

Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century

N. A. Rayner, D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, and D. P. Rowell
Hadley Centre for Climate Prediction and Research, Met Office, Bracknell, UK

E. C. Kent

James Rennell Division, Southampton Oceanography Centre, Southampton, UK

A. Kaplan

Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York, USA

Received 19 June 2002; revised 30 October 2002; accepted 11 March 2003; published 17 July 2003.

[1] We present the Met Office Hadley Centre's sea ice and sea surface temperature (SST) data set, HadISST1, and the nighttime marine air temperature (NMAT) data set, HadMAT1. HadISST1 replaces the global sea ice and sea surface temperature (GISST) data sets and is a unique combination of monthly globally complete fields of SST and sea ice concentration on a 1° latitude-longitude grid from 1871. The companion HadMAT1 runs monthly from 1856 on a 5° latitude-longitude grid and incorporates new corrections for the effect on NMAT of increasing deck (and hence measurement) heights. HadISST1 and HadMAT1 temperatures are reconstructed using a two-stage reduced-space optimal interpolation procedure, followed by superposition of quality-improved gridded observations onto the reconstructions to restore local detail. The sea ice fields are made more homogeneous by compensating satellite microwave-based sea ice concentrations for the impact of surface melt effects on retrievals in the Arctic and for algorithm deficiencies in the Antarctic and by making the historical in situ concentrations consistent with the satellite data. SSTs near sea ice are estimated using statistical relationships between SST and sea ice concentration. HadISST1 compares well with other published analyses, capturing trends in global, hemispheric, and regional SST well, containing SST fields with more uniform variance through time and better month-to-month persistence than those in GISST. HadMAT1 is more consistent with SST and with collocated land surface air temperatures than previous NMAT data sets. *INDEX TERMS:* 1610 Global Change: Atmosphere (0315, 0325); 1635 Global Change: Oceans (4203); 1827 Hydrology: Glaciology (1863); 3339 Meteorology and Atmospheric Dynamics: Ocean/atmosphere interactions (0312, 4504); *KEYWORDS:* sea surface temperature, sea ice, night marine air temperature, climate data reconstruction, bias correction, climate change

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

1. Introduction

[2] Much of this paper is devoted to the description and assessment of the Hadley Centre sea ice and sea surface temperature (SST) data set version 1 (HadISST1), which was developed at the Met Office Hadley Centre for Climate Prediction and Research. HadISST1 improves upon previous global sea ice and SST (GISST) data sets: GISST1 [Parker *et al.*, 1995a], GISST2 [Rayner *et al.*, 1996], and GISST3, all developed at the Hadley Centre. We give a full account of these improvements, and provide a range of diagnostics to assess HadISST1. The primary purpose of HadISST1 is to force atmospheric models

(AGCMs) in the simulation of recent climate and to evaluate coupled atmosphere-ocean models, thereby improving our understanding of natural and human-induced climatic variations and allowing evaluation of model performance. HadISST1 has also been used to supply information for the ocean surface for the period 1958 through 1981 in the 40-year ECMWF Reanalysis (ERA40), with 2DVAR and OI.v2 [Reynolds *et al.*, 2002] used thereafter. To fulfill these aims, HadISST1 has been made globally complete. Gaps in the SST data have been interpolated, and sea ice concentrations have been supplied in the ice zones. Care must be taken when using HadISST1 for studies of observed climatic variability, particularly in some data-sparse regions, because of the limitations of the interpolation techniques, although it has been done successfully

[Sheppard and Rayner, 2002]. It is recommended that the noninterpolated SST data set HadSST [Jones *et al.*, 2001] be used alongside HadISST1 for climate monitoring and climate change detection studies, as was done in the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) [Folland *et al.*, 2001a].

[3] At the time of writing, the GISST/HadISST family of (nominal) 1° latitude-longitude resolution monthly data sets is unique among available integrated SST and sea ice analyses in being globally complete while spanning well over a century. GISST has been widely used in AGCM simulations [e.g., Folland *et al.*, 1998; Rodwell *et al.*, 1999; Rowell and Zwiers, 1999; Zheng and Frederiksen, 1999] and HadISST1 has also already been used in this context [e.g., Hansen *et al.*, 2002; Rodwell and Folland, 2003]. Other SST data sets have been developed for different purposes and more restricted periods. The U.S. National Oceanic and Atmospheric Administration (NOAA) optimal interpolation (OI) SST data sets [Reynolds and Smith, 1994; Reynolds *et al.*, 2002] are globally complete, contain varying sea ice, have a spatial resolution of 1° latitude by 1° longitude (hereafter 1° area) and weekly temporal resolution. They utilize both in situ SSTs from ships and buoys, and bias-adjusted SSTs from the satellite-borne advanced very high resolution radiometer (AVHRR, as does HadISST1), but only start in late 1981 when AVHRR began. The best known noninterpolated gridded in situ-only historical SST data set is included within the Comprehensive Ocean-Atmosphere Data Set (COADS; Woodruff *et al.* [1998]) which begins in 1856, but does not include the pre-1942 bias corrections included in both HadSST and HadISST1 (Folland and Parker [1995]; also section 3.2). Other interpolated historical data sets [Kaplan *et al.*, 1998, 2003; Smith *et al.*, 1996, 1998; Smith and Reynolds, 2003] are at most quasi-global, do not contain varying sea ice (although T. M. Smith and R. W. Reynolds are currently adding our sea ice analysis to their SST fields) and have lower spatial resolution because of the relative lack of data before the satellite era. These historical data sets all use data reconstruction techniques based on empirical orthogonal functions (EOFs), which are used to capture the major modes of SST variability and are then projected onto the available gridded SST observations to form quasi-globally complete fields.

[4] In HadISST1, broad-scale fields of SST are reconstructed using one of these EOF-based techniques, reduced space optimal interpolation (RSOI). RSOI is described by Kaplan *et al.* [1997], who show that it is more reliable than EOF projection, which was used in GISST and by Smith *et al.* [1996]. We adapt RSOI into a two-stage process: first reconstructing the global pattern of long-term change and then the residual interannual variability. This results in a better representation of trends than does a single application of RSOI as used by Kaplan *et al.* [1998, 2003]. Also, we augment the reconstructions by blending with quality-improved in situ SST to recapture local variance lost in the broad-scale RSOI.

[5] Owing to the diversity of the input data, construction of HadISST1 was a complex process. Sea ice and SST data were collated separately, and biases were removed as far as possible. Where appropriate, gaps were interpolated before the SST and sea ice analyses were merged to form a globally

complete product. Creating a sea ice analysis for the last 130 years as a companion to the SST fields is itself an involved procedure, because of the varied data sources that have to be exploited and their inhomogeneities. Here we have attempted to remove the effects of these inconsistencies (described in section 2 and Appendix A) and the result is an integrated SST and sea ice analysis without the unphysical discontinuities seen in GISST data sets [Rayner and Parker, 1999]. The HadISST1 sea ice analysis has been used both for climate monitoring [Folland *et al.*, 2001a] and for model validation [Gregory *et al.*, 2002]. Section 3 documents the input SST data, and summarizes the theory and application of the methods used in the SST analysis (with details in Appendices B and C). Section 4 describes how the SST and sea ice analyses were combined (with details in Appendix D) and includes special consideration of the Southern Ocean.

[6] We also document, in section 5 and Appendix E, the Hadley Centre nighttime marine air temperature (HadMAT1) data set, which supersedes the Met Office historical marine air temperature (night) data set MOHMAT4N [Parker *et al.*, 1995b]. Monthly fields in HadMAT1 were interpolated using RSOI, in much the same way as HadISST1, but without the sea ice data; HadMAT1 is therefore not truly globally complete. We have revised the corrections applied to the historical NMAT data to remove the effects of changing ships' deck heights and extended them to the present, as ships have continued to become taller. These improved corrections have brought the NMAT data into better agreement with SST and collocated land air temperature, removing some of the discrepancies in the study by Folland *et al.* [2001a]. As in previous studies [e.g., Parker *et al.*, 1995b], we use variations in NMAT to corroborate those seen in SST.

[7] Section 6 presents key diagnostics used to verify HadISST1 and HadMAT1, along with comparisons with GISST and several published SST analyses. Our conclusions are summarized in section 7. Table 1 contains a glossary of some recent Met Office marine temperature analyses for ease of reference.

2. Sea Ice Analysis

[8] Sea ice data are important for forcing AGCMs, and can also be used to estimate SST in nearby open water. However, the available sea ice data are heterogeneous, because sea ice has been observed using a variety of methods and in very different levels of detail through the historical record. Although many data sets may provide an approximately homogeneous record of sea ice extent, i.e., the total size of the region at least partly covered by sea ice, the important parameter from the perspective of forcing a climate model is the variation in sea ice concentration, i.e., the relative fraction of sea ice in each grid box. This is more likely to be heterogeneous. For example, satellite-borne passive microwave retrievals of sea ice concentration are not consistent with historical charts based on in situ observations, aerial reconnaissance and infrared satellite images. So, GISST2.3b and 3.0 contained spurious decreases in sea ice area when microwave retrievals began to be utilized, particularly in the Southern Hemisphere (Figures 1 and 2; also see Rayner and Parker [1999]). This was caused by updating chart-derived fields set to 100% concentration poleward of the marginal ice zone (i.e., the area of partial

Table 1. Glossary of Met Office Marine Temperature Products

Data Set Name	Description
MOHSST6D	Gridded, quality-controlled and bias-adjusted in situ-only SST. 5° area, 1856 onward [Parker <i>et al.</i> , 1995b].
HadSST1	As MOHSST6D, but corrected to remove effect of data sampling on variance, 1870 onward [Jones <i>et al.</i> , 2001].
GISST2.2	Globally complete SST, reconstructed using EOF projection technique, includes bias-adjusted AVHRR SST and sea ice analysis. 1° area, 1903 onward [Rayner <i>et al.</i> , 1996].
GISST2.3b	As GISST2.2, but with improved analysis for 1870–1948.
GISST3.0	As GISST2.3b, but reconstruction blended with in situ SST as in Appendix E of this paper.
HadISST1	Globally complete SST, reconstructed using RSOI and blended with variance-corrected in situ data, includes bias-adjusted AVHRR SST and homogenized sea ice analysis. 1° area, 1871 onward (described here).
MOHMAT42N	Gridded, quality-controlled NMAT, corrected using adjustments documented in Parker <i>et al.</i> [1995b]. 5° area, 1856 onward.
MOHMAT43N	As MOHMAT42N, but with new deck height corrections (section 5 of this paper).
HadMAT1	Quasi-globally complete NMAT, reconstructed using RSOI and blended with in situ data. 5° area, 1856 onward (described here).

sea ice cover near the ice edge) with microwave-derived fields which included open water areas poleward of the marginal zone, particularly in summer (see below for further discussion). Such heterogeneous records must be manipulated to provide a self-consistent history of observed sea ice concentration without unrealistic trends or discontinuities. This was done for HadISST1 in collaboration with a group of international experts brought together by ECMWF to produce a homogenized sea ice data set for input to the ERA40 Reanalysis. Because of time constraints, it was necessary to adopt compromises to produce a workable, but inevitably still imperfect, data set. The OI.v2 data set [Reynolds *et al.*, 2002] uses the same sea ice analysis.

[9] For the most part sea ice extents were left as in the input data sets, so HadISST1 should provide a good record of sea ice extent change over the last century in the Northern Hemisphere and over the last three decades in the Southern Hemisphere. As data sources are limited for the Southern Hemisphere, HadISST1 gives only a general indication of sea ice extent variations there on decadal timescales prior to the 1970s.

[10] We summarize the sources of sea ice data used, in approximate chronological order. Details are in the cited references. To the best of our knowledge, we used all the digitized information readily available at the time (1999); additional data, however, have become available since then (see section 7) and some still reside in historical archives (V. Smolyanitsky, personal communication, 2002). Details of analysis methods used to create the HadISST1 sea ice fields can be found in Appendix A.

2.1. Digitized Sea Ice Charts

[11] This subsection details the data originally derived from hand-drawn charts. In some cases these charts were simply ice extents, in others some information about sea ice concentration was also available.

2.1.1. Walsh Northern Hemisphere Sea Ice Concentration Charts

[12] These are end-of-month sea ice concentration fields for the Northern Hemisphere for 1901–1995, covering the Arctic Ocean and peripheral seas, assembled from a variety of sources [Walsh, 1978; Walsh and Johnson, 1978; Walsh and Chapman, 2001] (hereinafter referred to as Walsh). The data were collated to [Walsh, 1978] “provide a relatively uniform set of sea ice extent for all longitudes as a basis for hemi-

spheric-scale studies of observed sea ice fluctuations.” Although the Walsh data set is based on passive microwave retrievals from satellites after October 1978, we only used it up to this time. The pre-satellite data are expressed as fields of sea ice concentration, but their information content is mainly sea ice extent: complete cover is assumed within the ice-pack.

[13] Because directly observed sea ice concentrations were not available and only the ice extent could be deduced from the sources, the characteristics of the marginal sea ice zone were imposed by Walsh using climatological seasonal ice concentration gradients calculated from passive microwave satellite observations. As there were no data at all for September–March 1901–1956, sea ice concentrations in the marginal ice zone in these months were temporally interpolated using available data for the summer half of each year, along with observed temporal intermonthly autocorrelations of sea ice concentration [Walsh, 1978].

[14] The resulting Walsh compilation includes measured or calculated data for all months in all years from 1901–1995. However, the pre-passive microwave data are not entirely complete for the Northern Hemisphere as the Great Lakes and the Caspian Sea are not included. Because Walsh amalgamated heterogeneous data types, there is a discontinuity in the total sea ice area in the full data set when the satellite microwave data began in 1978 (see documentation at the U.S. National Snow and Ice Data Center (NSIDC, <http://www.nsidc.org>)).

[15] In HadISST1, Northern Hemisphere Walsh fields for 1901–1978 were used as the main data source for that period. As the SST data are monthly averages (section 3), and the satellite-based sea ice data used later are monthly medians, consistency within the HadISST1 sea ice time series required conversion of the end-of-month Walsh data to monthly medians. So an end-of-month sea ice concentration climatology was calculated from calibrated passive microwave data (see Appendix A) for 1979–1996 and subtracted from the Walsh end-of-month fields to give fields of end-of-month sea ice concentration anomalies for 1901–1978, which were then linearly interpolated to mid-month. A monthly median concentration climatology of calibrated passive microwave data for 1979–1996 was finally added to give monthly median equivalent Walsh sea ice concentration fields.

2.1.2. Great Lakes Fields

[16] Assel [1983] assembled a set of half-monthly sea ice concentration fields for the Laurentian Great Lakes for

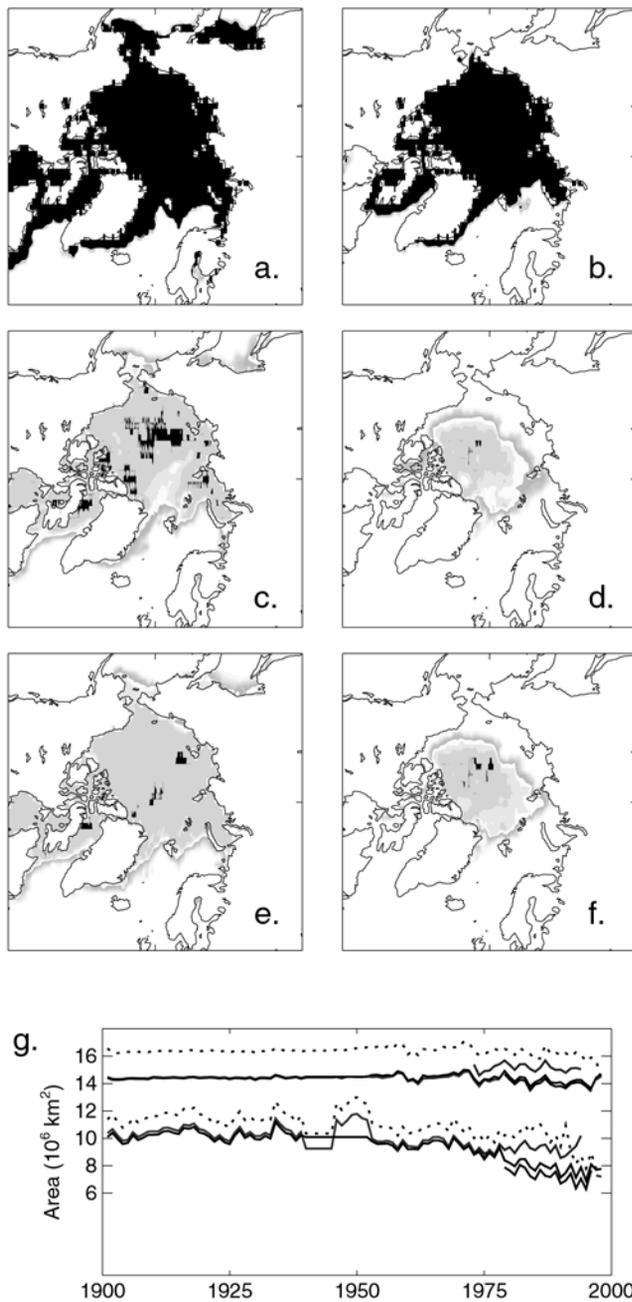


Figure 1. Northern Hemisphere sea ice. Concentration maps: (a) Walsh, January 1930; (b) Walsh, August 1930; (c) GSFC, January 1990; (d) GSFC, August 1990; (e) HadISST1, January 1930; and (f) HadISST1, August 1990. Values = 100% are colored brown. Sea ice area time series 1901–1998: (g) (top curves: January; bottom curves: July) from GISST3.0 (black dots), Walsh (light blue), NIC (green), GSFC (red), NCEP (dark blue) and HadISST1 (black). Appendix A gives details of procedures used to produce HadISST1 fields. Ghostly grey features in open water areas are artifacts of plotting only and do not feature in the data set. See color version of this figure at back of this issue.

1960–1979 from charts produced by Environment Canada, the Great Lakes Environmental Research Laboratory and the U.S. Coast Guard (see <http://www.nsidc.org>). Concentration is given to the nearest 10%. Fields are available from

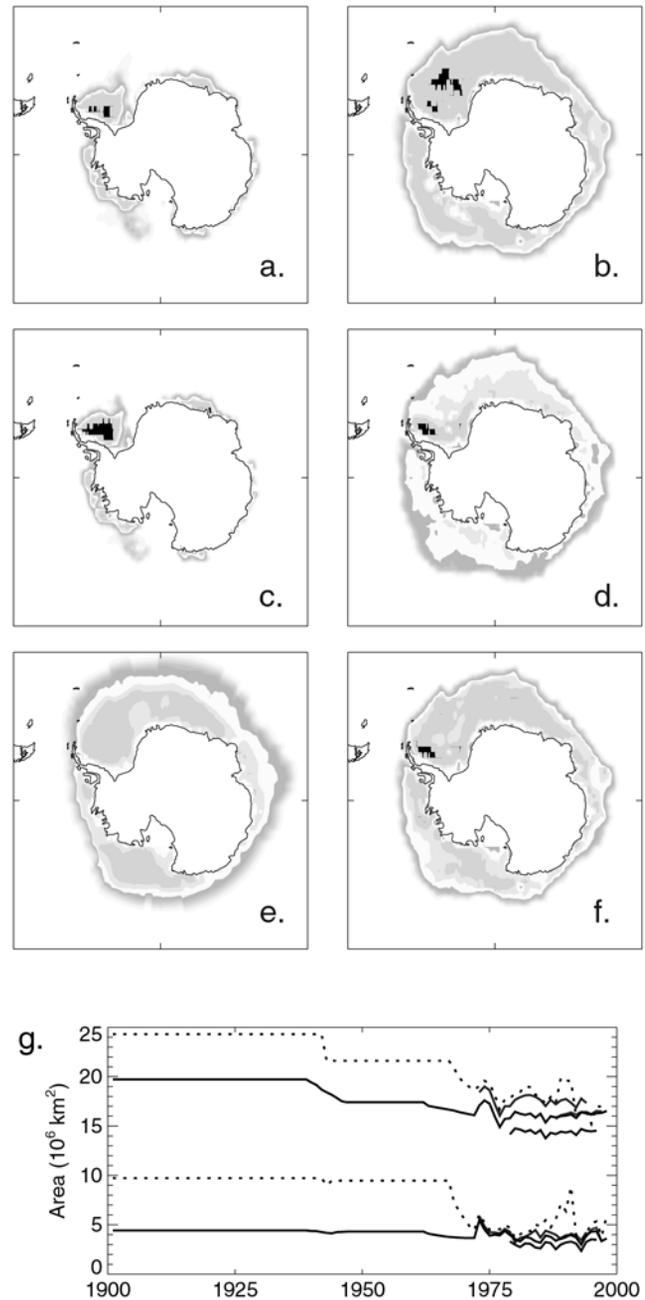


Figure 2. Southern Hemisphere sea ice. Concentration maps: (a) Bristol, January 1990; (b) Bristol, August 1990; (c) GSFC, January 1990; (d) GSFC, August 1990; (e) HadISST1, August 1930; and (f) HadISST1, August 1990. Values = 100% are colored brown. Sea ice area time series 1901–1998: (g) (top curves: September; bottom curves: January) from GISST3.0 (black dots), Bristol (light blue), NIC (green), GSFC (red), NCEP (dark blue) and HadISST1 (black). Appendix A gives details of procedures used to produce HadISST1 fields. See color version of this figure at back of this issue.

the second half of December to the end of April in each year. The rest of the year is taken to be ice-free. Years before 1960 were accorded this climatology in HadISST1 as described in Appendix A (section A1.3). Satellite-based passive microwave retrievals are used for 1980 onward.

2.1.3. Antarctic Atlas Climatologies

[17] Before the advent of satellite-based imagery in 1973, sea ice concentration data for the Antarctic are not available, and sea ice extent data are not readily available for individual months, seasons or years, although some visible and infrared data do exist for 1966–1972 [Zwally *et al.*, 1983] and some undigitized charts reside in national archives (e.g., V. Smolyanitsky, personal communication, 2002). Readily available information was limited to two historical climatologies of sea ice extent. Therefore our sea ice concentration analysis before 1973 is derived indirectly, and does not include any interannual variability, though there are some trends resulting from the differences between climatologies for different periods.

[18] Prior to 1973, we used the calendar monthly sea ice extent climatology for 1929–1939 published by the *Deutsches Hydrographisches Institut* [1950], and that of *Tolstikov* [1966], which summarizes ice extents observed during Russian expeditions between 1947 and 1962. The 1929–1939 climatology was repeated for all years 1871–1939 and the 1947–1962 climatology was used for all years 1947–1962. In periods for which no information was available, fields were interpolated (see Appendix A, section A2.4). Ice extent in these climatologies is generally greater than present-day extents, especially in winter (Figure 2), and there is some independent supporting evidence for this [de la Mare, 1997; Jones, 1990]. Climatological spatial variations of sea ice concentration within the ice edge were reconstructed using statistics of recent sea ice concentration and its gradient, as described in Appendix A (section A2.3).

2.1.4. National Ice Center Charts for Both Hemispheres

[19] Quasi-weekly sea ice concentration and extent for both the Northern (90° to 45°N) and Southern (90° to 50°S) Hemispheres for 1973–1994, were digitized by the U.S. National Climatic Data Center (NCDC) from hand-drawn U.S. National Ice Center (NIC) analyses [Knight, 1984]. The charts were based on U.S. Navy, Canadian and Danish aerial reconnaissance data and from retrievals from advanced very high resolution radiometer (AVHRR), passive microwave, and other satellite instruments. They were developed mainly for shipping purposes and are available from <http://www.nsidc.org>. They were gridded onto a 1° area grid, and medians of the weekly values at each grid point were taken as the monthly concentration fields (W. Chapman, personal communication, 1998).

[20] The information in the NIC charts is spatially heterogeneous. The most detailed information is in the marginal ice zone in areas of operational interest. Information for regions with few shipping operations, including in general the Southern Hemisphere, is of lower resolution and quality. Inland seas are excluded. In addition, differences in analysis arising from, for example, a change of analyst are non-reproducible (J. Maslanik, personal communication, 1998).

[21] In HadISST1, NIC fields for the Southern Hemisphere were used as the main data source for 1973–1978. In the Northern Hemisphere, NIC charts were used mostly to

calibrate the summer passive microwave data, although they are indirectly included through the use of the Walsh data set (section 2.1.1).

2.2. Passive Microwave Retrievals

[22] Passive microwave retrievals from the Scanning Multichannel Microwave Radiometer (SMMR) and Special Sensor Microwave/Imager (SSM/I) instruments carried on Nimbus 7 and Defense Meteorological Satellite Program (DMSP) satellites are available every other day from 25 October 1978 to July/August 1987 from the SMMR and daily thereafter from the SSM/I [Cavalieri *et al.*, 1997]. Several retrieved sea ice concentration data sets are available, based on different algorithms employing the large contrast in microwave emission from sea ice and open water [Burns *et al.*, 1987]. Each algorithm (e.g., NASA Team [Cavalieri *et al.*, 1997, 1999], Bootstrap [Comiso, 1986], Bristol [Smith, 1996; Hanna and Bamber, 2001], enhanced NASA Team [Markus and Cavalieri, 2000]) produces a data set with different characteristics according to the formulation of the algorithm and the method of filtering out noise arising from weather effects. In the main, we used the data set from the Goddard Space Flight Center derived using the NASA Team algorithm [Cavalieri *et al.*, 1999] (hereinafter referred to as GSFC) as it was a long “homogenized” record for both hemispheres and readily available.

2.2.1. GSFC Data for Both Hemispheres

[23] The SMMR (SSM/I) instruments monitored vertically and horizontally polarized radiances at 18 (19.4) GHz, and vertically polarized radiance at 37 GHz. The orbits of the carrier satellites (one for SMMR; 3 for SSM/I) were similar but not identical [Cavalieri *et al.*, 1999]. So Cavalieri *et al.* [1997, 1999] of the NASA Goddard Space Flight Center (GSFC) endeavored to construct a homogeneous sea ice concentration record from the SMMR and SSM/I retrievals using the NASA Team algorithm. The ready availability and apparently homogeneous nature of these data for October 1978 to December 1996 were the main reasons for basing our analysis for this period on them.

[24] These data are nearly global and include inland lakes and seas, except the Caspian Sea (see Appendix A, section A1.3, for how ice information for the Caspian Sea was obtained). They provide details of sea ice concentration variation within the ice pack [Barry and Maslanik, 1989]. Data are affected by weather, but climatological SST thresholds (278 K in the Northern Hemisphere and 275 K in the Southern [Cavalieri *et al.*, 1999]) are used to filter out unreasonable retrievals of sea ice where the sea is too warm. Land contamination resulting from the relatively large footprint of the instrument (of order 50 km) can lead to spurious sea ice appearing around the coasts, but careful use of land/sea masks helps to remove this, although difficulties remain in small areas such as the Great Lakes. A bigger problem, however, is that thin ice is not identified as such by the microwave retrievals: instead it is returned as a mixture of thick ice and open water [Emery *et al.*, 1994]. Also, ponds resulting from summer melting on top of the ice often cause the microwave instrument to return a 10–30% lower than actual concentration of sea ice [Comiso and Kwok, 1996]: this particularly affects the Arctic in summer, as the Antarctic sea ice breaks up and disintegrates in summer and melt-ponds are a less prominent feature. Wet

snow on top of ice also affects the microwave emission [Smith, 1998]. The GSFC data set is thought [e.g., Hanna and Bamber, 2001; Markus and Cavalieri, 2000] to estimate too low sea ice concentrations in the Southern Hemisphere, so we have adjusted the data to be consistent with the Bristol and NCEP data (Appendix A, section A2.1).

