Supplement of Earth Syst. Sci. Data, 17, 7079–7100, 2025
https://doi.org/10.5194/essd-17-7079-2025-supplement

Open Access Earth System Science Data

Supplement of

An observational record of global gridded near-surface air temperature change over land and ocean from 1781

Colin P. Morice et al.

Correspondence to: Colin P. Morice (colin.morice@metoffice.gov.uk)

S1 Data coverage without statistical infilling

Data coverage of the non-infilled grids of the GloSATref data set is shown in Fig. S1. For comparison, data coverage for the non-infilled grids of the HadCRUT5 data set (Morice et al., 2021) is shown in Fig. S2. These figures provide a comparison of data coverage for the underpinning observational data that contributes to each data set. Differences in data coverage between these two data sets result from (1) the use of marine air temperature observations in GloSATref and sea-surface temperature in HadCRUT5, (2) land data acquisitions in GloSATref that extend the CRUTEM5 station database and (3) the capability to grid additional land station series in GloSATref through the use of new station climatology estimates. Further discussion of data coverage is provided in the main article.

S2 Data coverage of the GloSATref analysis

Data coverage of the GloSATref statistically infilled analysis fields is shown in Fig. S3. For comparison, data coverage for the HadCRUT5 analysis (Morice et al., 2021) is shown in Fig. S4. The two data sets share common statistical analysis methods but differ in observation data used.

Differences in data coverage between the GloSATref and HadCRUT5 analyses result from (1) difference in underpinning observation data coverage described in S1 and (2) differences in fitted statistical model hyperparameters. The marine air temperature analysis parameters of GloSATref result in interpolation over shorter spatial length scales (marine length scale parameter ρ = 1300 km) than the sea-surface temperature analysis of HadCRUT5 (marine length scale parameter ρ = 1550 km). Land surface air temperature (LSAT) analyses for the the two data sets share the same parameter estimate values (LSAT length scale parameter ρ = 1300 km), hence interpolation distances from available observations are broadly equivalent across land and sea ice regions in the two data sets.

S3 LSAT/MAT representation in GloSATref timeseries

The area represented by LSAT and MAT anomalies in the GloSATref analysis varies over time caused by changes in LSAT and MAT observation coverage and changes in sea ice coverage. For global and regional averages computed from the GloSATref analysis this results in varying relative weightings towards land and ocean averages throughout the observed record. The relative weightings towards LSAT and MAT analyses are shown in Fig. S5 for global and hemispheric average timeseries, along with comparisons to weighted averages of LSAT and MAT timeseries using a static weighting for each domain. These comparisons are illustrative of the effects of changing LSAT/MAT weighting only for regions that are represented in the analysis. They do no account for unrepresented regions (although estimates of expected errors from unrepresented regions are included GloSATref uncertainty estimates for global and regional averages through the coverage uncertainty model (Morice et al., 2021)]).

Comparison of global average temperatures indicate that changes in LSAT/MAT weighting have an effect of typically less than 0.1°C for the observed region prior to the 1900s, with individual years prior to the 1850s showing differences approaching 0.3°C. Differences between GloSATref merged analysis series and the fixed weight series are less than 0.02°C following the 1900s. In the LSAT analysis, data is largely unavailable in the southern hemisphere prior to the 1850s. Southern hemisphere series at this time are representative of MAT averages, with differences between the regular merged SAT series and the fixed weight series tending to zero and unable to provide information about LSAT/MAT sampling error. For both hemispheres, prior to the 1900s, periods of persistent difference between merged analysis and fixed weight series of 0.1 to 0.2°C are coincident with large changes in data availability. For the northern hemisphere form the start on the 1900s onward and for the southern hemisphere from the 1950s onward, differences indicative of the effects of changing LSAT and MAT weighting are less than 0.01°C.

