations

Removing the effects of internal atmospheric circulation variability via the constructed analog technique greatly reduces the diversity of SAT trends across the CESM-LE and essentially eliminates any SLP trends (Fig. 4). Not only are the patterns and magnitudes of the SAT trends more similar between individual ensemble members, their amplitudes are generally smaller than their unadjusted counterparts (recall Fig. 1). Notably, regions of cooling are largely eliminated, and areas of extreme warming [$>4^{\circ}\text{C (50 yr)}^{-1}$] are reduced. The remaining spread across the dynamically adjusted SAT trend maps is largely attributable to thermodynamic processes (in particular, those associated with sea ice and snow cover: see section 4), as the circulation analog

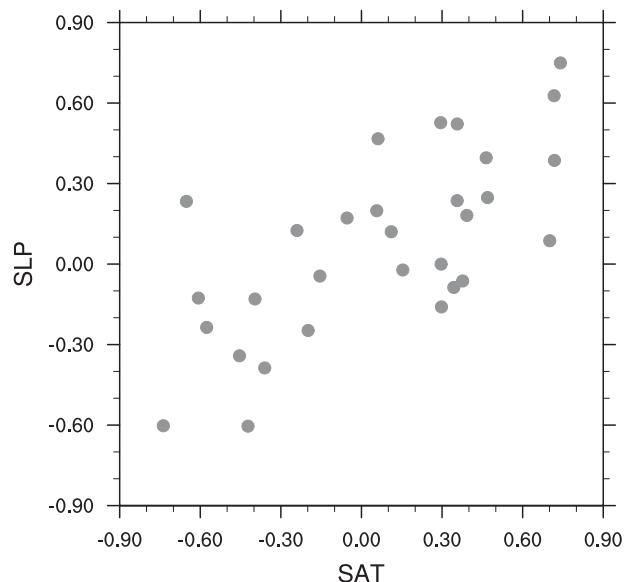


FIG. 3. Pattern correlations between simulated and observed 1963–2012 DJF SAT (x axis) and SLP (y axis) trends. Each dot represents a different CESM-LE run. Pattern correlations are based on the Pacific-North American domain for SLP and North America for SAT. SAT and SLP observation are from MLOST and 20CR, respectively.

technique removes more than 90% of the internal SLP trend variance (Fig. S2 in the supplemental material). Indeed, the SLP trend maps based on the PiCTL constructed analogs (Fig. S3 in the supplemental material) are very similar to the raw SLP trend maps, with pattern correlations >0.96 and RMS differences $<0.11\text{ hPa (50 yr)}^{-1}$ for all ensemble members.

In observations, the dynamically adjusted SAT trend map (bottom-right panel of Fig. 4) shows less spatial heterogeneity and a more dominant expression of poleward amplification compared to the raw SAT trend map. In particular, the warming across western Canada is considerably reduced, and the east–west contrast across the United States is largely alleviated.

To highlight the importance of internal circulation effects in observations, Fig. 5 compares maps of the total and (internal) dynamically induced SAT trends, along with the contribution of internal dynamics expressed as a fraction of the total SAT trend at each location. Unforced circulation trends contribute 40%–60% of the total warming over much of western Canada and the eastern United States and up to 70% in some locations. In other areas, such as the western United States, parts of Alaska, and far northeastern Canada, internal dynamics offset the warming trend by up to 20%–30%.

Dynamical adjustment improves the resemblance between the observed SAT trend distribution and those of the individual CESM-LE ensemble members (Fig. 4).

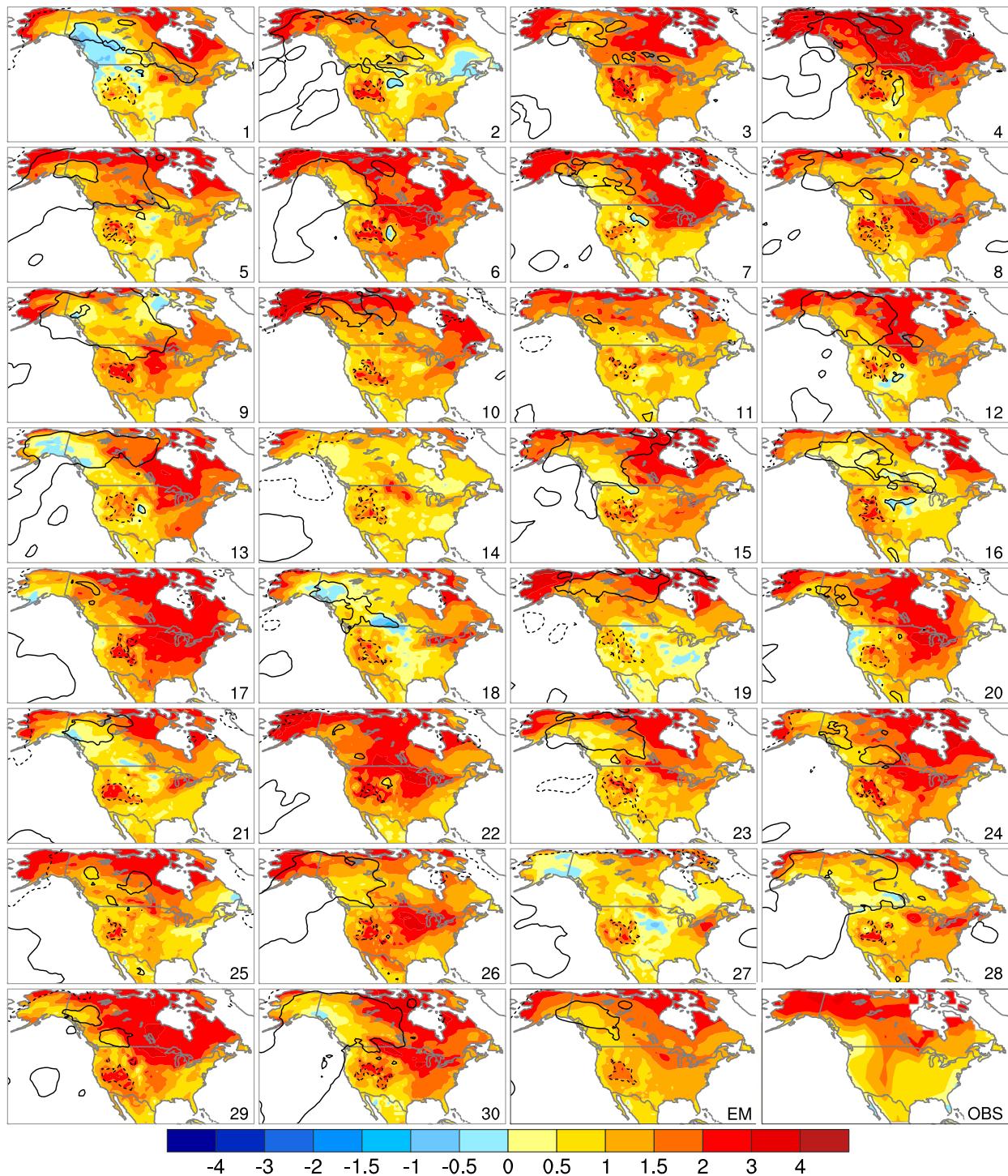


FIG. 4. As in Fig. 1, but internal circulation effects removed (see text).

In particular, the RMS differences are considerably lower after applying dynamical adjustment to the internal component of circulation trends, with values ranging from 0.32° to $0.71^{\circ}\text{C} (50\text{ yr})^{-1}$ and an average of $0.49^{\circ}\text{C} (50\text{ yr})^{-1}$ (blue circles in Fig. 2) compared to 0.62° – $1.75^{\circ}\text{C} (50\text{ yr})^{-1}$

and an average of $1.15^{\circ}\text{C} (50\text{ yr})^{-1}$ for the raw trends (gray circles in Fig. 2). The pattern correlations are also improved for the dynamically adjusted SAT trends, ranging from 0.01 to 0.72 with a mean of 0.37 compared to a range from -0.70 to 0.66 with a mean of 0.02 for the raw trends

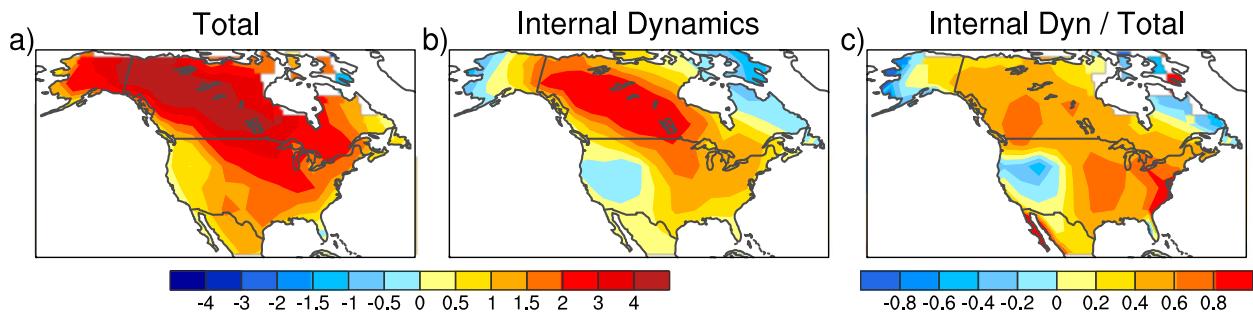


FIG. 5. Observed DJF SAT trends (1963–2012): (a) Total [$^{\circ}\text{C} (50 \text{ yr})^{-1}$], (b) contribution from internal dynamics [$^{\circ}\text{C} (50 \text{ yr})^{-1}$], and (c) fractional contribution from internal dynamics.

(Fig. 2; note that the spatial-mean trend is excluded when computing pattern correlations).

