

Hemispheric and Large-Scale Surface Air Temperature Variations: An Extensive Revision and an Update to 2001

P. D. JONES AND A. MOBERG

Climatic Research Unit, University of East Anglia, Norwich, Norfolk, United Kingdom

(Manuscript received 1 March 2002, in final form 15 July 2002)

ABSTRACT

This study is an extensive revision of the Climatic Research Unit (CRU) land station temperature database that is used to produce a gridbox dataset of 5° latitude \times 5° longitude temperature anomalies. The new database comprises 5159 station records, of which 4167 have enough data for the 1961–90 period to calculate or estimate the necessary averages. Apart from the increase in station numbers compared to the earlier study in 1994, many station records have had their data replaced by newly homogenized series that have been produced by several recent studies. New versions of all the gridded datasets currently available on the CRU Web site (<http://www.cru.uea.ac.uk>) have been developed. This includes combinations with marine (sea surface temperature anomalies) data over the oceans and versions with adjustment of the variance of individual gridbox series to remove the effects of changing station numbers through time.

Hemispheric and global temperature averages for land areas developed with the new dataset differ slightly from those developed in 1994. Possible reasons for the differences between the new and the earlier analysis and those from the National Climatic Data Center and the Goddard Institute for Space Studies are discussed. Differences are greatest over the Southern Hemisphere and at the beginnings and ends of each time series and relate to gridbox sizes and data availability. The rate of annual warming for global land areas over the 1901–2000 period is estimated by least squares to be $0.07^{\circ}\text{C decade}^{-1}$ (significant at better than the 99.9% level). Warming is not continuous but occurs principally over two periods (about 1920–45 and since 1975). Annual temperature series for the seven continents and the Arctic all show significant warming over the twentieth century, with significant (95%) warming for 1920–44 for North America, the Arctic, Africa, and South America, and all continents except Australia and the Antarctic since 1977. Cooling is significant during the intervening period (1945–76) for North America, the Arctic, and Africa.

1. Introduction

The purpose of this paper is to revise, improve, and update the gridded land-based Climatic Research Unit (CRU) temperature database, last documented by Jones (1994). There are two principal reasons for such an analysis at the present time. First, nearly a decade has passed since it was last undertaken and significant changes and improvements have been made to the availability of monthly average temperature data in real time. The majority of the world's countries have endorsed the initiatives of the Global Climate Observing System (GCOS) to improve the quality and quantity of monthly climate data routinely exchanged. The GCOS Surface Network (GSN; see Peterson et al. 1997) has been in operation for nearly two years and several countries have either made more historic climate data available or extended the real-time availability or both.

The second reason is that several national and other initiatives have also dramatically improved the quantity

and quality of monthly mean temperature data available. Several countries have extensively homogenized their entire national holdings, releasing the results for all to use. CRU over the last eight years has also received several national and other temperature datasets. The quality of the earlier CRU gridded data (Jones 1994; Jones et al. 1999) has been criticized as not being of the best quality over parts of Europe (Böhm et al. 2001; Moberg and Alexandersson 1997), particularly for the pre-1900 period. Inclusion of these more recent datasets will alleviate these problems as some of the homogenized series result, in part, from these studies. Despite many data series coming directly from the National Meteorological Services (NMSs) the degree of homogeneity assessments made is variable. It can be assumed though that the checks performed are better than for data routinely exchanged between countries. Although these initiatives have only led to a modest improvement in real-time monitoring (such that similar updating will be necessary in another decade or so) the homogeneity and the quantity of data for much of the period since 1850 has been significantly enhanced. In particular, data for the 1961–90 period is now almost complete for the majority of records.

Corresponding author address: Prof. P. D. Jones, Climatic Research Unit, University of East Anglia, Norwich, NR4 7TJ, United Kingdom.
E-mail: P.Jones@uea.ac.uk

Despite these improvements, it is not expected that major changes will occur in the hemispheric-average series, as the quality of station data in the earlier (Jones 1994) analysis was high, due to extensive homogeneity assessments that had been performed earlier (see Jones et al. 1985, 1986). The principal reason, however, for expecting only small changes is that the time series of the many thousands of station records are not statistically independent of each other. The number of statistically independent locations over the earth's surface has been estimated by several authors to be about 100 (see discussion in Jones et al. 1997). The improvements to data quality in the present study, though, should impact individual gridbox series and analyses of spatial patterns.

The paper is organized in the following way. Section 2 extensively discusses the sources of additional data and the many difficulties of merging, replacing, and updating the existing station-based records. Section 3 briefly discusses the gridding technique used to develop the improved gridbox datasets. Section 4 presents extensive comparisons of the new analyses with those already available, illustrating the improvements in coverage. Section 5 discusses global-scale temperature variations and patterns after combination with marine data from the Hadley Centre at the Met Office. Conclusions follow in section 6.

2. Data

To augment the temporal and spatial coverage of the monthly mean temperature database of Jones (1994, hereafter Jones), data from several other datasets were used. These datasets can be basically divided into five categories.

- 1) *Homogenized records, beginning before 1900, covering an individual country or a regional grouping.* Several datasets have been developed in the past 10 years where data, generally beginning in the late nineteenth century, have been subject to extensive quality control, inspection of metadata, and homogenized using modern statistical techniques (see discussion of various techniques in Peterson et al. 1998a). Different homogeneity methods, however, have been used for each dataset. Such datasets are available for northwestern Europe and Greenland (Frich et al. 1996; Tuomenvirta et al. 2001), Austria (Auer et al. 2001), and Canada (Vincent and Gullet 1999). A few even longer homogenized records, starting in the eighteenth century, have also recently been made available for some countries across Europe, within inter-European research projects (Moberg et al. 2000; Camuffo and Jones 2002) and by some NMSs.
- 2) *Records established by NMSs, covering a part or the whole of the twentieth century.* CRU has collected a number of temperature records through direct con-

tacts with the NMSs in Algeria, Croatia, Iran, Israel, South Africa, Syria, and Taiwan. Many of these records cover only the period 1961–90, but others extend over the entire twentieth century. Data for the whole of Australia for 1991–2000 have also been collected through direct contacts.

- 3) *Datasets with global coverage of long climatic time series.* The Global Historical Climatology Network (GHCN) is a monthly temperature, precipitation, and pressure database containing several thousand land surface temperature stations (Vose et al. 1992; Peterson and Vose 1997). Using Jones and GHCN as their primary source, New et al. (1999, 2000, hereafter, collectively, NHJ) developed mean monthly terrestrial climatologies and time series data, including surface temperatures, with global coverage for the periods 1961–90 and 1901–96. We selected data from GHCN and/or NHJ for Australia, Mexico, the United States, and several countries in Africa, South America, and Asia where they could help to improve the temporal or spatial coverage in Jones.
- 4) *Global datasets covering the period after 1981.* We used data from the World Weather Records (WWR) 1981–90, data internationally exchanged over the World Meteorological Organization (WMO) official international climate monitoring (CLIMAT) network 1991–2001 and through the WMO monthly publication Monthly Climatic Data for the World (MCDW), to fill in, for the last two decades, data that were not included in the other data sources. WWR data for 1981–90 are the latest of the decade publications developed for WMO by the National Climate Data Center (NCDC) in Asheville, North Carolina. In this source, data are listed together in six volumes (continent-wide for North America, Europe, Asia, Africa, South America, and Islands, which includes Australia and the Antarctic). WWR data for 1981–90 have yet to be included within GHCN. Data for previous volumes of WWR are incorporated within GHCN and hence in NHJ.
- 5) *Data for Antarctica.* Jones and Reid (2001) developed a databank of Antarctic surface temperature and pressure data, with data often beginning about 1957. This is the most comprehensive databank with Antarctic temperatures. Several Antarctic records are available through the Web sites of the national Antarctic institutes that maintain stations on the continent.

The Jones dataset contained over 3900 temperature records, of which 2961 stations were used in the gridding. By including data from the other sources, a new dataset has been developed that contains 5159 stations. The new dataset contains 1246 stations that were not represented in Jones. Of these extra stations, 623 come from various countries spread over all continents—excepting the contiguous United States. The remaining 623 stations come from sites in the contiguous United

States. All temperature records that were not represented in Jones were simply added to the Jones dataset here.

We also merged data from a large number of stations that were already represented in Jones. When merging these records with Jones, a decision had to be made whether to overwrite the data in Jones, or to use the other data only to fill in missing values in Jones and to add data before/after the corresponding series in Jones began/ended. When the other data were used to overwrite Jones data, we say that we merged with priority for the other record, whereas if they were used to include missing values we say that we merged with priority for Jones. Data in category 1 and 5 were always merged with priority for the other data as they were considered to be the best available version of the respective series. For the other categories, a decision was made concerning priority, after analyzing the month-to-month differences (other data minus Jones) for overlapping periods. Inspection of these differences revealed systematic or random discrepancies between the two versions of the same data, and a decision on priority was made for each station. This analysis was both laborious and time consuming. After completion of the merging of all new data, a check for outliers was made. This outlier check is described briefly in section 3.

A particular problem we encountered, occurred when the data in Jones for a certain station began before the corresponding series in the other dataset. The homogeneity adjustment in the early part of these new records sometimes differed between Jones and the other data, and, hence, we had to apply adjustments to a few early Jones series to ensure that no nonclimatic discontinuities were introduced if the other records were merged with priority. This problem was encountered for some stations with data in the nineteenth century, and corrections had to be applied to early Jones data for a number of Canadian stations, and for a few others in various continents.

In the following, we give some details on how the various data were merged and discuss some particular problems encountered. A summary of the number of stations merged, the datasets, countries, and periods involved is given in Table 1. Some of the merged data series have been adjusted for inhomogeneities, using a variety of techniques (see discussion in Peterson et al. 1998a). It is difficult to say exactly how many series have been adjusted as some of the component datasets do not include this information. Furthermore, only a part of a series, sometimes as few as 10 years, may be adjusted.

a. Homogeneous records for northwest Europe and Greenland

Within a cooperative program among the NMSs in Denmark, Sweden, Finland, Norway, Iceland, the United Kingdom, Ireland, the Netherlands, and Belgium, a dataset of homogenized instrumental climatic records,

generally covering the period 1890–1990, was developed for stations in these countries, additionally including Greenland, Svalbard, and the Faroe Islands (Frich et al. 1996). This dataset was named the North Atlantic Climatological Dataset (NACD). Recently, the NMSs in Finland, Norway, Sweden, Iceland, and Denmark released an updated version (to 1999), named NORDKLIM, of their part (Tuomenvirta et al. 2001). We selected 89 NACD/NORDKLIM temperature records, of which 66 were included in the NORDKLIM update. Some of the stations in NORDKLIM have data beginning before 1890, and we obtained these earlier data as well as some updates through 2000 directly from the respective NMSs. There were already 43 of the NACD/NORDKLIM stations represented in Jones, although often with shorter records and with less extensive homogeneity assessments. We merged the new versions with Jones, giving priority to NACD/NORDKLIM.

b. Very long European temperature records

The European project IMPROVE (Moberg et al. 2000; Camuffo and Jones 2002) has recently constructed daily temperature and pressure series for seven selected sites (Padua, Milan, Cadiz, Brussels, Stockholm, Uppsala, and St. Petersburg) with records beginning in the eighteenth century. Padua and Cadiz were not represented in Jones. All these series have undergone extensive homogeneity assessment on a daily basis. All daily IMPROVE series have been released on a CD-ROM together with the original observational data (Camuffo and Jones 2002). We calculated monthly mean temperatures from the homogenized daily records.

