1. TITLE

1.1 Data Set Identification

ISLSCP II Climate Research Unit CRU05 Monthly Climate Data

1.2 Database Table Name(s)

Not applicable to this data set.

1.3 File Name(s)

This data set contains a mean monthly *climatology* for several climate variables averaged over the period from 1961 to 1990 created by the Climatic Research Unit (CRU) at the University of East Anglia, U.K.

There are 23 data files with this data set which includes 1 changemap ASCII file and 22 *.zip data files named using the following convention:

cru5_*variable_Xdeg*_1961-*90.zip*

where:

cru5	Identifies these data as CRU version 5.
variable	This is the climate variable name. **See Table 1 below for a list of the variables and abbreviations used in the file names**.
Xdeg	This identifies the spatial resolution of the data: Can be "1d" for 1-degree or "hd" for half-degree resolutions in both latitude and longitude.
1961-90	This identifies the data as monthly averages for the entire period from 1961-1990. MM is the month from 01 to 12.

cru5_*mean_hd*_changemap.*asc:* Gridded ASCII maps showing the results of applying the 0.5 degree ISLSCP II land/water mask: all points added ("1"), all points unchanged ("0"), and all

points removed ("-1"). These files are only available for the original 0.5 degree data set because the 1.0 degree version was created by the ISLSCP II staff by averaging the 0.5 degree data (see Section 9.2.3).

All points removed from the original half-degree files are stored in "differences" files (with the extension ".dif"). These are ASCII tables of "differences", points that didn't match the ISLSCP Initiative II Land/Water mask, and were removed from or added to the ASCII map files (see Sections 8.4 and 9.2.3 for more details). These files are only available for the original 0.5 degree data set because the 1.0 degree version was created by the ISLSCP II staff by averaging the 0.5 degree data (see Section 9.2.3).

Table 1. List of climate variables included in this data set. Abbreviations for the variables as used in the filenames are also given. See Section 8.2 for units of measurement.

Climate Variables

Abbreviation	Variable
cloud_covr	Cloud cover
diurnl_tmp	Diurnal temperature range
frost_freq	Ground frost frequency
precip	Precipitation
radiation	Radiation
temp_max	Maximum temperature
temp_mean	Mean temperature
temp_min	Minimum temperature
vapor_pres	Vapor pressure
wetday_ frq	Wet day frequency
wind	Wind

1.4 Revision Date of this Document
June 13, 2011

2. INVESTIGATOR(S)

2.1 Investigator(s) Name and Title
Dr. Mark New
School of Geography & the Environment
University of Oxford, United Kingdom

Prof. Phil Jones
Climatic Research Unit
School of Environmental Sciences
University of East Anglia, United Kingdom

Dr. Mike Hulme
Tyndall Centre for Climate Change Research
School of Environmental Sciences
University of East Anglia, United Kingdom

2.2 Title of Investigation

CRU05 Monthly Climatology (1961-1990) and Monthly Climate Time Series (1901-1995/6) for Global Land Areas [part of the project, "The missing carbon sink", funded by the UK Natural Environment Research Council].

2.3 Contacts (For Data Production Information)

	Contact 1	**Contact 2**
2.3.1 Name	Dr. Mark New	Dr. Mike Hulme
2.3.2 Address	School of Geography & the Environment University of Oxford Mansfield Road	Tyndall Centre for Climate Change Research School of Environmental Sciences University of East Anglia
City/St.	Oxford	Norwich
Zip Code	OX1 3TB	NR4 7TJ
Country	United Kingdom	United Kingdom
2.3.3 Tel. No.	44-1865-271919	44-1603-593900
Fax No.	44-1865-271929	44-1603-593901
2.3.4 E-mail	mark.new@geog.ox.ac.uk	m.hulme@uea.ac.uk

	Contact 3
2.3.1 Name	Mr. David Landis
2.3.2 Address	Code 614.4 NASA/GSFC
City/St.	Greenbelt, MD
Zip Code	20771
Country	USA
2.3.3 Tel. No.	(301))614-6605
Fax No.	(301) 286-0239
2.3.4 E-mail	david.r.landis@gsfc.nasa.gov

2.4 Data Set Citation

New, M., P. Jones, and M. Hulme. 2011. ISLSCP II Climate Research Unit CRU05 Monthly Climate Data. In Hall, Forrest G., G. Collatz, B. Meeson, S. Los, E. Brown de Colstoun, and D. Landis (eds.). ISLSCP Initiative II Collection. Data set. Available on-line [http://daac.ornl.gov/] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. doi:10.3334/ORNLDAAC/1015

2.5 Requested Form of Acknowledgment

Users of the International Satellite Land Surface Climatology (ISLSCP) Initiative II data collection are requested to cite the collection as a whole (Hall et al. 2006) as well as the individual data sets. Please cite the following publications when these data are used:

Hall, F.G., E. Brown de Colstoun, G. J. Collatz, D. Landis, P. Dirmeyer, A. Betts, G. Huffman, L. Bounoua, and B. Meeson, The ISLSCP Initiative II Global Data sets: Surface Boundary Conditions and Atmospheric Forcings for Land-Atmosphere Studies, *J. Geophys. Res.*, 111, doi:10.1029/2006JD007366, 2006.

