GPCP ONE-DEGREE DAILY PRECIPITATION DATA SET DOCUMENTATION

George J. Huffman
David T. Bolvin

SSAI and Laboratory for Atmospheres,
NASA Goddard Space Flight Center

21 April 2004

i. CONTENTS

ii. KEYWORDS

1DD
accuracy
AGPI precipitation product
archive and distribution sites
ATOVS
data access policy
data file access technique
data file identifier
data providers
data set archive
data set creators
data set inventory
data set name
day definition
documentation creator
documentation revision history
estimate missing values
file date
GPCP
GPI precipitation product
GPROF fractional occurrence of precipitation
grid
intercomparison results
IR

IR data correction
IR Tb histogram data set
known anomalies
known errors
missing days
obtaining data
originating machine
period of record
processing change
production and updates
read a day of data
read a day of byte-swapped data
read the header record
references
satellite-gauge precipitation product
SG error correction
similar data sets
spatial coverage
spatial resolution
SSM/I
SSM/I error detection/correction
standard missing value
temporal resolution
TMPI
TOVS
TOVS data correction
TOVS precipitation estimate
units of the 1DD estimates

iii. ACRONYMNS

1DD	One Degree Daily
AGPI	Adjusted GPI
ATOVS	Adjusted TOVS
CPC	Climate Prediction Center
DMSP	Defense Meteorological Satellite Program
GARP	Global Atmospheric Research Programme
GATE	GARP Atlantic Tropical Experiment
Geo	Geosynchronous
GEWEX	Global Energy and Water Cycle Experiment
GMDC	GPCP Merge Development Centre
GMS	Geosynchronous Meteorological Satellite
GOES	Geosynchronous Operational Environmental Satellites
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GPI	Global Precipitation Index
GPROF	Goddard Profiling Algorithm
GSPDC	Geostationary Satellite Precipitation Data Centre
HIRS2	High-Resolution Infrared Sounder 2
IR	Infrared
lat/lon	latitude/longitude
Leo	Low-Earth-orbit
MB	megabytes
MSU	Microwave Sounding Unit
NASA	National Aeronautics and Space Administration
NCDC	National Climatic Data Center
NCEP	National Centers for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
SRDC	Surface Reference Data Center

SSM/I	Special Sensor Microwave/Imager
SSU	Stratospheric Sounding Unit
Ta	Antenna Temperature
Tb	Brightness Temperature
TIROS	Television Infrared Operational Satellite
TMPI	Threshold Matched Precipitation Index
TOVS	TIROS Operational Vertical Sounder
UTC	Universal Coordinated Time (same as GMT, Z)
WCRP	World Climate Research Programme
WMO	World Meteorological Organization

....

1. DATA SET NAMES AND GENERAL CONTENT

The formal *data set name* is the "GPCP One-Degree Daily Precipitation Data Set." It is also referred to as the "1DD Data Set."

The data set currently contains a product providing daily, global gridded values of precipitation totals for the 2.5-year period October 1996 through January 2004.

The main refereed citation for the data set is Huffman et al. (2001) (all references are listed in section 13).

....

2. RELATED PROJECTS, DATA NETWORKS, AND DATA SETS

The *data set creators* are G.J. Huffman, R.F. Adler, and D.T. Bolvin, working in the Laboratory for Atmospheres, NASA Goddard Space Flight Center, Code 912, Greenbelt, Maryland, 20771 USA, as the GPCP Merge Development Centre.

....

The work is being carried out as part of the Global Precipitation Climatology Project (*GPCP*), an international project of the World Meteorological Organization/World Climate Research Programme/ Global Energy and Water Experiment (WMO/WCRP/GEWEX) designed to provide improved long-record estimates of precipitation over the globe. The GPCP home page is located at

  http://orbit-net.nesdis.noaa.gov/arad/gpcp/

....

The 1DD Data Set contains data from several *data providers*:

1. GPCP Geostationary Satellite Precipitation Data Centre (IR Tb histograms),
2. GPCP Merge Development Centre (GPCP SG Merged Precipitation estimate and GPROF 6.0 SSM/I fractional occurrence), and
3. GSFC Satellite Data Utilization Office (TOVS precipitation estimates).

Some of these data sets extend beyond the 1DD period in their original archival locations.

....

There are numerous *similar data sets*, although no other matches all the attributes of being global, daily, and formed by combining multiple data sources. The Susskind et al. (1997) TOVS precipitation data set is available as global daily fields, but is based on a single data source. Several SSM/I-based data sets are available as daily, single-sensor data sets with significant data voids in cold-land, snow-covered, and ice-covered areas, including GPROF 6.0 (based on Kummerow et al 1996) and NOAA Scattering (Grody 1991), among others. Other daily, single-sensor data sets are available for open-water regions based on SSM/I data (Wentz and Spencer 1998) or MSU data (Spencer 1993). Numerous daily single-sensor or combination data sets are available at the regional scale, but are not really "similar."

....

3. STORAGE AND DISTRIBUTION MEDIA

The *data set archive* consists of unformatted binary files. The day files are appended into month files that have ASCII headers. The 1DD is distributed by FTP over the Internet and by special request on Exabyte tape media. Each month file occupies ~7.4 MB, and the whole data set contains ~430 MB (uncompressed).

....

4. READING THE DATA

A *data file identifier* is embedded in the data file name as

```
  gpcp_<version>_<variable><technique>.<date>
```

where

```
  <version>   is a short alphanmueric string giving the version:
              nosm = beta version (obsolete);
              1dd  = Version 1;
  <variable>  is the parameter:
              p    = precipitation; TMPI 40N-S, ATOVS outside that;
              e    = random error.
  <technique> is the algorithm:
              1d   = 1DD
  <date>      is the UTC date as YYYYMMDD (i.e., numerical year, month,
              day);
```

....

The *data file access technique* is the same for all files in a release. These files are accessible by standard third-generation computer languages (FORTRAN, C, etc.).

Each monthly file consists of a 1440-byte header record containing ASCII characters (which is the same size as one row of data), then a month (28, 29, 30, or 31) of daily grids of size 360x180 containing REAL*4 values. The header line makes the file nearly self-documenting, in particular spelling out the variable and version names, and giving the units of the variable. The header line may be read with standard text editor tools or dumped under program control. All the month's days of data are present, even if some have no valid data. Grid boxes without valid data are filled with the (REAL*4) "missing" value -99999. The data may be read with standard data-display tools (after skipping the 1440-byte header) or dumped under program control.

....

