Geophysical Research Letters

RESEARCH LETTER
10.1029/2020GL090873

Key Points:

- NOAA's global surface temperature analysis has limited polar coverage, resulting in a small cold bias in recent decades
- We create a spatially complete analysis for 1850–2018 using Arctic air temperatures and climate reanalysis fields
- Full coverage has a slight impact on global trends, but it significantly increases warming rates in the Arctic since at least the 1980s

Supporting Information:

- Supporting Information S1

Correspondence to:

R. Vose,
Russell.Vose@noaa.gov

Citation:

Vose, R. S., Huang, B., Yin, X., Arndt, D., Easterling, D. R., Lawrimore, J. H., et al. (2021). Implementing full spatial coverage in NOAA's global temperature analysis. *Geophysical Research Letters*, *48*, e2020GL090873. https://doi.org/10.1029/2020GL090873

Received 22 SEP 2020
Accepted 24 DEC 2020

Implementing Full Spatial Coverage in NOAA's Global Temperature Analysis

R. S. Vose[1], B. Huang[1], X. Yin[2], D. Arndt[1], D. R. Easterling[1], J. H. Lawrimore[1], M. J. Menne[1], A. Sanchez-Lugo[1], and H. M. Zhang[1]

[1]NOAA's National Centers for Environmental Information, Asheville, NC, USA, [2]Riverside Technology Inc, Fort Collins, CO, USA

Abstract The National Oceanic and Atmospheric Administration (NOAA) maintains an operational analysis for monitoring trends in global surface temperature. Because of limited polar coverage, the analysis does not fully capture the rapid warming in the Arctic over recent decades. Given the impact of coverage biases on trend assessments, we introduce a new analysis that is spatially complete for 1850–2018. The new analysis uses air temperature data in the Arctic Ocean and applies climate reanalysis fields in spatial interpolation. Both the operational analysis and the new analysis show statistically significant warming across the globe and the Arctic for all periods examined. The analyses have comparable global trends, but the new analysis exhibits significantly more warming in the Arctic since 1980 (0.598°C dec^{-1} vs. 0.478°C dec^{-1}), and its trend falls outside the 95% confidence interval of its operational counterpart. Trend differences primarily result from coverage gaps in the operational analysis.

Plain Language Summary NOAA provides a suite of climate services to government, business, academia, and the public to support informed decision-making. Among these services is the State of the Climate report, which is a collection of monthly summaries recapping climate-related occurrences across the globe. This report relies heavily upon NOAA's global surface temperature data set to depict recent monthly conditions and long-term changes. Our research introduces a new edition of this flagship data set that is based upon additional temperature data and improved scientific methods. The new data set extends back to 1850 and has complete coverage of all land and ocean areas for the first time. These improvements are particularly important in the Arctic, which has warmed more rapidly than the rest of the planet in recent decades, and the new data set likewise has larger trends than its predecessor in that part of the world. The introduction of this new data set is consistent with the NOAA practice of periodically developing improved versions of its foundational datasets, the goal being to ensure the best possible representation of historical conditions across the globe. The results of this study suggest that the new data set can substantially contribute to future NOAA monitoring and assessment activities.

1. Introduction

Global surface temperature (GST) is a key indicator of observed climate change. Several groups maintain GST analyses derived primarily from historical data; examples include Berkeley Earth (Rohde et al., 2013), Cowtan and Way (2014), NASA (Lenssen et al., 2019), NOAA (Zhang et al., 2019), and the Met Office Hadley Centre/Climatic Research Unit (Morice et al., 2012). These groups regularly track the status of GST as part of their climate monitoring operations, and their products support assessment activities such as those by the Intergovernmental Panel on Climate Change (IPCC), which document long-term GST changes to inform international policy discussions. Extending back to the 19th century, GST analyses traditionally employ air temperature data from weather stations over land and sea surface temperature (SST) data from ships and buoys over the ocean. Because data availability varies in space and degrades back in time, GST analyses often use statistical methods to improve spatial coverage and thereby minimize bias in global averages. There are a variety of statistical methods for addressing coverage biases, including distance-weighted averaging (Hansen et al., 2010); reduced-space methods (Vose, Arndt, et al., 2012); and hybrid kriging techniques that blend in situ and satellite data (Cowtan & Way, 2014).

Of the major GST products available, the NOAA analysis (NOAAGlobalTemp) is unique in that its statistical approach generates quasi-complete coverage of most tropical and midlatitude regions but very limited

coverage of polar latitudes. Because the Arctic has warmed faster than the rest of the planet, limited polar coverage implies NOAAGlobalTemp contains a cool bias, particularly in recent decades (Cowtan & Way, 2014; Karl et al., 2015). In a climate assessment context, it is important that GST reconstruction methods limit as much as possible the effects of coverage changes. Consequently, this paper introduces a new edition of NOAAGlobalTemp that is spatially complete since 1850. The approach for achieving a spatially complete product is threefold: the use of air temperature data in the Arctic Ocean, the application of climate reanalysis fields in spatial interpolation, and the elimination of geographic masking criteria from reconstructed temperatures. To quantify the impact of full spatial coverage, we compare trends in the new version (hereafter termed Interim) with trends in its operational counterpart (version 5) over several representative periods.