Acknowledgements to the Climatic Research Unit, University of East Anglia for provision of the data. The following two papers should be cited when referring to the CRU05 climate data:

New, M., M. Hulme and P. Jones (1999). "Representing twentieth-century space-time climate variability. Part I: Development of a 1961-90 mean monthly terrestrial climatology." *Journal of Climate* **12**: 829-856.

New, M., M. Hulme and P. Jones (2000). "Representing twentieth-century space-time climate variability. Part II: Development of 1901-1996 monthly grids of terrestrial surface climate." *Journal of Climate* **13**(13): 2217-2238.

3. INTRODUCTION

3.1 Objective/Purpose

These data were originally developed by the Climatic Research Unit (CRU) for application to F. Ian Woodward's global biogeochemical model to simulate the terrestrial carbon cycle over the 20th century, under the auspices of a project funded by the UK Natural Environment Research Council. Since then the data have been used in numerous studies where *spatially complete* climate data are required at regional, continental and global scales. For example, these data have been used for gridded biogeochemical, biological, agricultural and hydrological modelling, either over the 20th century, or as a baseline climate data set for climate change impacts assessments.

3.2 Summary of Parameters

Table 1 in Section 1.3 provides a complete listing of the climate variables provided here (also see Section 8.2 for more in depth information). This data set includes global, gridded, 30-year monthly means for the World Meteorological Organization (WMO) normal period, 1961-1990 for land areas excluding Antarctica. The data are provided at two spatial resolutions of 0.5 and 1.0 degree in both latitude and longitude.

3.3 Discussion

3.3.1 Monthly Mean Climatology (1961-1990)

This climatology was constructed from a data-set of station 1961-1990 climatological normals, numbering between 19,800 (precipitation) and 3615 (windspeed; see New et al, 1999 for details). The station data were interpolated as a function of latitude, longitude and elevation using thin-plate splines. The accuracy of the interpolations were assessed using cross-validation and by comparison with other climatologies. The climatology represents an advance over earlier published global terrestrial climatologies in that it is strictly constrained to the period 1961-1990, 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 has been used by researchers in the areas of ecosystem modelling, climate model evaluation and climate change impact assessment.

The primary variables, precipitation, mean temperature and diurnal temperature range, were interpolated directly from station observations. The resulting time-series were compared with other, coarser resolution, data sets of similar temporal extent. The remaining climatic elements, termed secondary variables, were interpolated from merged data sets, comprising station observations and, in regions where there were no station data, synthetic data estimated using predictive relationships with the primary variables.

This data set represents an advance over earlier products because (i) it has higher spatial resolution than other data sets of similar temporal extent, (ii) it has longer temporal coverage than other products of similar spatial resolution; (iii) it encompasses a more extensive suite of surface climate variables than available elsewhere, (iv) it includes an elevation-dependence in the underlying climatology and (v) the construction method ensures that strict temporal fidelity is maintained.

The data have been produced by CRU and modified by the staff of the ISLSCP II data collection to match the land/water boundaries of the ISLSCP II land/water mask. Files are provided that show the differences between the original CRU data and the ISLSCP II land/water mask. “Differences” files are also provided containing any data point in the original data that has been changed (see Section 9.2.3). This was done to ensure land/water consistency between the many different data sets in this collection. Further, the ISLSCP II staff has created a 1.0 degree version of the data set by averaging the original 0.5 degree data. These 1.0 degree data are given to provide a complete ISLSCP Initiative II collection at this resolution but are recommended for browse use only. Users should refer to the original 0.5 degree data for more in depth analyses.

4. THEORY OF ALGORITHM/MEASUREMENTS

All data sets are derived through interpolation of surface climate meteorological station observations onto a 0.5 degree lat/long Earth grid. The 1.0 degree data are derived through averaging of the 0.5 degree data.

4.1 Monthly Mean Climatology (1961-1990)

The monthly grids were derived in two stages. (1) Interpolation of station monthly anomalies (relative to the 1961-1990 mean) using an angular distance weighting (ADW) scheme, and (2) combination of the anomalies with the mean monthly climatology.accomplished by determining the amount of data smoothing that is required to minimise the generalised cross validation (GCV).

