Global Precipitation Analysis Products of the GPCC

U. Schneider, A. Becker, A. Meyer-Christoffer, M. Ziese, B. Rudolf
Global Precipitation Climatology Centre (GPCC)
Deutscher Wetterdienst, Offenbach a. M., Germany, Dec. 2010

Introduction

Precipitation plays an important role in the global energy and water cycle. Accurate knowledge of precipitation amounts reaching the land surface is of special importance for fresh water assessment and management related to land use, agriculture and hydrology, incl. risk reduction of flood and drought. High interest in long-term precipitation analyses arises from the needs to assess climate change and its impacts on all spatial scales. Based on this demand national and international organizations initiated and support many research and monitoring programmes.

In this framework, the Global Precipitation Climatology Centre (GPCC) has been established in 1989 on request of the World Meteorological Organization (WMO). It is operated by the Deutscher Wetterdienst (DWD, National Meteorological Service of Germany) as a German contribution to the World Climate Research Programme (WCRP). Mandate of the GPCC is the global analysis of monthly precipitation on earth's landsurface based on *in situ* raingauge data. Since its start, the centre is the *in situ* component of the WCRP Global Precipitation Climatology Project (GPCP). In 1994, the long-term operation of the GPCC has been requested by WMO in order to contribute to the climate monitoring activities of the Global Climate Observing System (GCOS). Since 1999, GPCC is one of the two global GCOS Surface Network Monitoring Centers (GSNMCs) with special emphasis on precipitation.

The aim of the GPCC is to serve user requirements esp. with regard to accuracy of the gridded precipitation analyses and timeliness of the product availability. The WCRP Global Energy and Water Experiment (GEWEX) for instance requests high spatial resolution and accuracy for the last two decades, while the priority of GCOS and IPCC is focused on long-term homogeneous time-series. Timeliness of products is ensured by cut-off dates for data processing and analysis. All GPCC analysis products (except the 50-year VASClimO data set) result from the same quasi-operational data management and analysis system. However, depending on the requirements they differ with regard to the number of stations (data sources) included and the level of data quality-control being performed.

All GPCC products, gauge-based gridded monthly precipitation data sets for the global land surface, are available in spatial resolutions of 1.0° x 1.0° and 2.5° x 2.5° geographical latitude by longitude, non real-time products based on the complete GPCC monthly rainfall station database (the largest monthly precipitation station database of the world with data from ca. 85,000 different stations) are also available in 0.5° x 0.5° resolution. GPCC's new global precipitation climatology (available in 2.5° x 2.5°, 1.0° x 1.0°, 0.5° x 0.5°, and 0.25° x 0.25° resolution) based on data from ca. 64,400 stations is used as background climatology for GPCC analyses. Corresponding to international agreement, station data provided by Third Parties are protected. However the gridded GPCC analysis products are freely available via Internet (http://gpcc.dwd.de). In 2007 ca. 100,000 accesses on GPCC's Website have been counted.

The different products of the GPCC are used world-wide by various institutions, in particular in context of water- and climate-related research and monitoring activities of WMO, WCRP, GCOS, FAO (UN Food and Agriculture Organisation), UNESCO (UN Educational, Scientific and Cultural Organization) and GEO (Group on Earth Observations, see Figure 11).

GPCC's suite of gridded precipitation analysis products

- The new Global **Precipitation Climatology** is focussing on the period 1951-2000 and consists of data from ca. 64,400 stations. The climatology comprises normals collected by WMO (CLINOs), delivered by the countries to GPCC or calculated from time-series of monthly data (with at least 10 complete years of data) available in our data base. In case that time series of sufficient length (more than 40 years) for the period 1951-2000 were not available from a specific station, then climatological normals have also been calculated for 30-year reference periods 1961-1990, 1951-1980 or 1971-2000 with at least 20 years of data. If even this was not possible for a station, then normals have been calculated for the period 1931-1960, or for any other period with at least 10 complete years of data. Figure 1 displays the climatological mean precipitation for July as an example.

Note: GPCC's monthly precipitation analysis products described below are based on anomalies from climatological normals at the stations. The anomalies are spatially interpolated by using the analysis method SPHEREMAP and the gridded anomaly analyses are then superimposed on GPCC's new background climatology (except for VASClimO data set V1.1 using Kriging interpolation and an older climatology version).

- The **First Guess Product** of the monthly precipitation anomaly is based on interpolated precipitation anomalies from more than 6,000 stations worldwide. Data sources are synoptic weather observation reports (SYNOP) received at DWD via the WMO Global Telecommunication System (GTS) and climatic mean (mainly 1951-2000, or other reference periods as described before) monthly precipitation totals at the same stations extracted from GPCC's global normals collection. An automatic-only quality-control (QC) is applied to these data. Since September 2003, GPCC First Guess monthly precipitation analyses are available within 5 days after end of an observation month. Main application purpose is to serve as input for near-realtime drought monitoring applications, e.g. by the FAO and the University of London Hazard Research Centre. Fig. 2 shows a drought monitoring application based on First Guess analyses.

- The **Monitoring Product** of monthly precipitation for global climate monitoring is based on SYNOP and monthly CLIMAT reports received near-realtime via GTS from ca. 7,000 –8,000 stations (after high level QC) and is available within about 2 months after observation month. This is the GPCC product with the longest history: Operational monthly analysis started with year 1986 and has continuously been updated every month since then. The analyses are based on automatic and intensive manual quality control of the input data. The GPCC Monitoring Product is the *in situ* component to the satellite-gauge combined precipitation analyses of GPCP (Huffman et al. 1995, Adler et al. 2003) and of CMAP (Xie and Arkin 1997). It also supports regional climate monitoring. Figure 3 illustrates the heavy rainfall in Pakistan during the La Niña event in summer 2010.

- The **Full Data Reanalysis Product** is of much higher accuracy compared to the GPCC near real-time products mentioned above. Therefore, its application is recommended for hydrometeorological model verification and water cycle studies, e.g. in context of UNESCO, GEWEX, and GTN-H (Global Terrestrial Network for Hydrology). This analysis product is based on all stations, near real-time and non real-time, in the GPCC data base supplying data for the individual month. The GPCC Full Data Reanalysis Product Version 5 covers the period from 1901 to 2009 (see example in Figure 4); this new extended product version using the new GPCC climatology as analysis background was generated in Dec. 2010. The data coverage per month varies from some 10,000 to more than 47,000 stations (Figure 7). The full data re-analyses will be updated at irregular time intervals subsequent to significant data base improvements.