2. Arctic Air Temperatures

NOAAGlobalTemp consists of monthly temperature anomalies on a 5° × 5° latitude-longitude grid. Land data are from the Global Historical Climatology Network-Monthly (GHCNv4) data set (Menne et al., 2018); NOAAGlobalTemp uses approximately 750 GHCNv4 stations in 1900, 5,000 in 1950, and 10,000 in 2000. Ocean data are from the Extended Reconstructed Sea Surface Temperature (ERSSTv5) data set (Huang et al., 2017), a statistical reconstruction based on ship and buoy observations from the International Comprehensive Ocean-Atmosphere Data Set-Release 3 (ICOADS R3; Freeman et al., 2017) and Argo observations from the Global Data Assembly Centre (Argo, 2000). GHCNv4 and ERSSTv5 contain homogeneity adjustments that address historical changes in temperature instrumentation, observing practice, and other factors (Huang et al., 2017; Menne & Williams, 2009). The purpose of these adjustments is to provide a consistent record that is free of the changes humans have introduced in the process of observing temperatures worldwide.

NOAAGlobalTemp version 5 and Interim employ different data over the Arctic Ocean, which is largely covered by sea ice in winter. In particular, version 5 uses SSTs over the Arctic, and when ice concentrations exceed 60%, SSTs are gradually damped toward the freezing point of seawater (–1.8°C). In contrast, Interim uses air temperatures in the Arctic, and regardless of ice concentration, there is no lower temperature bound, which improves consistency between temperatures over the terrestrial surface and sea ice. The Interim approach presumes that sea ice and snow cover insulate the air from the underlying marine environment, an assumption generally supported by observations and reanalyses (e.g., Comiso & Hall, 2014; Simmons et al., 2014). Notably, several other GST analyses (e.g., Cowtan & Way, 2014; Lenssen et al., 2019) already extrapolate air temperatures over sea ice. In Interim, air temperatures are used over all water and ice surfaces poleward of 65°N, which is a simple approximation of maximum southward sea ice extent in boreal winter.

ICOADS R3 is the primary source of air temperature data for the Arctic Ocean in Interim. The data set contains historical records from a variety of platforms, including ships, buoys, and ice stations. Ship data are sporadic until the mid-20th century and largely restricted to ice-free areas throughout the record. From 1954 to 1991, high-quality data are available for 1–3 ice stations per year in the High Arctic—i.e., manned "North Pole" stations operated by the Soviet Arctic and Antarctic Research Institute (Romanov et al., 1997). More recently, buoy data appear in earnest in the 1980s. All of the ship, buoy, and ice station data received quality assurance reviews as a part of ICOADS R3 processing, including the filtering of outliers more than 4.5 standard deviations from the smoothed median (Freeman et al., 2017). Reporting frequencies varied by platform type (usually every 3–6 h for ships and every hour for buoys); to account for this, we computed daily temperatures from the original subdaily data for each ship, buoy, and ice station.

The International Arctic Buoy Program (IABP) is a secondary source of air temperature data for the Arctic Ocean in Interim. This data set contains historical records for ice buoys deployed to support real-time operations as well as research applications. Very generally, air temperatures are available for at least 15 buoys per year from 1979 to 2000 and for at least 40 buoys per year thereafter. We applied a suite of quality assurance reviews to address potential data issues in the ice buoy data (see Martin & Munoz, 1997; Rigor et al., 2000). Modeled after the GHCN approach (Durre et al., 2008; Lawrimore et al., 2011), these reviews focused on random errors, such as data spikes resulting from electrical malfunctions, as well as systematic errors, such

as periods with extremely low variance resulting from snow-covered temperature sensors. In general, ice buoys with shorter records were somewhat more prone to data quality problems, so Interim only includes ice buoys having at least six months of data. Reporting frequencies varied by period (usually every 3 h before 2010 and hourly thereafter); once again, daily temperatures were computed from the subdaily data to compensate for this.

We derived monthly grid-box anomalies for the Arctic Ocean in Interim using the daily temperatures from ICOADS R3 and IAPB. The first step involved calculating the average temperature of each 5° × 5° grid box on each day using the data for all platforms located in the grid box on that day. The second step required computing the average temperature of each grid box in each year and month (assuming the box had at least 10 days of data in the month). The third step entailed calculating the grid-box anomaly in each year and month by subtracting the calendar-month grid-box normal from the actual grid-box temperature. Because few grid boxes had complete records, we estimated normals using atmospheric reanalysis fields for 1981–2000. Furthermore, because ensemble means often exhibit better performance than individual reanalyses (Hofer et al., 2012; Vose, Applequist, et al., 2012), we used the average of five reanalyses: the 20th Century Reanalysis version 3 (Slivinski et al., 2019); the Climate Forecast System Reanalysis version 2 (Saha et al., 2014); ERA5 (Hersbach et al., 2020); the Japanese 55-year Reanalysis (Kobayashi et al., 2015); and the Modern Era Reanalysis for Research and Applications version 2 (Gelaro et al., 2017). Since each reanalysis had a different spatial resolution, we computed normals on the native reanalysis grids first, then reinterpolated to the 5° × 5° grid. The final normal at each grid box was the average of the five interpolated values. As the last step in completing Arctic coverage for Interim, we merged the ocean grid boxes with land grid boxes from GHCNv4. Grid boxes over coastal areas were the area-weighted average of the ocean and land anomalies. (For additional details on the spatial and temporal coverage of Arctic temperature data described in this section, see the Supporting Information.)

3. Statistical Reconstruction Approach

Interim employs a statistical approach to create a spatially complete reconstruction of air temperature that covers all global land areas as well as all oceans poleward of 65° (i.e., areas with extensive seasonal ice coverage). In essence, the approach generates a continuous surface rather than an exact interpolation through the original grid-box data, thereby minimizing small-scale noise while allowing for the estimation of anomalies in grid boxes without data. This section briefly summarizes the approach and its few (but important) differences with the method used in version 5, particularly the application of climate reanalysis fields in spatial interpolation and the elimination of geographic masking criteria from reconstructed temperatures. For a more detailed treatment, see Vose, Arndt, et al. (2012), Smith et al. (2008), and Smith and Reynolds (2005).