Dynamical adjustment also improves the resemblance between the observed SAT trend pattern and the forced SAT trend pattern obtained from the average of the CESM-LE members (Fig. 4, bottom-right panels labeled OBS and EM, respectively). In particular, the pattern correlation between the CESM-LE ensemble mean trend and the observed trend increases from 0.12 to 0.54, and the RMS difference decreases from 0.90° to $0.34^{\circ} (50 \text{ yr})^{-1}$, after applying dynamical adjustment. Thus, dynamical adjustment provides an improved estimate of the anthropogenic signal in observed 50-yr trends.

c. Probability distributions of raw and dynamically adjusted SAT trends

The raw and dynamically adjusted SAT trends for a given location or region may be summarized in terms of probability distributions based on the set of 30 CESM-LE simulations. These distributions also provide context for the single estimate from observations. Figure 6 shows results for three regions of contrasting spatial scale: North America, western Canada (140° – 90°W), and the grid box containing Fairbanks, Alaska. Compared to the raw trend distributions (gray bars), the dynamically adjusted ones (blue bars) are considerably narrower and more Gaussian in character. Not surprisingly, the relative degree of narrowing of the trend distributions depends on spatial scale, with the greatest reductions for Fairbanks and the smallest for North America. For example, Fairbanks shows a broad range of SAT trend values [from -1.8° to $+6.1^{\circ} (50 \text{ yr})^{-1}$] with 75% of the simulations in the range from -0.8° to $+3.5^{\circ} (50 \text{ yr})^{-1}$. After removing the influence of internal atmospheric circulation variability, the distribution narrows to a range from $+0.1^{\circ}$ to $+2.5^{\circ} (50 \text{ yr})^{-1}$ and exhibits a more Gaussian character with $\sim 75\%$ of simulations within the range from $+0.8^{\circ}$ to $+2.2^{\circ} (50 \text{ yr})^{-1}$.

For all three regions, the observed trends (raw and dynamically adjusted) lie within their respective model distributions, with the raw values (dashed orange lines)

near the upper end of the model range and the dynamically adjusted values (solid orange lines) close to the center of the model range (e.g., the model's forced response; Fig. 6). Recall from section 2e that the only model information used in the calculation of the observed dynamically adjusted SAT trends is the CESM-LE ensemble-mean SLP trends, which are nearly zero (recall Fig. 1). For North America, the observed warming trend is reduced from 2.2° to $1.4^{\circ} (50 \text{ yr})^{-1}$ after dynamical adjustment, a decrease of 37%. Put another way, internal circulation variability has augmented the warming trend due to radiative forcing (plus a small contribution from internal thermodynamics; see below) by 57%. The effects are larger for western Canada, where dynamical adjustment reduces the warming trend by 48% [from 3.5° to $1.8^{\circ} (50 \text{ yr})^{-1}$]; that is, internal circulation variability has nearly doubled the forced (plus internal thermodynamic) warming trend. The effect of dynamical adjustment for Fairbanks is smaller than that for North America and western Canada but still substantial, with a reduction in warming of 34% [from 2.9° to $1.9^{\circ} (50 \text{ yr})^{-1}$].

The sensitivity of the observational results to the choice of dataset is summarized in Table 1 and Fig. S4 in the supplemental material. The contribution of internal dynamics to the total warming trend varies by 6%–7% across the 3 SAT datasets (Table 1). In an absolute sense, the contribution of internal dynamics ranges by $0.08^{\circ} (50 \text{ yr})^{-1}$ for North America, $0.17^{\circ} (50 \text{ yr})^{-1}$ for western Canada, and $0.29^{\circ} (50 \text{ yr})^{-1}$ for Fairbanks, Alaska (Table 1); that is, not surprisingly, observational uncertainty increases as the spatial scale decreases, both for the total SAT trend and the internal dynamically induced component (Table 1). However, the broad spatial patterns and magnitudes of the SAT trends (total, internal dynamically induced, and residual components) across the three observational datasets are very similar (Fig. S4).

d. Additional consequences of dynamical adjustment

Here we quantify the utility of dynamical adjustment in reducing ensemble spread, enhancing signal to noise,

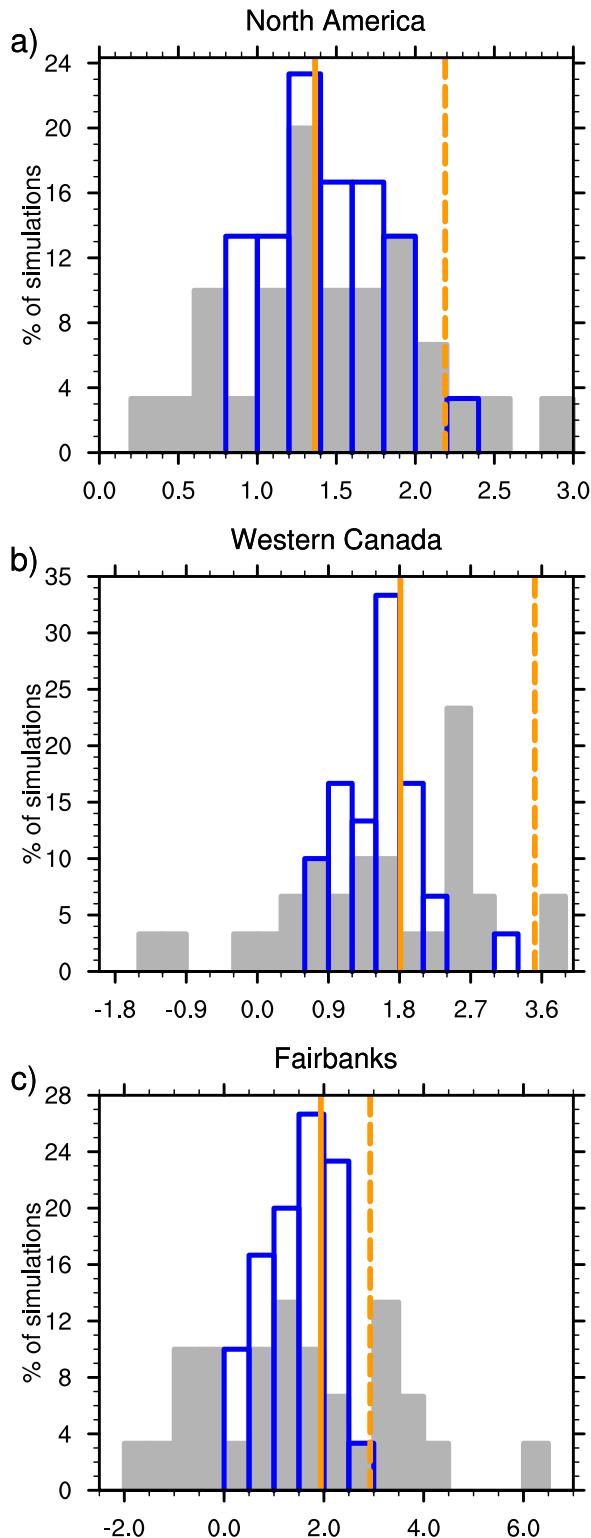


FIG. 6. Histograms of DJF SAT trends [$1963\text{--}2012; ^{\circ}\text{C} (50 \text{ yr})^{-1}$] for (a) North America, (b) western Canada, and (c) Fairbanks, Alaska. Gray (blue) bars denote raw (dynamically adjusted) trends from the CESM-LE. Dashed (solid) orange vertical lines denote raw (dynamically adjusted; internal dynamics removed) trends from observations. Note the different horizontal and vertical axis ranges in each panel.

and advancing the “time of emergence” of SAT trends in the CESM-LE. Ensemble spread, defined here as the variance of SAT trends across the individual ensemble members, is reduced by more than 80% over much of the western and southeastern portions of North America and by 50%–60% around Hudson Bay, the Great Lakes, central Rockies, and Mexican highlands (Fig. 7c). Averaged over the continent as a whole, removal of internal dynamical influences via our dynamical adjustment procedure reduces the variance of SAT trends (1963–2012) in the CESM-LE by 75%, more than in any other season (not shown). The actual SAT trend variance in the raw data is largest over western Canada and Alaska extending southeastward along a narrow band to the Great Lakes and over the U.S. Rocky Mountains, with maximum values $\sim 3\text{--}5 [^{\circ}\text{C} (50 \text{ yr})^{-1}]^2$ (Fig. 7a). In contrast, the SAT trend variance in the dynamically adjusted data is nearly uniform across the continent, with values $0.5\text{--}1 [^{\circ}\text{C} (50 \text{ yr})^{-1}]^2$ over Alaska and central Canada and $<0.5 [^{\circ}\text{C} (50 \text{ yr})^{-1}]^2$ elsewhere (Fig. 7b).

Related to the reduction in ensemble spread, dynamical adjustment increases the signal-to-noise ratio (SNR) of SAT trends, where the signal is defined as the forced (ensemble mean) SAT trend and the noise as the standard deviation of SAT trends across the 30 ensemble members (Fig. 7f). Notably, dynamical adjustment more than doubles the SAT trend SNR over the northwestern and southeastern portions of the continent and increases it by at least 50% elsewhere (Fig. 7f). The raw SAT trends show SNR values that are close to unity over most of the continent (and <1 in parts of the northwest): only Mexico and a small area east of Hudson Bay exhibit values >2 (Fig. 7d). On average, the raw SAT trend SNR over North America is 1.3, indicating that the forced climate change signal and internal variability are comparable for winter SAT trends over the past 50 yr. In contrast, dynamically adjusted SAT trends show SNR values generally above two over most of the continent, with higher values (>3) along the Arctic border, Florida, the southwestern United States, and central Mexico and lower values (1.5–2) over western Canada and the north-central United States (Fig. 7e). On average, the dynamically adjusted SAT trend SNR over North America is 2.3, compared to 1.3 for the raw trends.