The *originating machine* on which the data files where written is a Silicon Graphics, Inc. Unix workstation, which uses the "big-endian" IEEE 754-1985 representation of REAL*4 unformatted binary words. Some CPUs including PCs and DEC machines may require a change of representation before using the data.

....

For the monthly files, it is possible to *read the header record* with most text editor tools, although the size (1440 bytes) may be longer than some tools will support. Alternatively, the header record may be dumped out under program control, as demonstrated in the following programming segment. The header is written in a "PARAMETER=VALUE" format, where PARAMETER is a string without embedded blanks that gives the parameter name, VALUE is a string that gives the value of the parameter, and blanks separate each "PARAMETER=VALUE" set. To prevent ambiguity, no spaces or "=" are permitted as characters in either PARAMETER or VALUE.

```
C**********************************************************************
C       FORTRAN program segment to read the header record and file
C       arrays of KEYWORD and VALUE.
C
C     The header is written in a "PARAMETER=VALUE" format, where
C     PARAMETER is a string without embedded blanks that gives the
C     parameter name, VALUE is a string that gives the value of the
C     parameter, and blanks separate each "PARAMETER=VALUE" set.  To
C     prevent ambiguity, no spaces or "=" are permitted as characters
C     in either PARAMETER or VALUE.
C
C       The data arrays are dimensioned large enough that we don't have
C       to be careful about overflows; they could be reduced if space
C       is short.
C**********************************************************************
C
        IMPLICIT        NONE
        CHARACTER*1440  header
        CHARACTER*80    keywd (50), value (50), file
        INTEGER         neq   (50), kstrt (50), nvend (50)
        INTEGER         iret, i, l_header, ipt, in, numkey, j
C
C       Open the data file (using January 1997 as an example) with a
C       RECL of 1 data row.
C               ==>>    WARNING WARNING WARNING    <<==
C     The RECL is defined differently on different machines; it isn't
C     specified in the FORTRAN77 standard.  On SGI it's in 4-B words.
C     If you find that you only get 90 good values and then garbage
C     (either all zeros or random values) in the last 270 elements of
C       the row, your machine wants RECL in bytes, and you should say
C       RECL=1440 in the following OPEN.
C
      file = 'gpcp_1dd_p1d.199701'
        OPEN  ( UNIT=10, FILE=file, ACCESS='DIRECT',
     +          FORM='UNFORMATTED', STATUS='OLD', RECL=360,
     +            IOSTAT=iret )
        IF  ( iret .NE. 0 )  THEN
            WRITE (*, *) 'Error: open error', iret,
     +                   ' on file ', file
            STOP
        END IF
C
C       Read the header (the first record) and close the file.
C
        READ ( UNIT=10, REC=1, IOSTAT=iret )  header
```

```
      IF  ( iret .NE. 0 )  THEN
          WRITE (*, *) 'Error: read error', iret,
     +                 ' on file ', file
          STOP
      END IF
      CLOSE ( UNIT=10 )
C
C     Find the actual length of the header (as opposed to the
C     declared FORTRAN size) by parsing back from the end for the
C     first non-blank character (it was written blank-filled).
C
      DO  10 i = 1, 1440
          IF  ( header (1441-i:1441-i) .NE. ' ' )  GO TO 20
   10 CONTINUE
      WRITE (*, *) 'Error: found no non-blanks in the header'
      STOP
   20 l_header = 1441 - i
C
C     Parse for "=".
C
      ipt = 1
      DO  30 i = 1, l_header
          in = INDEX ( header (ipt:l_header), '=' )
          IF  ( in .EQ. 0 )  THEN
              GO TO 40
            ELSE
              neq (i) = ipt + in - 1
              ipt     = ipt + in
          END IF
   30 CONTINUE
      WRITE (*, *) 'Error: ran through header without ending parsing'
      STOP
   40 CONTINUE
      numkey = i - 1
C
C     Now find corresponding beginning of each keyword by parsing
C     backwards for " ".  The first automatically starts at 1.  We
C     assume that there are at least 2 keywords!
C
      kstrt (1) = 1
      DO  60 i = 2, numkey
          DO  50 j = 1, neq (i) - 1
              IF  ( header (neq(i)-j:neq(i)-j) .EQ. ' ' )  GO TO 55
   50     CONTINUE
   55     kstrt (i) = neq (i) - j + 1
   60 CONTINUE
C
C     The end of the value string is the 2nd character before the start
C     of the next keyword, except the last is at l_header.
C
      DO  70 i = 1, numkey - 1
          nvend (i) = kstrt (i+1) - 2
   70 CONTINUE
      nvend (numkey) = l_header
C
C     Now use these indices to load the arrays.  We assume that null
C     strings will not be encountered.
C
      DO  80 i = 1, numkey
          keywd (i) = header (kstrt(i):neq(i)-1)
```

```
            value (i) = header (neq(i)+1:nvend(i))
   80   CONTINUE
C
C       Now there are "numkey" keywords with corresponding values ready
C       to be manipulated, printed, etc.  For example, print them:
C
        DO  85 i = 1, numkey
            WRITE (*, *) '"', keywd (i) (1:neq(i)-kstrt(i)), '" = "',
     +                   value (i) (1:nvend(i)-neq(i)), '"'
   85   CONTINUE
        STOP
        END
```

....