Another project (ALOCLIM) concentrated on creating a homogenized multivariable dataset of long Austrian records, which was also released on a CD-ROM together with extensive metadata information (Auer et al. 2001). This dataset contains 15 monthly temperature records, of which 3 (Kremsmünster, Wien, and Innsbruck) start in the eighteenth century and most of the rest in various years in the nineteenth century. Nine stations in ALOCLIM were not represented in Jones.

A few other European monthly temperature records beginning in the eighteenth century have been homogeneity assessed by their respective NMSs; De Bilt (the Netherlands; van Engelen and Nellestijn 1996), Hohenpeissenberg (Germany), and Warsaw and Krakow (Poland). These four long European NMS records were already represented in Jones. The records of IMPROVE, ALOCLIM, and the long NMS records were merged with priority for these data. The newly released versions are more complete, almost always longer, and more extensively homogenized.

A particular problem connected with future updating of the Austrian ALOCLIM series needs to be pointed out. The ALOCLIM mean temperatures are adjusted to the corresponding mean values calculated from 24-hourly observations. Some of the ALOCLIM stations are

TABLE 1. Overview of the various datasets merged with Jones (1994), priority in merging, main countries and regions involved, and generalized time periods of temperature records merged. Dataset abbreviations and their proper references are given in the main text. The number of stations denotes the number of stations selected from each dataset used in the merging. Asterisks (*) after number of stations in the columns labeled "Already in Jones" denote that the stations were either already represented in Jones (1994) or that our new dataset had received data for some/several of the stations after the merging of datasets placed at a higher level in the table. The numbers of stations in the various table cells cannot be added to get the total number of stations merged because several stations are represented in more than one cell.

Dataset	No. of stations		Main countries or regions	Approximate time periods
	Not in Jones	Already in Jones, priority: other	Already in Jones, priority: Jones	
NACD and NORD-KLIM	46	43		Scandinavia, northwest Europe, Greenland
IMPROVE and NMSs very long	2	9		Italy, Spain, Belgium, Sweden, Russia, Germany, the Netherlands, Poland
ALOC LIM	9	6		Austria
CHTD	98	112		Canada
Jones and Reid	3	26		Antarctica
(2001) Web sites				
NHJ	237	69		Australia
Bureau of Meteorology	173	71		Australia
NHJ	623	438	138	United States
NHJ	150	52		Former Yugoslavia, other European, Israel
				Several countries in east and Southeast Asia, Africa, Alaska, Central America, South America
NHJ			49	Various countries in Europe, Africa, South America, south, central, and Southeast Asia
				Islands in Pacific and Atlantic Oceans
NMSs	54	19		Iran
				Taiwan, Algeria
NMSs			36*	Croatia, Syria, Israel, South Africa
GHCN	24			Bangladesh, Mexico, Central America, Philippines
GHCN		235		Former Soviet Union, Mexico, Brazil, Peru
GHCN			236*	Several countries in Asia except former Soviet Union, all of Africa, Central and South America, islands in Pacific and Atlantic Oceans
WWR		1023*	217*	All countries that have stations with data already in the new dataset, except stations in NACD, NORDKLIM, IMPROVE, ALOCLIM, CHTD
CLIMAT		1548*		All countries that have stations with data already in the new dataset, except stations in NORDKLIM, IMPROVE, ALOCLIM, CHTD, and some NMS-direct data
GHCN			1335*	All countries that have stations with data already in the new dataset
CLIMAT and a few other sources			~1300	All countries that have stations with data already in the new dataset

also among the stations for which data are available in real time through the CLIMAT network. Austrian mean temperatures on CLIMAT, however, are calculated by taking the average of the maximum and minimum temperatures, and, hence, they will differ slightly from the values calculated in ALOCLIM (Auer et al. 2001). For our future updating of the ALOCLIM stations, we will have to obtain data directly from the Austrian NMS once a year, which will replace real-time temperatures obtained from CLIMAT.

c. Homogeneous Canadian temperature data

Vincent (1998) and Vincent and Gullet (1999) released a set of 210 Canadian minimum and maximum monthly temperature records, called the Canadian Historical Temperature Database (CHTD). Their records from southern Canada often start in 1895, whereas records from northern parts are shorter due to the later start of observing stations there (often in the 1930s–50s). Many records come from stations with continuous

measurements at one and the same site, whereas some records have been assembled by joining a few shorter segments from nearby stations to form one long record, generally covering the entire period 1895–1999. All records have been homogenized and generally adjusted to the conditions of the most recent location. Mean temperatures were calculated from the monthly maxima and minima, and merged with Jones with priority for CHTD. Data for 112 of the 210 stations were already represented in Jones, but the Jones data were overwritten by the more extensively homogenized CHTD data during the merging. We had access to the February 2001 version of CHTD (provided personally by Vincent), which was updated through 1999.

In a few cases, Jones had records that started earlier than CHTD (i.e., often before 1895). Some of these early data in Jones had to be adjusted to a level corresponding to that of the earliest CHTD data, because some early CHTD data had been adjusted by Vincent and Gullet (1999). We accomplished this adjustment by adding constant values for each month to the Jones pre-CHTD data. Vincent provided us (L. A. Vincent, 2001, personal communication) with the necessary adjustment values. In some of the actual cases, Jones had already applied corrections to his early data, derived during homogeneity checks made in the 1980s (Jones et al. 1985). In these latter cases we adjusted the Jones pre-CHTD data by adding the difference between the Vincent–Gullet and Jones et al. 1985 corrections.

Vincent and Gullet (1999) noted that the Canadian minimum temperature series exhibit an artificial discontinuity in 1961, due to a change of hour when the minimum temperatures were observed. This change introduced a bias in minimum temperatures of ca. -0.6°C for stations in eastern Canada after 1961, whereas the effect was smaller (less than -0.2°C) in the western part. Vincent and Gullet (1999) corrected the minimum temperatures after 1961 for 36 stations in eastern Canada, so that they correspond, on average, to the conditions before 1961. The fact that minimum temperatures for these 36 CHTD records are adjusted in their most recent period, means that mean temperatures calculated from CHTD will differ from mean temperatures in Canadian CLIMAT messages. This, in turn, means that any future updating of the CHTD series using CLIMAT data must involve the application of a correction. L. A. Vincent (2001, personal communication) has provided us with a table of corrections that we use when updating with CLIMAT data.

d. Australia

Temperature records from 306 stations in Australia were obtained from NHJ. Of these stations, 237 were not represented in Jones. The Australian records in NHJ began in various years, some in the late nineteenth century, but most in the twentieth century at some point before 1960. The vast majority of records end in 1992.

Merging was undertaken with priority for NHJ. To update the Australian records, we used data for the period 1991–2000 from 244 stations contained on a CD-ROM obtained directly from the Australian NMS (Bureau of Meteorology 2000).

As with the cases of Austrian and Canadian data mentioned above, there is a discrepancy between the recent data we used and the CLIMAT messages. Prior to November 1994, the Australian CLIMAT mean temperatures were calculated as the average of the monthly means of daily temperature maxima and minima. In November 1994, the averaging method was changed to the use of averages of all available synoptic observations, which is in general at 3-hourly intervals. This change has introduced an inhomogeneity in the CLIMAT data that varies from station to station, but for the whole of Australia is about a 0.2°C cooling (B. Trewin et al. 2001, personal communication) for the post-November 1994 data compared to before. The Australian rationale for changing the method of mean temperature calculations used for CLIMAT was that from November 1994, the CLIMAT messages additionally included the mean monthly maxima and minima. It is possible, therefore, to calculate consistent monthly mean temperatures, but it appears that the change was not widely reported and only a few within the Bureau of Meteorology in Australia knew that the method of mean temperature calculation in the CLIMAT message had been changed.

The mean temperatures we obtained from the CD-ROM were calculated according to the earlier method. Hence our Australian records are homogeneous until the latest date on the CD, which is generally in the austral autumn 2000. We will obtain future updates of Australian mean monthly maximum and minimum temperatures directly from the Australian NMS, while this problem persists.

e. Contiguous United States

The temperature data coverage over the contiguous United States throughout the twentieth century is denser than almost anywhere else in the world, and there were already a great number of stations in Jones. Despite this, we augmented this part of the database by adding, from NHJ, data for 597 stations without WMO-number designations and 26 stations with WMO numbers that were not represented in Jones. Additional data for 438 non-WMO stations, which were already represented in Jones, were merged with priority for NHJ. We also merged (from NHJ) 138 stations with WMO numbers that were already in Jones. Merging of the latter series was made with priority for Jones, so the NHJ data were in these cases only used to fill in missing values and to add early/late data that Jones did not already have.

f. Other countries

We selected from NHJ data for 251 stations from various countries in all continents, of which 150 were

not represented in the previous dataset of Jones. Of the 101 stations that were already represented in Jones, priority was given to NHJ in 52 cases. These records principally came from South Africa, Alaska, countries in the Middle East, eastern Asia, South America, and a few countries in eastern and central Europe. The remaining 49 records were merged with priority for Jones. The latter mainly came from Mongolia, various countries in Africa and South America, and a few islands in the Atlantic and Pacific Oceans.

CRU has received data from a number of countries through direct contacts with the respective NMSs. Records for 54 stations that were not represented in Jones were now included. Data for an additional 19 stations already represented in Jones were merged with priority for the NMS source. Most of these NMS records came from Iran, Algeria, Taiwan, Croatia, Israel, South Africa, and Syria. The merging of the NMS data was made after those from NHJ had been merged. As some stations were represented both in the NMS and NHJ data, the priority during merging was sometimes given to the data from NHJ rather than to data originating from Jones.

The other main source of data from various countries was the GHCN dataset. GHCN is released in two versions: “v1” (Vose et al. 1992) and “v2” (Peterson and Vose 1997). Version 2 includes data for a larger number of stations compared to v1, and has also more detailed assessments of data quality, more rigorous homogeneity adjustments, and more detailed metadata. Furthermore, GHCN v2 is available both in an unadjusted and an adjusted version. In the adjusted version of v2, adjustments have been made for various kinds of biases and inhomogeneities (Peterson et al. 1998a). The adjusted version contains a smaller number of stations than the unadjusted version. Both the adjusted and unadjusted versions have passed quality controls outlined by Peterson and Vose (1997).