Many studies have examined the time of emergence (TOE) of the anthropogenic climate change signal, estimated by evaluating when the forced response first exceeds a given amplitude of internal variability [typically, one or two standard deviations (e.g., Mahlstein et al. 2011; Diffenbaugh and Scherer 2011; Deser et al. 2012a)]. Here we assess the effect of dynamical adjustment on TOE. We compute TOE from the CESM-LE based on 10-yr running means of DJF SAT anomalies,

TABLE 1. Sensitivity of observed DJF SAT trends [1963–2012; $^{\circ}\text{C} (50 \text{ yr})^{-1}$] to choice of dataset for North America, western Canada, and Fairbanks, Alaska (same regions used in Fig. 6). “Total” indicates the total trend; “internal dynamics” indicates the trend due to internal dynamics; and “residual” indicates their difference (total minus internal dynamics). Percent change indicates the residual expressed as a fraction of the total (%). Datasets are MLOST, CRUTEM4 (CRUT), and GISTEMP (GIST). See text for details.

	North America			Western Canada			Fairbanks, Alaska		
	MLOST	CRUT	GIST	MLOST	CRU	GIST	MLOST	CRUT	GIST
Total	2.19	2.53	2.44	3.51	3.95	3.83	2.93	2.86	3.16
Internal dynamics	0.82	0.78	0.86	1.70	1.61	1.78	0.99	1.17	1.30
Residual	1.37	1.75	1.58	1.81	2.34	2.05	1.94	1.68	1.86
% change	−37%	−31%	−35%	−48%	−41%	−46%	−34%	−41%	−41%

where anomalies are computed relative to the period 1920–49. Specifically, we find the year when the forced (ensemble mean) 10-yr running mean SAT anomaly first exceeds and remains above one standard deviation of the internal variability (computed across the 30 SAT anomalies at each time step after applying a 10-yr running mean). We also compute TOE based on dynamically adjusted data (internal dynamics removed). This “dynamically adjusted” TOE will generally be earlier

than the raw TOE because of the reduced amplitude of internal variability in the dynamically adjusted SAT compared to the raw SAT.

The TOE map based on the raw 10-yr running mean DJF SAT anomalies shows that the forced signal has not yet emerged above the internal variability over a wide swath of the western portion of the continent (gray shading) and has only recently emerged (since 2005) over much of the intermountain west, the central portion of

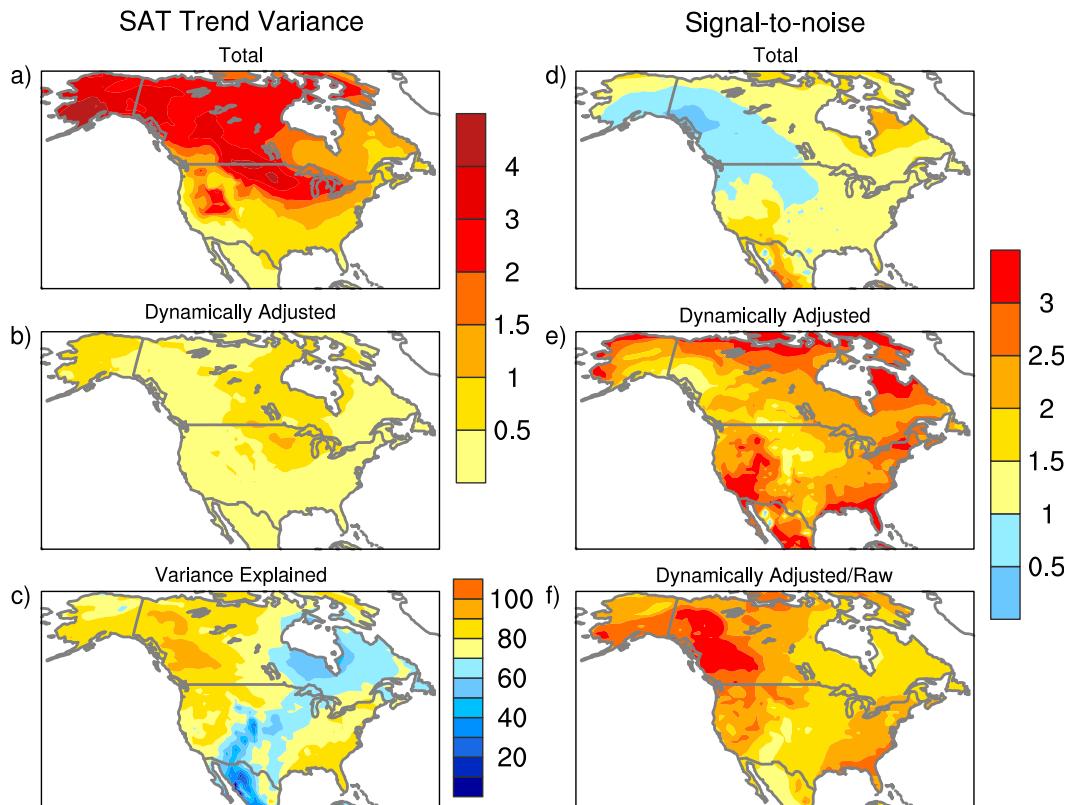


FIG. 7. (Left) Variance of DJF SAT trends $\{1963\text{--}2012; [^{\circ}\text{C} (50 \text{ yr})^{-1}]^2\}$ across the CESM-LE based on: (a) total and (b) dynamically adjusted (internal dynamics removed) data. (c) The proportion of total trend variance (%) accounted for by dynamical adjustment [e.g., (a) minus (b) divided by (a)]; red (blue) shading indicates that internal dynamics contributes to a high (low) proportion of the total SAT trend variance. (right) SNR for DJF SAT trends (1963–2012) from the CESM-LE based on (d) total and (e) dynamically adjusted (internal dynamics removed) data. (f) The ratios of the SNR values in (e) divided by those in (d). Signal is defined as the ensemble-mean trend, and noise is defined as the standard deviation of the trends across the 30 members.

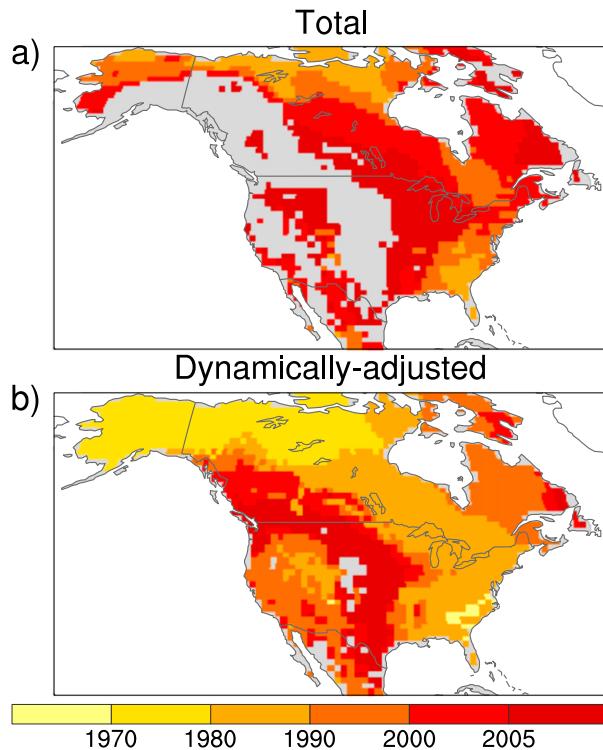


FIG. 8. Maps of TOE of forced winter SAT anomalies based on 10-yr running means from the CESM-LE for (a) total and (b) dynamically adjusted (internal dynamics removed) data. See text for details. Gray areas denote grid boxes where the forced signal has not yet emerged by 2012.

North America, and far eastern Canada (Fig. 8a). The earliest TOE values (1980s) occur along the Arctic border and the southeastern United States (Fig. 8a). Dynamical adjustment advances the TOE considerably at all locations, and nearly all areas are now “emergent” (Fig. 8b). Specifically, the dynamically adjusted TOE values are in the 1970s over Alaska and northern Canada, the 1980s over much of the eastern half of the continent, the 1990s over the western United States, and after 2005 over the midsection of the United States and along the western portion of the Canadian–U.S. border (Fig. 8b). This advancement of TOE is entirely due to the reduction of interannual variance in the dynamically adjusted data compared to the raw data, since the forced signal is the same in both calculations.

e. Decomposition of forced and internal SAT trends into dynamic and thermodynamic components in the CESM-LE and observations

Having established the central role of internal atmospheric dynamics to the CESM-LE and observed SAT trend patterns over North America, we conclude our analysis with a complete decomposition of the SAT trends into forced and internal dynamical and thermodynamic

components. This full decomposition is performed for observations and for an illustrative member of the CESM-LE (run 7 whose SAT trends happen to bear the closest resemblance to observations; recall Fig. 1). The results are shown in Fig. 9a for the CESM-LE and Fig. 9b for observations. The model results are discussed first. The top row shows the total SAT trends and their constituent internal and forced parts, the latter obtained from the CESM-LE ensemble mean and the former obtained by subtracting the forced component from the total. The middle row shows the contribution of dynamics to the total, internal, and forced components. As mentioned earlier, the total dynamical contribution is obtained directly by applying the constructed analog technique to the total SLP trend field for that particular run; the forced dynamical contribution is obtained by averaging the total dynamical contributions for all 30 ensemble members; the internal dynamical contribution is obtained by subtracting the forced dynamics contribution from the total dynamical contribution. Finally, the bottom row shows the contribution of thermodynamics to the total, internal, and forced components, obtained as residuals between the total and dynamical contributions. Strictly speaking, the residuals also contain dynamical contributions not accounted for in the circulation analog technique (as a result of both inadequate sampling as well as methodological uncertainty). However, as discussed in the appendix, these errors are generally very small given the length of control run used to identify the constructed analogs as well as the repeated random sampling and averaging of the optimal linear combinations of constructed analogs.