It is possible to *read a day of data*, i.e., one grid of data, with many standard data-display tools. In the monthly file the 1440-byte header is designed to be exactly the size of one row of data, so the header may be bypassed by skipping 1440 bytets or 360 REAL*4 data points or one row. Alternatively, the data may be dumped out under program control as demonstrated in the following programming segment. Once past the header, there is always a month of daily grids. In all cases the grids are of size 360x180 containing REAL*4 values. All days of data in the month are present, even if some have no valid data. Grid boxes without valid data are filled with the (REAL*4) "missing" value -99999. Days in the month that lack data are entirely filled with "missing."

```
C**********************************************************************
C       FORTRAN program segment to read a month of data.
C
C       Once the header of size 1440 bytes (one data row) is skipped
C       (for the official release), there is always a month of grids.
C       All grids are of size 360x180 containing REAL*4 values.  All
C       days of data in the month are present, even if some have no
C       valid data.  Grid boxes without valid data are filled with
C       the (REAL*4) "missing" value -99999.
C**********************************************************************
C
        IMPLICIT        NONE
        CHARACTER*80    fname
        REAL*4          data (360, 180)
        INTEGER         nday, nskip, iret, i, j
C
C       Set the user input for file name and day number (using 19970121
C       as an example).
C
        fname = 'gpcp_1dd_p1d.199701'
        nday  = 21
C
C       Open the data file with a RECL of 1 data row.
C               ==>>    WARNING WARNING WARNING    <<==
C     The RECL is defined differently on different machines; it isn't
C     specified in the FORTRAN77 standard.  On SGI it's in 4-B words.
C     If you find that you only get 90 good values and then garbage
C     (either all zeros or random values) in the last 270 of the row,
C     your machine wants RECL in bytes, and you should say RECL=1440
C     in the following OPEN.
C
        OPEN  ( UNIT=10, FILE=fname, ACCESS='DIRECT',
     +          FORM='UNFORMATTED', STATUS='OLD', RECL=360,
     +            IOSTAT=iret )
```

```
      IF  ( iret .NE. 0 )  THEN
          WRITE (*, *) 'Error: open error', iret,
   +                   ' on file ', fname
          STOP
      END IF
C
C     Compute the number of records to skip, namely 1 for the
C     header and 180 for each intervening day.
C
      nskip = 1 + ( nday - 1 ) * 180
C
C     Read the 180 rows of data and close the file.
C
      DO  10 j = 1, 180
          READ ( UNIT=10, REC=j+nskip, IOSTAT=iret )
   +           ( data (i, j), i = 1, 360 )
          IF  ( iret .NE. 0 )  THEN
              WRITE (*, *) 'Error: read error', iret,
   +                       ' on file ', fname
              STOP
          END IF
 10   END DO
      CLOSE ( UNIT=10 )
C
C     Now array "data" is ready to be manipulated, printed, etc.
C     For example, dump the single day as unformatted direct:
C
      OPEN  ( UNIT=10, FILE='junk', ACCESS='DIRECT',
   +          FORM='UNFORMATTED', RECL=360, IOSTAT=iret )
      IF  ( iret .NE. 0 )  THEN
          WRITE (*, *) 'Error: open error', iret,
   +                   ' on file junk'
          STOP
      END IF
      DO  20 j = 1, 180
          WRITE ( UNIT=10, REC=j, IOSTAT=iret )
   +            ( data (i, j), i = 1, 360 )
          IF  ( iret .NE. 0 )  THEN
              WRITE (*, *) 'Error: write error', iret,
   +                       ' on file junk'
              STOP
          END IF
 20   END DO
      CLOSE ( UNIT=10 )
      STOP
      END
```

....

It is possible to *read a day of byte-swapped data*. The GPCP data are generated using a Silicon Graphics, Inc. Unix workstation, which uses the "big-endian" IEEE 754-1985 representation of REAL*4 unformatted binary words. To read this data on machines which use the IEEE "little-endian" format such as
IBM-compatible PCs, the user will need to reverse the order of the bytes (i.e.,
byte-swap the data). The code segment below performs this byte swapping. Note
that the code segment below is the same as given above, but with the added feature of swapping the bytes.

```
C*******************************************************************
C       FORTRAN program segment to read a month of data and perform
C     byte swapping.
C
C       Once the header of size 1440 bytes (one data row) is skipped
C       (for the official release), there is always a month of grids.
C       All grids are of size 360x180 containing REAL*4 values.  All
C       days of data in the month are present, even if some have no
C       valid data.  Grid boxes without valid data are filled with
C       the (REAL*4) "missing" value -99999.  The bytes are swapped
C     after the data has been read and before it is output.  The GPCP
C     data are generated using a Silicon Graphics, Inc. Unix
C     workstation, which uses the "big-endian" IEEE 754-1985
C     representation of REAL*4 unformatted binary words.  To read this
C     data on machines which use the IEEE "little-endian" format such
C     as IBM-compatible PCs, the user will need to reverse the order
C     of the bytes (i.e., byte-swap the data).
C*******************************************************************
C
        IMPLICIT        NONE
        CHARACTER*80    fname
      REAL*4            varin, var
        REAL*4          datain (360, 180), data (360, 180)
        INTEGER         nday, nskip, iret, i, j
      CHARACTER*1 cvarin (4), cvar (4)
C
      EQUIVALENCE (cvarin, varin)
      EQUIVALENCE (cvar,   var)
C
C       Set the user input for file name and day number (using 19970121
C       as an example).
C
        fname = 'gpcp_1dd_p1d.199701'
        nday  = 21
C
C       Open the data file with a RECL of 1 data row.
C               ==>>    WARNING WARNING WARNING    <<==
C     The RECL is defined differently on different machines; it isn't
C     specified in the FORTRAN77 standard.  On SGI it's in 4-B words.
C     If you find that you only get 90 good values and then garbage
C     (either all zeros or random values) in the last 270 of the row,
C     your machine wants RECL in bytes, and you should say RECL=1440
C     in the following OPEN.
C
        OPEN  ( UNIT=10, FILE=fname, ACCESS='DIRECT',
     +          FORM='UNFORMATTED', STATUS='OLD', RECL=360,
     +            IOSTAT=iret )
        IF  ( iret .NE. 0 )  THEN
            WRITE (*, *) 'Error: open error', iret,
     +                   ' on file ', fname
            STOP
        END IF
C
C       Compute the number of records to skip, namely 1 for the
C       header and 180 for each intervening day.
C
        nskip = 1 + ( nday - 1 ) * 180
C
C       Read the 180 rows of data and close the file.
C
```

```
        DO  10 j = 1, 180
            READ ( UNIT=10, REC=j+nskip, IOSTAT=iret )
     +           ( datain (i, j), i = 1, 360 )
            IF  ( iret .NE. 0 )  THEN
                WRITE (*, *) 'Error: read error', iret,
     +                       ' on file ', fname
                STOP
            END IF
   10   END DO
        CLOSE ( UNIT=10 )
C
C     Now that the day of data has been read into the array, swap
C     the byte order.
C
      DO i = 1, 360
          DO j = 1, 180
            varin = datain (i, j)
            cvar (1) = cvarin (4)
            cvar (2) = cvarin (3)
            cvar (3) = cvarin (2)
            cvar (4) = cvarin (1)
            data (i, j) = var
          END DO
      END DO
C
C       Now array "data" is ready to be manipulated, printed, etc.
C       For example, dump the single day as unformatted direct:
C
        OPEN  ( UNIT=10, FILE='junk', ACCESS='DIRECT',
     +          FORM='UNFORMATTED', RECL=360, IOSTAT=iret )
        IF  ( iret .NE. 0 )  THEN
            WRITE (*, *) 'Error: open error', iret,
     +                   ' on file junk'
            STOP
        END IF
        DO  20 j = 1, 180
            WRITE ( UNIT=10, REC=j, IOSTAT=iret )
     +            ( data (i, j), i = 1, 360 )
            IF  ( iret .NE. 0 )  THEN
                WRITE (*, *) 'Error: write error', iret,
     +                       ' on file junk'
                STOP
            END IF
   20   END DO
        CLOSE ( UNIT=10 )
        STOP
        END
```
....