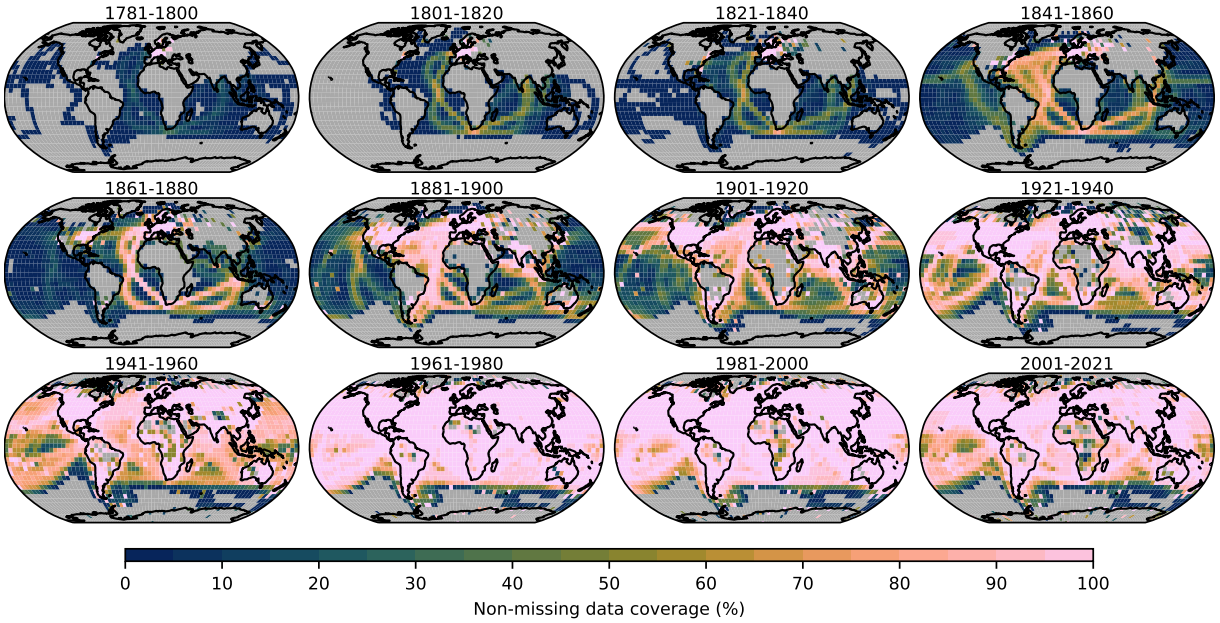

Figure S1. Data coverage for the unfilled GloSATref data set, showing the percentage of months with non-missing data for 20-year periods (final panel shows 21-year period). Grey shaded regions have no data in the corresponding period.

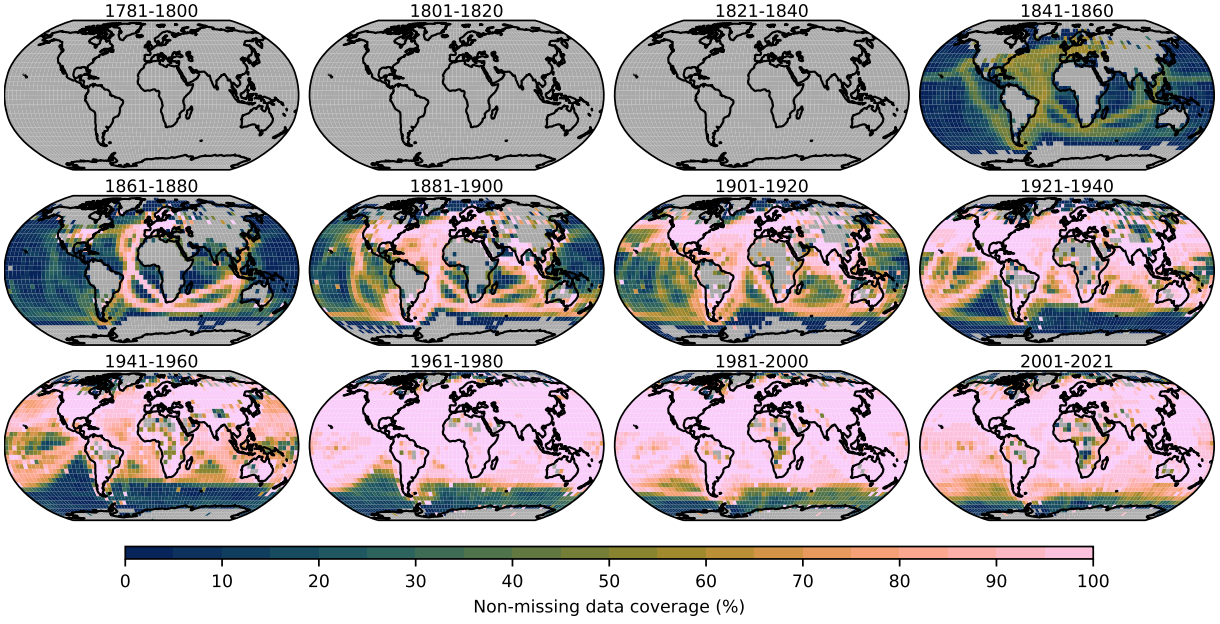

Figure S2. Data coverage in the unfilled HadCRUT5 data set for comparison to Fig. S1, showing the percentage of months with non-missing data for 20-year periods (final panel shows 21-year period). Note that the HadCRUT5 grids contain data from 1850 only.