4.1.1 Anomaly Interpolation

Primary Variables (precipitation, mean temp, diurnal temp range)

Prior to interpolation, each station time-series was converted to anomalies relative to the 1961-1990 mean. Series with less than 20 years of data during 1961-1990 were excluded from the analysis. Anomalies for mean temperature and diurnal temperature range were expressed in absolute units (i.e. degrees C), while precipitation was expressed as percentage of the 1961-1990 mean. We used percentage units for precipitation because the variance of precipitation is closely related to the mean, and interpolation in percentage units preserves this relationship better than interpolation in absolute units. The inverse-distance weighting scheme used to estimate grid-point values used anomalies from the eight nearest stations, regardless of direction or distance, resulting in a radius of influence that varies with station density. A consequence of this varying influence radius is that the grid roughness can increase in areas where the station density is high and the spatial variability of the anomalies is high (if the anomaly data have a lower spatial variability, the grid roughness will not be as great, even where the station density is high). Weights for the eight stations were determined in a two-stage process. All stations are first weighted by distance from the grid point, with an empirically derived correlation decay distance (CDD) controlling the rate at which the weight decreases with distance from the grid point. A correlation decay function can be defined (Jones *et al.*1997):

$$r = e^{-x/x_0}$$

where x_0 is the CDD and x is the distance from the grid point of interest. Thus for any grid point and any station, *k*, a distance weight can be defined by:

$$w_k = r^m$$

where a value of 4 for *m* was adopted, after experimentation showed that this produced the lowest cross validation error.

The second component of the distance weight was determined by the directional (angular) isolation of each of the *nj* data points selected:

$$a_k = \sum_{l=1}^{n_j} w_l \, [1 - \cos Ø_j \,(k, l)] \, / \sum_{l=1}^{n_j} w_l$$

where θ *j(k,l)* is the angular separation of data points *k* and *l* with the vertex of the angle defined at grid point *j*, calculated in spherical co-ordinates, and *wl* is the distance weight at data point *l*.

The angular and distance weights are then combined to arrive at an angular distance weight:

$$W_k = w_k \left[1 + a_k \right]$$

Interpolation as a function of latitude and longitude, as in ADW, ignores the influence of elevation. A large proportion of the spatial variation in monthly temperature anomalies is a function of large-scale circulation features, and is relatively independent of topography (New et al. 1999). Interpolation of mean temperature and diurnal temperature range as a function of only latitude and longitude is therefore adequate. This is not necessarily true for precipitation, where inclusion of elevation as a co-predictor has been shown to improve the accuracy of the anomaly interpolation in some situations (M. F. Hutchinson, personal communication, 1997). However, the ADW gridding employed in this study did not permit the inclusion of elevation as a predictor. As discussed earlier, a station is unlikely to provide useful information about the variable of interest at gridpoints beyond its CDD. To prevent extrapolation to unrealistic values the interpolated anomaly fields were forced towards zero at grid-points beyond the influence of any stations. This was accomplished by creating synthetic stations with anomaly values of zero in regions where there were no stations within a predefined distance, chosen to be equal to the global-mean CDD. These distances were 450 km for precipitation, 750 km for diurnal temperature range and 1200 km for mean temperature.

Secondary Variables

The patchy distribution of stations with secondary variable data, particularly prior to 1960, meant that interpolation of anomalies directly from station data was not feasible. This is despite the large CDDs determined for cloud cover and, particularly, vapor pressure. We therefore used the existing data to develop and/or test empirical (in the case of cloud cover and ground frost frequency) or conceptual (vapor pressure and wet-day frequency) relationships with the primary variables: for wet-day frequency this was precipitation; for vapor pressure, mean temperature and diurnal temperature range; for cloud, diurnal temperature range; for ground frost frequency, mean temperature and diurnal temperature range. The details of the predictive relationships with primary variables are described in detail in New *et al* (2000).

These relationships were used to calculate grids of synthetic monthly anomalies. In the case of cloud cover, wet-day frequency and vapor pressure, the synthetic grids were then blended with station anomalies in the regions where such data were available. Frost frequency anomaly grids are entirely synthetic.

4.1.2 Combination with Climatology

Monthly anomaly fields were combined with the 1961-1990 monthly climatology, either through simple addition or, precipitation, through back-transformation of the percentage anomaly field.

In some areas with sparser station coverage, the 1961-1990 average of the monthly anomaly grids diverged from zero, for example over Angola and the Democratic Republic of the Congo. This arose directly from the interpolation error in the individual anomaly fields which, not unexpectedly, did not sum to zero. To maintain consistency, individual fields from 1961-1990 were adjusted so that their 1961-1990 mean was zero by subtracting this mean interpolation error.

