Representing Twentieth-Century Space–Time Climate Variability. Part I: Development of a 1961–90 Mean Monthly Terrestrial Climatology

MARK NEW, MIKE HULME, AND PHIL JONES

Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom

(Manuscript received 22 December 1997, in final form 3 April 1998)

ABSTRACT

The construction of a 0.5° lat × 0.5° long surface climatology of global land areas, excluding Antarctica, is described. The climatology represents the period 1961–90 and comprises a suite of nine variables: precipitation, wet-day frequency, mean temperature, diurnal temperature range, vapor pressure, sunshine, cloud cover, ground frost frequency, and wind speed. The climate surfaces have been constructed from a new dataset of station 1961–90 climatological normals, numbering between 19 800 (precipitation) and 3615 (wind speed). The station data were interpolated as a function of latitude, longitude, and elevation using thin-plate splines. The accuracy of the interpolations are assessed using cross validation and by comparison with other climatologies.

This new climatology represents an advance over earlier published global terrestrial climatologies in that it is strictly constrained to the period 1961–90, describes an extended suite of surface climate variables, explicitly incorporates elevation as a predictor variable, and contains an evaluation of regional errors associated with this and other commonly used climatologies. The climatology is already being used by researchers in the areas of ecosystem modelling, climate model evaluation, and climate change impact assessment.

The data are available from the Climatic Research Unit and images of all the monthly fields can be accessed via the World Wide Web.

1. Introduction

Concern about anthropogenic climate change has stimulated much research into the likely response of the ocean–atmosphere system to greenhouse gas forcing and the impacts of resultant climate change on the earth surface environment (IPCC 1996). Accurate representation of the mean state and variability of the present climate is important for a number of purposes in global change research. These include monitoring and detection of climate change (Jones 1994); evaluation of General Circulation Models (GCMs) (Hulme 1994a; Airey et al. 1996) and regional climate simulations (Christensen et al. 1997); ground truthing, calibration, or merging with satellite climatologies (Huffman et al. 1995); understanding the role of climate in biogeochemical cycling (Sellers et al. 1997; Cao and Woodward 1998); and construction of climate change scenarios (Carter et al. 1994).

These applications have historically had differing but converging priorities in terms of spatio-temporal resolution and accuracy. Coarse resolution datasets such as those of Jones (1994) for temperature and Hulme (1992a, 1994a) for precipitation have been adequate for monitoring and detection of climate change and GCM evaluation, where capturing temporal variability is as important as the representation of spatial detail. The development of higher resolution GCMs and regional climate models has, however, meant that the spatial resolution required for model evaluation has increased, and is likely to continue to do so. Biophysical modelers and the climate impacts community have favored finer spatial resolution (commonly 0.5° lat × 0.5° long) and a wider range of surface climate variables. Consequently they have tended to use long-term mean climatologies such as that of Leemans and Cramer (1991; W. Cramer 1997, personal communication). Yet, as these climate studies impact progress from equilibrium to transient simulations, researchers are increasingly requiring data that accurately represent climate variability continuously in space *and* time (Cramer and Fischer 1996; Piper and Stewart 1996). Similarly, there has been a growing demand for the representation of temporal variability in climate change scenarios (Hulme and Brown 1998).

Although these data requirements are converging toward high-resolution representation of climate in both space and time, there are few datasets that satisfy this demand. Notable exceptions are the monthly time step Global Precipitation Climatology Project (GPCP) dataset (Xie and Arkin 1996; Xie et al. 1996), the monthly

Corresponding author address: Dr. Mike Hulme, Climate Research Unit, School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom.
E-mail: m.hulme@uea.ac.uk

TABLE 1. List of the number of normals obtained from each source.

	NMA	WMO	CRU	CIAT[a]	Müller[b]	FAO[c]	Other	Total
Precipitation	8916	331	5878	3162	52	132	824	19 295
Wet days > 0.1 mm	6038	236	132	1174	91	0	700	8371
Wet days > 1.0 mm	3055	613	140	0	5	0	531	4344
Mean temperature	7210	612	2006	1599	56	155	454	12 092
Diurnal temperature range	7059	467	1464	940	67	189	541	10 727
Vapor pressure	2446	516	134	287	0	209	295	3887
Relative humidity	2593	449	97	1011	125	0	67	4342
Sunshine	2536	339	201	573	57	277	57	4040
Cloud cover	1618	426	167	140	0	0	327	2678
Ground frost days	1108	24	40	0	0	0	0	1172
Air frost days	3698	205	111	0	0	0	786	4800
Wind speed	2297	230	194	169	84	332	309	3615

[a] Centro Internacional de Agricultura Tropical (1997); [b] Müller (1982); [c] Food and Agriculture Organisation (1984).

1900–88, 2.5° lat × 2.5° long precipitation dataset of Dai et al. (1997, hereafter Dai), and the 0.5° lat × 0.5° long daily time step dataset being developed by Piper and Stewart (1996, henceforth PS). However these products either cover relatively short periods (1970s to present; GPCP, PS), are limited to precipitation (GPCP, PS, Dai) and minimum and maximum temperature (PS), do not include an elevation dependence in their interpolation schemes, or are coarse resolution (Dai). A further limitation is that GPCP and PS interpolate directly from station time series and are therefore restricted in the number of stations they can use.

An alternative approach to direct interpolation of station time series is to separate the time and space components by first constructing a high-resolution (0.5° × 0.5°) mean climatology for the climate variables of interest, subsequently deriving gridded monthly anomalies relative to the period for which the mean climate is defined. The anomaly and mean fields are then combined to arrive at gridded time series of each variable. The advantage of this approach is that the number of archived *and* easily obtainable station normals is far greater than that of station time series. This is particularly so in the case of variables other than precipitation and temperature, and for all variables as one goes back in time. Using as many stations as possible to generate the mean fields, together with an explicit treatment of elevation dependency, maximizes the representation of spatial variability in mean climate. Monthly anomalies, on the other hand, tend to be more a function of large-scale circulation patterns and relatively independent of physiographic control. Therefore, a comparatively less extensive network is sufficient to describe the month-to-month departures from the mean climate.

In this, the first of two papers, we describe the construction of a new 0.5° × 0.5° mean monthly terrestrial climatology, strictly constrained to the period 1961–90, for a suite of climate variables: precipitation, wet-day frequency, mean temperature, diurnal temperature range, vapor pressure, cloud cover, sunshine duration, ground frost frequency, and wind speed. In a companion paper, we describe the construction of monthly anomaly grids and their merging with this baseline climatology to produce 1901–present-day monthly time step climate fields (New et al. 1999). The present paper is structured as follows. In section 2 the new Climatic Research Unit (CRU) dataset of 1961–90 climate normals is described. Section 3 describes the interpolation methodology and construction of the mean climatology fields. In section 4, interpolation errors are assessed, and our new climatology is compared with some existing climatologies. Finally, section 5 contains some concluding remarks.

2. Observational dataset

a. Sources

The 1961–90 station normals used to construct the climatology were collated from a number of sources. Although the World Meteorological Organisation (WMO) 1961–90 global standard normals were released in May 1997 through the National Climate Data Center (NCDC), we began our own data collection in 1994 through direct contact with national meteorological agencies (NMAs), personal contacts, and other published sources. Initially, this endeavor was confined to European and neighboring countries (Hulme 1994b; Hulme et al. 1995), but was subsequently extended to obtain global coverage, excluding Antarctica. In many countries we have gained access to more stations and variables than were made available to NCDC. The NCDC dataset has, however, complimented ours in countries where we were unable to obtain data. The main data sources are described below and summarized in Table 1.

In some cases, published sources did not provide information on the normal period (e.g., CIAT, see below) or the normals represented a period other than 1961–90 (e.g., Müller 1982; FAO 1984). These stations were used in areas where no other data were available as it has been shown that, in the absence of marked interdecadal variability, the improvement in interpolation accuracy gained by including additional stations outweighs any penalty associated with relaxing temporal

fidelity (Willmott et al. 1996; Hulme and New 1997). These normals were assigned a low weight during the interpolation so that they did not unduly bias the true 1961–90 normals (see section 3).

1) NATIONAL METEOROLOGICAL AGENCIES

Data supplied by NMAs in direct response to our request for 1961–90 normals comprise by far the largest single data source, ranging between 75% (wet days) and 44% (precipitation) of the total. Most of these data were supplied on diskette, but about 25% were supplied either as published volumes of 1961–90 normals or on NMA paper copies. Data on the latter two media were scanned or keyed in, with independent checks against the originals.

2) WMO 1961–90 GLOBAL STANDARD NORMALS

The WMO recently published the 1961–90 climatological normals (WMO 1996). These data, for about 4000 stations, were purchased and merged with the CRU dataset. This resulted in an additional 690 stations that had data for at least one variable not previously in the CRU dataset.

3) CRU GLOBAL DATASETS OF STATION TIME SERIES

Global datasets of monthly time series of precipitation (Eischeid et al. 1991, updated), mean temperature (Jones 1994, updated), and maximum and minimum temperature (Easterling et al. 1997, with additions) for several thousand stations worldwide were searched for additional stations. CRU also holds smaller datasets of monthly time series of the other variables. These data have been quality controlled and checked for inhomogeneities. Station means for 1961–90 were derived from these time series and added to the global normals datasets.

4) CIAT SOUTH AMERICAN DATABASE

The Centro Internacional de Agricultura Tropical (CIAT) has collated several thousand climatological means for South and Central America (CIAT 1997). Unfortunately, the period each mean represents is unspecified, although the number of years of record contributing to the mean is usually supplied. These were assigned a low priority during the interpolation.

5) PUBLISHED SOURCES

Several countries in Africa (e.g., Zaire and Angola) and Southeast Asia (e.g., Cambodia) provided few or no 1961–1990 normals to either CRU or the WMO. In these cases data were extracted from one of two sources: Müller (1982) or FAO (1984). Generally, these means were calculated using data from the period 1931–60 and, in the case of the FAO publications, the number of years contributing to a mean was unknown. These data were also assigned a low weighting during the interpolation.

6) OTHER SOURCES

A small number of stations from several other sources were used. These included the U.S. Air Force Climatological Data Volume (USAF 1987) and a number of personal contacts where data for between one and several tens of stations were obtained.

b. Quality control

Data from the WMO collection were subjected to a fairly comprehensive series of quality control (QC) checks by the National Climatic Data Center (NCDC 1997). All WMO data that failed the QC were flagged as such by NCDC. We excluded all WMO stations that failed the NCDC QC tests from our analysis. Data obtained directly from NMAs were assumed to have been quality checked at source. Nonetheless, all data were subjected to a two-stage quality control process. In the first stage, prior to interpolation, a standard series of automated tests were performed on individual station normals. These tests were essentially the same as those used by the NCDC during the collation of the WMO 1961–90 climatological normals, namely:

- internal consistency checks, for example, ensuring that the monthly means follow a consistent seasonal cycle and that predefined absolute limits are not exceeded; and
- between-variable consistency tests, for example, ensuring that monthly minimum, mean, and maximum temperatures are consistent and that months with zero precipitation have zero wet days.

The second stage of QC occurred during the interpolation of station data, where the interpolation diagnostics enabled identification of station-months with large residuals (see section 3b).

As a general rule, data that failed these QC tests were removed from the interpolation. In some cases, however, the data could be compared and replaced with normals calculated from the CRU monthly station time series described above.

c. Variables

The final numbers of stations used in the interpolation are listed in Table 2. Five variables [wet day frequency, humidity (relative humidity and vapor pressure), sunshine, cloud cover, and frost days] had more than one definition and required standardization to a common unit. The procedures used to standardize these variables are described in the sections that follow.

The remaining variables (mean temperature, diurnal

TABLE 2. Number of station normals used in the construction of the climatology.

	Before standardization	After standardization
Precipitation	19 295	
Wet days greater than 0.1 mm	8371	9237
Mean temperature	12 092	
Diurnal temperature range	10 727	
Vapor pressure	3887	5940
Sunshine percent	4040	5181
Cloud cover	2678	4839
Ground frost days	1172	10 644
Wind speed	3615	

temperature range, precipitation, and wind speed) were not modified and the spatial distributions of the available stations are shown in Figs. 1–4. Although the way the variables are measured both between and within countries differ, it was felt such problems were neither possible nor important enough to correct for. Precipitation measurements can be influenced by several factors, most notably gauge type, the ratio of solid to liquid precipitation, and wind conditions/turbulence. Various attempts have been made to correct such biases in precipitation (e.g., Groisman et al. 1991; Legates and DeLiberty 1993), but we made no attempt to correct our precipitation data; in most cases there was insufficient information to attempt this. Differences in temperature measurement times have been shown to induce disparities of several tenths of a degree Celsius (Karl et al. 1986; Andersson and Mattison 1991) and different countries calculate mean temperature in various ways (Jones et al. 1985; Jones et al. 1986a; Jones et al. 1986b). Where possible, mean temperature was defined as the average of mean maximum and minimum temperature, which are measured more uniformly across the world. At the 1607 stations where only mean temperature was available, these values were used, despite the uncertainty about their derivation. Wind speed is measured at heights above the surface of between 2 and 20 m. Measurement height varies both within and between countries, and in many cases the heights were not specified. Consequently, no corrections were made to wind data. The large majority of known heights was 10 m, and the interpolated wind field should be assumed to represent speed at this height.

1) WET DAYS

These data were generally expressed as number of days per month with precipitation greater than 0.1 mm or 1.0 mm. The number of 0.1-mm normals exceeded 1.0-mm normals by a factor of two; consequently normals with a 1.0-mm threshold were converted to a 0.1-mm threshold. A small number of normals (e.g., United Kingdom and Australia) used some other threshold, typically 0.2 mm, and no adjustment was made for these more moderate differences.

Normals for both 0.1-mm and 1.0-mm thresholds were available for stations in a number of countries and these were used to derive an empirical conversion formula. The dataset of common normals was divided into two sections comprising stations in (roughly) temperate–polar (45°–90° south–north) and tropical–subtropi-

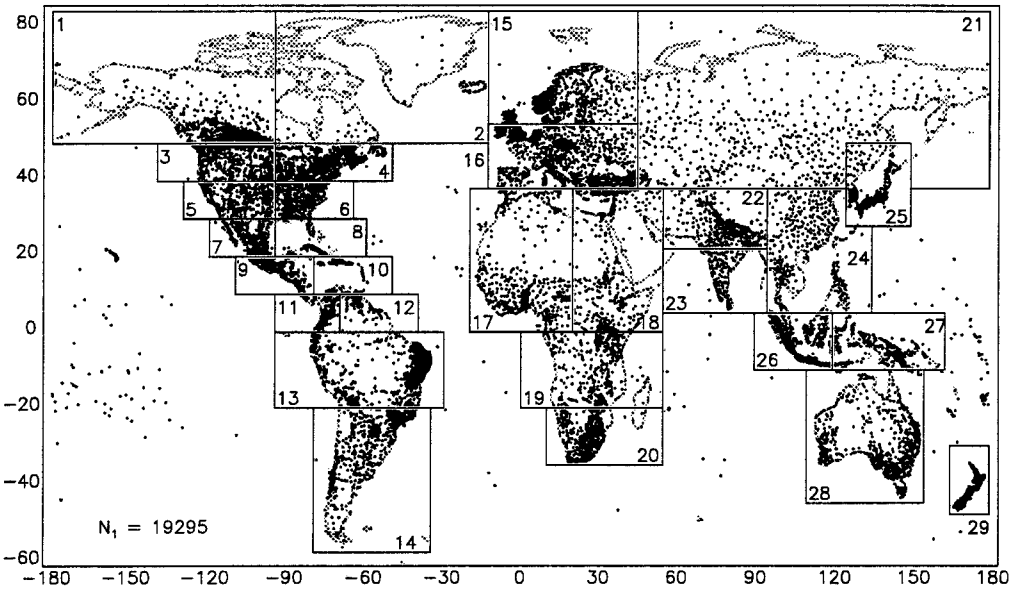

FIG. 1. Location of stations with precipitation normals. Geographic tiles used in the interpolation are shown and *N* signifies the total number of stations used. Note that (i) for all variables, oceanic stations were used during the interpolation of a global ''background'' tile and (ii) tile numbers and sizes differ between variables.