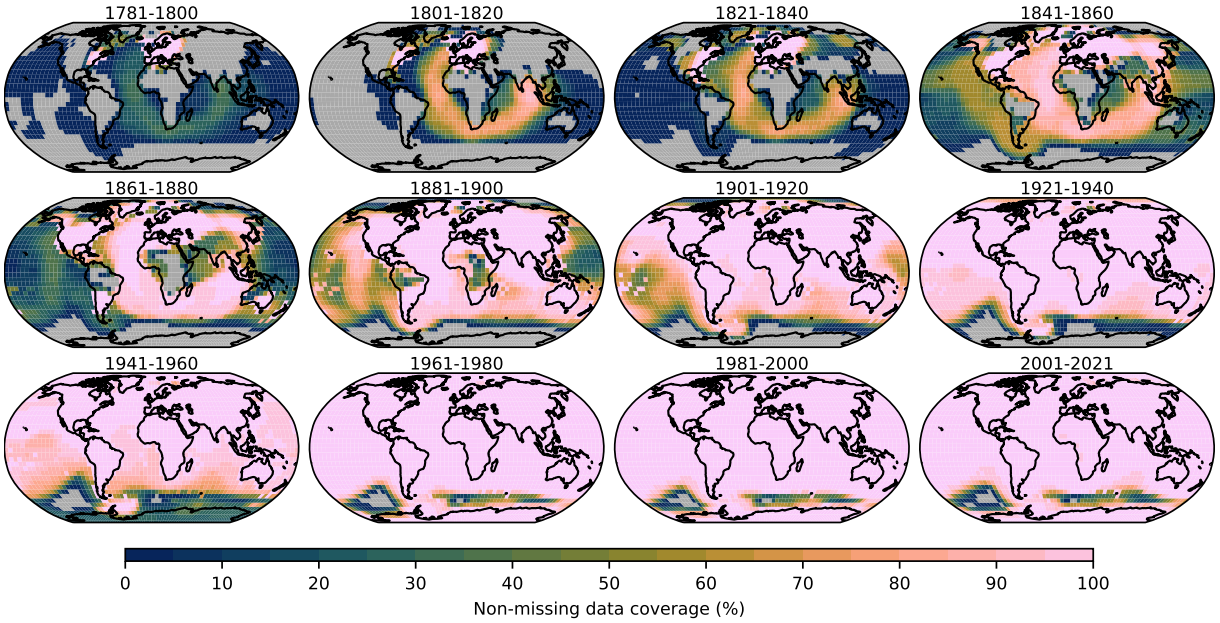

Figure S3. Data coverage for the GloSATref infilled analysis, showing the percentage of months with non-missing data for 20-year periods (final panel shows 21-year period). Grey shaded regions have no data in the corresponding period.