In GHCN there can exist more than one duplicate (in their terminology) of data for a single station. Different duplicates refer to different data sources, which may have used different methods for calculating monthly mean temperatures. The duplicate corresponding to the longest record is always labeled with “0” (their terminology). As we were essentially interested in obtaining long records, we selected data with duplicate number 0. These records, however, often end in 1990 or slightly earlier, so data for the last decade were missed in this selection. A separate selection was therefore made for GHCN data in the last decade. This selection is described in section 2i.

Records for 495 stations were selected from the unadjusted GHCN v2. Only 24 of these stations were not represented in Jones. Data for 235 stations were merged with priority for GHCN. The main regions where these records came from were countries of the former Soviet Union, Mexico, Brazil, and Peru. We also merged 236 records with priority for Jones, mainly from Saudi Arabia, Pakistan, Bangladesh, Mongolia, countries in south-

eastern Asia, Africa, and Central and South America. As the merging of GHCN data was made after the merging of the NHJ data, and as some stations occurred both in our selections of GHCN and NHJ data, the priority in a few cases was given to data from NHJ rather than to data in Jones. As NHJ used the adjusted v2 of GHCN as one of their primary data sources, the way we merged GHCN data here gave priority to the adjusted GHCN v2 rather than to the unadjusted v2 in some cases.

g. *Antarctica*

The databank for Antarctica developed by Jones and Reid (2001) contains data from 29 meteorological stations, many including over 40 years of data. Most of the stations are located around the Antarctic coastline, but there are a few from research stations on the ice sheet (Siple, Byrd, Amundsen-Scott, Vostok). The database has been made available online by the Carbon Dioxide Information Analysis Center (CDIAC; <http://cdiac.esd.ornl.gov/epubs/ndp/ndp032/ndp032.html>). The primary data sources used by Jones and Reid were WWR (1951–90) and the MCDW (since 1961). Jones and Reid (2001) updated each series by requesting data directly from the responsible NMSs and national Antarctic institutions. The entire databank of Jones and Reid was merged with Jones data, giving priority to Jones and Reid. Additional data were obtained for a few stations, principally in the Antarctic Peninsula region from a Web page of the British Antarctic Survey (<http://www.antarctica.ac.uk/met/gjma>).

h. *World Weather Records (WWR) 1981–90*

After having merged all the datasets mentioned above, the total number of stations had reached its final value of 5159. To further improve the quality and temporal coverage in the decade 1981–90, we used data from WWR. This contains data that have generally been delivered directly from each NMS to NCDC, Asheville. A great number of temperature values in WWR 1981–90 had not been included in the other datasets. We selected WWR 1981–90 data for 1240 stations from all over the world. In this selection we were only interested in getting data for stations that were already included in our new dataset, so all other stations in WWR 1981–90 were discarded. WWR data for stations that were included in the NACD, NORDKLIM, IMPROVE, AL-OCLIM, or CHTD datasets were ignored. Data for 1023 stations were merged with priority for WWR, meaning that the WWR data overwrote the other data, regardless of the earlier source. By giving priority to WWR for all these stations, a number of outliers and erroneous data were replaced by better values.

In our analyses of the differences (WWR minus Jones in the 1981–90 period), we observed a systematic error in the WWR files for some countries in the year 1990. Comparison of WWR data for other recent years re-

vealed that temperature data for 1992 had erroneously been inserted instead of data for 1990. Among the countries where this occurred were Saudi Arabia, Afghanistan, Laos, Western Sahara, Morocco, Tunisia, Senegal, and Zimbabwe. The source of data where this occurred was given as the World Monthly Surface Station Climatology from the National Center for Atmospheric Research (NCAR) Boulder, Colorado. This implies that they are essentially CLIMAT data and have not been sent in delayed mode from the NMSs. So for 217 series we merged WWR data with priority for Jones. In these cases we considered our data (after having merged all other datasets) more reliable because of this 1990 problem. We excluded WWR data for 1990 from all affected stations in the merging, in order not to fill in any missing data we had for 1990 with 1992 data.

i. CLIMAT and GHCN data for 1991 onward

Additional data for the last decade were obtained from the international exchanges of the WMO CLIMAT network. We only selected data for those stations that were already represented in our new dataset, but excluded those where data had been obtained from NORDKLIM, IMPROVE, ALOCLIM, and CHTD and also some of the directly delivered NMS records. In total, 1548 stations with data for the last decade were selected and merged with priority; that is, they overwrote the data we already had in the new dataset.

By including CLIMAT temperatures, most records were made as complete as possible in the last decade. Some stations, however, were not included in the CLIMAT files. To enable more stations to be updated, a final selection of data was made from GHCN (unadjusted v2 updated through July 2001). We obtained data from 1991 onward, by selecting data with GHCN duplicate numbers larger than 0 for stations that were already represented in our new dataset. For several stations, there was more than one duplicate available. In these cases we chose the one with the largest amount of data in the period 1991–2001, and filled in missing values with data from the other duplicates. The resulting selection of GHCN updates consisted of 1335 stations, which were merged with our new dataset with priority given to our data. The GHCN 1991–2001 updates were therefore mainly used to include missing values, but they also filled in nearly 90 stations for which we had found no data in the other sources.

Despite our attempts to improve coverage from 1991 onward, the number of available stations reduces significantly compared to 1990. Another major reduction occurs after 1996 when part of our network for the United States stops. The effect these reductions have on gridbox averages will be discussed in section 3. Reductions in station numbers are greatest during 2000 and 2001, years for which we do not have access to all data from sources such as NORDKLIM, IMPROVE, ALOCLIM, and CHTD, for example, nor from the di-

rectly supplied NMS sources. It is clear, however, that the number of available CLIMAT messages has increased slightly after 1995 compared to 1991–94 as a result of countries including real-time data from their newly designated GCOS Surface Network (GSN) stations. The rationale behind the initial selection of stations in the GSN is given in Peterson et al. (1997). The availability of extra GSN stations is patchy and increases are only evident for a limited number of countries. For a few countries, such as Australia, they are important because they mean more grid boxes have data.

3. Interpolation to a regular grid

All analyses of large-scale temperatures recognize that the irregular distribution of the available station data requires that some form of gridding is necessary, in order for analyses (e.g., hemispheric averages, principal component patterns) not to be biased. Three different methods have been proposed to interpolate station temperature data to a regular grid (Peterson et al. 1998b). The most widely used method is termed the climate anomaly method (CAM; e.g., Jones), with the other two being the reference station method (RSM; Hansen and Lebedeff 1987) and the first difference method (FDM; Peterson et al. 1998b).

Analyses of direct comparisons of the three approaches with the same basic data have been discussed by Peterson et al. (1998b). Possible differences between the techniques and advantages/disadvantages of each are also discussed by Jones et al. (1999). In this study we will use the CAM approach, which requires reducing all the station temperature data to anomalies, from a common period such as 1961–90 on a monthly basis. Gridbox anomaly values will then be produced by a simple averaging of the individual station anomaly values within each grid box.

The major disadvantage of CAM is that stations must have enough years with data within the 1961–90 period in order to be used, although even this constraint can be overcome by judicious use of neighboring series and other periods (e.g., 1951–70 and 1951–80; see Jones). Improvements to coverage, particularly since 1980 for many regions of the world (discussed in section 2), now mean that many more stations have near-complete monthly data for the 1961–90 period (often referred to as *normals*). For some stations with incomplete data for 1961–90 it will be possible to use published 1961–90 normals (WMO 1996), although care is required. We noticed several (~10) instances where WMO normals appeared unrelated to both our time series data and that of GHCN. Part of the problem relates to changes to WMO numbers in some countries and to renumbering in the countries of the former Yugoslavia.

Improvements to the base period should also mean that monthly averages for 1961–90 will sum to zero, for many more grid boxes than was evident in the earlier (Jones) analyses. In this earlier study normals were cal-

culated based on at least 20 years within the 30-yr period. As the 1981–90 period was more likely to contain missing temperature data, normals calculated for 1961–90 were often biased slightly cold so anomaly averages (for 1961–90) calculated for the hemispheres were slightly positive (by 0.01° – 0.05°C depending on the month).

The potential advantages of RSM and FDM are that they can use all series that have at least 10 years of data, because normals are not required. RSM achieves this by selecting the longest station in each grid box and then adding shorter duration series adjusted so that their averages equal the composite of all stations already incorporated over the overlap with the composite. FDM uses more data than CAM by working with the station first difference time series [e.g., $\text{Jan}(t)$ minus $\text{Jan}(t - 1)$], averaging series together within grid boxes. Comparable series to CAM and FDM are produced by cumulating the first difference averages.

Peterson et al. (1998b), in comparisons of the results (in terms of hemispheric and global averages) of the three techniques, indicate that they all produce similar series. Differences between estimates of trends at these scales, over the 1880–1990 period, are comparable to differences that might result due to calculating linear trends using different methods. They also noticed that CAM and FDM produced similar variances of hemispheric averages, but that variances of RSM were significantly smaller. This is likely due to differences in gridbox size ($5^{\circ} \times 5^{\circ}$ in this application of CAM and FDM compared to the 80 equal-area boxes for the world used by Hansen et al. 1999, 2001). Folland and Karl (2001), in comparisons of the three approaches for the third Intergovernmental Panel on Climate Change (IPCC) report, also noted significant and potentially more important differences between the three analyses. Since the late 1980s (see their Fig. 2.1b on p. 107) the FDM global series (Peterson et al. 1998b, also referred to as the NCDC dataset) warms compared to CAM (Jones) and RSM [Hansen et al. 1999, 2001, also referred to as the Goddard Institute for Space Studies (GISS) dataset]. Given that data availability during the 1990s is likely to be similar, being based on the CLIMAT system from the Global Telecommunication System (GTS) of WMO and the publication MCDW, what is the likely cause of the differences? We will return to this issue in section 4.

a. Development of 1961–90 normals and outlier checks

Monthly averages for 1961–90 (the latest WMO normal period) were calculated from the enhanced station dataset, accepting an average if at least 20 years of data were available, 4 of which must be in each of the three decades. For stations where this was not possible WMO (1996) normals were used if available. For a further set of stations, 1961–90 normals were estimated using the

earlier Jones gridded analyses. Here, 1951–70 averages (requiring at least 15 years of data) from the enhanced dataset were calculated and adjusted by the difference of the gridbox averages for 1961–90 and 1951–70 from the earlier Jones analyses. Finally 1961–90 normals were used directly from Jones for a small number of stations. Altogether 1961–90 normals were developed for 4167 stations, of which 3348 were calculated directly, 291 from WMO (1996), 482 using 1951–70 averages, and 46 from the earlier Jones analysis. Temperature data for any station without 1961–90 normals will not be used in subsequent gridding.