[25] For HadISST1, the daily GSFC data were gridded from their 25 km polar stereographic grid to a regular 1° area grid, by forming a simple area-weighted average. However, the satellites on which the SMMR and SSM/I instruments are flown do not travel directly over the North Pole. This led to a data void from the Pole to 84°N in SMMR data and from the Pole to 87°N in SSM/I data. Where data void grid boxes had more than one neighbor with a concentration value, the voids were interpolated using inverse-square distance weighting [Cressman, 1959] with zero weighting for distances of 160 km or greater. Other grid boxes were filled using linear interpolation between 100% at the Pole and the average sea ice concentration at the most northerly latitude to contain data in a 31° longitude band centered on the target grid box. Median values of all available daily concentrations in each grid box in a given month were used to create monthly fields.

2.2.2. NCEP Data for Both Hemispheres

[26] The NCEP operational sea ice data set [Grumbine, 1996], available from <http://polar.wwb.noaa.gov>, is used to update HadISST1. It is based on SSM/I data and the NASA Team algorithm, like the GSFC data. However it is available only from 1997 onward, and the processing of the data differs from that of the GSFC data set. The data set was provided daily with a 0.5° area resolution (R. Grumbine, personal communication, 1999). An area-weighted average was used to produce 1° area fields, and the median value of all available daily concentrations in a given grid box over a month was used to create monthly fields. 1° area grid boxes made up of four 0.5° area coastal boxes were filled using the average of data in their nearest seaward neighbors, thereby avoiding footprints which may have been affected by land.

2.2.3. Bristol Algorithm Data for the Antarctic

[27] A passive microwave-derived sea ice data set for the Southern Hemisphere has recently been developed [Hanna and Bamber, 2001] from SSM/I retrievals (September 1987 onward) using the Bristol algorithm [Smith, 1996]. There are several reasons to believe that this is a more realistic data set than the GSFC in this region. The algorithm itself is a development of the Bootstrap algorithm [Comiso, 1986] and maximizes the sea ice signal by using information from both the polarization of the microwave radiation and the relationship between the different microwave frequencies to determine sea ice concentration. This is different from the approach taken by the Bootstrap algorithm, which uses either polarization or frequency information depending on the circumstances but not both (see Smith [1996] for further details). In addition “100% sea ice” and “open water” calibration points are obtained for each season and each year to allow for the effect of changing surface conditions on the brightness temperature of sea ice. This also helps to remove differences resulting from the use of a succession of instruments over the years. By contrast, the GSFC data set uses a single set of calibration points for all seasons and all years [Cavalieri et al., 1997], which is thought to lead to erroneous retrievals of sea ice concentration (see, for example, sensitivity tests given by Hanna and Bamber [2001]). In compari-

sons of Bristol- and NASA Team algorithm-derived concentrations against those from a limited sample of contemporaneous thermal AVHRR measurements during September and October 1994 in the Ross and Weddell Seas [Hanna and Bamber, 2001], the Bristol concentrations correlated more highly with those from higher-resolution, but less complete, AVHRR data than did the NASA Team data.

[28] Preliminary comparisons between SSM/I-based sea ice concentrations and in situ observations from British Antarctic Survey ships in the Weddell Sea have been made by E. Hanna (personal communication, 2000). The Bristol algorithm data were found to be more compatible with the in situ data than were SSM/I data processed using other algorithms. The in situ observations tended to give higher concentrations than the SSM/I, but this may be partly the result of a systematic overestimation, by even these highly trained observers, of the concentration of sea ice in their vicinity, owing to perspective (similar to the effect on surface-based observations of cloud cover [New et al., 1999]).

[29] The homogenized GSFC data set (see above) has the advantage of merging data from both the SMMR and SSM/I instruments. However, there was a discrepancy between the sea ice extents indicated by the GSFC and the NCEP data, and the Bristol algorithm data appeared from the limited available overlap period to agree better with the NCEP data. So we used the Bristol data, which cover a longer period than the NCEP data, to adjust the GSFC data (see Appendix A, section A2.1). The latter remain an essential data source for HadISST1 because they extend back 9 years before the Bristol data.

[30] For HadISST1, monthly mean Bristol algorithm fields for September 1987 through 1997 were obtained from E. Hanna on polar stereographic grids and converted to the regular 1° area latitude/longitude grid in the same way as for the daily GSFC fields.

[31] The apparent differences between the icier Bristol means and the less icy NASA medians may slightly underestimate the algorithm-related differences when the sea ice concentration approaches 100%, because the distribution of concentrations is likely to be negatively skewed so the mean will be less than the median. The opposite will apply for concentrations near zero.

3. Sea Surface Temperature Analysis

[32] Our objective is a spatially complete, monthly SST analysis for 1871 to date, preserving real climate signals on global, ocean-basin and subregional scales, while minimizing random errors, sampling noise, and systematic biases.

[33] To achieve this, we based our SST analysis for 1871–1981 on gridded, quality-controlled in situ SST observations; the gridded data for 1871–1941 were bias-adjusted following Folland and Parker [1995] (section 3.2). To extend the analysis over most of the data-sparse oceanic regions, we applied reduced space optimal interpolation (RSOI, section 3.4) [Kaplan et al., 1997] to gridded in situ anomaly data for 1871–1981, on 4° area resolution between 1871 and 1948 and on 2° area resolution thereafter. The RSOI utilized empirical orthogonal functions (EOFs) based on the above-mentioned gridded in situ data and on combined in situ/remotely sensed SSTs (section 3.6). We adjusted the satellite SSTs to be unbiased relative to the in

situ data (Appendix C). For 1982 onward, the RSOI technique was applied to the in situ/satellite SST combination; an additional analysis of the Southern Ocean was performed (section 4.2) for this period. Because, like all optimal interpolation schemes, RSOI tends to the first guess value when data are sparse [Kaplan *et al.*, 2003], special measures were taken to preserve the “trend” in the global mean (section 3.4).

[34] The bias-adjusted noninterpolated gridded in situ data were meanwhile further quality-controlled to homogenize their grid-scale variance following Jones *et al.* [2001]. The RSOI-reconstructed fields were then blended with these data (section 3.5) to restore some of the variance on ~ 500 km scales which was not captured by the large-scale RSOI analysis. Thus HadISST SST fields are not purely reconstructed from EOFs, but are a blend of the reconstruction with the original data.

[35] At the ice margins, 1° area SST values were specified using statistical relationships between sea ice concentration and SST (section 4.1). These formed the outer boundary condition for the completion of the global fields on 1° area resolution (section 4.2) using the Poisson blending technique [Reynolds, 1988], which extended the observed and reconstructed data over the remaining data-void regions.

3.1. In situ Input Data

[36] Individual ships’ observations from the Met Office Marine Data Bank (MDB), which from 1982 onward also includes data received through the Global Telecommunication System (GTS), were quality controlled by the methods used for the Met Office Historical SST data set (MOHSST6) [Bottomley *et al.*, 1990; Parker *et al.*, 1995b] and gridded onto a 1° area grid; the number of constituent observations used in each gridded mean was recorded. In order to enhance data coverage, we also used monthly median SSTs for 1871–1995 from the Comprehensive Ocean-Atmosphere Data Set (COADS) [Woodruff *et al.*, 1987, 1998]. (Where there was a choice, we used the “enhanced” version of COADS, which has 4.5 s.d. trimming and includes data from all in situ platforms, rather than just ships.) The pre-1960 COADS SSTs were on a 2° area resolution, so we converted them to anomalies from the 1961–1990 GISST2.2 climatology and then assigned them to a 1° area, according to the mean latitude and longitude of each 2° area datum. Data for 1960–1995 were taken from the COADS Pan American Climate Studies program 1° area summaries. COADS SSTs were inserted into 1° area grid boxes that did not have an MDB value. Hereafter, we refer to this combined data set as MDB/COADS. For each 1° area where we used COADS data, the total number of observations in a month was read from the COADS monthly statistics.

3.2. Bias Adjustment of in situ Data

[37] For climate research, SST analyses should be unbiased to order 0.1° [Taylor, 1999]. So bias adjustments are expected to be necessary, since we have combined SST data obtained by diverse methods, each with different biases and even different definitions of SST.

[38] For example, SSTs derived from samples collected in buckets prior to 1942 were generally, although not exclusively, too low relative to modern data [Folland and Parker,

1995], owing to the noninsulated or partially insulated fabric of those buckets and their exposure to the wind on deck. Folland and Parker [1995] developed adjustments for these biases on a 5° area, monthly resolution. We ascribed these to each constituent 1° area and smoothed the resulting fields twice using a 1:2:1 filter (both east-west and north-south) to remove discontinuities at the edges of 5° areas. The smoothed adjustments were applied to the blended 1° area MDB/COADS anomalies from 1871 through 1941. Published tests have given support to the Folland and Parker [1995] adjustments [Folland and Salinger, 1995; Folland *et al.*, 1997, 2001b; Hanawa *et al.*, 2000; Smith and Reynolds, 2002]; see also section 7.

[39] Also, buckets used for much of the historical SST record to collect seawater, as well as fixed or drifting buoys, sample the bulk temperature of the water around a meter or less below the sea surface. However, many measurements taken by modern voluntary observing ships use engine room intake water thermometers, or, much less often, hull contact sensors; these are representative of the mixed layer temperature down to about 10 m below the surface [World Meteorological Organization (WMO), 1955–1999] and so arise from a rather ill-defined depth. In the daytime, with low wind speeds, the temperature at depths sampled by SST buckets and buoys can be systematically higher than that at the typical depths of engine intakes and hull sensors [Taylor, 1999]. In general, however, modern insulated buckets were found to have a cold bias (0.08°C) relative to engine intake data [Folland *et al.*, 1993]. This may be a result of the influence of ships’ heating on the water sampled by engine intake thermometers: Kent *et al.* [1993] found that SSTs derived from engine room intake measurements were biased warm relative to those from hull contact sensors by 0.35°C on average. We did not apply any bias adjustments to post-1941 in situ SSTs, because our SST anomalies may not be greatly biased overall since they are expressed relative to the mix of measurements made during 1961–1990. However, the possible need for adjustments to modern data is being investigated (section 7).

3.3. Preparation of Data for Reduced Space Optimal Interpolation (RSOI)

[40] In this section we document the regridding and additional quality control of the SST data prior to our implementation of RSOI. The same procedures were applied both to the data used to develop the EOFs for the RSOI, and to the data input to the RSOI.

3.3.1. Regridding

[41] Owing to the greater availability of SST data since the mid 20th century (Figure 3), we reconstructed SST anomalies on a 2° area resolution from 1949 onward but on a 4° area resolution for earlier years. The combined, bias-adjusted MDB/COADS 1° area monthly SST anomalies (sections 3.1 and 3.2) were therefore averaged into 4° areas centered on the equator, 4°N , 4°S , 8°N , 8°S , etc., for 1871–1948 and into 2° areas centered on the equator, 2°N , 2°S , 4°N , 4°S , etc., for 1949 onward. Having equatorially centered boxes enables a better representation of the enhanced equatorial Pacific cold tongue during La Niña events than was possible with the equator-flanking grids of Bottomley *et al.* [1990] or Parker *et al.* [1995b]. The numbers of observations in each 1° area were used as weights to

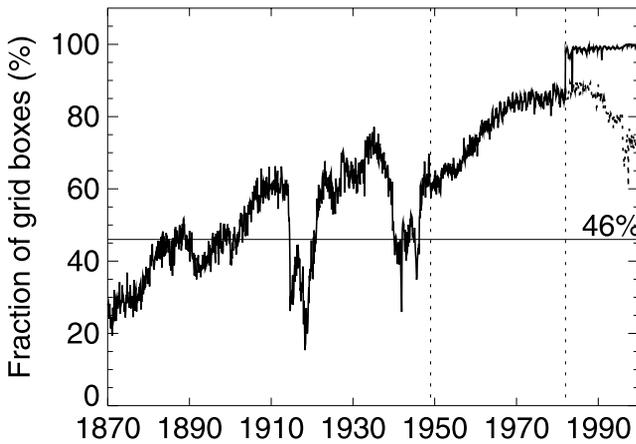


Figure 3. Percentages of grid boxes with SST data in regions where the EOFs were defined. Grid boxes are 4° areas through 1948 and 2° areas thereafter. The dotted time series from 1982 onward is for in situ data sources only: the solid line includes AVHRR data. In months falling below the horizontal line at 46%, three-month running mean SST anomalies were used in the construction of HadISST1 (see text).

calculate the 2° or 4° area average anomaly. (At the time of the analysis, there was no count of numbers of observations readily available from 1982 onward, so all grid boxes were weighted equally thereafter.)

3.3.2. Additional Quality Control of Gridded Data

[42] Despite the quality control applied to the individual observations and “winsorisation” [Bottomley et al., 1990] into 1° area monthly means, some obviously erroneous values still remained in the gridded fields. So we applied additional neighbor-based quality controls. If there was a mean anomaly in more than one of the eight grid boxes neighboring a 2° or 4° area SST anomaly, it was compared against a weighted average of these neighbors. The weighting was proportional to the number of observations contributing to the anomaly in each neighboring grid box. Values deviating from the average of their neighbors by more than a predetermined threshold value were replaced by their neighbors’ average. The threshold value was set to $\bar{\Delta} \pm 1.64\sigma_{\Delta}$, where $\bar{\Delta}$ and σ_{Δ} are the mean and standard deviation of the difference of the value, in that grid box and calendar month, from its neighbors’ average over the whole period of the data set. The threshold was chosen empirically to preserve as many data as possible while removing grid-scale noise. We used the whole period of the data set to obtain stable statistics. Grid boxes with one or no neighboring value were compared against a threshold, set to $\overline{anom} \pm 1.64\sigma_{anom}$ (where \overline{anom} and σ_{anom} are the mean and standard deviation of the anomaly), calculated for each grid box and calendar month over the whole period. If they exceeded this threshold, they were removed. This somewhat ad hoc procedure removed the worst of the remaining unreliable grid boxes prior to reconstruction.

[43] These quality control procedures were also applied separately to the combined MDB/COADS in situ/bias-adjusted advanced very high resolution radiometer (AVHRR) satellite SST data (sections 3.6 and 3.7) used from 1982 onward.

3.4. Reconstruction of in situ Data-Only Fields, 1871–1981

3.4.1. Reduced Space Optimal Interpolation (RSOI) Technique

[44] In the second and third versions of the GISST data set, we used an EOF projection technique to reconstruct monthly SST anomaly fields from incomplete observed data [Rayner et al., 1996]. For HadISST1, we used the similar, but more rigorous, RSOI [Kaplan et al., 1997].

[45] In both methods, a set of fixed EOFs, E , describing the characteristic spatial patterns of SST anomaly variations in a generally well-observed period, was defined from the spatial covariance matrix of the gridded SST data in that period. Use of these EOFs, formed from all available data over large areas, is an improvement on the use of localized analytical correlation-versus-distance functions, which cannot readily take account of the complex spatial characteristics of the data. It is assumed in both EOF projection and RSOI that the same set of patterns dominated throughout the period of reconstruction, and that the magnitude of their amplitudes remained the same [Kaplan et al., 1997]. This is a fundamental assumption of all such reconstruction methods and its validity was tested by Kaplan et al. [1998]. The set of EOFs, E , truncated to admit real signals and exclude noise, was used to reconstruct past fields of SST, T , via a vector of time coefficients, a : $T = Ea$. The reconstruction was fitted to the incomplete, gridded observed data, T_o , for each historical month in a least squares sense in EOF space, determining a for all EOFs used.

[46] The difference between EOF projection and RSOI is found in the expressions for a :

$$a = (E_o^T E_o)^{-1} E_o^T T_o \quad (1)$$

for EOF projection, and

$$a = (E_o^T R^{-1} E_o + \Lambda^{-1})^{-1} E_o^T R^{-1} T_o \quad (2)$$

for RSOI.

[47] Here R is a matrix representing a combination of estimated data error and EOF-truncation error variances and Λ is a diagonal matrix of eigenvalues, i.e., expected variances of each element of a . The subscript o on E denotes the use of only those spatial weights of each EOF that coincide with observed values, T_o .

[48] The EOF projection method (equation (1)) gives stable results when data are plentiful and well distributed. However, if data are sparse and noisy, the least squares fit is unrestrained from yielding large EOF amplitudes, so spurious large anomalies can be reconstructed. RSOI (equation (2)) is superior to EOF projection in that the matrix Λ , which has small terms for those EOF patterns which account for small amounts of covariance in the matrix E , acts to suppress the contribution to the reconstruction from lower eigenvalue (and hence less important) EOFs whose variance contributions are small compared to data error, R . In addition, through the matrix R , reduced weight is accorded to data with greater estimated error variance: thus noisy or sparse data are restrained from yielding spurious, high-amplitude patterns [Kaplan et al., 1997]. The absence of these safeguards in the GISST analyses necessitated some manual intervention to remove noise from the reconstruc-

tions or to vary, subjectively, the number of EOFs used. RSOI has avoided these difficulties for HadISST1. Furthermore, RSOI provides error estimates for the reconstructed field; however, as the RSOI reconstruction is only one step in the creation of HadISST1, these error estimates are not sufficiently accurate to release as part of the analysis.

[49] Although in HadISST we have only used the space-wise RSOI, we note that *Kaplan et al.* [1997, 1998] use reduced space optimal smoothing (RSOS), which also incorporates a model for the time-wise evolution of SST anomalies. We felt that insufficient historical in situ data were available to properly determine the appropriate time-dependent model.

3.4.2. Application of Reduced Space Optimal Interpolation

[50] Optimal interpolation techniques tend to the first guess (in this case, zero anomaly, or the 1961–1990 climatology) in areas where there is no information [*Reynolds and Smith*, 1994]. So, unless the long-term changes of the mean of the data are removed prior to optimal interpolation of the fields, the trends in the resultant global and regional mean time series will be too weak [*Hurrell and Trenberth*, 1999; *Kaplan et al.*, 2003]. We therefore took account of the known long-term changes of SST since the late 19th century [*Parker et al.*, 1995b] first. To do this, we averaged monthly 4° area combined in situ/satellite gridded SSTs (quality-controlled and bias-adjusted) for 1901–1997 into seasonal (JFM, AMJ, JAS, OND) anomalies, low-pass filtered them using a Chebyshev filter [*Cox and Hayes*, 1973] removing variations of period less than 8 years, and calculated the first covariance EOF. We defined this EOF wherever data coverage was at least 50% of 3-month seasons in 1901–1997. The EOF closely represents the pattern of century-long global warming (Figure 4; see also *Folland et al.* [1999]), with a correlation of 0.96 between the time series of its principal component and the low-pass filtered global average SST anomaly. The RSOI reconstruction of this EOF was subtracted from all the 4° area (before 1949) and 2° area (from 1949) quality-controlled monthly gridded data before they were averaged to seasons and used to create the covariance EOFs for the RSOI of the interannual variability (see below), and before they were input to the RSOI. Afterward, these “global change” and residual “interannual” reconstructions were added together. Figure 5 demonstrates the overall effect this procedure has on the resultant global mean SST anomaly. *Kaplan et al.* [1998, 2003] used no such preliminary analysis step and changes in the global mean of their analysis are seen to be smaller than those of HadISST1 when compared to the noninterpolated HadSST data set [*Jones et al.*, 2001]. Through 1981, the “Kaplan” time series in Figure 5 is the global average of the *Kaplan et al.* [1998] analysis, based on the MOHSST5 in situ data set [*Parker et al.*, 1994], and thereafter it is the updated analysis of *Kaplan et al.* [2003], based on the *Reynolds and Smith* [1994] OI data set. It exhibits a weaker warming trend than actually observed. Part of this difference (roughly one third prior to 1920) is caused by differences in the in situ data sets on which the HadISST1 and *Kaplan et al.* [1998, 2003] analyses are based. The global mean SST in HadISST1 may not be perfect, but our extra reconstruction step (or an improvement upon it) does appear to be crucial to achieving a realistic trend.

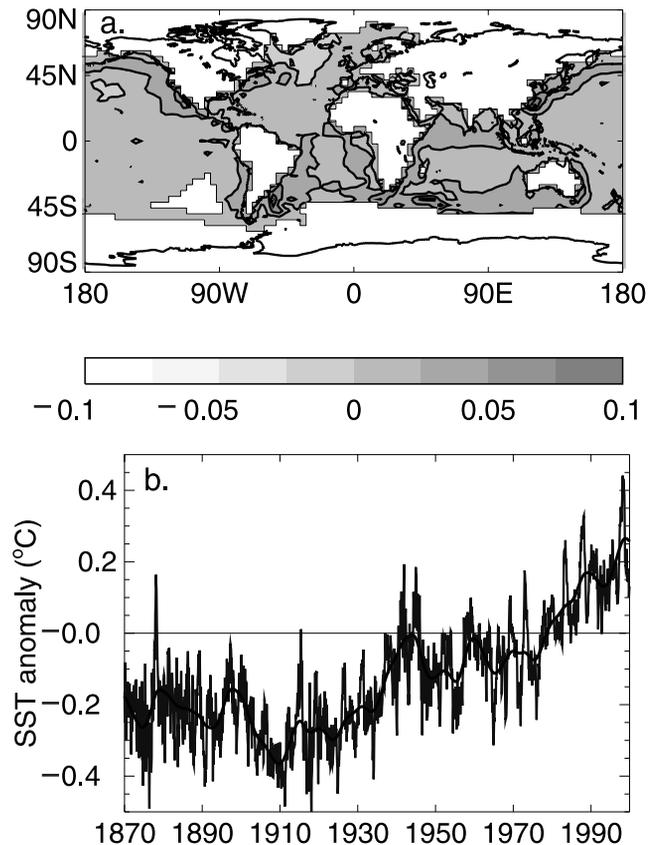


Figure 4. (a) The leading EOF of low-pass filtered (>8 years) 4° area resolution combined in situ and AVHRR data for 1901–1997. (b) Global average of the field reconstructed using only Figure 4a for 1870–1999 (see text). Blue curve is the monthly average, while the red curve has a 21pt binomial (near decadal) filter applied. See color version of this figure at back of this issue.

[51] We used RSOI to reconstruct “interannual” SST anomalies on a 4° area resolution in 1871–1948 and on a 2° area resolution thereafter. For this purpose, two sets of covariance EOFs were created, using seasonal, detrended (by subtracting the global change reconstruction), quality-controlled in situ and bias-adjusted satellite data for 1958–1997: one set on 4° area resolution and one set on 2° area resolution (Appendix B gives details of the truncated subset used). The grids used were as described in section 3.3.1. Again, the EOFs were only defined in grid boxes that had data in at least 50% of 3-month seasons in the input data period (roughly the area covered by the EOF in Figure 4). A greater spatial extent could be achieved for the reconstruction if data for only the last 20 years were used. However, this would mean capturing and utilizing fewer of the common interannual to decadal modes of variability of the data. The 4° area EOFs were calculated for the globe as a whole, but the 2° area EOFs, owing to computing limitations, were created for two separate regions: Atlantic/Mediterranean/Black Sea (hereafter AMB) and Indian/Pacific Oceans (hereafter IP). The AMB and IP regions overlapped by a few grid boxes, to enable a smooth reconstruction.

[52] Some of the monthly SST anomaly fields to be reconstructed at 4° area resolution contained very sparse data

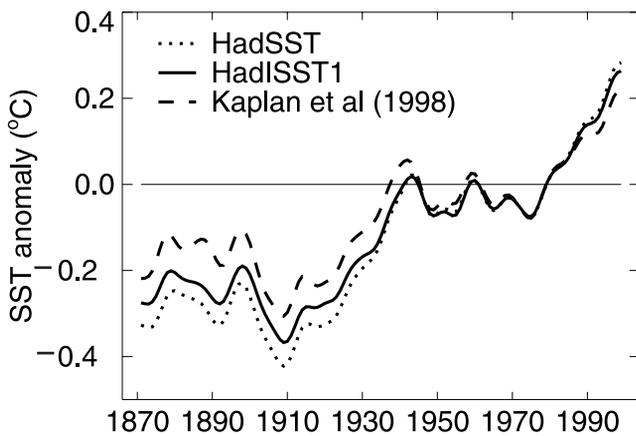


Figure 5. Global (excluding the Southern Ocean), annual average SST anomaly (relative to 1961–1990), 1871–2000. Annual averages smoothed using a 21pt binomial filter.

(Figure 3). Following empirical tests, if fewer than 46% of the 4° area grid boxes in the region encompassed by the 4° area resolution EOFs (similar to the near-global region shown in Figure 4) contained data values in a given month, reconstructions were based on running-three-month-averaged SST anomalies (with a minimum of one month out of three required to form an average in a grid box) instead of on monthly SST anomalies. This step was required to ensure that the temporal progression of the reconstructed fields was always reasonable.

[53] EOFs and data error variance estimates (Appendix B) were used along with detrended monthly (or three-month-running) bias-adjusted SST anomaly fields to produce sets of monthly varying “interannual” reconstructions using the RSOI technique described in section 3.4.1. To ensure a smooth reconstruction from 1949 onward, the AMB and IP reconstructions on 2° area resolution were averaged where they overlapped (i.e., south of South America and South Africa). Finally, the global change reconstruction was added to the assembled reconstructions to give complete fields over the region depicted in Figure 4.

3.5. Blending of Reconstructed SSTs With Refined Gridded SSTs

[54] As described above, ocean-basin or larger-scale EOFs were used to reconstruct quasi-global fields of SST anomaly. The use of such large-scale EOFs exploited teleconnection patterns within and between ocean basins. However, this led to reconstructions with reduced local variance. Therefore localized variability was reintroduced by blending the reconstructions with noninterpolated gridded in situ SST anomalies: the observations were superimposed onto the reconstruction, then the fields were smoothed. In GISST3.0, these observations were from a version of the monthly MOHSST6 data set [Parker *et al.*, 1995b] and the blend required substantial data-adaptive smoothing to remove discontinuities between reconstructed and observed data. In HadISST1, we incorporated fields of gridded MDB in situ SST data, improved in the same way as used in the noninterpolated in situ-only HadSST data set (for details, see Jones *et al.* [2001]) to minimize sampling and random measurement error.