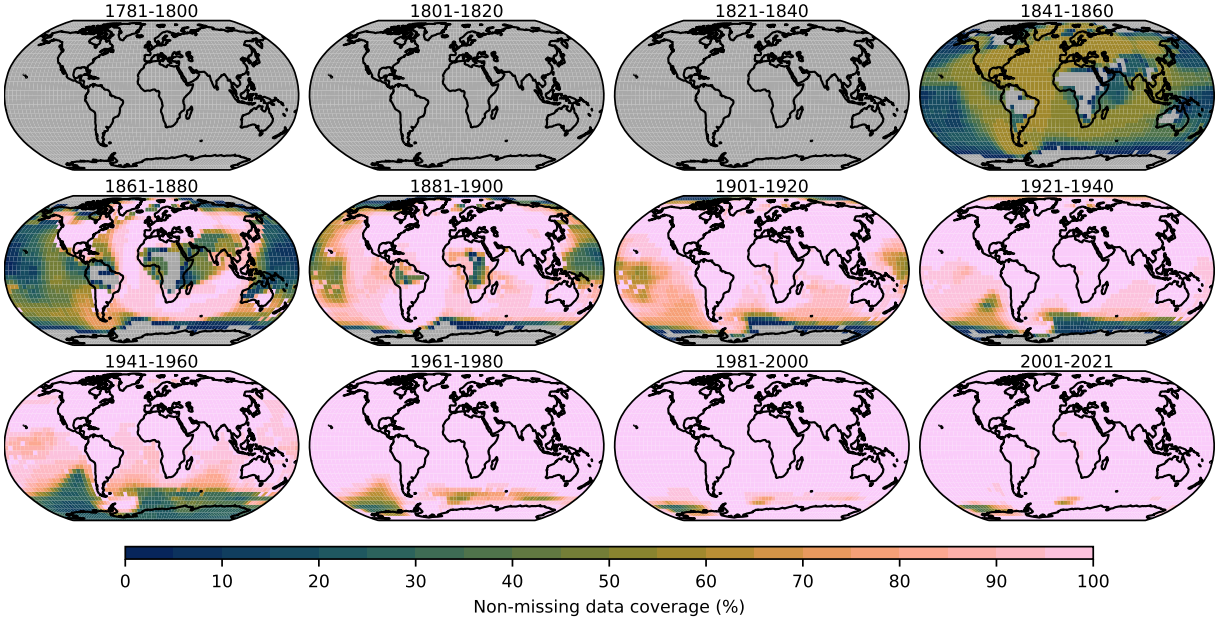

Figure S4. Data coverage in the HadCRUT5 infilled analysis for comparison to Fig. S1, showing the percentage of months with non-missing data for 20-year periods (final panel shows 21-year period). Note that the HadCRUT5 grids contain data from 1850 only.