5. DEFINITIONS AND DEFINING ALGORITHMS

The *day definition* for this data set is the sum of all images given a nominal time in the specified Universal Coordinated Time (UTC, also known as GMT or Z) day.

Because the IR data are provided at 00Z, 03Z, ..., 21Z, the data set most closely describes the period starting at 2230Z the previous day to 2230Z of the day named. The alternative of splitting the 00Z images between days was more cumbersome and not clearly better in practice.
....

The *IR Tb histogram data set* is produced by the Geostationary Satellite Precipitation Data Centre (GSPDC) of the GPCP under the direction of J. Janowiak, located in the Climate Prediction Center, NOAA National Centers for Environmental Prediction, Washington, DC, 20233 USA. Each cooperating geostationary (geo) satellite operator (the Geosynchronous Operational Environmental Satellites, or GOES, United States; the Geosynchronous Meteorological Satellite, or GMS, Japan; and the Meteorological Satellite, or Meteosat, European Community) accumulates three-hourly infrared (IR) imagery. These are forwarded to GSPDC as 24-class histograms of Tb on a 1x1-deg lat/lon grid. The global geo-IR are then merged on a global grid.

In parallel, the NOAA-series low-earth-orbit (leo) satellite operator (United States) provides GPI values on a 1x1-deg lat/lon grid accumulated to the nearest 3-hrly time.

These data are used as input to 1DD processing.
....

The *GPROF fractional occurrence of precipitation* is computed as the ratio of the number of pixels with precipitation to the total number of valid pixels, both accumulated on a 0.5x0.5-deg lat/lon grid swath by swath.

The Goddard Profiling Algorithm (GPROF) Version 6.0 is based on Kummerow et al. (1996) and Olson et al. (1999). Summarizing, GPROF is a multichannel physical approach for retrieving rainfall and vertical structure information from satellite-based passive microwave observations (here, SSM/I). Version 6 applies a Bayesian inversion method to the observed microwave brightness temperatures using an extensive library of cloud-model-based relations between hydrometeor profiles and microwave brightness temperatures. Each hydrometeor profile is associated with a surface precipitation rate. GPROF includes a procedure that accounts for inhomogeneities of the rainfall within the satellite field of view. Over land and coastal surface areas the algorithm reduces to a scattering-type procedure using only the higher-frequency channels. This loss of information arises from the physics of the emission signal in the lower frequencies when the underlying surface is other than all water.

These data are used as input to 1DD processing.
....

The *TOVS precipitation estimate* is computed from a regression relationship between collocated rain gauge measurements and several TOVS-based parameters that relate to cloud volume: clout-top pressure, fractional cloud cover, and relative humidity profile. The relation-ship is allowed to vary seasonally and latitudinally. The data are delivered for daily nodes (ascending|descending) on a 1x1-deg grid. Susskind and Pfanendtner (1989) and Susskind et al. (1997) contain more details.

These data are used as input to 1DD processing.
....

The *GPI precipitation product* is produced by the Geostationary Satellite Precipitation Data Centre of the GPCP under the direction of J. Janowiak, located in the Climate Prediction Center, NOAA National Centers for Environmental Prediction, Washington, DC, 20233 USA. Each cooperating geostationary satellite operator (the Geosynchronous

Operational Environmental Satellites, or GOES, United States; the Geosynchronous Meteorological Satellite, or GMS, Japan; and the Meteorological Satellite, or Meteosat, European Community) accumulates three-hourly infrared (IR) imagery which are forwarded to GSPDC. The global IR rainfall estimates are then generated from a merger of these data using the GOES Precipitation Index (GPI); Arkin and Meisner, 1987) technique, which relates cold cloud-top area to rain rate.

The GPI technique is based on the use of geostationary satellite IR observations. Colder IR brightness temperatures are directly related to higher cloud tops, which are loosely related to increased precipitation rates. From data collected during the Global Atmospheric Research Programme (GARP) Atlantic Tropical Experiment (GATE), an empirical relationship between brightness temperature and precipitation rate was developed. For a brightness temperature <= 235K, a rain rate of 3 mm/hour is assigned. For a brightness temperature > 235K, a rain rate of 0 mm/hour is assigned. The GPI works best over space and time averages of at least 250 km and 6 hours, respectively, in oceanic regions with deep convection.

For the period 1986-March 1998 the GPI data are accumulated on a 2.5x2.5-deg lat/lon grid for pentads (5-day periods), preventing an exact computation of the monthly average. We assume that a pentad crossing a month boundary contributes to the statistics in proportion to the fraction of the pentad in the month. For example, given a pentad that starts the last day of the month, 0.2 (one-fifth) of its samples are assigned to the month in question and and 0.8 (four-fifths) of its samples are assigned to the following month.

Starting with October 1996 the GPI data are accumulated on a 1x1-deg lat/lon grid for individual 3-hrly images. In this case monthly totals are computed as the sum of all available hours in the month.

In both data sets gaps in geo-IR are filled with low-earth-orbit IR (leo-IR) data from the NOAA series of polar orbiting meteorological satellites. However, the 2.5x2.5-deg data only contain the leo-IR used for fill-in, while the 1x1-deg data contain the full leo-IR. The latter allows a more accurate AGPI (see "AGPI precipitation product"). The Indian Ocean sector routinely lacked geo-IR coverage until Meteosat-5 was repositioned in June 1998.

The Version 2 GPI product is based on the 2.5x2.5-deg IR data for the period 1988-1996, and the 1x1-deg beginning in 1997. The boundary is set at January 1997 to avoid making the change during the 1997-1998 ENSO event.

These data are used as input to the GPCP Satellite-Gauge Precipitation Product and the leo-GPI are used as input to the TMPI.
....

The *AGPI precipitation product* is produced as part of the GPCP Version 2 Combined Precipitation Data Set by the GPCP Merge Development Centre (see section 2). The technique follows the Adjusted GPI (AGPI) of Adler et al. (1994).