[55] Briefly, this variance-correction technique utilized the preliminary HadISST1 reconstruction as an estimate of the large-scale SST anomaly field. Adjustments were applied to subgrid residuals from this background field, to remove the effect of random and sampling error, while retaining true subgrid variance. (Thus HadSST stemmed from a preliminary step of HadISST1, so was not input to the RSOI.)

[56] Jones *et al.* [2001] show that the variance of area-averaged SST anomalies is smaller in HadSST than in the original gridded data, especially in periods of data sparsity, because random sampling and measurement error variance has been reduced. Overall, the variance is more homogeneous in HadSST than in MOHSST6. So only light data-adaptive smoothing was required after the refined data had been superposed on the reconstructed fields: we calculated climatological standard deviations, σ_n , of the difference between gridded SST anomalies and their neighbors during 1956–1995, and anomalies that differed from the average of their neighbors by more than $3\sigma_n$ were replaced by the average of these neighbors.

[57] The combination of reconstructed and refined gridded data has captured the variability of the SST anomaly fields better than the reconstruction alone in well-observed regions. For example, the sea was exceptionally cold around the eastern and southern UK during the severe winter of 1962–1963, and sea ice was reported locally off the southeastern coast of the UK. The RSOI reconstruction alone yielded a muted cold anomaly pattern, but this was considerably enhanced by the blending procedure (not shown). This procedure also contributed to the improved intermonthly persistence in HadISST1, compared with GISST3 (see section 6).

3.6. Satellite-Based Input Data

[58] At no stage in the observational record have in situ SSTs covered the entire ocean [Parker *et al.*, 1995b]. In particular, the Southern Ocean has generally not been monitored. We therefore made use of satellite-based SSTs in HadISST1 to give almost complete observational coverage for recent years and a firmer basis for the EOFs used to interpolate the earlier, in situ data. We chose to use SSTs from the advanced very high resolution radiometer (AVHRR) because of their greater coverage and longer record than SSTs from the more recent Along-Track Scanning Radiometer (ATSR) [Delderfield *et al.*, 1986]. We used monthly SSTs for January 1982 onward from the operational U.S. National Oceanographic and Atmospheric Administration (NOAA) satellite-borne AVHRR instruments, provided by R. W. Reynolds and D. C. Stokes. These data took the form of 1° area monthly superobservations, separately estimated for day and night. The monthly superobservations were calculated by taking averages of weekly superobservations for all weeks falling totally or mainly in the target month. Fields of numbers of constituent observations were also supplied for use in estimating data errors for the RSOI (section 3.4 and Appendix B).

[59] Satellite-borne radiometers estimate the surface skin temperature of the sea, if atmospheric properties are fully accounted for. However, the algorithms used to retrieve the AVHRR SSTs had been tuned by regression of brightness temperatures from the different infrared AVHRR channels onto in situ SST data from a set of drifting buoys. So the

AVHRR SSTs are in principle equivalent to in situ measurements of bulk SST (although away from the buoys used in the regression their reliability is uncertain). Nevertheless, these SSTs will still be biased if the atmospheric conditions are unrepresentative of those used to tune the algorithms. Also the biases may differ between day and night owing to the different combination of channels used.

[60] Particular causes of bias are the presence of clouds or aerosols and satellite calibration errors. The AVHRR instrument measures radiation from cloud tops where clouds obscure the surface. The inferred temperatures are usually too low and readily detected and rejected [Reynolds, 1993], but this can be difficult for low, warm clouds. Sea ice within the field of view may cause a similar cold bias. Stratospheric aerosols resulting from violent volcanic eruptions have a similar effect on the retrievals to clouds, generating large cold biases [Reynolds *et al.*, 1989; Reynolds, 1993]. Tropospheric aerosols such as Saharan dust can also bias AVHRR SSTs, to an extent depending on their temperature, and therefore their altitude, as well as their optical depth. There was a large negative bias in midlatitudes of the Southern Hemisphere during the last few months of 1991. This was mainly due to the interaction between an instrument calibration error and the algorithm used to correct for the effect of aerosols resulting from the eruption of Mount Pinatubo in June 1991 [Reynolds, 1993]. The problem affected one infrared channel, only used in the nighttime retrieval algorithm. The bias did not affect the austral summer of 1992, but reappeared in 1993. The general difference between the day and night retrievals widened after 1995 when the NOAA-14 satellite was brought into use (not shown). (Since April 2001, the operational AVHRR fields have come from the NOAA-16 satellite and (to date) we use these data to update HadISST1 in near real time.) The recent day AVHRR data appear to follow the in situ data very well, whereas the night data have become colder. However, because the relative cooling of night data was only of order 0.2°C , it may not have been large enough to trigger an operational algorithm change (R. W. Reynolds, personal communication, 2000). Figure 6 shows the zonal average of a smoothed analysis (Appendix C) of this night bias relative to in situ data.

[61] Because the biases in the AVHRR data can be considerable and variable in both space and time, these data cannot be used in combination with in situ data in HadISST1 and similar data sets without adjustment. We describe our adjustment procedure in Appendix C. Although both day and night AVHRR SST data were available, we only used night data in HadISST1, because biases in nighttime retrievals are more easily removed than those in daytime data which are affected by reflected solar radiation and geographically varying diurnal warming. The night data are not biased cold by the diurnal cycle, because the tuning algorithm uses buoy SSTs for all hours. The night AVHRR data used in HadISST1 appear to have been successfully corrected where we have in situ data (not shown). However, the general cooling of the night data relative to the daytime and in situ data in recent years may still have resulted in a cold bias in the Southern Ocean where there are few in situ data. Nevertheless, in this region HadISST1 is slightly warmer than the OI.v2, which incorporates bias-corrected AVHRR SST data through a different method [Reynolds *et al.*, 2002].

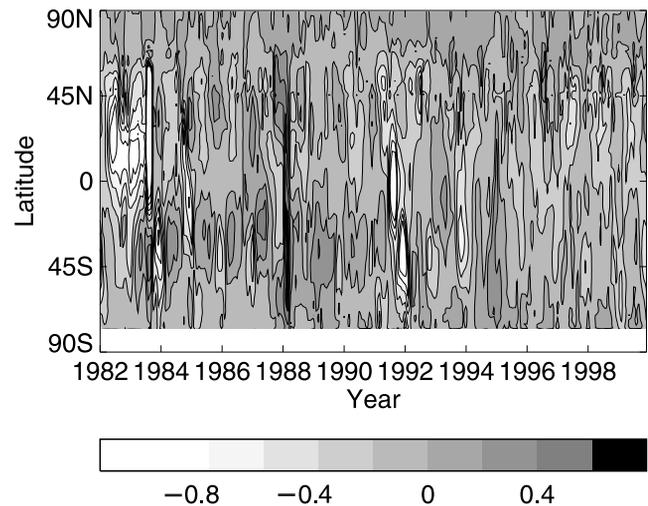


Figure 6. Latitude-time section of the zonal average of smoothed nighttime AVHRR minus in situ SST difference ($^{\circ}\text{C}$), 1982–1999. See color version of this figure at back of this issue.

3.7. Combination of in situ SSTs With Bias-Adjusted AVHRR SSTs, 1982 Onward

[62] In situ and bias-adjusted (Appendix C) AVHRR SSTs were used together from 1982 onward in the RSOI of the “global change” component on a 4° area resolution and the RSOI of the residual “interannual” variability on 2° area resolution, as described in section 3.4.2. The separate 1° area in situ and satellite SST anomaly fields were combined into 2° and 4° area grid box weighted means, using the number of observations contributing to each 1° area value as the weights. After averaging onto these grids, the data were subjected to neighbor-based quality controls as described in section 3.3.2.

[63] The data errors input to the RSOI were defined as in Appendix B, but using the total number of in situ and satellite observations as a divisor. The two components of the reconstruction, covering the areas where the EOFs were defined (approximately mapped in Figure 4), were added to give quasi-global fields which were then blended with the refined in situ data (not satellite data) as in section 3.5. In addition, a reconstruction of the Southern Ocean was performed, described in section 4.2.

4. Combining SST and Sea Ice Analyses

[64] We have constructed a sea ice concentration analysis (section 2) and a quasi-globally complete SST analysis (section 3). In order to make the SST fields globally complete, we use statistical relationships between sea ice concentration and SST to specify temperatures in partially ice-covered grid boxes and bridge the gap between these SSTs and the low-latitude and midlatitude analysis.

4.1. Ice Zone Temperature Analysis

[65] In regions affected by sea ice, there are few in situ observations of SST, especially in the Southern Hemisphere. We therefore specified SST from the sea ice concentration fields developed in section 2. To do this, we used

recent in situ and AVHRR SST observations to develop monthly and geographically varying statistical relationships between collocated sea ice concentration and SST in collaboration with R. W. Reynolds and D. C. Stokes. See Appendix D for details. This work is a development of the method used earlier in GISST2 and 3 [Rayner *et al.*, 1996]. It is known from field experiments (J. Maslanik, personal communication, 1999) that summertime SSTs near sea ice can be several degrees higher than freezing when there is high insolation and light winds, so simply setting SST to -1.8°C at sea ice concentrations of at least 50%, as in the Reynolds and Smith [1994] OI data set biases these SST fields too cold. We used the relationships, along with the sea ice concentration fields developed in section 2, to specify SST in grid boxes partially covered by sea ice, throughout the HadISST record (this method is also used in the OI.v2 analysis [Reynolds *et al.*, 2002]). SST was specified in this way wherever the sea ice concentration was less than 90% (for 90% concentration and above, SST was set to a fixed value (Appendix D)) and at least 15%, the minimum sea ice concentration in HadISST1.

[66] Figure 7 shows fitted relationships between sea ice concentration and SST around 180°W in selected three-month periods. The fits are not always close. However, the quadratic form is physically more realistic than a linear fit, because the increase of SST with decreasing sea ice concentration is generally more rapid when sea ice concentration is high than when it is low.

[67] SSTs in the ice zones were specified using the sea ice concentration in each grid box and the relationship for the longitude band centered on the target location and three month season centered on the target month, or by using the sea ice concentration and the relationship for the target

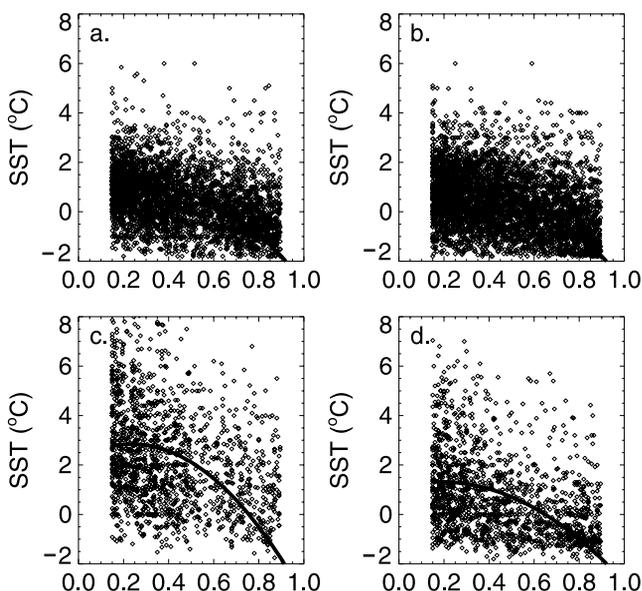


Figure 7. Fitted relationships (quadratic curves) between pairs of sea-ice concentration and sea surface temperature data (diamonds) for the 31° longitude band centered on 179.5°W in the Northern Hemisphere for three-month seasons centered on (a) January, (b) April, (c) July, and (d) October. See color version of this figure at back of this issue.

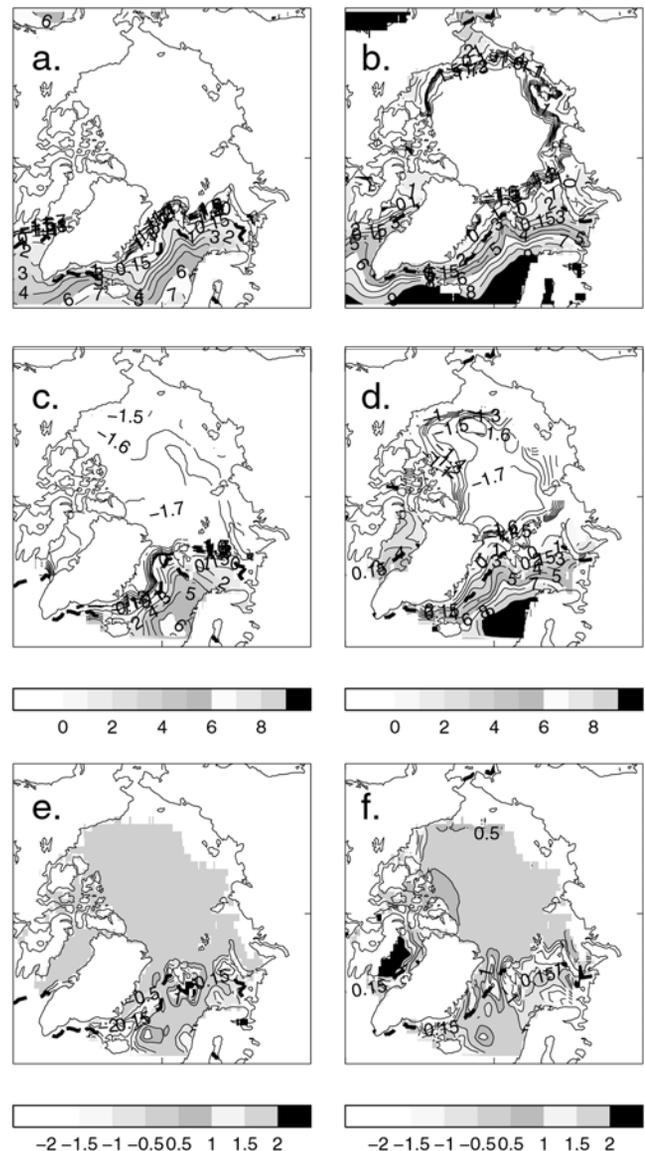


Figure 8. Arctic SST ($^{\circ}\text{C}$), averaged over 1920–1999, from HadISST1 ((a) January and (b) July) and the Generalised Digital Environment Model SST climatology ((c) January and (d) July). Difference GDEM - HadISST1 also shown ((e) January and (f) July). See color version of this figure at back of this issue.

peripheral sea area and three month season, as described in Appendix D.

[68] A comparison (Figure 8) between the 1920–1999 average Arctic SST in HadISST1 and in the top layer of the independent Generalised Digital Environment Model [GDEM] climatology, based on a fit of analytical functions to profile data [Teague *et al.*, 1990], shows good agreement in the winter, but some large discrepancies in the summer. New work (see section 7) using ATSR-2 SST to develop relationships between SST and sea ice concentration should improve our SST fields near sea ice.

4.2. Completion of SST Fields

[69] Prior to 1982, the SST analysis for the area covered by the EOF in Figure 4, was interpolated across the

remaining gaps up to the ice zones using the Poisson technique of *Reynolds* [1988], assuming that the two-dimensional spatial second derivative of the final analysis over the gaps was that of the globally complete 1° resolution GISST2.2 climatology [*Parker et al.*, 1995c].

[70] For analyzing the Southern Ocean in the period 1982 onward, we had three options. First, we could have used the bias-adjusted AVHRR SSTs where available, in the gaps between the analyzed areas (Figure 4) and the ice zones, interpolating the smaller gaps between the AVHRR data using the Poisson technique and the GISST2.2 climatology. However, this would have left small-scale noise in the analyses (Appendix C) giving temporal incoherence. Second, we could have ignored the AVHRR SSTs in the gaps between the analyzed areas and the ice zones, and interpolated using the second derivative of the GISST2.2 climatology. This method was used in a preliminary version (HadISST1.0), but when tested, was found to result in very little variability in the Southern Ocean and removed the signal of the Antarctic Circumpolar Wave (ACW) [*White and Peterson*, 1996; *Peterson and White*, 1998]. So we preserved this variability in the definitive version (HadISST1.1, referred to as HadISST1 here) by carrying out a separate RSOI analysis of the extratropical Southern Hemisphere using the in situ and bias-adjusted AVHRR SSTs for 1982 onward, and merging it with the quasi-global fields already created (section 3.7) from the in situ and AVHRR data further north. We used 2° area covariance EOFs based on data for 1982–98, but could not extend the Southern Ocean reconstruction prior to 1982 in this way, because the “global change” EOF (section 3.4.2) did not extend this far south.

[71] EOFs defined from 20°S to the ice-edge yielded reconstructions that were the most compatible in the area of overlap with the analysis for the rest of the globe. They also showed space-time variability in the Southern Ocean comparable with the results of *Reynolds and Smith* [1994] (see also section 6.3) as well as with the characteristics of the ACW reported by *White and Peterson* [1996]. The EOFs (e.g., Figure 9) filled the data void in the southeastern Pacific (Figure 4), as well as the Southern Ocean. Estimates of data error were as in Appendix B. Thirty-one EOFs representing around 80% of the variance of the input data set were selected. The Southern Ocean reconstructions were combined with the existing quasi-global reconstructions by averaging areas of overlap.

[72] Finally, the Poisson technique was used to fill small gaps between the RSOI/variance corrected in situ blend (after adding to the 1° area background climatology) at its high-latitude limits and SSTs associated with the sea ice edge.

5. Night Marine Air Temperature Analysis

[73] Night marine air temperature (NMAT) is air temperature measured from the decks of ships and from buoys at times when it is thought that solar heating of the deck is unimportant. NMAT has been used to monitor climate and detect its changes and to corroborate estimates of climatic variations made using land air temperature and/or SST [e.g., *Folland et al.*, 2001a]. It is useful for evaluating model simulations and assessing modeled ocean-atmosphere heat

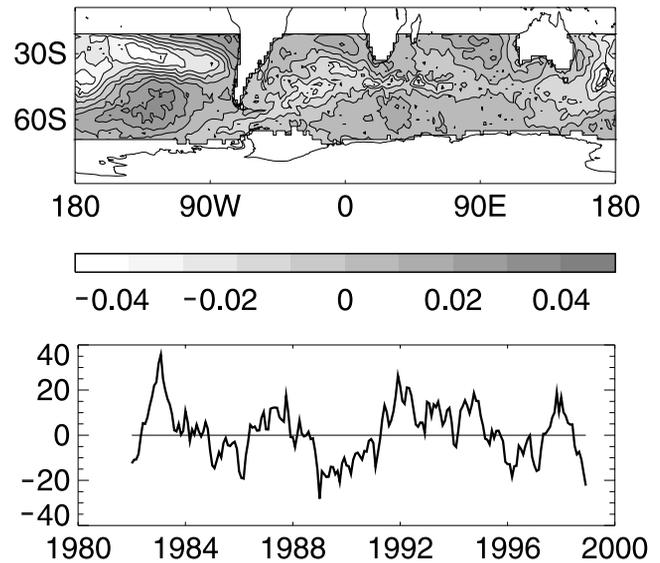


Figure 9. The leading EOF of SST variability south of 20°S in 1982–1998 and its projection onto the HadISST1 reconstruction for the same period. See color version of this figure at back of this issue.

fluxes. NMAT has been used with SST and land air temperature in regional studies of climatic variability [*Folland and Salinger*, 1995; *Folland et al.*, 1997].

[74] NMAT data are more difficult to work with than SSTs: there are substantially fewer of them; their temporal coherence within a month is less than that of SST, there are many uncertainties surrounding NMAT data collection practices and conditions, and changing practices have introduced biases to the data. Bias corrections have been developed for nonstandard observing practices during the second World War and during the 19th century [*Bottomley et al.*, 1990; *Parker et al.*, 1995b] and applied to the data set upon which our analysis is based. We develop a refinement and extension to the *Bottomley et al.* [1990] adjustment made to the data for changing ships’ deck heights (and therefore the heights of thermometer screens or aspirated psychrometers above the sea surface).

[75] To derive the maximum benefit from the available data, quasi-global monthly fields were reconstructed using the RSOI technique (see section 3.4.1 and *Kaplan et al.* [1997, 1998]), as used in HadISST1. We have preserved the local variability of NMAT, where there are sufficient observations, as we did with SST, by blending the reconstruction with the original gridded data and smoothing in a data-adaptive way (see Appendix E).

[76] Section 5.1 summarizes the data and the bias adjustments applied. Section 5.2 describes the analysis. Diagnostics of the resulting data set, HadMAT1, and comparisons with SST and land surface air temperatures, are presented in section 6.

5.1. Data and Bias Adjustment

[77] The Met Office historical marine air temperature (MOHMAT) data set [*Bottomley et al.*, 1990; *Parker et al.*, 1995b] is partitioned into day and night marine air temperature. Our analysis builds upon the monthly gridded

5° latitude-longitude (i.e., 5° area) NMAT data set MOH-MAT4N [Parker *et al.*, 1995b]. Night is defined to be the period between one hour after sunset and one hour after sunrise. This reduces bias in the data due to the evening persistence of warmth from solar heating of ships' decks.

[78] Daily normals linearly interpolated from pentad normals, created by harmonic synthesis of the monthly climatology, were used to quality control the observations within MOH-MAT4N: see Appendix I of Parker *et al.* [1995b]. These data were then corrected for systematic changes in the height of ships' decks, which have risen over time. All data were adjusted to the local average height over the 1961–1990 climatology period. The local height was assumed equal to the global average (Figure 10a) through 1970, but geographically varying heights were used thereafter based on 5-year average fields of deck heights centered on 1982 through 1995. Heights were linearly interpolated in 1971–81 and augmented at 0.14m/year after 1995 following global average trends based on the “height of the observing platform” field of WMO No. 47 [WMO, 1955–1999]. So, we differ from Bottomley *et al.* [1990] and Parker *et al.* [1995b], by applying continued adjustments up to the present (Figure 10b), entailing a global average adjustment of +0.05°C to values for the late 1990s, relative to 1961–1990 climatology. The adjustments are based on boundary-layer similarity theory [Fairall *et al.*, 1996; A. Grant, personal communication, 2001]. They are smaller than the cooling in tropical NMAT, relative to SST, reported by Christy *et al.* [2001] using data without this deck-height adjustment. The new adjustments add about 0.1°C to the overall warming of NMAT since the 1860s (Figure 10b) and reduce the divergence between Southern Hemisphere SST and NMAT trends in the most recent decade (section 6.5.1).

[79] Warm biases in NMAT due to nonstandard observing practices at night during 1939–1945 were corrected using day MAT (DMAT, where here day was defined to be from sunrise to sunset). NMAT data up to December 1941 were adjusted so that (NMAT - DMAT) equaled its local average for 1929–1938; NMAT data from 1942–1945 were adjusted so that (NMAT - DMAT) equaled its average for 1946–1955. This remains the most uncertain aspect of HadMAT1.

[80] During the period 1876–1893 in the Mediterranean Sea and North Indian Ocean, NMAT anomalies were high compared to SST anomalies and the rest of the historical NMAT record. This was probably due to the practice of piling cargo on deck rather than in the hold to avoid taxes at the Suez Canal, thus restricting the air flow around the air temperature thermometer [Bottomley *et al.*, 1990]. Bias-adjusted SST anomalies from MOHSST6 [Parker *et al.*, 1995b] were therefore used instead of NMATs. Similarly, between 1856 and 1885 in the Atlantic Ocean, NMAT anomalies were found to be high, particularly in windy conditions [Bottomley *et al.*, 1990]. Here, the calendar monthly average NMAT for 1856–1885 in each grid box was constrained to equal the same average for MOHSST6.

[81] NMAT data corrected as described above and upon which the HadMAT1 analysis was based are referred to hereafter as MOH-MAT43N; we also later show comparisons to MOH-MAT42N, which differs from MOH-MAT43N by including the earlier deck height corrections of Parker *et al.* [1995b].

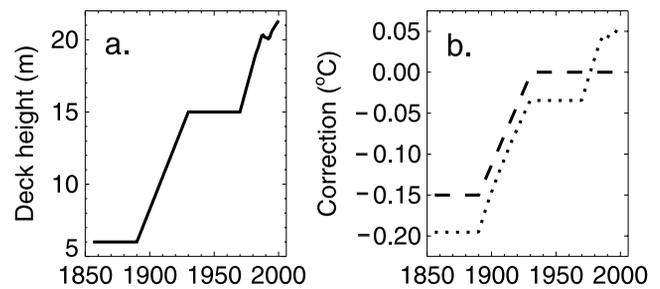


Figure 10. (a) History of globally averaged ships' decks' heights (in m). (b) Consequent global average adjustments to NMAT (°C) given by Bottomley *et al.* [1990] (dashes) and this study (dots). The new adjustments average to zero over 1961–1990.

5.2. Analysis Methodology

[82] The HadMAT1 analysis methodology is broadly similar to that of the HadISST1 SST fields. To create HadMAT1, monthly 5° area MOH-MAT4N anomalies were interpolated using the RSOI technique (see section 3.4). The 5° resolution allowed reconstructions to be made for the whole globe in one step. However, the EOFs were only defined over grid boxes that contained data both in at least 50% of the months during 1953–1997 and in 50% of the seasons during 1901–1997, the periods used to calculate the “global change” and “interannual” EOFs used. Many grid boxes remain unreconstructed: these are mainly in the Arctic, the southeast Pacific and the Southern Ocean and in parts of the equatorial Pacific. In these unreconstructed regions, HadMAT1 contains only un-interpolated MOH-MAT4N data, smoothed as described in Appendix E.

[83] The “global change” signal is described by the first EOF of a seasonal version of MOH-MAT4N for 1901–97 from which variability on timescales shorter than eight years had been removed. Its time series is correlated at 0.94 with the low-pass filtered global mean NMAT anomaly time series and has a spatial pattern that is similar to that of SST (Figure 4), but with negative weights in the majority of the north Pacific. This EOF was used to reconstruct and remove the global change signal from the NMAT anomalies before the interannual EOFs were calculated from these detrended monthly data for 1953–1997. The first 71 of these interannual EOFs, explaining 80% of the variance, were used to reconstruct complete fields from the detrended MOH-MAT4N. The trend component as defined by the global change EOF was added back to give the finished reconstructions.

[84] The original gridded MOH-MAT4N data were then superimposed on the RSOI reconstruction. The data-adaptive smoothing step (Appendix E) produced fields similar or equal to the original MOH-MAT4N data in areas of good data coverage, and a blend between the RSOI reconstruction and the original data elsewhere.

6. Diagnostics of HadISST1 and HadMAT1 and Comparisons With Other Analyses

[85] In this section we test the quality of HadISST1 and HadMAT1 and illustrate the advantages of HadISST1 over GISST and other published SST data sets. For SST we

include comparisons of climatologies and analyses of trends, standard deviations and autocorrelations of anomalies. For both SST and NMAT, we present time series of global and regional means. We also include a few selected analyses of AGCM simulations using HadISST1 and compare them against simulations using GISST3 to make specific points.

6.1. Comparison of SST Climatologies

[86] We compare the monthly climatology for 1971–2000 derived from HadISST1 with the adjusted OI.v2 1971–2000 climatology [Reynolds *et al.*, 2002]. Figure 11 shows the January and July differences, OI.v2 minus HadISST1.