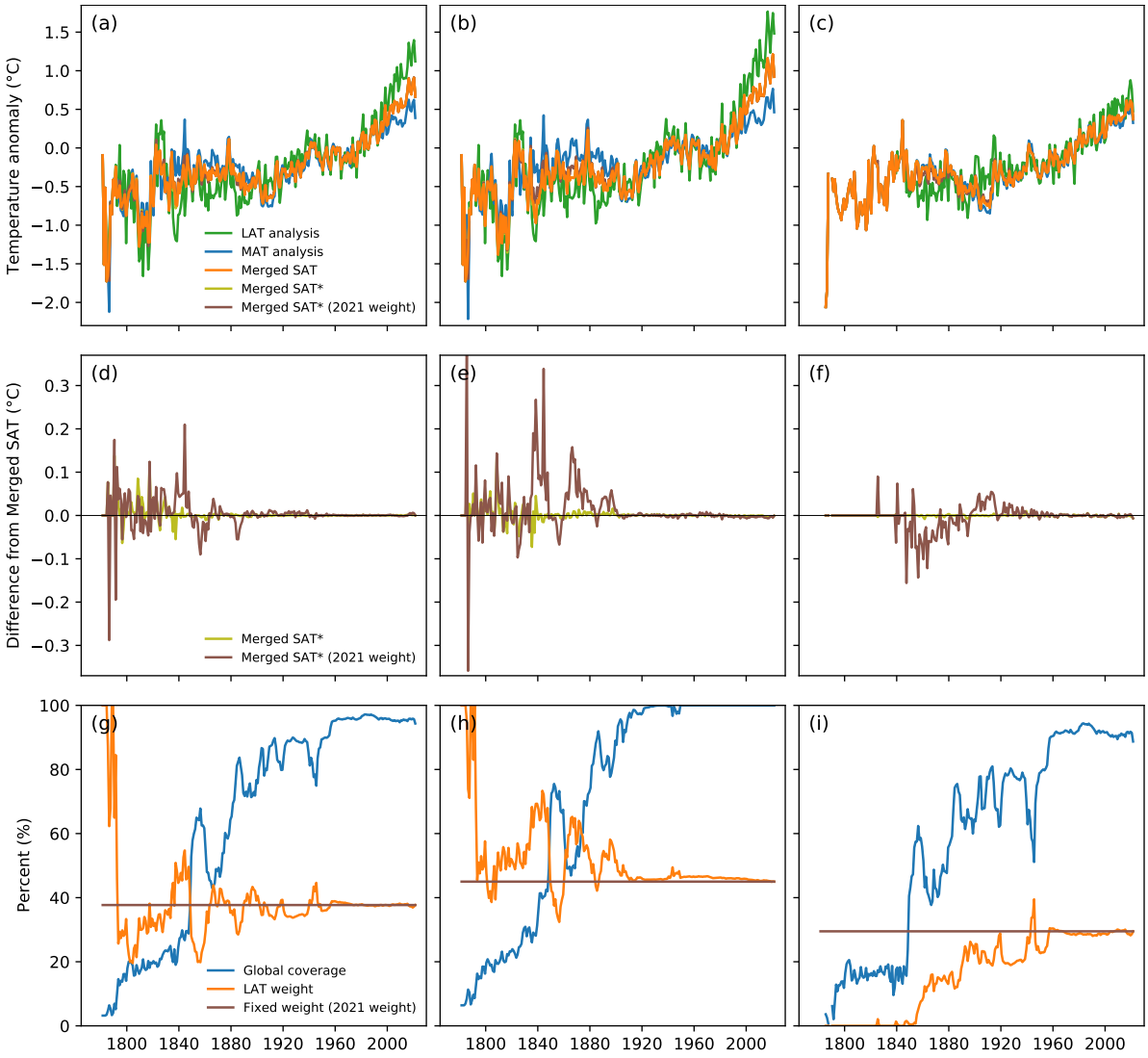

Figure S5. Global and hemispheric annual temperature anomaly time series demonstration the effects of the relative weighting of LSAT and MAT series in the merged analysis. Panels correspond to the following time series: left column (a,d,g), global averages without hemispheric weighting; middle column (b,e,h) northern hemisphere averages, and (c,f,i) southern hemisphere averages. Top row: temperature anomalies, relative to 1961-1990, for MAT and LSAT analysis (including sea ice regions), the merged GST analysis, and variations of the a merged GSAT using different weighting methods. Merged series named with an asterisk are computed as a weighted average of LSAT and MAT series, reverting to single domain analysis series if no data is available for the other domain. Merged SAT* (2021 weighting) series are constructed using a fixed LSAT/MAT weighting for every year applying the weighting from 2021. Merged SAT* applies the varying LSAT/MAT for each analysis year. Differences between Merged SAT and Merged SAT* series demonstrate differences in computing average anomaly series from the merged analysis grids or from a weighted average of SAT/MAT series. Middle row: Differences in SAT series from the Merged SAT series in the corresponding above panels. Bottom row: percentage of total area coverage by analysis grids and relative weighting of LSAT in the Merged SAT series.

S4 Comparisons to paleoclimate reconstructions

Global annual average temperature anomaly series are shown in Fig. S6 in comparison to the PAGES2k multi-method ensemble paleoclimate reconstructions (PAGES2k, 2019). The GloSATref ensemble mean shows greater variability than the PAGES2k median in the early instrumental record. Central estimates exceed the 2.5% and 97.5% confidence limits of each ensemble. Confidence intervals for the GloSATref and PAGES2k ensembles overlap for most years, while departures from confidence intervals occur around the times of the unidentified 1809 volcanic eruption and the 1815 eruption of Mount Tambora. The individual ensemble medians for each paleoclimate reconstruction method have a differing response to these early 19th century volcanoes, with weaker responses than is apparent in the GloSATref analysis global average time series. The GloSATref analysis does, however, have reduced global coverage in this period (see Fig S3), with greater variability resulting from regional sampling.