During the SSM/I period (starting July 1987), separate monthly averages of approximately coincident GPI and merged SSM/I--TOVS precipitation estimates are formed by taking cut-outs of the 3-hourly GPI values that correspond most closely in time to the local overpass time of the DMSP platform. The ratio of merged SSM/I--TOVS to GPI averages is computed and controlled to prevent unstable answers. In regions of light precipitation an additive adjustment is computed as the difference between smoothed merged SSM/I--TOVS and ratio-adjusted GPI values when the merged SSM/I/TOVS is greater, and zero otherwise. The spatially varying arrays of adjustment coefficients are then applied to the full set of GPI estimates. In regions lacking geo-IR data, leo-GPI data are calibrated to the merged SSM/I--TOVS, then these calibrated leo-GPI are calibrated to the geo-AGPI. This two-step process tries to mimic the information contained in the AGPI, namely the local bias of the SSM/I and possible diurnal cycle biases in the geo-AGPI. The second step can only be done in regions with both geo- and leo-IR data, and then smooth-filled across the leo-IR fill-in. In the case of the 2.5x2.5-deg IR, which lacks leo-IR in geo-IR regions, the missing calibrated leo-GPI is approximated by smoothed merged SSM/I--TOVS for doing the calibration to geo-AGPI.

The period before mid-1987 is handled differently, but is not relevant to the 1DD.

These data are used as input to the GPCP Satellite-Gauge Precipitation Product.
....

The GPCP *satellite-gauge precipitation product* (SG) is produced as part of the GPCP Version 2 Combined Precipitation Data Set by the GPCP Merge Development Centre (GMDC). The technique is similar to the Version 1 described in Huffman et al. (1997). The monthly data are delivered on a 2.5x2.5-deg grid.

The combination of satellite data into a multi-satellite (MS) product is carried out differently during 3 periods according to data availability. Strong efforts were made to homogenize the data record:

The period mid-1987 to the present uses geo-IR, leo-IR, TOVS, and SSM/I data. This is the combination that contributes to the 1DD. TOVS is merged in with SSM/I where the SSM/I is suspect (outside about 45N-S) or missing. Then SSM/I and geo-IR are approximately time-matched to compute local coefficients to adjust the full geo-IR GPI to the bias of the SSM/I in the 40N-S band. As well, leo-IR GPI is approximately scaled to the SSM/I. The AGPI is built from geo-IR AGPI where possible and leo-IR AGPI elsewhere. The MS is composed of AGPI in the band 40N-S and the merged TOVS-SSM/I elsewhere.

The period before mid-1987 is handled differently, but is not relevant to the 1DD.

In each of the periods the MS and a gauge analysis are linearly combined into a satellite-gauge (SG) combination using weighting by inverse estimated mean-square error for each.

The 1DD values are calibrated to sum to the SG on a monthly basis.
....

The Threshold Matched Precipitation Index (*TMPI*) provides GPI-like precipitation estimates in which both the IR Tb threshold and the conditional rain rate for raining pixels are set locally in time and space:

- The geo-IR Tb threshold (Tb0) is set locally from month-long
  accumulations of time/space coincident geo-IR Tb and GPROF-SSM/I-

  based fractional occurrence of precipitation.
- The conditional rain rate is set locally from the resulting frequency of Tb<Tb0 for the ENTIRE month and the GPCP SG.
- Separately, leo-IR GPI are processed like the TOVS (see "ATOVS") because they suffer a similar over-estimate of precipitation frequency.

The available geo-IR histograms in each 3-hrly global image are processed into precipitation estimates, and the adjusted leo-GPI data are used to fill holes in the individual 3-hrly geo-IR images. Then all the available images in a UTC day (00Z, 03Z, ..., 21Z) are averaged to produce the daily estimate (on a 1x1-deg grid).

These data are produced as intermediate files in 1DD processing.
....

The Adjusted TOVS (*ATOVS*) is produced outside 40N-S to make the 1DD globally complete. The Susskind et al. (1997) precipitation estimates from TOVS were considered to have too large a number of rain days, and we wanted to maintain consistency with the monthly GPCP SG. Accordingly, we revise the TOVS estimates by:

- computing the ratio of TMPI rain days to TOVS rain days separately for 39-40N and 39-40S;
- using the corresponding ratio in each grid box over the entire hemisphere to reduce the occurrence of TOVS precipitation by zeroing the (1-ratio) smallest daily TOVS rain accumulations; and
- rescaling the remaining (non-zero) TOVS rain days to sum to the monthly SG.

The daily data are gridded at 1x1 deg. These data are produced as intermediate files in 1DD processing.
....

The One-Degree Daily (*1DD*) precipitation data set is a first approach to estimating global daily precipitation at the 1x1-deg scale strictly from observational data. It is composed of TMPI where available (40N-S) and ATOVS elsewhere. The data boundaries at 40N and 40S do not exhibit serious problems, probably because both the TMPI and ATOVS are responding to CLOUD features. Nevertheless, smoothing was performed at the data boundaries as follows:

- The (TMPI-ATOVS) difference was computed at all grid boxes in the bands 39-40N and 39-40S.
- The difference field is smooth-filled between the actual differences at 40N and S and a zero value at 45N and S, respectively.
- The difference field is added to the ATOVS field.

The smoothing improves continuity across the data boundary at the expense of modifying the fractional occurrence field and causing some "feathering" along sharp features.
....

The *units of the 1DD estimates* are mm/day.
....

6. TEMPORAL AND SPATIAL COVERAGE AND RESOLUTION

The *file date* is the UTC month in which nominal times of the input data occur. Thus, a 00Z image is NOT divided between the two days it borders. All dates are UTC.

....

The *temporal resolution* of the products is one day. The temporal resolution of the original single-source data sets varies: 3-hrly for geo-IR, 3-hr accumulations for leo-IR, instantaneous for GPROF fractional occurrence, monthly for GPCP SG, and daily for TOVS. Some of the single-source data sets are available from other archives at a finer resolution.

....

The *period of record* for the GPCP 1DD is October 1996 through January 2004. The start is based on the availability of 3-hrly IR data. The end is based on the availability of input analyses, and will be extended in future releases. Some of the single-source data sets have longer periods of record in their original archival sites.

....

The *grid* on which each field of values is presented is a 1x1-deg latitude--longitude (Cylindrical Equal Distance) global array of points. It is size 360x180, with X (longitude) incrementing most rapidly West to East from the Prime Meridian, and then Y (latitude) incrementing North to South. Whole- and half-degree values are at grid edges:

```
First point center  = (89.5N,0.5E)
Second point center = (89.5N,1.5E)
Last point center   = (89.5S,0.5W)
```

....

The *spatial resolution* of the products is 1x1-deg lat/long. Some of the single-source data sets are available from other archives at a finer resolution.

....