[87] The most noticeable difference is in the Northern Hemisphere summer in areas that are well sampled by in situ data. Here, HadISST1 is systematically warmer (by up to about 0.7°C in places). This is symptomatic of a general difference between our data sets and COADS or OI.v2, manifested by a relatively enhanced annual cycle in our data sets; the northern Indian Ocean tends to be cooler in our data sets in these months.

[88] The HadISST1 and OI.v2 climatologies also differ in high-SST-gradient areas such as the Gulf Stream and the Malvinas Current regions, because of the different analysis resolutions. Here, HadISST1 is cooler than OI.v2 by over 1°C in summer. In the Southern Ocean, the highly variable differences between HadISST1 and OI.v2 may be due to the relatively short-term nature of the OI.v2 climatology: it is based on data for the satellite era with an adjustment to make it equivalent to a 1971–2000 climatology. As a result, the OI.v2 climatology can be expected to have more spatial detail than the HadISST1 climatology. The relative warmth of the OI.v2 climatology at high northern latitudes in summer away from sea ice is likely due to inadequate adjustment of this climatology to the 1971–2000 period here. HadISST1 and OI.v2 used the same sea ice analysis and sea ice to SST algorithm and so agree very well in sea ice covered regions.

6.2. Local Trends

[89] We analyzed four periods according to the character of the global mean temperature curve: 1871–1909, slight cooling; 1910–1945, warming; 1946–1975, little overall change of temperature; and 1982–1999, warming. Strictly speaking, the final warming period began in 1976 [Karl *et al.*, 2000], but we chose to begin our last period in 1982 because satellite SSTs, and hence OI.v2, are available from that date. Four data sets are compared: GISST2.3b, Kaplan *et al.* [1998, 2003], OI.v2 and HadISST1. The GISST2.3b anomalies are relative to the GISST2.2 1961–1990 climatology, the Kaplan *et al.* [1998, 2003] anomalies are relative to the Parker *et al.* [1995a] 1951–80 climatology and the OI.v2 and the HadISST1 anomalies are relative to the HadISST1 1961–1990 climatology. Trends and autocorrelations (section 6.4) are unaffected by these choices; standard deviations (section 6.3) are calculated using absolute SST.

[90] Figure 12 depicts restricted maximum likelihood [Diggle *et al.*, 1999] trends in $^{\circ}\text{C}$ per decade in each 5° area grid-box. The four periods exhibit very different patterns. 1871–1909 generally displays little spatial structure in the GISST2.3b, Kaplan *et al.* [1998, 2003] (neither are shown) or HadISST1 trends (Figure 12a). The two periods of pronounced warming: 1910–1945 and 1982–1999, show

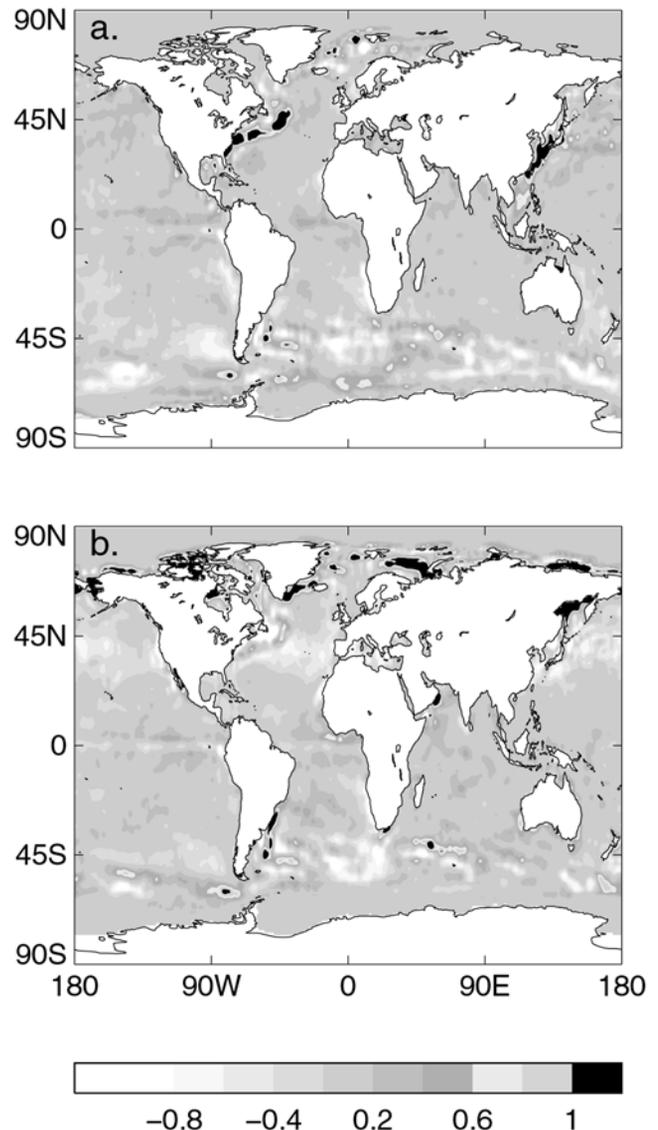


Figure 11. Sea surface temperature climatology (1971–2000) differences ($^{\circ}\text{C}$): adjusted OI.v2 minus HadISST1, for (a) January and (b) July. See color version of this figure at back of this issue.

some similarities (Figures 12b–12d and 12f–12i), but the magnitudes of the warming trends are often much larger in the more recent period. Despite the larger warming trends, there are areas of cooling in all data sets in the 1982–1999 period but no consistent cooling regions in 1910–1945. The trends are in close agreement with the oceanic trends from HadSST presented by Folland *et al.* [2001a].

[91] The warming between 1910 and 1945 is reduced in HadISST1 relative to GISST and is more like that of Kaplan *et al.* [1998, 2003] in the Atlantic, but does not contain so many areas of negative trend in the Pacific.

[92] In HadISST1 between 1946 and 1975 (Figure 12e) we see large areas of cooling to the south and southeast of Greenland. However, the warming here between 1982 and 1999 has been very strong (Figures 12f–12i). This is less well represented by Kaplan *et al.* [1998, 2003] (Figure 12j). The spatial pattern of the first EOF of the low-pass filtered 4°

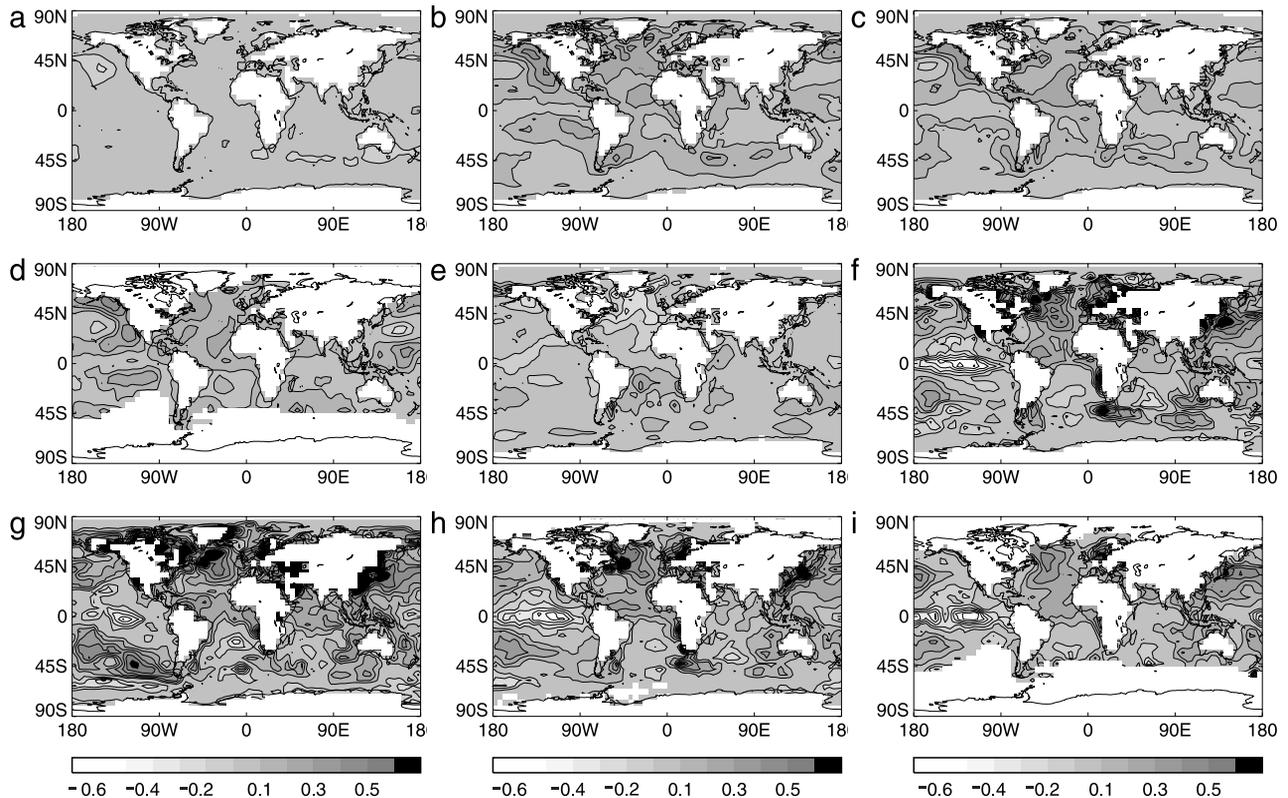


Figure 12. Restricted maximum likelihood linear trends in $^{\circ}\text{C}$ per decade in each 5° area grid-box in: (a) 1871–1909 (HadISST1); (b–d) 1910–1945 (GISST2.3b, HadISST1, and *Kaplan et al.* [1998, 2003], respectively); (e) 1946–1975 (HadISST1); (f–i) 1982–1999 (OI.v2, GISST2.3b, HadISST1, and *Kaplan et al.* [1998, 2003], respectively). See color version of this figure at back of this issue.

area in situ/AVHRR blend for 1901–1997 shown in Figure 4 has an area of cooling relative to much of the rest of the world’s oceans to the southeast of Greenland, resulting mainly from the cooling between 1946 and 1975 [*Parker et al.*, 1994; *Folland et al.*, 1999; *Hansen et al.*, 1999].

[93] The marked area of cooling along the equator in the eastern Pacific since 1982 in OI.v2 is poorly reproduced in GISST2.3b (Figures 12f and 12g, respectively), which did not have an equatorially centered grid to resolve the cold tongue. The cooling has a more realistic shape in HadISST1 (Figure 12h).

[94] The large warming trend in GISST3.0 in the southeast Pacific since 1982 may have resulted from sparsity of in situ data to adjust the AVHRR locally. The Southern Ocean EOF analysis in HadISST1 (section 4.2) appears to have rectified this and yielded a result similar to OI.v2.

6.3. Standard Deviations

[95] We examine fields of grid box standard deviation in the same multidecadal periods and a time series of global average standard deviation to investigate how homogeneous HadISST1 is through time (Figure 13).

[96] Fields of standard deviations were calculated separately for each calendar month in each of the four periods, using detrended data from OI.v2, GISST3.0 and HadISST1. In the Gulf Stream, Kuroshio and Malvinas Current regions, variability is much smaller in HadISST1 than in either GISST2.3b or 3.0 in both 1871–1909 and 1910–1945

(not shown). In later periods the reduction is less, but the standard deviation is less than that of the OI.v2 in 1982–1999 (Figure 13). In the Indian Ocean, HadISST1 has less variability than GISST 2.3b or 3.0 on monthly timescales throughout the record. Standard deviations may have differed because of: damping by the RSOI to prevent overfitting to sparse data; blending of only variance-corrected MDB in situ data with the reconstruction in HadISST1, whereas in GISST3.0 unadjusted MDB and COADS data were used; using large-scale EOFs to exploit teleconnections between ocean basins but thereby losing small-scale variability which was not adequately replaced by the blend with variance corrected in situ data (in GISST, the EOFs were based on individual ocean basins [*Rayner et al.*, 1996]); use of quasi-global or Indian/Pacific Ocean EOFs, leading to reduced variability in the Indian Ocean through truncation of a greater proportion of the variability in the Indian Ocean (but, the apparent Indian Ocean dipole pattern shown by *Saji et al.* [1999] and *Webster et al.* [1999] is evident in HadISST (not shown)); variability in data-sparse periods of GISST2.3b and 3.0 may have been artificially enhanced by the EOF projection technique, which overfits to outliers (section 3.4.1); projection of the EOFs onto three month average data when data were sparse (section 3.4.2), in lieu of the temporal model in the Reduced Space Optimal Smoothing (RSOS) technique [*Kaplan et al.*, 1998] (but, we did the same in GISST2.3b and 3.0, so any changes in standard deviation due to this procedure should have been minimal).

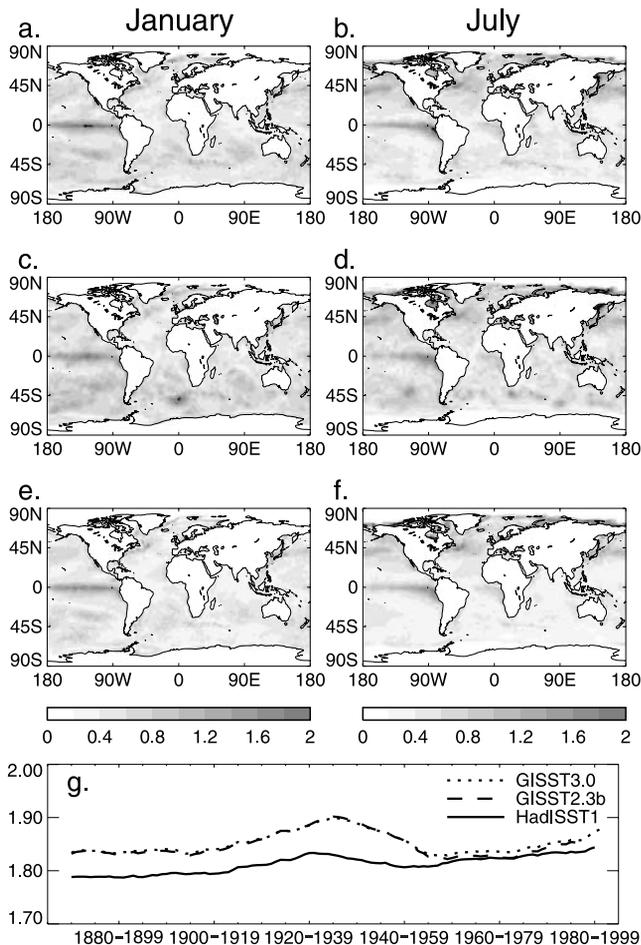


Figure 13. SST standard deviations ($^{\circ}\text{C}$), 1982–1999 in: (a–b) OI.v2; (c–d) GISST3.0; and (e–f) HadISST1 (left panels: January and right panels: July). (g) Time series of global root mean square 1° area standard deviation in running 20-year periods for GISST2.3b (dashed), GISST3 (dotted), and HadISST1 (solid). See color version of this figure at back of this issue.

[97] On the other hand, the variability in the eastern equatorial Pacific appears to have been maintained in the period 1871–1909 in HadISST1 (not shown). The shape of the eastern equatorial Pacific variance maximum in HadISST1 in 1982–1999 is more like that of OI.v2 than that of GISST (Figure 13).

[98] Before 1982, both HadISST and GISST lack variability in the Southern Ocean relative to later years owing to lack of data, but the true variability cannot be assessed. In general, the standard deviation fields for HadISST1 for 1982–1999 show an improvement over GISST3.0, as they are more coherent and sharply defined. However, especially in the Southern Hemisphere winter, variability in HadISST1 appears reduced relative to the OI.v2. The locally high values in Figures 13c and 13d in GISST in the Southern Ocean resulting from poor data are not seen in HadISST1.

[99] Figure 13g shows time series of global root mean square average standard deviation for GISST2.3b, GISST3 and HadISST1. The averages are calculated from fields of 1° area grid box standard deviation for overlapping 20-year periods: 1871–1890, 1872–1891, ..., 1980–1999. The

time series for HadISST1 is remarkably homogeneous and more consistent through time than those of GISST2.3b and 3.0. HadISST1 has reduced variance prior to the 1920s, possibly as a result of the tendency of the RSOI analysis to reconstruct zero anomaly when data are particularly sparse. The peak in standard deviation in the middle part of the GISST record has been replaced by a weak maximum, which may relate to the relative abundance of data in the 1920s and 1930s (Figure 3). Because HadISST1 anomaly fields were reconstructed on 4° area resolution prior to 1949 and on 2° area resolution thereafter, slightly higher standard deviations would be expected in recent years.

6.4. Autocorrelation of SST Anomalies

[100] Hurrell and Trenberth [1999] observed that in 1982–1997 the month-to-month persistence of SST anomalies in GISST2.3b (identical with GISST3 in this period) was much lower than that of the Reynolds and Smith [1994] OI SST. Figure 14 shows the one-month lag autocorrelation of 2° area detrended SST anomalies in 1982–1999 in GISST3.0, OI.v2 and HadISST1. The greatest lack of temporal coherence in GISST (Figure 14a) is found in the Southern Ocean where in situ data are sparse. The fields of autocorrelation for GISST3.0 in the other periods: 1871–1909, 1910–1945 and 1946–1975 (not shown) have rather similar characteristics to that for 1982–1999. Here the incoherence is likely to be due to the blending of the reconstructions with the sparse non-variance-corrected in situ data.

[101] The increase in persistence in HadISST1 relative to GISST3.0 is great in 1982–1999 (Figure 14c). Autocorrelations exceed 0.8 in much of the tropics and in many parts of the Southern Ocean. The increase in autocorrelation is smaller in the North Atlantic where in situ data are most plentiful. (The autocorrelation of detrended fields comprised only of reconstructed SST was far too high as the RSOI reconstruction alone contains insufficient small-scale variance.) HadISST1 is now slightly more persistent than OI.v2 (Figure 14b), particularly in the Indian Ocean, where intermonthly variance is reduced in HadISST1 (section 6.3).

[102] The intermonthly autocorrelations in HadISST1 are weakest in 1910–1945 (not shown). A contributing factor may be the very data-sparse periods 1914–1920 and 1940–1945. Coverage in 1871–1909 was often sparse also, but no year had as few data as 1918 (Figure 3). In addition, the El Niño–Southern Oscillation phenomenon, which engenders strong monthly persistence in the tropics and some extratropical regions, was strong and coherent in the late 19th century and generally weak and less coherent between roughly 1920 and 1940 [e.g., Allan *et al.*, 1996].

6.5. Global and Regional Average Time Series

[103] We now compare annual and near-decadal averages of several SST and NMAT data sets over a number of regions. Unless otherwise stated, we use all available data, i.e., data are not collocated. However, we exclude grid boxes partially covered by sea ice. All time series are expressed relative to their respective averages over 1961–1990.

[104] We also calculate equivalent linear trends in the global and hemispheric averages, using a similar method to that used by Folland *et al.* [2001a], but excluding explicit consideration of the uncertainties in the annual values.

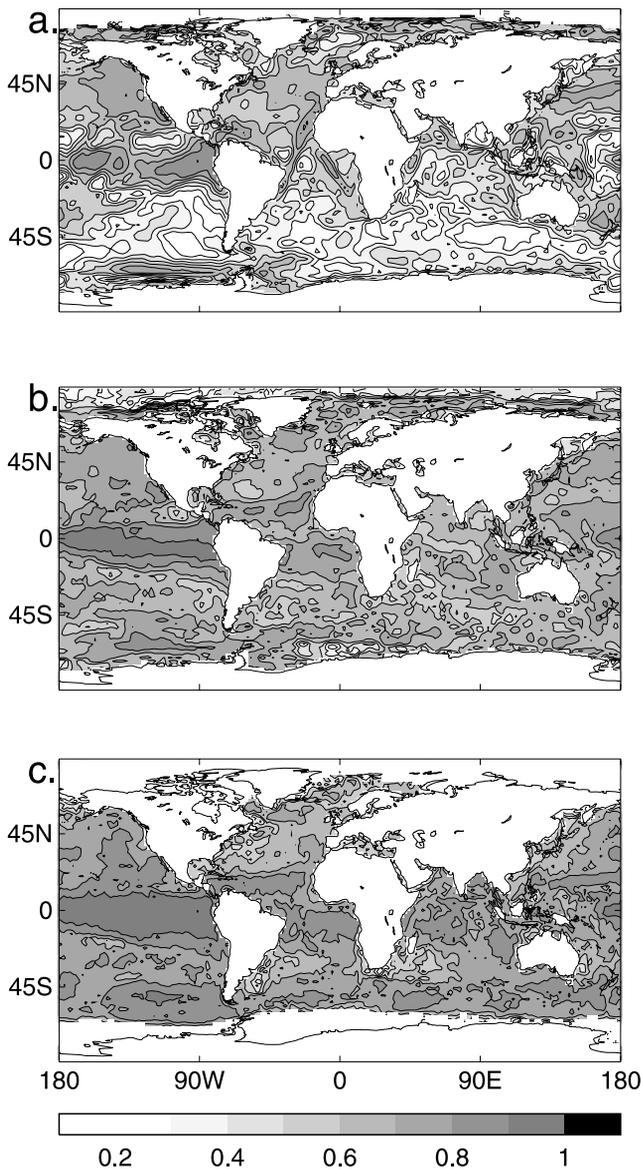


Figure 14. One-month lag autocorrelation of detrended monthly sea surface temperature anomalies in each 2° area grid-box, 1982–1999, for (a) GISST3.0, (b) OI.v2, and (c) HadISST1. See color version of this figure at back of this issue.

6.5.1. Globe and Hemispheres

[105] The global temperature trend over the period 1901–1999 is slightly, but not significantly, weaker in HadISST1 than in MOHSST6D, GISST3.0 and HadSST1 (Table 2).

However, the global, Northern Hemisphere, Southern Hemisphere and Atlantic trends for HadISST1 are very close to those for HadSST1 when only collocated data are included (not shown).

[106] HadISST1 is generally around 0.05°C warmer than MOHSST6, GISST3.0, and HadSST1 in the late 19th and early 20th centuries (Figure 15a), although generally cooler than Kaplan *et al.* [1998, 2003] in this period (see section 3.4.2).

[107] In line with the global and hemispheric means, HadISST1 zonal mean anomalies for 1982–1999 are generally less than in GISST, but agree very well with those of OI.v2. So our treatment of AVHRR SST (section 3.7) may have given cooler SSTs relative to those in MOHSST6 and GISST during this period. The trend in the global average SST anomaly in both GISST3.0 and HadISST1 between 1982 and 1999 increases by 0.02°C/decade when the Southern Ocean is excluded (see final two rows of Table 2). The Southern Ocean is an area for which there are few in situ data to provide anchor points for the adjustment of AVHRR data so there could be a cold bias here. However, the differences between the two sets of trends for 1982–99 in Table 2 are not statistically significant.

[108] In the data-sparse periods 1914–1920 and 1940–1945, HadISST1 succeeds in avoiding some unrealistically large negative monthly anomalies that affected GISST, owing to overfitting of the EOFs to the few available data in the EOF projection method. In particular, 1941 was poorly represented in GISST, but looks more realistic in HadISST1 (not shown).

[109] On a global average (Figure 15b), one of the most striking differences between HadMAT1 and MOHMAT43N was during the 1900s and 1910s, when the global average temperature anomaly was at its minimum; the HadMAT1 global average is about 0.1°C warmer and looks much more like that of HadISST1 and GISST3.0. However, when a collocated comparison was made, no difference was seen between the data sets in the global or hemispheric curves. We included only values between 80°N and 65°S, as the MOHMAT4 climatology was of very poor quality at high latitudes, resulting in spurious large anomalies.

[110] The HadMAT1 anomaly is about 0.1°C warmer than SST in the early 1940s in the Northern Hemisphere (Figure 15d), but this is a reduction of around 0.05°C from the relative anomaly in MOHMAT42N remarked upon by Folland *et al.* [2001a]. HadMAT1 is seen to warm relative to SST over the last two decades of the 20th century in the Northern Hemisphere and global averages, in line with global land air temperature changes.

[111] MOHMAT43N cools relative to HadMAT1 in the Southern Hemisphere in the 1990s (Figure 15f). We include

Table 2. Global Temperature Trends (and Their 2σ Uncertainties) in Surface Temperature Data Sets^a

Period	MOHSST6	GISST3.0	HadSST1	HadISST1	MOHMAT42N	MOHMAT43N	HadMAT1
1871–1999	0.04 ± 0.02	0.04 ± 0.02	0.04 ± 0.02	0.04 ± 0.00	0.04 ± 0.02	0.05 ± 0.02	0.04 ± 0.02
1901–1999	0.06 ± 0.02	0.06 ± 0.02	0.06 ± 0.02	0.05 ± 0.02	0.06 ± 0.02	0.07 ± 0.02	0.06 ± 0.02
1871–1909	−0.04 ± 0.06	−0.03 ± 0.04	−0.02 ± 0.04	−0.02 ± 0.04	−0.04 ± 0.08	−0.04 ± 0.08	−0.03 ± 0.06
1910–1945	0.16 ± 0.04	0.12 ± 0.02	0.15 ± 0.04	0.12 ± 0.04	0.18 ± 0.06	0.18 ± 0.06	0.15 ± 0.04
1946–1975	0.01 ± 0.08	0.03 ± 0.08	0.01 ± 0.04	0.02 ± 0.04	−0.02 ± 0.04	−0.01 ± 0.04	−0.01 ± 0.04
1982–1999	0.16 ± 0.06	0.16 ± 0.06	0.15 ± 0.06	0.13 ± 0.06	0.10 ± 0.08	0.11 ± 0.08	0.15 ± 0.08
1982–1999 (EOF region)	0.17 ± 0.08	0.18 ± 0.08	0.15 ± 0.08	0.15 ± 0.08	0.12 ± 0.08	0.13 ± 0.08	0.15 ± 0.08

^aValues are in °C per decade and are given to two decimal places. The trends were calculated from annual global averages of gridded temperature anomalies (relative to 1961–1990) using the restricted maximum likelihood technique [Diggle *et al.*, 1999]. All available data are used unless otherwise stated.