As already described, run 7 features strong warming over the northern two-thirds of the continent, with maximum values exceeding $4^{\circ}\text{C} (50\text{yr})^{-1}$ over large regions of Canada and the northeast United States [Fig. 9a(i)]. Atmospheric circulation trends contribute much of the warming over western Canada by virtue of anomalous southerly flow associated with a deepened Aleutian low [recall Fig. 4; Fig. 9a(ii)]. Thermodynamic processes are responsible for the enhanced warming in the vicinity of Hudson Bay, the Canadian archipelago, and northern Alaska [Fig. 9a(iii)], in association with diminished snow cover and sea ice cover (not shown). The total dynamical component [Fig. 9a(ii)] is mainly a result of internal circulation trends [Fig. 9a(v)], with forced circulation trends contributing mainly to warming along the Arctic border [Fig. 9a(viii)]. On the other hand, the total thermodynamic contribution [Fig. 9a(iii)] is largely forced [Fig. 9a(ix)]; the internal thermodynamic component [Fig. 9a(vi)] augments the forced warming over eastern Canada, the Great Lakes region, and Alaska and partially offsets the forced warming elsewhere. In summary, both dynamics

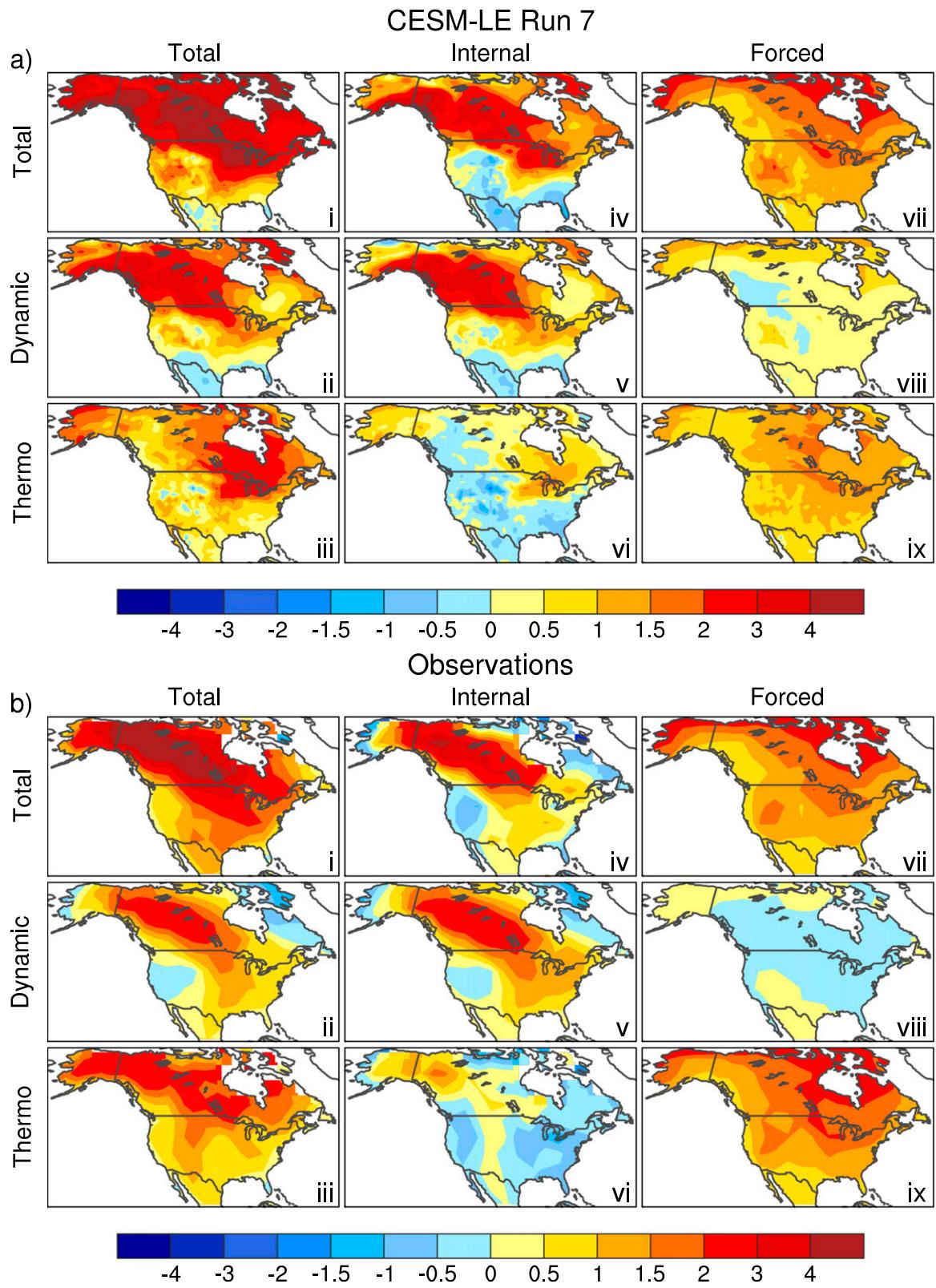


FIG. 9. Decomposition of DJF SAT trends [1963–2012; $^{\circ}\text{C} (50 \text{ yr})^{-1}$] into internal, forced, dynamical, and thermodynamic components for (a) run 7 of the CESM-LE and (b) observations (MLOST). See text for details.

and thermodynamics contribute to the total SAT trends; further, the dynamical contribution is almost entirely internal, whereas the thermodynamic contribution is both internal and forced.

The same decomposition is performed for the observed SAT trends in Fig. 9b. Strictly speaking, the only partitioning that can be made based purely on observations is the separation of the raw trends into dynamical and thermodynamic components (left column). Using the model's forced SLP response in conjunction with observed SLP–SAT relationships enables separation of the dynamical contribution into internal and forced components (middle row). Specifically, we have performed a separate dynamical adjustment for observations based on the internal component of the SLP anomalies at each time step, where the internal component is obtained by subtracting the ensemble-mean SLP anomaly from the CESM-LE from the observed SLP anomaly. The forced dynamical contribution in observations is then obtained by subtracting the internal dynamical contribution from the total dynamical contribution. The remaining components require use of the model's forced SAT and were determined as follows (note that, for this purpose, the model's forced SAT was regridded to the mesh of the observed SAT data). The total forced component of the observed SAT is equated to the model's total forced SAT, and the total internal component of the observed SAT is obtained by subtracting the forced SAT from the observed SAT. The forced thermodynamic component of the observed SAT is obtained by subtracting the forced dynamical component of the observed SAT from the total forced SAT. Finally, the internal thermodynamic component of the observed SAT is obtained by subtracting the internal dynamic component from the total internal component.

Dynamics contributes to warming [2° – 3°C (50 yr^{-1})] across western Canada and weak cooling over northeastern Canada, western Alaska and the northwestern United States [Fig. 9b(ii)] while thermodynamics accounts for a broad pattern of warming over the entire continent, with largest amplitudes over Canada [Fig. 9b(iii)]. These patterns are remarkably similar to those from CESM-LE run 7 [Figs. 9a(ii), 9a(iii)], as are the patterns of the total trends [Fig. 9a(i)], although they are based on completely independent datasets. This, in turn, lends confidence to the robustness of the decomposition and interpretation of the results in both the model and observations. The circulation-induced component of the observed SAT trends [Fig. 9b(ii)] is almost entirely internal [Fig. 9b(v)], similar to that found for run 7 of the CESM-LE except along the Arctic border, where the model shows modest warming from forced circulation changes [Fig. 9a(viii)]. It is difficult to assess whether the

differences along the Arctic border are a result of model shortcomings or observational data constraints, since this is a region of limited data coverage. Further, this region may be subject to thermodynamic influences from dynamically forced changes in sea ice cover that are implicitly included in the circulation analogs from the model's control run but may not be adequately sampled in the short observational record. Finally, it is evident that the thermodynamic component of the observed SAT trends [Fig. 9b(iii)] is mainly forced [Fig. 9b(ix)], with secondary contributions from internal processes that augment the warming over western Canada and Alaska and cool much of the United States [Fig. 9b(vi)], in line with the results from run 7. In summary, observed DJF SAT trends over the past 50 years are largely the result of internal dynamics and forced thermodynamics, with secondary contributions from internal thermodynamics.

A complementary, temporal view of the relative contributions of internal and forced dynamics and thermodynamics to North American SAT anomalies during 1920–2012 in run 7 of the CESM-LE is shown in Fig. 10 (observational results are shown in Fig. S5 of the supplemental material). The top pair of curves contrasts the raw (black) and dynamically adjusted (magenta) SAT records. While the raw time series displays prominent interannual fluctuations superimposed upon a long-term warming trend beginning in the early 1960s, the dynamically adjusted record exhibits more muted variability and a smaller but more monotonic rise in SAT starting in the late 1970s (similar results are obtained when both forced and internal dynamics are removed; not shown). Thus, not only does dynamical adjustment alter the timing and amplitude of the long-term warming trend, it also provides for a more stable estimate of these parameters because of the reduction in interannual variance (noise in this context). The second pair of curves contrasts the forced (red) and internal (blue) components of North American SAT. The forced time series bears a close resemblance to the dynamically adjusted record, with even less interannual variability, while the internal time series is dominated by high-frequency fluctuations. The third pair of curves compares the thermodynamic (brown) and dynamical (orange) contributions to forced SAT. It is clear that thermodynamics dominates, with dynamics making a small but nonnegligible contribution to the forced component of warming in recent decades. The final pair of curves compares the thermodynamic (green) and dynamical (cyan) contributions to the internal SAT record. Both are primarily high frequency in character, with larger amplitudes for the dynamical component than the thermodynamic one. In summary, dynamical adjustment leads to an improved estimate of the forced component of the North American SAT time series and