- The **VASClimO 50-Year Data Set** supplying gridded time-series of monthly precipitation for climate variability and trend studies is based on data being selected with respect to a (mostly) complete temporal data coverage and homogeneity of the time-series. The current version 1.1 is based on time-series of 9,343 stations covering the period 1951-2000

(Beck, Grieser and Rudolf, 2005). These long-term climatological analyses of homogenised area-averaged precipitation time-series are of special interest for GCOS and supported the Intergovernmental Panel on Climate Change (IPCC) Working Group I Fourth Assessment Report (FAR), published in year 2007. Figure 5 shows an application example: Linear trends of annual total precipitation for 1951-2000. The VASClimO Data Set will be replaced in 2011 by the new homogenized precipitation analysis HOMPRA covering the period 1951-2005.

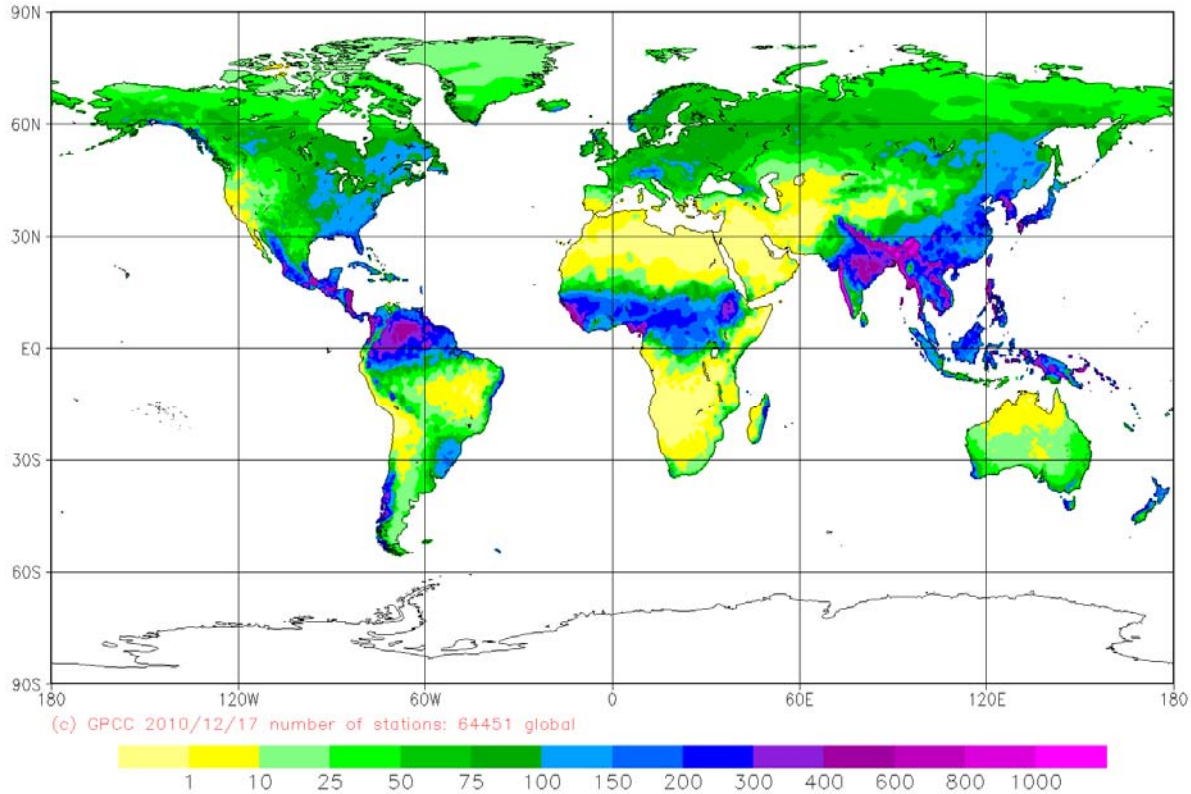

Figure 1: *Climatological mean precipitation for July (based on GPCC's new Precipitation Climatology V.2010 focussing on the period 1951-2000, 0.25° x 0.25° resolution)*

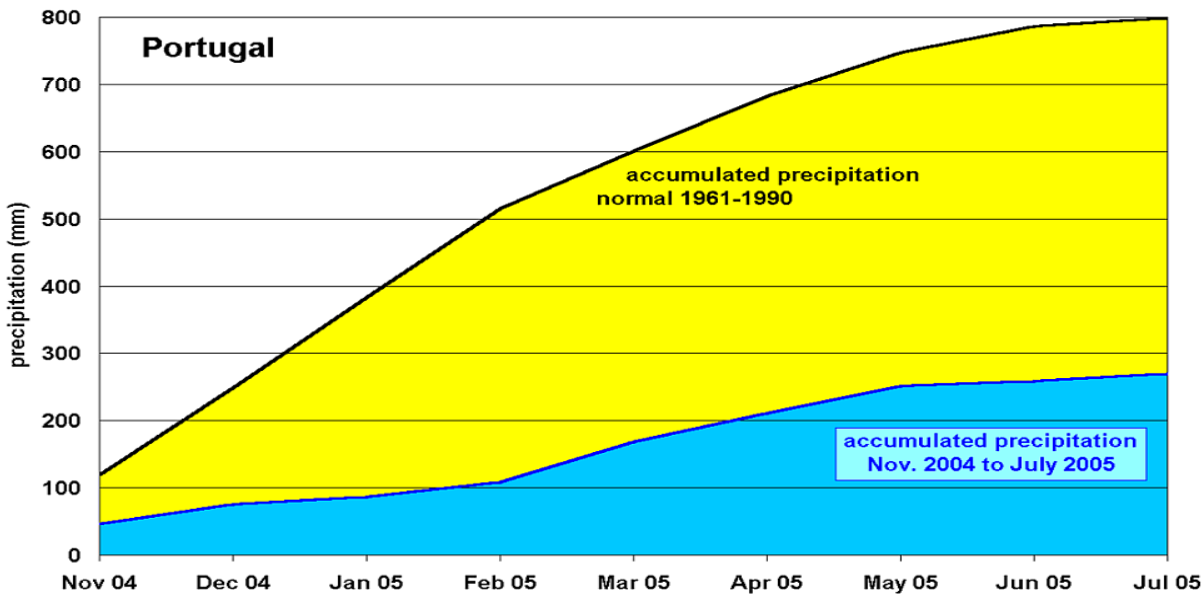

Figure 2: *Accumulated precipitation totals (based on GPCC First Guess) and accumulated GPCC precipitation normals 1961-1990 indicating an increasing precipitation deficiency in year 2005 in South West Europe*