In brief, the reconstruction approach extracts the primary low- and high-frequency signals in the historical temperature record, then blends them to produce a final reconstruction. Low-frequency signals are longer-term, broader-scale features such as the rapid warming over most land areas since the mid-1970s. High-frequency signals are shorter-term, smaller-scale features such as the strong warm anomaly over Russia during the boreal winter of 2020. The first step involves creating low-frequency grids through a nonparametric smoothing process—i.e., by averaging the original anomaly grids over broad areas and compositing across multiyear periods (Smith & Reynolds, 2005). The second step entails creating high-frequency grids using a pattern-recognition technique—i.e., by fitting spatial covariance modes known as Empirical Orthogonal Teleconnections (EOTs; Van den Dool et al., 2000) to anomaly grids from which the low-frequency signal has been subtracted. The selection of each EOT mode is based on the availability of observations within the areas covered by the mode, which ensures that the reconstruction accurately represents conditions in data-rich areas while filling gaps in data-void areas. The third step simply involves adding together the low- and high-frequency grids into composite fields for each year and month.

The reconstruction process uses a suite of input data sets for the identification of low- and high-frequency signals. For low-frequency signals over land areas, both version 5 and Interim use GHCNv4. For low-frequency signals over polar oceans, Interim employs observed air temperatures over the Arctic Ocean (see Section 2) and extrapolated air temperatures from Antarctic stations over the Southern Ocean (which is necessary because there are a paucity of ship and buoy air temperatures over this area, which represents

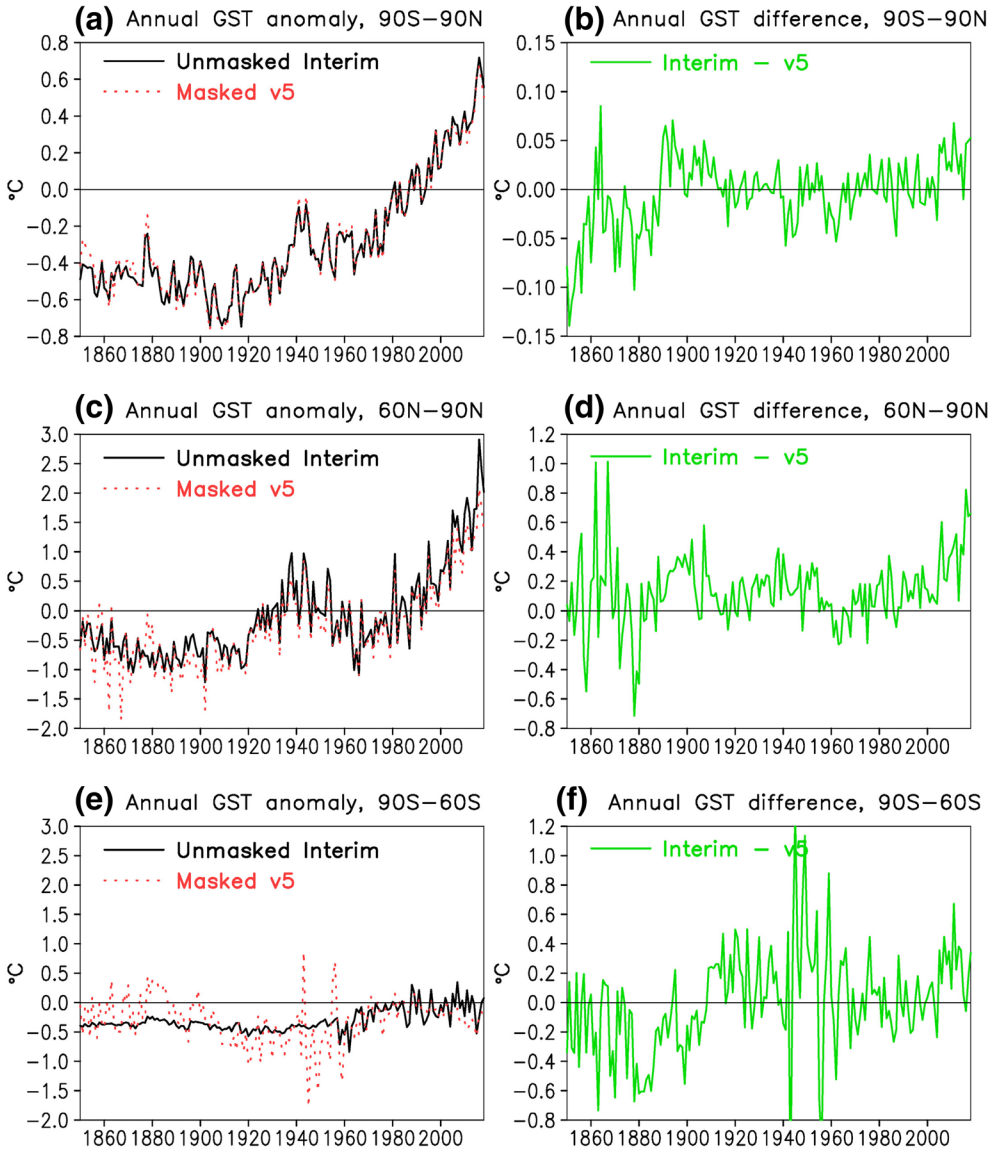

Figure 1. Annual anomaly time series (left column) and difference time series (right column): (a) 90°S–90°N anomalies, (c) 60°N–90°N anomalies, (e) 90°S–60°S anomalies, (b) 90°S–90°N differences, (d) 60°N–90°N differences, and (f) 90°S–60°S differences. Anomalies are relative to a 1971–2000 base period.

