Data Center Report
Global Precipitation Climatology Centre (GPCC)

U. Schneider, A. Becker, M. Ziese, Elke Rustemeier, P. Finger

Global Precipitation Climatology Centre
Deutscher Wetterdienst
Department of Hydrometeorology

Frankfurter Straße 135, 63067 Offenbach am Main, Germany
E-Mail: gpcc@dwd.de ; Web: gpcc.dwd.de
Outline

1. GPCC background
2. Data base and data sources available (near and non real-time)
3. Quality-control processing at GPCC
4. GPCC's portfolio of precipitation analysis products
5. Outlook
GPCC background

- **GPCC** was established at the beginning of 1989 at Deutscher Wetterdienst (DWD) on invitation by **WMO** (the need for global precipitation data sets); in operation for almost 30 years now.

- **GPCC**’s main task is the archiving and the analysis of precipitation on the basis of in-situ data for the land-surface.

- **GPCC** is contributing to **GEWEX** (Global Energy and Water Exchanges Project) via GHP and GDAP and to the Global Climate Observing System (**GCOS**).
GPCC background

Data delivered in different formats

All data in same format

Data stored in data bank

Extracted data for analyses
Outline

1. GPCC background
2. Database and data sources available (near and non real-time)
3. Quality-control processing at GPCC
4. GPCC’s portfolio of precipitation analysis products
5. Outlook

GEWEX GHP Meeting, 25 Oct 2018, Santiago de Chile, Chile
GPCC data sources

Near real-time (GTS):
- GTS SYNOP (DWD RTH Offenbach)
- GTS CLIMAT (DWD RTH Offenbach)
- GTS CLIMAT (JMA RTH Tokyo)
- GTS CLIMAT (UKMO RTH Exeter)
- SYNOP-based (NOAA RTH Washington)

Non real-time:
- Additional data from ca. 190 countries
- International project data (GEWEX-related and other)
- Historical data collections (CRU, FAO, GHCN, ECA&D and compilations from research projects)
GPCC monthly data base (source specific)

GPCC Monthly Precipitation Database
accumulated number of records and sources, status April 2018

- **GPCC Full**
- **National**
- **GHCN** (GHCN V.2 + suppl. from GHCN daily)
- **FAO**
- **CRU**
- **Regional**
- **ECAD**
- **GTS**

GTS includes SYNOP-based GPCC and CPC data and CLIMAT

GEWEX GHP Meeting, 25 Oct 2018, Santiago de Chile, Chile
With the beginning of 2012 the GPCC started with the acquisition, processing and analysis of daily precipitation data.

- Almost from the start of this new activity the GPCC ran into the problem “How is the day defined?”

- Daily precipitation - generally being observed at about 07:00 local time - should be assigned to the previous day (most of the accumulation period is lying in the previous day).

- Unfortunately this is not done consistently in the different countries; most countries are assigning the daily totals to the previous day, but others are doing this differently (i.e. assigning precip to the day when the observation is taken).
GPCC daily data base

GPCC calculates the daily precip totals from the SYNOP reports in a consistent way (used for First Guess Daily and Full Data Daily)

GPCC is checking and correcting this, as far as possible, upon integration of GHCN daily and of national data sets into its data base
GPCC daily data base (source specific)

GPCC Daily Precipitation Database
(accumulated number of records and sources, status May 2018)

Number of stations

Full Data Daily V.2 (June 2018)
Full Data Daily V.1 (Sept. 2015)

For First Guess Daily

GEWEX GHP Meeting, 25 Oct 2018, Santiago de Chile, Chile
Outline

1. GPCC background
2. Data base and data sources available (near and non real-time)
3. Quality-control processing at GPCC
4. GPCC’s portfolio of precipitation analysis products
5. Outlook
Schematic diagram of QC processing at GPCC (part I)

Multi-level QC processing

Near real-time (GTS data)
- SYNOP (DWD, CPC)
  - calculation of monthly totals
- CLIMAT (DWD, JMA, UKMO)

Non real-time (data acquisition by WMO support letters and bilateral contacts)
- Historical data collections (i.e., FAO, GHCN, CRU)
  - individual data deliveries from the countries

Precontrol and reformatting
- uniform data sets

Data loading module and metadata identification, harmonization
- Check against background statistics (since 2003)

RDBMS "data bank"