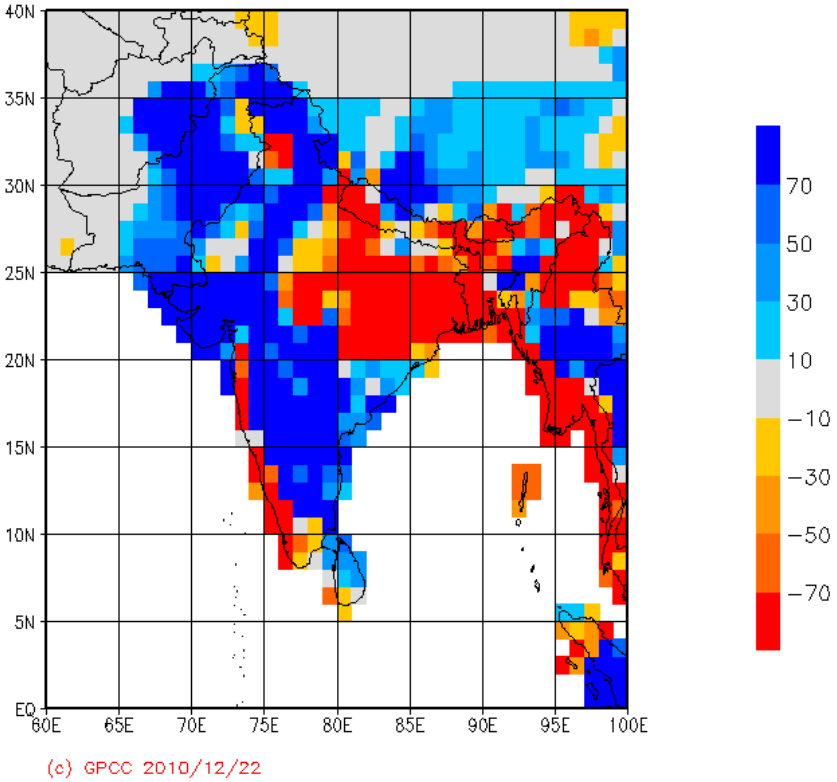

Figure 3: *Example of monthly precipitation anomalies in August 2010 (La Niña event) for Pakistan/ India (based on GPCC Monitoring product, 1.0° x 1.0° resolution)*

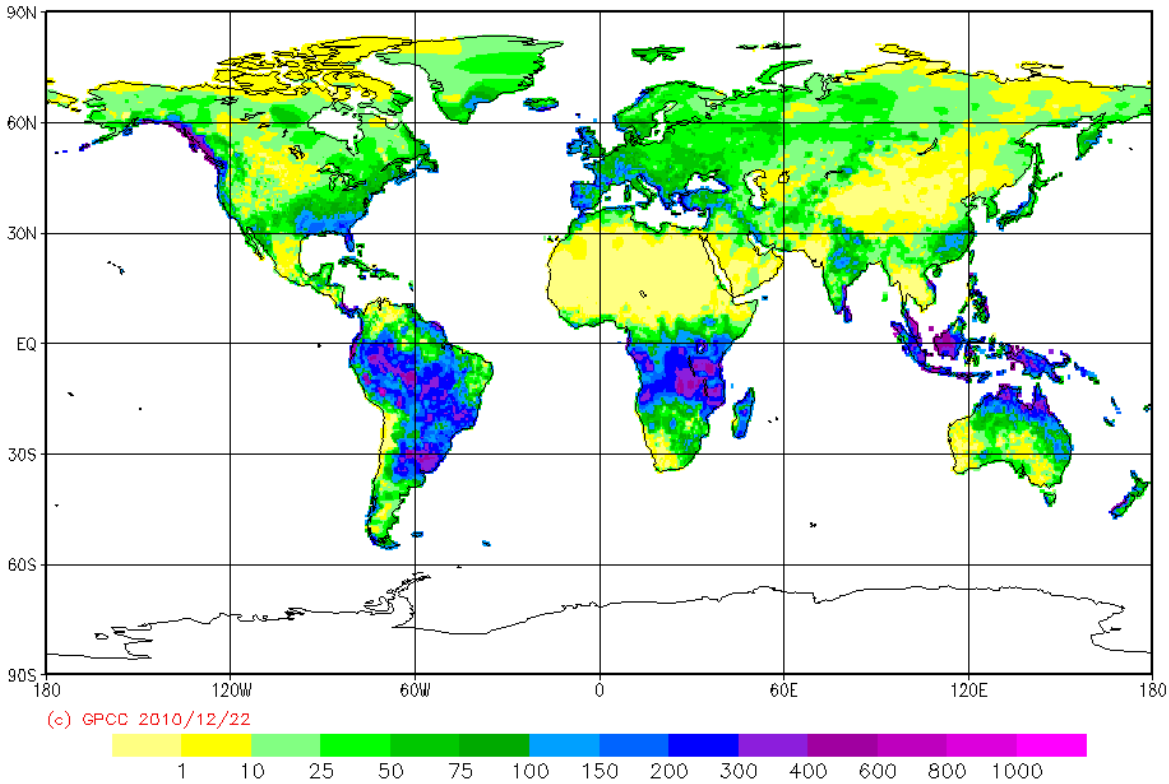

Figure 4: *Total Precipitation for Dec. 1997 in mm/month (based on GPCC Full Data Product Version 5, 0.5° x 0.5° resolution).*