about 2% of the planetary surface). For high-frequency signals (i.e., EOTs), version 5 uses data for 1982–1991 from GHCNv2 (Peterson & Vose, 1997). In contrast, Interim employs data for 1981–2000 from the ERA-Interim reanalysis (Dee et al., 2011), which has complete global coverage and has demonstrated efficacy in depicting surface air temperature variations worldwide (Simmons et al., 2014), including the Arctic (Simmons & Poli, 2015).

To obtain full global coverage, we merged the Interim reconstruction with ERSSTv5. This was a relatively simple process: grid boxes over the land surface and the polar oceans came from the Interim reconstruction, grid boxes over the nonpolar oceans came from ERSSTv5, and grid boxes over coastal areas were the area-weighted average of the land and ocean anomalies. The resulting global product is spatially complete for the entire period (1850–2018). In contrast, version 5 includes a subsequent postprocessing step that masks out high-latitude grid-boxes (i.e., north of 75°N and south of 65°S) that have no underlying data in a given month or that have more than 50% sea ice coverage in a particular month. The implication of these

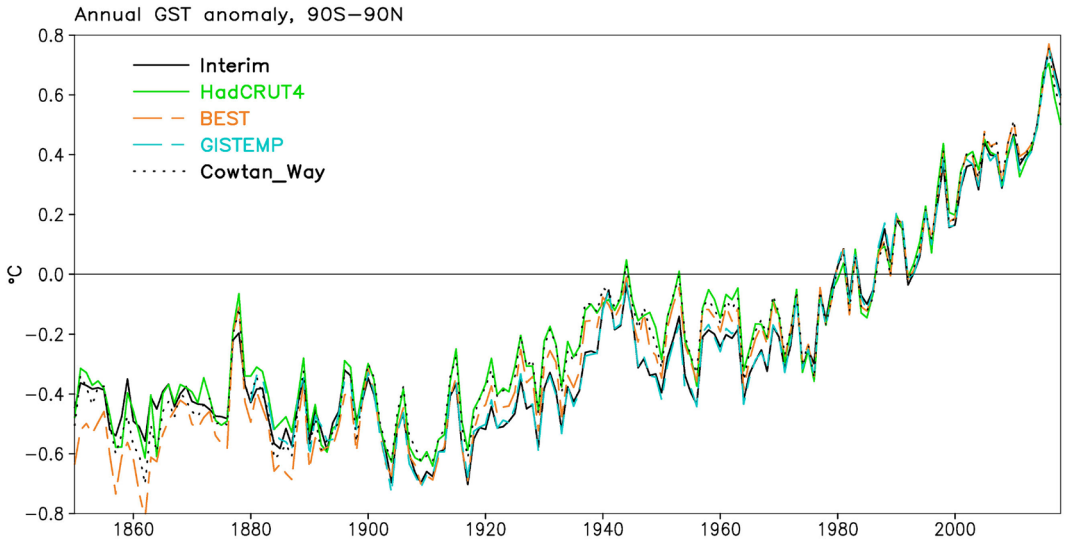

Figure 2. Annual anomaly time series for 90°S–90°N for various GST analyses: Interim, the Met Office Hadley Center (HadCRUT4; Morice et al., 2012); Berkeley Earth (BEST; Rohde et al., 2013); NASA (GISTEMP; Lenssen et al., 2019); and Cowtan and Way (2014).

masking criteria is that version 5 has very limited coverage in the Arctic throughout the record even though (1) many ship and buoy observations (see Section 2) are actually available within the region, and (2) many more observations are available from high-latitude land stations. The situation is similar in the Antarctic (i.e., version 5 has very limited coverage even though GHCNv4 has data for many land stations in Antarctica since the late 1950s). Version 5 also includes a postprocessing step that masks out grid boxes in data-sparse areas (i.e., grid boxes in which the sampling rate is less than 20% within a 25° latitude/longitude buffer; Vose, Arndt, et al., 2012). In general, this masking is more prevalent early in the record and occurs more frequently in areas such as the Amazon Basin, the Sahara Desert, the Congo Basin, and the Tibetan Plateau.

4. Temperature Trends

Version 5 and Interim have nearly identical global time series. The high degree of similarity is not surprising because the reconstructions use the same underlying data over most of the globe (i.e., GHCNv4 and ERSSTv5). Both series depict a decrease in temperature through the early 1900s, an increase through the early 1940s, a decrease through the early 1960s, and an increase through the end of the record (Figure 1a). Over the full period (1850–2018), the least squares trend in Interim is trivially larger than the trend in version 5 (0.052°C dec^{-1} vs. 0.049°C dec^{-1}). The series differ somewhat more in the 19th century because polar

Table 1
Annual Least-Squares Trends and 95% Confidence Intervals (°C Dec^{-1}) in Version 5 and Interim

Region	Analysis	1850–2018	1960–2018	1980–2018
90°S–90°N	Interim	0.052 ± 0.010	0.164 ± 0.018	0.179 ± 0.032
	Version 5	0.049 ± 0.010	0.155 ± 0.019	0.168 ± 0.032
60°N–90°N	Interim	0.109 ± 0.025	0.452 ± 0.078	0.598 ± 0.129
	Version 5	0.100 ± 0.022	0.358 ± 0.063	0.479 ± 0.111
90°S–60°S	Interim	0.021 ± 0.007	0.049 ± 0.033	0.007 ± 0.058
	Version 5	−0.003 ± 0.016	0.004 ± 0.036	−0.094 ± 0.035

Note. Confidence intervals computed with the student's t-test and an effective sampling number based on lag^{-1} autocorrelations (von Storch & Zwiers, 1999).