The *spatial coverage* of the products is global in the sense that they are provided on a global grid. However, a small scattering of points are missing at the poles due to the forward regridding used in TOVS. In the current version these missing values are filled by average-filling the final fields. This step also fills in small areas of missing, particularly near the poles, but they are rare and the filling is considered beneficial. See "processing change" and "known anomalies" for more discussion of the average-fill issue.

....

7. PRODUCTION AND UPDATES

The GPCP is responsible for managing *production and updates* of the GPCP 1DD Data Set (WCRP 1986). The 1DD is produced by the GPCP Merge Development Centre (GMDC), located at NASA Goddard Space Flight Center in the Laboratory for Atmospheres.

Various groups in the international science community are given the tasks of preparing precipitation estimates from individual data sources, then the GMDC is charged with combining these into a "best" global product. This activity takes place after real time, at a pace governed by agreements about forwarding data to the individual centers and activities designed to ensure the quality in each processing step. The techniques used to compute the combined estimates are described in section 5.

Updates will be released to (1) extend the data record, (2) take advantage of improved combination techniques, or (3) correct errors. Updates resulting from the last two cases will be given new version numbers.

```
-->  NOTE: The changes described in this section are typical of the <--
-->  changes that are required to keep the GPCP Combined            <--
-->  Precipitation Data Set abreast of current requirements and     <--
-->  science.  Users are stongly encouraged to check back routinely <--
-->  for additional upgrades, and to refer other users to this site <--
-->  rather than redistributing data that are potentially out of    <--
-->  date.                                                          <--
```

At the present the GMDC and the GPCP are working to improve the 1DD by directly incorporating high-quality microwave estimates into the individual 3-hourly fields. The high-quality estimates will likely include SSM/I, TRMM, Advanced Microwave Sounding Unit (AMSU), and Advanced Microwave Scanning Radiometer (AMSR) estimates.
....

To date, one *processing change* has been implemented in the 1DD code. The algorithm to fill in the gaps in daily TOVS data has been modified from an iterative smooth-fill scheme, which competely fills all gaps, to a boxcar average which preserves large gaps. The new boxcar averaging scheme uses a latitude-based 3x3 grid on which the east-west size of the box approximately matches the north-south size. This change was motivated by large gaps that arose in the TOVS data during the span September 11-18, 2001. Users should be aware that gaps may occur throughout the 1DD data record.
....

8. SENSORS

The Special Sensor Microwave/Imager (*SSM/I*) is a multi-channel passive microwave radiometer that has flown on selected Defense Meteorological Satellite Program (DMSP) platforms since mid-1987. The DMSP is placed in a sun-synchronous polar orbit with a period of about 102 min. The SSM/I provides vertical and horizontal polarization values for 19, 22, 37, and 85 GHz frequencies (except only vertical at 22) with conical scanning. Pixels and scans are spaced 25 km apart at the suborbital point, except the 85-GHz channels are collected at 12.5 km spacing. Every other high-frequency pixel is co-located with the low-frequency pixels, starting with the first pixel in the scan and the first scan in a pair of scans. The channels have resolutions that vary from 12.5x15 km for the 85 GHz to 60x75 km for the 19 GHz. Pixels are oval due to the slanted viewing angle.

The polar orbit provides nominal coverage over the latitudes 85N-S, although limitations in retrieval techniques prevent useful precipitation estimates in cases of cold land (scattering), land (emission), or sea ice (both scattering and emission).

The SSM/I is an operational sensor, so the data record suffers the usual gaps in the record due to processing errors, down time on receivers, etc. Over time the coverage has improved as the operational system has matured. As well, the first 85 GHz sensor to fly degraded quickly due to inadequate solar shielding. After launch in mid-1987, the 85 GHz vertical- and horizontal-polarization channels became unusable in 1989 and 1990, respectively.

Further details are available in Hollinger et al. (1990).

The GPROF fractional occurrence estimates are based on all available estimates, which have included:

Satellite	Record Period	Months
F08	1987/07/09 - 1987/11/30	31
	1988/01/13 - 1990/02/28	
F10	1991/01/01 - 1997/11/14	83
F11	1992/01/01 - 1996/12/31	60
F13	1995/05/03 - Current	53
F14	1997/05/08 - Current	29

....

The infrared (*IR*) data are collected from a variety of sensors. The primary source of IR data is the international constellation of geosynchronous-orbit meteorological satellites -- the Geosynchronous Operational Environmental Satellites (GOES, United States), the Geosynchronous Meteorological Satellite (GMS, Japan), and the Meteorological Satellite (Meteosat, European Community). There are usually two GOES platforms active, GOES-EAST and -WEST, which cover the eastern and western United States, respectively. Gaps in geosynchronous coverage (most notably over the Indian Ocean prior to June 1998) must be filled with IR data from the NOAA-series polar-orbiting meteorological satellites. The geosynchronous data are collected by scanning (parts of) the earth's disk, while the polar-orbit data are collected by cross-track scanning. The data are accumulated for processing from full-resolution (4x8 km) images.

Starting with October 1996 the IR data are accumulated on a 1x1-deg lat/lon grid for individual 3-hrly images. In parallel, all low-earth orbit IR (leo-IR) data from the NOAA series of polar orbiting meteorological satellites are accumulated to the nearest 3-hr time on a 1x1-deg grid.

The various IR instruments are operational sensors, so the data record suffers the usual gaps in the record due to processing errors, down time on receivers, sensor failures, etc.

Further details are available in Janowiak and Arkin (1991).

....

The Television-Infrared Operational Satellite (TIROS) Operational Vertical Sounder (*TOVS*) data are collected from the High-Resolution Infrared Sounder 2 (HIRS2), Microwave Sounding Unit (MSU), and Stratospheric Sounding Unit (SSU) instruments on the NOAA series of polar orbiting meteorological satellites. There are usually two such satellites with orbits roughly in quadrature. The data are accumulated for processing at varying resolutions, but the effective resolution of the retrievals is about 60x60 km.

The various instruments contributing to TOVS are operational sensors, so the data record suffers the usual gaps in the record due to processing errors, down time on receivers, sensor failures, etc.

Further details are available in Susskind et al (1997).

....

9. ERROR DETECTION AND CORRECTION

SSM/I error detection/correction has several parts. Built-in hot- and cold-load calibration checks are used to convert counts to Antenna Temperature (Ta). An algorithm has been developed to convert Ta to Brightness Temperature (Tb) for the various channels (eliminating cross-channel leakage). As well, systematic navigation corrections are performed. All pixels with non-physical Tb and local calibration errors are deleted.