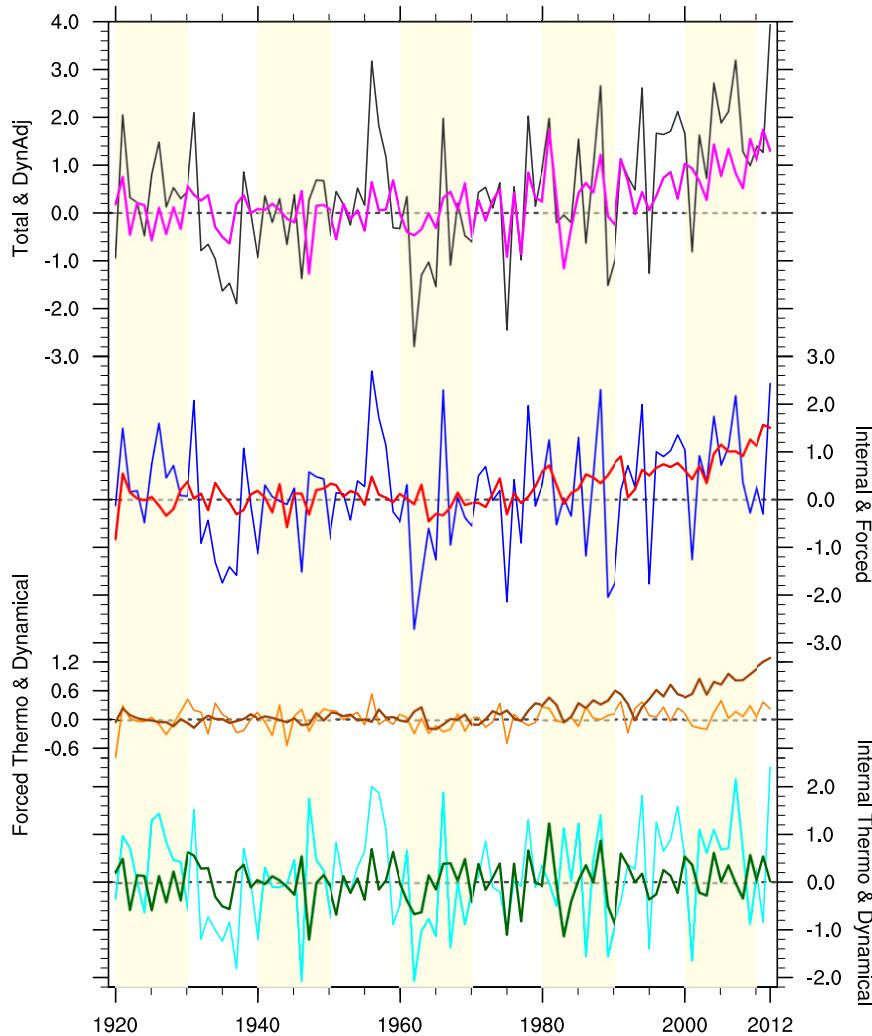


FIG. 10. Time series decomposition of DJF SAT anomalies ($^{\circ}\text{C}$) averaged over North America from run 7 of the CESM-LE into internal, forced, dynamical, and thermodynamic components. (top) The total (black) and dynamically adjusted (magenta; internal dynamics removed) components. (upper middle) The internal (blue) and forced (red) components. (lower middle) The forced thermodynamics (brown) and forced dynamics (orange) components. (bottom) The internal thermodynamics (green) and internal dynamics (cyan) components. Note the different vertical scales for each set of curves. See text for explanation.

provides a more stable estimate (e.g., less subject to sampling fluctuations) of the timing and amplitude of low-frequency trends.

4. Discussion

The utility of dynamical adjustment for attributing SAT trends was demonstrated in the seminal studies of Wallace et al. (2012) and Smoliak et al. (2015) using a variety of techniques. These studies focused primarily on the period 1965–2000, an interval characterized by a prominent trend in the northern annular mode with

attendant effects on winter SAT over North America and Eurasia. Similar to our results, both studies found that dynamics account for approximately 40% of the total cold season SAT trend over land poleward of 40°N during 1965–2000. As shown in Smoliak et al. (2015), the spatial pattern of the dynamical contribution to SAT trends over North America during 1965–2000 resembles that found in our study based on 1963–2012, except for the lack of cooling over the western United States. Neither Wallace et al. (2012) nor Smoliak et al. (2015) distinguished between internal and forced contributions to atmospheric circulation trends during 1965–2000.

A key aspect of our study was to make use of the forced SLP response as given by the ensemble mean of the CESM-LE. This, in conjunction with the observed relationships between SLP and SAT anomalies, enabled us to evaluate both the internal and forced dynamical contributions to observed SAT trends. In addition, by incorporating the model's forced SAT response, we were able to further decompose the remaining portion of the observed SAT trends into internal and forced thermodynamic components. It is noteworthy that the forced component of winter SLP anomalies over the North Pacific and North America in any given year, and in trends over the last 50 years, is small (nearly zero) compared to the internal component, both according to the CESM-LE (Fig. 1, EM panel) and the CMIP5 multimodel archive (Fig. S6 in the supplemental material, EM panel). The extent to which this result is model dependent remains to be assessed.

Although dynamical adjustment greatly reduces the diversity of SAT trends across the 30-member CESM-LE, differences remain, especially in the vicinity of Hudson Bay, the Great Lakes, and Alaska (recall Fig. 7b). For example, dynamically adjusted warming trends over much of Canada exceed 4°C in run 4 but are $<1.5^{\circ}\text{C}$ in run 27 (recall Fig. 4). What processes contribute to the remaining spread in dynamically adjusted SAT trends? The pattern of residual SAT trend variance (Fig. 11a) suggests that snow cover and sea ice changes may play a role. Indeed, a similar pattern is obtained by regressing trends in dynamically adjusted SAT onto trends in sea ice concentration averaged over Hudson Bay (Fig. 11b). Our physical interpretation of this regression pattern is that internal trends in Hudson Bay sea ice (possibly initiated by atmospheric forcing) feed back onto SAT locally and over adjacent land areas via thermodynamic processes. In addition, associated changes in snow cover west of the Canadian and U.S. Great Lakes may amplify these thermodynamic SAT feedbacks. We speculate that similar physical mechanisms are at work over Alaska. Further investigation of the processes contributing to the spread in the thermodynamic component of internal SAT trends is warranted.

As discussed in Hawkins and Sutton (2009) and other studies, the CMIP5 archive contains three sources of uncertainty: model response uncertainty due to structural differences between models, radiative forcing uncertainty, and uncertainty due to internal variability. How large is the contribution of structural model uncertainty versus internal variability uncertainty to the spread of SAT trends within CMIP5 for a given radiative forcing scenario, and does internal atmospheric circulation variability play a role? This question is difficult to address without a sufficient number of ensemble members to define the forced response in each model,

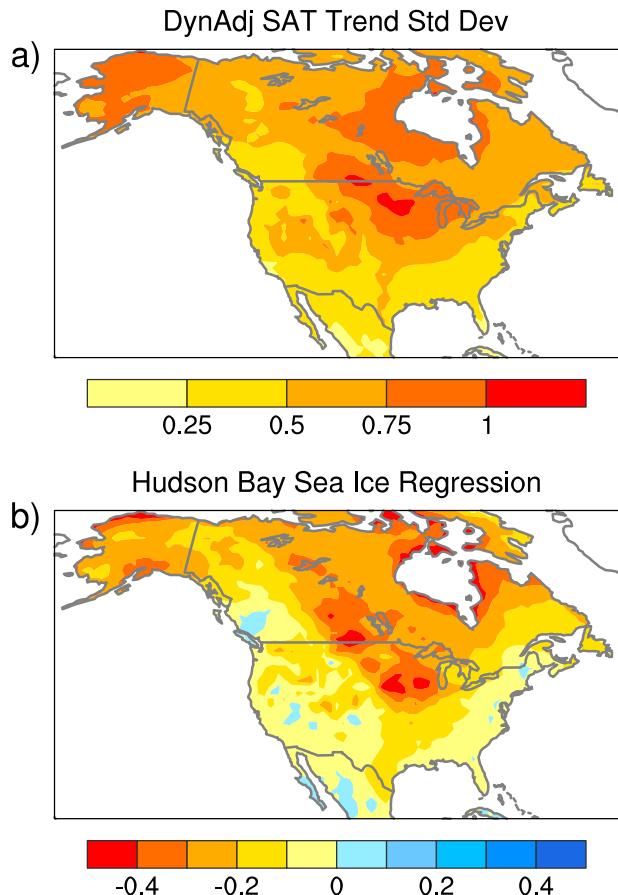


FIG. 11. (a) Standard deviation of dynamically adjusted DJF SAT trends [$1963\text{--}2012; ^{\circ}\text{C} (50\text{yr})^{-1}$] across the CESM-LE. (b) Regression map of dynamically adjusted SAT trends onto the Hudson Bay sea ice trend index. See text for details.

although results based on a previous 40-member initial-condition ensemble with the Community Climate System Model, version 3 (CCSM3), suggested an important role for internal variability in the spread within CMIP3 (Deser et al. 2012b).