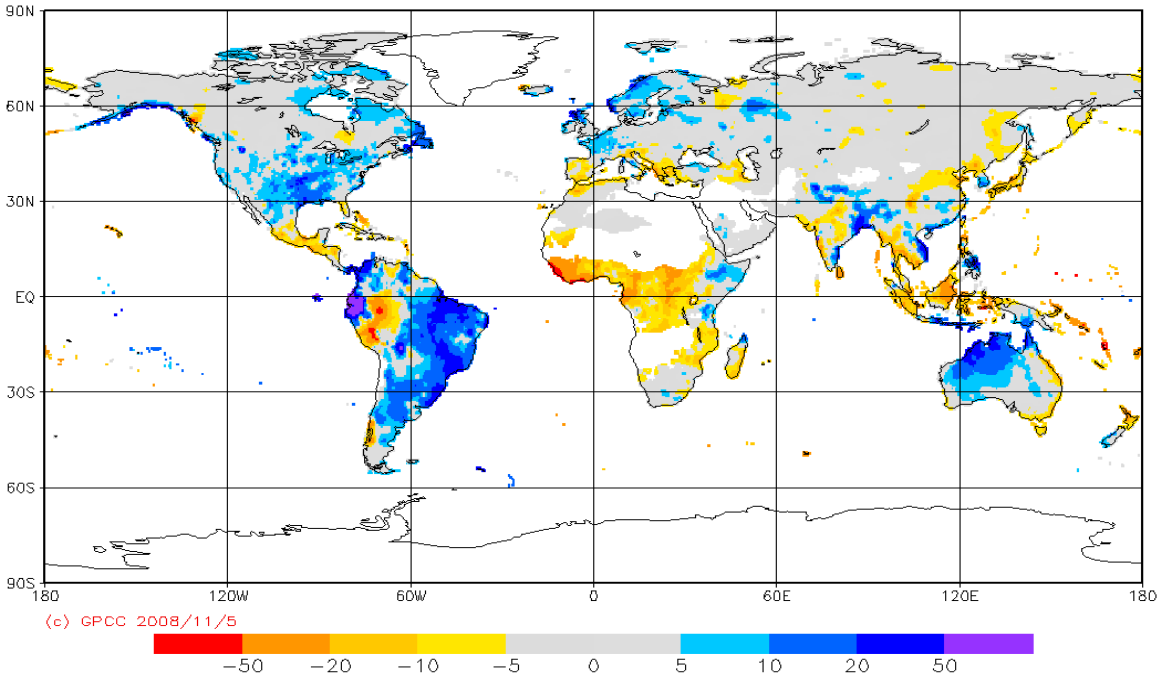

Figure 5: *Linear trends (in mm/month/50 yrs) of annual total precipitation for 1951-2000 (based on GPCC VASClimO Product Version 1.1, 0.5° x 0.5° resolution)*

The GPCC Data Base

The accuracy of raingauge based precipitation analyses mainly depends on the spatial density of stations being used. In order to calculate monthly area-mean precipitation on 2.5° gridboxes with a sampling error of not more than 10%, between 8 and 16 stations per gridbox are needed (WMO 1985, Rudolf et al. 1994). To cover the global land-surface by gridded data of this accuracy, as requested by the GPCP plan (WMO, 1990), this requirement adds up to 40,000 equally distributed stations worldwide.

We distinguish two types of observed precipitation data with regard to their timeliness: data being available near real-time (based on synoptic weather observation data and climate reports exchanged among the national meteorological services via the WMO GTS), and additional data being obtained with a larger time delay. The reason for GPCC to supply a set of different products is that a near real-time analysis is requested by international programmes for various applications, but the near real-time available data base is insufficient in many regions with regard to the requested product accuracy (in addition to the spatial density of stations the quality-control (QC) of station meta information and precipitation data performed is crucial (see section GPCC data processing).

GPCC near real-time data base

The data base for GPCC's Monitoring Product is merged from three sources: monthly precipitation totals derived from synoptical weather reports (SYNOP) at the DWD, Germany, and at NOAA/NCEP, USA, and monthly precipitation totals extracted from CLIMAT-bulletins received at the DWD, JMA (Japan Met. Agency) and UKMO (UK Met. Office). The data base available near real-time comprises ca. 7,000–8,000 stations and provides in some regions a sufficient data base for quantitative precipitation estimates, if the grid resolution is not too high. Users are advised to carefully take into account the number of stations per grid, which is provided as additional information to every GPCC product. Within the data pool, the CLIMAT data – after a quality check – are of higher quality and provide a reference for quality assessment of the SYNOP-based data. The GPCC First Guess Product includes the DWD SYNOP-based monthly precipitation totals from ca. 6,000 stations.

GPCC Full data base

With respect to the limited real-time availability of raingauge data, additional data from dense national observation networks of individual countries are collected at the GPCC. The data acquisition is supported by recommendation letters of the WMO. So far, National Meteorological and/or Hydrological Services (NMHSs) from more than 185 countries of the world contributed data to the GPCC. However, the delay of the deliveries varies between 1-5 years or even more due to the processing time needed by the originators (Figure 7).