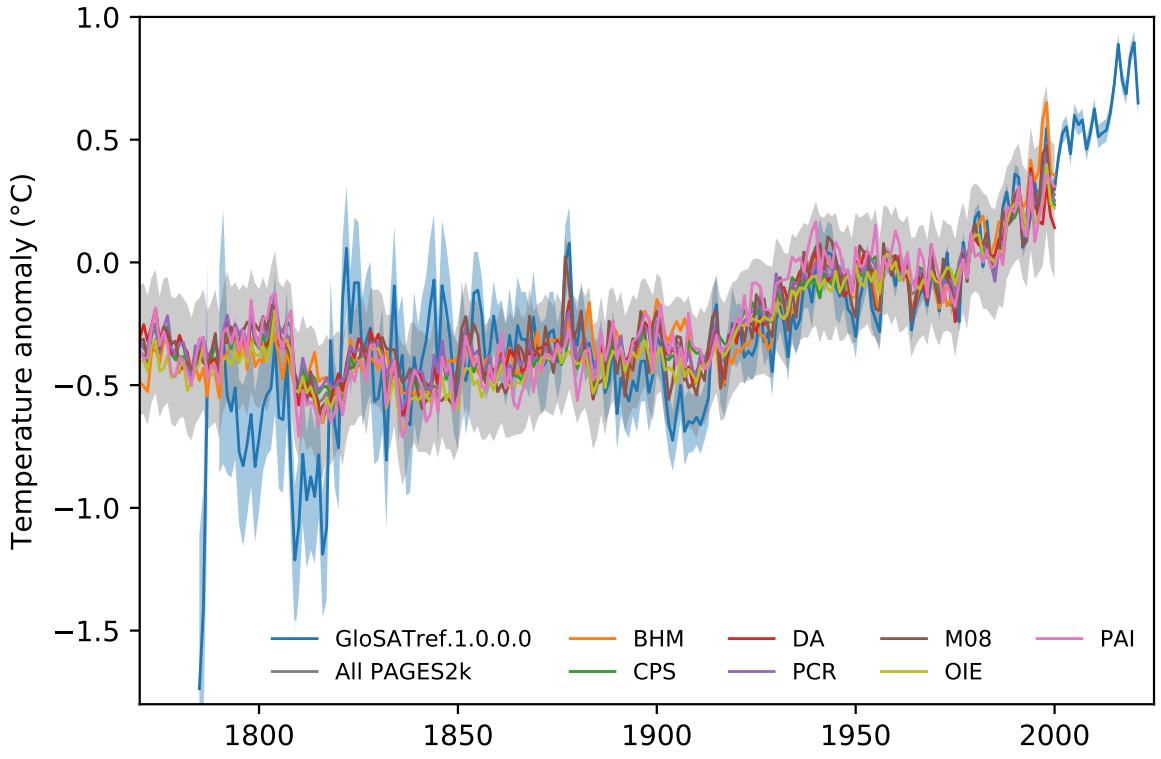

Figure S6. Annual global average temperature anomaly time series for the GloSATref analysis and PAGES2k paleoclimate proxy reconstructions (°C relative to 1961-1990). Uncertainties shown as 95% confidence intervals. Ensemble means for each PAGES2k reconstruction method shown: BHM, CPS, DA, PCR, M08, OIE and PAI.

S5 Extended LSAT and SST/NMAT/MAT comparisons

As in the main article Fig. 4, global average temperature anomaly series over land and marine domains are shown in Fig. S7 and S8, here with additional data sets included that are less closely related to the constituent data sets of the GloSATref analysis. The spread of LSAT data sets is typically greater than the difference between CRUTEM5 and GloSATLAT, likely

due to the greater range of input data and processing methods adopted. The Berkeley Earth land data set extends back to 1750 (Rohde et al., 2013a, b), allowing comparison with early record GloSATLAT. The two data sets show similar decadal variability albeit with substantially greater differences between time series prior to the mid-19th century than following. Berkeley Earth uses spatial interpolation while GloSATLAT does not. This contributes to differences between times series for the two data sets. However, as data coverage prior to the mid-19th century is largely confined to Europe and north America, the two data sets are subject to similar regional sampling errors at this time. SST/NMAT/MAT comparisons indicate cooler anomalies in SST data sets in the late 19th century then in GloSATMAT. Early 20th century NMAT/MAT series are cooler than SST, particularly so in comparison to DCSST and COBE-SST3, with recent studies indicating that prexisting SST data sets may be biased cold in this period (Sippel et al., 2024; Chan et al., 2024). The SST data sets exhibit greater warming from SST data sets from the mid-20th century in comparison to CLASSNMAT v1/v2 and GloSATMAT data sets.