It should be noted that a direct consequence of the relaxation of the anomaly surfaces to zero in regions with no data coverage is that the resulting monthly climate relaxes towards the 1961-1990 climatology in such areas.

5. EQUIPMENT

5.1 Instrument Description

5.1.1 Platform (Satellite, Aircraft, Ground, Person)

All data are ground-based measurements from meteorological instruments, and observer estimates for clouds.

5.1.2 Mission Objectives

To create a spatially continuous gridded data set at 0.5 degree lat/long resolution of key surface monthly climate variables important for modelling earth surface processes.

5.1.3 Key Variables

Primary variables (based solely on observations)

- precipitation
- temperature
- diurnal temperature range

Secondary variables (based on observations and primary variables)

- wet-day frequency
- vapor pressure
- cloud cover

5.1.4 Principles of Operation

Standard instrumentation and observing practice, on a national and/or organizational basis. In most instances instrument metadata were not available, and various assumptions about instrument type, height etc are made (see New *et al*. 1999 for details).

5.1.5 Instrument Measurement Geometry

See Section 5.1.4.

5.1.6 Manufacturer of Instrument

See Section 5.1.4.

5.2 Calibration

5.2.1 Specifications

Not applicable to this data set.

5.2.1.1 Tolerance

Not applicable to this data set.

5.2.2 Frequency of Calibration

Not applicable to this data set.

5.2.3 Other Calibration Information

Not applicable to this data set.

6. PROCEDURE

6.1 Data Acquisition Methods

Station data were obtained from numerous sources over many years by various personnel at the Climatic Research Unit, most notably Mike Hulme, Phil Jones and Mark New (see New *et al.* 1999, 2000). Major sources include National Meteorological Agencies, Global Telecommunications System/Global Climate Observing System (GTS/GCOS), the Global Historical Climatology Network (GHCN), WMO 1961-1990 climate normals, Centro Internacional de Agricultura Tropical (CIAT) and Institut de Recherche pour le Développement (IRD), France. A full list of sources is available in New *et al* (1999, 2000).

6.2 Spatial Characteristics

6.2.1 Spatial Coverage

The coverage is for global land areas, excluding Antarctica. The land mask is derived from 5-minute lat/long Terrainbase (http://www.ngdc.noaa.gov/mgg/topo/) digital elevation/bathymetry data set, area averaged to 0.5 degrees lat/long 0.5 degree Cells are considered land if more than 25 (25%) of Terrainbase cells are land. Additionally, small ocean islands are also considered to be land, even when their area is less than 25 Terrainbase cells. For consistency, the Initiative II staff applied the ISLSCP land-water mask and differences were handled as described in Section 9.2.3.

6.2.2 Spatial Resolution

The data are provided on two equal angle Earth grids with spatial resolutions of 0.5 and 1.0 degree in both latitude and longitude.

6.3 Temporal Characteristics

6.3.1 Temporal Coverage

The monthly climatology represents average conditions from 1961 to 1990.

6.3.2 Temporal Resolution

Monthly totals or averages.

7. OBSERVATIONS

7.1 Field Notes

None.

8. DATA DESCRIPTION

8.1 Table Definition with Comments

Not applicable to this data set.

8.2 Type of Data

8.2.1 Parameter/ Variable Name	8.2.2 Parameter/ Variable Description	8.2.3 Data Range[1]	8.2.4 Units of Measurement	8.2.5 Data Source
Monthly Monthly Climatology (1961-1990) Maps (*.asc)				
cloud_covr	Cloud cover	min=0 max=100	percent	cloud observations sunshine measurements
diurnal_tmp	Diurnal temperature range	min=0	degrees C	Air temperature thermometers
precip	Monthly precipitation as analyzed from precip-gauge measurements.	min = 0	Mm/day	Rain-gauge
radiation	Shortwave Radiation, averaged over 24 hours (including 0 value for nighttime).		W/m2	
frost_freq	Ground frost frequency	min=0 max=all days in month	days	ground thermometer; air thermometer; synthetic
wind	Wind speed		m/s	
temp_max	Monthly maximum temperature	N/A	degrees C	Air temperature thermometers
temp_min	Monthly minimum temperature	N/A	degrees C	Air temperature thermometers
temp_mean	Monthly mean temperature	N/A	degrees C	Air temperature thermometers
vapor_pres	Vapor pressure	min=0.1	Hpa	wet/dry bulb electronic synthetic
wetday_frq	Wet day frequency (number of days per month with >=0.1mm precipitation)	min=0 max=all days in month	days	Rain-gauge (days with >0.1mm)
Differences Tables (*.dif)				
Lat	Latitude for the center of a cell. South latitudes are negative.	Min=-90 Max=90	Decimal Degrees	Earth Grid
Lon	Longitude for the center of a cell. West longitudes are negative.	Min=-180 Max=180	Decimal Degrees	Earth Grid
Data_Removed	Value in each original file that doesn't match the land/water mask, and was removed.	Varies by Parameter	Varies by Parameter	Original Data
Data_Added	Value added to the original file because the land/water mask indicated land, so an interpolated point was added.	Varies by Parameter	Varies by Parameter	Computed