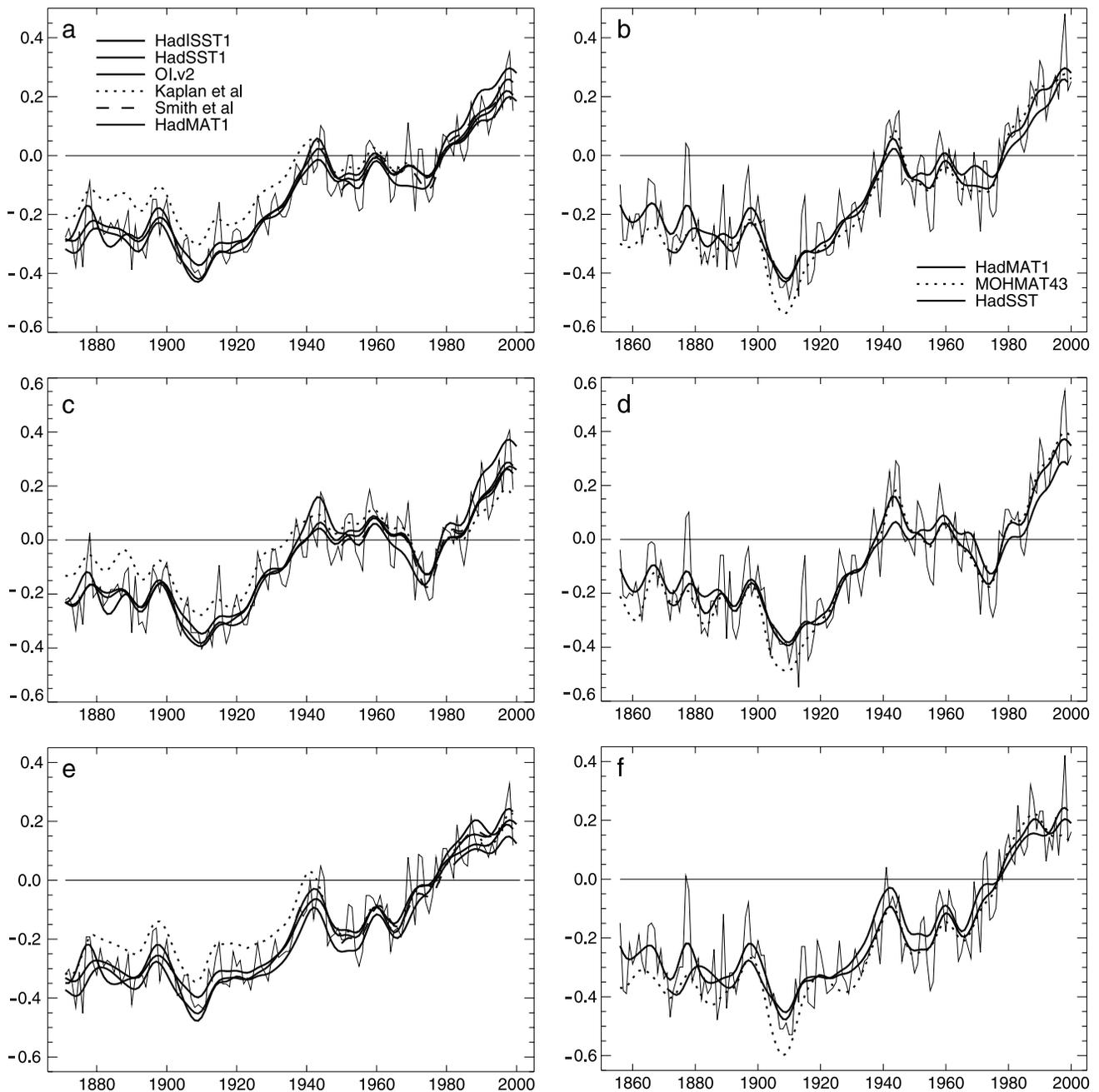


Figure 15. Smoothed hemispheric and global average temperature anomalies, relative to 1961–1990. Left panels: HadISST1, other SST data sets and HadMAT1, 1871–2000. Right panels: HadMAT1, MOHMAT43N and HadSST, 1856–2000. (a and b) Globe; (c and d) N.Hem.; and (e and f) S.Hem. All grid-boxes with data are used. A 21-point binomial (near-decadal) filter was applied to annual anomalies. Thin curves are unsmoothed annual averages: HadISST1 (left) and HadMAT1 (right). “Smith et al.” refers to *Smith et al.* [1996], and “Kaplan et al.” refers to *Kaplan et al.* [1998, 2003]. See color version of this figure at back of this issue.

all available data in our time series, which for MOH-MAT43N includes values between 55° and 65° S where HadMAT1 has no data. Global trends for the period 1982–1999 (last two rows of Table 2) show that if these high-latitude data are excluded, the trends in the MOH-MAT43N global averages increase by $0.02^{\circ}\text{C}/\text{decade}$. Accounting for the effect of rising deck heights on NMAT has not removed all of the divergence of the Southern Hemisphere NMAT

and SST time series after 1991 seen in the study by *Folland et al.* [2001a] (compare HadSST and HadMAT1 curves in Figure 15f), but the residual differences look less unusual.

6.5.2. Selected Regions

[112] In the Gulf Stream region (Figure 16a), the data sets appear to be in generally good agreement overall, except for *Kaplan et al.* [1998, 2003], which exhibits a weaker trend

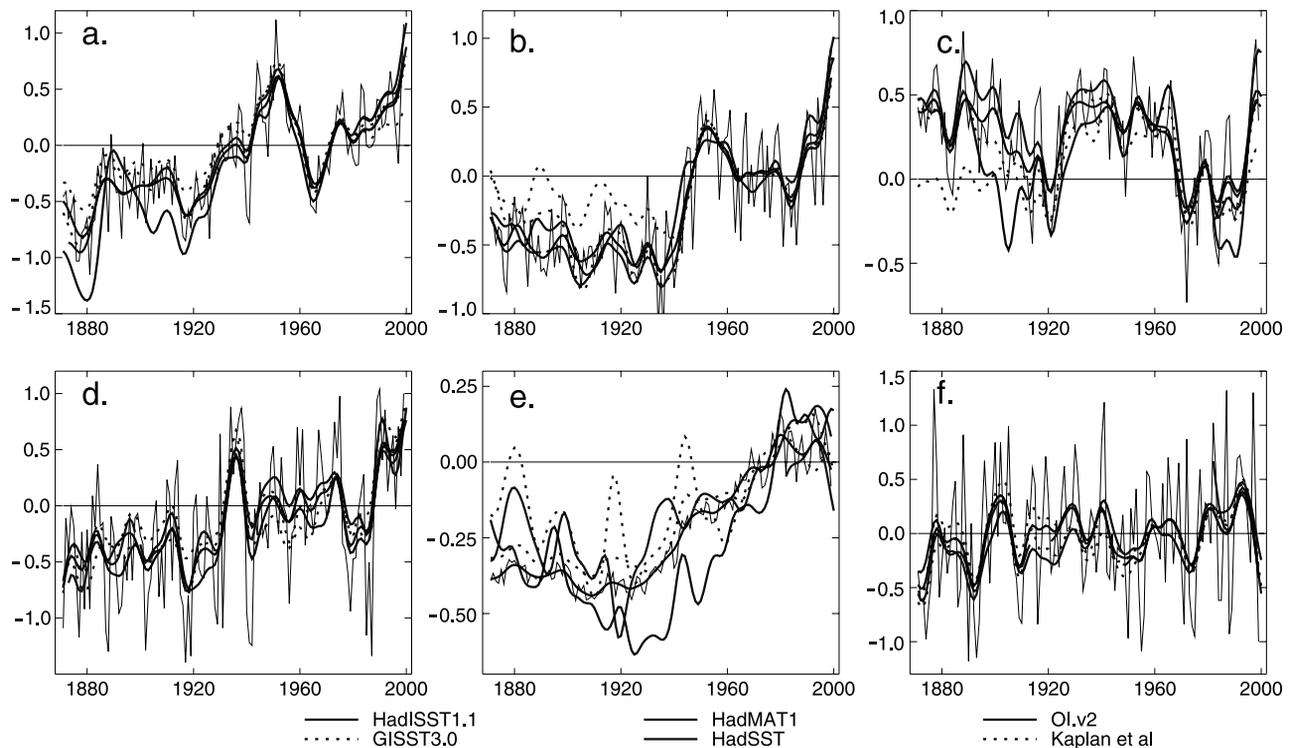


Figure 16. As left panels of Figure 15 (i.e., using all ice-free grid boxes with data) but for: (a) Gulf Stream region (35°N to 45°N, 50°W to 70°W); (b) Kuroshio region (30°N to 40°N, 125°E to 160°E); (c) Greenland region (50°N to 70°N, 30°W to 70°W); (d) Baltic Sea; (e) Southern Ocean south of 50°S; (f) Pacific Niño 3.4 region (5°N to 5°S, 120°W to 170°W). Thin black curves are unsmoothed HadISST1 annual anomalies. “Kaplan et al.” refers to *Kaplan et al.* [1998, 2003]. See color version of this figure at back of this issue.

and HadMAT1, which has a stronger trend. OI.v2 is noticeably warmer than the other SST data sets in the Kuroshio region (Figure 16b). Here, again, *Kaplan et al.* [1998, 2003] is much warmer than HadISST1 before 1940.

[113] In the Greenland region, HadISST1 agrees more closely with HadSST than does GISST3.0 (Figure 16c). There is a high level of consistency between data sets in the Baltic region (Figure 16d).

[114] In the high-latitude Southern Hemisphere (50° to 90°S, Figure 16e), HadISST1 compares generally well with HadSST, but is cooler pre-1900 and, like OI.v2, is cooler from 1982 onward, possibly owing to the use of under-corrected AVHRR SSTs. Note the increased variability from 1982 onward when the Southern Ocean was explicitly reconstructed.

[115] In tropical Pacific regions, exemplified by Niño 3.4 (Figure 16f), GISST3.0 is generally about 0.2°C cooler than HadISST1 between 1915 and 1949, although the timing of this difference changes with region owing to the different temporal distribution of the availability of data. In the west tropical Pacific, HadISST1 and *Kaplan et al.* [1998, 2003] are warmer than GISST3.0 from the early 1980s onward, and match MOHSST6 better (not shown). In Niño 3.4, HadISST1 is warmer in 1878 than in GISST3.0 and about as warm as *Kaplan et al.* [1998, 2003]. These differences reflect the varying influences of patchy, sparse data on the analyses.

[116] The unfiltered annual Southern Hemisphere average shows a cooling of around 0.1°C in HadISST1 relative to GISST3.0 between 1990 and 1994 (not shown). This offset

is most marked in the Indian and South Pacific Oceans and in the area around New Zealand. The offset is concentrated in the 30° to 55°S region and ends abruptly at the end of 1994. A very similar sequence arises if MOHSST6 is subtracted from HadISST1, or if GISST3.0 is subtracted from OI.v2, or MOHSST6 subtracted from bias-adjusted AVHRR SST anomalies (although here the signal is much larger, see Figure 6). Several of the main bias peaks appear in the austral summer months and match the nighttime AVHRR calibration error in this region at this time discussed in section 3.6. This suggests that the bias-correction procedures applied to HadISST1 and the OI.v2 were unable to adequately correct for these particular biases, which occurred in regions with extremely sparse in situ data. However, if so, what remains in HadISST1 is a relatively small residual bias.

6.6. Comparison of SST and NMAT With Collocated Land Air Temperature

[117] Decadally averaged differences between coastal and island land air temperatures (LAT) [*Jones et al.*, 2001], and MOHMAT42N, HadMAT1 and HadISST1 are shown in Figure 17. We use 5° area grid boxes that have both NMAT or SST and LAT values. The LAT minus HadMAT1 Northern Hemisphere differences are smaller than LAT minus MOHMAT42N, particularly prior to about 1890, when they are now near zero. In the Southern Hemisphere, LAT minus HadMAT1 differences remain slightly positive before 1960 (due principally to problems in Australian LAT [*Folland et*

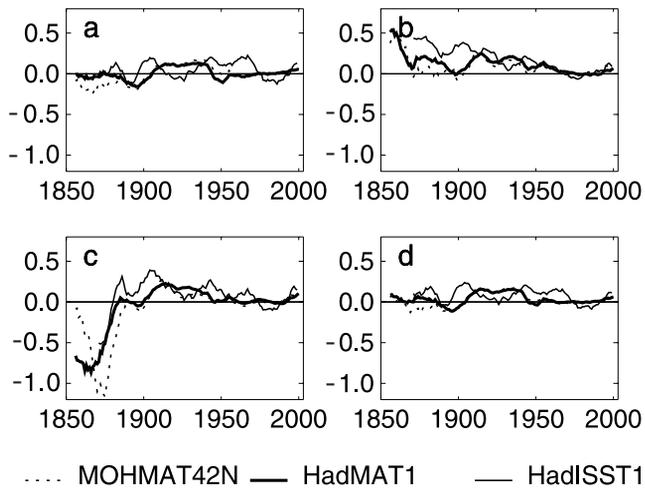


Figure 17. 10-year running mean differences between coastal and island land air temperature anomalies and collocated SST and NMAT anomalies ($^{\circ}\text{C}$, relative to 1961–1990) for 1856–2000: (a) Northern Hemisphere; (b) Southern Hemisphere; (c) Tropics (20°N to 20°S); and (d) Globe.

al., 2001b]) but the difference in the early 1990s is reduced to almost zero, lending weight to the argument that the residual relative differences between NMAT and SST then are real. In the Tropics, HadMAT1 matches LAT better (by about 0.5°C) than MOHMAT42N does between 1865–1885 but the reverse holds before 1865; these results are very uncertain because of scarcity of data. The global average LAT minus HadMAT1 and LAT minus MOHMAT42N differences are very similar after 1885, but the former are smaller prior to this date, following the result for the relatively data-rich Northern Hemisphere. Differences between SST and collocated land air temperature are broadly similar, though some details of their temporal evolution differ.

6.7. Test of SST Reconstructions Using AGCM Simulations of the SOI and ACW

[118] As the primary purpose of HadISST is to force AGCMs, it is important to verify whether the performance of an AGCM has been improved compared to forcing by earlier data sets. Forcing an AGCM with a new analysis and assessing its performance is an important complementary way of evaluating such a data set.

[119] Two six-member ensembles of simulations of climate since 1871 were performed with the HadAM3 [Pope *et al.*, 1999] atmosphere-only GCM. The first ensemble was forced using the GISST3.1 data set (a modification of GISST3.0 with more homogeneous sea ice), the second with the HadISST1 data set. Here we discuss two features of these simulations: the Southern Oscillation Index (SOI) and the Antarctic Circumpolar Wave (ACW) [White and Peterson, 1996].

6.7.1. Southern Oscillation Index

[120] The SOI has a high signal-to-noise ratio relative to any extra-tropical indicator. Figure 18 shows the observed SOI time series (the pressure at Tahiti minus that at Darwin, then normalized for each calendar month with respect to

1961–1990) and the ensemble mean simulated SOIs from our two sets of model runs. The correlation between the observed and ensemble mean simulated time series is 0.59 (0.51) for the HadISST1 (GISST3.1) runs. This improvement is significant at the 1% level, assessed using a Monte Carlo test.

[121] Thick horizontal bars in Figure 18 indicate the events where the largest improvement in the skill of the ensemble mean SOI was obtained. Replacing the ensemble mean time series derived from the GISST runs by that from the HadISST1 runs just over these periods almost equalized the correlations. Thus an improvement in the simulation of several individual events has created the improved correlation. In particular, the La Niña of 1917/18 has been improved in HadISST1 by placing less weight on suspect observations in the north Pacific. The improvement in the simulation of the protracted warm event in the 1990s is likely to be a consequence of the improved temporal persistence in HadISST1 over the last two decades.

6.7.2. Antarctic Circumpolar Wave

[122] We took particular steps to improve our analysis of the Southern Ocean in HadISST1 after 1981 (section 4.2). Figures 19a and 19b depict the signal of the ACW in the HadISST1 and OI.v2 analyses. The meridional average SST anomalies for 55° to 57°S are band-pass filtered to show variations in the range 3–7 years and show a very similar traveling wave in each case. The variance is less in HadISST1, because of the higher resolution of the OI.v2 analysis. Figure 19c shows the ACW signal in the ensemble mean MSLP field from the HadISST1-forced simulations. Comparison of this with Figure 1 of White and Peterson [1996], which was derived from ECMWF analyses, shows qualitative improvements in the size of the MSLP anomalies relative to those in the GISST3.1-forced simulations (not

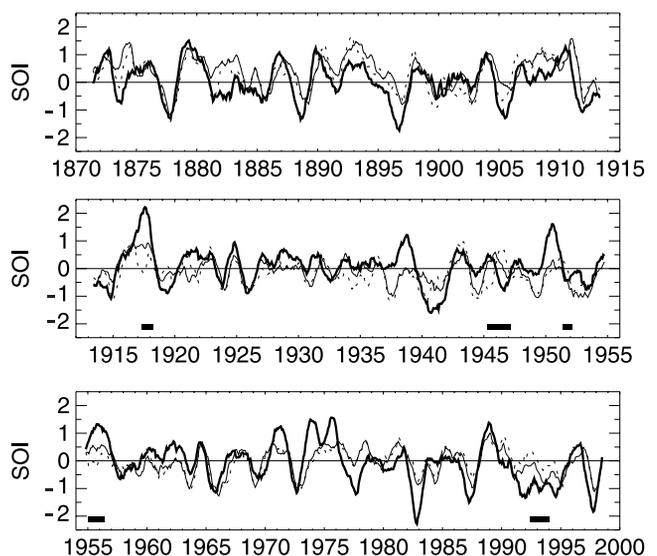


Figure 18. Time series of seasonal Southern Oscillation Index 1871–1998, using an 11-point filter: observed (thick solid), and simulated by HadAM3 using GISST3.1 (dotted) and HadISST1 (thin solid). Horizontal bars identify events where greatest improvement in skill is gained by using HadISST1 (see text).

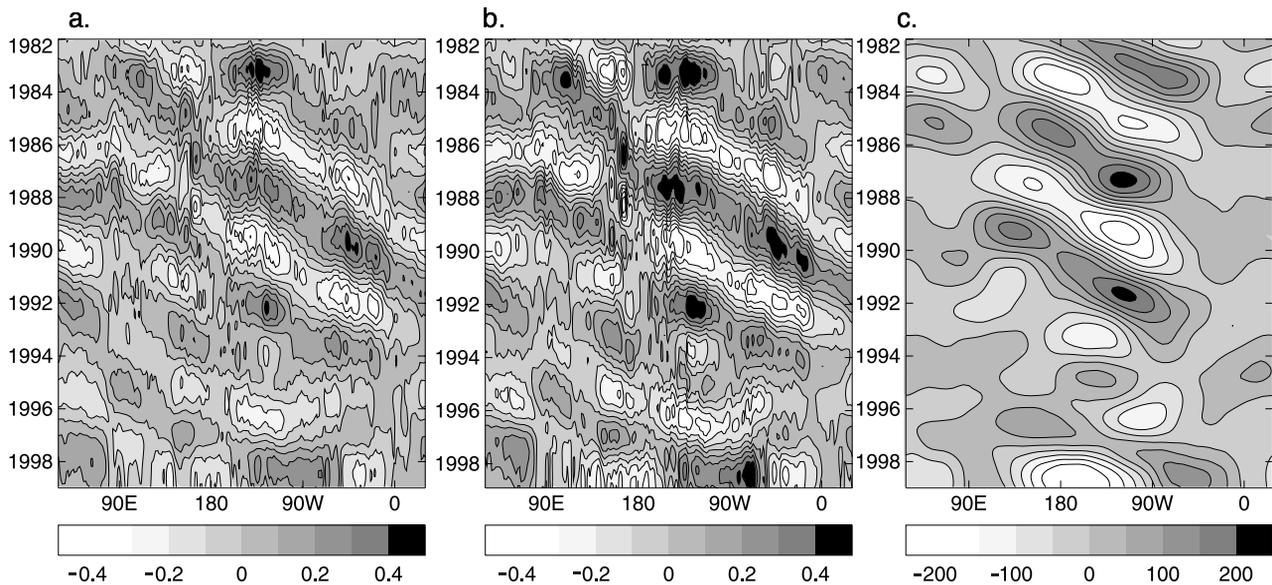


Figure 19. Monthly anomalies averaged over the belt 55° to 57° S, 1982–1998, 3–7 year band-pass filtered to show the Antarctic Circumpolar wave [White and Peterson, 1996]: (a) HadISST1 SST ($^{\circ}$ C); (b) OI.v2 SST ($^{\circ}$ C); and (c) mean sea level pressure (hPa) simulated by HadAM3 forced with HadISST1. See color version of this figure at back of this issue.

shown). Prior to 1982, the ACW is poorly represented in HadISST1 owing to data sparsity.

7. Discussion and Conclusions

[123] HadISST1 incorporates major improvements over its GISST predecessors. The broad-scale SST analysis in HadISST is based on RSOI instead of the less stable EOF projection technique, used in GISST2 and 3. We have improved the representation of local detail of SST, by again superimposing the original gridded SST data, but now with suppressed random sampling and measurement errors. This provides a unique globally complete analysis of historical SST data since 1871. Recent satellite-based estimates of sea ice concentration have been modified to compensate for the impact of melt ponds and wet snow on passive microwave sensors. The earlier sea ice concentration record was adjusted to be homogeneous with the modified satellite record, largely removing artificial jumps from the time series. Finally, the statistical procedure for estimating SST in sea ice zones has been refined, improving consistency with limited observational data for the central Arctic ice-pack and providing a close integration of nearby SST and sea ice information. HadISST1 succeeds in capturing both regional and large-scale variations in SST trends, through the use of the two-stage analysis process. The intermonthly autocorrelation of the HadISST1 SST fields has been improved over that in GISST and their variance is remarkably homogeneous throughout the record. An ensemble of AGCM simulations forced using HadISST1 produces an SOI time series that correlates significantly better with observations than the SOI from an ensemble of simulations forced with GISST3.1. The Antarctic Circumpolar Wave in SST is well represented in HadISST1 since 1982 and the resultant pattern in mean sea level pressure is well captured by the HadISST1-forced AGCM runs. Variance “bull’s-

eyes” seen in GISST in the Southern Hemisphere have been avoided in HadISST1. Differences between HadISST1 and the OI.v2 from 1982 onward are mostly due to the higher spatial and temporal resolution of the OI.v2 analysis. However, there is an enhanced annual cycle in HadISST1 relative to the OI.v2 which is likely due to the differing quality control procedures applied to the input data; it is not clear which is better. HadISST1 SST fields have reduced variance relative to the OI.v2 and GISST in some areas, especially the Indian Ocean. There is likely to be a small residual cool bias in HadISST1 (and in OI.v2) in situ data-sparse areas resulting from the inclusion of under-corrected cool-biased AVHRR data.

[124] The use of RSOI in HadMAT1 suppresses much of the excessive noise in sparsely observed periods and regions found in MOHMAT43N and allows coverage of much of the oceans. Increased data coverage has been obtained without compromising the real variability of the original observed data. We have also applied revised corrections to HadMAT1 derived from new information detailing how ships’ deck heights have changed through time. The resulting global and hemispheric time series in HadMAT1 agree more closely with SST than did MOH-MAT42N through the early part of the record. Latterly, HadMAT1 shows faster warming than MOH-MAT42N, globally much the same as SST. However, there remains a small cooling in NMAT relative to SST in the Southern Hemisphere from the early 1990s onward, only partially ameliorated by the revised deck height corrections, which needs more investigation.

[125] Both HadISST1 and HadMAT1 are now updated in near real time every month.

[126] Planned developments to HadISST and HadMAT include a substantial strengthening of the basic data through the use of the new I-COADS data set [Diaz *et al.*, 2002]. This is expected to bring particular improvements to the

analyses around the 1910s and may also allow the extension of HadISST back to the 1850s. Work is in hand to allow robust estimates of analysis error in each grid box and to explore the production of a submonthly analysis from around 1950 onward.

[127] Corrections to the pre-1942 SSTs for the use of uninsulated buckets will be modified if the newly incorporated data require this. However, recent atmospheric model simulations of global land surface air temperature forced with GISST3.1 with and without the existing bucket corrections suggest that these corrections have very good skill on an annual average [Folland *et al.*, 2001b]; the validity of their seasonal variation is currently being tested in the same way. So far, no adjustments have been made to post-1941 SSTs. Research is under way to assess the need for adjustments to modern SSTs and marine air temperatures to compensate for possible biases in the measurement techniques, although we believe these problems are considerably less serious than those currently corrected before 1942.

[128] HadISST1 is not of sufficiently high spatial resolution to resolve very localized SST features or the meanderings of the Gulf Stream. These aspects may be improved in recent, data-rich years when bulk-adjusted ATSR-1, -2 and Advanced ATSR (AATSR) SST data since 1991 are assimilated into HadISST. These (A)ATSR data, along with newly developed retrievals of SST from microwave radiometers [Wentz *et al.*, 2000], may also help to resolve some of the remaining biases in the AVHRR data. Work is also currently under way to use 1 km resolution ATSR-2 data to investigate the detailed variations of the surface temperature of open water between ice floes in both the Arctic and the Antarctic. This will help to improve the SST/sea ice concentration relationships currently used to specify SST in sea ice-covered regions in HadISST.

[129] The early sea ice data will be augmented. Data sources have been digitized by the Norwegian Polar Institute for some regions in the North Atlantic prior to 1901 [Walsh and Chapman, 2001; Løyning *et al.*, 2003] (see documentation and data archive available at <http://acsys.n-polar.no/ahica/intro.htm>). The WMO Global Digital Sea Ice Data Bank project digitized data for the Russian and Canadian Arctic, and data for other areas have also recently become available. Antarctic data for the 1950s and 1960s in Russian archives (V. Smolyanitsky, personal communication, 2002) may also be utilized. Efforts will be made to improve the bias-corrections applied to the summertime passive microwave data in the Arctic, while including information on surface melt to make fluxes into AGCMs more realistic.

[130] HadISST1 and HadMAT1 are freely available to researchers worldwide. Details of access to the data are given at <http://www.metoffice.com>.

Appendix A: Homogenization and Combination of Sea Ice Concentration Data in HadISST1

[131] The constituent sea ice concentration data sets were intercalibrated with the aim of removing, as far as possible, any spurious trends, and to ensure that SSTs derived from these concentrations would be self-consistent throughout the record.

[132] In all data sets, we define the ice edge such that only grid boxes with sea ice concentrations of at least 15% are retained. Hence, in the resultant HadISST1 fields, areas with concentrations of less than 15% appear as open water. This threshold was chosen because it was felt that the 15% ice edge would be consistent amongst all data sets and would avoid spurious passive microwave sea ice retrievals at low concentrations.

[133] In the Northern Hemisphere, we combined the Walsh, Assel, NIC and passive microwave data. Prior to 1901, we used a calendar monthly climatology of Walsh data, because digitized observed data for this period were lacking. In the Southern Hemisphere, we only have the NIC data, passive microwave data, and the two historical climatologies described above. Consequently, HadISST1 has no interannual variation in sea ice in the Southern Hemisphere before the early 1970s.

A1. Northern Hemisphere

[134] The main problems affecting the data sets for the Northern Hemisphere are depressed concentrations in the summertime passive microwave fields due to effects of surface melt on the retrievals, lack of within-pack concentration variability in the chart-derived fields and the varying availability of data for peripheral sea ice regions. All of these issues have the potential to introduce spurious trends in time series of sea ice area.

A1.1. Correction of Summer Melt Bias

[135] We assumed that the winter concentrations contained in the passive microwave data sets were correct, but that the summer concentrations were biased low because of the effects of surface melt and ponding. We used comparisons with NIC charts to homogenize the summer concentrations.