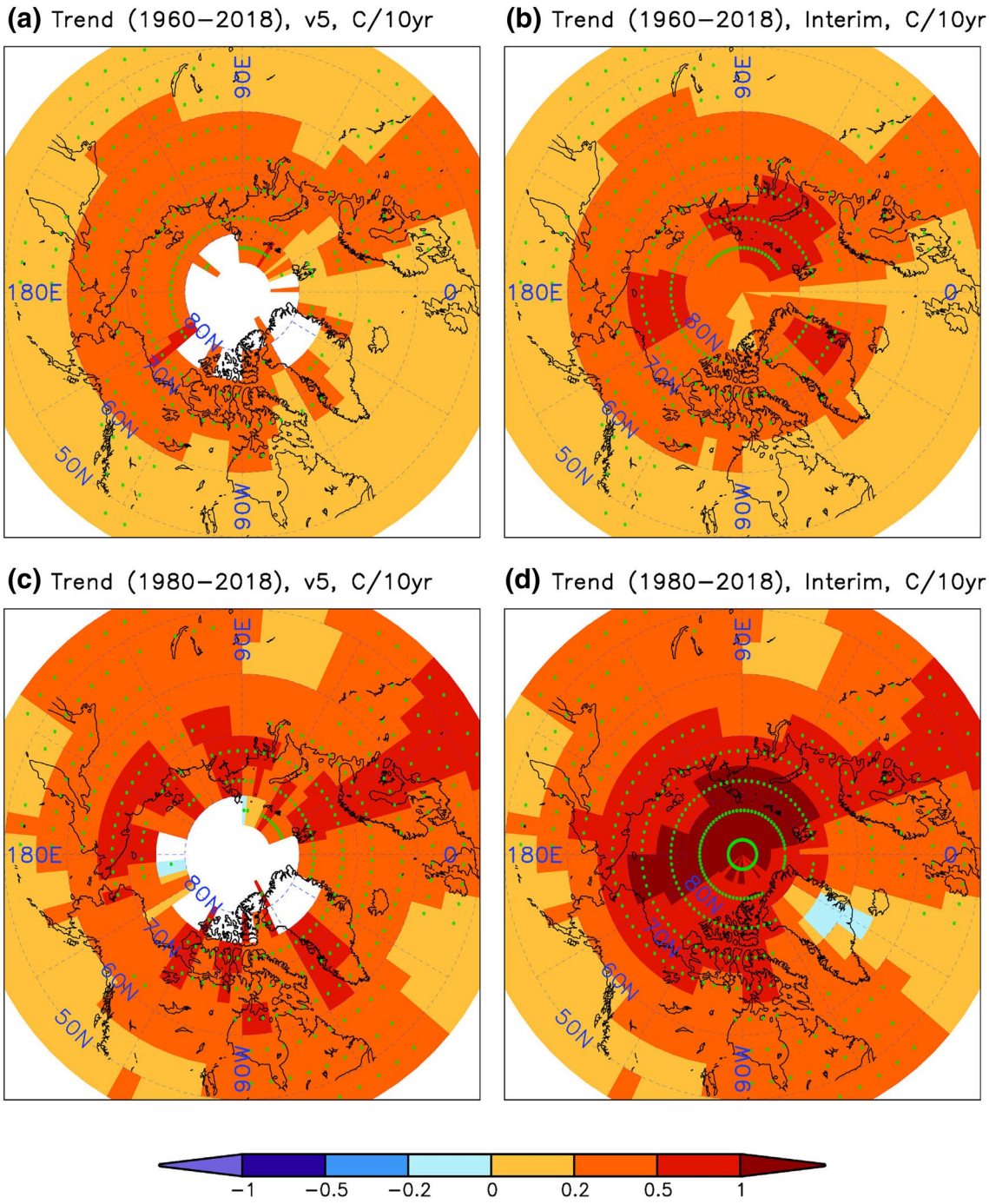

Figure 3. Grid-box trends (°C dec^{-1}) in the northern high latitudes: (a) 1960–2018 from version 5, (b) 1960–2018 from Interim, (c) 1980–2018 from version 5, and (d) 1980–2018 from Interim. Grid-box trends smoothed with a 9-point local filter. Stippling denotes 95% significance.

regions and data-sparse areas are masked out in version 5 whereas Interim has full coverage. There is suggestive evidence that Interim is slightly warmer in the last 2 decades (Figure 1b). Likewise, the global trend in Interim slightly exceeds the trend in version 5 from 1960 to 2018 (0.164°C dec^{-1} vs. 0.155°C dec^{-1}) and 1980–2018 (0.179°C dec^{-1} vs. 0.168°C dec^{-1}); however, each Interim trend falls within the 95% confidence interval of its version 5 counterpart (Table 1).