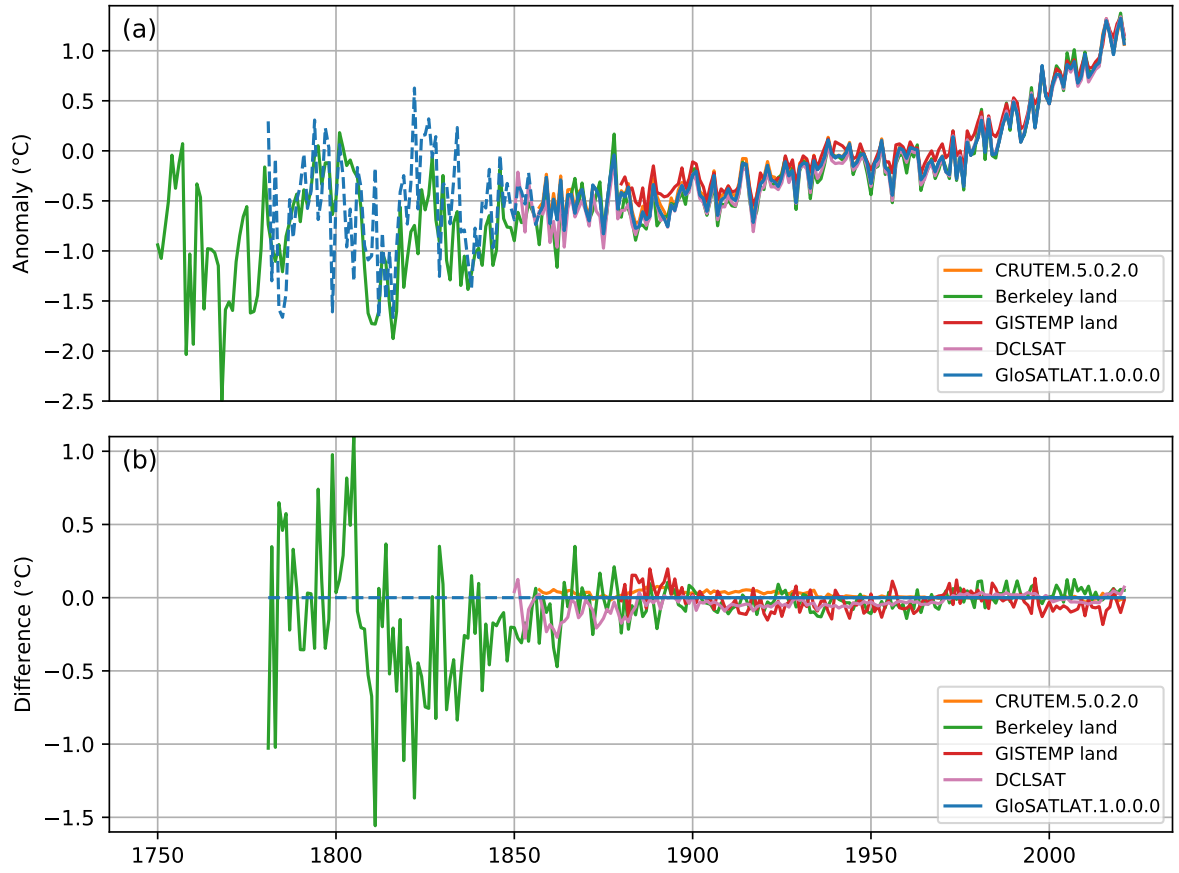

Figure S7. (a) Global average LSAT anomalies (°C) relative to 1961-1990 and (b) differences between each global average LSAT anomaly time series and that of GloSATLAT. The dashed line for GloSATLAT extends the series prior to 1856 by omitting the requirement for at least 5 populated grid cells in each hemisphere, with only Northern Hemisphere station data present in GloSATLAT grids prior 1854. GISTEMP and Berkeley land data sets use spatial interpolation while other data sets do not. DCLSAT series are computed from the gridded fields applying the CRUTEM5 hemispheric weighting scheme of 3/2 northern hemisphere to 1/3 southern hemisphere weighting. Data are not co-located to a common data coverage across data sets.