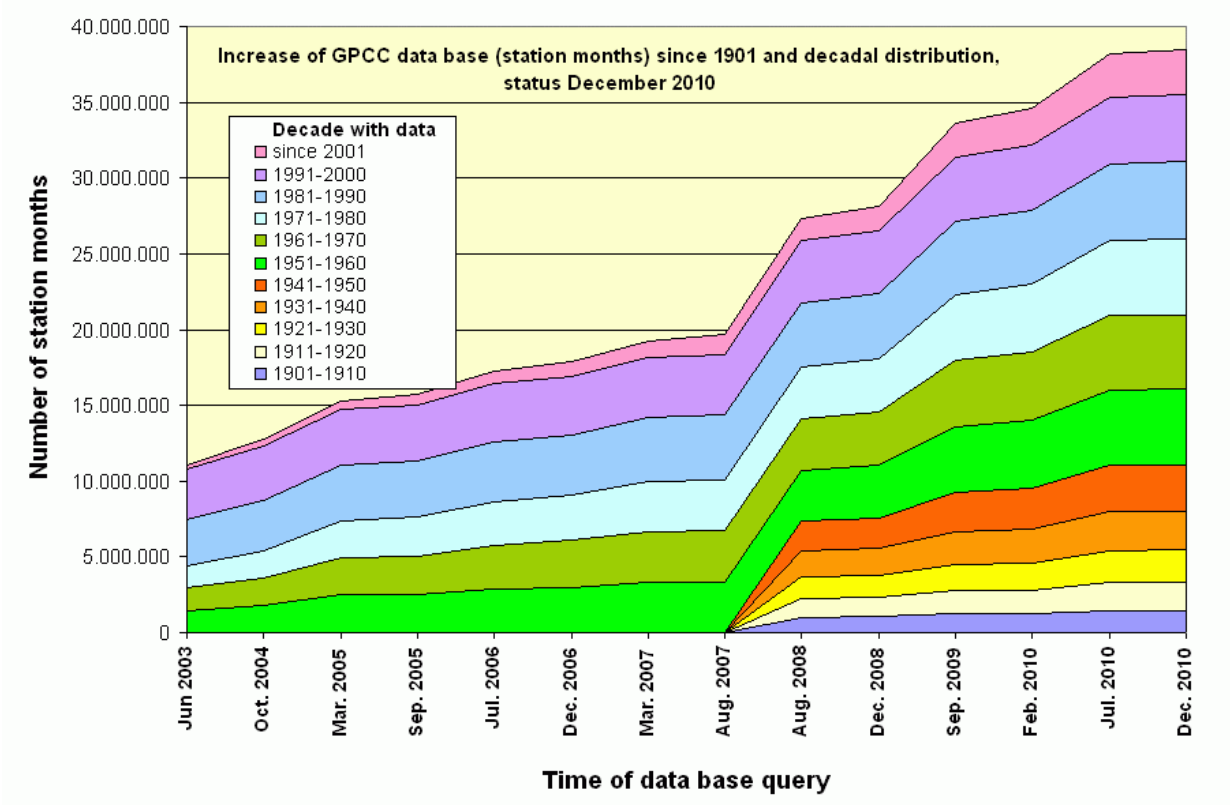

Figure 6: *Temporal evolution of GPCC Monthly Precipitation Database between 2003/06 and 2010/12*

In addition, other available global and regional collections of climate data (Global Historical Climatology Network, GHCN; University of East Anglia Climate Research Unit, CRU; FAO; GEWEX related projects; Asia-Pacific/Matsumoto, etc.) have been integrated in the GPCC data base. Thereby GPCC has compiled the most comprehensive global collection of monthly precipitation data from *in situ* observations (Figures 6-8). With respect to the interests and conditions given by the originators (NMHSs), the GPCC cannot redistribute the station related precipitation data to other parties.

The temporal data coverage of the GPCC products is illustrated by Fig. 7. The the near-real-time First-Guess analysis is including the monthly totals accumulated from SYNOP reports received at DWD. For GPCC's Monitoring Product all SYNOP and CLIMAT data are used if available within ca. one month after observation. The Full Data Reanalysis Product includes all data being supplied later by the individual countries, if the stations have at least 10 years of data (fragments of time-series aren't used any more). The year with the best data coverage is 1987 with data for more than 47,000 stations. A gradual decrease of the number of stations from more than 40,000 in 1986-1991 down to 8,000 stations after 2009 is caused by the delay of the data delivery to and by post-processing at GPCC. The data base continuously increases by delivery of updates for recent years, supplements with additional stations and complementation by long time-series of data. All data suppliers are encouraged to provide annual updates to GPCC. GPCC updates its non real-time products at irregular time intervals subsequent to significant data base improvements.

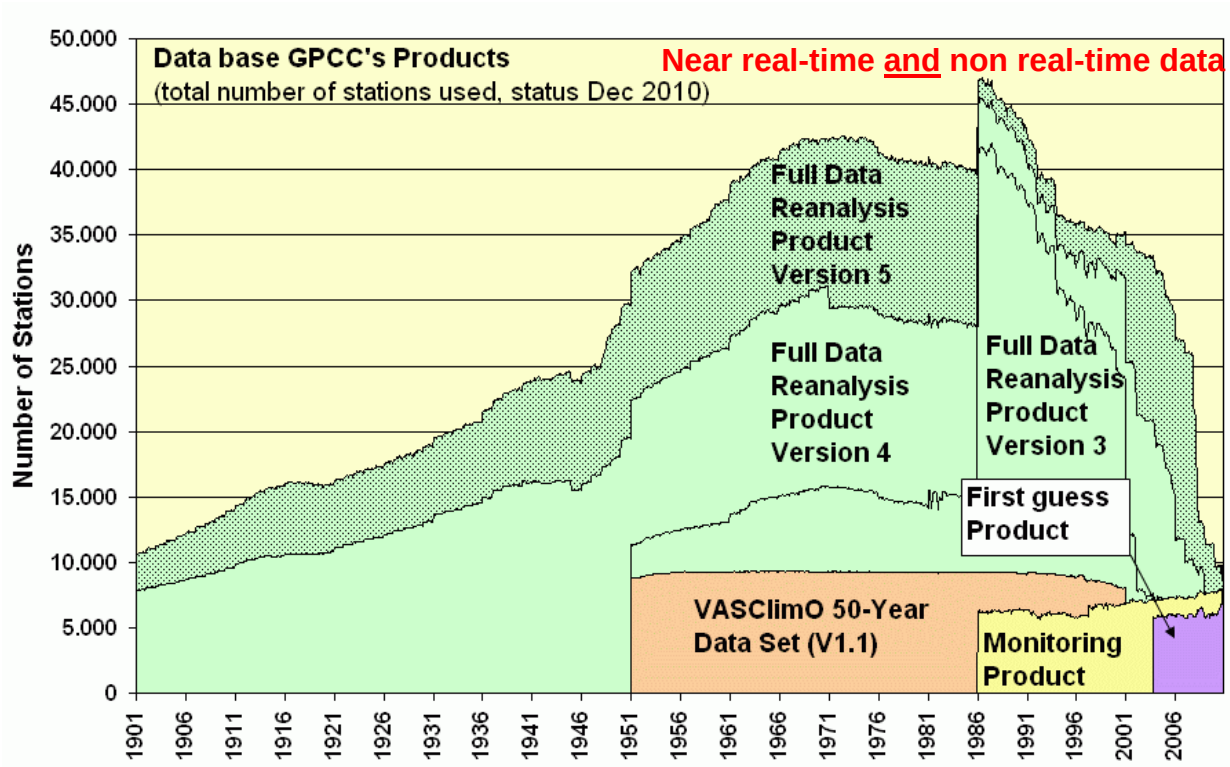

Figure 7: *Total number of stations used for the GPCC products (Near real-time First-Guess Product, Monitoring Product; non real-time 50 Year Data Set (Version 1.1) and Full Data Reanalysis Product (Version 5))*