[136] First, a field of mean winter (December to March) differences between the NIC and the GSFC passive microwave data was calculated for 1979–1994. This was used to define the local climatological offset between these fields, at times when melt does not occur. This gave an estimate of the bias in the NIC data, as winter passive microwave data are assumed to be unbiased. Next, for each month when both NIC and GSFC data were available, i.e., October 1978 through December 1994, the local difference between the monthly median NIC and passive microwave fields was taken. (This was done for all months individually to allow for variations in the timing of melt from year to year. However, effectively no adjustments were made to the winter passive microwave fields.) After 1994, when monthly NIC fields were not available, the NIC monthly climatology for 1979–1994 was used to create corresponding differences; NCEP passive microwave data were used instead of GSFC data for 1997 to 1999, as the GSFC data set ended in 1996. The NIC-GSFC winter bias field was then subtracted from the NIC-passive microwave difference fields for each individual month to give an estimated bias-adjustment field for the summer surface melt problem. Partly to allow the calculation of weekly sea ice concentration fields for use in the OI.v2 analysis [Reynolds *et al.*, 2002], the monthly bias adjustment fields were interpolated to a daily time-scale and used to bias-correct the daily satellite sea ice fields. The monthly mean biases were

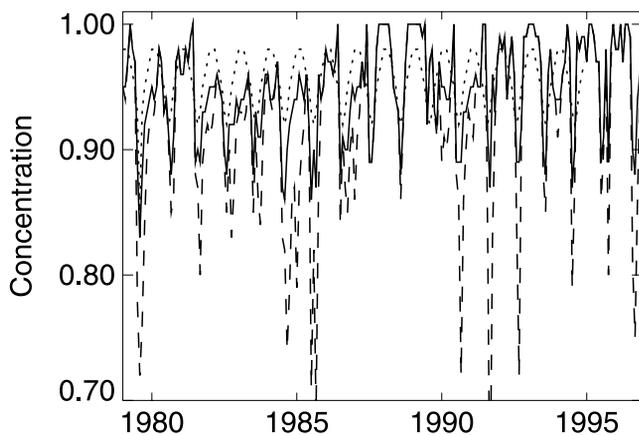


Figure A1. Time series of sea ice concentration at 0.5°E , 84.5°N , 1979–1996 in: NIC (dots); GSFC (dashes); and HadISST1 (solid).

assumed to apply to the middle day of each month and the daily bias fields were derived by linear interpolation between these mid-month values. To preserve the structure of the marginal ice zone, where ponding is less extensive and so the passive microwave data less biased, only those sea ice grid boxes which contained at least 50% sea ice and were at least three grid boxes away from the ice edge were adjusted. In these grid boxes, the daily bias adjustments, where positive, were added to the passive microwave concentrations. If this led to concentrations exceeding 100%, the grid box value was set to 100%. Where the adjustments were negative, no change was made, as we did not wish to reduce the passive microwave concentrations, as they are known to be biased low. The effect of this procedure can be seen in Figure A1 for one location within the ice pack. Wintertime concentrations are the same as in the GSFC data set, but in the summer, concentrations in HadISST1 are higher. Despite our adjustment, a seasonal cycle of concentration variation is preserved.

[137] The above process was applied to both the GSFC and NCEP daily sea ice concentration data for November 1978 onward. Monthly medians of these adjusted two-daily (to 1987) or daily (1987 onward) values provided the “corrected” sea ice concentration fields used in HadISST1.

A1.2. Addition of Spatial Variability to the Walsh Sea Ice Concentration Values

[138] The Walsh chart-derived data set prior to 1979 contains large areas having 100% sea ice concentration. The corrected passive microwave data have many fewer 100% grid boxes than this owing to the presence of leads and other fractures in the ice. In order that the general characteristics of the monthly HadISST1 sea ice fields are as consistent as possible through time, extra spatial variability was added to the Walsh-derived sea ice concentration fields. To achieve this, a 1979–1996 climatology of bias-corrected passive microwave data was used to define a set of typical calendar monthly sea ice concentration fields. In grid boxes where the Walsh-derived data had sea ice concentrations of 100% and the corrected passive microwave climatology had values of at least 90%, these climatological values were

used. This cutoff was chosen to avoid making severe reductions to the Walsh concentrations in areas that may truly have been in the historical ice pack but are now in the marginal ice zone.

A1.3. Assembly of Sea Ice Fields and Addition of Data for Peripheral Regions

[139] The data used for the Northern Hemisphere fields were as follows: (1) 1871–1900: a calendar-monthly climatology of adjusted mid-monthly Walsh data (see section A1.2) for 1901–1930. (2) 1901 to October 1978: mid-monthly adjusted Walsh data. However, fields for the period 1940–1952 were set to the calendar monthly 1940–1952 climatology, as the Walsh data set appears to be a sequence of two different climatologies during that period (Figure 1). (3) November 1978–1996: monthly median bias-adjusted GSFC data. Fields for the SSM/I data-void of December 1987 and January 1988 [Cavalieri *et al.*, 1999] were filled by linear temporal interpolation of anomalies for the previous and following months, and adding the result to the 1978–1996 climatology of the bias-adjusted GSFC data. (4) 1997 onward: monthly median bias-adjusted NCEP data.

[140] The Walsh data set contains no information on ice concentration on the Laurentian Great Lakes. The monthly Great Lakes data for 1960–79 collected by Assel [1983] (see section 2.1.2) were used to define a calendar monthly climatology, which was used in all fields prior to 1960. The Assel monthly varying Great Lakes data were used for 1960–1979. Thereafter, the passive microwave data sets were used for the Great Lakes, but owing to land contamination effects we removed spurious out of season (May through November) ice.

[141] The Sea of Japan is another area without Walsh data. The Walsh information for the Gulf of Saint Lawrence and the Baltic also appears intermittent. So, passive microwave sea ice climatologies for 1979–1996 were used in the Sea of Japan and Gulf of Saint Lawrence prior to 1979, and in the Baltic Sea between 1953 and 1971.

[142] None of the constituent data sets contains information for the Caspian Sea, which has some sea ice in winter. Unfortunately, no additional information for this region could be found, so we used a climatology for 1982–1994, first used in GISST1 [Parker *et al.*, 1995a].

[143] The overall effect of the homogenization procedures is illustrated in Figure 1. Comparison of Figures 1d and 1f illustrates the effect of the bias correction on the passive microwave-derived field for August 1990: the extent remains the same, the marginal ice zone is preserved, but the reduced sea ice concentrations within the ice pack have been increased. The addition of within-pack variability to the Walsh-derived fields can be seen when Figures 1a and 1e are compared: the concentrations have been reduced in the field for January 1930 from 100% everywhere away from the marginal ice zone to between 90 and 100%, more like the distribution seen in Figure 1c for January 1990. Figure 1g illustrates the sudden drop in summertime sea ice area seen when passive microwave data were introduced in the early 1990s in the GISST2.3b and 3.0 data sets (the same sea ice fields were used in both). The large jumps in the summertime Walsh time series (and in GISST) in the 1940s and 1950s have been replaced by one climatology for the whole period. The reduction in summertime sea ice area in

HadISST1 relative to Walsh is caused by the addition of within-pack variability discussed above and the increase in summertime sea ice area in HadISST1 relative to GSFC is due to the corrections applied to remove the effect of surface melt on the passive microwave retrievals. Note that the trend in summertime sea ice area in the last two decades remains the same in HadISST1 as in GSFC, but the discontinuity between the time series of the Walsh and GSFC data sets has been removed in HadISST1.

A2. Southern Hemisphere

[144] Creating a homogeneous set of sea ice fields for the Southern Hemisphere was particularly challenging. The basic data used were two atlas climatologies pertaining to 1929–1939 and 1947–1962, NIC data for 1973–1994, GSFC data for November 1978–1996 and NCEP data for 1997 onward. The two atlas climatologies (section 2.1.3) depicted the mean position of the ice edge, with no information about concentration variations within the ice edge. This information had to be interpreted in as realistic a way as possible. The 1929–1939 climatology was used to define the fields for 1871–1939. Between 1939 and the start of the second climatology in 1947, the sea ice concentration fields based on the two climatologies were linearly interpolated, as were the fields between 1962 and the start of the monthly varying data in 1973.

[145] The NIC chart-derived data had very high concentrations within the ice edge relative to the GSFC and NCEP passive microwave derived data. These also had to be reconciled. In addition, comparison of the GSFC and other passive microwave fields revealed a relative bias, which is treated first below.

A2.1. Homogenization of Passive-Microwave-Derived Fields

[146] The GSFC and NCEP data sets both utilize the NASA Team algorithm, so should have very similar types of concentration variation and compatible ice edges. However, although there was no overlap period between them, it was evident that the NCEP data were “icier” (i.e., were of generally larger concentration) than the GSFC. In addition, the NCEP sea ice extent appeared to be greater, although that might have reflected a real increase since 1996. So the Bristol algorithm data set [*Hanna and Bamber, 2001*] was used as an independent cross-check. The sea ice extent defined by the 15% concentration threshold was found to be much smaller in the GSFC data than in the other three data sets: NIC, NCEP and Bristol algorithm. When the ice edge was taken at a concentration of 1% in the GSFC data, extents appeared to be more compatible. This appeared to confirm that GSFC concentrations in the Antarctic are biased low. Therefore we recalibrated the GSFC data to align them with the Bristol data set and hence with the NCEP data used to update HadISST semi-operationally.

[147] Monthly difference fields were calculated for 1988–1996 between the Bristol algorithm data where that indicated some ice, i.e., a concentration of at least 15%, and those GSFC data having concentrations of at least 1%. Calendar monthly means of these difference fields were then linearly interpolated to a daily resolution, using the same method as for the Northern Hemisphere sea ice

concentration bias adjustment, i.e., the monthly mean was assumed to be applicable to the middle of each month. Values less than 1% in the interpolated difference fields were set to zero. The interpolated difference fields were added to all daily GSFC concentration values greater than zero. Resulting concentrations greater than 100% were set to 100%. If a grid box had a nonzero concentration, but no difference value was available, the box was filled using bilinear interpolation. Monthly median fields were calculated from the adjusted daily values and the ice edge reset to a concentration of 15%. (Since HadISST1 was completed, an error in the regridding of the NCEP fields from polar stereographic to the regular lat/long grid has been discovered (R. Grumbine, personal communication, 2000). This error had led to excessive sea ice extent, particularly in the Southern Hemisphere where sea ice is found at lower latitudes than in the Arctic. The effect is of the order of 1% (5%) in sea ice extent in the Northern (Southern) Hemisphere, but is smaller than the differences between the Bristol and GSFC algorithms.)

A2.2. Addition of Spatial Variability to National Ice Center Fields

[148] The relative high bias in the NIC data was removed by calibration with the recalibrated GSFC data (section A2.1). NIC fields for 1979–1994 with an ice edge at a concentration of 15% were compared with bias-corrected GSFC data with the ice edge at 1%. NIC minus bias-adjusted GSFC difference fields were calculated for each month in 1979–1994 and calendar monthly mean difference fields calculated. These calendar monthly biases were subtracted from each month in the NIC data set where both fields had a concentration greater than zero. Where NIC values were greater than zero, but there were no difference fields, the corrected NIC fields were filled by bilinear interpolation. Concentrations less than 15% were set to zero.

A2.3. Incorporation of Sea Ice Concentration Information Into Atlas Data

[149] We assumed that the mean ice edges depicted by the German and Russian Antarctic atlas climatologies corresponded to a concentration of 15%. Calibrated NIC and GSFC fields for 1973–1998 (sections A2.1 and A2.2) were used to create a realistic concentration climatology. Wherever both the atlas and modern climatologies had sea ice data and the concentration of the modern climatology was at least 80%, these climatological concentrations were inserted within the atlas ice-pack.

[150] Two approaches to filling the remainder of the concentration values were taken, depending on the season. In the austral summer, i.e., December to March, any intervals between the $\geq 80\%$ concentrations and the atlas ice edges were filled using bilinear interpolation between these $\geq 80\%$ concentrations and the assumed 15% concentration at the atlas ice edge: interpolation was carried out first north-south then east-west. Because of the rather complicated shape of the atlas ice edge in these months, this procedure results in concentrations that are rather low in the tongue of ice extending out into the Weddell Sea and Indian Ocean sector. It is difficult to justify infilling a complex ice field using linear techniques, but evidence for low concentrations can be found in the documentation of

the German climatology (*Deutsches Hydrographisches Institute* [1950], translated from the German by P. Frich).

[151] For months April to November, we used a different method that resulted in slightly more monotonic concentration gradients. Using the recalibrated satellite concentration climatology for 1973–1998, we calculated the concentration gradient between 1° area grid boxes bordering the ice edge and the concentration three grid boxes south of the ice edge, and averaged this gradient in each of 360 running 31° longitude sectors. Grid boxes from one to three boxes south of the atlas ice edges were then filled using the mean gradient for the sector centered on the grid boxes, assuming 15% concentration at the atlas ice edge. If any resulting values exceeded 80%, they were set to 80%. Finally, the regions between the climatological concentration of at least 80% and the gradient-filled boxes were filled using bilinear interpolation, north-south then east-west. Gradient and climatology values for May were used to fill atlas climatology fields for April as the shape of the April atlas fields were more like the modern field for May.

A2.4. Combination of Ice Sources

[152] The German 1929–1939 climatology was used to define the fields for 1871–1939. Between 1939 and the start of the Russian climatology in 1947, the sea ice concentration fields based on the two climatologies were linearly interpolated. A fifteen-year calendar monthly mean of the homogenized fields for 1973–1987 was used as the end point for the linear interpolation of concentration between the end of the Russian climatology in 1962 and the start of the monthly varying data in 1973.

[153] We summarize the results of our procedures by showing reconstructed Southern Hemisphere sea ice concentration fields for February and August for three selected years in the HadISST1 record (Figure 2). Like Figure 1 for the Northern Hemisphere, Figure 2 illustrates both the problem of inconsistent data sets and the resulting homogenized time series of sea ice area. Figures 2b and 2d illustrate the wintertime differences between two passive microwave-derived data sets: Bristol and GSFC. Figure 2f shows the corrected version of the GSFC field used in HadISST1. Figure 2e is the result of adding the modern concentration climatology to the German atlas-derived ice edge for August. The apparent overall decline in sea ice area in Figure 2g has been substantially reduced by our homogenization procedures from that shown in the GISST2.3b and 3.0 data sets.

Appendix B: EOF Subset and Data Error Used in RSOI of SST

[154] The number of “interannual” EOFs used was determined by discontinuities in plots of log eigenvalue versus EOF number, representing the separation of real variability from noise. The 46 4° latitude by 4° longitude (hereafter 4° area) EOFs used before 1949 represented 83% of the input variance. The 42 AMB and 44 IP (see section 3.4.2) 2° area EOFs used from 1949 represented 80% of the variance of the input data. Tests using 2° area EOFs containing approximately 90% of the variance in the 1958–1997 period yielded no significant improvement in the reconstructions.

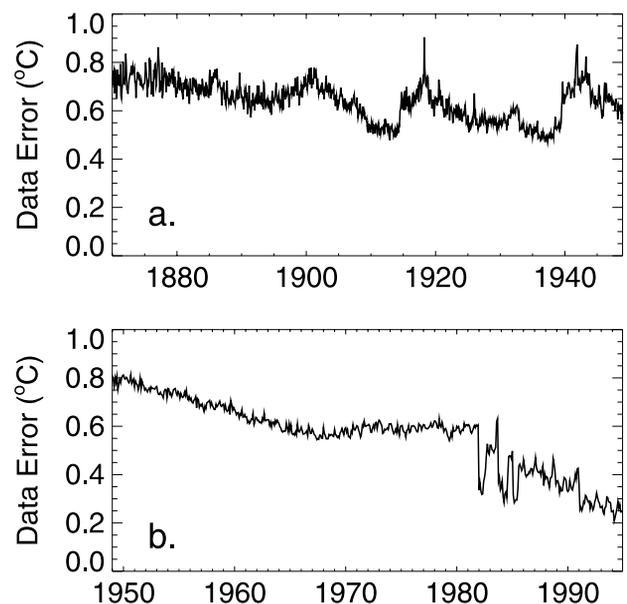


Figure B1. Time series of global root mean square average of the monthly data error used in the HadISST1 reconstruction in: (a) 4° area grid boxes, 1871–1948 and (b) 2° area grid boxes, 1949 onward.

[155] RSOI requires the calculation of monthly fields of data error with which to inversely weight the data. To create these fields for the 4° area low-frequency “global change” reconstruction, monthly 5° area fields of variance of 1° area pentad anomalies comprising each 5° area monthly MOHSST6 anomaly were bilinearly interpolated to the 4° grid, and divided by the monthly number of observations in each 4° area. The square root of this ratio gave the required grid box error fields. Data error estimates were made for the residual interannual analysis (on 2° or 4° area resolution) in the same way as for the global change reconstruction, but with one addition. The SST measurement error variance estimates of *Kent et al.* [1999] were added to the intrabox variances, and the result multiplied by 0.75 to give the estimated total measurement plus sampling error variance of each 2° or 4° area box. The factor 0.75 takes account of expected measurement error contributions to the intrabox variability values [*Jones et al.*, 2001]. For the 2° area reconstruction, fields of squared intragrid box standard deviation and monthly numbers of observations taken from COADS were used as above. Figure B1 shows the variation with time of the global root mean square average monthly data error used.

Appendix C: Bias Adjustment of Satellite-Based Input SST

[156] In GISST, the influence of AVHRR data biases on the analysis was minimized by interpolating gaps between in situ SST data using the spatial second derivative of the AVHRR SST values [*Reynolds*, 1988]. However, *Hurrell and Trenberth* [1999] showed that the month-to-month persistence of SST in GISST2.3b (and 3.0) from 1982 onward was lower than it should have been over much of the ocean, especially in the Southern Hemisphere. The

problem arose because of cloud- or aerosol-related small-scale, temporally incoherent, variations in the biases in AVHRR, which may not have been neutralized by the above procedure, especially in the data-sparse Southern Ocean.

[157] To create reliable, smooth estimates of the biases in the AVHRR SSTs, we created a smooth, complete in situ analysis using RSOI and subtracted it from a completed AVHRR analysis. The difference was then smoothed. The in situ analysis was based on the quality-controlled MDB/COADS blend (section 3.3.2). These SST anomalies were reconstructed using RSOI as described in section 3.4, except that the EOFs were based entirely on in situ data for 1982–1998, and were computed separately for each of the Indian, Pacific and Atlantic (including the Mediterranean and Black Seas) oceans, the global change and interannual components were not separated. They covered a wider area than that shown by the EOF in Figure 4 and included the parts of the Southern Ocean adjacent to each above-mentioned ocean basin. Data error estimates were derived as described in Appendix B. The noise-cutoff criterion, based on log (eigenvalue) plots as in Appendix B, was chosen to ensure that at least 80% of the variance of the input data set was captured. Ocean-basin fields were reconstructed using RSOI and combined; then the original in situ data were reinserted in areas too sparsely observed to create EOFs. Fields of absolute SST values were created by adding back the 1° GISST2.2 1961–1990 climatology [Parker *et al.*, 1995c]. Finally, we added SSTs in areas of partial sea ice cover (similar to section 4.1 but based on relationships developed for GISST2.3b), and remaining gaps were filled using the Poisson technique, preserving the second derivative of the GISST2.2 climatology. We completed the monthly 1° area fields of night AVHRR SST data in the same way. The time-varying bias of the AVHRR fields relative to the in situ fields was then calculated by subtracting the in situ fields from the AVHRR fields. The difference fields were smoothed using a moving window average with radius 2224 km (20 degrees of latitude). The smoothed bias fields were then subtracted from the monthly AVHRR SST.

Appendix D: Derivation of Relationships Used to Specify SST Near Sea Ice

[158] Both in situ and bias-adjusted AVHRR SST data were used, where available, to develop the relationships. (Note, however, that it is unlikely that AVHRR data at these high latitudes will have been adequately bias-adjusted by the sparse in situ data (Appendix C). So the resultant relationships may specify SSTs too low owing to residual contamination of the AVHRR SSTs by atmospheric aerosols and, in particular, by sea ice within the field of view.) Relationships were calculated separately for each hemisphere and calendar month in each of 360 overlapping 31° longitude sectors. In the Northern Hemisphere, additional relationships were formed for peripheral regions: the Great Lakes, the Baltic Sea, the Seas of Okhotsk and Japan and the Gulf of Alaska. Relationships for the Northern Hemisphere as a whole were used for the Caspian Sea because there are few reliable data there.

[159] Coincident pairs of 1° area monthly median sea ice concentration and mean SST were collated for each sector

or region and three-calendar-month period, using data for 1961 to 1998 except for the Great Lakes where we used data for 1961–1979. Twelve overlapping three-month periods were used to develop the relationships, to ensure a smooth transition between calendar months. Quadratic equations relating SST to sea ice concentration (SIC) were fitted to these data using ordinary least squares (see Figure 7):

$$SST = a \times SIC^2 + b \times SIC + c \quad (D1)$$

where $0.15 \leq SIC < 0.90$. These relationships were constrained such that, at sea ice concentrations of 90% or more, SST was set to its freezing value of -1.8°C , assuming a salinity of 35 parts per thousand. (In reality, the salinity of the upper layers of the Arctic Ocean can be much less than 35 parts per thousand, owing to river runoff and to melting of desalinated sea ice or its snow cover, so using -1.8°C here may bias the result too cold.) In the freshwater Great Lakes, we used 0°C as the freezing limit. We tested constraining the relationships to the freezing point at sea ice concentrations of 80% or 100%. The use of 90% achieved the best fit to the independent Generalised Digital Environment Model (GDEM) [Teague *et al.*, 1990] climatology of SST for the Arctic (Figure 8). When there were fewer than 100 pairs of SST and sea ice data on which to base a relationship, coefficients from equation (D1) for neighboring areas or months were linearly interpolated.

Appendix E: Data-Adaptive Smoothing of NMAT Data

[160] To provide a spatial consistency benchmark, fields of mean and standard deviation, σ_n , of the difference between the GISST3.0 temperature anomaly in each 5° grid box and the average of that in its eight near-neighbors were produced for each calendar month during the period 1982–1997. GISST3.0 was used because it is globally complete and such statistics for SST and NMAT were found to be similar: because the HadMAT1 analysis was performed first, HadISST1 was not yet available. The period 1982–1997 was chosen to include information from satellite SSTs and large deviations from recent El Niño events.

[161] Each 5° grid box in the NMAT analysis was compared with the average of its neighbors. If the target grid box had no neighbors with data, or if it and its neighbors contained only interpolated values, no smoothing was applied. Also, if the total number of real observations in the nine grid boxes was equal to at least 135 and the number of real observations contributing to the target grid box value was at least 15 (15 was taken as the minimum number of observations required to produce a reliable monthly 5° area average), no smoothing was applied. This ensured that well observed, but unusual NMAT anomalies, such as occurred around the British Isles in February 1947 and January 1963, were not smoothed unnecessarily. Otherwise, if the difference between the target box and the average of its neighbors was greater than $3.3\sigma_n$, the target value was smoothed. The smoothed value was the mean of the (up to) nine values, weighted by the number of observations within each box. In this calculation, boxes containing more than 1500 observations were assigned exactly 1500 observations, and boxes containing reconstructed data were assigned just 15 obser-

vations (these 15 observations were not counted for the purpose of determining whether or not a box should be smoothed). The smoothing process was repeated until fewer than 2% of boxes required smoothing according to these criteria.

[162] **Acknowledgments.** The sea ice fields contained in HadISST1 were developed with invaluable help from a sea ice working group set up by ECMWF for ERA40. We would like to thank all members of that working group for their contributions: J. E. Walsh, R. W. Reynolds, M. C. Serreze, J. A. Maslanik, R. Grumbine, and W. Chapman. We would also like to thank E. Hanna, S. Worley, R. W. Reynolds, and D. C. Stokes for providing the Bristol algorithm, COADS, and AVHRR data sets, respectively. Goddard Space Flight Center sea ice fields were provided by the National Snow and Ice Data Center, University of Colorado, Boulder, Colorado, USA. Marine air temperature profiles used to improve NMAT corrections were supplied by A. Grant. Thanks are also due to M. Fiorino, P. Viterbo, D. M. Smith, D. M. H. Sexton, D. Cavalieri, J. W. Hurrell, R. W. Reynolds, D. C. Stokes, and T. Ansell for their help and/or useful comments. Technical help was provided by A. Brady. This paper was improved following the constructive suggestions of three reviewers. This work was supported by the U.K. Government Meteorological Research contract and by the U.K. Department of the Environment, Food and Rural Affairs contract PECD/7/12/37. Through the contribution of the Met Office authors, this work is British Crown Copyright.