Interpolation_Level	The number of times the interpolation routine was run to get a value for this point. The higher the number, the less reliable the value is.	Min=1 Max=5	Unitless	Computed
Change Map (*_changemap.asc)				
Point Changed	Differences between the ISLSCP II land/water mask and the original data: -1 = ISLSCP II mask is water and original data is land (data removed) 0 = Data sets agree over land or water (data unchanged) ≥1 = ISLSCP II mask is land and original data is water or missing (data interpolated)	Min=-1 Max=1	See 8.2.2	Computed

***NOTE:** Data range values are not available at this revision. A value of –99 is assigned to water bodies and –88 to missing data over land.

8.3 Sample Data Record

The "differences" file is an ASCII table with some header lines, then the lat and long coordinates of each removed point, plus the value of that point. See the sample below.

```
ISLSCP II Differences for file 'cru5_cloud_covr_hd_198601.asc'.
Contains Lat-Lon coordinates and data for each point in the original file
that differed from the ISLSCP II Land/Water mask, and thus was removed.
Points added using interpolation are listed at the bottom of this file.

Lat,Lon,Data_Removed
83.75,-37.75,3.7
83.75,-37.25,3.6
83.75,-36.75,3.6
83.75,-36.25,3.6
83.75,-35.75,3.6
83.75,-35.25,3.5
83.75,-34.75,3.5
83.75,-34.25,3.5
83.75,-33.75,3.5
83.75,-33.25,3.5
83.75,-32.75,3.5
83.75,-32.25,3.5
   "      "
   "      "
   "      "
Lat,Lon,Data_Added,Interpolation_Level
81.25,60.25,-88,1
80.25,54.75,5.83,1
78.75,-74.25,2.9,1
78.25,-113.75,3.45,1
68.75,52.25,7,1
64.75,-172.25,5.8,1
64.75,-65.25,4.67,1
```

8.4 Data Format

All of the files in the ISLSCP Initiative II data collection are in the standard ArcGIS ASCII Grid format. The file format consists of numerical fields of varying length, which are delimited by a single space and arranged in columns and rows. All values in these files are written as integer numbers. However, note the very large maximum values given in Section 8.2.3. In all layers areas with no data over oceans are assigned the value of –99. Areas with no data over land (i.e. Antarctica) are assigned the value of –88.

The files at different spatial resolutions each contain the following numbers of columns and rows:

One degree: 360 columns by 180 rows

1/2 degree: 720 columns by 360 rows

All files are gridded to a common equal-angle lat/long grid, where the coordinates of the upper left corner of the files are located at 180 degrees W, 90 degrees N and the lower right corner coordinates are located at 180 degrees E, 90 degrees S. Data in the map files are ordered from North to South and from West to East beginning at 180 degrees West and 90 degrees North.

The ASCII map files (with the extension of ".asc") have all had the ISLSCP II land/water mask applied to them. All points removed from the original half-degree files are stored in "differences" files (with the extension ".dif"). These ASCII files contain the Latitude and Longitude location of the cell-center of each removed point, and the data value at that point. At the bottom of these files are also a list of all points added to the file through "nearest neighbor averaging" interpolation, where the land/water mask indicated land but there was no data in the original file. There is also a column called "Interpolation_Level" that contains the number of times the interpolation routine was run to get a value for that point. The higher the number, the less reliable the value is. There is one ".dif" file for each ASCII map file. Note: the added points for Antarctica were not included in the ".dif" files because there is no real data there.

The "change map" files show the results of applying the land/water mask, as a viewable ASCII map: all points added (positive number, containing the "Interpolation_Level", see above), all points unchanged ("0"), and all points removed ("-1"). There is only a file for the half-degree data, as the 1-degree data was created through averaging.

WARNING: The 1.0 x 1.0 degree product is for browse use only. These data files are averaged from the original 0.5 x 0.5 degree pixels. Thus the data values at specific pixels are not exact. Use this data with caution and refer to the original half-degree data files for specific information.