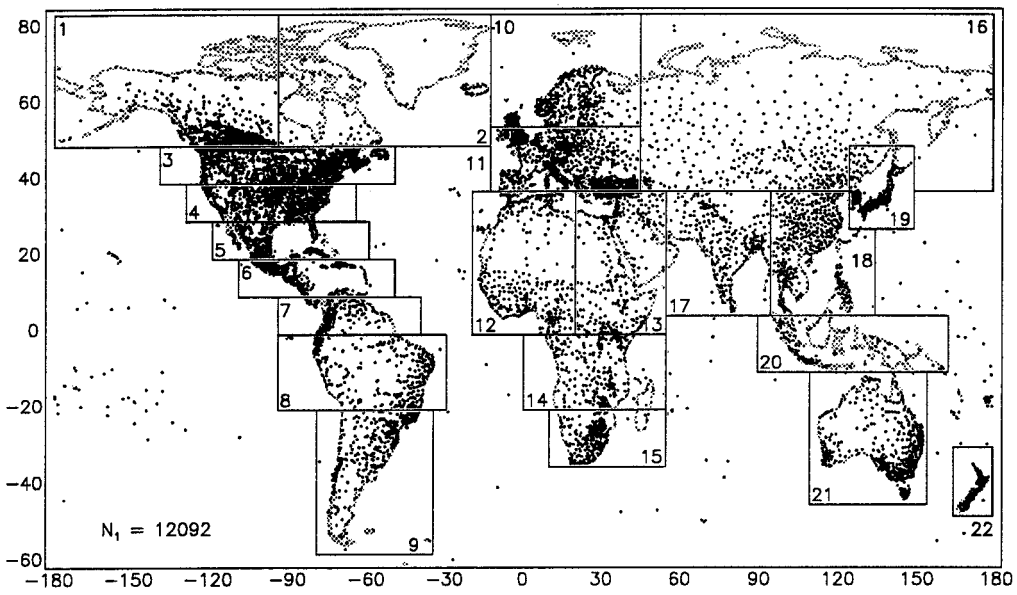

FIG. 2. Location of stations with mean temperature normals. Geographic tiles used in the interpolation are also shown, as are the total number of stations (*N*) used in the construction of the mean temperature fields.

cal (45°S–45°N) zones. Relationships between the two thresholds differed somewhat in the two domains, particularly during high-latitude winters. At low latitudes, monthly correlations between the two variables ranged from 0.86 in December–February (DJF) to 0.96 in June–August (JJA) (vice versa in the Southern Hemisphere), whereas in high latitudes, the correlation ranged between 0.65 in DJF and 0.94 in JJA (vice versa for Southern Hemisphere data). In each domain, the data were randomly split into equally sized calibration and validation datasets. The calibration datasets were used to derive linear regressions between the two thresholds on a month-by-month basis (Fig. 5). In the case of high-latitude stations, it was found that by including minimum temperature as a second predictor, the multiple correlation in winter months was improved to around 0.93. This temperature dependence in winter is probably related to the poor gauge catch during snowfall being

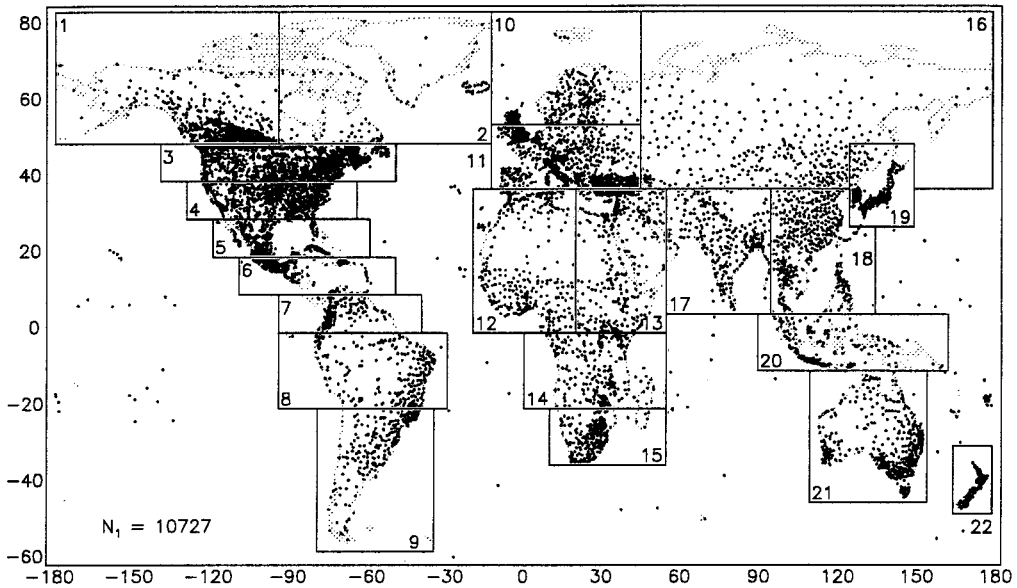

FIG. 3. Location of stations with diurnal temperature range normals. Geographic tiles used in the interpolation are also shown, as are the total number of stations (*N*) used in the construction of the diurnal temperature range fields.

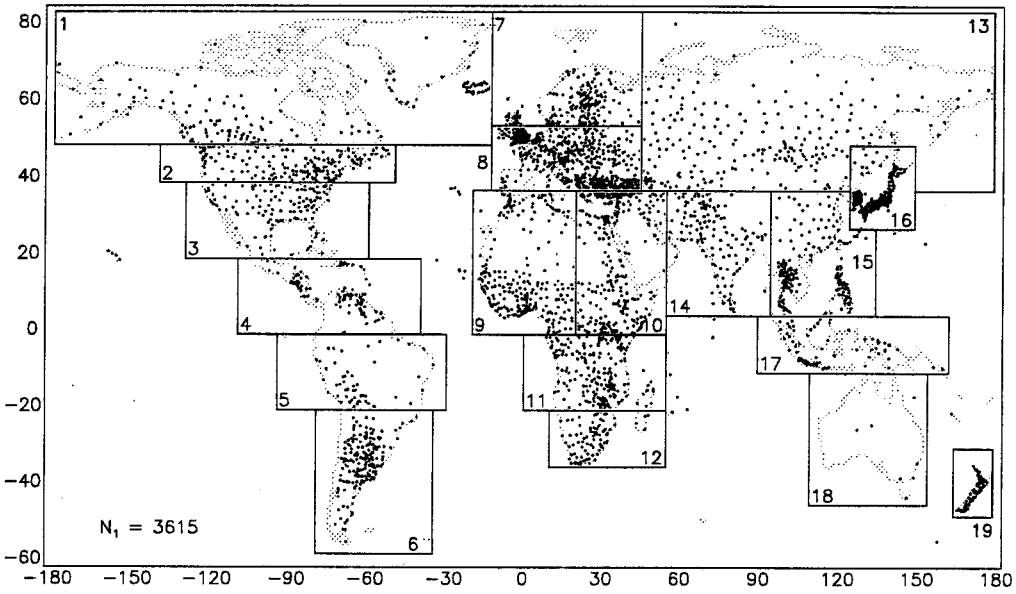

FIG. 4. Location of stations with wind speed normals. Geographic tiles used in the interpolation are also shown, as are the total number of stations (N) used in the construction of the monthly wind speed fields. No interpolation was done over tile 18, as there were insufficient data; however, the normals within this tile were used to interpolate the global background tile.

exaggerated when the fall is extremely light. In the summer half-year, the use of minimum temperature at high latitudes did not improve the regression. The final distribution of wet-day normals is shown in Fig. 6.

2) SUNSHINE AND CLOUD COVER

The distribution of cloud cover and sunshine normals is shown in Fig. 7 and Fig. 8, respectively. Sunshine

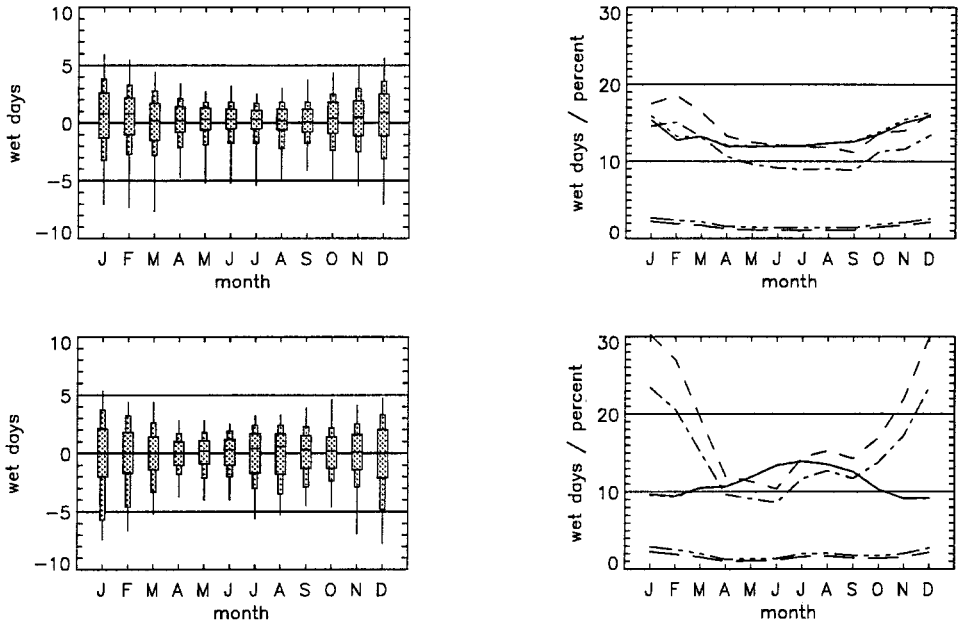

FIG. 5. Validation statistics for the >1.0 mm to >0.1 mm wet-day frequency conversion for high, i.e., >45°, (top) and low, i.e., <45°, (bottom) latitudes. Box and whisker plots show the distribution of residuals: median, interquartile range, 10th and 90th percentiles, and extremes. Line plots show mean observed and predicted wet-day frequency (solid and dotted), mean absolute and rms error in wet days (long dash and dash–3 dots) and these same errors expressed as the percentage of the observed mean (short dash and dash–dot).

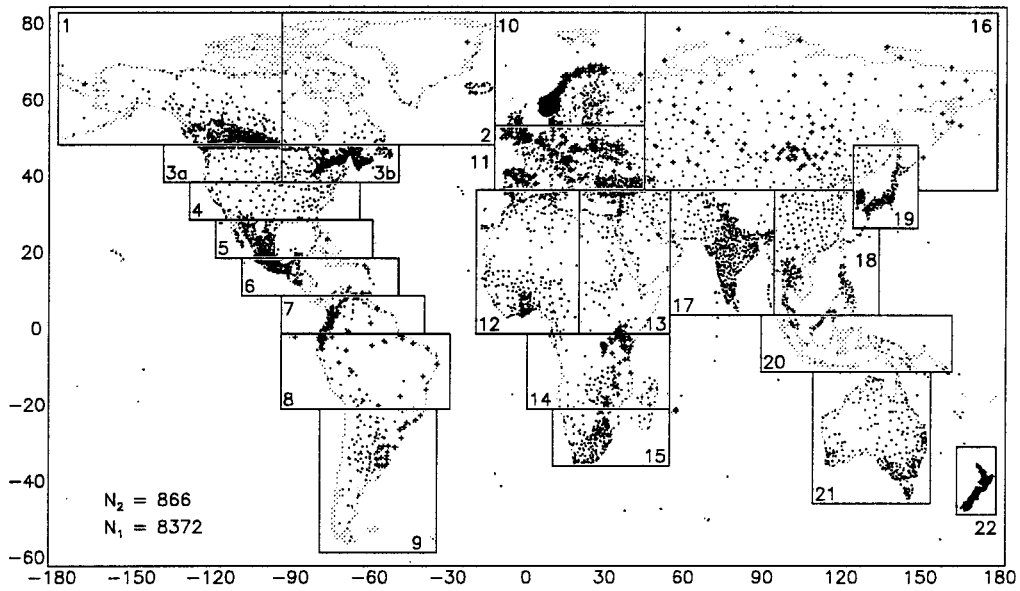

FIG. 6. Location of stations with wet-day frequency normals. Dots represent stations where the threshold was 0.1 mm and plus signs represent stations where frequencies were converted from the 1.0-mm threshold (e.g., northern Russia). Geographic tiles used in the interpolation are also shown, whereas N_1 and N_2 are the numbers of stations in the correct units and requiring transformation, respectively.

normals were supplied as either mean hours per month or percent of maximum possible bright sunshine. Total cloud cover normals were mostly provided in oktas and sometimes in tenths. Normals in units of sunshine hours were converted to percent of possible, and cloud cover normals were standardized to oktas. For some countries both cloud cover and sunshine data were available, but in most instances either one or the other was provided and it was necessary to convert cloud cover to sunshine (and vice versa) to obtain more complete coverage. Although this is elementary in principle, the method can produce only approximate results in practice due to fundamental problems in accurately measuring cloud cover and, in certain cases, sunshine (UKMO 1969). For example, Hulme et al. (1995) found that, at high latitudes, bright sunshine predicted from cloud was grossly over-

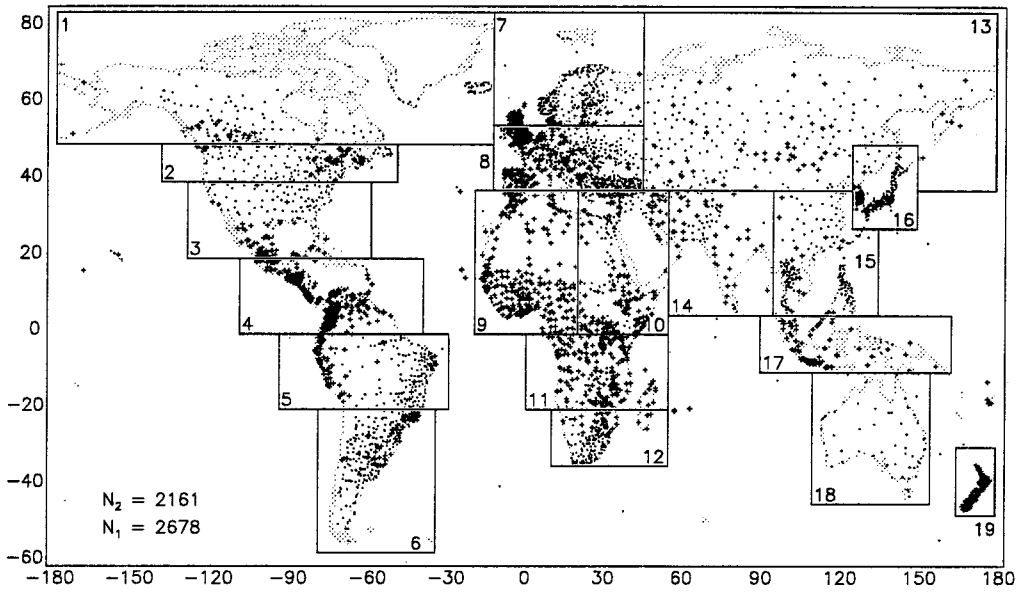

FIG. 7. Location of stations with cloud cover normals (dots, N_1) and those derived from sunshine (plus signs, N_2). Geographic tiles used in the interpolation are also shown.

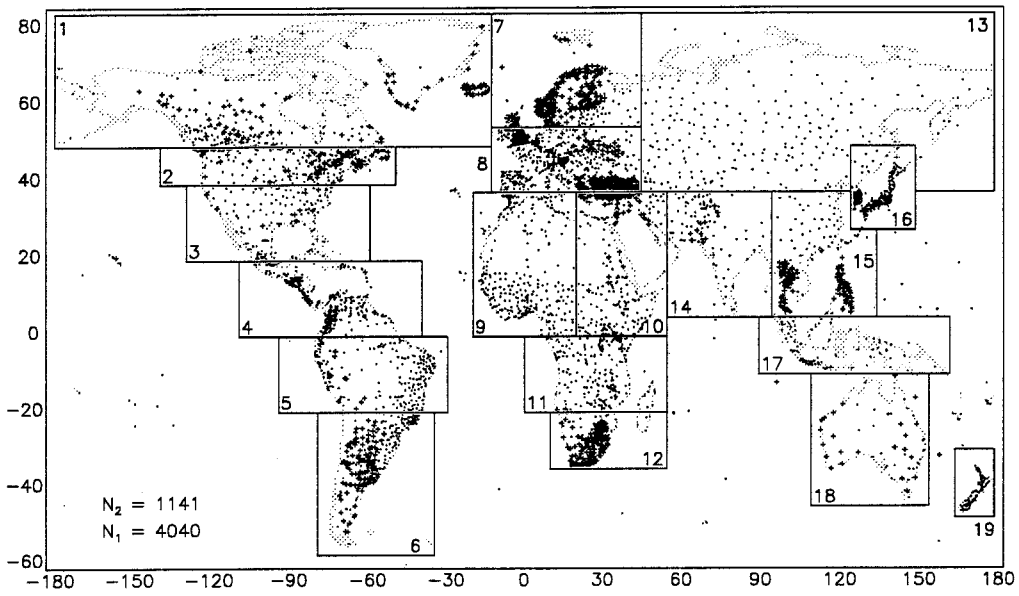

FIG. 8. Location of stations with sunshine percent normals (dots, N_1) and those derived from cloud cover (plus signs, N_2). Geographic tiles used in the interpolation are also shown.

estimated using standard conversion procedures such as those of Doorenbos and Pruit (1984); this is most likely due to weak sunshine at low sun angle not being recorded by sunshine instruments. Similarly, cloud cover estimated from sunshine would be overestimated. Conversely, reported cloud cover at low latitudes is frequently too high (B. Weare 1997, personal communication) because observers tend not to look vertically upward and, particularly in the presence of deep cumulus clouds, overestimate the amount of cloud cover. The cloud–sunshine conversion is further complicated by the fact that cloud cover is often a 24-h average, whereas sunshine relates only to daylight hours. To convert between the two, we therefore assumed that there was no diurnal cycle in cloud cover.