References

- Allan, R., J. Lindesay, and D. E. Parker, *El Niño Southern Oscillation and Climatic Variability*, 416 pp., CSIRO, Collingwood, Victoria, Australia, 1996.
- Assel, R. A., Description and analysis of a 20-year (1960–79) digital ice-concentration database for the Great Lakes of North America, *Ann. Glaciol.*, 4, 14–18, 1983.
- Barry, R. G., and J. A. Maslanik, Arctic sea ice characteristics and associated atmosphere-ice interactions in summer inferred from SMMR data and drifting buoys, 1979–1984, *GeoJournal*, 18, 35–44, 1989.
- Bottomley, M., C. K. Folland, J. Hsiung, R. E. Newell, and D. E. Parker, *Global Ocean Surface Temperature Atlas*, 20 + iv pp. and 313 color plates, Her Majesty's Stn. Off., Norwich, UK, 1990.
- Burns, B. A., D. J. Cavalieri, M. R. Keller, W. J. Campbell, T. C. Grenfell, G. A. Maykut, and P. Gloersen, Multisensor comparison of ice concentration estimates in the marginal ice zones, *J. Geophys. Res.*, 92, 6843–6856, 1987.
- Cavalieri, D. J., P. Gloersen, C. L. Parkinson, J. C. Comiso, and H. J. Zwally, Observed hemispheric asymmetry in global sea ice changes, *Science*, 278, 1104–1106, 1997.
- Cavalieri, D. J., C. L. Parkinson, P. Gloersen, J. C. Comiso, and H. J. Zwally, Deriving long-term time series of sea ice cover from satellite passive-microwave multisensor data sets, *J. Geophys. Res.*, 104, 15,803–15,814, 1999.
- Christy, J. R., D. E. Parker, S. J. Brown, I. Macadam, M. Stendel, and W. B. Norris, Differential trends in tropical sea surface and atmospheric temperatures since 1979, *Geophys. Res. Lett.*, 28, 183–186, 2001.
- Comiso, J. C., Characteristics of Arctic winter sea ice from satellite multi-spectral satellite observations, *J. Geophys. Res.*, 91, 975–994, 1986.
- Comiso, J. C., and R. Kwok, Surface and radiative characteristics of the summer Arctic sea ice cover from multisensor satellite observations, *J. Geophys. Res.*, 101, 28,397–28,416, 1996.
- Cox, M. G., and J. G. Hayes, Curve fitting: A guide and suite of algorithms for the non-specialist user, *Rep. NAC26*, Natl. Phys. Lab., Teddington, UK, 1973.
- Cressman, G. P., An operational objective analysis system, *Mon. Weather Rev.*, 87, 367–374, 1959.
- de la Mare, W. K., Abrupt mid-twentieth century decline in Antarctic sea ice extent from whaling records, *Nature*, 389, 57–59, 1997.
- Delderfield, J., D. T. Llewellyn-Jones, R. Bernard, Y. de Javel, E. J. Williamson, I. M. Mason, D. R. Pick, and I. J. Barton, The Along-track Scanning Radiometer (ATSR) for ERS-1, *Proc. SPIE Int. Soc. Opt. Eng.*, 589, 114–120, 1986.
- Deutsches Hydrographisches Institute, Atlas of ice conditions in the North Atlantic Ocean and general charts of the ice conditions of the North and South Polar regions, *Publ. 2335*, Hamburg, Germany, 1950.
- Diaz, H. F., C. K. Folland, T. Manabe, D. E. Parker, R. W. Reynolds, and S. D. Woodruff, Workshop on advances in the use of historical marine climate data (Boulder, Co., USA, 29th Jan–1st Feb 2002), *WMO Bull. 51*, pp. 377–380, World Meteorol. Organ., Geneva, 2002.
- Diggle, P. J., K.-Y. Liang, and S. L. Zeger, *Analysis of Longitudinal Data*, 253 pp., Clarendon, Oxford, UK, 1999.
- Emery, W. J., C. Fowler, and J. Maslanik, Arctic sea ice concentrations from special sensor microwave imager and advanced very high resolution radiometer satellite data, *J. Geophys. Res.*, 99, 18,329–18,342, 1994.
- Fairall, C. W., E. F. Bradley, D. P. Rodgers, J. B. Edson, and G. D. Young, Bulk parameterization of air-sea fluxes for TOGA-COARE, *J. Geophys. Res.*, 101, 3744–3764, 1996.
- Folland, C. K., and D. E. Parker, Correction of instrumental biases in historical sea surface temperature data, *Q. J. R. Meteorol. Soc.*, 121, 319–367, 1995.
- Folland, C. K., and M. J. Salinger, Surface temperature trends and variations in New Zealand and the surrounding ocean, 1871–1993, *Int. J. Climatol.*, 15, 1195–1218, 1995.
- Folland, C. K., R. W. Reynolds, M. Gordon, and D. E. Parker, A study of six operational sea surface temperature analyses, *J. Clim.*, 6, 96–113, 1993.
- Folland, C. K., M. J. Salinger, and N. A. Rayner, A comparison of annual South Pacific island and ocean surface temperatures, *Weather Clim.*, 17(1), 23–42, 1997.
- Folland, C. K., D. M. H. Sexton, D. J. Karoly, C. E. Johnson, D. P. Rowell, and D. E. Parker, Influences of anthropogenic forcing, sea surface temperature and sea ice extent on recent climate change, *Geophys. Res. Lett.*, 25, 353–356, 1998.
- Folland, C. K., D. E. Parker, A. W. Colman, and R. Washington, Large scale modes of ocean surface temperature since the late nineteenth century, in *Beyond El Niño: Decadal and Interdecadal Climate Variability*, edited by A. Navarra, pp. 73–102, Springer-Verlag, New York, 1999.
- Folland, C. K., T. R. Karl, J. R. Christy, R. A. Clarke, G. V. Gruza, J. Jouzel, M. E. Mann, J. Oerlemans, M. J. Salinger, and S.-W. Wang, Observed climate variability and change, in *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report the Intergovernmental Panel on Climate Change*, edited by J. T. Houghton et al., pp. 99–181, Cambridge Univ. Press, New York, 2001a.
- Folland, C. K., et al., Global temperature change and its uncertainties since 1861, *Geophys. Res. Lett.*, 28, 2621–2624, 2001b.
- Gregory, J. M., P. A. Stott, D. J. Cresswell, N. A. Rayner, C. Gordon, and D. M. H. Sexton, Recent and future changes in Arctic sea ice simulated by the HadCM3 AOGCM, *Geophys. Res. Lett.*, 29(24), 2175, doi:10.1029/2001GL014575, 2002.
- Grumbine, R. W., Automated passive microwave sea ice concentration analysis at NCEP, U.S. Department of Commerce, *NCEP Tech. Note*, Natl. Weather Serv., Natl. Oceanic and Atmos. Admin., Camp Springs, Md., 1996.
- Hanawa, K., S. Yasunaka, T. Manabe, and N. Iwasaka, Examination of correction to historical SST data using long-term coastal SST data taken around Japan, *J. Meteorol. Soc. Jpn.*, 78, 187–195, 2000.
- Hanna, E., and J. Bamber, Derivation and optimisation of a new Antarctic sea ice record, *Int. J. Remote Sens.*, 22, 113–139, 2001.
- Hansen, J., R. Ruedy, J. Glascoe, and M. Sato, GISS analysis of surface temperature change, *J. Geophys. Res.*, 104, 30,997–31,022, 1999.
- Hansen, J., et al., Climate forcings in Goddard Institute for Space Studies SI2000 simulations, *J. Geophys. Res.*, 107(D18), 4347, doi:10.1029/2001JD001143, 2002.
- Hurrell, J. W., and K. E. Trenberth, Global sea surface temperature analyses: Multiple problems and their implications for climate analysis, modeling and reanalysis, *Bull. Am. Meteorol. Soc.*, 80, 2661–2678, 1999.
- Jones, P. D., Antarctic temperatures over the present century: A study of the early expedition record, *J. Clim.*, 3, 1193–1203, 1990.
- Jones, P. D., T. J. Osborn, K. R. Briffa, C. K. Folland, E. B. Horton, L. V. Alexander, D. E. Parker, and N. A. Rayner, Adjusting for sampling density in grid-box land and ocean surface temperature time series, *J. Geophys. Res.*, 106, 3371–3380, 2001.
- Kaplan, A., Y. Kushnir, M. Cane, and M. Blumenthal, Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures, *J. Geophys. Res.*, 102, 27,835–27,860, 1997.
- Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan, Analyses of global sea surface temperature 1856–1991, *J. Geophys. Res.*, 103, 18,567–18,589, 1998.
- Kaplan, A., M. A. Cane, and Y. Kushnir, Reduced space approach to the optimal analysis of historical marine observations: Accomplishments, difficulties and prospects, in *WMO Guide to the Applications of Marine Climatology*, World Meteorol. Organ., Geneva, in press, 2003.
- Karl, T. R., R. W. Knight, and B. Baker, The record breaking global temperatures of 1997 and 1998: Evidence for an increase in the rate of global warming?, *Geophys. Res. Lett.*, 27, 719–722, 2000.
- Kent, E. C., P. K. Taylor, B. S. Truscott, and J. S. Hopkins, The accuracy of Voluntary Observing Ships' meteorological observations: Results of the VSOP-NA, *J. Atmos. Oceanic Technol.*, 10, 591–608, 1993.

- Kent, E. C., P. G. Challenor, and P. K. Taylor, A statistical determination of the random observational errors present in Voluntary Observing Ships' meteorological reports, *J. Atmos. Oceanic Technol.*, *16*, 905–914, 1999.
- Knight, R. W., Introduction to a new sea ice database, *Ann. Glaciol.*, *5*, 81–84, 1984.
- Løyning, T. B., C. Dick, H. Goodwin, O. Pavlova, T. Vinje, G. Kjærli, and T. Villinger, ACSYS historical ice chart archive (1553–2002), *IACPO Informal Rep.* 8, Int. ACSYS/CLIC Proj. Off., Norw. Polar Inst., Tromsø, Norway, 2003.
- Markus, T., and D. J. Cavalieri, An enhancement of the NASA Team sea ice algorithm, *IEEE Trans. Geosci. Remote Sens.*, *38*, 1387–1398, 2000.
- New, M., M. Hulme, and P. D. Jones, Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology, *J. Clim.*, *12*, 829–856, 1999.
- Parker, D. E., P. D. Jones, C. K. Folland, and A. Bevan, Interdecadal changes of surface temperature since the late 19th century, *J. Geophys. Res.*, *99*, 14,377–14,399, 1994.
- Parker, D. E., C. K. Folland, A. Bevan, M. N. Ward, M. Jackson, and K. Maskell, Marine surface data for analysis of climatic fluctuations on interannual to century timescales, in *Natural Climate Variability on Decade-to-Century Time Scales*, edited by D. G. Martinson et al., pp. 241–250 + color figures on pp. 222–228, Natl. Acad. Press, Washington, D. C., 1995a.
- Parker, D. E., C. K. Folland, and M. Jackson, Marine surface temperature: Observed variations and data requirements, *Clim. Change*, *31*, 559–600, 1995b.
- Parker, D. E., M. Jackson, and E. B. Horton, The GISST2.2 sea surface temperature and sea ice climatology, *Clim. Res. Tech. Note CRTN 63*, Hadley Cent., Met Office, Bracknell, UK, 1995c.
- Peterson, R. G., and W. B. White, Slow oceanic teleconnections linking the Antarctic Circumpolar Wave with the tropical El Niño–Southern Oscillation, *J. Geophys. Res.*, *103*, 24,573–24,583, 1998.
- Pope, V. D., M. L. Gallani, P. R. Rowntree, and R. A. Stratton, The impact of new physical parametrizations in the Hadley Centre climate model HadAM3, *Clim. Dyn.*, *16*, 123–146, 1999.
- Rayner, N. A., and D. E. Parker, Techniques used in the construction of GISST, in *Proceedings of the OOPC/AOPC Workshop on Global Sea Surface Temperature Data Sets, IRI/LDEO, NY, USA, 2–4 November 1998*, Annex III, pp. 13–16, Lamont-Doherty Earth Obs., Palisades, N. Y., 1999.
- Rayner, N. A., E. B. Horton, D. E. Parker, C. K. Folland, and R. B. Hackett, Version 2.2 of the global sea ice and sea surface temperature data set, 1903–1994, *Clim. Res. Tech. Note CRTN74*, Hadley Cent., Met Office, Bracknell, UK, 1996.
- Reynolds, R. W., A real-time global sea surface temperature analysis, *J. Clim.*, *1*, 75–86, 1988.
- Reynolds, R. W., Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures, *J. Clim.*, *6*, 768–774, 1993.
- Reynolds, R. W., and T. M. Smith, Improved global sea surface temperature analyses using optimum interpolation, *J. Clim.*, *7*, 929–948, 1994.
- Reynolds, R. W., C. K. Folland, and D. E. Parker, Biases in satellite-derived sea-surface-temperature data, *Nature*, *341*, 728–731, 1989.
- Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, An improved in situ and satellite SST analysis for climate, *J. Clim.*, *15*, 1609–1625, 2002.
- Rodwell, M. J., and C. K. Folland, Atlantic air-sea interaction and model validation, *Ann. Geophys.*, *46*, 47–56, 2003.
- Rodwell, M. J., D. P. Rowell, and C. K. Folland, Oceanic forcing of the wintertime North Atlantic Oscillation and European climate, *Nature*, *398*, 320–323, 1999.
- Rowell, D. P., and F. W. Zwiers, The global distribution of sources of atmospheric decadal variability and mechanisms over the tropical Pacific and southern North America, *Clim. Dyn.*, *15*, 751–772, 1999.
- Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, A dipole mode in the tropical Indian Ocean, *Nature*, *401*, 360–363, 1999.
- Sheppard, C., and N. A. Rayner, Utility of the Hadley Centre sea ice and sea surface temperature data set (HadISST1) in two widely contrasting coral reef areas, *Mar. Pollut. Bull.*, *44*, 303–308, 2002.
- Smith, D. M., Extraction of winter total sea ice concentration in the Greenland and Barents Seas from SSM/I data, *Int. J. Remote Sens.*, *17*, 2625–2646, 1996.
- Smith, D. M., Recent increase in the length of the melt season of perennial Arctic sea ice, *Geophys. Res. Lett.*, *25*, 655–658, 1998.
- Smith, T. M., and R. W. Reynolds, Bias corrections for historical sea surface temperatures based on marine air temperatures, *J. Clim.*, *15*, 73–87, 2002.
- Smith, T. M., and R. W. Reynolds, Extended reconstruction of global sea surface temperatures based on COADS data (1854–1997), *J. Clim.*, *16*, 1495–1510, 2003.
- Smith, T. M., R. W. Reynolds, R. E. Livezey, and D. C. Stokes, Reconstruction of sea surface temperature using empirical orthogonal functions, *J. Clim.*, *9*, 1403–1420, 1996.
- Smith, T. M., R. E. Livezey, and S. S. Shen, An improved method for analysing sparse and irregularly distributed SST data on a regular grid: The tropical Pacific Ocean, *J. Clim.*, *11*, 1717–1729, 1998.
- Taylor, P. K., Report of Working Group 1: Observations, in *Proceedings of the OOPC/AOPC Workshop on Global Sea Surface Temperature Data Sets, IRI/LDEO, NY, USA, 2–4 November 1998*, pp. 6–16, Lamont-Doherty Earth Obs., Palisades, N. Y., 1999.
- Teague, W. J., M. J. Carron, and P. J. Hogan, A comparison between the Generalized Digital Environmental Model and Levitus climatologies, *J. Geophys. Res.*, *95*, 7167–7184, 1990.
- Tolstikov, E. I., (Ed.), *Atlas of the Antarctic*, vol. 1, 126 pp., Glav. Uprav. Geod. Kart. MGSSSR, Moscow, 1966.
- Walsh, J. E., A data set on Northern Hemisphere sea ice extent, 1953–76, *Glaciol. Data Rep. GD-2, Arctic Sea Ice, Part 1*, pp. 49–51, <http://nsidc.org/>, World Data Cent. for Glaciol., Boulder, Colo., 1978.
- Walsh, J. E., and W. L. Chapman, Twentieth-century sea ice variations from observational data, *Ann. Glaciol.*, *33*, 444–448, 2001.
- Walsh, J. E., and C. M. Johnson, Analysis of Arctic sea ice fluctuations 1953–1977, *J. Phys. Oceanogr.*, *9*, 580–591, 1978.
- Webster, P. J., A. M. Moore, J. P. Loschnigg, and R. R. Leben, Coupled ocean-atmosphere dynamics in the Indian Ocean during 1997–98, *Nature*, *401*, 356–360, 1999.
- Wentz, F. J., C. Gentemann, D. Smith, and D. Chelton, Satellite measurements of sea surface temperature through clouds, *Science*, *288*, 847–850, 2000.
- White, W. B., and R. G. Peterson, An Antarctic circumpolar wave in surface pressure, wind, temperature and sea ice extent, *Nature*, *380*, 699–702, 1996.
- Woodruff, S. D., R. J. Slutz, R. L. Jenne, and P. M. Steurer, A comprehensive ocean-atmosphere data set, *Bull. Am. Meteorol. Soc.*, *68*, 1239–1250, 1987.
- Woodruff, S. D., H. F. Diaz, J. D. Elms, and S. J. Worley, COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields, *Phys. Chem. Earth*, *23*, 517–526, 1998.
- World Meteorological Organization (WMO), International list of selected, supplementary and auxiliary ships, *WMO Publ.* 47, Geneva, 1955–1999.
- Zheng, X., and C. S. Frederiksen, Validating interannual variability in an ensemble of AGCM simulations, *J. Clim.*, *12*, 2386–2396, 1999.
- Zwally, H. J., C. L. Parkinson, and J. C. Comiso, Variability of Antarctic sea ice and changes in carbon dioxide, *Science*, *220*, 1005–1012, 1983.

L. V. Alexander, C. K. Folland, E. B. Horton, D. E. Parker, N. A. Rayner, and D. P. Rowell, Hadley Centre for Climate Prediction and Research, Met Office, London Road, Bracknell, Berkshire RG12 2SY, UK. (nick.rayner@metoffice.com)

A. Kaplan, Lamont-Doherty Earth Observatory, Columbia University, P. O. Box 1000, Palisades, NY 10964, USA.

E. C. Kent, James Rennell Division, Southampton Oceanography Centre, Southampton SO14 3ZH, UK.

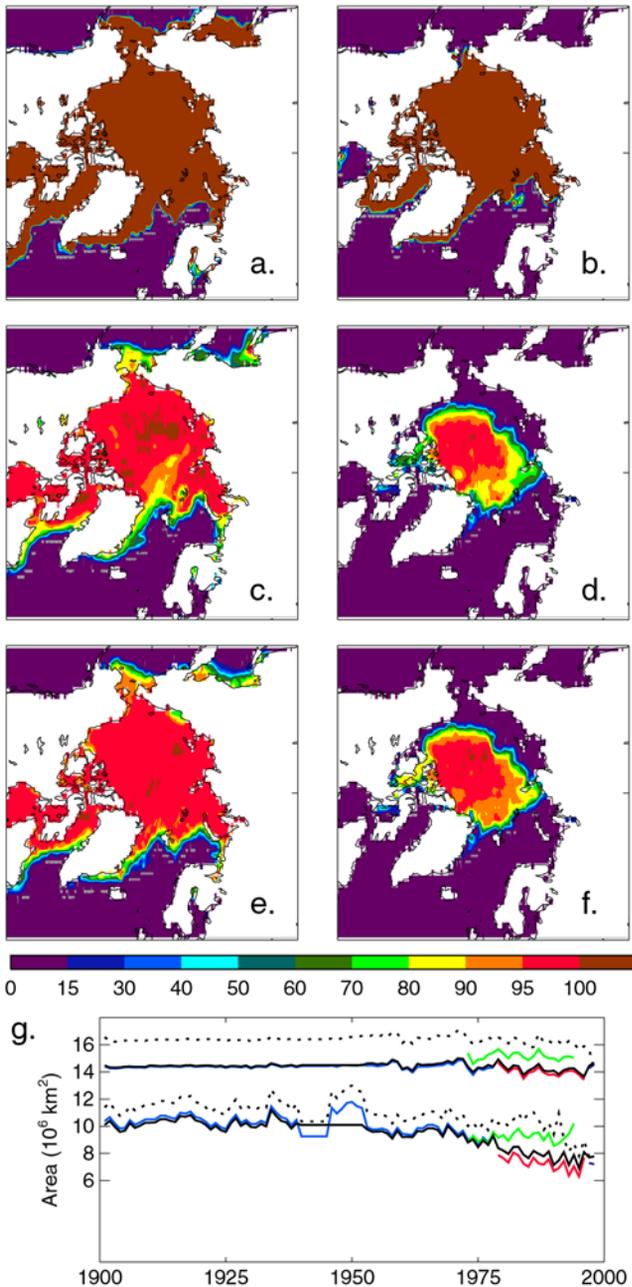


Figure 1. Northern Hemisphere sea ice. Concentration maps: (a) Walsh, January 1930; (b) Walsh, August 1930; (c) GSFC, January 1990; (d) GSFC, August 1990; (e) HadISST1, January 1930; and (f) HadISST1, August 1990. Values = 100% are colored brown. Sea ice area time series 1901–1998: (g) (top curves: January; bottom curves: July) from GISST3.0 (black dots), Walsh (light blue), NIC (green), GSFC (red), NCEP (dark blue) and HadISST1 (black). Appendix A gives details of procedures used to produce HadISST1 fields.

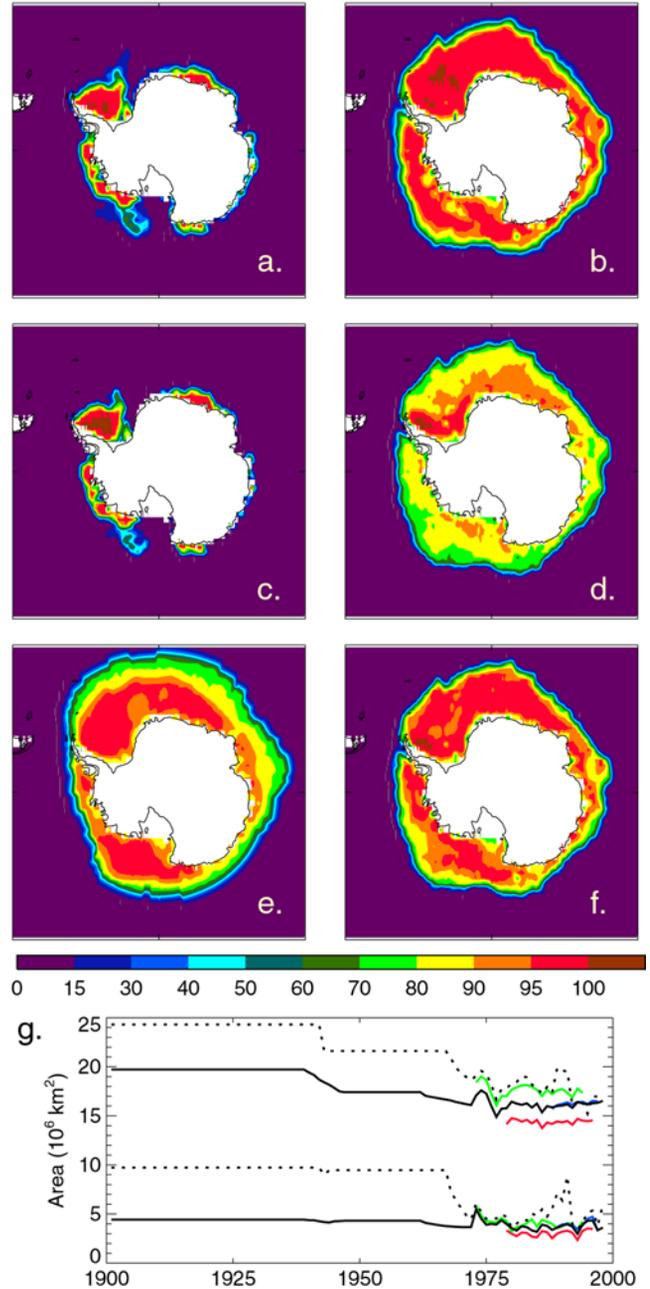


Figure 2. Southern Hemisphere sea ice. Concentration maps: (a) Bristol, January 1990; (b) Bristol, August 1990; (c) GSFC, January 1990; (d) GSFC, August 1990; (e) HadISST1, August 1930; and (f) HadISST1, August 1990. Values = 100% are colored brown. Sea ice area time series 1901–1998: (g) (top curves: September; bottom curves: January) from GISST3.0 (black dots), Bristol (light blue), NIC (green), GSFC (red), NCEP (dark blue) and HadISST1 (black). Appendix A gives details of procedures used to produce HadISST1 fields.

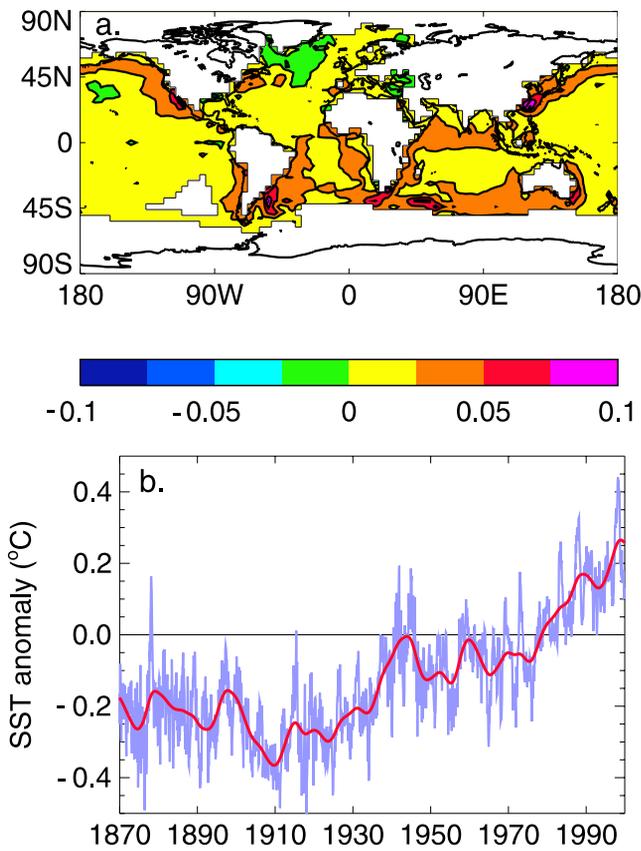


Figure 4. (a) The leading EOF of low-pass filtered (>8 years) 4° area resolution combined in situ and AVHRR data for 1901–1997. (b) Global average of the field reconstructed using only Figure 4a for 1870–1999 (see text). Blue curve is the monthly average, while the red curve has a 21pt binomial (near decadal) filter applied.

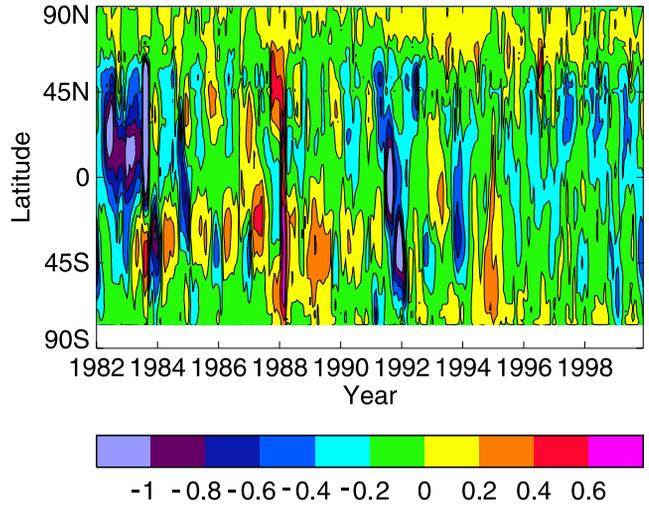


Figure 6. Latitude-time section of the zonal average of smoothed nighttime AVHRR minus in situ SST difference ($^\circ\text{C}$), 1982–1999.

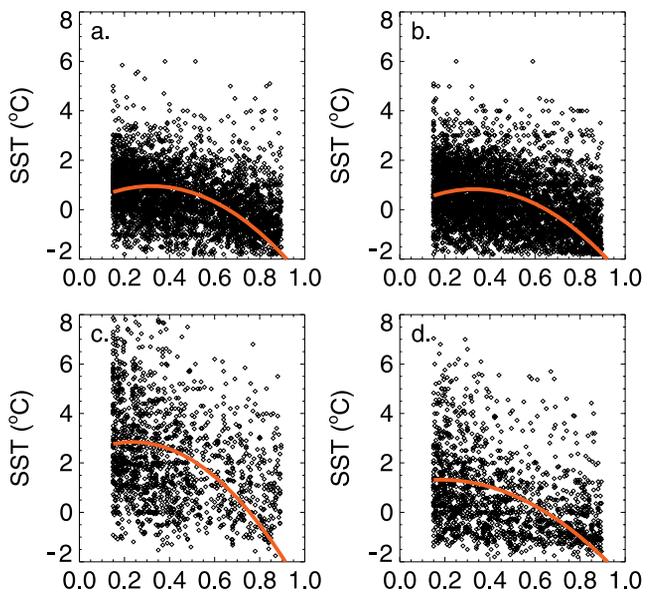


Figure 7. Fitted relationships (quadratic curves) between pairs of sea-ice concentration and sea surface temperature data (diamonds) for the 31° longitude band centered on 179.5°W in the Northern Hemisphere for three-month seasons centered on (a) January, (b) April, (c) July, and (d) October.