Section 2 has extensively discussed the sources of the additional temperature data and the priorities given to the different sources in the merging of the new datasets. Although many of the sources have undergone detailed homogeneity testing there is still the possibility of outliers, which might induce a larger influence if they occur during the 1961–90 period. To assess outliers we have also calculated monthly standard deviations for all stations with at least 15 years of data during the 1921–90 period. All outliers in excess of five standard deviations from the 1961–90 mean were compared with neighbors and accepted, corrected, or set to the missing code. Correction was possible in many cases because the sign of the temperature was wrong or the temperature was clearly exactly 10° or 20°C too warm or cold. After this step the 1961–90 normals and the 1921–90 standard deviations were recalculated. Exactly how many values were corrected or set to missing is impossible to determine, as we do not know how many were corrected in the component datasets. We made changes to about 500 monthly averages, with approximately 80% corrected and 20% set to missing. In terms of the total number of monthly data values in the dataset this is a very small amount ($<0.01\%$).

b. Gridding

Each station was first associated with its $5^{\circ} \times 5^{\circ}$ grid box and gridbox anomaly values calculated by simple averaging of all available station anomaly values for all months 1851–2001, exactly as in Jones. All outliers in excess of five standard deviations were omitted from the analysis. This latter condition only affects years before 1961 and after 1990.

Figure 1 illustrates both the number of stations used each year and the percent coverage this produces for each hemisphere. The results are compared with the earlier Jones analysis. For station numbers significant improvements are seen in both hemispheres, particularly for the later years of the 1970s to the 1990s. For the Southern Hemisphere (SH) this increase approaches a doubling of the number of stations. The results can be a little misleading as many of these result from the increase in Australian data. For the Northern Hemisphere (NH) the increases partly occur through the use of more U.S. data, which also explains the large drop

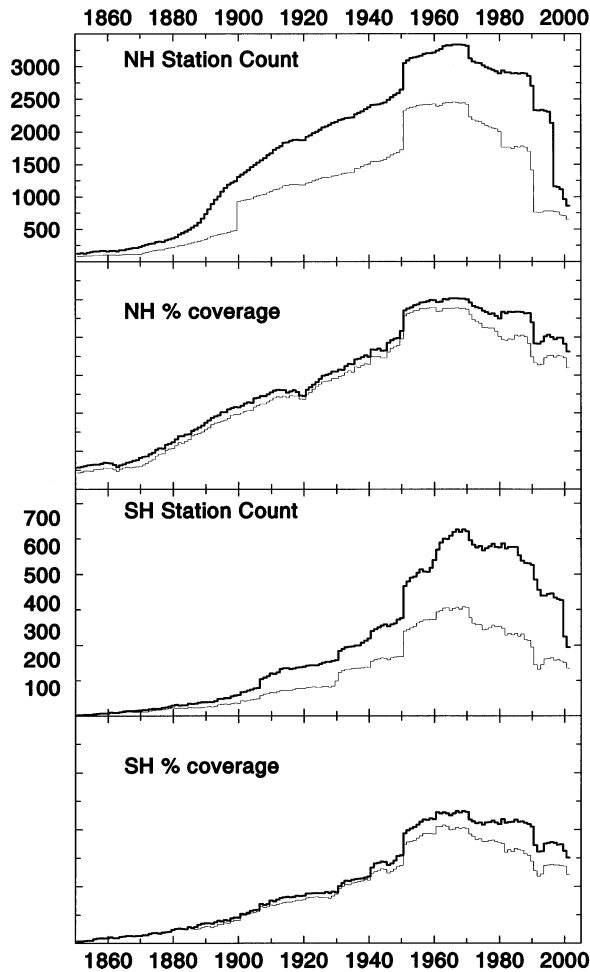


FIG. 1. Station availability by hemisphere and percent hemispheric coverage for the dataset developed in the present study (CRUTEM2) compared to Jones (1994). The present study is represented by the thicker lines with Jones (1994) being thinner. The top and third panels are station counts for the NH–SH, respectively, with the second and fourth panels being NH–SH percent coverage.

in station counts after 1996. Despite this, all contiguous U.S. grid boxes still contain at least two stations after 1996, a number that is more than in most other regions of the world. It would have been possible to access post-1996 data, but we have concentrated our efforts, in locating additional data, in the more poorly sampled regions of the world. Both hemispheric station counts still exhibit sharp increases and reductions at decade starts and ends, principally due to the availability of data in the WWR decade yearbooks. Reductions after 1990 will be reduced in the future when the decade books for the 1991–2000 period become available. It is to be hoped that these will be produced more speedily than in recent decades.

In terms of percent coverage, the improvements in coverage have had a smaller effect than in terms of station numbers, with the increase being greater in the SH compared to the NH. Both hemispheres show a slight

reduction in the percent coverage in the early 1990s, compared to after 1994, illustrating the slightly improved data availability from the CLIMAT network after 1994. A reduction of about 7% for the NH and 5% for the SH occurs during 2001 compared to 1999. Late data, available during 2002, will likely reduce this slightly. For both hemispheres the optimum percent coverage is larger than the total land area, because we count the total gridbox area even if the data is from a single, small island. The slight reduction around 1920 in the NH is due to more missing data from Russia at this time.

In the dataset terminology of Jones et al. (2001) the gridded version developed here will be referred to as CRUTEM2. Also following the method outlined in Jones et al. (2001) a variance-corrected version has also been developed (CRUTEM2v). Variance correction adjusts each gridbox series so that time series values are not influenced by changing numbers of contributing stations through time, each gridbox series being reduced to the variance of the infinitely sampled gridbox series (see Jones et al. 2001 for details).

CRUTEM2 and CRUTEM2v have been combined with unadjusted and variance-adjusted Met Office sea surface temperature (SST; versions MOHSST6D and HadISSTv, respectively) anomaly data (producing HadCRUT2 and HadCRUT2v) using the method detailed in Jones et al. (2001). A few brief results from this combination are discussed later in section 5, but the main emphasis of the rest of this study (section 4) is the analysis of the land-only data.

4. Analysis of the enhanced gridded land data

In this section we consider continental and hemispheric averages, comparing the latter with the analyses from NCDC (Peterson et al. 1998b) and GISS (Hansen et al. 1999, 2001). The enhanced dataset used in this section is CRUTEM2v, which is considered the preferable version, compared to CRUTEM2, because analyses from periods with sparser data should be more reliable as a result of the variance adjustment. Spatial patterns of trends are discussed in section 5, where the combined land and marine dataset, HadCRUT2v, is used to provide more complete spatial coverage.

a. “Continental”-scale analyses

Eight regions are defined (see Fig. 2 and Table 2) for the seven continents of the world, with an additional one for the Arctic. The “continents” are defined by rectangles in 5° latitude \times 5° longitude space and they do not overlap. They are only termed continents as a naming convenience. None conforms to a strict geographical definition of a continent.

Each continental average is the weighted (using the cosine of the midgridbox latitude) average of all grid boxes with data in a given month. The rectangular nature of the areas implies that they include all islands with

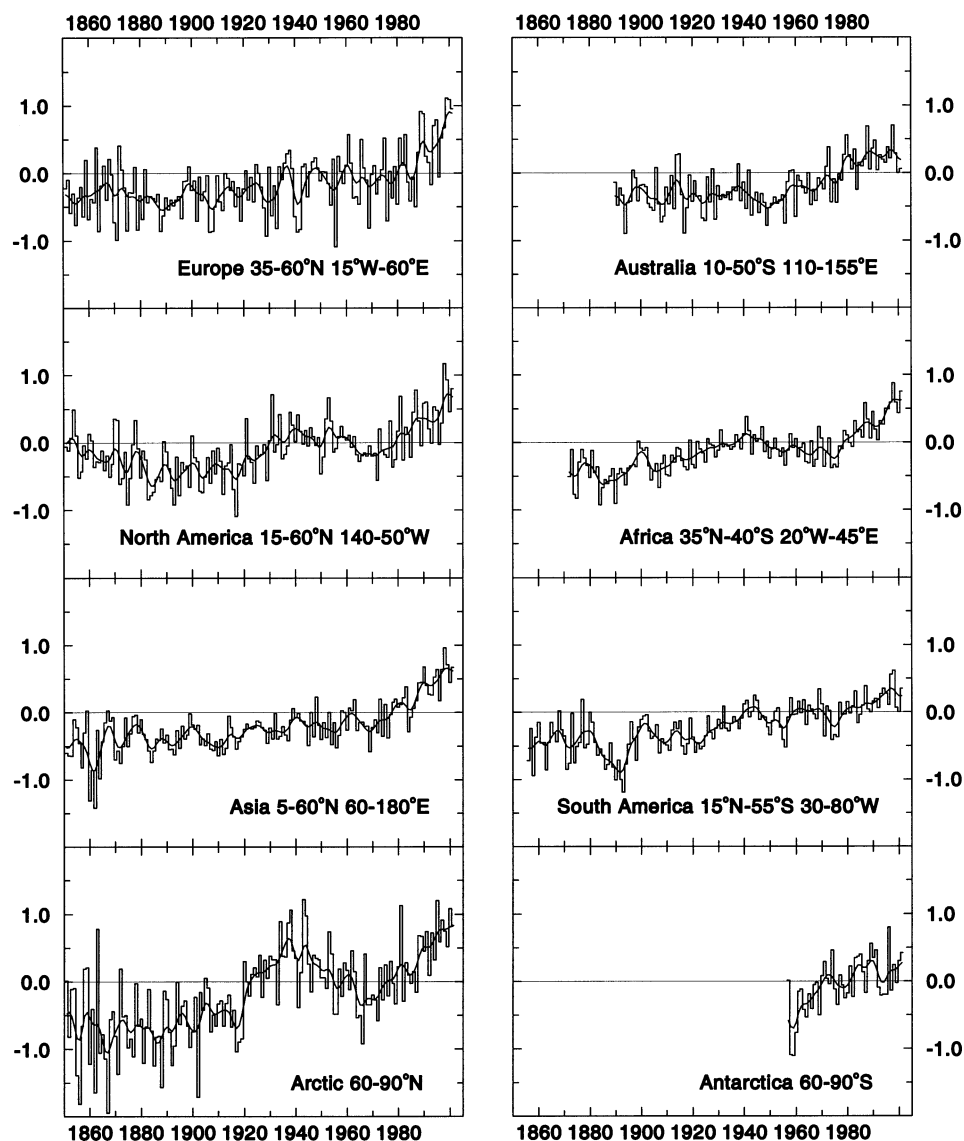


FIG. 2. Annual temperature anomaly estimates for eight continental areas, derived using CRUTEM2v. The areas are defined in each of the panels and in Table 2. NH areas are on the left of the figure with SH and tropical regions on the right. All series in this and subsequent plots are expressed as anomalies from 1961–90 with the smooth line being a 10-yr Gaussian filter, padded sufficiently at each end to extend the smooth line to the beginning/ending year of each series.