We can gain some insight into these questions, however, by comparing the SLP and SAT trends simulated by the CMIP5 models with those simulated by the CESM-LE. Qualitatively, the diversity of SAT and SLP trend patterns and magnitudes within the set of CMIP5 model runs is reminiscent of that within the CESM-LE, although there appear to be fewer cases of strong negative SAT trends (Fig. S6). Particularly noteworthy is the wide range of SLP trends whose multimodel mean (panel labeled EM in Fig. S6) is near zero, just as in the CESM-LE. A more quantitative comparison is provided in Fig. 12a, which shows pattern correlations and RMS differences between the SAT trends from each of the 38 CMIP5 models against observations (red dots). The set of CMIP5 model runs shows a similar range of pattern

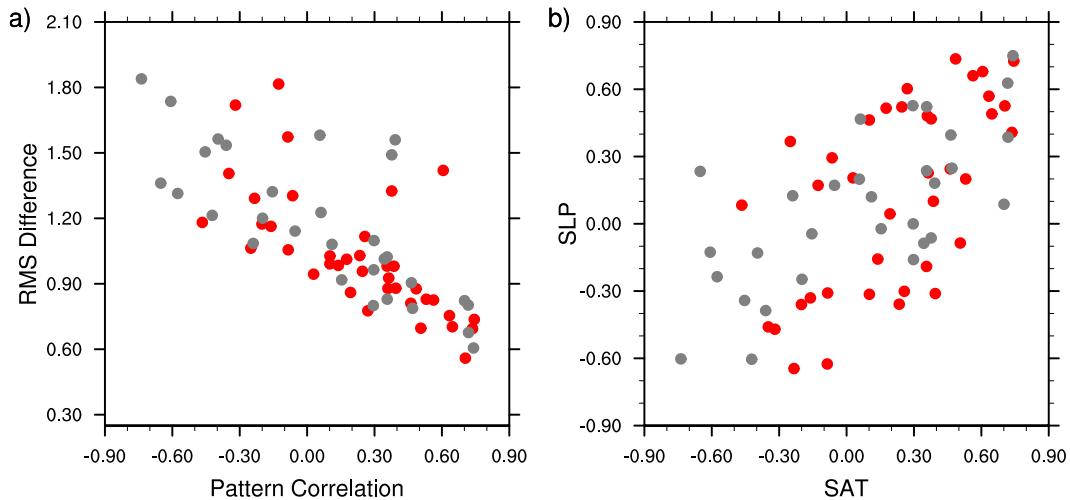


FIG. 12. (a) Pattern correlation vs RMS difference [$^{\circ}\text{C} (50 \text{ yr})^{-1}$] of DJF SAT trends (1963–2012) over North America for each member of the CESM-LE (gray dots) and each member of the CMIP5 archive (red dots) against observations (MLOST). (b) Pattern correlations between simulated and observed 1963–2012 winter SAT and SLP trends. Dots are for the CESM-LE (gray) and CMIP5 (red) simulations. Pattern correlations are based on the Pacific-North American domain for SLP and North America for SAT.

correlations (albeit with fewer large negative values) and RMS differences as the CESM-LE (gray dots in Fig. 12a), suggesting that internal variability may be an important source of SAT trend spread within the CMIP5 archive. The range of pattern correlations between the observed and simulated SLP trends is also comparable across the CMIP5 models (red dots) and the CESM-LE ensemble (gray dots), and importantly, the SLP and SAT pattern correlations show a similar relationship in the two model ensembles (Fig. 12b). Taken together, these results are suggestive of an important role for internal dynamics (as opposed to model response uncertainty) in the diversity of North American SAT trends within the CMIP5 archive. Further investigation of the role of internal variability versus model uncertainty in the CMIP5 archive is clearly warranted.

Another important implication of our results pertains to model evaluation. A common metric for climate model assessment is the skill with which observed trends during recent decades are simulated. However, as this study highlights, a single simulation with a credible model need not match the observed climate trajectory at local/regional scales if the contribution from internal variability is comparable or larger than that from external forcing. Conversely, a single integration with a model that is lacking in realism may show fortuitous agreement with the observed trends. The methodology outlined in this paper, including the application of dynamical adjustment in conjunction with an estimate of the forced circulation response, can be used to provide a more informative and reliable assessment of a model's ability to simulate observed climate trends.

Thompson et al. (2015) have recently argued that climate model evaluation would be well served to focus on two simple statistical properties of the unforced interannual variability: standard deviation σ and auto-correlation. These two parameters (especially σ) dictate the confidence intervals that can be placed on trends of any length due to internal variability, provided the data are normally distributed and stationary in time [see discussion in Thompson et al. (2015)]. In this context, it is worth noting that the CESM-LE simulates generally realistic magnitudes (within 10%–20% of observations in most regions) and spatial patterns of SAT and SLP σ for both unfiltered and 8-yr low-pass filtered data (Fig. S7 in the supplemental material).

Our methodology, based on a combination of observations, a large initial-condition ensemble of historical simulations, and a technique for estimating the dynamical contribution, has general applicability beyond the specific purpose of this study. In particular, it can be used to inform attribution of observed climate anomalies on a near-real-time basis, as well as improve the physical understanding of differences in climate changes simulated by different models. As with any empirical method, inherent uncertainties due to limited sampling of circulation statistics in the short (~ 100 yr) observational record must be taken into account. This problem is ameliorated to a great extent in models for which lengthy control simulations exist. There is also the caveat that, as models improve and resolve finer spatial and vertical scales, their forced SLP (and SAT) responses may be altered, leading to new insights regarding the role

of the atmospheric circulation in anthropogenic climate change.

5. Summary

By combining observations with a 30-member initial-condition ensemble of CESM1 coupled model simulations (the CESM-LE) and making use of “dynamical adjustment,” we have provided insight into the mechanisms of internal and forced components of winter SAT trends over North America during the past 50 years. Our results show that, at local/regional scales, the simulated SAT trends are strongly influenced by internal variability, with an average signal-to-noise ratio of only 1.3. The importance of internal variability is evidenced by the variety of spatial patterns, amplitudes, and polarities of trends among the 30 simulations, each of which is subject to the identical scenario of historical radiative forcing. The range of model solutions spans the single realization of the real world, providing important context for the interpretation of observed long-term trends.

The internally generated component of SAT trends within the CESM-LE is largely dynamically induced, whereas the forced component is primarily thermodynamically controlled, either directly via radiative effects from increased GHGs or indirectly via changes in SSTs, sea ice, and snow cover. This follows from the fact that the simulated winter SLP trends in the CESM-LE over the past 50 years are almost entirely a result of internal climate variability, with a negligible forced component. (It remains to be seen whether other climate models with sufficient ensemble size also show negligible forced SLP trends for the time period and region examined here; however, we note that the CMIP5 multimodel mean agrees with the CESM-LE in this respect.) For the real world, we estimate that internal circulation trends account for approximately one-third of the observed wintertime warming over North America during the past 50 years and more than half locally over parts of Canada and the United States. In a few areas, such as the central Rocky Mountains, far western Alaska, and northeastern Quebec province, internal dynamics has offset the warming trend by 10%–30%. Removing the effects of internal atmospheric circulation variability via a constructed analog technique narrows the spread within the CESM-LE, thereby enhancing the signal-to-noise ratio of the simulated SAT trends (by a factor of 2, on average) and advancing the “time of emergence” of the forced SAT component (by approximately a decade in many locations). Dynamical adjustment also brings the observed trends closer to the model’s forced response, both in terms of pattern and amplitude, facilitating their interpretation.

The methodological framework proposed in this study provides a general template for improving physical understanding and interpretation of observed and simulated climate trends worldwide. Application to other seasons, regions, time periods, and parameters will be pursued in future work.

Acknowledgments. We have benefited from useful discussions with Drs. Julien Boé, Brian V. Smoliak, David W. J. Thompson, and John M. Wallace during the course of this work. We thank the three reviewers for their constructive comments that helped to improve the manuscript. We also acknowledge the efforts of all those who contributed to producing the CESM-LE and thank the Climate and Global Dynamics Division at NCAR for supporting L. Terray’s sabbatical visit. NCAR is sponsored by the National Science Foundation.

APPENDIX

Constructed Analog Technique for Dynamical Adjustment

Here we provide additional details on our constructed analog methodology for dynamical adjustment. As described in the main text, we first determine the N_a closest analogs \mathbf{X}_c from the PiCTL for each time step of a particular CESM-LE run \mathbf{X}_h , and then randomly draw N_s of the N_a analogs. We then estimate a constructed SLP analog \mathbf{X}_{ca} as a linear combination of these N_s analogs according to

$$\mathbf{X}_h \approx \mathbf{X}_{ca} = \mathbf{X}_c \boldsymbol{\beta}, \quad (\text{A1})$$

where \mathbf{X}_c is a matrix of column vectors comprising the selected N_s closest analogs of the column vector \mathbf{X}_h , and $\boldsymbol{\beta}$ is a column vector of the fitted regression coefficients that are the linear proportions of the contributions of each column of \mathbf{X}_c to the constructed analog \mathbf{X}_{ca} . The dimensions of \mathbf{X}_h are $m \times 1$, where m is the number of grid points contained in the SLP pattern. The dimensions of \mathbf{X}_c are $m \times N_s$ and that of $\boldsymbol{\beta}$ are $N_s \times 1$. The $\boldsymbol{\beta}$ coefficients can be estimated by using the Moore–Penrose pseudoinverse of \mathbf{X}_c :

$$\boldsymbol{\beta} = [(\mathbf{X}_c^T \mathbf{X}_c)^{-1} \mathbf{X}_c^T] \mathbf{X}_h. \quad (\text{A2})$$

In practice, we calculate $\boldsymbol{\beta}$ using a singular value decomposition of \mathbf{X}_c . The $\boldsymbol{\beta}$ coefficients are then applied to the N_s SAT patterns from the PiCTL simulation corresponding to the selected SLP analogs to reconstruct the dynamically induced component of SAT in the CESM-LE run.