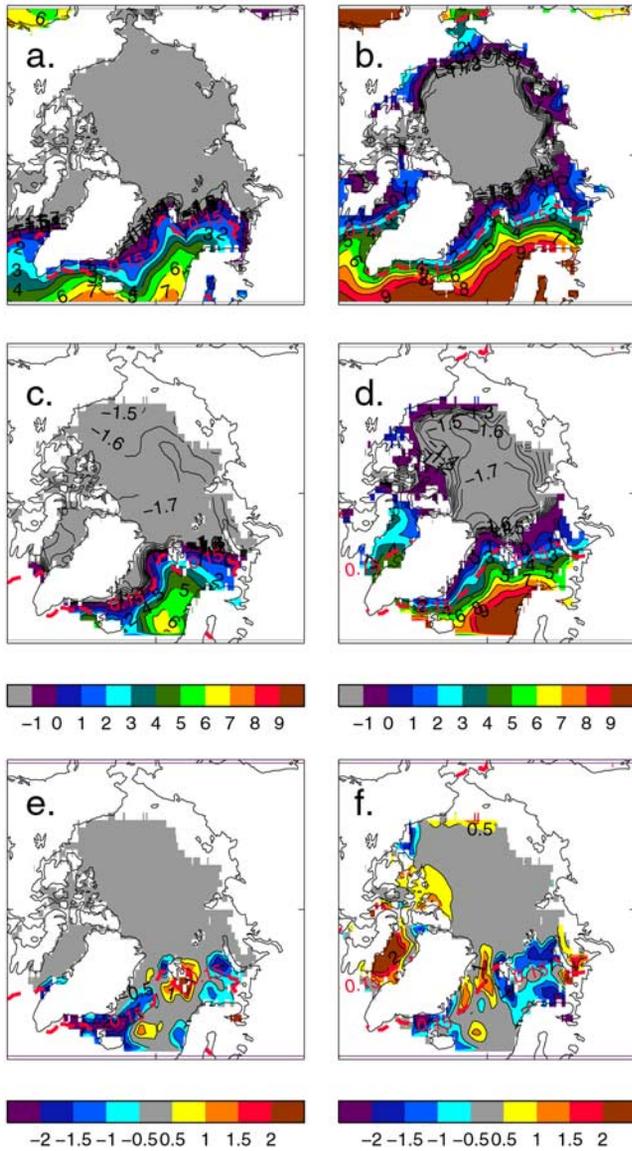


Figure 8. Arctic SST ($^{\circ}\text{C}$), averaged over 1920–1999, from HadISST1 ((a) January and (b) July) and the Generalised Digital Environment Model SST climatology ((c) January and (d) July). Difference GDEM - HadISST1 also shown ((e) January and (f) July).

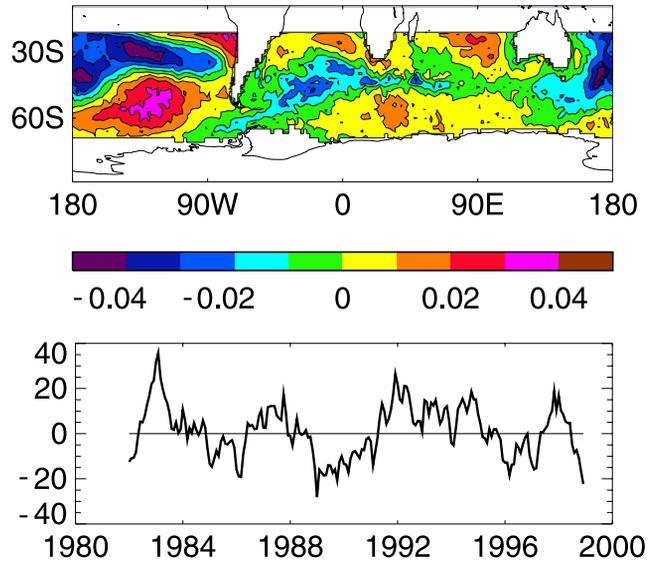


Figure 9. The leading EOF of SST variability south of 20°S in 1982–1998 and its projection onto the HadISST1 reconstruction for the same period.

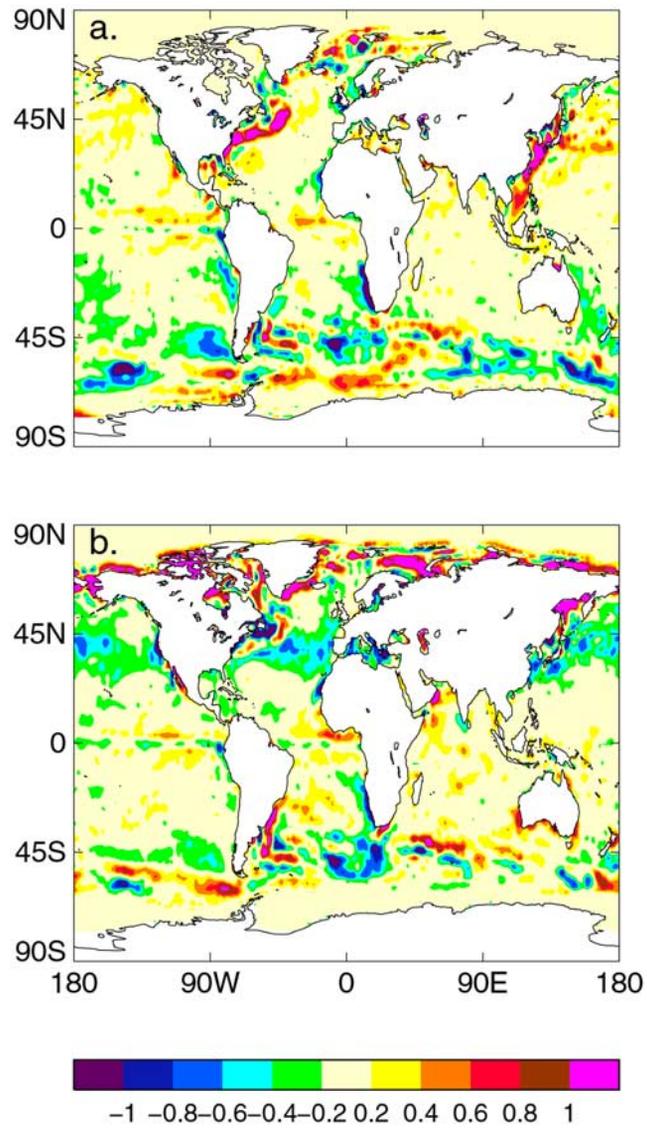


Figure 11. Sea surface temperature climatology (1971–2000) differences (°C): adjusted OI.v2 minus HadISST1, for (a) January and (b) July.

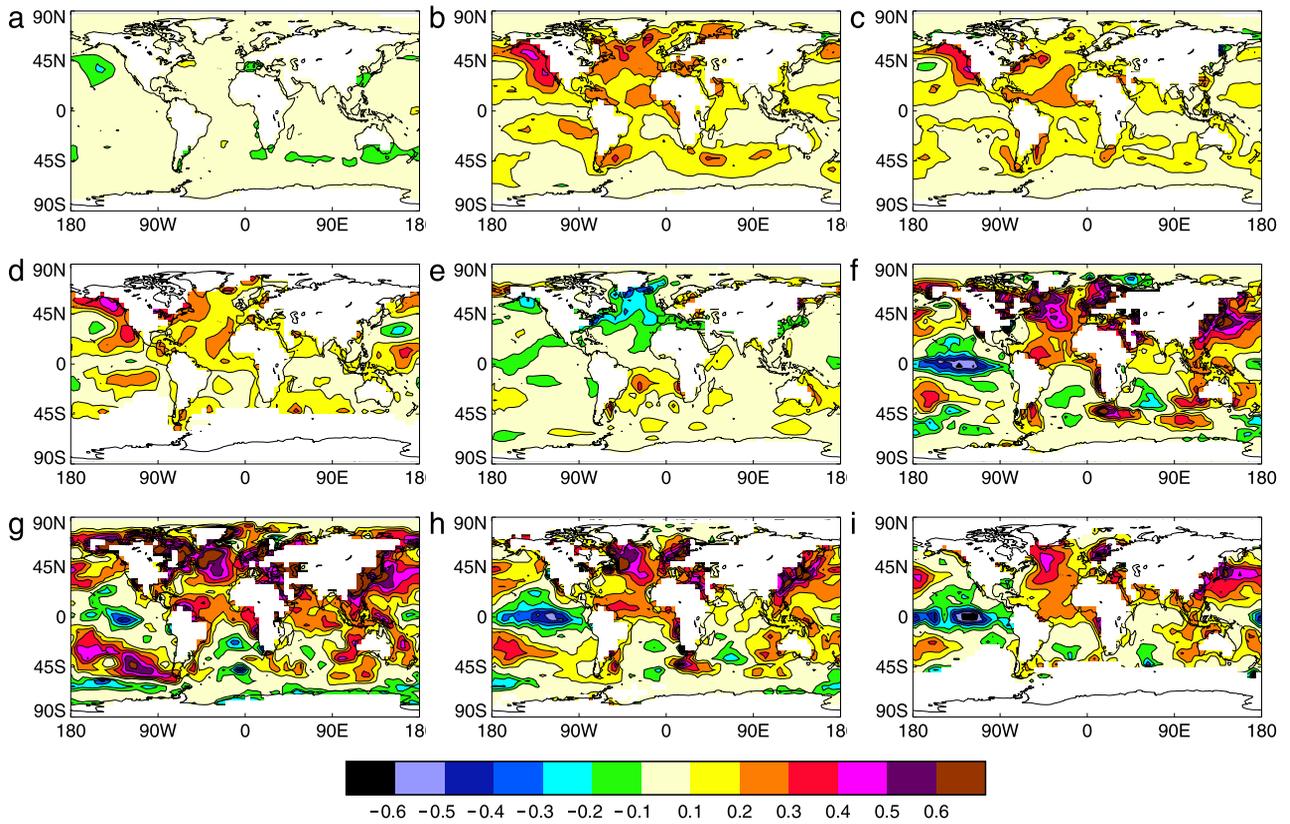


Figure 12. Restricted maximum likelihood linear trends in $^{\circ}\text{C}$ per decade in each 5° area grid-box in: (a) 1871–1909 (HadISST1); (b–d) 1910–1945 (GISST2.3b, HadISST1, and Kaplan *et al.* [1998, 2003], respectively); (e) 1946–1975 (HadISST1); (f–i) 1982–1999 (OL.v2, GISST2.3b, HadISST1, and Kaplan *et al.* [1998, 2003], respectively).

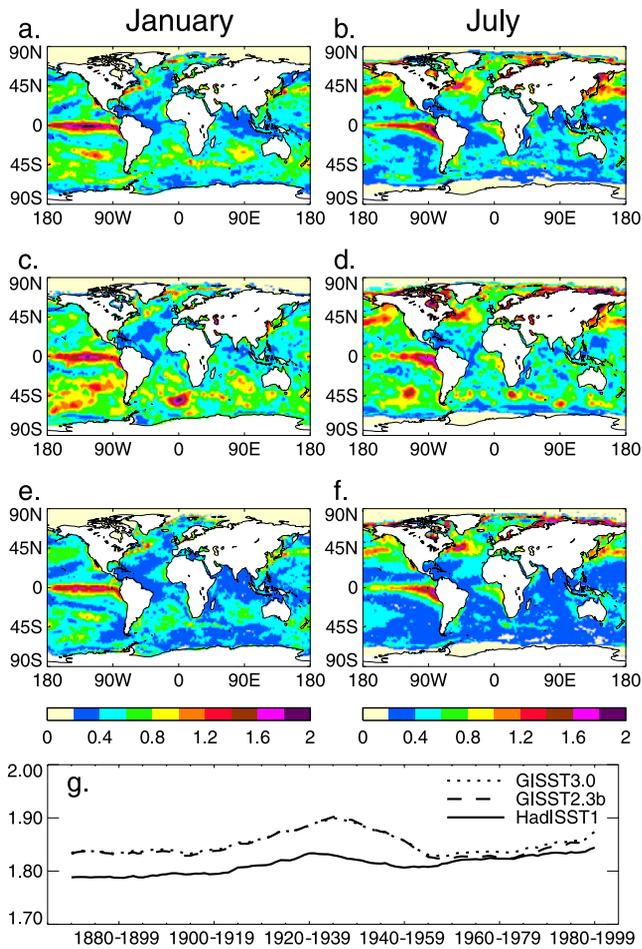


Figure 13. SST standard deviations ($^{\circ}\text{C}$), 1982–1999 in: (a–b) OI.v2; (c–d) GISST3.0; and (e–f) HadISST1 (left panels: January and right panels: July). (g) Time series of global root mean square 1° area standard deviation in running 20-year periods for GISST2.3b (dashed), GISST3 (dotted), and HadISST1 (solid).

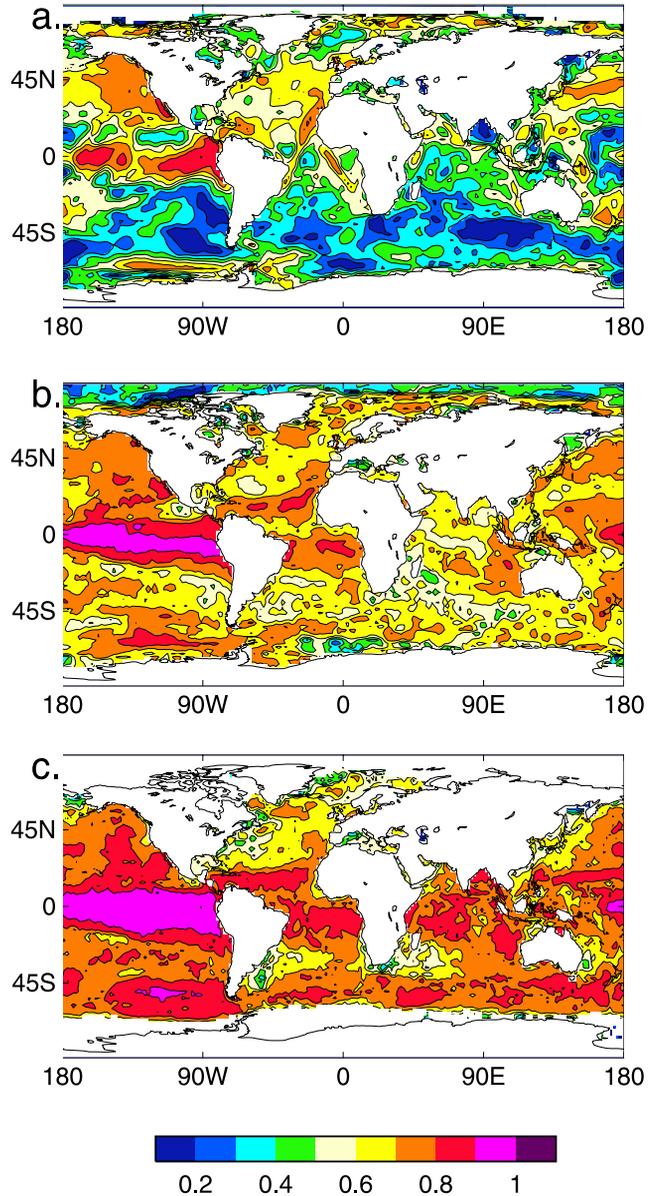


Figure 14. One-month lag autocorrelation of detrended monthly sea surface temperature anomalies in each 2° area grid-box, 1982–1999, for (a) GISST3.0, (b) OI.v2, and (c) HadISST1.

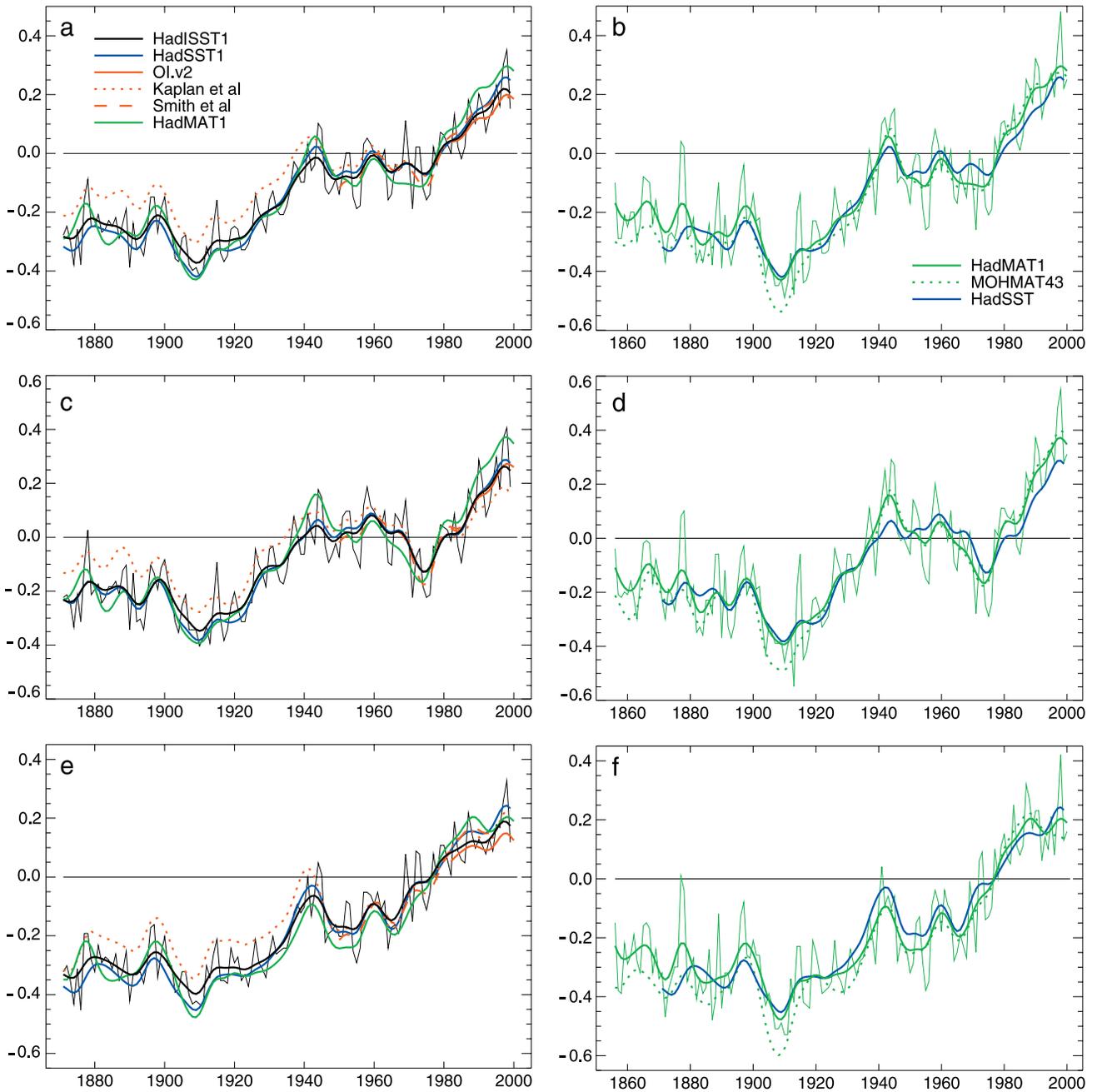


Figure 15. Smoothed hemispheric and global average temperature anomalies, relative to 1961–1990. Left panels: HadISST1, other SST data sets and HadMAT1, 1871–2000. Right panels: HadMAT1, MOHMAT43N and HadSST, 1856–2000. (a and b) Globe; (c and d) N.Hem.; and (e and f) S.Hem. All grid-boxes with data are used. A 21-point binomial (near-decadal) filter was applied to annual anomalies. Thin curves are unsmoothed annual averages: HadISST1 (left) and HadMAT1 (right). “Smith et al.” refers to *Smith et al.* [1996], and “Kaplan et al.” refers to *Kaplan et al.* [1998, 2003].

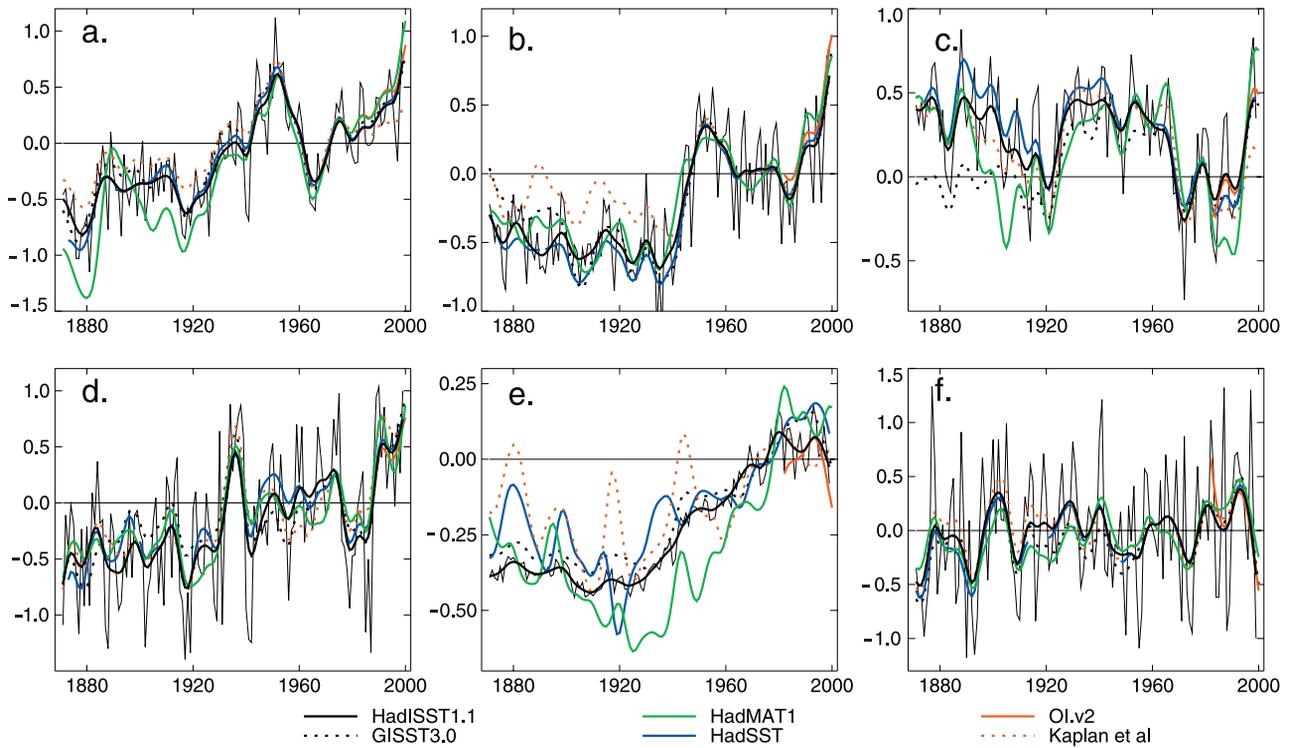


Figure 16. As left panels of Figure 15 (i.e., using all ice-free grid boxes with data) but for: (a) Gulf Stream region (35°N to 45°N, 50°W to 70°W); (b) Kuroshio region (30°N to 40°N, 125°E to 160°E); (c) Greenland region (50°N to 70°N, 30°W to 70°W); (d) Baltic Sea; (e) Southern Ocean south of 50°S; (f) Pacific Niño 3.4 region (5°N to 5°S, 120°W to 170°W). Thin black curves are unsmoothed HadISST1 annual anomalies. “Kaplan et al.” refers to *Kaplan et al.* [1998, 2003].

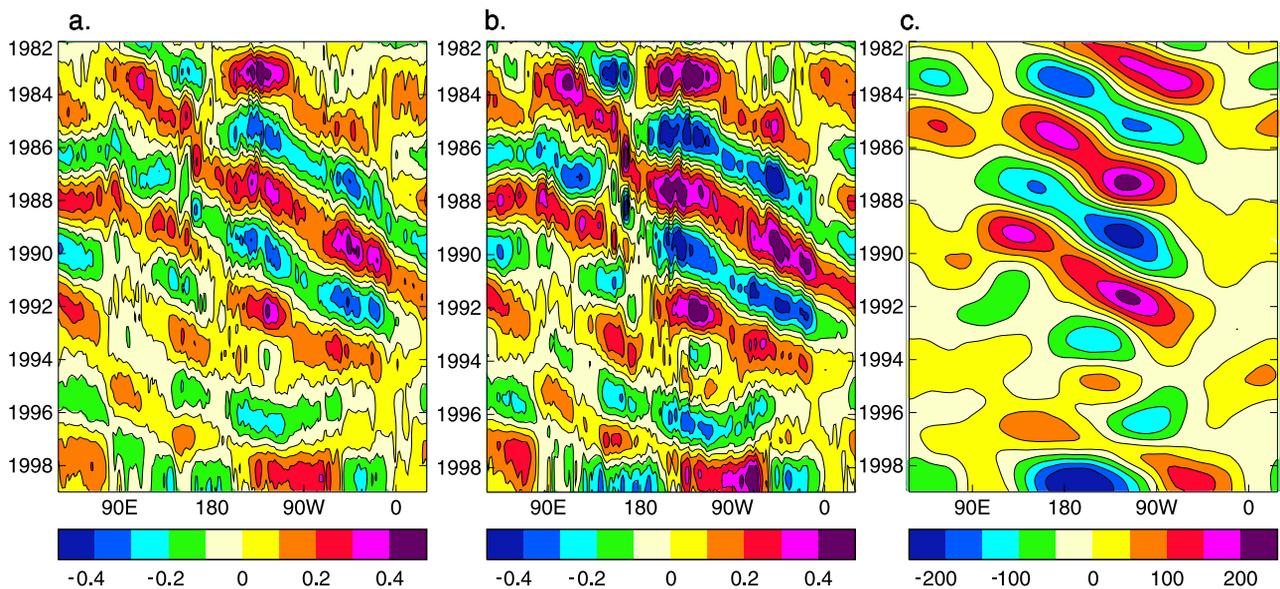


Figure 19. Monthly anomalies averaged over the belt 55° to 57°S, 1982–1998, 3–7 year band-pass filtered to show the Antarctic Circumpolar wave [*White and Peterson, 1996*]: (a) HadISST1 SST (°C); (b) OI.v2 SST (°C); and (c) mean sea level pressure (hPa) simulated by HadAM3 forced with HadISST1.