TABLE 2. Temperature change explained for each continent by the linear trend over several periods. Trends are expressed in $^{\circ}\text{C decade}^{-1}$ and in bold if significant at the 95% significance level. All start years are before 1861 except for Australia, Africa, and Antarctica, and their start years are given in parentheses.

Continent	1861–2000	1901–2000	1920–44	1945–76	1977–2001
Europe 35°–60°N, 15°W–60°E	0.049	0.076	0.006	–0.029	0.425
North America 15°–60°N, 140°–50°W	0.054	0.075	0.202	– 0.103	0.289
Asia 5°–60°N, 60°E–180°	0.057	0.082	0.047	0.020	0.283
Arctic 60°–90°N	0.098	0.079	0.274	– 0.156	0.364
Australia 10°–50°S, 110°–155°E (1890)	0.060	0.067	0.036	0.126	0.055
Africa 35°N–40°S, 20°W–45°E (1872)	0.060	0.059	0.136	– 0.073	0.284
South America 15°N–55°S, 30°–80°W	0.055	0.060	0.249	–0.006	0.125
Antarctica 60°–90°S (1957)	—	—	—	—	0.080

data within each region. The series are plotted in Fig. 2 on an annual basis (average of the 12 monthly values). In this and subsequent figures, the histogram shows values for individual years and the smooth lines highlight variations on decadal timescales using a 10-yr Gaussian filter. Table 2 gives trends in $^{\circ}\text{C decade}^{-1}$ for several periods, calculated using least squares regression. The periods were chosen based on the major trends in global and hemispheric temperature series and will be used consistently in all later tables.

All eight series are plotted on the same scale to illustrate differences in interannual variability, although this is partly confounded by the different sizes of the continents and from the gridbox sampling. Interannual variability is generally higher in polar, compared to tropical, regions. Each is plotted from the point in time enough grid boxes were available to define realistic values. Up to 40 of the earliest years were discarded for two of the three tropical/southern temperate continents, because the distribution of available stations, and hence grid boxes, is restricted to only parts of continents during these times. Antarctica is only plotted from 1957. Significant data exist for earlier years but coverage is restricted to a few boxes around the Antarctic Peninsula back to the mid-1940s and one grid box (which uses the station on the South Orkney Islands) that starts in 1903.

All seven century-long series show warming and the 1990s is the warmest decade of the record. The long-term warming generally occurs during the first part of the twentieth century and since the mid-1970s. The early warming (partly captured by 1920–44 in Table 2) is statistically significant (at the 95% level) in all areas except Europe, Asia, and Australia. For the recent period (1977–2001) all continents except Australia and the Antarctic show significant warming. North America, Africa, and especially the Arctic show significant cooling over the 1945–76 period. Australia, in contrast, shows significant warming during this period. The relative warmth of Australia in the pre-1890 period is potentially erroneous (see discussion in Folland et al. 2001) and these data have been ignored. They are probably related to the poorer exposure of the thermometers due to the use of nonstandard screens (see discussion in Collins and Della Marta 1999).

b. Hemispheric-scale averages

Cosine weighting is again used to develop seasonal [traditional climatological ones, December–January–February (DJF) is boreal winter/austral summer etc.] and annual temperature averages for the two hemispheres. Figure 3 shows the series for the NH (from 1851) and Fig. 4 for the SH (from 1858, the first year when each of the three temperate southern continents are represented by at least two grid boxes). Table 3 gives trends for the same periods as used earlier. The percent coverage of data by hemisphere can be seen in Fig. 1.

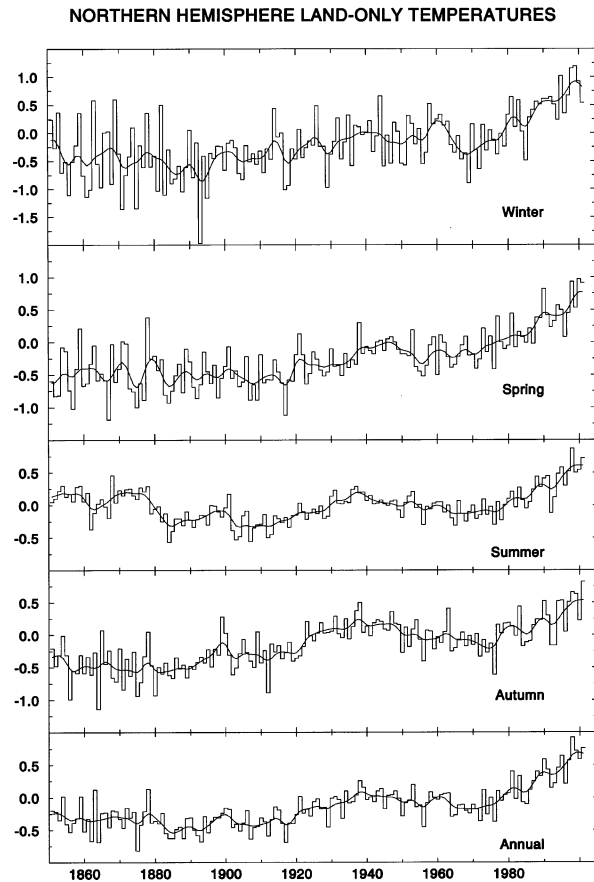


FIG. 3. Seasonal and annual time series (1851–2001) for the NH. Seasons are traditional climatological (DJF is winter, etc.). Further details (dataset used, base period, and smoothing) are given in Fig. 2.

For the NH, year-to-year variability is greatest during winter and least in summer. The slightly greater variability prior to 1880 in all seasons (except summer) is more likely to be due to sparser coverage than a real feature, despite the gridbox variance reductions incorporated within CRUTEM2v. Variance correction adjusts individual gridbox time series for changing station data contribution, but it is impossible to correct regional averages for reduced gridbox availability in a realistic way. All seasons show comparable century-scale warming during the twentieth century, but there are differences in timing between them, evidence that can also be seen in the trends for the periods given in Table 3. Warming is significant in all seasons and annually for 1861–2000, 1901–2000, and 1977–2001 and during summer [June–July–August (JJA)], autumn [September–October–November (SON)], and the year during 1920–44. Cooling is apparent in all seasons and annually for 1945–76 and is significant for summer and autumn.

The relative warmth of the summer season prior to 1880 may be artificial and due to the possible impacts of direct sunlight on the instruments, prior to the development of Stevenson-type screens, in higher northern

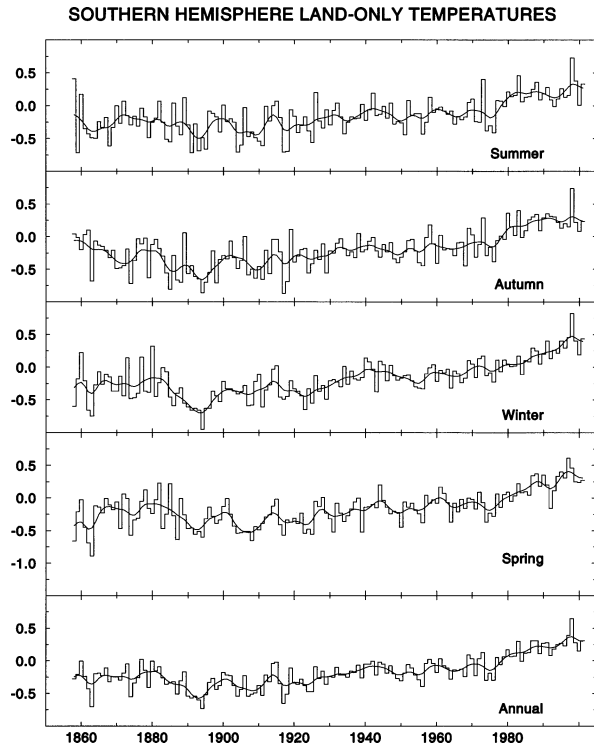


FIG. 4. Seasonal and annual time series (1858–2001) for the SH. Seasons are traditional climatological (DJF is summer, etc.). Further details (dataset used, base period, and smoothing) are given in Fig. 2.

latitudes. This possibility cannot yet be fully discounted, despite the more extensive homogeneity assessments over Fennoscandia and Canada, as these only fully extend back to 1890 and 1895, respectively (see section 2). If this seasonal contrast is real, then it implies a marked change in continentality (greater winter/summer temperature differences) over part of the NH prior to 1880. Further work is required, possibly using circulation data, to determine the validity of this feature.

The SH (Fig. 4) in contrast shows less seasonal contrasts in both trends and year-to-year variability than the NH, a result that might be expected given its greater ocean fraction. Warming is statistically significant for all seasons and annually for 1861–2000, 1901–2000, all seasons except summer (DJF) for 1920–44, and all seasons except summer and autumn (MAM) for 1977–2001. There is no evidence of cooling during the 1945–76 period. For all periods the winter (JJA) always warms by the greatest magnitude.

The relative warmth of the seasons prior to 1890 may be partly spurious. This period is relatively cool over South America and southern Africa (see Fig. 2), but the greatest areal contribution to the SH series comes from Australia and doubt has been cast on pre-1910 temperatures by several studies (see earlier discussion in section 4a). Much of our data for this continent for these years comes indirectly from GHCNv2, which has undergone extensive homogeneity assessment (Peterson

TABLE 3. Temperature change explained by the linear trend over several periods for both hemispheres for specified seasons and annually. Trends are expressed in $^{\circ}\text{C decade}^{-1}$ and in bold if significant at the 95% significance level.

Hemisphere/ season	1861–2000	1901–2000	1920–44	1945–76	1977–2001
NH/DJF	0.075	0.092	0.171	−0.011	0.416
NH/MAM	0.066	0.098	0.106	−0.019	0.310
NH/JJA	0.022	0.054	0.124	−0.058	0.289
NH/SON	0.058	0.048	0.158	−0.115	0.202
NH/annual	0.055	0.074	0.129	−0.055	0.298
SH/DJF	0.037	0.058	0.092	−0.001	0.070
SH/MAM	0.042	0.066	0.087	0.030	0.093
SH/JJA	0.046	0.069	0.189	0.032	0.199
SH/SON	0.036	0.068	0.129	0.014	0.160
SH/annual	0.040	0.065	0.130	0.020	0.129

and Vose 1997) but this assessment has only been undertaken using objective comparisons between neighbors. This procedure would be inadequate to detect a nationwide bias caused by a systematic change in the exposure of the instruments.

c. Comparison of annual hemispheric series with the results of other analyses

In this section the two hemispheric averages are compared with three other analyses: the earlier analysis of Jones and the series developed by the NCDC (Peterson et al. 1998b) and GISS (Hansen et al. 1999, 2001). Both our present and the earlier Jones study use a base period of 1961–90. NCDC currently use 1880–2000 and GISS 1951–80 for their published series. For direct comparison we have adjusted both series to our 1961–90 base. Figure 5 shows hemispheric and global series, plotting annual values for the present study and decadal smoothed series for all four analyses. Trends for all four series (calculated using annual values) for the selected time periods are given in Table 4. The global land average for 1901–2000 warms by $0.07^{\circ}\text{C decade}^{-1}$.