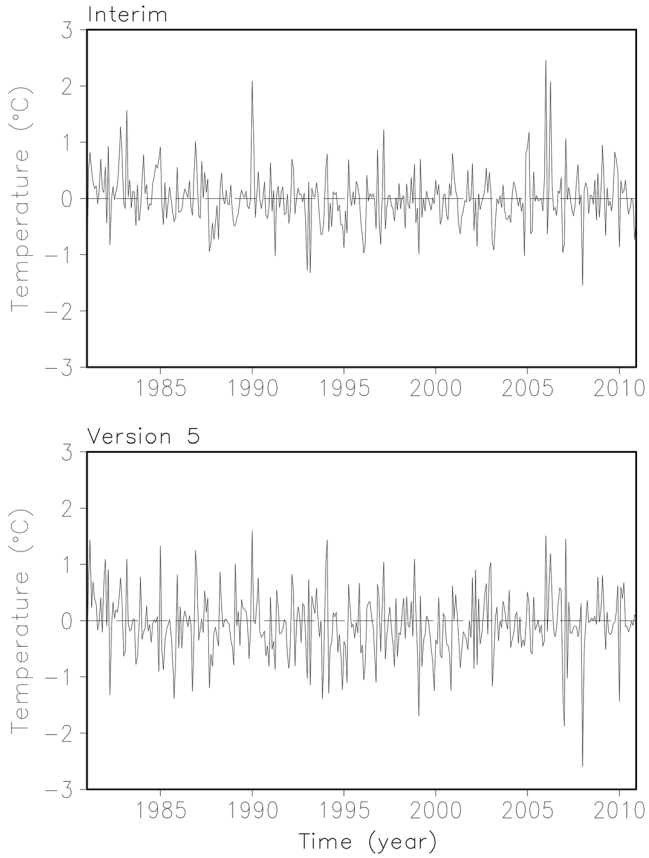

Figure 4. Monthly anomaly time series for 70°N–90°N derived from ERA-Interim. The top panel depicts anomalies derived using the coverage of Interim (i.e., the data set described in Section 2). The bottom panel depicts anomalies derived using the coverage of version 5 (i.e., GHCNv4 and ERSSTv5).

On the global scale, Interim is consistent with other GST analyses derived from in situ data (Figure 2). From 1970 to present, Interim is largely indistinguishable from the other analyses, which collectively exhibit extremely good agreement (with the occasional exception, such as HadCRUT4 being a bit cooler in some recent years owing to less Arctic coverage). From 1910 to 1970, Interim and GISTEMP are nearly identical whereas the others are slightly warmer. This reflects the underlying SST reconstruction employed in each case; Interim and GISTEMP use ERSSTv5 whereas the others use HadSST (Kennedy et al., 2011a, 2011b), which is slightly warmer during that period (Huang et al., 2017).

Differences between version 5 and Interim are more apparent in the high latitudes of the Northern Hemisphere (60°N–90°N). For example, the reconstruction process generated 65 EOTs in Interim (vs. 60 in version 5); the additional EOTs all capture Arctic variations. (Similarly, ERSSTv5 contains 10 more Arctic EOTs than its predecessor, which slightly reduced root mean square error according to several cross-validation experiments [Huang et al., 2017].) In general, both version 5 and Interim are quasi-stable until about 1920, then exhibit an increase until the late 1930s, a decrease until the early 1960s, and an increase thereafter (Figure 1c). Over the full period (1850–2018), the trend in Interim is slightly larger than the trend in version 5 (0.109°C dec^{-1} vs. 0.100°C dec^{-1}). Version 5 exhibits considerable interannual variability in the 19th century because most areas north of 75°N are masked out (i.e., its average is based on a smaller portion of the Arctic than Interim, which is spatially complete). Thereafter, Interim is generally about 0.1°C–0.2°C warmer in most years, with growing differences in recent decades (Figure 1d). The trend in Interim exceeds the trend in version 5 from 1960 to 2018 (0.452°C dec^{-1} vs. 0.358°C dec^{-1}) and 1980–2018 (0.598°C dec^{-1} vs. 0.479°C dec^{-1}); furthermore, each Interim trend falls outside the 95% confidence interval of its version 5 counterpart (Table 1). (For additional details on the Arctic time series, including a discussion of the anomaly patterns behind Interim's warmth relative to version 5, see the Supporting Information.)

We performed two analyses to evaluate the importance of sampling error in the northern high latitudes. To estimate the impact on recent trends, we applied the version 5 masking criteria to Interim (i.e., by removing grid boxes north of 75°N that had no underlying data). This reduces the trend in Interim by 0.048°C dec^{-1} since 1960 and 0.108°C dec^{-1} since 1980; in other words, discrepancies in spatial sampling account for 50% of the trend difference since 1960 and 90% since 1980, indicating that smaller trends in version 5 mainly result from its incomplete coverage of areas with rapid warming in recent decades (Figure 3). To estimate the impact of spatial sampling on area-average anomalies, we resampled the ERA-Interim reanalysis according to the actual year-by-year grid-box coverage of version 5 (i.e., GHCNv4 and ERSSTv5) and Interim (i.e., the data set described in Section 2). For 1981–2010, the coverage of version 5 results in an average cool bias of -0.08°C for 70°N–90°N and −0.13°C for 75°N–90°N whereas the Interim coverage has an average bias of zero in both latitude bands. The coverage of version 5 also results in a time series for 70°N–90°N (Figure 4) with higher variability (e.g., standard deviation of 0.55°C vs. 0.48°C).