Following the approach of Hulme et al. (1995), the Doorenbos–Pruit procedures were used to estimate sunshine from cloud cover at the 1088 stations where both sunshine and cloud cover were available, and vice versa. The predicted values were then compared to observed normals and an empirical adjustment was derived to correct for the high-latitude and tropical biases described earlier (Fig. 9). The adjustments reduced the mean absolute prediction error in winter by 55% and 47% for cloud and sunshine, respectively (Fig. 10). The Doorenbos–Pruit cloud–sunshine and sunshine–cloud conversions, plus empirical adjustments, were then applied to those stations that had only one of the variables. This produced an additional 1141 and 2161 estimated sunshine and cloud cover normals, respectively. Normals estimated in this way were accorded a low weight during the interpolation (see section 3)

3) VAPOR PRESSURE

Humidity normals in the CRU dataset comprised roughly equal numbers of relative humidity (RH) and

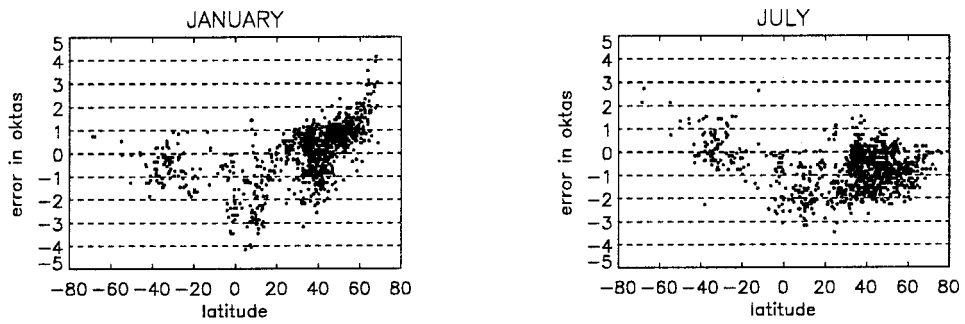

FIG. 9. Residuals (predicted minus observed) from the Doorenbos–Pruit uncorrected sunshine to cloud conversion.

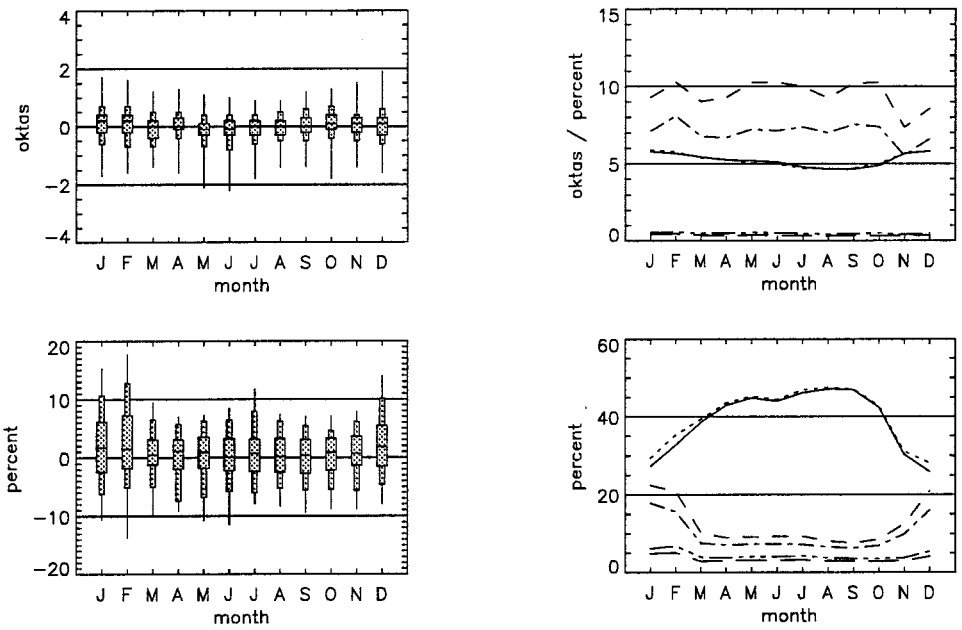

FIG. 10. (top) Validation statistics for the sunshine to cloud and (bottom) cloud to sunshine conversion, both with empirical correction. See Fig. 5 for explanation.

vapor pressure (e). The RH was converted to vapor pressure using the relationship

$$e = (\mathrm{RH}/100)e_{\mathrm{sat}},$$

where, after Shuttleworth (1992, 4.3),

$$e_{\mathrm{sat}} = 6.108 \exp\left(\frac{17.27T}{(237.3 + T)}\right) \quad \mathrm{hPa},$$

where e_{sat} is the saturated vapor pressure at the mean air temperature T.

This estimation is reliable provided that temperature and RH are measured simultaneously. In some cases the mean temperatures at the times of measurement were available and the estimated vapor pressure can be considered reliable. However, in most cases only the mean, or minimum and maximum, temperature were available, whereas RH was the mean of one or several daily measurements. Vapor pressure derived from RH and mean temperature must be considered approximate, particularly at stations with a large diurnal temperature range. The RH to vapor pressure conversion was validated on 978 stations, which had normals for both stations (Fig. 11). Vapor pressure at nearly all stations in the validation dataset was predicted to within 1 hPa of the observed value. Figure 12 shows the final distribution of vapor pressure normals used in the interpolation.

4) GROUND FROST DAY FREQUENCY

Most ground frost normals were defined as the frequency of grass minimum temperatures below 0°C. Some normals, however, were defined as the frequency of minimum air temperatures below 0°C and these had to be converted to ground frost frequency. As there was no straightforward theoretical basis for this conversion, the empirical formula derived by Hulme et al. (1995) was used:

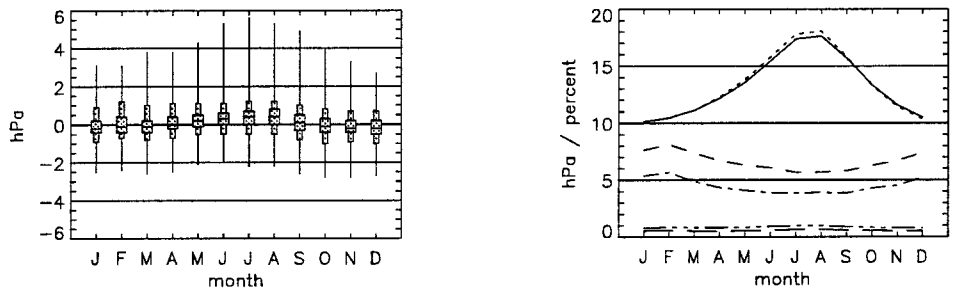

FIG. 11. Validation statistics for the RH to vapor pressure conversion. See Fig. 5 for explanation.

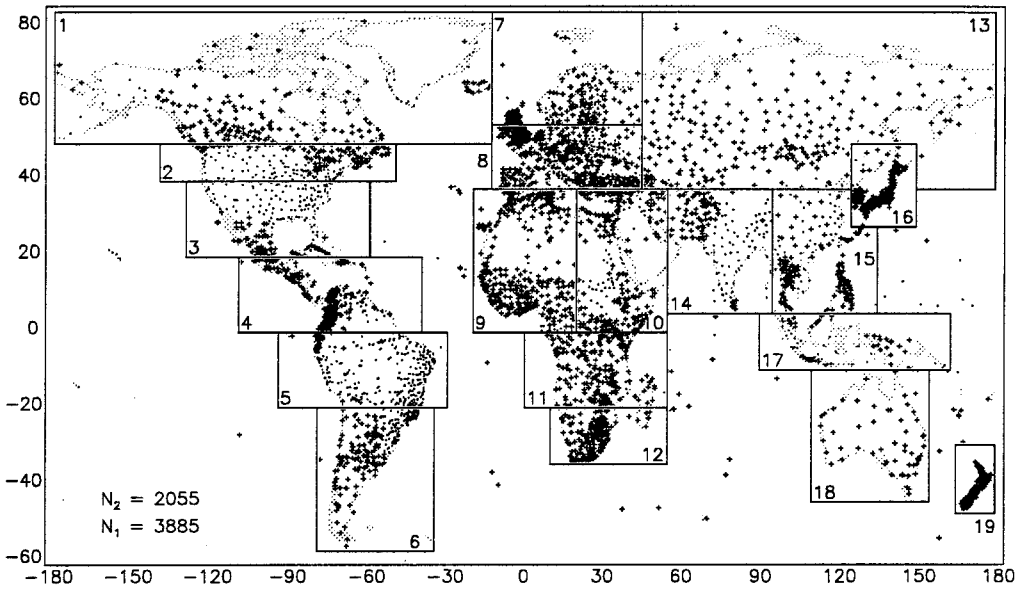

FIG. 12. Location of stations with vapor pressure normals (dots, N_1) and those converted from RH (plus signs, N_2). Interpolation tiles are also shown.

$$F_{gr} = 5.2 + (0.738F_{air}) - (0.284T_{mn}) \qquad T_{mn} < 9°C$$

$$F_{gr} = 0 \qquad T_{mn} \geq 9°C,$$

where F_{gr} is in days.

This conversion was validated using the 245 (non-European) stations that had both air and ground frost normals and that were not used by Hulme et al. (1995) in their derivation (Fig. 13). The validation stations show a mean negative bias (underprediction) of up to two days in the northern summer (the majority of stations occur in the Northern Hemisphere). This suggests that the relationship between ground and air frost is less distinct in warmer weather, perhaps being more site specific.

For much of the Tropics, no ground frost or air frost normals were available for the obvious reason that temperatures rarely fall below freezing point. This absence of data presented problems for the interpolation of frost

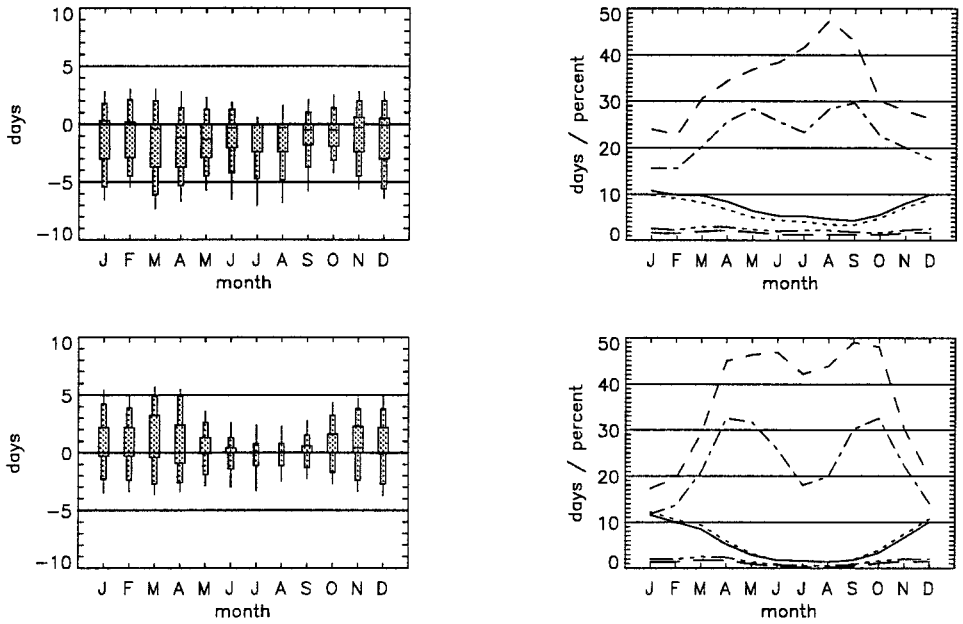

FIG. 13. (top) Validation statistics for the air frost and minimum temperature to ground frost and (bottom) the minimum temperature only to ground frost conversions. See Fig. 5 for explanation.

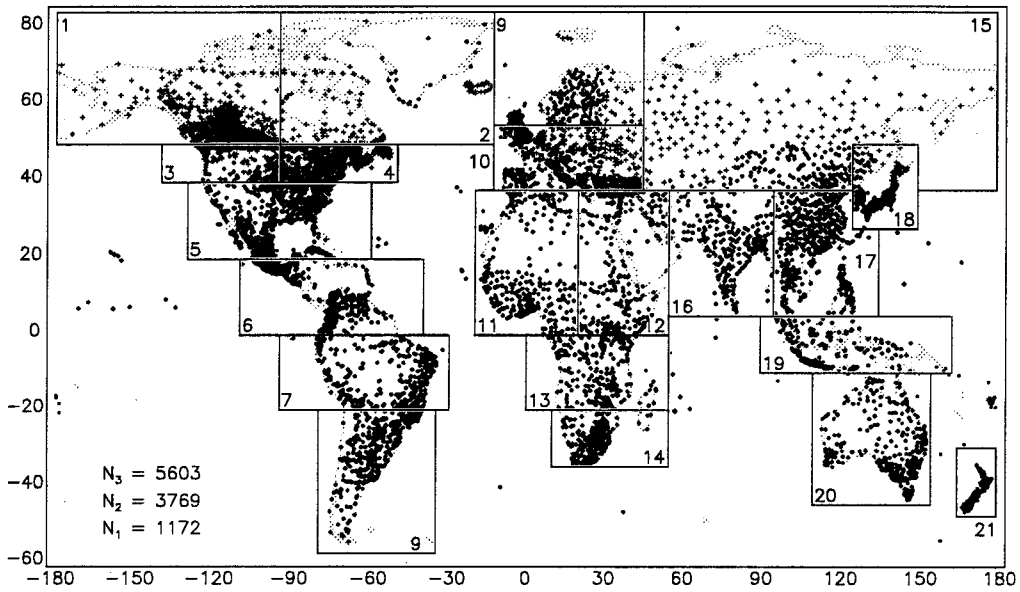

FIG. 14. Location of stations with ground frost normals (dots, N_1), those converted from air frost and minimum temperature (plus signs, N_2), and those converted from minimum temperature only (crosses, N_3). Interpolation tiles are also shown.

frequency to high elevations in the Tropics. To surmount this shortcoming, a second empirical formula, using only minimum temperature as the predictor, was derived using a randomly selected subset of half the 1094 stations, which had both minimum temperature and ground frost normals:

$$F_{gr} = 0.581 - 0.00670T_{mn} - 2.31 \times 10^{-5}T_{mn}^2 + 2.67 \times 10^{-7}T_{mn}^3 + 1.73 \times 10^{-9}T_{mn}^4 - 4.07 \times 10^{-9}T_{mn}^5 \qquad T_{mn} < 9°C$$

$$F_{gr} = 0 \qquad T_{mn} \geq 9°C,$$

where F_{gr} is in percent (i.e., frost days were standardized to avoid problems due to months with different numbers of days).

The temperature-only conversion performs almost as well as the air frost plus temperature conversion, but in contrast to the latter, shows a positive bias for most months of the year, particularly in winter. The smaller mean error in summer compared to the air frost method arises from the large number of normals from warmer regions used in the derivation and validation; in most cases T_{mn} was greater than 9°C, resulting in an exact prediction. The minimum temperature-only conversion was used at stations that had minimum temperature but no frost normals. The eventual distribution of ground frost normals used in the interpolation is shown in Fig. 14.

5) ELEVATION DATA

The digital elevation model (DEM) used in the study originated from the TBASE 5-min lat–long global DEM (NGDC 1996), modified to exclude all major water bodies (W. Cramer 1997, personal communication). These data were processed to a 0.5° lat × 0.5° long resolution by calculating the mean of the 36 5-min pixels present in each 0.5° cell. A 0.5° cell was considered to be ocean only if all 5-min pixels were ocean pixels. Land cells over Antarctica were also excluded, but ocean islands were included where the 5-min DEM indicated there was some land (i.e., at least one 5-min pixel). The resulting 0.5° DEM has 62 300 land cells, out of a possible 259 200 land and sea cells over the globe.

3. Interpolation methodology

The interpolation of irregular gauge data onto a uniform grid has been the focus of much research and a large number of methods have been proposed, ranging from the relatively simple Thiessen (Thiessen 1911) and distance weighting methods (Shepard 1968; Willmott et al. 1985) to geostatistical methods such as kriging (Phillips et al. 1992) or splines (Hutchinson 1995) and locally varying regression techniques (e.g., PRISM; Daly et al. 1994). For many climate applications it is important that elevation is included as a covariate or independent variable because the climate variable is dependent on elevation in some manner (Willmott and Matsuura 1995; Briggs and Cogley 1996).

a. Thin-plate splines

We used thin-plate splines to interpolate the climate surfaces as a function of latitude, longitude, and elevation. The original thin-plate spline-fitting technique

was described by Wahba (1979), whereas Hutchinson (1995) provides a theoretical description of their application to surface climate variables such as precipitation. The technique is robust in areas with sparse or irregularly spaced data points. Splines and kriging are formally equivalent, but are formulated differently (Wahba 1990; Hutchinson and Gessler 1994). Thin-plate splines are defined by minimizing the roughness of the interpolated surface, subject to the data having a predefined residual. This is usually accomplished by determining the amount of data smoothing that is required to minimize the generalized cross validation (GCV). This is calculated by removing each data point in turn and summing, with appropriate weighting, the square of the difference between the omitted point and a surface fitted using all the other points. The GCV is calculated implicitly and hence without recourse to computationally demanding iterative procedures. Kriged surfaces are defined by minimizing the variance of the error of estimation, which is normally dependent on a preliminary variogram analysis. The main advantage of splines over many other geostatistical methods is that prior estimation of the spatial autocovariance structure is not required (Hutchinson 1995).