The global series for the present study and for Jones is calculated by averaging those of the two hemispheres, while for NCDC the global average is calculated as a single domain. GISS (Hansen et al. 2001) has recently changed to a zonal-averaging approach, which ensures the global mean is the average of the two hemispheres. Estimates by NCDC are likely to be biased to the NH value due to its greater percent coverage. We contend that averaging the hemispheres is the more logical choice when calculating the global average (see also the discussion in Wigley et al. 1997). The SH has the greater percentage of missing land areas, principally because Antarctic data are unavailable until the late 1950s, and this region should be more reliably inferred from other SH land areas. The implicit assumption made by all calculations of large-area averages is that the temperature average of the unavailable areas is equal to that for the areas with data. Whether Antarctic temperatures can be estimated (or assumed) from any other part of

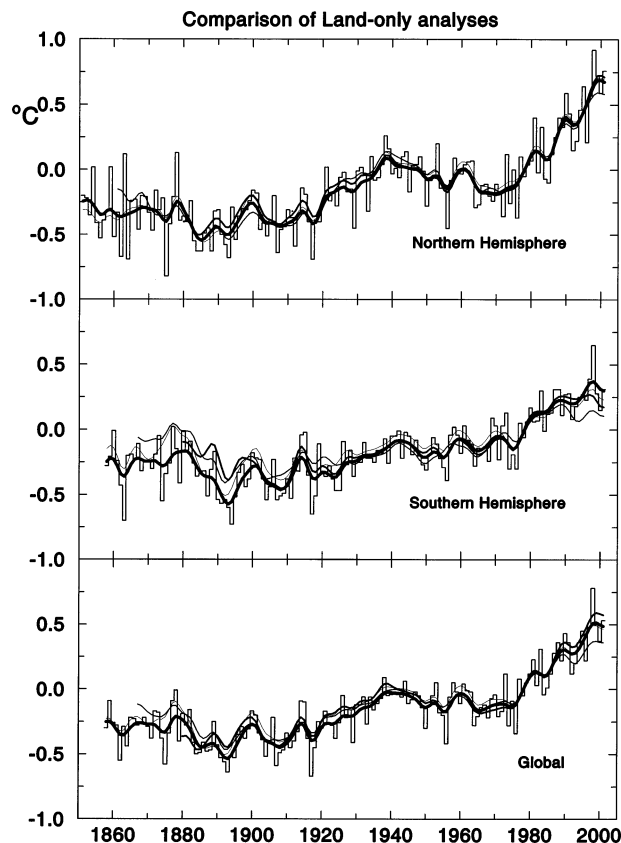


FIG. 5. Comparison of annual hemispheric and global temperature averages for four different datasets (this study CRUTEM2v: Jones 1994; NCDC: Peterson et al. 1998b; and GISS: Hansen et al. 1999, 2001). The histograms show annual values for CRUTEM2v. The four smooth lines show 10-yr Gaussian filtered series for each dataset. The thickest line is the present study, Jones (1994) the thinnest with NCDC being slightly thicker than GISS. The first years of each dataset are given in Table 4.

the world is doubtful and highly debatable (see discussion in Jones 1990 and Raper et al. 1984). Intuitively the use of the rest of the SH for this purpose seems better than incorporating influences from the NH. Differences between the two assumptions are in the range $\pm 0.05^{\circ}\text{C}$ before 1920 and $\pm 0.03^{\circ}\text{C}$ since then (Wigley et al. 1997). Such small differences can affect ranking, particularly with respect to some of the recent record-breaking years (see also the discussion in section 5).

In Fig. 5, the greater differences between analyses occur over the SH, particularly at the beginnings and the ends of the records. The principal difference between the present study and the earlier analysis (Jones) is that the warming of the SH is greater in the present study (e.g., 0.65°C warming, cf. 0.53°C for the 1901–2000 period). For the other two datasets (NCDC and GISS) warming over the SH land areas is much greater for NCDC than GISS (e.g., 0.61°C , cf. 0.48°C over 1901–2000) and almost comparable with the present study. The GISS pattern of warming and cooling over the SH is of markedly lower amplitude, particularly recently.

TABLE 4. Temperature change explained by the linear trend over several periods for four different analyses of hemispheric and global temperature. Trends are expressed in $^{\circ}\text{C}$ decade $^{-1}$ and in bold if significant at the 95% significant level. The first year of each series is given in parentheses, when it is after 1861.

Series (first year)	1861– 2000	1901– 2000	1920–44	1945–76	1977– 2001
NH (this study)	0.055	0.074	0.129	–0.055	0.298
NH (Jones)	0.057	0.073	0.120	–0.050	0.292
NH (NCDC-1880)	—	0.068	0.109	–0.057	0.316
NH (GISS)	0.052	0.070	0.154	–0.073	0.250
SH (this study)	0.040	0.065	0.130	0.020	0.129
SH (Jones)	0.032	0.053	0.110	–0.000	0.097
SH (NCDC-1880)	—	0.061	0.108	0.001	0.068
SH (GISS-1867)	—	0.048	0.067	0.044	0.012
GL (this study)	0.048	0.070	0.128	–0.017	0.206
GL (Jones)	0.045	0.063	0.115	–0.025	0.195
GL (NCDC-1880)	—	0.067	0.111	–0.039	0.253
GL (GISS-1867)	—	0.059	0.109	–0.015	0.137

This is possibly related to the gridbox size and the greater influence given to island temperatures (especially in the South Pacific) due to the large equal-area boxes used. During the nineteenth century, the present study gives the coolest temperatures of all four series.

For the NH, the pattern of warming is remarkably similar between all four analyses. For the globe, the different methods used to produce global averages might explain the slight differences in trends, although this is unlikely (see Wigley et al. 1997). As with the SH, there are greater divergences of the series at the beginning and ends of the records. The GISS global series is slightly warmer during the nineteenth century compared to the other three and tends to cooler temperatures since the mid-1980s. This latter tendency is clearly coming from the SH and is barely noticeable in the NH. The NCDC global series begins to diverge from the two CRU and the GISS series in the mid-1980s and is about 0.1°C warmer by 2001 (also evident in Folland and Karl 2001, their Fig. 2.1b, p. 107). No positive differences for the two hemispheres are apparent and the divergence is difficult to reconcile, especially as the NCDC SH series is slightly cooler than the present study after 1995.

5. Analyses of the combined land and marine dataset

Combination of land and marine datasets produces much higher percentage coverages. In this analysis we use HadCRUT2v where the variance-adjusted version of the present land analysis (CRUTEM2v) is combined with a marine (SST) analysis (HadISSTv) using the method detailed in Jones et al. (2001). Only in situ SST data are used. Coverage percentages reach 96% for the NH (in the 1970s) and 80% for the SH (in summers during the 1970s and 1980s). Coverage varies little from season to season over the NH, but for the SH it can drop by up to 15% during the winters of the 1970s and

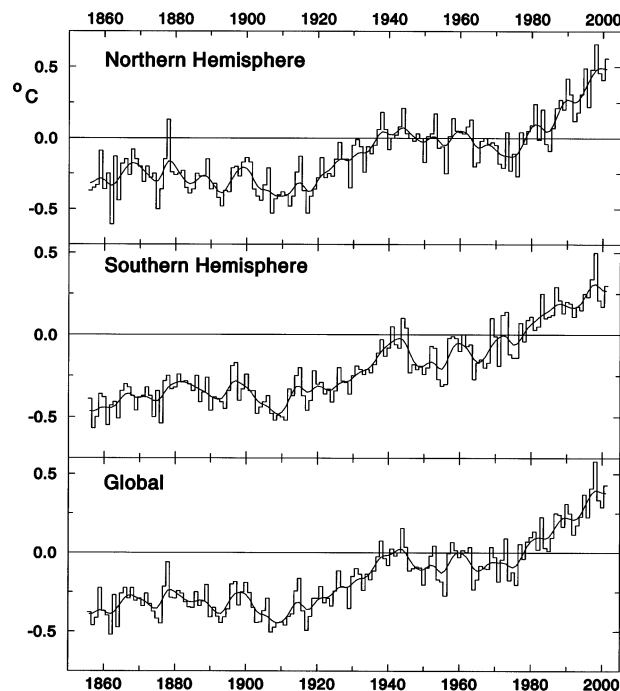


FIG. 6. Annual hemispheric and global temperature averages using the combined land and marine dataset (HadCRUT2v), 1856–2001. HadCRUT2v is variance corrected for the years 1870–2001 and HadCRUT2 for 1856–69.

1980s. The drop in coverage occurs due to lack of reports from the Southern Oceans during winters.

a. Hemispheric- and global-scale averages

Figure 6 shows annual hemispheric and global averages calculated using HadCRUT2v. The trends of temperature over the selected periods are given in Table 5. These series have been extensively discussed elsewhere (e.g., Wigley et al. 1997; Jones et al. 1999; Folland and Karl 2001) and the results using the improved land component are not significantly different from HadCRUTv.

Both hemispheric series are significantly correlated to indices of the El Niño–Southern Oscillation (ENSO) phenomenon (Jones 1989; Wigley 2000) and the NH winter series is positively correlated to the North Atlantic Oscillation (NAO; Hurrell 1996). Discussion has also highlighted the sometimes variable nature of these links (e.g., for ENSO in Wigley and Santer 2002, manuscript submitted to *J. Climate*, and for the NAO in Jones et al. 2002), with both appearing weaker during the mid-twentieth century than before or since. The above studies indicate that the correlations are strongest at higher frequencies and do not explain a significant fraction of the twentieth century annual timescale trends.

The most recent year (2001) was the second warmest year (0.43°C above the 1961–90 average) but still 0.15°C cooler than 1998. The warmth of the exceptional

TABLE 5. Temperature change explained by the linear trend over several periods for the combined land and marine hemispheric and global dataset (HadCRUT2v). Trends are expressed in $^{\circ}\text{C decade}^{-1}$ and in bold if significant at the 95% significance level.