8.5 Related Data Sets

A major update of the CRU05 data to 2001 is presently underway. Updates will be made available through the Climatic Research Unit (http://iridl.ldeo.columbia.edu/SOURCES/.UEA/.CRU/.New/.CRU05/.monthly/) and/or the Intergovernmental Panel on Climate Change (IPCC) Data Distribution Centre (DDC) (http://www.ipcc-data.org/obs/cru_climatologies.html). ISLSCP II project information and data sets may also be obtained from the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) http://daac.ornl.gov/ISLSCP_II/islscpii.shtml.

9. DATA MANIPULATIONS

9.1 Formulas

9.1.1 Derivation Techniques/Algorithms

See Section 4.

9.2 Data Processing Sequence

9.2.1 Processing Steps and Data Sets

Station data were received in a variety of paper and digital forms, and in a variety of formats and measurement units. In general, the processing steps for the data set are:

- reformatting to standard CRU format and units of measurement.
- quality control of station information –name, country, station ID (matching with existing data at CRU), coordinates and elevation.
- quality control of station data (see New *et al*. 1999, 2000 for details).
- merging with existing station data (either normals or time-series).
- interpolation.
- visual quality control through inspection of interpolated fields.
- re-interpolation after corrections to data.
- data storage.

9.2.2 Processing Changes

None given.

9.2.3 Additional Processing by the ISLSCP Staff

The original data files submitted to the ISLSCP Initiative II staff contained one year of a variable in each file, with 12 separate monthly maps concatenated together, one after the other. Each value was in a 5-character fixed-length field, and there was considerable 'touching' of the fields (no delimiting space between them). The Monthly Time Series data files had an origin at the Greenwich Meridian and the South Pole while the Monthly Mean climatology data files had an origin at the Greenwich Meridian and the North Pole. Missing data were encoded as -9999. This data was at the half-degree resolution only.

The data files were reprocessed and each yearly variable file was broken up into 12 monthly files, and each line was parsed for the correct number of values and each value was given a one space-delimiter. The missing data values of -9999 were changed to -99. All data files were processed so that the origin (upper-left corner) of each file was at the dateline and the North Pole.

Also, the units for many of the original files were 10 times normal units ("degrees C x 10" for example). The only parameters in normal units were Monthly Time Series precipitation, and Monthly Mean cloud cover and radiation. The "10x" files were processed by dividing each value by 10. Missing data values were kept at -99.

The half-degree files were then further processed by comparing the data files for consistency against the ISLSCP II land/water mask. Missing land data (usually in

Antarctica) in these files were added in the form of -88 values. All added values outside of the Antarctic had these -88 missing values replaced with "nearest neighbor averaging" interpolation. New ASCII table files containing the removed and added points (points that didn't match the land/water mask), also called "differences" files with the extension ".dif", were also created. These files contain the Latitude and Longitude of the cell-center of each removed or added point, and the data value at that point. The added points also contain a column called "Interpolation_Level" that contains the number of times the interpolation routine was run to get a value for that point. The higher the number, the less reliable the value is. Note: the added points for Antarctica were not included in the ".dif" files because there is no real data there.

Finally, a "change map" was created for the half-degree data, showing the results of applying the land/water mask, as a viewable ASCII map: all points added (positive number, containing the "Interpolation_Level", see above), all points unchanged ("0"), and all points removed ("-1").

Seeing a need for a 1.0 degree product, the ISLSCP Initiative II Staff took the 0.5 data files and "averaged" them down to one degree. For each 1.0 cell, 4 0.5 degree cells were averaged, ignoring missing data cells, and filling the new cell with -99 if three or more 0.5 cells were -99 as well. Then the data files were renamed to the current naming scheme (see Section 1.3). Thus, there are no ".dif" files or "change maps" for the 1.0 degree data.

WARNING: The 1.0 x 1.0 degree product is for browse use only. These data files are averaged from the original 0.5 x 0.5 degree pixels. Thus the data values at specific pixels are not exact. Use this data with caution and refer to the original 0.5 degree data files for specific information.

9.3 Calculations

9.3.1 Special Corrections/Adjustments

Data in different formats (units of measurement, instrument type) were converted to a standard format (see New *et al*, 1999, 2000 for details):

- precip – inches to mm
- temp_mean – none, but where available, temperature is calculated as mean of maximum and minimum temperature; where tmax and tmin were not available, mean temperature as provided by source is used.
- wetday_frq – wet day frequencies for >1mm threshold were converted to an equivalent for a >0.1mm threshold, using an empirical conversion algorithm.
- vapor_pres – approximately 40% of data were provided as relative humidity; these were converted to vapor pressure at dew point using standard formulations (New et al, 1999, 2000)
- cloud_covr – nearly half the “cld” data were provided as sunshine hours of sunshine percent; these were converted to cld. Nearly half the “sunp” data were provided as cloud; these were converted to sunp
- wind – all wnd data are assumed to be at 10m height.