During the GCV minimization, data points are weighted by σ^2/n, where n is the number of years of record contributing to the mean and σ^2 is the variance estimate obtained from the data. Where σ^2 is unknown, the data may be weighted by $1/n$, on the assumption that means derived from incomplete records are less likely to represent the true (in our case 30-yr) mean. Thus errors at stations with large variance or short records are assigned less weight. The use of σ^2/n is particularly appropriate for precipitation, which can show marked interannual variability (e.g., Hulme and New 1997). For other variables such as temperature, $1/n$ can be used because interannual variability is not particularly important.

Where the number of years of data for a normal was unknown it was assumed to be 10. Weights for normals from periods other than 1961–90 were reduced by 0.5 for precipitation and 0.75 for other variables. Similarly, the weights of normals that were derived from other variables (e.g., vapor pressure from RH) were reduced by 0.75. The reductions were multiplicative. Thus normals with a short or unknown record length, not sampling 1961–90 and derived from another variable were assigned the lowest weighting under this scheme. This procedure resulted in normals with the best temporal fidelity being assigned the highest weight.

Elevation is scaled in kilometers, which essentially exaggerates its influence by 100 times in relation to horizontal position. This has been shown to be of critical importance in ensuring that the interpolation is truly three-dimensional (Hutchinson 1995; Hutchinson 1998). Nonetheless, three-dimensional interpolation does require there to be sufficient elevation spread in the input data to enable smaller-scale elevation dependencies to be captured (in our case, smaller-scale implies resolution greater than 0.5° lat × 0.5° long). Therefore, in mountainous areas with sparse data coverage, the trivariate thin-plate spline interpolation is likely to underestimate lapse rates. The plots of station distributions (Figs. 1–4, 6, 7, 8, 12, 14) provide an indication of areas where sampling density is low.

The splining procedure returns a set of surface coefficients that can, in practice, be interrogated at any resolution provided a suitable digital elevation model is available. However, the fitted surface is only as good as the station data that were used to define it. If, as is the case here, interstation distances are of the order to 10–100 km, the fitted surface will reflect only spatial and elevation relationships captured at this resolution. Using a higher-resolution DEM to create a climatology can be misleading, as users might mistakenly assume the resulting climate fields adequately describe variability at this higher resolution. We feel that the 0.5° resolution used here represents the limit at which such large-scale interpolations can reliably be interrogated.

b. Interpolation procedure

The interpolation of a large number of data points becomes computationally demanding. In addition, fitting the same spline function to areas with markedly different station densities (e.g., Europe vs North Africa) can result in too much smoothing in data-rich areas and too little smoothing in data-poor areas.

For each variable, therefore, the terrestrial surface was divided into a number of geographic tiles over which separate spline functions were fitted (Figs. 1–4, 6, 7, 8, 12, 14). The size of the tiles varied primarily according to station density, but also as a function of spatial complexity of the climate variable. Where necessary, tiles were forced to overlap by at least 5° lat × 5° long so as to minimize edge effects. The number of stations in a tile varied between about 200 and 1000.

A ''background'' tile was also interpolated, encompassing the entire globe between 60°S and 85°N, using a subset of the available normals that included all stations not within the tiles mentioned above (i.e., oceanic islands) and some 750 evenly distributed continental stations.

No interpolations of ground frost were possible over tiles 12 (eastern North Africa) and 19 (Indonesia) because no stations were cold enough to have ground frost; this precluded interpolation to higher elevations where there might well be frost days. Consequently, ground frost frequency surfaces were derived from minimum temperature surfaces (mean temperature minus half of diurnal temperature range) using the relationship with minimum temperature described in section 2. Similarly, wind speed was not interpolated over Australia, as there were only seven normals for this region. Here, the global background field (see below) was used.

Output diagnostics from the spline programs were

TABLE 3. Percentage of grid cells where it was necessary to constrain the interpolated fields to realistic values. Only those variables where unrealistic values were detected are listed. Here Nmonth is the number of days in the month.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Precipitation < 0	4	3	3	2	3	4	4	3	3	3	4	5
Wet days < 0	3	2	1	1	1	2	2	2	1	1	2	2
Wet days > Nmonth	<1	<1	<1	<1	<1	—	<1	<1	—	—	—	—
Vapor > Saturation	9	7	4	2	1	<1	<1	<1	1	2	3	7
Sunshine < 0	2	<1	—	—	—	—	—	—	—	—	<1	4
Cloud < 0	—	—	—	—	—	<1	<1	—	—	—	—	—
Ground frost < 0	14	15	16	17	18	21	27	24	18	17	16	14
Ground frost > Nmonth	32	32	30	21	8	2	<1	<1	4	16	28	32

used to identify erroneous data. Most common errors were typographic and locational. Where possible, the errors were corrected; stations that could not be corrected were removed from the dataset.

The final fitted spline functions for each tile were applied to the portion of the 0.5° DEM falling within the tile to derive the climate grids for each variable. The tiles were then merged to produce a global land field (excluding Antarctica). Where there was overlap between tiles, for example, between tiles one and two in Fig. 1, grid values were calculated as a weighted average from contributing tiles. Weights were simply a linear function of the inverse-distance in grid points between a particular point and the edge of its tile.

Finally, the grids were constrained as follows to avoid unrealistic values. For all variables except temperature, negative values were converted to zero. Wet-day and frost-day frequency were limited to the number of days in the month under consideration. In areas with zero rainfall, wet-day frequency was set to zero. Vapor pressure was constrained to be less than the saturated vapor pressure at the corresponding mean temperature. Sunshine and cloud were set to upper limits of 100% and 8 oktas, respectively. The number of grid points that had to be constrained is listed, for each relevant variable, in Table 3. These primarily occurred throughout the year in the Sahara and Arabia, but were also present during the austral winter over drier regions of Namibia, Botswana, Angola, and Peru. Wet-day frequency exceeds the numbers of days in the month over the Amazon Basin, particularly between February and April. This arises from a combination of a sparse station network and the fact that during this period (the wet season) the true wet-day frequency is high; thus the likelihood of values being interpolated beyond the upper limit is enhanced. The wet-day grid points with values less than zero are similar in distribution to precipitation, as might be expected.

The number of vapor pressure grid cells with values greater than saturated vapor pressure was quite large for winter months. The cells are all situated in regions with extremely low temperatures, namely, Canada, Greenland, and Siberia. Humidity measurements at these temperatures are subject to considerable uncertainty because vapor pressure is so low, often below 0.1 hPa, the precision of most of the vapor pressure normals. When interpolating such low values, small absolute errors will result in rather large errors relative to the similarly low saturated vapor pressure derived from mean temperature. Thus, vapor pressure in cold regions should be considered as approximate. In summer, the few vapor pressure points that are greater than saturated vapor pressure occur in the Greenland interior. These are most likely due to the problem of low vapor pressure mentioned above. However, station control for both the mean temperature and vapor pressure interpolation is limited over the Greenland interior, so there may be some additional error in both the vapor pressure estimates and the temperature-based estimates of saturated vapor pressure used to constrain the vapor pressure.

All the sunshine grid points with negative values occur at high northern latitudes in winter and result from the extrapolation of station normals with values close to or equal to zero into regions without station control, namely northern Russia and Greenland. The few cloud grid cells with values less than zero occur in the Arabian desert, some distance from input station control.

There were numerous cells where it was necessary to constrain the ground frost estimates. This is because the interpolation is unable to handle lower and upper limits. Thus in areas where the station data indicate ground frost is close to the upper (number of days in the month) or lower (no days) limit, gradients extrapolated away from the control can result in many grid points being assigned values outside the limits. This is reflected in the seasonal variation of the number of cells greater (less) than the upper (lower) limits (Table 3).

4. Validation and reliability of interpolated fields

An issue that has received little attention in recent years is the frequent use of climatologies interpolated from station data without reference to the predictive error associated with the field (Hulme and New 1997; Legates 1997). End-users of climate data are, more often than not, unfamiliar with the details of the interpolation and do not have a feel for the overall accuracy of the interpolated field or are unaware of areas where the interpolation is more or less reliable. In this section we assess the accuracy and reliability of the interpolated fields in our new climatology.

One method for independently assessing the predic-

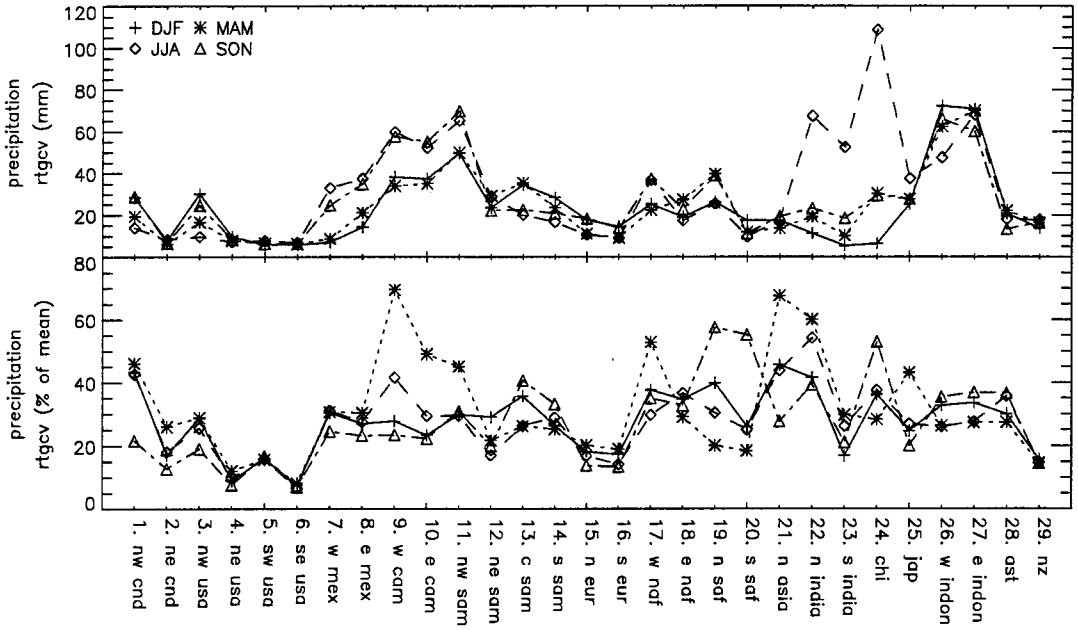

FIG. 15. Seasonal RTGCV for precipitation over each interpolation tile. For each season, (top) is in mm units and (bottom) in percent of the mean of stations falling within the tile. Tile indices are shown in Fig. 1.

tive error of an interpolated field is cross validation. These validation errors can then be appraised in a number of ways, for example, by looking at spatial patterns of error (Legates 1987) or in terms of average behavior, such as mean-square error and mean bias (Hulme et al. 1995; Hulme et al. 1996). Although this can be computationally demanding, it complements the standard error estimates of geostatistical methods, which are largely dependent on the statistical model being suitable for the data. A major shortcoming of cross validation is that it provides error information at station locations only and can therefore be misleading if the station network is inadequate. The thin-plate spline interpolation we used calculates the generalized cross validation (equivalent to the mean-square cross validation error) implicitly during the surface fitting procedure, and we assess these in some detail in section 4a below.

Another way of evaluating the predictive error associated with an interpolated field is by comparison with other datasets that nominally describe the same variable. This has been done over the United States by Pan et al. (1996), who compared several precipitation and temperature climatologies—those of Leemans and Cramer (1991; W. Cramer 1997, personal communication), Legates and Willmot (1990a,b), and the VEMAP project (Kittel et al. 1995)—and found large differences between the datasets, of the order of 20%–40% for precipitation in some places. These differences were due to several factors: different observing networks; different interpolation schemes; and for precipitation, differences in corrections made for gauge biases. Similarly, Hulme and New (1997) compared the Leemans and Cramer and Legates and Willmott precipitation climatologies to their own over North Africa and Europe, and found differences of between 10% and 30%. In section 4b below we compare the new CRU 1961–90 climatology to these other commonly used global climatologies.

a. Spline diagnostics

The GCV and its square root (RTGCV) provide an estimate of the mean predictive error (and hence power) of the fitted surface and as such permit an assessment of the relative accuracy of a fitted surface. Although the RTGCV from different tiles is partly dependent on the sampling adequacy of the station network, we compared the GCV results from different tiles. This provided an indication of areas where the interpolation is more or less accurate. The RTGCV for each variable and each interpolation tile are summarized in Figs. 15–17.

In general, precipitation, wet-day frequency, and wind speed have the largest relative RTGCV. In the case of precipitation, this is despite being the variable with the largest number of stations ($N = 19\,295$) and is a reflection of its greater spatial variability, even when averaged over a large number of monthly totals. Precipitation prediction errors are greatest where the station network is too sparse to capture the spatial variability of this variable. Examples include mountainous regions such as the North and South American cordillera and the Himalayan region, all of which suffer from relatively poor station coverage and complex precipitation patterns. In northern Asia, where precipitation is less spatially variable but the network is extremely sparse, relative errors are similarly large. In well-sampled regions,

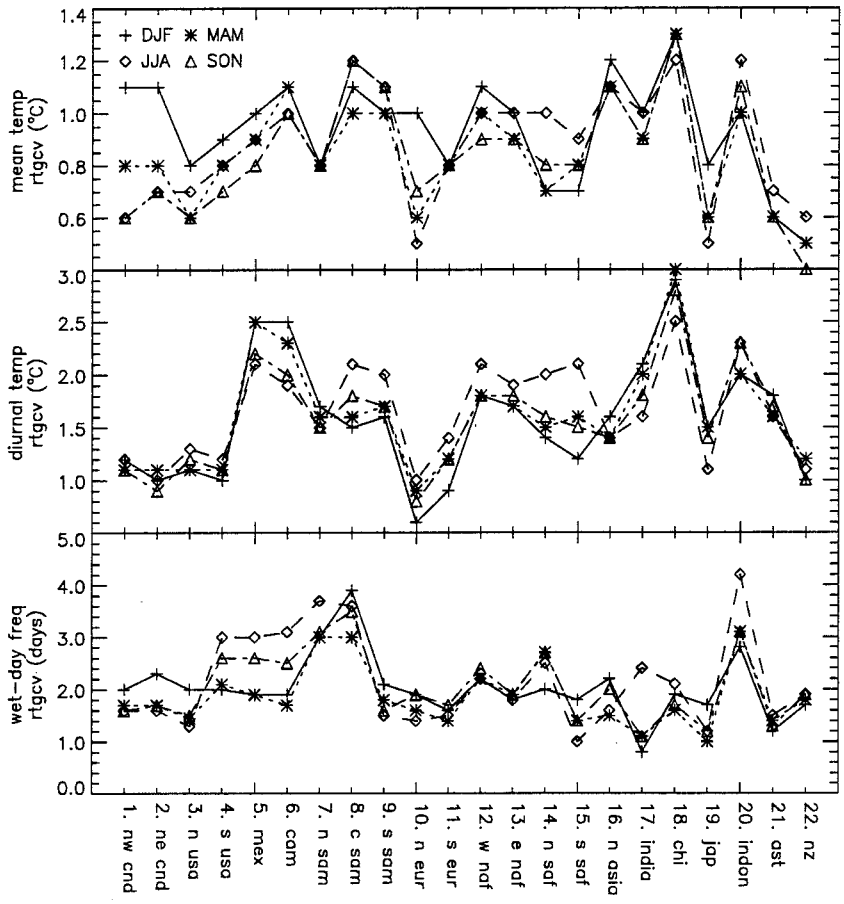

FIG. 16. Seasonal RTGCV for (top) mean temperature, (middle) diurnal temperature range, and (bottom) wet-day frequency over each interpolation tile. Tile indices are shown in Figs. 2, 3, and 6.

such as Europe and the United States, the RTGCV of 10%–25% agrees well with error statistics from other regional precipitation climatologies (e.g., Daly et al. 1994; Hulme et al. 1995). The largest absolute RTGCVs occur in the Indonesian region (tiles 26 and 27), a reflection of the high precipitation that occurs there.

Wet-day frequency errors are generally between 1 and 3 days (Fig. 16), while relative errors are between 10% and 30% of the mean of all stations over each tile (not shown). The exceptions are northern and central South America, where the RTGCV is between 3 and 4 days, and the JJA season in Indonesia. Both these regions have high precipitation and wet-day frequency and relatively sparse networks, so high absolute errors might be expected. It is only in the central South American dry season that relative errors are appreciably larger than in other areas. Large relative errors also occur over the Indian subcontinent during the dry season (DJF and March–May), but these have low absolute errors of 0.8 and 1.1 days, respectively.

RTGCVs for wind speed are also large (Fig. 17), between 0.5 and 2.0 m s^{-1} (between 20% and 77% of the mean of all stations in individual tiles). As with precipitation, wind speed is influenced by small-scale physiographic features and also has quite steep coast to inland gradients. The sparser observing network (N = 3615), combined with possible uncertainties with regard to instrument type and height, means that the resultant monthly grids capture only spatial variability associated with larger-scale circulation features.