Time series for the high latitudes of the Southern Hemisphere (90°S–60°S) exhibit moderate agreement after the establishment of permanent observing stations across Antarctica in the late 1950s. Both version 5 and Interim generally depict a slight warming from about 1960 to 1980; thereafter, version 5 exhibits cooling whereas Interim is quasi-stable (Figure 1e). The trend in Interim exceeds the trend in version 5 from 1960 to 2018 (0.049 vs. 0.004°C dec^{-1}) and 1980–2018 (0.007 vs.−0.094°C dec^{-1}), and each Interim trend falls outside the 95% confidence interval of its version 5 counterpart (Table 1). To quantify the impact of spatial sampling, we applied the version 5 masking criteria to Interim (i.e., by removing grid boxes south of 65°S that had no underlying data). This reduces the trend in Interim by 0.043°C dec^{-1} since 1960 and 0.099°C

dec^{-1} since 1980; in other words, the version 5 masking criteria account for more than 80% of the trend difference in both periods. (For additional details on the Antarctic time series, including a discussion of the large differences before the late 1950s, see the Supporting Information.)

5. Summary and Future Work

This paper introduced a spatially complete edition of NOAAGlobalTemp (termed Interim), which includes three improvements over its operational counterpart (version 5): the use of air temperature data in the Arctic Ocean, the application of climate reanalysis fields in spatial interpolation, and the elimination of geographic masking criteria from reconstructed temperatures. Both versions show warming across the globe and the Arctic for all periods examined. Global time series are nearly indistinguishable during most of the record, but discrepancies exist in the Arctic in recent decades, Interim having greater warming because it captures ongoing changes throughout that region. Time series for the southern high latitudes show moderate agreement since the late 1950s. Looking toward the future, Interim could be enhanced through a formal analysis of parametric and reconstruction uncertainties as in Huang et al. (2020).

Data Availability Statement

Data supporting the conclusions are available from https://www.ncei.noaa.gov/pub/data/cmb/ersst/v5/2020.grl.dat/.

Acknowledgments
This work was partially supported by the Climate Observations and Monitoring Program of the NOAA Climate Program Office.

References

Argo. (2000). Argo float data and metadata from Global Data Assembly Centre. *SEANOE*, https://doi.org/10.17882/42182

Comiso, J. C., & Hall, D. K. (2014). Climate-trends in the Arctic as observed from space. *WIREs Climate Change*, *5*(3), 389–409. https://doi.org/10.1002/wcc.277

Cowtan, K., & Way, R. G. (2014). Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Quarterly Journal of the Royal Meteorological Society*, *140*(683), 1935–1944. https://doi.org/10.1002/qj.2297

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, U., et al. (2011). The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Durre, I., Menne, M. J., & Vose, R. S. (2008). Strategies for evaluating quality assurance procedures. *Journal of Applied Meteorology and Climatology*, *47*(6), 1785–1791. https://doi.org/10.1175/2007JAMC1706.1

Freeman, E., Woodruff, S. D., Worley, S. J., Lubker, S. J., Kent, E. C., Angel, W. E., et al. (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

Gelaro, R., McCarty, W., Saurez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era retrospective analysis for research and applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global surface temperature change. *Reviews of Geophysics*, *48*, RG4004. https://doi.org/10.1029/2010RG000345

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horanyi, A., Munoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*. *146*(730), 1–51. https://doi.org/10.1002/qj.3803

Hofer, M., Marzeion, B., & Molg, T. (2012). Comparing the skill of different reanalyzes and their ensembles as predictors for daily air temperature on a glaciated mountain (Peru). *Climate Dynamics*, *39*, 1969–1980. https://doi.org/10.1007/s00382-012-1501-2

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., et al. (2017). Extended reconstructed sea surface temperature version 5 (ERSSTv5), Upgrades, validations, and intercomparisons. *Journal of Climate*, *30*(20), 8179–8205. https://doi.org/10.1175/JCLI-D-16-0836.1

Huang, B., Menne, M. J., Boyer, T., Freeman, E., Gleason, B. E., Lawrimore, J. H., Liu, C., et al. (2020). Uncertainty estimates for sea surface temperature and land surface air temperature in NOAA GlobalTemp version 5. *Journal of Climate*, *33*, 1351–1379. https://doi.org/10.1175/JCLI-D-19-0395.1

Karl, T. R., Arguez, A., Huang, B., Lawrimore, J. H., McMahon, J. R., Menne, M. J., et al. (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

Kennedy, J. J., Rayner, N. A., Smith, R. O., Saunby, M., & Parker, D. E. (2011a). Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 1: Measurement and sampling errors. *Journal of Geophysical Research*, *116*, D14103. https://doi.org/10.1029/2010JD015218

Kennedy, J. J., Rayner, N. A., Smith, R. O., Saunby, M., & Parker, D. E. (2011b). Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 2: Biases and homogenization. *Journal of Geophysical Research*, *116*, D14104. https://doi.org/10.1029/2010JD015220

Kobayashi, S., Yukinari, O. T. A., Harada, Y., Ebita, A., Moriya, M., Onada, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, *93*(1), 5–48. https://doi.org/10.2151/jmsj.2015-001

Lawrimore, J. H., Menne, M. J., Gleason, B. E., Williams, C. N., Wuertz, D. B., Vose, R. S., & Rennie, J. (2011). An overview of the global historical climatology network monthly mean temperature data set, version 3. *Journal of Geophysical Research*, *116*, D19121. https://doi.org/10.1029/2011JD016187

Lenssen, N. J. L., Schmidt, G. A., Hansen, J. E., Menne, M. J., Persin, A., Ruedy, R., & Zyss, D. (2019). Improvements in the GISTEMP uncertainty model. *Journal of Geophysical Research: Atmospheres*, *124*, 6307–6326. https://doi.org/10.1029/2018JD029522