Editor
- Check of file structure: format identification
- Data set visualization: consistency checks
- Review of questionable metadata: individual check and correction

Editor
-GIS
Schematic diagram of QC processing at GPCC (part II)

GPCC’s analysis products are generated on the basis of the qc‘ed data in the RDBMS (relational data base management system)
Checks of precipitation data in the data base

Examples for typical errors detected in the QC processing:

• Data sometimes shifted by a month, or even a year
• Factor*10 errors in precipitation rates
• Typing or coding errors
• Errors in the conversion of inch, mm etc. (mostly with historical data)
• Incorrect flagging of missing precipitation observations (might be misinterpreted as „0“)

Data from different sources are stored separately in the RDBMS, being very helpful in the QC processing
QC of precipitation data - Summary

• Almost every large data set is containing more or less frequently erroneous data

• “Bad data” should not simply be thrown away, but corrected where possible (data errors are often obvious and thus can be corrected (data maybe important in data sparse areas)

Important:

• True extreme values must not be eliminated by “QC” (therefore semi-automatic QC at GPCC; automatic pre-checks and visual control)

• Corrected data always archived together with the original data

➢ Careful data QC is absolutely necessary !!
GPCC Function/Capability
2. QA and QC to make data reliable and usable
Since 2\textsuperscript{nd} of May 2017, Guest Scientist Dr. Yu Yu (CMA) stays with GPCC for 6 Months to study CMA precipitation data contributions and GPCC methods. Step 1: Compare V1 & V2 of CMA’s contribution (Yu Yu, 2017)

<table>
<thead>
<tr>
<th>NO.</th>
<th>Number of Stations</th>
<th>Temporal Extent</th>
<th>Production Time</th>
<th>Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>212</td>
<td>1981-2010</td>
<td>Dec. 2011</td>
<td>NMIC, CMA</td>
</tr>
<tr>
<td>V2</td>
<td>641</td>
<td>1951-2016</td>
<td>Apr. 2017</td>
<td></td>
</tr>
</tbody>
</table>

Use Case
1. Global precipitation and drought monitoring (also for Copernicus EMS)
2. Assessment of hydro-climate environment and developments also in context of political crises
3. Seasonal prediction of ENSO related precipitation
4. Centennial analysis of variability and trends
5. Reference data for satellite monitoring across land
6. Potable water resources
7. Arctic Precipitation
8. Wet Deposition

Top: Joint and unique V1 stations
Left: Comparison of joint collective

GRDC-SC XIII, 22-23 June 2017, Koblenz, Germany
Status Report & Use Cases of the Global Precipitation Climatology Centre (GPCC)

GPCC Function/Capability

2. QA and QC to make data reliable and usable
Since, 2nd of May 2017, Guest Scientist Dr. Yu Yu (CMA) stays with GPCC for 6 Months to study CMA precipitation data contributions and GPCC methods. Step 2: Dr. Yu Yu & GPCC QC jointly process CMA data.

Use Case

1. Global precipitation and drought monitoring (also for Copernicus EMS)
2. Assessment of hydro-climate environment and developments also in context of political crises
3. Seasonal prediction of ENSO related precipitation
4. Centennial analysis of variability and trends
5. Reference data for satellite monitoring across land
6. Potable water resources
7. Arctic Precipitation
8. Wet Deposition

Step 3: A GPCC analysis across China, serving also CMA’s requirements

Distribution of 53184 AWS
Outline

1. GPCC background
2. Data base and data sources available (near and non real-time)
3. Quality-control processing at GPCC
4. GPCC’s portfolio of precipitation analysis products
5. Outlook
User requirements

- Features of gridded precipitation data as required by the users:
  - Timeliness (for drought monitoring)
  - High resolution (for regional structures in global maps)
  - High accuracy (for verification of model results)
  - Homogeneity (for climate change and variability analysis)

All of these requirements cannot be met by one single gridded data set

==> A portfolio of different analysis products has been designed and optimized with respect to the application purposes
Data base for different GPCC products

Data base GPCC's Products
(total number of stations used, status April 2018)

- Full Data Monthly V. 8
- Full Data Monthly V. 7
- Full Data Monthly V. 6
- Full Data Monthly V. 5
- Full Data Monthly V. 4
- Full Data Monthly V. 3

Number of Stations

Years:

GEWEX GHP Meeting, 25 Oct 2018, Santiago de Chile, Chile
Near real-time monthly precipitation products

- **GPCC** is providing the following gridded data sets on a quasi-operational basis (GTS data base):
  - A **First Guess** Analysis of monthly **Precipitation** available within 5 days after the end of the month via internet, used for **drought monitoring** (e.g. by FAO),
    - **Period:** Oct. 2003 to present
    - **Data base:** ~6,000-7,600 stations
  - The Precipitation **Monitoring Product** available within 2 months, used by GDAP/GPCP as **early in-situ reference** for the near real-time (**new V.6 !!**)
    - **Period:** Jan. 1982 (!) up to present
    - **Data base:** ~7,000-9,000 stations
Non real-time precipitation products

- From time to time (after significant enlargements / improvements of the data base) reanalyses of the following gridded products are being generated:
  - The Global Precipitation Climatology (over land) (V. May 2018), background for all other GPCC precipitation analysis products, based on ca. 79,200 stations with climatological normals (overall data of more than 116,000 stations in GPCC data base)
Station data base of GPCC’s Precipitation Climatology V.2018 as basis for anomaly analyses (number of stations: ca. 79,200)
GPCC's Precipitation Climatology V.2018 for July on a 0.25° grid
GPCC's Precipitation Climatology V.2018 for July on a 0.25° grid

Refined Land-Sea Mask and 0.25° resolution resolve
• Topographically induced patterns of the climatology
• Many pacific islands
Non real-time precipitation products

From time to time (after significant enlargements / improvements of the data base) reanalyses of the following gridded products are being generated:

- The **Global Precipitation Climatology (over land) (May 2018)**, background for all other GPCC precipitation analysis products, based on ca. 79,200 stations with climatological normals (overall data of more than 116,000 stations in GPCC data base)

- The **Full Data Monthly V.2018 (V.8, June 2018)**, optimized for high spatial resolution and accuracy, used for model verification and hydrological studies
  
  **Period:** Jan. 1891 up to Dec. 2016  
  **Data base:** 6,000-53,000 stations per month (overall 79,200 stations included)
Non real-time precipitation products

- Every data product is DOI referenced and features uncertainty information
  - Number of stations per grid cell supporting the analysis
  - Interpolation error according to Yamamoto (2000)
  - Standard deviation
  - An operator field to undo the statistical infilling applied

Statistical infilling illustration for May 1891

(c) GPCC 2018/05/16 number of stations: 5352

Without infilling
Statistical infilling illustration for May 1891

With infilling

(c) GPCC 2018/05/16 number of stations: 5352 + 14997
Daily precipitation analysis products

- GPCC is providing the following daily gridded data sets:

  • A First Guess Daily Analysis available within 5 days after the end of the month via internet
    Period: Jan. 2009 to present
    Data base: ca. 7,000-8,100 stations
    described in Schamm et al. (2014)

  • The Full Data Daily Analysis (V.2) updated from time to time
    Period: Jan. 1982 to 2016
    Data base: ca. 25,000-35,000 stations
Extreme events: Peru and Colombia

March/April 2017

Totals - 03.2017
Normals March

% of March Normals

1. Global precipitation and drought monitoring
2. Assessment of hydro-climate environment and developments also in context of political crises

Coastal El Nino in eastern equatorial Pacific’s SST

Limited representativeness of local scale heavy precipitation events like in Mocoa, Colombia

GRDC-SC XIII, 22-23 June 2017, Koblenz, Germany
GPCC has assembled a unique bi-decadal daily precipitation analysis data set together with CM_SAF@DWD

Use Case
5. Reference data for satellite monitoring across land
6. Potable water resources
7. Wet deposition of (radioactive) aerosols
8. Arctic Precipitation
GPCC Analysis to identify world-wide regions of ENSO sensitive average precipitation

Use Case
1. Global precipitation and drought monitoring
2. Assessment of hydro-climate environment and developments also in context of political crises
3. Seasonal prediction of ENSO related precipitation
4. Centennial analysis of variability and trends
5. Reference data for satellite monitoring across land
6. Potable water resources
7. Wet deposition of (radioactive) aerosols
8. Arctic Precipitation
GPCC Drought Index indicating water stress in California & Brazil (Sao Paolo)

GPCC drought analysis since early 1952 allows for cross-checks against historic data in the fields of economy, ecology, and migrations (climate refugees)