Mean temperature is far better predicted than diurnal temperature range, despite the two variables having similar station networks (Fig. 16). The RTGCV for mean temperature ranges from 0.5° to 1.3°C, and is typically about half that for diurnal temperature range. Errors at the high end of the range occur in less well-sampled regions. The poorer results for diurnal temperature range primarily reflect the more complex character of minimum temperature, which can be strongly influenced by local topography and land–sea interactions.

Cross validation of the other variables (Fig. 17) resulted in RTGCVs that were less than 20% of the mean of stations used in the interpolations. The largest vapor pressure relative errors of 11%–16% occur over the Indian subcontinent, which has a comparatively sparse network. Cloud cover and sunshine percent have a sim-

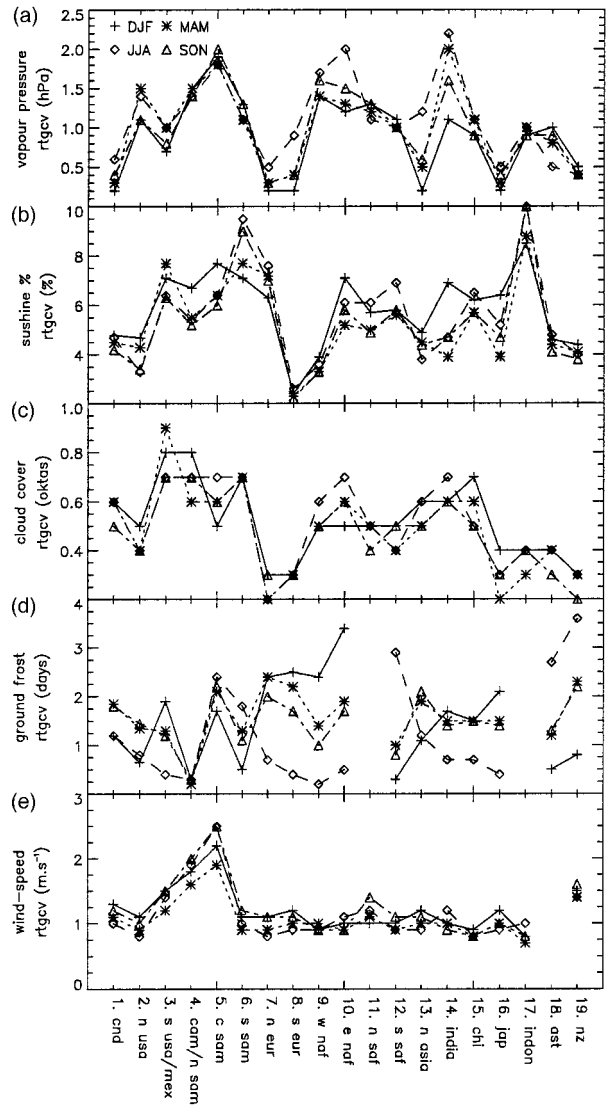

FIG. 17. Seasonal RTGCV for (a) vapor pressure, (b) sunshine, (c) cloud cover, (d) ground frost frequency, and (e) wind speed over each interpolation tile. Tile indices are shown in Figs. 4, 7, 8, 12, and 14.

ilar range in RTGCV, once the difference in units is taken into account. Relative RTGCVs are inversely related because of the inverse relationship between the two variables. Some of the error in the RTGCV for cloud and sunshine is likely to stem from the problems in converting from sunshine to cloud (and vice versa), discussed in section 2b.2. Ground frost frequency RTGCVs vary between 0.3 and 3.4 days, but are generally greater than 1.0 in regions where ground frost is a significant factor.

b. Comparison with other climatologies

Two commonly used 0.5° global surface climatologies are those of Legates and Willmott (1990a,b, henceforth LEG) and Leemans and Cramer (1991, henceforth CRA), who have recently released an updated version (Version 2.1; W. Cramer 1997, personal communication). LEG constructed a mean monthly climatology for precipitation and mean temperature, with complete global coverage. LEG has predominantly been used in climatological applications, particularly for the evaluation of GCM precipitation fields, and represents both land areas and oceans, for the period 1921–80, with greater emphasis on the latter years. The monthly fields over land were interpolated from about 25 000 precipitation and 18 000 mean temperature stations worldwide, using a spherical angular distance weighting algorithm (Willmott et al. 1985). CRA have produced mean monthly fields of precipitation (N = 18 927 stations), wet-day frequency (N = 6262), mean temperature (N = 13 656), diurnal temperature range (N = 9927), and sunshine percent (N = 5263). Their spatial domain is the same as CRU's, namely global land areas excluding Antarctica, but represents (roughly) the period 1931–60. CRA has been derived using the same thin-plate spline interpolation routines that were used for construction of the CRU climatologies. CRA was developed for use in global ecosystem modeling and has received most application in these disciplines.

We compared the CRU climatologies to those of LEG (for precipitation and mean temperature) and CRA (for precipitation, mean temperature, diurnal temperature range, wet-day frequency, and sunshine). LEG required resampling to arrive at grid cells with the same centroid as CRU and CRA. LEG was further processed so that only areas that were in common with the CRU and CRA land-only fields were assigned values. Subsequently, all three climatologies were degraded to produce 2° lat × 2° long grids, in order to improve viewability. Neither LEG or CRA have climatologies of cloud cover, so we compared the CRU cloud cover fields to those of Hahn et al. (1994) and the International Cloud Climatology Satellite Project (ISCCP; Rossow and Schiffer 1991) mean climatology.

We are not aware of any global climatologies that describe vapor pressure, ground frost, and wind speed, so these fields remain unevaluated.

1) PRECIPITATION

CRU precipitation fields were compared to those of LEG and CRA. It should be noted that the LEG precipitation fields we used are the *unadjusted* version, in that they contain no adjustments for gauge biases.

Zonal-mean precipitation through the seasonal cycle is shown in Fig. 18, as are the zonal-mean absolute error (MAE) and bias between each pair of climatologies. The zonal-mean precipitation in all three climatologies agrees closely, except at high latitudes. Along the Pacific coastline of southern South America (50°S–60°S), CRU is drier than both LEG and CRA. This can mostly be explained by examination of regional time series con-

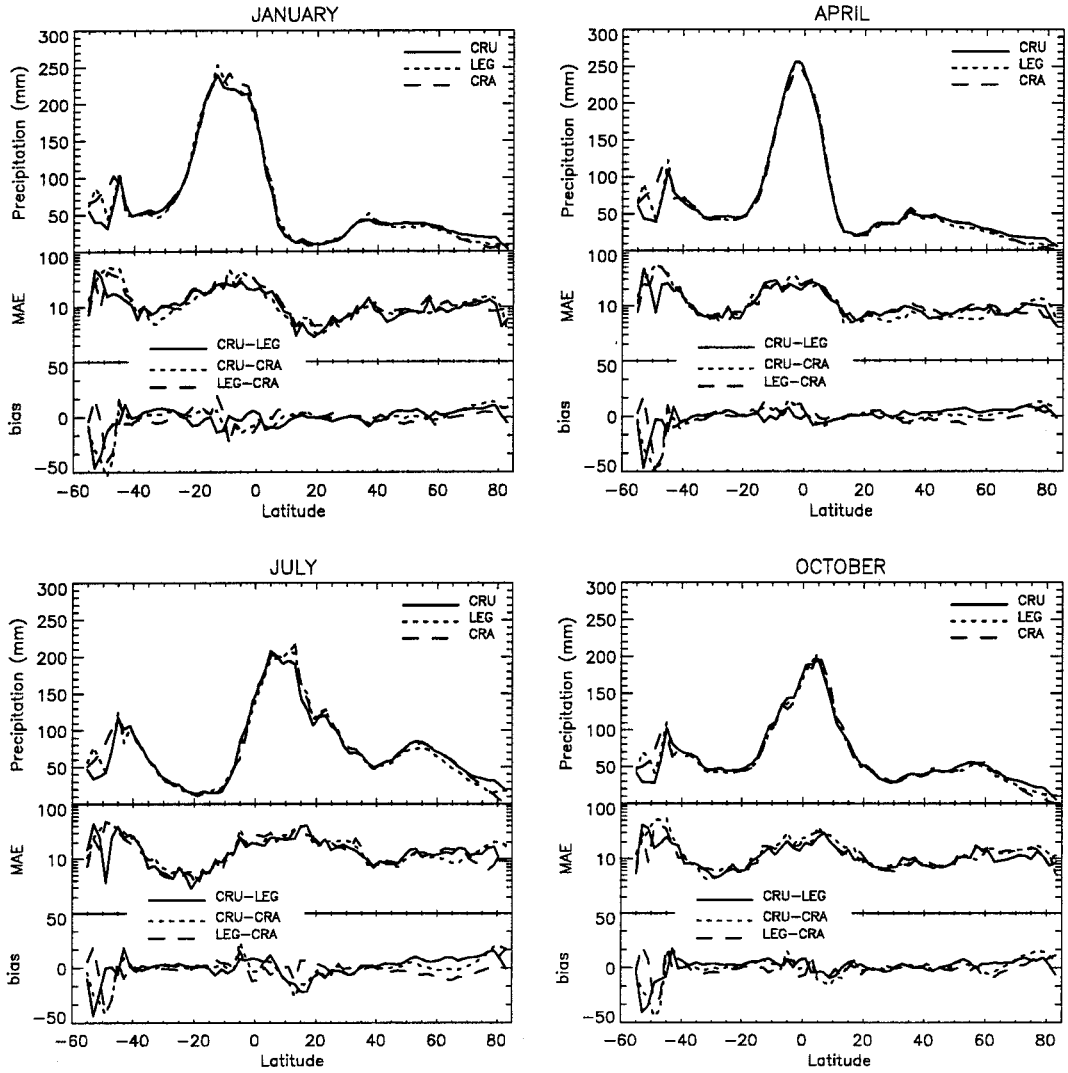

FIG. 18. Zonal-mean precipitation, mean absolute differences (MAE), and mean bias between climatologies, for CRU, LEG, and CRA. The midseason months of Jan, Apr, Jul, and Oct are shown.

structed from CRU station time series (not shown), where it can be seen that mean precipitation for the period 1931–60 and 1921–80 is wetter than the period 1961–90 by between 20% and 50%. If these effects are removed, MAE between CRU and the other two climatologies would be smaller. In the Arctic, CRU precipitation is, on average, greater than both LEG and CRA by about 5 mm month^{-1}, but CRA and LEG do not agree particularly well here either. These differences are understandable given the sparseness of station data at these latitudes. In particular, the absence of gauges has resulted in fairly large differences between climatologies over the Greenland interior.

We attempt to quantify the significance of differences between the CRU, LEG, and CRA precipitation climatologies in Fig. 19, which shows gridpoint Student's t-statistics for the differences between each pair of climatologies. The t-statistic was calculated at individual gridpoint pairs as follows (Hulme and New 1997):

$$t = \frac{m_1 - m_2}{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^{1/2}},$$

where m and s are the gridpoint means and standard deviations, respectively, and n is the number of years the mean climatologies represent. Monthly standard deviations were calculated from the CRU 0.5° lat × 0.5° long gridded monthly precipitation time series (New et al. 1999), using the periods 1931–60, 1940–80, and 1961–90 for CRA, LEG, and CRU, respectively. The 0.5° gridded monthly standard deviation fields were subsequently degraded to 2° lat × 2° long resolution. Values for n were assumed to be 30 for CRU and CRA, but

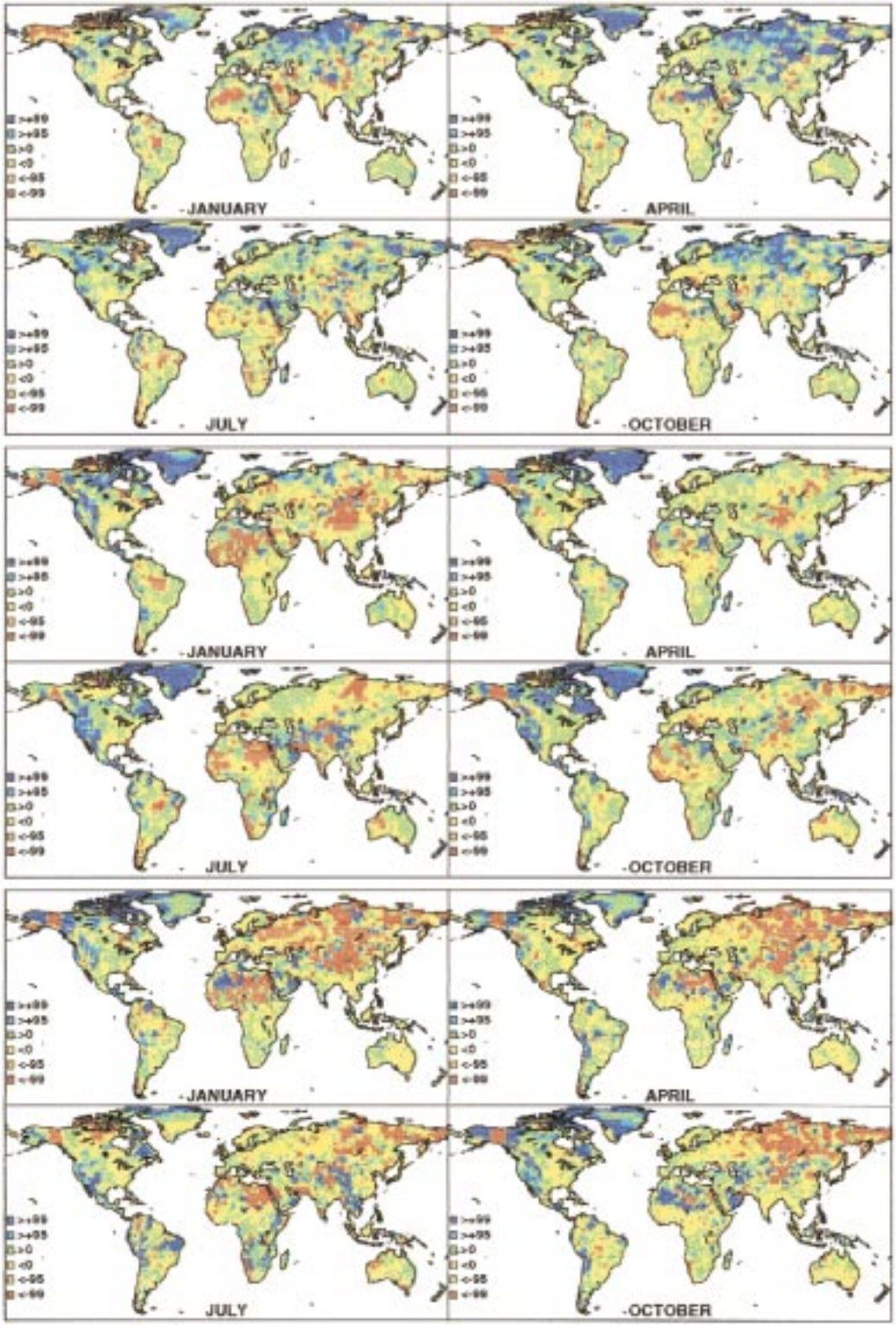

FIG. 19. Student's t-statistic for the differences (former minus latter) between the (top) CRU and LEG, (middle) CRU and CRA, and (bottom) LEG and CRA mean monthly precipitation climatologies (regridded to 2° lat × 2° long). Values in the legend correspond to the significance of the t-statistic, with green-blue representing areas where the former is wetter than the latter and yellow-red where the former is drier than the latter.

40 for LEG, because the stations are described by the authors as being more heavily sampled in the latter years of the nominal 1921–80 normal period.

There are well-documented shortcomings in the use of t-statistics for assessing differences between fields of climate data (e.g., Wigley and Santer 1990). In addition, monthly precipitation is not normally distributed and, in this instance, assuming constant degrees of freedom at each grid point is not strictly valid. Nonetheless, the plots of gridpoint t-statistics highlight areas where differences between the climatologies are more or less significant. A further advantage of this approach is that it avoids the necessity of looking at difference fields in combination with fields of absolute precipitation amounts, or of having to use relative and absolute difference fields in areas of high and low precipitation, respectively.

Comparison of the t-statistic fields indicates that most of the significant differences occur in data-poor regions, such as Amazonia, the Sahara, eastern Russia, and the Arctic, or in regions with complex topography, such as the American cordillera. CRU is drier than the other two climatologies over the Sahel region, which is to be expected, given the well-documented drought that occurred here in the 1970s and 1980s and the fact that LEG and CRA are sampling earlier, wetter decades (Hulme and New 1997). Over Russia, CRU and CRA show no consistent pattern of agreement, but both exhibit higher precipitation than LEG at most grid points. In the case of CRU, this is because most of the precipitation normals are derived from time series that have been adjusted for gauge undercatch (Groisman et al. 1991), whereas the LEG version used for this comparison is not adjusted for gauge biases. Other areas where CRU shows significant differences from LEG are over northern Canada and Alaska, where CRU is drier; both CRU and LEG are significantly different from CRA in this region, with similar sign over Alaska, but opposite sign in the Arctic.