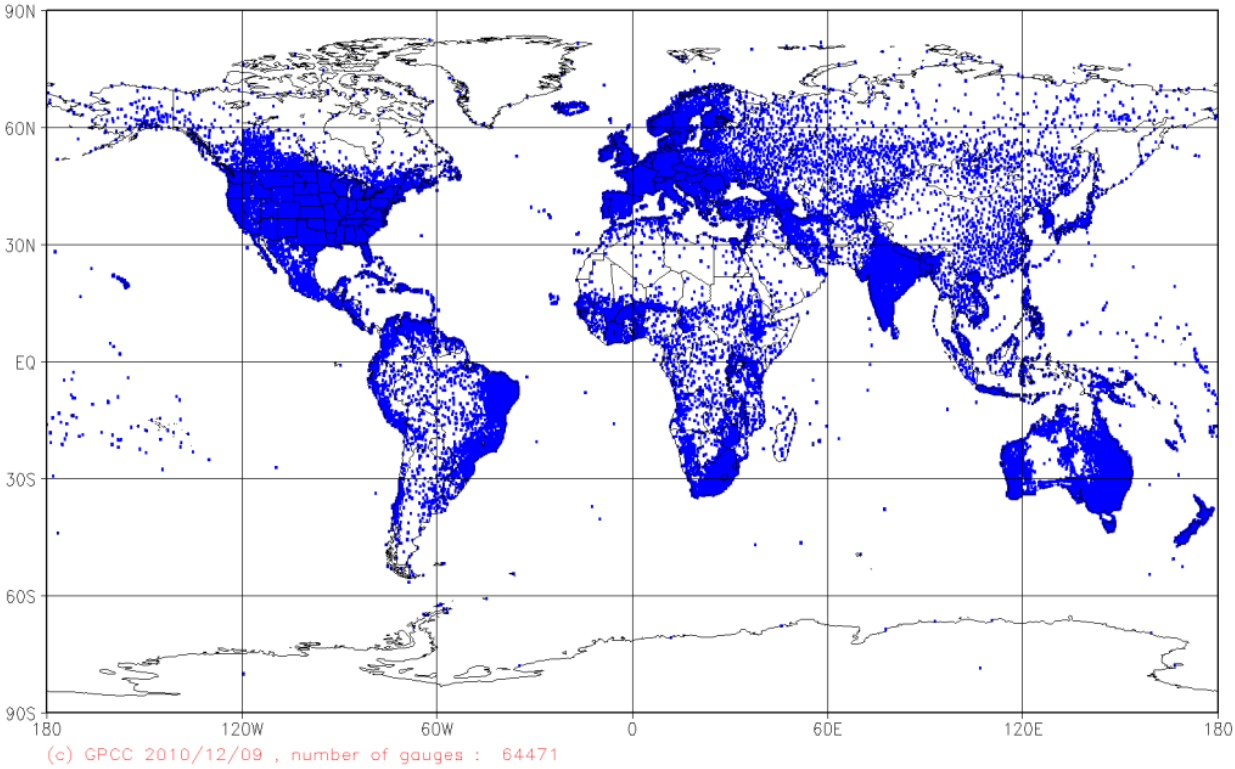

Figure 8: *Spatial distribution of monthly in-situ stations with a climatological precipitation normal, based on at least 10 years of data, available in the GPCC data base (number of stations in June: 64,471)*

The GPCC Data Processing

All data reaching the GPCC are checked, processed, reformatted and integrated in a Relational Data Base Management System. Within this data bank, the records from the different sources (SYNOP, CLIMAT, national data etc.) are stored in parallel (source specific slots) under addition of quality flags indicating the results of data processing. By this an intercomparison and cross-check is possible, which is very helpful in the quality-control (QC) and product generation process.

The data processing steps include QC and harmonization of the meta data (station identification), quality-assessment of the precipitation data, selection and intercomparison of the data from the different sources for the particular products, interpolation of the station-related data to a regular mesh system, and calculation of the spatial means on the 2.5° respectively 1.0° latitude/longitude gridbox area. The Full Data Reanalysis as well as the 50-year VASClimO data set are also available in 0.5° resolution, the new global precipitation climatology is available on a 0.25° resolution, too. The basic information about the methods used is published by Rudolf et al. (1994) and Rudolf and Schneider (2005), additional information is given on GPCC's website (http://gpcc.dwd.de).

The near-realtime Monitoring Product and the non-realtime Full Data Reanalysis Product provide the following variables calculated on the grid:

- Monthly precipitation totals for the individual month
- Mean monthly precipitation totals focussing on the period 1951-2000 ("normals")
- Monthly precipitation anomaly i.e. deviation from the mean 1951-2000
- Monthly precipitation percentage related to the mean 1951-2000
- Number of gauges used per gridcell for the individual month
- Systematic gauge-measuring error per gridcell for the individual month (since Jan. 2007)
- Fraction of liquid and solid precipitation in % of total precipitation per gridcell for the individual month (since Jan. 2007)

In GPCC's reanalyses of non-realtime products (Full Data Reanalysis V.5) as well as in the processing of near-realtime products (Monitoring Product V.3, First Guess Analysis) the normals as in GPCC's precipitation climatology (focussing on the period 1951-2000) are used.

About the accuracy of the gridded results

The two major error sources are: (1) The systematic measuring error which results from evaporation out of the gauge and aerodynamic effects, when droplets or snow flakes are drifted by the wind across the gauge funnel, and (2) The stochastic sampling error due to a sparse network density. The GPCC provides a gridded quantification for the following errors:

The systematic gauge-measuring error is – except for very specific situations – an undercatch of the true precipitation. Parameters affecting the efficiency of measurement are features of the instrument used (size, shape, exposition etc.) and the meteorological conditions (wind, precipitation type, air temperature, humidity, radiation) during the precipitation event. This information is not available for most of the precipitation stations. The global distribution of the error has been estimated for long-term mean precipitation (Legates and Willmott, 1990) and is provided as climatological mean for each calendar month. The error is large in snow regions respectively in cold seasons. Since the GPCC analysis for January 2007, a new on-event correction method for systematic gauge measuring errors is available at GPCC (Fuchs et al., 2001). This event-based correction is usually smaller than the climatological correction, however it is still a rough bias estimate based only on wind, weather, temperature and humidity data from synoptic observations of ca. 6,000 stations available worldwide.

The sampling error of gridded monthly precipitation data has been quantified by GPCC for various regions of the world. Based on statistical experiments using data from very dense networks, the relative sampling error of gridded monthly precipitation is between +/- 7 to 40%

of the true area-mean, if 5 raingauges are used, and with 10 stations the error can be expected within the range of +/- 5% and 20% (Rudolf et al. 1994). The error range for a given number of stations represents the spatial variability of precipitation in the considered region.