Martin, S., & Munoz, E. A. (1997). Properties of the Arctic 2-meter air temperature field for 1979 to the present derived from a new gridded dataset. *Journal of Climate*, *10*(2), 1428–1440. https://doi.org/10.1175/1520-0442(1997)010<1428:POTAMA>2.0.CO

Menne, M. J., & Williams, C. N. (2009). Homogenization of temperature series via pairwise comparisons. *Journal of Climate*, *22*, 1700–1717. https://doi.org/10.1175/2008JCLI2263.1

Menne, M. J., Williams, C. N., Gleason, B. E., Rennie, J. J., & Lawrimore, J. H. (2018). The Global Historical Climatology Network monthly temperature dataset, version 4. *Journal of Climate*, *31*(24), 9835–9854. https://doi.org/10.1175/JCLI-D-18-0094.1

Morice, C. P., Kennedy, J. J., Rayner, N. A., & Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research*, *117*(D8), D08101. https://doi.org/10.1029/2011JD017187

Peterson, T. C., & Vose, R. S. (1997). An overview of the global historical climatology network temperature database. *Bulletin of the American Meteorological Society*, *78*, 2837–2849.

Rigor, I. G., Colony, R. L., & Martin, S. (2000). Variations in surface air temperature observations in the Arctic, 1979–97. *Journal of Climate*, *13*, 896–914. https://doi.org/10.1175/1520-0442(2000)013<0896:VISATO>2.0.CO

Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S., Rosenfeld, A., & Wurtele, J. (2013). Berkeley Earth temperature averaging process. *Geoinformatics Geostatistics: An Overview*, *1*. https://doi.org/10.4172/gigs.1000103

Romanov, I. P., Konstantinov, Y. B., & Kornilov, N. A. (1997). *Drifting ice stations "North Pole" (1937-1991)* (p. 225). St. Peterburg, FL: Gidrometeoizdat.

Saha, S., Moorthi, S., Pan, H., Wu, X., Wang, J., Nadiga, S., et al. (2014). The NCEP Climate Forecast System Version 2. *Journal of Climate*, *27*(6), 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1

Simmons, A. J., & Poli, P. (2015). Arctic warming in ERA-Interim and other analyses. *Quarterly Journal of the Royal Meteorological Society*, *141*(689), 1147–1162. https://doi.org/10.1002/qj.2422

Simmons, A. J., Poli, P., Dee, D. P., Berrisford, P., Hersbach, H., Kobayashi, S., & Peubey, C. (2014). Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim. *Quarterly Journal of the Royal Meteorological Society*, *140*, 329–353. https://doi.org/10.1002/qj.2317

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., et al. (2019). Toward a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, *145*(784), 2876–2908. https://doi.org/10.1002/qj.3598

Smith, T. M., & Reynolds, R. W. (2005). A global merged land–air–sea surface temperature reconstruction based on historical observations (1880–1997). *Journal of Climate*, *18*(12), 2021–2036. https://doi.org/10.1175/JCLI3362.1

Smith, T. M., Reynolds, R. W., Peterson, T. C., & Lawrimore, J. (2008). Improvements to NOAA's historical merged land–ocean surface temperatures analysis (1880–2006). *Journal of Climate*, *21*, 2283–2296. https://doi.org/10.1175/2007JCLI2100.1

Van den Dool, H. M., Saha, S., & Johansson, Å. (2000). Empirical orthogonal teleconnections. *Journal of Climate*, *13*, 1421–1435.

von Storch, H., & Zwiers, F. W. (1999). *Statistical analysis in climate research* (p. 484).Cambridge, UK: Cambridge University Press.

Vose, R. S., Arndt, D., Banzon, V. F., Easterling, D. R., Gleason, B., Huang, B., et al. (2012). NOAA's merged land-ocean surface temperature analysis. *Bulletin of the American Meteorological Society*, *93*, 1677–1685. https://doi.org/10.1175/BAMS-D-11-00241.1

Vose, R. S., Applequist, S., Menne, M. J., Williams, C. N., & Thorne, P. (2012). An intercomparison of temperature trends in the U.S. historical climatology network and recent atmospheric reanalysis. *Geophysical Research Letters*, *39*, L10703. https://doi.org/10.1029/2012GL051387

Zhang, H.-M., Lawrimore, J. H., Huang, B., Menne, M. J., Yin, X., Sánchez-Lugo, A., et al. (2019). Updated temperature data give a sharper view of climate trends. *Eos*, *100*. https://doi.org/10.1029/2019EO128229

References From the Supporting Information

Jones, P. D. (1990). Antarctic temperatures over the present century—A study of the early expedition record. *Journal of Climate*, *3*, 1193–1203.

Marshall, J., Armour, K. C., Scott, J. R., Kostov, Y., Hausmann, U., Ferreira, D., et al. (2014). The ocean's role in polar climate change: Asymmetric Arctic and Antarctic responses to greenhouse gas and ozone forcing. *Philosophical Transactions of the Royal Society A*, 372., 20130040. https://doi.org/10.1098/rsta.20130040

Nicolas, J. P., & Bromwich, D. H. (2014). New reconstructions of Antarctic near-surface temperatures: Multidecadal trends and reliability of global reanalyzes. *Journal of Climate*, *27*, 8070–8093. https://doi.org/10.1175/JCLI-D-13-00733.1

Serreze, M. C., & Barry, R. C. (2011). Processes and impacts of Arctic amplification: A research synthesis. *Global and Planetary Change*, *77*, 85–96. https://doi.org/10.1016/j.gloplacha.2011.03.004