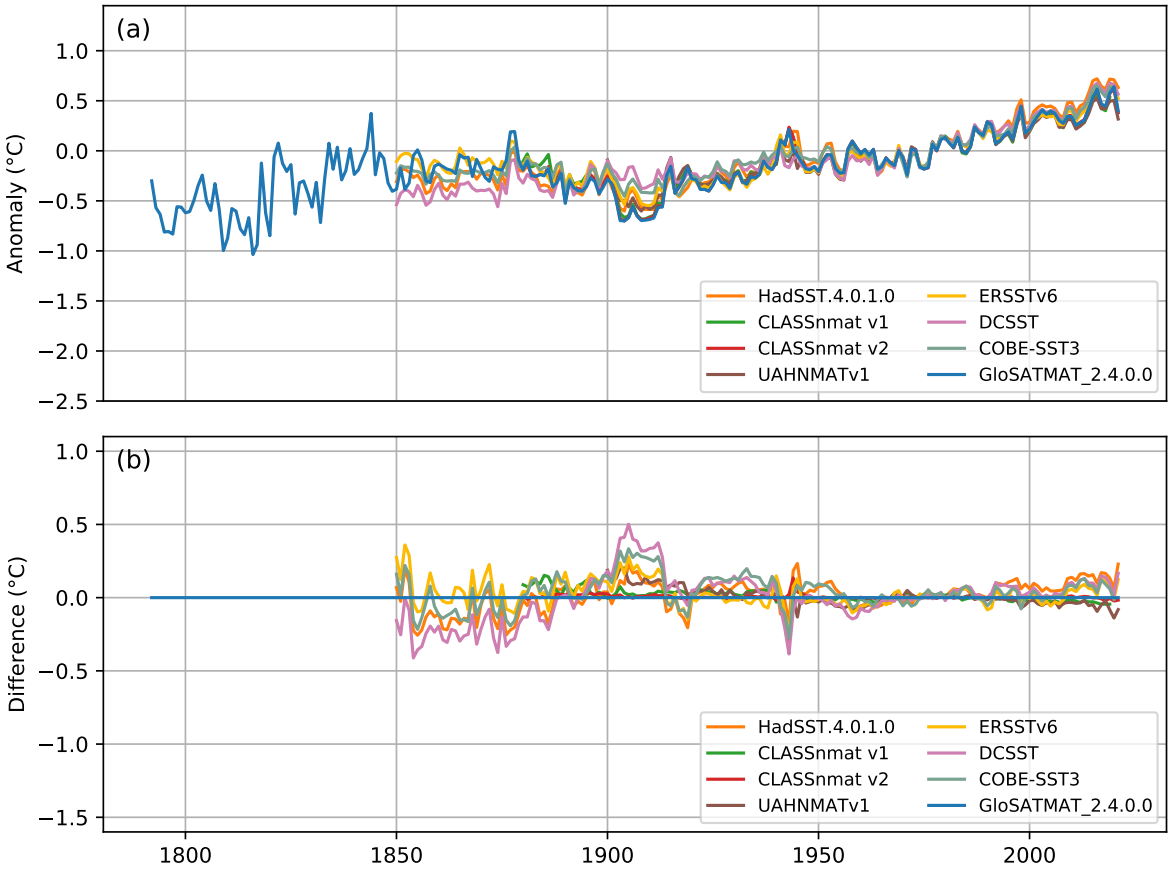

Figure S8. (a) Global average SST/NMAT/MAT anomalies (°C) relative to 1961-1990 and (b) differences between each global average SST/NMAT/MAT anomaly time series and that of GloSATMAT. The plotted data sets do not use spatial interpolation. DCSST series are computed with an area weighted average of the DCSST gridded fields. Data are not co-located to a common data coverage across data sets.

References

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

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J. H., Osborn, T. J., Jones, P. D., and Simpson, I. R.: An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set, Journal of Geophysical Research: Atmospheres, 126, https://doi.org/10.1029/2019jd032361, 2021.

PAGES2k: Consistent multidecadal variability in global temperature reconstructions and simulations over the Common Era, Nature Geoscience, 12, 643–649, https://doi.org/10.1038/s41561-019-0400-0, 2019.

Rohde, R., Muller, R., , Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., Wurtel, J., Donald, G., and Wickham, C.: A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011, Geoinformatics & Geostatistics: An Overview, 01, https://doi.org/10.4172/2327-4581.1000101, 2013a.

Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S., and Mosher, S.: Berkeley Earth Temperature Averaging Process, Geoinformatics & Geostatistics: An Overview, 01, https://doi.org/10.4172/2327-4581.1000103, 2013b.

Sippel, S., Kent, E. C., Meinshausen, N., Chan, D., Kadow, C., Neukom, R., Fischer, E. M., Humphrey, V., Rohde, R., de Vries, I., and Knutti, R.: Early-twentieth-century cold bias in ocean surface temperature observations, Nature, 635, 618–624, https://doi.org/10.1038/s41586-024-08230-1, 2024.