9.4 Graphs and Plots

The data can visualized at the Intergovernmental Panel on Climate Change (IPCC) Data Distribution Centre (http://www.ipcc-data.org/). Also see New et al. (1999, 2000) for figures.

10. ERRORS

See section 11 notes for additional discussion of known errors and problems.

10.1 Sources of Error

Errors arise from several sources, and have not been quantified:

- Input data – some erroneous station *will* escape quality control, particularly when the process is automated, as is necessary for large data sets. Errors in station input data will therefore "cascade" through to the final product.
- Interpolation error – even with perfect station data, errors in interpolation away from station control is inevitable, and vary according to climate element and location. Average interpolation error for climate means is assessed through implicit Generalised Cross-Validation (GCV) during interpolation and are reported in New *et al* (1999). Error for anomaly fields are not reported, but are highly dependent on station density and the spatial coherence of climate anomalies.

10.2 Quality Assessment

10.2.1 Data Validation by Source

See Section 9.

10.2.2 Confidence Level/Accuracy Judgement

1961-1990 Climatology Primary variables in the 1961-1990 climatology were compared against the climate data sets of Legates and Willmott (1990, henceforth LW; Legates and Willmott 1990) and Leemans and Cramer (1991, updated). Agreement is generally good, but marked differences are apparent in some regions; reasons for this are a combination of differing data sets and interpolation methods, where the elevation dependence in the CRU data result in improved estimates in high elevation areas. There are no exiting global climatologies of secondary variables to permit intercomparison. See New et al. (1999) for complete discussion.

10.2.3 Measurement Error for Parameters and Variables

See Section 10.1, 10.2 and New *et al*. (1999,2000).

10.2.4 Additional Quality Assessment Applied

None.

11. NOTES

11.1 Known Problems with the Data

The precipitation measurements have not been corrected for the systematic gauge-measuring error (in general an underestimation of the true precipitation which varies between 5% and more than 100% of the measured data for monthly accumulations depending on the weather conditions during the month). Correction of the gridded data is recommended, but left to the users, by making use of, for example, Legates (1987) climatological correction factors. The correction factors from Legates (1987) are available in this ISLSCP INITIATIVE II data collection as a part of the Global Precipitation Climatology Centre (GPCC) Rain gauge-Only Global Precipitation data set.

Since being used, a number of specific problems have been discovered, and are likely to continue to come to light as researchers focus on specific aspects of the data set. A list of "known problems" is maintained on an ad-hoc basis on the IPCC-DDC Web site, which is a source for the full 1901-1996 data set (http://ipcc-ddc.cru.uea.ac.uk/). As of September 2002, the list is as follows:

Wet Days

Area affected	**Reason for error**
Brazil (Amazonia)	Conversion was required from available data (the threshold was 1.0mm). The method used probably gave a positive bias and thus overestimated the number of wet days.
Spain and Spanish stations in N. Africa	Error over definition for approximately 40 stations. Threshold was assumed to be 0.1mm but was in fact 1.0mm.
Syria	Error over definition for all stations. Threshold was assumed to be 0.1mm but was in fact 1.0mm.

Diurnal Temperature Range

Area affected	**Reason for error**
Greenland *(see below)	Lack of station data in central Greenland has caused (too) high values to be interpolated to the region.
Poland	Some stations were found to have their range-values based on monthly extreme max. and min. temperature values instead of average values - therefore values too large.

Wind Speed

Area affected	**Reason for error**
Bolivia	Values for several stations found to be in knots – therefore too high
Greece	Values for six stations found to be in knots – therefore too high.
Honduras	Values for all stations were found to be in knots – therefore too high.
Sierra Leone	Values for all stations found to be in knots – therefore too

	high.
Sudan	Values for all stations originally in miles per hour but incorrect conversion to m/s produced values too low
Peru (Iquitos & Cajamarca)	Values for both stations found to be in error (much higher than is feasible in the area)

Relative Humidity

Station data relating to humidity was split roughly half-and-half between pressure (VP) or relative humidity (RH). The conversion of VP to RH, or vice versa, (see New *et al*.) does pose problems in some parts of the world – notably in the coldest areas in winter months. This is due to a loss of instrumental precision at very low temperatures. Small errors are magnified when conversion (using mean temperature) is undertaken. For this reason, systematic differences in winter RH are apparent according to political divisions in areas like the northern Russia and northern Canada. In addition, mean monthly RH is affected by the timing of daily readings. For this reason, mean RH may be biased if mean monthly values are not based on true daily mean values, or if the time of measurement of daily temperatures and RH do not coincide.

Greenland – all variables

The interior of Greenland is poorly covered by meteorological observation. This coupled with the presence of the high elevation ice cap makes interpolation of climate normals very difficult due to the potential for unusual lapse rates. It is likely that significant bias may be present with all variables for the interior of the landmass (e.g. diurnal temperature range and precipitation too high).