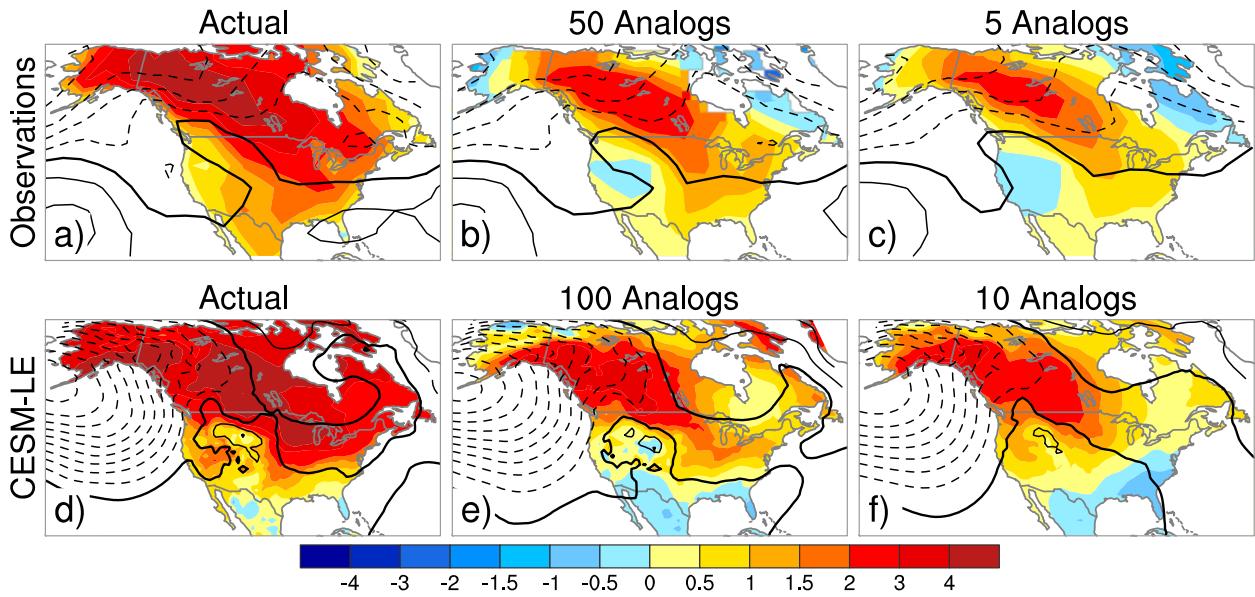


FIG. A1. Total estimated dynamical contribution to SLP and SAT trends (1963–2012) as a function of the number of closest N_s analogs used. (top) Observational results and (bottom) model results based on run 7 of the CESM-LE. (a),(d) The actual trends; (b),(c),(e),(f) the constructed analog trend estimates [SAT in color shading, $^{\circ}\text{C} (50 \text{ yr})^{-1}$, and SLP in contours, contour interval = $1 \text{ hPa} (50 \text{ yr})^{-1}$] based on different choices of N_s , as stated in the panel titles.

For each month, we repeat the drawing of the N_s analogs N_r times. We then end up with N_r different samples of dynamically induced SAT. Note that these differ mainly because of thermodynamically induced internal variability. We then take the mean of the N_r samples as the final dynamically induced SAT. The spread among the N_r samples gives a lower bound for the spread because of thermodynamically induced internal variability. We repeat all the steps for each month of the CESM-LE simulation to obtain a complete reconstruction

of dynamically induced SAT in that run. Finally, we repeat the entire algorithm for each member of the CESM-LE.

Figure A1 shows the sensitivity of the results to the choice of closest N_s analogs, for both observations and run 7 of the CESM-LE. The SLP trends obtained from the constructed analog methodology closely match the actual SLP trends in both observations and CESM-LE run 7 for a wide range of N_s values (5–50 for observations and 10–100 for the model run; note that the observations were only computed for $N_s \leq 50$). Similarly, only minor

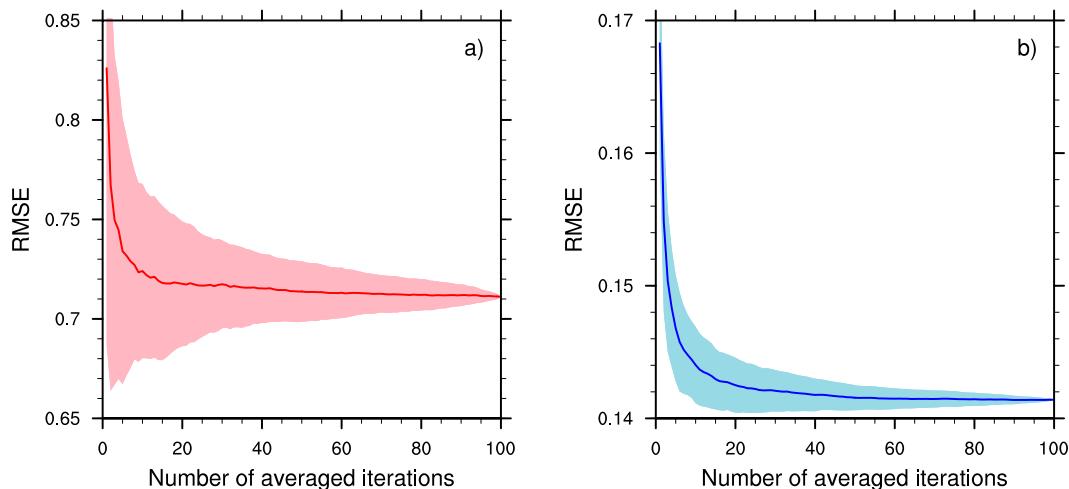


FIG. A2. RMS error (RMSE) as a function of N_r for (a) SAT [$^{\circ}\text{C} (50 \text{ yr})^{-1}$] and (b) SLP [$\text{hPa} (50 \text{ yr})^{-1}$] DJF trends (1963–2012) from the CESM-LE. Shading indicates the range across an ensemble of 200 averaging iteration steps that differ by a random initial sequence of N_r SAT and SLP trends. The colored curve denotes the average over the ensemble.

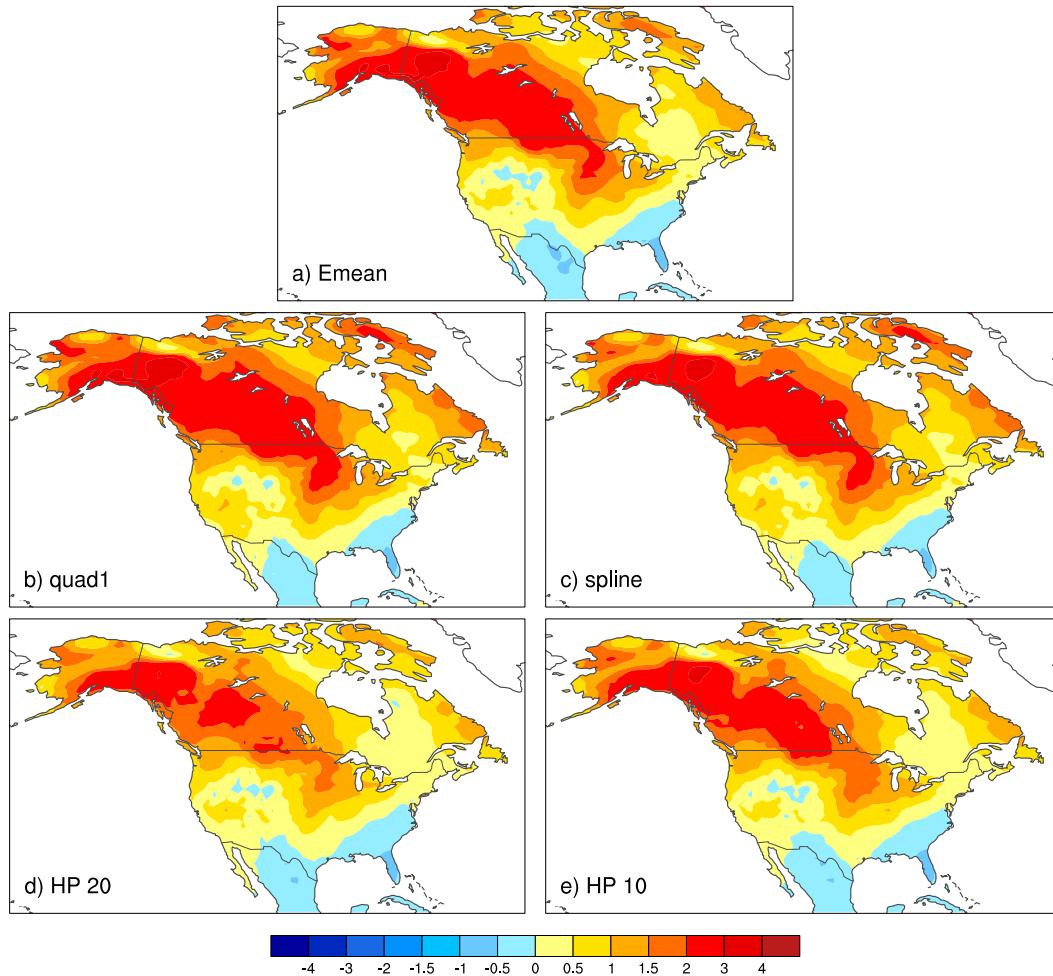


FIG. A3. Dynamical contribution to SAT trends [1963–2012; $^{\circ}\text{C} (50 \text{ yr})^{-1}$] in run 7 of the CESM-LE. Shown are (a) the “true” estimate and the estimates based on (b) quadratic trend removal, (c) spline trend removal, (d) 20-yr high-pass filter, and (e) 10-yr high-pass filter. See text for explanation.

differences in the total estimated dynamical contribution to SAT trends are found between $N_s = 50$ and $N_s = 5$ for observations (Figs. A1b,c, respectively) and between $N_s = 100$ and $N_s = 10$ for the model run (Figs. A1e,f, respectively), although the dynamical contribution to simulated SAT trends is slightly greater when more closest analogs are used.

The sensitivity of the results to N_r , quantified in terms of RMS error, is shown in Fig. A2 for both SAT and SLP. The shading indicates the range across all 30 CESM-LE simulations, and the colored curve denotes the average over the ensemble. It is clear that the results converge for $N_r > 20$ and that there is less spread for SLP compared to SAT across all N_r .

As discussed in section 2e of the main text, we have used the framework of the CESM-LE to determine which high-pass filter to use for dynamically adjusting the observations. Using a single run from the CESM-LE,

we can subtract the forced SLP and SAT responses (obtained from the CESM-LE ensemble mean) to obtain the unforced “residual” component of variability. We can then find circulation analogs and associated SAT anomalies from this residual record and apply our dynamical adjustment procedure to obtain the “true” unforced dynamically induced component in that run. We can then compare this true dynamical estimate to that obtained without knowledge of the forced response (as is the case when dealing with observations). The latter is found by high-pass filtering the SAT record over the period 1920–2012 and then applying our constructed circulation analog protocol.