The 50-Year Precipitation Data Set

Based on a well-selected subset of the full data base, a homogenised gridded monthly precipitation dataset for the period 1951 to 2000 has been developed at the GPCC in the research project VASClimO (Beck, Grieser and Rudolf, 2005). This climatological data set provides gridded monthly data for 50 years covering the global land areas with a spatial resolution of 0.5° x 0.5° latitude and longitude.

Only station time series with a minimum of 90% data availability during the analysed period 1951 – 2000 are used for interpolation to a regular 0.5° x 0.5° grid in order to minimise the risk of generating temporal inhomogeneities in the gridded data due to varying station densities. Prior to gridding, all available station-data are subjected to a multi-stage quality control of observed values as well as of station-meta data. The analysis is based on relative precipitation (ratio of monthly and long-term mean) time series from 9,343 stations and on the background field of mean monthly precipitation from 28,600 stations. The interpolation method used (Ordinary Kriging) is proved to result in the least interpolation error of several methods tested. Thus, the resulting gridded data-set is highly suitable for the application in studies concerning long-term aspects of climate variability.

Figure 5 gives the spatial distribution of the annual precipitation trends 1951-2000 based on the 50-year data set. Significant negative trends (decrease of precipitation) are marked by the colors yellow and green, significant positive trends (increase of precipitation) by magenta and blue. The negative trend in precipitation over tropical Africa might be caused by a shift of the northern extension of the ITCZ. The results for the calendar months or seasons show varying patterns of long-term precipitation changes, but are still subject to detailed evaluation and interpretation.

The gridded 50-year data set VASClimO, incl. documentation and analysis results (e.g. linear trend magnitude, relative trend, trend/noise ratio, Mann-Kendall significance and more, all for the year and the 12 months), are available from the GPCC Website.

Access to GPCC's gridded products

The different gridded monthly precipitation data sets of GPCC as well as the GPCP Version 2.1 Combined Data Set are freely available. They can be visualized in maps like Figs. 1, 3 or 4 or downloaded in ASCII format using the GPCC-Visualizer (Fig. 9) from our Website http://gpcc.dwd.de. The 50-Year VASClimO data set is also available from this site. Other products can be provided on email request.

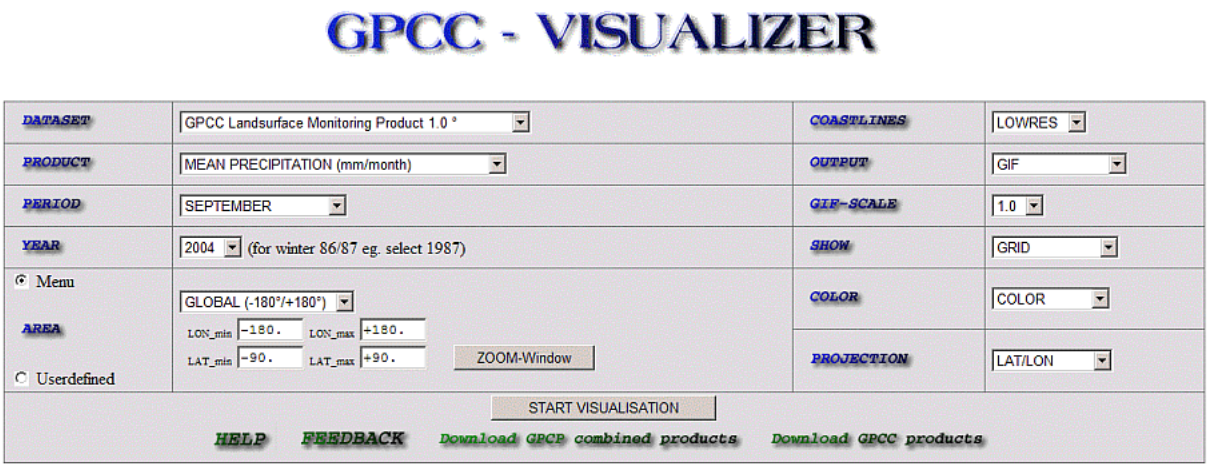

Figure 9: *GPCC Visualizer for online visualization and download of gridded GPCC products*

Some hints and recommendations to GPCC product users

- Check which product is most suitable for the application purpose with regard to the priority of timeliness, regional accuracy or homogeneity.
- Pay attention to the accuracy-related information provided by the GPCC (number of stations per grid, systematic error). Check the error range by consideration of the systematic error estimates and the regional number of stations used.
- Do not compare regional area-means which are calculated from data sets on different grid resolutions. The rough approximation of coastlines may cause relevant deviations between 2.5° and 1.0° based area means.
- When analysing long–term climate variability and changes do not combine different GPCC products available for different periods, which may cause discontinuities in time. Only the GPCC VASClimO product has been adjusted to support long-term precipitation variability and trend analyses.
- Gridded anomalies can be generated in two different ways: (#1) calculation of the anomaly on the stations which requires the availability of both, data from the considered month and normal values, and (#2) by the relation of gridded data sets, which were separately generated for the considered month and for the normal precipitation totals. Method #1 is consistent with regard to the stations used, method #2 includes a much larger number of stations. For technical reasons, method #2 is used by the Visualizer, results based on the anomaly interpolation are available on email request.
- Reference to the GPCC is requested from the users, and feedback about the application of the products is very welcome. You might provide your feedback to gpcc@dwd.de.

The GPCC kindly requests all responsible national agencies to follow the WMO call and to provide the GPCC with the required precipitation and meta data. The analysis results are of high importance e.g. concerning the verification of global climate models and climate variability studies based on observed data. The analysis results of the GPCC are published and freely accessible. But the station-related data delivered by the countries will not be distributed to third parties, in order to respect and protect the ownership of the originators.

Acknowledgements:
The GPCC is operated by Deutscher Wetterdienst (DWD, National Meteorological Service of Germany) under the auspices of the World Meteorological Organization (WMO). The research project VASClimO, contributing to the development of the 50 year data set was supported by the German Federal Ministry of Education and Research (BMBF) within the German Climate Research Programme.

A special thank is addressed to the data contributors, which mostly are National Meteorological and/or Hydrological Services of the world but also some other institutes. Their data contributions enable the GPCC to do its global precipitation analyses described in this document.