9 months look-backs since early 2014 (left), October 2014 (right)

Use Case
1. Global precipitation and drought monitoring
2. Assessment of hydro-climate environment and developments also in context of political crises
3. Seasonal prediction of ENSO related precipitation
4. Centennial analysis of variability and trends
5. Reference data for satellite monitoring across land
6. Potable water resources
7. Wet deposition of (radioactive) aerosols
8. Arctic Precipitation
GPCCs globally applicable drought index diagnosed since 1952
Change of drought risk 1952-2013

increased
decreased
Outline

1. GPCC background
2. Database and data sources available (near and non real-time)
3. Quality-control processing at GPCC
4. GPCC's portfolio of precipitation analysis products
5. Outlook
Outlook

• The QC of a huge data set from Brazil (>10000 stations) comes to completion

• The INTENSE data set (sub-daily data!) was provided to the GPCC and will also be integrated into GPCC‘s relational database management system (will enable studies of extreme events)
So far, collected hourly data from ~25,000 stations...

**Global dataset:** HadISD, approx. 10,000 stations (varying data quality, more useful data at 3h and 6h), freely available sub-daily precipitation data. Plus access to additional datasets (i.e. E Europe, China) to calculate indices.
Outlook

• The QC of a huge data set from Brazil (>10000 stations) comes to completion

• The INTENSE data set (sub-daily data!) was provided to the GPCC and will also be integrated into GPCC’s relational database management system (will enable studies of extreme events)

• Integration of satellite and radar information to improve interpolation (e.g. autocorrelation function for kriging)

• Intensified data acquisition in so far data scarce areas in collaboration with scientific community (Chris Funk, Sharon Nicolson, Hailey Fowler, Andreas Fink, etc…)

• BTW: GPCC serves as a backup for any scientific project that has raised data records of 10yrs length or longer
## Visualize and Download GPCC Products

<table>
<thead>
<tr>
<th>GPCC Product</th>
<th>Spatial Resolution</th>
<th>Time Coverage</th>
<th>Possible Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Guess Monthly</td>
<td>1.0°</td>
<td>2004 - present</td>
<td>drought monitoring</td>
</tr>
<tr>
<td>First Guess Daily</td>
<td>1.0°</td>
<td>2009 - present</td>
<td>analysis of extremes</td>
</tr>
<tr>
<td>Monitoring Version 6</td>
<td>1.0°, 2.5°</td>
<td>1982 - present</td>
<td>calibration of satellite data</td>
</tr>
<tr>
<td>Full Data Monthly Version 2018</td>
<td>0.25°, 0.5°, 1.0°, 2.5°</td>
<td>1891 - 2016</td>
<td>hydrological studies</td>
</tr>
<tr>
<td>Full Data Daily Version 2018</td>
<td>1.0°</td>
<td>1982 - 2016</td>
<td>analysis of extremes</td>
</tr>
<tr>
<td>HOAPS/GPCC global daily precipitation Version 1</td>
<td>0.5°, 1.0°, 2.5°</td>
<td>1988 - 2008</td>
<td>analysis of extremes</td>
</tr>
<tr>
<td>HOMPRA Europe Version 1</td>
<td>0.5°, 1.0°, 2.5°</td>
<td>1951 - 2005</td>
<td>trend analysis</td>
</tr>
<tr>
<td>Precipitation Climatology Version 2018</td>
<td>0.25°, 0.5°, 1.0°, 2.5°</td>
<td>1951/2000</td>
<td>for application as foundation, and for utilization of the area-to-point interpolation method</td>
</tr>
<tr>
<td>Interpolation Test Dataset</td>
<td>1.0°</td>
<td>1988</td>
<td>comparison of interpolation schemes</td>
</tr>
<tr>
<td>Drought Index Version 1</td>
<td>1.0°</td>
<td>2013 - present</td>
<td>drought monitoring</td>
</tr>
<tr>
<td>Drought Index Version 1.1</td>
<td>1.0°</td>
<td>1952 - 2013</td>
<td>drought monitoring</td>
</tr>
<tr>
<td>GPCC V</td>
<td></td>
<td></td>
<td>access to the GPCC Visualizer,</td>
</tr>
</tbody>
</table>

Thank you for your attention!


GEWEX GHP Meeting, 25 Oct 2018, Santiago de Chile, Chile