Series	1861–2000	1901–2000	1920–44	1945–76	1977–2001
NH	0.042	0.065	0.152	−0.051	0.223
SH	0.045	0.067	0.157	0.041	0.106
Global	0.044	0.066	0.154	−0.005	0.165

year 1998 was clearly related to the record-breaking 1997/98 El Niño event. Estimates of the influence from the Southern Oscillation Index (SOI) would suggest that $0.15^{\circ}\text{--}0.2^{\circ}\text{C}$ of the warmth was due to this cause. The rate of recent global temperature rise (see Tables 4 and 5) means that the next El Niño event is likely to lead to a new record, even if the event were only of moderately strong intensity. The globe (the underlying value minus the ENSO influence) should be about 0.08°C warmer in 2003 than in 1998, by simply projecting the trend of the 1977–2001 period from Table 5. At the time of writing (June 2002), the value for 2002 is on course to be the second warmest year.

b. Patterns of warming

So far, almost all the discussion has considered trends in continental- and hemispheric-scale temperature averages. Here, we show the spatial patterns of trends, contrasting seasonal and annual trends for two 25-yr periods when the world has warmed by significant amounts (Fig. 7 for 1920–44 and Fig. 8 for 1977–2001). The two figures additionally illustrate the improvements in coverage in recent years, with near-complete coverage from 70°N to 50°S (apart from some parts of central Africa). For the earlier period less of the land areas have data and the shipping routes (from where the SST data originate) can be clearly seen over the Pacific and the Southern Oceans.

For both periods, hemispheric and global averages indicate highly significant warming (Table 5). Despite this, significant temperature trends are present in only 10%–20% of the available grid boxes. Warming does not occur everywhere, nor in every season, but many more grid boxes indicate significant warming than cooling. Cooling is evident for some regions during 1920–44 (Fig. 7) and is significant for a few grid boxes during some seasons. For the recent period, 1977–2001, warming clearly dominates and there is a tendency for it to be greater over land areas, although only in some seasons. Warming is of greater magnitude in the DJF season, particularly over parts of North America and Eurasia. Some regions (e.g., Siberia) show marked seasonal patterns: strong warming everywhere in DJF, warming in the east and cooling in the west in MAM, opposite trends to MAM in JJA, and slight cooling everywhere in SON. The greater magnitude of winter trends in this region dominates the annual map.

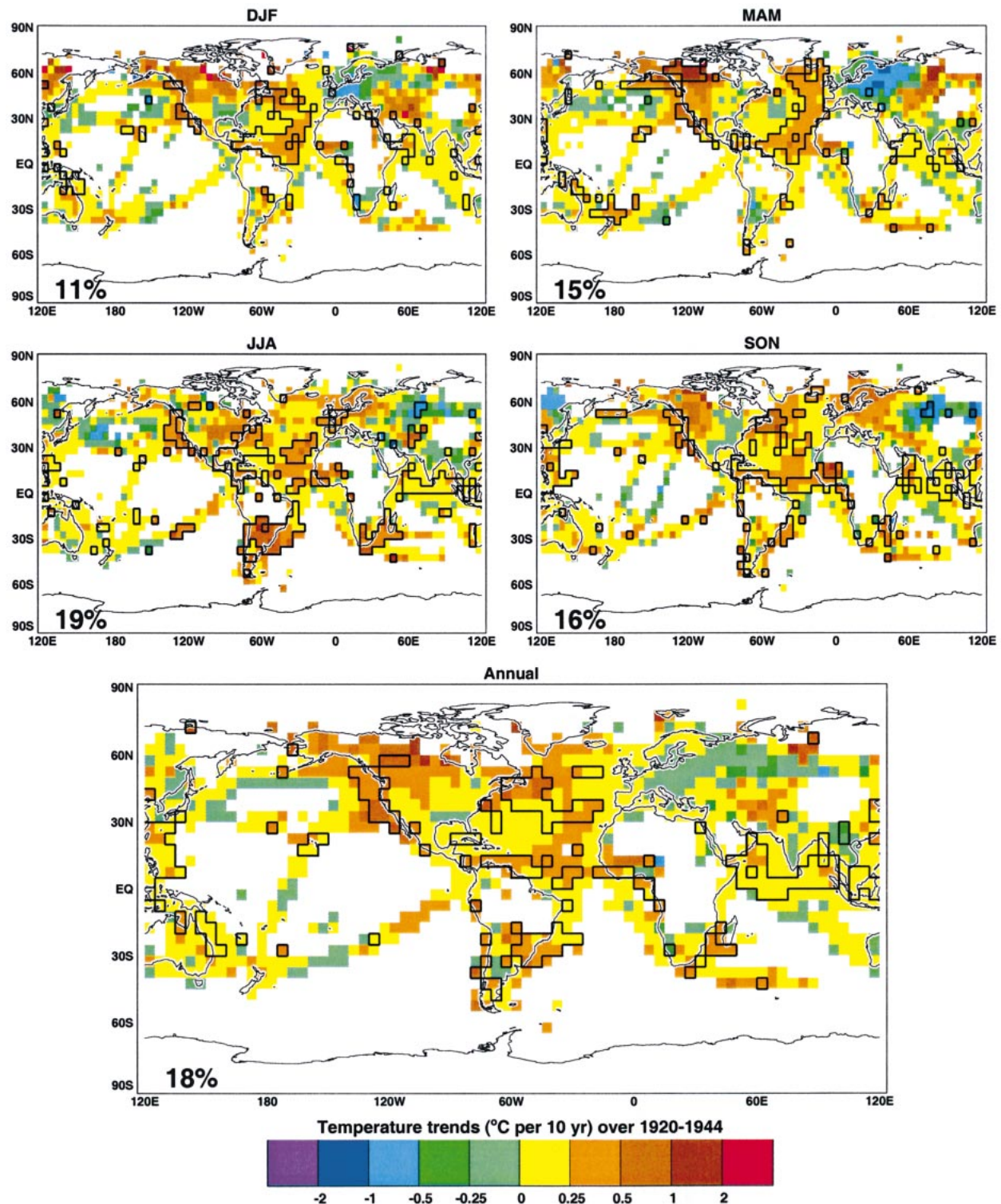


FIG. 7. Spatial patterns of seasonal and annual trends of surface temperature (using HadCRUT2v) for the 1920–44 period. Values are plotted in $^{\circ}\text{C decade}^{-1}$. The thicker line around some grid boxes/areas denote those regions where the temperature trend is significant at the 95% significance level. At least 15 yr within the period is required to calculate a trend. The number in the bottom left-hand corner of each plot is the percentage of grid boxes with significant trends.

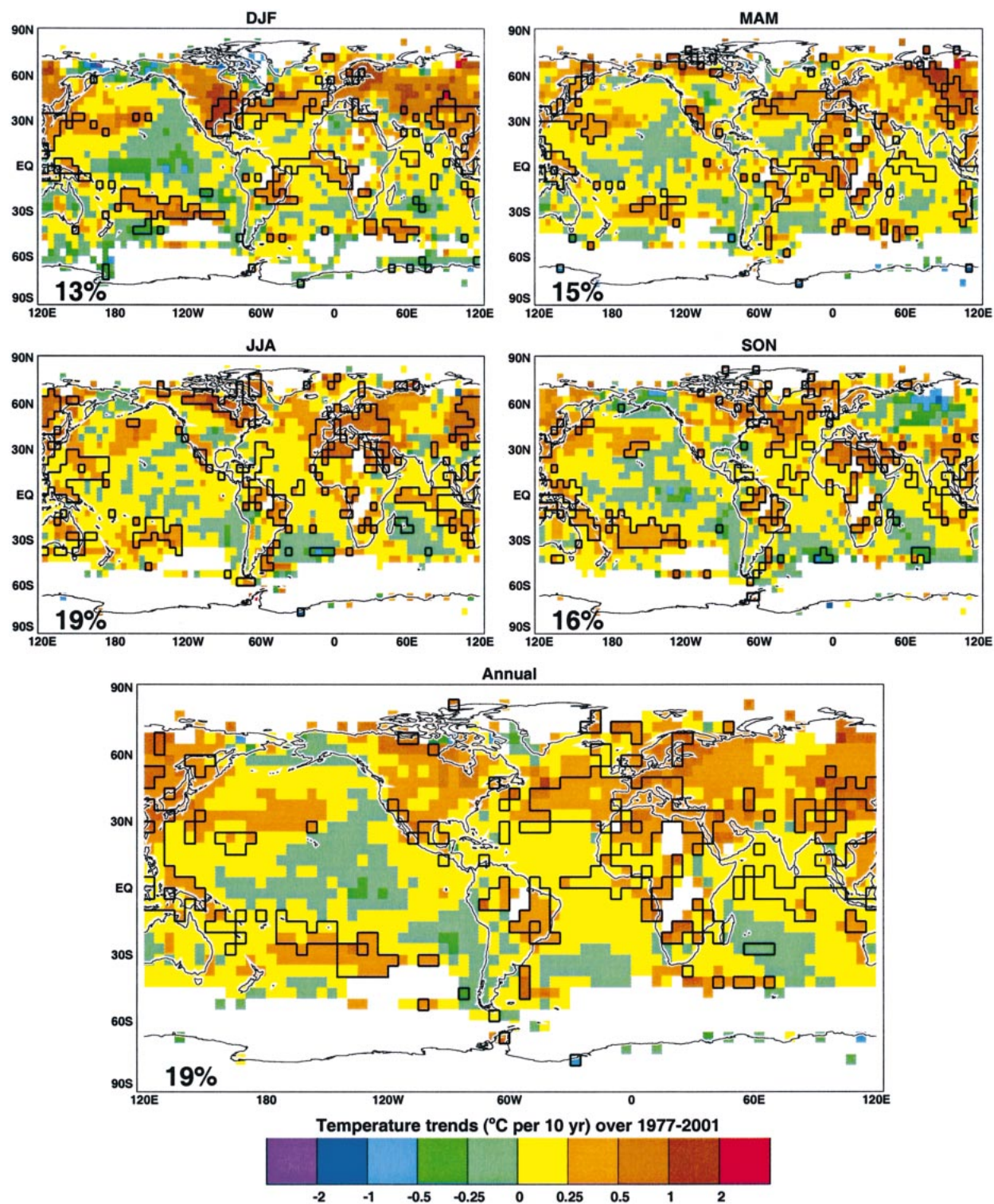


FIG. 8. Spatial patterns of seasonal and annual trends of surface temperature (using HadCRUT2v) for the 1977–2001 period. Values are plotted in $^{\circ}\text{C decade}^{-1}$. The thicker line around some grid boxes/areas denote those regions where the temperature trend is significant at the 95% significance level. At least 15 yr within the period is required to calculate a trend. The number in the bottom left-hand corner of each plot is the percentage of grid boxes with significant trends.

6. Conclusions

In this paper we have detailed the improvements to the gridded temperature datasets available from the Climatic Research Unit. The improvements to the quality of the basic gridbox dataset have been made possible by the work of many scientists and institutions around the world, through more extensive and comprehensive homogeneity analyses of the individual station temperature series. Merging these diverse datasets with our original database has not been a trivial task. In many cases decisions had to be made on a station-by-station basis to determine which source had the best-quality station data. The task was also very labor intensive. For many stations several sources contain similar but rarely exactly the same basic data. Many of the difficulties arise because national datasets, developed by both NMSs and individuals, are maintained to different standards than data routinely exchanged between countries.