11.2 Usage Guidance

In data void/sparse continental areas, the quality of the analysis results will be poor. Maps of station coverage are available in New *et al*. (1999, 2000) and lists showing the location of stations used to construct both the "base" 1961-1990 mean climate grids and the monthly anomaly fields are available from the following web site:
http://www.cru.uea.ac.uk/~timm/grid/CRU_CL_1_0_text.html#Intro

The 1.0 x 1.0 degree version created by the ISLSCP Initiative II staff is recommended for browse use only. These data files are averaged from the original 0.5 x 0.5 degree pixels. Thus the data values at specific pixels are not exact. Use these data with caution and always refer to the original 0.5 data files for specific information.

11.3 Other Relevant Information

None.

12. REFERENCES

12.1 Satellite/Instrument/Data Processing Documentation

None.

12.2 Journal Articles and Study Reports

Dai, A., I. Fung and A. Del Genio (1997). "Surface observed global land precipitation variations during 1900-1988." *Journal of Climate* **11**: 2943-2962.

Easterling, D. R., B. Horton, P. D. Jones, et al. (1997). "Maximum and minimum temperature trends for the globe." Science **277**: 364-367.

Hulme, M. (1994). Validation of large-scale precipitation fields in General Circulation Models. *Global precipitation and climate change*. M. Desbois and F. Desalmand. Berlin, Springer-Verlag**:** 387-405.

Hutchinson, M. F. (1995). "Interpolating mean rainfall using thin plate smoothing splines." *International Journal of Geographical Information Systems* **9**: 385-403.

Jones, P. D., M. New, D. E. Parker, S. Martin and I. G. Rigor (1999). "Surface air temperature and its changes over the past 150 years." *Reviews of Geophysics* **37**(2): 173-199.

Jones, P. D., T. J. Osborn and K. R. Briffa (1997). "Estimating sampling errors in large scale temperature averages." *Journal of Climate* **10**: 2548-2568.

Leemans, R. and W. Cramer (1991). The IIASA database for mean monthly values of temperature, precipitation and cloudiness on a global terrestrial grid. Laxenburg, IIASA**:** 63pp.

Legates, D. R. and C. J. Willmott (1990). "Mean seasonal and spatial variability in gauge-corrected, global precipitation." *International Journal of Climatology* **10**: 111-127.

Legates, D. R. and C. J. Willmott (1990). "Mean seasonal and spatial variability in global surface air- temperature." *Theoretical and Applied Climatology* **41**: 11-21.

New, M., M. Hulme and P. Jones (1999). "Representing twentieth-century space-time climate variability. Part I: Development of a 1961-90 mean monthly terrestrial climatology." *Journal of Climate* **12**: 829-856.

New, M., M. Hulme and P. Jones (2000). "Representing twentieth-century space-time climate variability. Part II: Development of 1901-1996 monthly grids of terrestrial surface climate." *Journal of Climate* **13**(13): 2217-2238.

Wahba, G. (1979). "How to smooth curves and surfaces with splines and cross-validation." 24th Conference on the Design of Experiments, Research Triangle Park, North Carolina, U.S. Army Research Office.

13. DATA ACCESS

13.1 Data Access Information

The ISLSCP Initiative II data are archived and distributed through the Oak Ridge National Laboratory (ORNL) DAAC for Biogeochemical Dynamics at http://daac.ornl.gov.

13.2 Contacts for Archive

E-mail: uso@daac.ornl.gov

Telephone: +1 (865) 241-3952

13.3 Archive/Status/Plans

The ISLSCP Initiative II data are archived at the ORNL DAAC. There are no plans to update these data.

14. GLOSSARY OF ACRONYMS

ADW	Angular Distance Weighting
CDD	Correlation Decay Distance
CIAT	Centro Internacional de Agricultura Nacional
CRU	Climatic Research Unit (University of East Anglia)
DAAC	Distributed Active Archive Center
DDC	Data Distribution Centre (IPCC)
GCOS	Global Climate Observation System
GCV	Generalized Cross Validation
GHCN	Global Historical Climatology Network
GPCC	Global Precipitation Climatology Centre
GSFC	Goddard Space Flight Center (NASA)
GTS	Global Telecommunications System
IPCC	Intergovernmental Panel on Climate Change
IRD	Institut de Recherche pour le Developpement
ISLSCP	International Satellite Land Surface Climatology Project
NASA	National Aeronautics and Space Administration
ORNL	Oak Ridge National Laboratory
RH	Relative Humidity
VP	Vapor Pressure
WMO	World Meteorological Organization