Accuracies in the Tb's are within the uncertainties of the precipitation estimation techniques. For the most part, tests show only small differences among the SSM/I sensors flying on different platforms.

Some satellites experienced significant drifting of the equator-crossing time during their period of service. There is no direct effect on the accuracy of the SSM/I data, but it is possible that the systematic change in sampling time could introduce biases in the resulting precipitation estimates.
....

The dominant *IR data correction* is for slanted paths through the atmosphere. Referred to as "limb darkening correction" in polar-orbit data, or "zenith-angle correction" (Joyce et al. 2001) in geosynchronous-orbit data, this correction accounts for the fact that a slanted path through the atmosphere increases the chances that (cold) cloud sides will be viewed, rather than (warm) surface, and raises the altitude dominating the atmospheric emission signal (almost always lowering the equivalent Tb). In addition, the various sensors have a variety of sensitivities to the IR spectrum, usually including the 10-11 micron band. Inter-satellite calibration differences are documented, but they are not implemented in the current version. They are planned for a future release. The TMPI largely corrects inter-satellite calibration, except for small effects at boundaries between satellites. The satellite operators are responsible for detecting and eliminating navigation and telemetry errors.
....

The *TOVS data correction* mostly depends on the retrieval system generating self-consistent sets of parameters. The requirement for consistency is intended to limit the effect of calibration drift and other faults in the data.
....

There is no *SG error correction* per se. The procedure is designed to minimize bias and throw out outliers, but no post-processing is done. An estimate of RMS random error is attached to each field.
....

Three types of *known errors* are contained in part or all of the current data set, and will be corrected in a future general re-run. They have been uncovered by visual inspection, but are considered too minor or insufficiently understood to provoke an immediate reprocessing.

1. The 270K maximum Tb bin limit is too cold in the subtropical highs.
2. Smoothing at the 40N and S data boundaries is primitive and affects the fractional occurrence over 40-50N and S.
3. There is no inter-satellite calibration applied to the IR data.
....

Some *known anomalies* in the data set are documented and left intact at the discretion of the data producers. The only current anomaly occurs due to a change in the treatment of gaps in the TOVS data. Prior to the September 2001 data, the gaps in the TOVS were filled by an iterative smooth-fill algorithm. These gaps tended to be small and a smooth-fill scheme produced realistic precipitation estimates. During the span September 11-18, 2001, large gaps appeared as a result of a known satellite attitude problem. The smooth-fill algorithm was replaced with a latitude-dependent boxcar average scheme which filled smaller gaps but maintained the larger gaps. These gaps are preserved in the final 1DD product. Currently, the TOVS data producer is assessing the feasibility of recovering the missing September data. If the data are recovered, the September GPCP data will be regenerated and distributed.
....

10. MISSING VALUE ESTIMATION AND CODES

There is generally no effort to *estimate missing values* in the single-source data sets.
....

All products in the GPCP 1DD Data Set use the *standard missing value* '-99999.'
....

All *missing days* are filled entirely with the standard missing value, so that the day number and the position of the day in the file always agree, and there are always complete sequences of day files for the month.
....

11. QUALITY AND CONFIDENCE ESTIMATES

The *accuracy* of the precipitation products can be broken into systematic departures from the true answer (bias) and random fluctuations about the true answer (sampling), as discussed in Huffman (1997). The former are the biggest problem for climatological averages, since they will not average out. However, on the daily time scale the low number of samples and/or algorithmic inaccuracies tend to present a more serious problem. That is, for leo-IR and TOVS the sampling is spotty enough that the collection of values over one day may not be representative of the true distribution of precipitation over the day. For geo-IR, leo-IR, and TOVS the algorithms likely have substantial RMS error.

Accordingly, the "random error" is assumed to be dominant, and estimates could be computed as discussed in Huffman (1997). Random error cannot be corrected.

The "bias error" should average to that of the GPCP SG for the month, and we assume the latter is small, or at least contained. This is particularly true over land, where the SG is adjusted to the large-area bias of the rain gauges. Studies of the submonthly bias have not yet been carried out. Note that regions suffering Tb threshold saturation (mostly subtropical highs) will likely suffer a low bias in order to keep the conditional precipitation rate reasonable. However, totals should be modest, so this is considered tolerable.

For the period January 1979 - February 1999, the TOVS precipitation estimates are based on at least two NOAA satellites flying in quadrature. Beyond February 1999, the TOVS estimates are based only on NOAA-14. The quantitative impact of using a single NOAA satellite for the TOVS Precipitation estimates on the 1DD product is currently unknown. Qualitatively, it does lead to salt-and-pepper missings over the globe, again due to forward regridding. When the NOAA-15 data processing stream is implemented, the TOVS precipitation beginning February 1999 will be reprocessed forward using two satellites and incorporated into the 1DD product.

....

The 1DD *intercomparison results* are still being developed. The time series of the global images shows good continuity in time and spatially across the data boundaries. An early validation against the Oklahoma Mesonet by the Surface Reference Data Center appears to show underestimation during the spring and fall (by about 20 and 15%, respectively), and overestimation during the summer (by about 20%). Mean absolute error (correlation) peaks (is minimum) in summer and is a minimum (peaks) in winter. An independent study of large-area averages over the Baltic drainage basin show reasonable behavior in all seasons (Rubel and Rudolph, 1999). Overall, the 1DD appears to be working as expected in both the TMPI and TOVS data (Oklahoma and the Baltic, respectively).

Huffman et al. (2001) and recent issues of the "Validator" (available from the SRDC) contain additional statistics.

....

12. DATA ARCHIVES

The *archive and distribution sites* for the official release of the GPCP 1DD Precipitation Data Set are:

Mr. David Smith
World Data Center A (WDC-A)
National Climatic Data Center (NCDC)
Rm 120
151 Patton Ave.
Asheville, NC 28801-5001 USA
Phone: 828-271-4053
Fax: 828-271-4328
Internet: dsmith@ncdc.noaa.gov
WDC-A Home Page: http://lwf.ncdc.noaa.gov/oa/wmo/wdcamet.html#GPCP

Dr. Bruno Rudolf
Global Precipitation Climatology Centre (GPCC)
Deutscher Wetterdienst (DWD)
Postfach 10 04 65
D-63004 Offenbach a.M., Germany
Phone: +49-69-8062-2765
Fax: +49-69-8062-2880
Internet: brudolf@dwd.d400.de
GPCC Home Page: http://www.dwd.de/research/gpcc

George J. Huffman
Code 912
NASA Goddard Space Flight Center
Greenbelt, MD 20771 USA
Phone: 301-614-6308
Fax: 301-614-5492
Internet: huffman@agnes.gsfc.nasa.gov

MAPB Precipitation Page: http://precip.gsfc.nasa.gov

Independent archive and distribution sites exist for the single-source data sets, and a current list may be obtained by contacting Dr. Huffman.
....