Despite our best efforts to improve the quality and quantity of data used for large-scale temperature estimation, it is likely that our new database still contains some problematic data, but its influence should have been significantly reduced. The new database contains 5159 station records of which 4167 could be used in gridding because they contain enough data for the 1961–90 reference period to calculate or estimate the necessary normals. Temperature data for these 4167 stations were used to develop $5^\circ \times 5^\circ$ gridbox anomalies for the period 1851 to 2001 for the world's land areas. Coverage is limited during the earliest years, reaching a maximum during the 1951–90 period. A reduction in availability occurs during recent years, but this only reduces coverage of the land areas that can be gridded by a few percentage points.

Compared to earlier analyses by CRU (Jones 1994), the improvements to the percent of the land area with data are marginal, but the quality of the gridbox values has been improved both by using the more homogeneous data and by greater station density within individual grid boxes. The trend of temperature change during the twentieth century does not significantly change for the NH, but the warming in the new analysis is increased for the SH.

The hemispheric average series, developed from land-only data, were also compared with the results from two other centers (NCDC: Peterson et al. 1998b and GISS: Hansen et al. 1999, 2001). As reported elsewhere (Peterson et al. 1998b; Jones et al. 1999; Folland and Karl 2001), the trends for the NH agree very well. Larger differences occur in the SH, with our new analysis indicating greater twentieth-century warming than our earlier analysis (Jones 1994) and those of NCDC and GISS. For the global average the NCDC analysis warms relative to this study, Jones and GISS by about 0.1°C since the mid-1980s.

Finally, we have merged the new land dataset with marine data from the Hadley Centre of the Met Office,

for both variance-adjusted and original versions (HadCRUT2v and HadCRUT2). All the datasets (both the gridded and the station series) have been made available through the Climatic Research Unit Web site (<http://www.cru.uea.ac.uk>).

Acknowledgments. This paper would not have been possible without many individuals and institutions making their new homogenized data freely available for us to use. In this respect, we thank Hans Alexandersson (SMHI), Per Øyvind Nordli (DNMI), Heikki Tuomenvirta and Achim Drebs (FMI) for NACD/NORDKLIM; Lucie Vincent (AES) for CHTD; Reinhard Böhm and Ingeborg Auer (ZAMG) for ALOCLIM; Neil Plummer, Michael Shevchenko, and Blair Trewin (NCC/Bureau of Meteorology) for Australian data; Tom Peterson (NCDC) for GHCNv2; G. Muller-Westermeier (DWD) for German data; Aryan van Engelen (KNMI) for Dutch data; Joanna Wibig for Polish data; and Gareth Marshall (BAS) for Antarctic data. All also speedily answered many questions about data quality and related issues that arose during the merging process. Jim Hansen and Reto Ruedy supplied the latest GISS hemispheric averages, while Dave Easterling and Jay Lawrimore performed the same task for the NCDC averages. We also acknowledge the comments of the anonymous reviewers. David Parker, Peter Thorne, and Chris Folland (all working at the Hadley Centre) provided comments on an earlier version of this paper, and the Hadley Centre provided the gridded sea surface temperature dataset used in section 5. This work has been supported by the U.S. Dept. of Energy under Grant DE-FG02-98ER62601.

REFERENCES

- Auer, A., R. Böhm, and W. Schöner, 2001: Austrian long-term climate 1767–2000: Multiple instrumental climate time series from Central Europe. Publ. 397, Österreichische Beiträge zu Meteorologie und Geophysik Heft 25, Zentralanstalt für Meteorologie und Geodynamik, Vienna, Austria, 147 pp.
- Böhm, R., I. Auer, M. Brunetti, M. Maugeri, T. Nanni, and W. Schöner, 2001: Regional temperature variability in the European Alps: 1760–1998 from homogenized instrumental time series. *Int. J. Climatol.*, **21**, 1779–1801.
- Bureau of Meteorology, 2000: *Climate Data: Australia, Version 1.01d*. CD-ROM, Climate Services, Bureau of Meteorology, Kent Town, Australia.
- Camuffo, D., and P. D. Jones, Eds., 2002: Improved understanding of past climatic variability from early daily European instrumental sources. *Climatic Change*, **53** (1–3), 1–392.
- Collins, D. A., and P. M. Della Marta, 1999: Annual climate summary 1998: Australia's warmest year on record. *Aust. Meteor. Mag.*, **48**, 273–283.
- Folland, C. K., and T. R. Karl, 2001: Observed climate variability and change. *Climate Change 2001: The Scientific Basis*, J. T. Houghton et al. Eds., Cambridge University Press, 99–181.
- , and Coauthors, 2001: Global temperature change and its uncertainties since 1861. *Geophys. Res. Lett.*, **28**, 2621–2624.
- Frich, P., and Coauthors, 1996: North Atlantic Climatological Dataset (NACD Version 1). Scientific Rep. 96-1, Danish Meteorological Institute, Copenhagen, Denmark, 47 pp. plus appendixes and data files.

- Hansen, J. E., and S. Lebedeff, 1987: Global trends of measured surface air temperature. *J. Geophys. Res.*, **92**, 13 345–13 372.
- , R. Ruedy, J. Glascoe, and M. Sato, 1999: GISS analysis of surface temperature change. *J. Geophys. Res.*, **104**, 30 997–31 022.
- , R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, and T. Karl, 2001: A closer look at United States and global surface temperature change. *J. Geophys. Res.*, **106**, 23 947–23 963.
- Hurrell, J. W., 1996: Influence of variations in extratropical winter-time teleconnections on Northern Hemisphere temperature. *Geophys. Res. Lett.*, **23**, 665–668.
- Jones, P. D., 1989: The influence of ENSO on global temperatures. *Climate Monit.*, **17**, 80–89.
- , 1990: Antarctic temperatures over the present century—A study of the early expedition record. *J. Climate*, **3**, 1193–1203.
- , 1994: Hemispheric surface air temperature variations: A reanalysis and an update to 1993. *J. Climate*, **7**, 1794–1802.
- , and P. A. Reid, cited 2001: Antarctic surface data 2000. CDIAC Numeric Data Package NDP 032. [Available online at <http://cdiac.esd.ornl.gov/epubs/ndp/ndp032/ndp032.html>.]
- , and Coauthors, 1985: A grid point surface air temperature data set for the Northern Hemisphere. U.S. Dept. of Energy, Carbon Dioxide Research Division, Tech. Rep. TR022, 251 pp.
- , S. C. B. Raper, B. S. G. Cherry, C. M. Goodess, and T. M. L. Wigley, 1986: A grid point surface air temperature data set for the Southern Hemisphere. U.S. Dept. of Energy, Carbon Dioxide Research Division, Tech. Rep. TR027, 73 pp.
- , T. J. Osborn, and K. R. Briffa, 1997: Estimating sampling errors in large-scale temperature averages. *J. Climate*, **10**, 2548–2568.
- , M. New, D. E. Parker, S. Martin, and I. G. Rigor, 1999: Surface air temperature and its changes over the past 150 years. *Rev. Geophys.*, **37**, 173–199.
- , T. J. Osborn, K. R. Briffa, C. K. Folland, B. Horton, L. V. Alexander, D. E. Parker, and N. Rayner, 2001: Accounting for sampling density in grid-box land and ocean surface temperature time series. *J. Geophys. Res.*, **106**, 3371–3380.
- , —, and —, 2002: Pressure-based measures of the NAO: A comparison and an assessment of changes in the strength of the NAO and in its influence on surface climate parameters. *The North Atlantic Oscillation*, J. W. Hurrell et al., Eds., Amer. Geophys. Union., in press.
- Moberg, A., and H. Alexandersson, 1997: Homogenization of Swedish temperature data. Part II: Homogenized gridded air temperatures compared with a subset of global air temperatures since 1861. *Int. J. Climatol.*, **17**, 35–54.
- , and Coauthors, 2000: Day-to-day temperature variability trends in 160- to 275-year-long European instrumental records. *J. Geophys. Res.*, **105**, 22 849–22 868.
- New, M., M. Hulme, and P. D. Jones, 1999: Representing twentieth century space–time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *J. Climate*, **12**, 829–856.
- , —, and —, 2000: Representing twentieth century space–time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *J. Climate*, **13**, 2217–2238.
- Peterson, T. C., and R. S. Vose, 1997: An overview of the Global Historical Climatology Network Temperature Database. *Bull. Amer. Meteor. Soc.*, **78**, 2837–2848.
- , H. Daan, and P. D. Jones, 1997: Initial selection of a GCOS surface network. *Bull. Amer. Meteor. Soc.*, **78**, 2145–2152.
- , and Coauthors, 1998a: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.*, **18**, 1493–1517.
- , T. R. Karl, P. F. Jamason, R. Knight, and D. R. Easterling, 1998b: The first difference method: Maximizing station density for the calculation of long-term global temperature change. *J. Geophys. Res.*, **103**, 25 967–25 974.
- Raper, S. C. B., T. M. L. Wigley, P. R. Mayes, P. D. Jones, and M. J. Salinger, 1984: Variations in surface air temperature. Part 3: The Antarctic, 1957–82. *Mon. Wea. Rev.*, **112**, 1341–1353.
- Tuomenvirta, H., A. Drebs, E. Førland, O. E. Tveito, H. Alexandersson, E. Vaarby Laursen, and T. Jonsson, 2001: Nordklim data set 1.0—Description and illustrations. KLIMA Rep. 08/01, Norwegian Meteorological Institute, Oslo, Norway, 26 pp. [CD-ROM available from A. Drebs, Finnish Meteorological Institute, P.O. Box 503, FIN-00101 Helsinki, Finland.]
- van Engelen, A. F. V., and J. W. Nellestijn, 1996: Monthly, seasonal and annual means of the air temperature in tenths of centigrades in De Bilt, Netherlands, 1706–1995. KNMI Report, Climatological Services Division, 6 pp.
- Vincent, L. A., 1998: A technique for the identification of inhomogeneities in Canadian temperature series. *J. Climate*, **11**, 1094–1104.
- , and D. W. Gullet, 1999: Canadian historical and homogeneous temperature datasets for climate change analyses. *Int. J. Climatol.*, **19**, 1375–1388.
- Vose, R. S., R. L. Schmoyer, P. M. Steurer, T. C. Peterson, R. Heim, T. R. Karl, and J. Eischeid, 1992: The Global Historical Climatology Network: Long-term monthly temperature, precipitation, sea level pressure, and station pressure data. Tech. Rep. ORNL/CDIAC-53, NDP-041, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN, 325 pp.
- Wigley, T. M. L., 2000: ENSO, volcanoes and record-breaking temperatures. *Geophys. Res. Lett.*, **27**, 4101–4104.
- , P. D. Jones, and S. C. B. Raper, 1997: The observed global warming record: What does it tell us? *Proc. Natl. Acad. Sci. USA*, **94**, 8314–8320.
- WMO, 1996: Climatological Normals (CLINO) for the period 1961–1990. World Meteorological Organization Doc. WMO/OMM-No. 847, Geneva, Switzerland, 768 pp.