CRU is wetter than both LEG and CRA over most of the Greenland interior through all seasons, but particularly in summer. This disagreement is understandable given the absence of station data away from the coast. We therefore compared our estimate of annual precipitation over the region (not shown) to maps from several sources reproduced by Chen et al. (1997). The various climatologies agree well along the coast where station control is reasonable. Over the northern interior, however, the CRU annual precipitation is between 40 and 60 mm, markedly higher than the 10–30 mm reported in Chen et al. (1997).

2) WET-DAY FREQUENCY

The only other global-mean monthly climatology of wet-day frequency of which we are aware is that of Leemans and Cramer (Version 2.1; W. Cramer 1997, personal communication) and we compare this to the CRU wet-day frequency climatology. Zonal-mean wet-day frequencies agree to within 3 days, except at high latitudes and during the wet season (austral summer) in the Southern Hemisphere Tropics (Fig. 20).

The large zonal-mean differences between 45°S and 55°S are due to CRA exceeding CRU by between 5 and 18 days over South America. Station coverage in CRU is particularly poor in this area (Fig. 6) and, although no information was available on CRA's station network, it is likely to be similarly sparse. This, combined with the fact that precipitation in 1961–90 was up to 50% less than in 1931–60, probably explains much of the difference.

Differences in the southern Tropics arise mainly because CRU wet-day frequency is greater than CRA over South America and Indonesia. Again, both these regions have sparse observational networks, which may account for some of the differences. A second factor could be the fact that many of the CRA wet-day frequencies may be for a higher threshold than the CRU normals (W. Cramer 1997, personal communication). It is worth noting that because CRU precipitation is lower than CRA and CRU wet-day frequency is higher over much of tropical South America, rainfall intensities derived from the CRU precipitation and wet-day frequency grids will be much less intense than those from CRA in this region.

The large differences in the Arctic, where wet days in CRU are less frequent than in CRA, occur in all seasons except in summer, where the two climatologies agree closely. Possible reasons for this include the sparse observing network or changes in reporting procedures and/or gauge catch efficiency between 1931–60 and 1961–90. The latter explanation seems likely because the differences are largest and extend further south in winter, and are negligible in summer, suggesting that they are related to snowfall in some way. Precise reasons are difficult to determine because of the lack of documentation that accompanied the CRA dataset; we do not know the threshold or whether precipitation is liquid or solid plus liquid. In the case of CRU station data, Canadian wet-day frequency normals supplied were explicitly described as days with precipitation (solid and liquid) greater than 0.1 mm. Normals from Russia used by CRU were derived directly from 6-h or 3-h observations of precipitation and so also represent days with solid and liquid precipitation greater than 0.1 mm.

3) MEAN TEMPERATURE

As with precipitation, we compared the CRU mean temperature climatology to LEG and CRA. CRU and CRA both make use of thin-plate splines for interpolation, in which elevation is an independent predictor variable. Thus, lapse rates are derived during the interpolation and these rates vary in space. Conversely, LEG does not incorporate elevation effects in their inverse-distance weighting interpolation.

Zonal statistics for January, April, July, and October

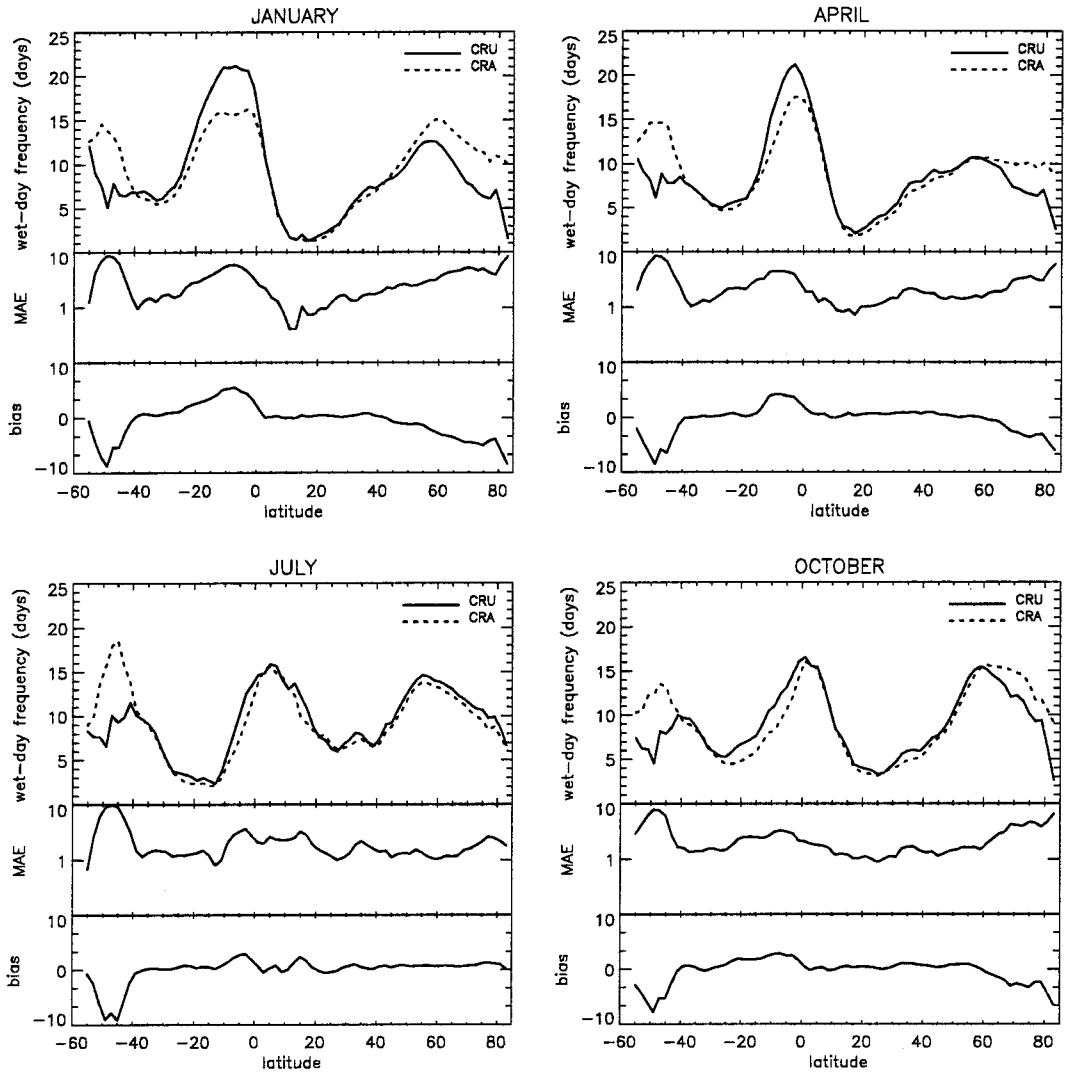

FIG. 20. Zonal-mean wet-day frequency for CRU and CRA, and zonal MAE and bias between CRU and CRA, for each of the four peak-season months.

are displayed in Fig. 21, where the strong latitudinal control of zonal-mean temperature patterns is evident. Mean absolute differences between CRU and CRA are consistently less than those between LEG and either CRU or CRA, by between 0.5°C and 2°C. The closer agreement between CRU and CRA is most likely a consequence of the use of the same interpolation method producing similar elevation-dependent gradients. This is most evident at 30°–40°N, where LEG is much warmer than CRU and CRA over the Tibetan plateau, by up to 15°C. Similarly CRU and CRA are colder than LEG in other areas with high elevation: the American cordillera, New Zealand, and the European Alps.

Examination of the spatial patterns in difference fields (Fig. 22) shows that CRU is generally slightly warmer than CRA and, to a lesser extent, LEG, by about 1°C over most of the Southern Hemisphere and Asia, whereas patterns elsewhere are more variable. Over the Greenland interior, CRU is warmer than CRA in winter, by up to 15°C, but colder in summer by a similar amount. The absence of long-term measuring stations here is clearly a major cause of discrepancy. CRU and CRA are both colder than LEG over the Greenland interior, except for a small area toward the south. There, LEG is colder, suggesting the presence of a station that was used in the LEG interpolation, but not in either CRU or CRA.

CRU is distinct from both LEG and CRA along the Arctic coasts of Russia and North America. In Russia, CRU is colder in winter and warmer in summer, whereas the reverse occurs over North America, albeit less distinctly. This is most likely due to differences in the way mean temperature was calculated in each climatology. For example, in colder regimes mean temperatures de-

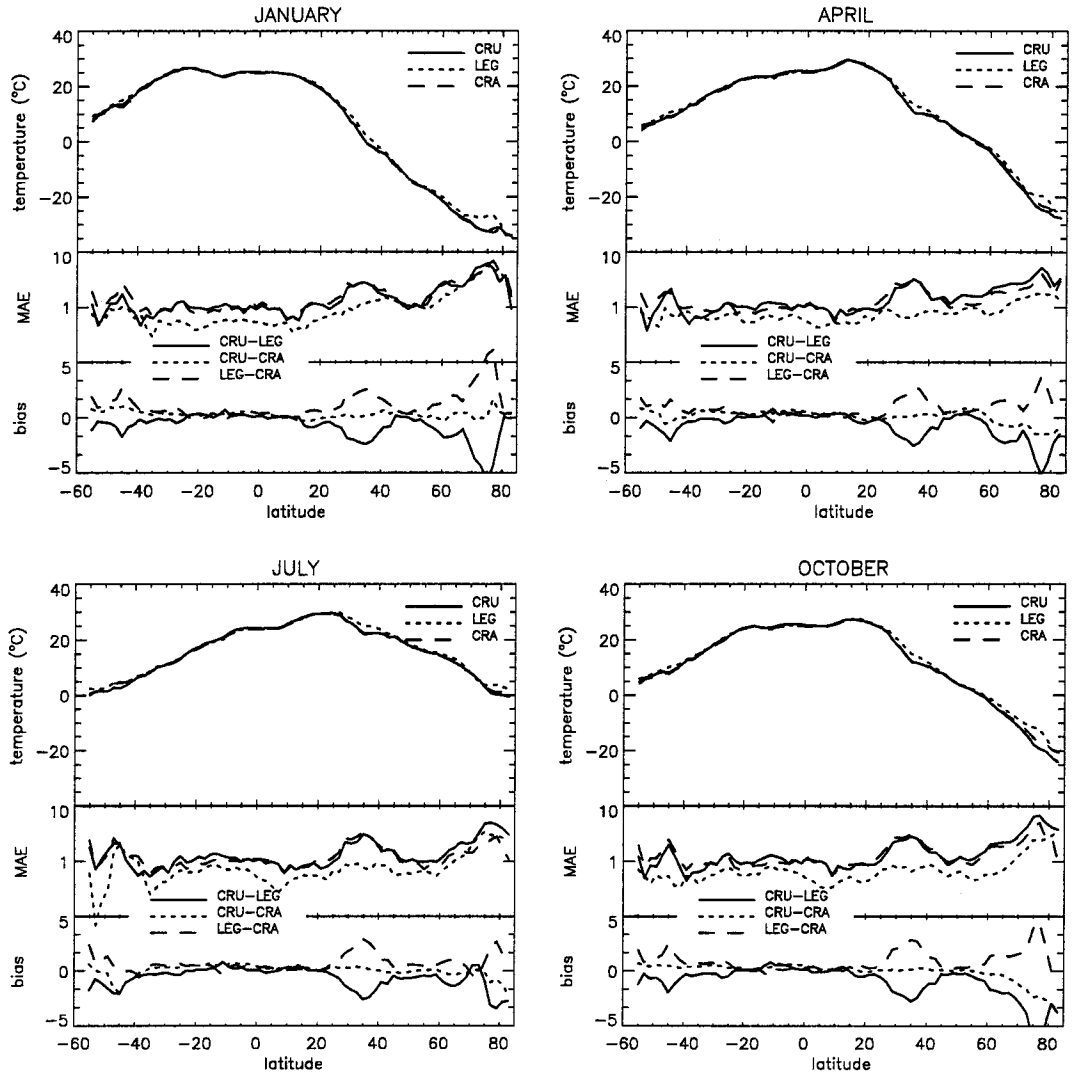

FIG. 21. Zonal-mean temperature, MAE, and mean bias between the CRU, LEG, and CRA mean temperature climatologies for Jan, Apr, Jul, and Oct.

fined as the average of maximum and minimum temperature, as in CRU, will be several degrees less than the mean calculated from 3-h or 6-h measurements, as is possibly the case in LEG and CRA over Russia. These differences have likely been exaggerated along the Arctic coast of Russia where there were no stations, and gradients interpolated from stations further to the south may well be different in the three climatologies.

4) DIURNAL TEMPERATURE RANGE

The only other global-mean monthly climatology of diurnal temperature range (or maximum and minimum temperature, from which diurnal temperature range can be derived) we are aware of is that of CRA (Version 2.1; W. Cramer 1997, personal communication). The CRA climatology represents the period 1931–60 and was derived in a different way to CRU: maximum and minimum temperature were interpolated separately (using thin-plate splines) and subsequently differenced to derive a diurnal temperature range climatology. Differences between CRU and CRA, for January, April, July, and October, are displayed in Fig. 23.

Over about 50% of the globe, CRU and CRA agree to within 1°C, which is similar to the RTGCV reported during interpolation of the CRU data over most tiles (Fig. 16). Over much of central Asia, CRU and CRA differ in a sporadic manner, but there is an overall pattern of CRU diurnal temperature range being lower than CRA. Over central Greenland, however, CRU diurnal temperature range is greater than CRA, except in July. Elsewhere, CRU is consistently different to CRA in several domains.

- CRU diurnal temperature range is greater than CRA

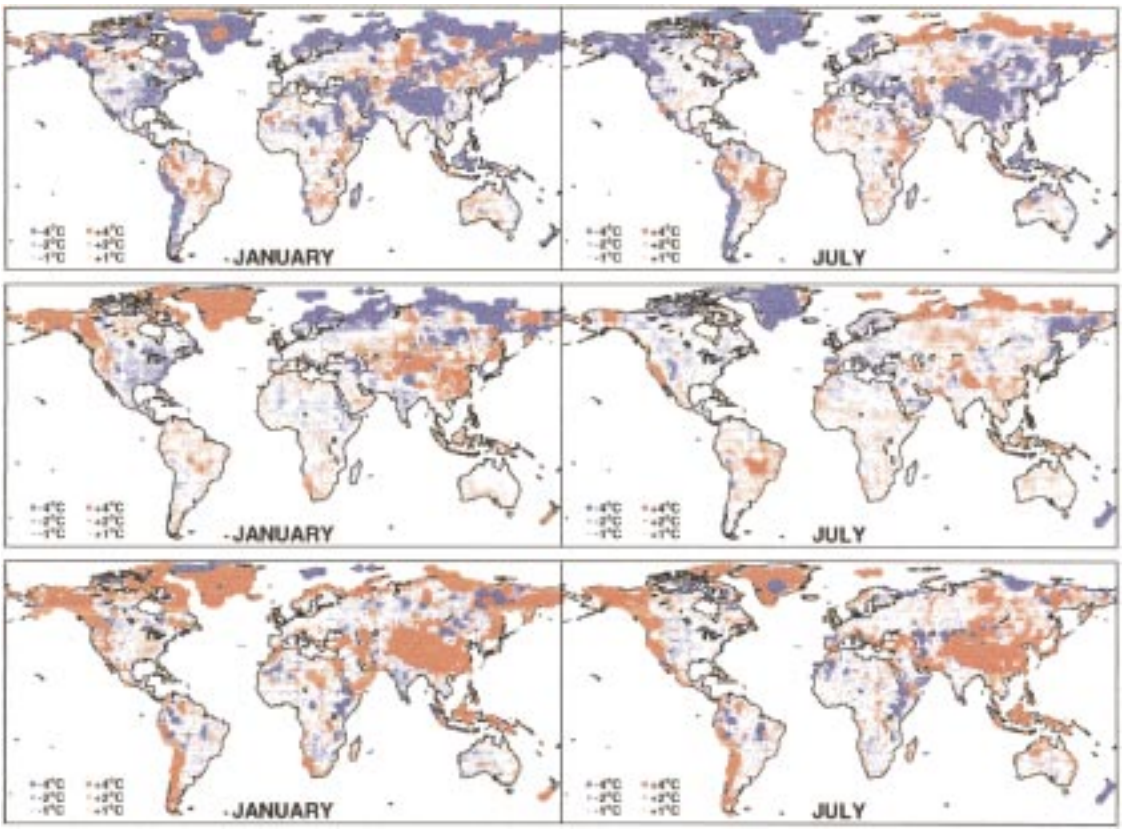

FIG. 22. Differences between CRU, LEG, and CRA (degraded to 2° lat × 2° long) mean temperature grids for January and July. The size of the dots (continuous scale) indicates the magnitude of the differences as shown in the legend. Top CRU minus LEG, middle CRU minus CRA, and bottom LEG minus CRA.