References:

Adler, R. F., G. J. Huffman, A. Chang, R. Ferraro, P.-P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin and E. Nelkin, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation analysis (1979-present). J. Hydrometeorol., 4, 1147–1167.

Beck, C., J. Grieser and B. Rudolf (2005): A New Monthly Precipitation Climatology for the Global Land Areas for the Period 1951 to 2000. DWD, Klimastatusbericht 2004, 181-190.

Fuchs, T., J. Rapp, F. Rubel and B. Rudolf (2001): Correction of Synoptic Precipitation Observations due to Systematic Measuring Errors with Special Regard to Precipitation Phases. Phys.Chem. Earth (B), Vol. 26, No. 9; pp. 689-693.

Huffman, G. J., R. F. Adler, B. Rudolf, U. Schneider, and P. R. Keehn, 1995: Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information. J. Climate, 8, 1285-1295.

Legates, D. R., and C. J. Willmott, 1990: Mean seasonal and spatial variability in gauge-corrected, global precipitation. Int. J. Climatol., 10, 111-127.

Rudolf, B., H. Hauschild, W. Rueth, U. Schneider (1994): Terrestrial Precipitation Analysis: Operational Method and Required Density of Point Measurements. NATO ASI I/26, Global Precipitations and Climate Change (Ed. M. Desbois and F. Desalmand), Springer Verlag Berlin, 173 - 186.

Rudolf, B., and U. Schneider (2005): Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations. Proceedings of the 2nd Workshop of the International Precipitation Working Group IPWG, Monterey, October 2004, 231-247.

WMO, 1985: Review of requirements for area-averaged precipitation data, surface based and space based estimation techniques, space and time sampling, accuracy and error, data exchange. WCP-100, WMO/TD-No. 115.

WMO (1990): The Global Precipitation Climatology Project - Implementation and Data Management Plan. WMO/TD-No. 367, 47 pp. and 6 Appendices, Geneva, 1990.

Xie, P. and P.A. Arkin (1997): Global Precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. Bull. Amer. Meteorol. Soc. 78, 2539 - 2558.

Global Precipitation Climatology Centre
c/o Deutscher Wetterdienst
P.O. Box 10 04 65
63004 Offenbach a. M.
Germany

Tel.: +49 - 69 - 8062 2900
Fax: +49 - 69 - 8062 3987
email: gpcc@dwd.de
Web: http://gpcc.dwd.de

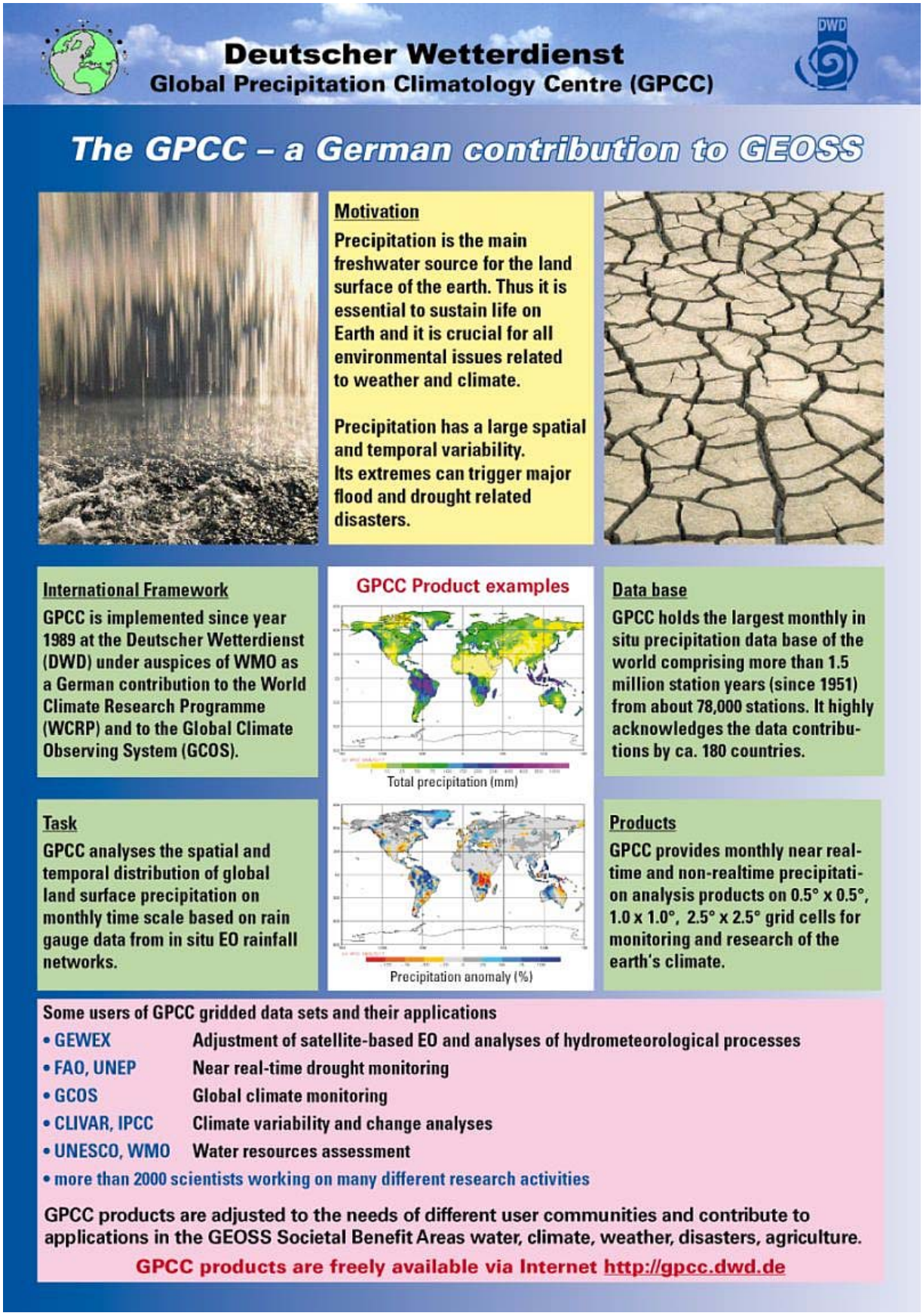

Fig. 11 GPCC poster presented at the GEO-IV plenary meeting Cape Town, South Africa, 28-30 Nov 2007