13. DOCUMENTATION

The *documentation creator* is:

Dr. George J. Huffman
Code 912
NASA Goddard Space Flight Center
Greenbelt, MD 20771 USA
Phone: 301-614-6308
Fax: 301-614-5492
Internet: huffman@agnes.gsfc.nasa.gov
MAPB Precipitation Page: http://precip.gsfc.nasa.gov
....

The *documentation revision history* is:

November 12, 1999	Draft 1 by GJH
January 11, 2000	Rev. 1 by DTB
March 14, 2000	Rev. 2 by DTB
October 4, 2000	Rev. 3 by GJH
March 1, 2001	Rev. 4 by DTB
March 28, 2001	Rev. 5 by DTB
April 19, 2001	Rev. 6 by DTB
June 4, 2001	Rev. 7 by DTB
June 28, 2001	Rev. 8 by DTB
August 1, 2001	Rev. 9 by DTB
August 18, 2001	Rev. 10 by DTB
September 4, 2001	Rev. 11 by DTB
October 17, 2001	Rev. 12 by DTB
November 2, 2001	Rev. 13 by DTB
February 2, 2002	Rev. 14 by DTB
February 5, 2002	Rev. 15 by DTB
February 14, 2002	Rev. 16 by GJH/DTB
March 29, 2002	Rev. 17 by DTB
April 4, 2002	Rev. 18 by DTB
May 22, 2002	Rev. 19 by DTB
July 3, 2002	Rev. 20 by GJH
July 16, 2002	Rev. 21 by DTB
July 31, 2002	Rev. 22 by DTB
August 26, 2002	Rev. 23 by DTB
September 3, 2002	Rev. 24 by DTB
September 10, 2002	Rev. 25 by DTB
September 26, 2002	Rev. 26 by DTB
October 31, 2002	Rev. 27 by DTB
January 10, 2003	Rev. 28 by DTB
March 3, 2003	Rev. 29 by DTB
July 28, 2003	Rev. 30 by DTB
January 14, 2004	Rev. 31 by DTB
January 27, 2004	Rev. 32 by DTB
February 3, 2004	Rev. 33 by DTB
February 20, 2004	Rev. 34 by DTB
February 25, 2004	Rev. 35 by DTB
March 15, 2004	Rev. 36 by DTB
April 21, 2004	Rev. 37 by DTB

The latest version includes data through January 2004.
....

The list of *references* used in this documentation is:

Grody, N.C., 1991: Classification of snow cover and precipitation using the Special Sensor Microwave/Imager (SSM/I). J. Geophys. Res., 96, 7423-7435.

Hollinger, J.P., J.L. Pierce, and G.A. Poe, 1990: SSM/I instrument evaluation. IEEE Trans. Geosci. Remote Sens., 28, 781-790.

Huffman, G.J., 1997: Estimates of root-mean-square random error contained in finite sets of estimated precipitation. J. Appl. Meteor., 36, 1191-1201.

Huffman, G.J., R.F. Adler, P.A. Arkin, A. Chang, R. Ferraro, A. Gruber, J.J. Janowiak, R.J. Joyce, A. McNab, B. Rudolf, U. Schneider, and P. Xie, 1997: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Data Set. Bull. Amer. Meteor. Soc., 78, 5-20.

Huffman, G.J., R.F. Adler, M.M. Morrissey, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multi-satellite observations. J. Hydrometeor., 2, 36-50.

Janowiak, J.E., and P.A. Arkin, 1991: Rainfall variations in the tropics during 1986-1989. J. Geophys. Res., 96, 3359-3373.

Kummerow, C., W.S. Olsen, and L. Giglio, 1996: A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors. IEEE Trans. Geosci. Remote Sens., 34, 1213-1232.

Olsen, W.S., C.D. Kummerow, Y. Hong, and W.-T. Tao, 1999: Atmospheric latent heating distributions in the tropics from satellite passive microwave radiometer measurements. J. Appl. Meteor., 38, 633-644.

Rubel, F., and B. Rudolf, 1999: Verification of GPCP's satellite-based daily precipitation analyses using corrected rain gauge data of BALTEX for the Baltic Sea catchment area. Veterinärmedizinische Universität Wien, http://www-med-physik.vu-wien.ac.at/staff/rub/pro/GPCC/applications.htm

Spencer, R.W., 1993: Global oceanic precipitation from the MSU during 1979-92 and comparisons to other climatologies. J. Climate, 6, 1301-1326.

Susskind, J., and J. Pfaendtner, 1989: Impact of interactive physical retrievals on NWP. Report on the Joint ECMWF/EUMETSAT Workshop on the Use of Satellite Data iomn Operational Weather Prediction: 1989-1993, Vol. 1, T. Hollingsworth, Ed., ECMWF, Shinfield Park, Reading RG2 9AV, U.K., 245-270.

Susskind, J., P. Piraino, L. Rokke, L. Iredell, and A. Mehta, 1997: Characteristics of the TOVS Pathfinder Path A Dataset. Bull. Amer. Meteor. Soc., 78, 1449-1472.

WCRP, 1986: Report of the workshop on global large scale precipitation data sets for the World Climate Research Programme. WCP-111, WMO/TD - No. 94, WMO, Geneva, 45 pp.

Wentz, F.J., and R.W. Spencer, 1998: SSM/I rain retrievals within a unified all-weather ocean algorithm. J. Atmos. Sci., 55, 1613-1627.
....

14. INVENTORIES

The *data set inventory* may be obtained by accessing the home pages or contacting the representatives listed in section 12.
....

15. HOW TO ORDER DATA AND OBTAIN INFORMATION ABOUT THE DATA

Users interested in *obtaining data* should access the home pages or contact the representatives listed in section 12.
....

The *data access policy* is "freely available" with three common-sense caveats:

1. The data set source should be acknowledged when the data are used. [One possible wording is: "The 1DD data were provided by the NASA/Goddard Space Flight Center's Laboratory for Atmospheres, which develops and computes the 1DD as a contribution to the GEWEX Global Precipitation Climatology Project."]
2. New users should obtain their own current, clean copy, rather than taking a version from a third party that might be damaged or out of date.
3. Errors and difficulties in the data set should be reported to the data set creators.
....