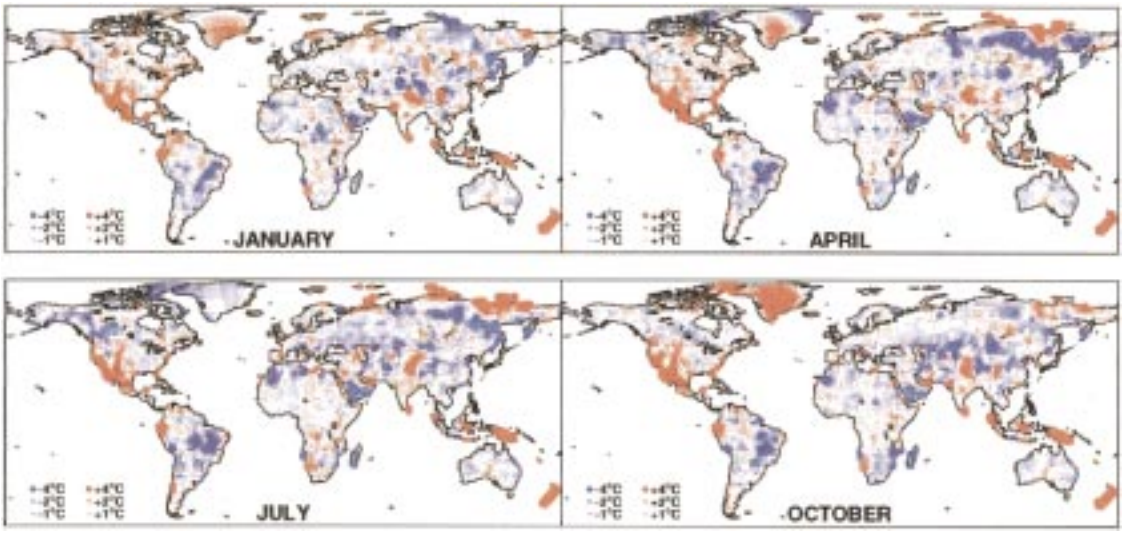

FIG. 23. Difference between CRU and CRA diurnal temperature range climatologies in peak season months. The size of the dots (continuous scale) shows the magnitude of the differences as indicated in the legend.

over Mexico and the southwest United States, Sri Lanka, New Guinea-Irian Jaya, New Zealand, and the Tibetan plateau, by several degrees Celsius in all seasons, and over Arctic Russia in spring and summer.
- CRU diurnal temperature range is less than CRA by more than 1°C over the Amazon Basin and Arabia throughout the year, and over parts of Canada and Alaska in the Northern Hemisphere summer.

Some of the differences described above can be explained by the fact that they occur in areas with sparse observing networks, where gradients arising from the two interpolation approaches are most likely to disagree. Examples here include Greenland, Amazonia, the Tibetan plateau, and New Guinea. Of more concern are the large differences over Mexico and the southwest United States and New Zealand, where the CRU dataset has a dense station network and the differences are up to 6°C. The differences in the USA are not due to changes in mean diurnal temperature range between the two normal periods, as Karl et al. (1993, their Figs. 2 and 4) show that diurnal temperature range in 1961–90 was about 1°C lower than in 1931–60. Similarly, Easterling et al. (1997) show a decreasing trend in diurnal temperature range of 1°–3°C (100 yr)$^{-1}$ (based on 1951–90 data) over New Zealand. Moreover, the contrasts are not due to the different interpolation methodologies, as we reinterpolated the CRU data to produce maximum and minimum temperature fields. These were then differenced to derive a new diurnal temperature range field (i.e., same approach as CRA) that exhibited similar differences to CRA. Closer examination of the patterns of differences indicate that they are greatest in, though by no means restricted to, mountainous regions (i.e., the Rocky Mountains, the Sierra Nevada, the Sierra Madre, and Baja California). A likely explanation for the differences is that the CRA station network inadequately captures topographic controls of diurnal temperature range (or maximum and minimum temperature) in these regions. However, at the time of writing no information on the CRA network was available from the authors, so this could not be addressed further.

5) CLOUD

We compared the CRU cloud cover climatology with two other frequently cited mean monthly climatologies:

- the 1984–91 ISCCP (Rossow and Schiffer 1991; ISCCP 1997) C1 mean monthly cloud fraction climatology; and
- the 1982–91 land-only cloud climatology of Hahn et al. (1994; henceforth Hahn).

Both these datasets have a resolution of 2.5° lat × 2.5° long. Therefore, the CRU cloud climatology was degraded to the same resolution by taking the average of all 0.5° grid points falling within the larger 2.5° cells. If fewer than five 0.5° values were present, then a 2.5° cell was assigned missing value (i.e., considered to be ocean).

The differences between the three climatologies are summarized in Fig. 24, which shows zonal-mean cloud fractions, MAE, and mean biases between the three climatologies for January, April, July, and October. In general, CRU zonal means tend to fall between ISCCP and Hahn, but agree more with Hahn, particularly in northern midlatitudes where the ground-based observing network is most extensive. MAEs are predominantly in the range 5%–15%, similar to results reported elsewhere (Rossow et al. 1993; Mokhov and Schlesinger 1994). The underestimation of ISCCP cloud amount in polar regions has also been well documented (Rossow et al. 1993; Mokhov and Schlesinger 1994).

CRU differs from Hahn most noticeably between 10° and 20°N in October and January, where Hahn exceeds both CRU and ISCCP by up to 28%. Examination of the geographical distribution of cloud amounts (not shown), revealed that the differences arise because Hahn cloud amounts are larger than CRU and ISCCP over the Indian/Southeast Asian monsoon region (by up to 40%), and over much of the Sahel. Some of the differences over the Sahel may be explained by the period of extended drought that occurred in the 1970s and 1980s, relative to earlier decades (Hulme 1992b; Hulme and New 1997) that would reduce cloud cover in Hahn more than CRU (which samples 1961–90 and, at some stations, 1931–60). A second possible cause of the differences over the Sahel could be that most of the CRU station data over this region are derived from sunshine, which introduces considerable uncertainty to the station values (see section 2 and Fig. 8). CRU and ISCCP agree quite well over the Sahel, however, which suggests that the CRU data are not inordinately biased. Over the Indian/Southeast Asian monsoon region, ISCCP and CRU agree well, although CRU tends to be 5%–10% more cloudy than ISCCP. Over India, CRU cloud cover normals are once again derived from sunshine, which may help explain differences between CRU and ISCCP, but is unlikely to explain the 25% differences between CRU and Hahn.

6) SUNSHINE

We compared the CRU and CRA sunshine climatologies. Although CRU and CRA use a similar number of stations, there is no information on the spatial distribution of the CRA stations, or the number that were derived from cloud cover.

Differences between CRU and CRA are summarized for the four midseason months in Fig. 25. The two climatologies agree well in mid and low latitudes, but diverge at high latitudes, particularly in the Northern Hemisphere, where stations are sparsely distributed. Examination of the spatial patterns of these differences (not shown) reveal that the differences at southern high latitudes arise primarily because CRU sunshine is higher

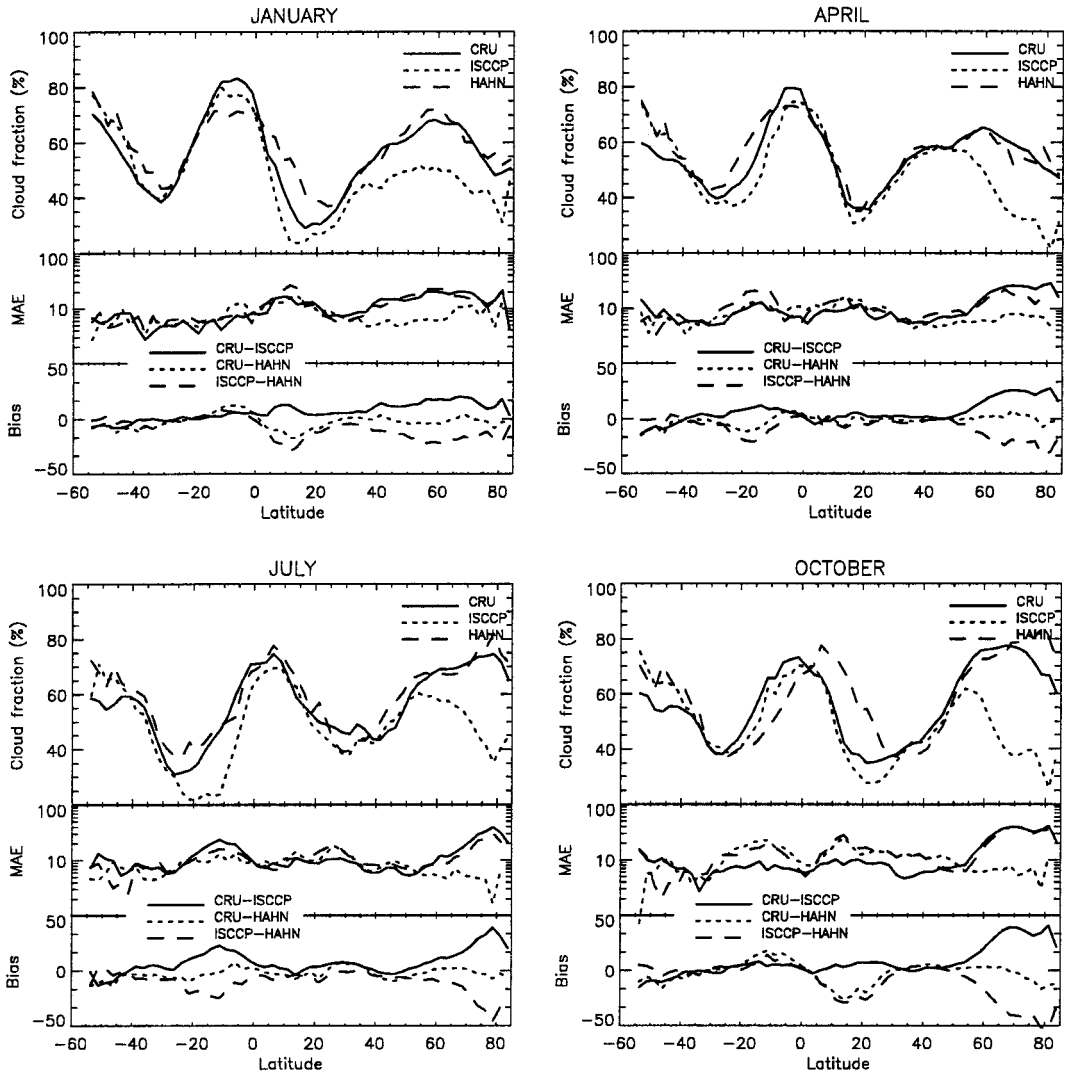

FIG. 24. Zonal-mean cloud cover, MAE, and bias between the CRU, ISCCP, and Hahn cloud climatologies in each of the peak-season months.

over the Pacific coasts and adjacent interiors of South America and Australia. This may be due to the sparse observing network (only a few stations in the CRU dataset) or differences arising from the estimation of sunshine from cloud data at high latitudes. It is difficult to assess the relative merits of these arguments because we have no information on the CRA station network.

At northern high latitudes the differences occur because CRU exceeds CRA over Russia and Greenland in spring, summer, and autumn, and vice versa in winter. In winter, CRU shows the expected strong latitudinal control on sunshine (due to zero hours at very high latitudes, and the effect of low sun angles on the amount of bright sunshine recorded), which CRA does not. Leemans and Cramer (1991) acknowledge that the 1991 version of CRA is limited by poor station coverage at high latitudes, and it is likely that the latest version suffers from similar shortcomings. The opposite trend over Russia in summer is difficult to explain. It is possible that CRA uses cloud cover measurements, converted to sunshine, and that their conversion to sunshine has resulted in systematic biases at high latitudes in summer. Without further documentation of CRA the causes of these differences cannot be resolved.

5. Conclusions

We have constructed a new high-resolution ($0.5°\times0.5°$) global-mean monthly climatology from surface observations. The dataset describes the spatial characteristics over land areas (excluding Antarctica) of a suite of nine variables: precipitation, wet-day frequency, mean temperature, diurnal temperature range, vapor pressure, sunshine, cloud cover, ground frost frequency,

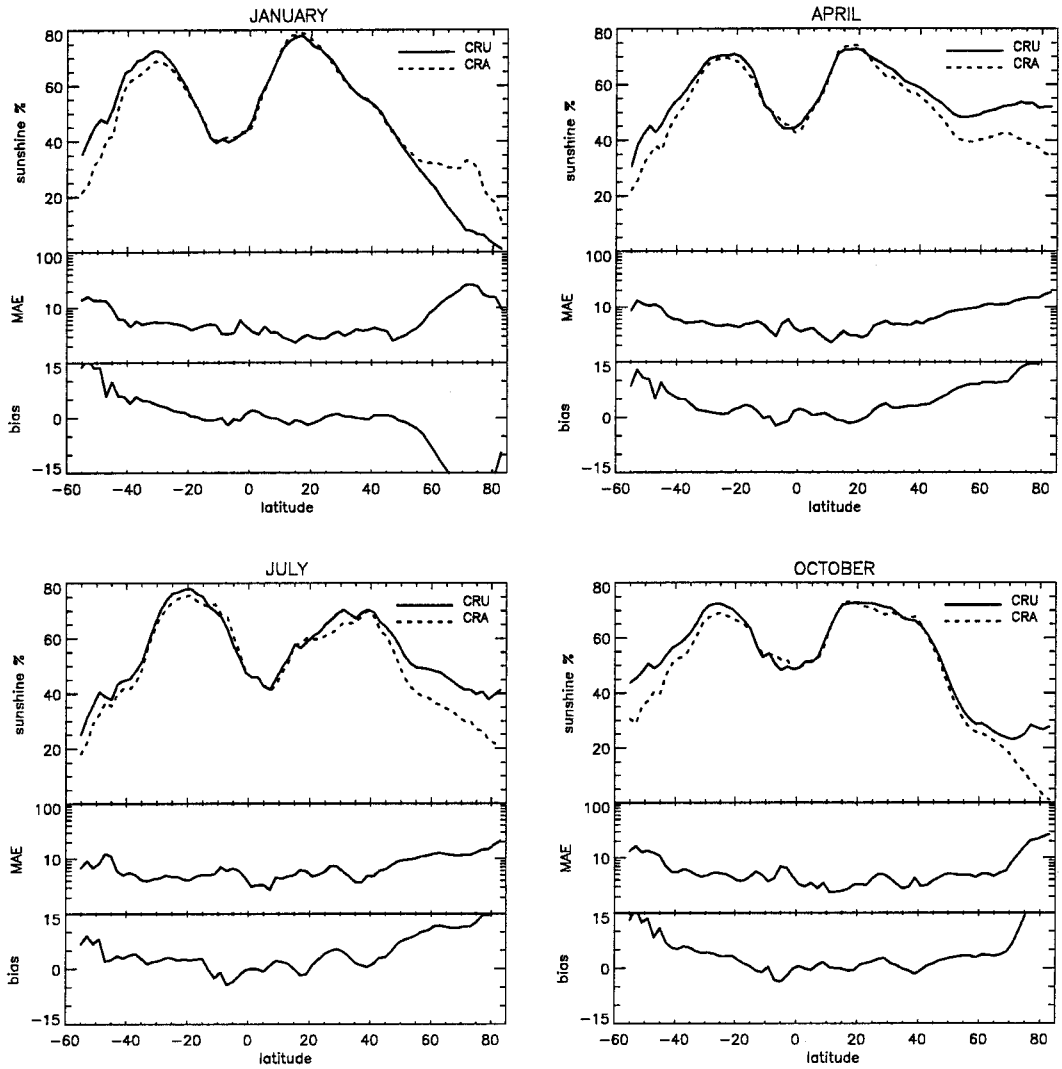

FIG. 25. Zonal-mean sunshine, MAE, and bias for the CRU and CRA sunshine climatologies.

and wind speed. The climate grids are derived from as extensive a database of time-dependent normals that exists anywhere.

The accuracy of the monthly grids varies both regionally and between variable. In general, cross-validation results indicate that mean temperature, diurnal temperature range, and vapor pressure are the variables most reliably interpolated from the available station data, followed, in decreasing order by sunshine, cloud cover, ground frost frequency, wet-day frequency, precipitation, and wind speed. For all variables, the surfaces are least reliable in areas with poor data coverage and/or complex topography. This is particularly so for the Greenland Ice Sheet, where only one or two short-term normals were available. Other regions with large cross-validation errors include the American cordillera, the Himalaya–Tibetan region, and Southeast Asia during the monsoon season. Some of the cross-validation errors for the surfaces can be ascribed to uncertainties inherent in the conversion of variables to a common base unit (vapor pressure, sunshine, cloud cover, wet-day frequency, and ground frost frequency), errors in station location, or data values that were not identified during the interpolation. However, most of the error is likely to arise from inadequate station networks.