We have compared the true estimate of the dynamical SAT trend contribution with those based on quadratic and spline trend removal for six of the CESM-LE ensemble members [these cases (runs 1, 7, 16, 21, 28, and 30) were chosen to span a broad range of patterns; recall

Fig. 1]. Both methods yield RMS amplitudes that differ from the true RMS amplitude by 3.7% on average for the six test cases. The 10-yr and 20-yr high-pass filters perform less well (these were tested on only one CESM-LE ensemble member, since their RMS errors were found to be much larger than any of those based on the quadratic and spline trend removals: 16% and 23%, respectively). Figure A3 illustrates the results for ensemble member 7 of the CESM-LE, comparing the true estimate of the dynamical SAT trend contribution (Fig. A3a) with those based on the four different choices of high-pass filter: quadratic trend removal (Fig. A3b), spline trend removal (Fig. A3c), 20-yr high-pass filter (Fig. A3d), and 10-yr high-pass filter (Fig. A3e). All four choices yield similar patterns of dynamical contribution, but the quadratic and spline trend removals are the closest to the true estimate.

REFERENCES

- Abatzoglou, J. T., D. E. Rupp, and P. W. Mote, 2014: Seasonal climate variability and change in the Pacific Northwest of the United States. *J. Climate*, **27**, 2125–2142, doi:10.1175/JCLI-D-13-00218.1.
- Cattiaux, J., R. Vautard, C. Cassou, P. Yiou, V. Masson-Delmotte, and F. Codron, 2010: Winter in 2010 in Europe: A cold extreme in a warming climate. *Geophys. Res. Lett.*, **37**, L20704, doi:10.1029/2010GL044613.
- Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, doi:10.1002/qj.776.
- Deser, C., R. Knutti, S. Solomon, and A. S. Phillips, 2012a: Communication of the role of natural variability in future North American climate. *Nat. Climate Change*, **2**, 775–779, doi:10.1038/nclimate1562.
- , A. S. Phillips, V. Bourdette, and H. Teng, 2012b: Uncertainty in climate change projections: The role of internal variability. *Climate Dyn.*, **38**, 527–546, doi:10.1007/s00382-010-0977-x.
- , M. A. Alexander, and B. V. Smoliak, 2014: Projecting North American climate over the next 50 years: Uncertainty due to internal variability. *J. Climate*, **27**, 2271–2296, doi:10.1175/JCLI-D-13-00451.1.
- Diffenbaugh, N. S., and M. Scherer, 2011: Observational and model evidence of global emergence of permanent, unprecedented heat in the 20th and 21st centuries. *Climatic Change*, **107**, 615–624, doi:10.1007/s10584-011-0112-y.
- Fischer, E. M., U. Beyerle, and R. Knutti, 2013: Robust spatially aggregated projections of climate extremes. *Nat. Climate Change*, **3**, 1033–1038, doi:10.1038/nclimate2051.
- Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. *Rev. Geophys.*, **48**, RG4004, doi:10.1029/2010RG000345.
- Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.*, **90**, 1095–1107, doi:10.1175/2009BAMS2607.1.
- , R. S. Smith, J. M. Gregory, and D. A. Stainforth, 2015: Irreducible uncertainty in near-term climate projections. *Climate Dyn.*, 1–13, doi:10.1007/s00382-015-2806-8.
- Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 663–746. [Available online at <https://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter9.pdf>.]
- Hoerling, M. P., J. Eischeid, and J. Perlwitz, 2010: Regional precipitation trends: Distinguishing natural variability from anthropogenic forcing. *J. Climate*, **23**, 2131–2145, doi:10.1175/2009JCLI3420.1.
- Hu, A., and C. Deser, 2013: Uncertainty in future regional sea level rise due to internal climate variability. *Geophys. Res. Lett.*, **40**, 2768–2772, doi:10.1002/grl.50531.
- Hurrell, J. W., 1996: Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature. *Geophys. Res. Lett.*, **23**, 665–668, doi:10.1029/96GL00459.
- IPCC, 2007: *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, 996 pp.
- , 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, 1535 pp., doi:10.1017/CBO9781107415324.
- Johnstone, J. A., and N. J. Mantua, 2014: Atmospheric controls on northeast Pacific temperature variability and change, 1900–2012. *Proc. Natl. Acad. Sci. USA*, **111**, 14 360–14 365, doi:10.1073/pnas.1318371111.
- Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) Large Ensemble Project: A community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.*, **96**, 1333–1349, doi:10.1175/BAMS-D-13-00255.1.
- Kelley, C., M. Ting, R. Seager, and Y. Kushnir, 2012: The relative contributions of radiative forcing and internal climate variability to the late 20th Century drying of the Mediterranean region. *Climate Dyn.*, **38**, 2001–2015, doi:10.1007/s00382-011-1221-z.
- Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646, doi:10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2.
- Mahlstein, I., R. Knutti, S. Solomon, and R. W. Portmann, 2011: Early onset of significant local warming in low latitude countries. *Environ. Res. Lett.*, **6**, 031002, doi:10.1088/1748-9326/6/3/034009.
- Meehl, G., A. Hu, J. Arblaster, J. Fasullo, and K. Trenberth, 2013: Externally forced and internally generated decadal climate variability associated with the interdecadal Pacific oscillation. *J. Climate*, **26**, 7298–7310, doi:10.1175/JCLI-D-12-00548.1.
- Monier, E., X. Gao, J. R. Scott, A. Sokolov, and C. A. Schlosser, 2015: A framework for modeling uncertainty in regional climate change. *Climatic Change*, **131**, 51–66, doi:10.1007/s10584-014-1112-5.
- Osborn, T. J., and P. D. Jones, 2014: The CRUTEM4 land-surface air temperature data set: Construction, previous versions and dissemination via Google Earth. *Earth Syst. Sci. Data*, **6**, 61–68, doi:10.5194/essd-6-61-2014.
- Räsänen, J., 2007: How reliable are climate models? *Tellus*, **59A**, 2–29, doi:10.1111/j.1600-0870.2006.00211.x.
- Smoliak, B. V., J. M. Wallace, P. Lin, and Q. Fu, 2015: Dynamical adjustment of the Northern Hemisphere surface air temperature field: Methodology and application to observations. *J. Climate*, **28**, 1613–1629, doi:10.1175/JCLI-D-14-00111.1.
- Sriver, R. L., C. E. Forest, and K. Keller, 2015: Effects of initial conditions uncertainty on regional climate variability: An analysis using a low-resolution CESM ensemble. *Geophys. Res. Lett.*, **42**, 5468–5476, doi:10.1002/2015GL064546.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.
- Thompson, D. W. J., J. M. Wallace, P. D. Jones, and J. J. Kennedy, 2009: Identifying signatures of natural climate variability in time

- series of global-mean surface temperature: Methodology and insights. *J. Climate*, **22**, 6120–6140, doi:10.1175/2009JCLI3089.1.
- , E. A. Barnes, C. Deser, W. E. Foust, and A. S. Phillips, 2015: Quantifying the role of internal climate variability in future climate trends. *J. Climate*, **28**, 6443–6456, doi:10.1175/JCLI-D-14-00830.1.
- Ting, M., Y. Kushnir, R. Seager, and L. Cuihua, 2011: Robust features of Atlantic multi-decadal variability and its climate impacts. *Geophys. Res. Lett.*, **38**, L17705, doi:10.1029/2011GL048712.
- Van den Dool, H. M., 1994: Searching for analogues, how long must one wait? *Tellus*, **46A**, 314–324, doi:10.1034/j.1600-0870.1994.t01-2-00006.x.
- , J. Huang, and Y. Fan, 2003: Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981–2001. *J. Geophys. Res.*, **108**, 8617, doi:10.1029/2002JD003114.
- Van Oldenborgh, G. J., S. Drijfhout, A. van Ulden, R. Haarsma, A. Sterl, C. Severijns, W. Hazeleger, and H. Dijkstra, 2009: Western Europe is warming much faster than expected. *Climate Past*, **5**, 1–12, doi:10.5194/cp-5-1-2009.
- Vose, R. S., and Coauthors, 2012: NOAA's Merged Land–Ocean Surface Temperature Analysis. *Bull. Amer. Meteor. Soc.*, **93**, 1677–1685, doi:10.1175/BAMS-D-11-00241.1.
- Wallace, J. M., Y. Zhang, and J. A. Renwick, 1995: Dynamic contribution to hemispheric mean temperature trends. *Science*, **270**, 780–782, doi:10.1126/science.270.5237.780.
- , Q. Fu, B. V. Smoliak, P. Lin, and C. M. Johanson, 2012: Simulated versus observed patterns of warming over the extratropical Northern Hemisphere continents during the cold season. *Proc. Natl. Acad. Sci. USA*, **109**, 14 337–14 342, doi:10.1073/pnas.1204875109.
- , C. Deser, B. V. Smoliak, and A. S. Phillips, 2014: Attribution of climate change in the presence of internal variability. *Climate Change: Multidecadal and Beyond*, C. P. Chang et al., Eds., World Scientific Series on Asia-Pacific Weather and Climate, Vol. 6, World Scientific, 1–29.
- Wettstein, J. J., and C. Deser, 2014: Internal variability in projections of twenty-first-century Arctic sea ice loss: Role of the large-scale atmospheric circulation. *J. Climate*, **27**, 527–550, doi:10.1175/JCLI-D-12-00839.1.
- Zorita, E., J. P. Hughes, D. P. Lettenmaier, and H. von Storch, 1995: Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *J. Climate*, **8**, 1023–1042, doi:10.1175/1520-0442(1995)008<1023:SCORCP>2.0.CO;2.