We compared our climate fields to several other commonly used climatologies, namely Legates and Willmott (LEG), Leemans and Cramer (CRA, Version 2.1), as well as the ISCCP and Hahn cloud climatologies. The CRU precipitation and mean temperature fields show large systematic differences to LEG in high-elevation regions. The main reason for these is that LEG does not include elevation as a predictor variable in their interpolation and we suggest that the CRU grids are more accurate in such regions. The two climatologies also disagree markedly in high latitudes. Despite this, CRU

precipitation is more similar to LEG than CRA, which may well be due to the fact that the CRU and LEG normal periods overlap (1961–90 and 1921–80, respectively), whereas CRU and CRA (nominally) sample entirely different periods.

The CRU diurnal temperature range fields show some systematic differences to CRA that cannot be attributed to differences in interpolation methodology, most notably over the southwestern United States, Mexico, Indonesia, and New Zealand, where errors are of the order of 3°–6°C. These regions exhibited the largest cross-validation errors for this variable during the interpolation of the CRU surfaces. This suggests that different station networks in these regions may well be the cause of the relatively large differences between CRU and CRA. Contrasts between CRU and CRA for sunshine and wet-day frequency are of similar magnitude to, but slightly larger than, the cross-validation errors reported during the interpolation of the CRU surfaces, suggesting that these fields agree well given the available station networks. Largest errors occur where the network is sparsest, namely, in the Tropics and at high latitudes.

The CRU cloud cover fields agree best with the observed climatology of Hahn, but generally falls between ISCCP and Hahn. Zonal-mean errors are around 10%, similar to the cross-validation errors from the CRU interpolation.

Thus the picture that emerges from cross validation and intercomparison is that there are nonnegligible prediction errors associated with the CRU climate fields and, more generally, in all observed climatologies of this type. In specific instances, the differences between climatologies can be attributed to either contrasting interpolation methodologies or temporal sampling strategies. However, much of the difference between fields is due to interpolation from different station networks, and it is probably impossible to provide an absolute answer as to which climatology, if any, is superior. It may well be that for specific applications in certain regions, one climatology is more appropriate than another. A more robust approach would be to use more than one climate field that describes the same variable, using the spread of results to arrive at a qualitative assessment of the uncertainty associated with the observed climate data (e.g., Arnell 1995; Pan et al. 1996).

Nonetheless, several features of our new climatology suggest that it should be widely applicable in climate and related activities. These include the following:

- it is derived from the largest possible collection of station normals that are constrained to a specific time period, namely, 1961–1990;
- it encompasses a suite of nine surface variables, several of which have not previously been interpolated at the global scale;
- we have used a consistent interpolation methodology that includes specific treatment of elevation as a predictor variable; and
- we provide regional error estimates and comparisons with other climatologies so that potential users are aware of the uncertainties associated with individual surfaces.

This paper describes the first step in an attempt to construct a high-resolution terrestrial climate dataset that describes month-by-month space–time variability of surface climate from 1901 to the present day. The normal climatology described here provides the background fields onto which monthly anomaly fields can be superimposed to arrive at monthly climate fields. A subsequent paper (New et al. 1999) will describe construction of the monthly fields and provide an assessment of their accuracy via comparison with other, more limited, gridded time series.

The data are available from the Climatic Research Unit and enquiries should be directed to the Climate Impacts LINK Project, Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ (d.viner@uea.ac.uk or http://www.cru.uea.ac.uk/~link). Full color plots of each monthly field can be accessed via the World Wide Web at http://ipcc-ddc.cru.uea.ac.uk.

Acknowledgments. This work was undertaken with funding for MN from the U.K. Natural Environment Research Council (Grant No. GR3/09721) and MH from U.K. Department of the Environment, Transport and the Regions (DETR; Contract EPG 1/1/48). Data supplied by numerous national meteorological agencies, research organizations, and private individuals made this research possible and these contributions, while too numerous to mention by name, are gratefully acknowledged. The CRU precipitation, mean temperature, and diurnal temperature range datasets have been compiled over a number of years with support from the U.K. DETR and the U.S. Department of Energy. Mike Hutchinson kindly supplied the interpolation software. Wolfgang Cramer and David Legates generously provided the CRA and LEG datasets, respectively. Louise Bohn, David Lister, Nick Brooks, Chris Turney, Martin Airey, and Ian Harris all spent many hours on data entry and quality control. The Climate Impacts LINK Project (U.K DETR. Contract EPG 1/1/16) at the Climatic Research Unit provided computing facilities for this research. The comments of two anonymous reviewers resulted in improvements in the final version of the manuscript.

REFERENCES

Airey, M. J., M. Hulme, and T. C. Johns, 1996: Evaluation of simulations of terrestrial precipitation in UK Met. Office/Hadley Centre climate change experiments. *Geophys. Res. Lett.,* **23,** 1657–1660.

Andersson, T., and I. Mattison, 1991: A field test of thermometer screens. SMHI/RMK 62, Swedish Meteorological and Hydrological Institute, 43 pp. [Available from Swedish Meteorological and Hydrological Institute, Oceanographical Laboratory, P.O. Box 2212, S-403 14 Goteborg, Sweden.]

Arnell, N. W., 1995: Grid mapping of river discharge. *J. Hydrol.*, **167,** 39–56.
Briggs, P. R., and J. G. Cogley, 1996: Topographic bias in mesoscale precipitation networks. *J. Climate,* **9,** 205–218.
Cao, M., and F. I. Woodward, 1998: Net primary and ecosystem production and carbon stocks of terrestrial ecosystems and their responses to climate change. *Global Change Biol.,* **4,** 185–198.
Carter, T. R., M. L. Parry, H. Harasawa, and S. Nishioka, 1994: IPCC technical guidelines for assessing climate change impacts and adaptations. 59 pp. [Available from Center for Global Environmental Research, National Institute for Environmental Studies, 16-2 Onogawa, Tsukaba, Ibaraki 305, Japan.]
Chen, Q. S., D. H. Bromwich, and L. S. Bai, 1997: Precipitation over Greenland retrieved by a dynamic method and its relation to cyclonic activity. *J. Climate,* **10,** 839–870.
Christensen, J. H., B. Machenhauer, R. G. Jones, C. Schar, P. M. Ruti, M. Castro, and G. Visconti, 1997: Validation of present-day regional climate simulations over Europe: LAM simulations with observed boundary conditions. *Climate Dyn.,* **13,** 489–506.
CIAT, cited 1997: Climate data base. [Available online at http://gis.ciat.cgiar.org/joseh/climate/climate2.html.]
Cramer, W., and A. Fischer, 1996: Data requirements for global terrestrial ecosystem modelling. *Global Change and Terrestrial Ecosystems,* B. Walker and W. Steffen, Eds., Cambridge University Press, 530–565.
Dai, A., I. Fung, and D. G. Genio, 1997: Surface observed global land precipitation variations during 1900–1988. *J. Climate,* **10,** 2943–2962.
Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.,* **33,** 140–158.
Doorenbos, J., and W. O. Pruitt, 1984: Guidelines for predicting crop water requirements. Rep. 24, Food and Agriculture Organisation, Rome, 143 pp. [Available from FAO, Viale delle Terme di Caracalla, 00100 Rome, Italy.]
Easterling, D. R., and Coauthors, 1997: Maximum and minimum temperature trends for the globe. *Science,* **277,** 364–367.
Eischeid, J. K., H. F. Diaz, R. S. Bradley, and P. D. Jones, 1991: A comprehensive precipitation data set for global land areas. DOE/ER-6901T-H1, U.S. Department of Energy, Washington, DC, 81 pp. [Available from U.S. Department of Energy, Office of Energy Research, Washington, DC 20585.]
FAO, 1984: *Agroclimatological Data for Africa.* Food and Agriculture Organisation of the United Nations, 723 pp.
Groisman, P. Y., V. V. Koknaeva, T. A. Belokrylova, and T. R. Karl, 1991: Overcoming biases of precipitation measurement: A history of the USSR experience. *Bull. Amer. Meteor. Soc.,* **72,** 1725–1733.
Hahn, C. J., S. G. Warren, and J. London, 1994: Climatological data for clouds over the globe from surface observations, 1982–1991: The total cloud edition. NDP026A, Carbon Dioxide Analysis Center, Oak Ridge National Laboratory, Oak Ridge, 39 pp. [Available from Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6335.]
Huffman, J. H., R. F. Adler, B. Rudolf, B. F. Schneider, and P. R. Keehn, 1995: Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information. *J. Climate,* **8,** 1284–1295.
Hulme, M., 1992a: A 1951–80 global land precipitation climatology for the evaluation of General Circulation Models. *Climate Dyn.*, **7,** 57–72.
——, 1992b: Rainfall changes in Africa—1931–1960 to 1961–1990. *Int. J. Climatol.,* **12,** 685–699.
——, 1994a: Validation of large-scale precipitation fields in General Circulation Models. *Global Precipitations and Climate Change,* M. Desbois and F. Desalmand, Eds., Springer-Verlag, 387–405.
——, 1994b: The cost of climate data—A European experience. *Weather,* **49,** 168–175.
——, and M. G. New, 1997: Dependence of large-scale precipitation climatologies on temporal and spatial sampling. *J. Climate,* **10,** 1099–1113.
——, and O. Brown, 1998: How likely is tolerable regional climate change? *Climate Res.,* **10,** 1–14.
——, D. Conway, P. D. Jones, T. Jiang, E. M. Barrow, and C. Turney, 1995: Construction of a 1961–1990 European climatology for climate change modelling and impact applications. *Int. J. Climatol.,* **15,** 1333–1363.
——, ——, A. Joyce, and H. Mulenga, 1996: A 1961–90 climatology for Africa south of the equator and a comparison of PET estimates. *S. Afr. J. Sci.,* **92,** 334–343.
Hutchinson, M. F., 1995: Interpolating mean rainfall using thin plate smoothing splines. *Int. J. Geogr. Inf. Sys.,* **9,** 385–403.
——, 1998: Interpolating mean rainfall using thin plate smoothing splines. Part II: Analysis of topographic dependence. *J. Geogr. Inf. Decis. Anal.,* in press.
——, and P. E. Gessler, 1994: Splines—More than just a smooth interpolator. *Geoderma,* **62,** 45–67.
IPCC, 1996: *Climate Change 1996: The Science of Climate Change.* Cambridge University Press, 453 pp.
ISCCP, cited 1997: International Satellite Cloud Climatology Project. [Available online at http://isccp.giss.nasa.gov/isccp.html.]
Jones, P. D., 1994: Hemispheric surface air temperature variability—A reanalysis and update to 1993. *J. Climate,* **7,** 1794–1802.
——, and Coauthors, 1985: A grid point surface air temperature data set for the Northern Hemisphere. DOE/EV/10098-2, U.S. Department of Energy, Washington, DC, 251 pp. [Available from U.S. Department of Energy, Office of Energy Research, Washington, DC 20585.]
——, S. C. B. Raper, R. Bradley, H. F. Diaz, P. M. Kelly, and T. M. L. Wigley, 1986a: Northern Hemisphere surface air temperature variations: 1851–1984. *J. Climate Appl. Meteor.*, **25,** 161–79.
——, ——, C. M. Goodess, B. S. G. Cherry, and T. M. L. Wigley, 1986b: A grid point surface air temperature data set for the Southern Hemisphere. DOE/EV/10098-6, U.S. Department of Energy, Washington, DC, 73 pp. [Available from U.S. Department of Energy, Office of Energy Research, Washington, DC 20585.]
Karl, T. R., C. N. Williams, and P. J. Young, 1986: A model to estimate the time of observation bias associated with monthly mean maximum, minimum and mean temperatures for the United States. *J. Climate Appl. Meteor.*, **25,** 145–160.
——, and Coauthors, 1993: A new perspective on recent global warming—Asymmetric trends of daily maximum and minimum temperature. *Bull. Amer. Meteor. Soc.*, **74,** 1007–1023.
Kittel, T. G. F., N. A. Rosenbloom, T. H. Painter, D. S. Schimel, and VEMAP Modelling Participants, 1995: The VEMAP integrated database for modelling United States ecosystem/vegetation sensitivity to climate change. *J. Biogeog.,* **22,** 857–862.
Leemans, R., and W. Cramer, 1991: The IIASA database for mean monthly values of temperature, precipitation and cloudiness on a global terrestrial grid. RR-91-18, 63 pp. [Available from IIASA, A-2361 Laxenburg, Austria.]
Legates, D. R., 1987: A climatology of global precipitation XL(1), 85 pp. [Available from Department of Geography, University of Delaware, Newark, DE 19716.]
——, 1997: ''Global and terrestrial precipitation: A comparative assessment of existing climatologies:'' A reply. *Int. J. Climatol.*, **17,** 779–783.
——, and C. J. Willmott, 1990a: Mean seasonal and spatial variability in gauge-corrected, global precipitation. *Int. J. Climatol.,* **10,** 111–127.
——, and ——, 1990b: Mean seasonal and spatial variability in global surface air-temperature. *Theor. Appl. Climatol.,* **41,** 11–21.
——, and T. L. DeLiberty, 1993: Precipitation measurement biases in the United States. *Water Resour. Bull.,* **29,** 855–861.
Mokhov, I. I., and M. E. Schlesinger, 1994: Analysis of global cloudiness. Part 2: Comparison of ground-based and satellite-based cloud climatologies. *J. Geophys. Res.*, **99** (D), 17 045–17 065.

Müller, M. J., 1982: *Selected Climatic Data for a Global Set of Standard Stations for Vegetation Science.* W. Junk, 306 pp.

NCDC, 1997: Data documentation for data set World Meteorological Organisation 1961–1990 global standard normals. Rep. TD-9641, 34 pp. [Available from National Climate Data Center, Federal Building, 151 Patton Avenue, Asheville, NC 28801-5001.]

New, M. G., M. Hulme, and P. D. Jones, 1999: Representing twentieth-century space–time climate variability. Part II: Development of 1901–1996 monthly terrestrial climate fields. *J. Climate,* in press.

NGDC, cited 1996: Terrainbase digital terrain data. [Available online at http://www.ngdc.noaa.gov/seg/topo/topo.shtml.]

Pan, Y. A., A. D. McGuire, D. W. Kicklighter, and J. D. Melillo, 1996: The importance of climate and soils for estimates of net primary production: A sensitivity analysis with the terrestrial ecosystem model. *Global Change Biol.,* **2,** 5–23.

Phillips, D. L., J. Dolph, and D. Marks, 1992: A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. *Agric. For. Meteor.,* **58,** 119–141.

Piper, S. C., and E. F. Stewart, 1996: A gridded global data set of daily temperature and precipitation for terrestrial biosphere modeling. *Glob. Biogeochem. Cycles,* **10,** 757–782.

Rossow, W. B., and R. A. Schiffer, 1991: ISCCP cloud data products. *Bull. Amer. Meteor. Soc.,* **72,** 2–20.

——, A. W. Walker, and L. C. Garder, 1993: Comparison of ISCCP and other cloud amounts. *J. Climate,* **6,** 2394–2418.

Sellers, P. J., and Coauthors, 1997: Modelling the exchanges of energy, water, and carbon between the continents and the atmosphere. *Science,* **275,** 502–509.

Shepard, D., 1968: A two-dimensional interpolation function for irregularly spaced data. *Twenty-Third ACM National Conference,* Brabdon Systems Press, 517–524.

Shuttleworth, J. W., 1992: Evaporation. *Handbook of Hydrology,* D. R. Maidment, Ed., McGraw-Hill, 4.1–4.53.

Thiessen, A. H., 1911: Precipitation averages for large areas. *Mon. Wea. Rev.,* **39,** 1082–1084.

UKMO, 1969: *Observers Handbook.* U.K. Meteorological Office, 242 pp.

USAF, 1987: *Station Climatic Summaries: Europe.* United States Air Force, 376 pp.

Wahba, G., 1979: How to smooth curves and surfaces with splines and cross-validation. *24th Conf. on the Design of Experiments,* U.S. Army Research Office, 167–192. [Available from U.S. Army Research Laboratory, Attn: AMSRL-CS-EA, 2800 Powder Mill Rd., Adelphi, MD 20783-1197.]

——, 1990: *Spline Models for Observational Data.* Society for Industrial and Applied Mathematics, 169 pp.

Wigley, T. M. L., and B. D. Santer, 1990: Statistical comparison of spatial fields in model validation, perturbation, and predictability experiments. *J. Geophys. Res.,* **95,** 851–865.

Willmott, C. J., and K. Matsuura, 1995: Smart interpolation of annually averaged air temperature in the United States. *J. Appl. Meteor.,* **34,** 2577–2586.

——, C. M. Rowe, and W. D. Philpot, 1985: Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid point interpolation and contouring. *Amer. Cartogr.,* **12,** 5–16.

——, S. M. Robeson, and M. J. Janis, 1996: Comparison of approaches for estimating time-averaged precipitation using data from the United States. *Int. J. Climatol.,* **16,** 1103–1115.

WMO, 1996: Climatological normals (CLINO) for the period 1961–1990. WMO/OMM 847, 768 pp. [Available from World Meteorological Organisation, 41, avenue Giuseppe-Motta, CP-2300-1211 Geneva 2, Switzerland.]

Xie, P., and P. A. Arkin, 1996: Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Climate,* **9,** 840–858.

——, B. Rudolf, U. Schneider, and P. A. Arkin, 1996: Gauge-based monthly analysis of global land precipitation from 1971 to 1994. *J. Geophys. Res.,* **101,** 19 023–19 034.