Global Precipitation Analysis Products of the GPCC

B. Rudolf

Precipitation plays an important role in the global energy and water cycle. With regard to land use, agriculture and hydrology, accurate knowledge of precipitation amounts reaching the land surface is of special importance for fresh water assessment and management on all spatial scales. With respect to global climate change and the needs for assessment, International Organizations initialized a variety of 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 Deutscher Wetterdienst (DWD, National Meteorological Service of Germany) as a German contribution to the World Climate Research Programme (WCRP). From the origin, the Centre is a component of the WCRP Global Precipitation Climatology Project (GPCP). Later (1994), the long-term operation of the GPCC has been requested by WMO with regard to the Global Climate Observing System (GCOS). The products of the GPCC, gauge-based gridded precipitation data sets for the global land surface, are world-wide used by various institutions, in particular within water-related projects of WMO, FAO, UNESCO and UNEP. Ana- lyses of time-series of area-averaged precipitation covering the periods from 1901 (resp. 1951) to present are currently carried out. They are of special interest for CLIVAR and GCOS and will support the IPCC assessments.

While a full global coverage by observed precipitation data can only be obtained using satellite observations, conventionally “in situ“ measured data are still necessary to improve the quality and accuracy by calibration, validation or verification of remotely sensed precipitation. The gauge-based “Monitoring Product” of the GPCC delivers the in situ component for both of the two world-wide used satellite-gauge combinations of the GPCP (Huffman et al. 1995, Adler et al. 2003) and of NCEP (Xie and Arkin 1997). For the investigation of climatic precipitation variability, e.g. long-term trends, time-series are necessary reaching retrospectively beyond the satellite aera.

The aim of the GPCC is to serve different user requirements concerning on the one side the accuracy of the gridded precipitation results depending on the number of stations used, and on the other side the timeliness of the product availability. GEWEX for instance requests high spatial resolution and accuracy for the recent decade, while the priority of GCOS and IPCC is focused on long homogeneous time-series. Suitable products are realized by cut-off dates for data extraction and analysis corresponding to the application type. All gauge-based analysis products (except of the 50-year climatology) result from the same quasi-operational data management and analysis system, but they differ with regard to the number of the included stations and the level of data quality control being performed. The gridded data sets are available in the spatial resolutions of 1.0° by 1.0° and 2.5° by 2.5° geographical latitude by longitude. Corresponding to international agreement, the gridded products are freely available on the Internet. In the year 2002, more than 200,000 accesses on GPCC’s Website have been counted.

Characteristics of the gridded analysis products being availble from the GPCC

- The *First Guess* of the monthly precipitation anomaly is based on interpolated precipitation anomalies at about 4,500 stations. The data sources are synoptic data received via GTS for the considered month, and the climatic mean monthly precipitation at the same stations (1961-1990) which are selected from the GPCC global normals collection. The fully automatical quality-control (QC) has been performed for the synoptic data. The first month published with a First Guess was September 2003. Since that time, a First Guess is available within 5 days after end of the observation month.

- The *Monitoring Product* of monthly precipitation for global climate monitoring is based on SYNOP (after high level QC) and monthly CLIMAT reports from totally 7,000 stations and is available within about 2 months after observation month. The operational production started with the year 1996 and is going on to near-present. An Interim Version of the Monitoring Product covering the period 1986-1995 has been derived from similar input data in 1994/1995 after GPCC's development phase. The series has been complemented backwards to 1979 by another preliminary gauge product using the same analysis method but a reduced input data set (Xie, Rudolf, Schneider and Arkin, 1996). The Monitoring Product supplies the in situ component to the satellite-gauge combinations of GPCP (Huffman et al. 1995, Adler et al. 2003) and of CMAP (Xie and Arkin 1997). Figures 1 and 2 illustrate exemplarily GPCC gauge-based products in map format.

- The *GPCP Satellite-Gauge Combined Data Set Version 2* is a full global data field of monthly precipitation on a 2.5° grid available within 3 months after observation month. This product is jointly created by the GPCP participants (DWD/GPCC, Eumetsat, JMA, NOAA, NASA and others). It includes multi satellite-based precipitation estimates and the gauge-based analyses of the GPCC for the period 1979 to near-present (Adler et al. 2003).

- The *Full Data Product* is of much higher accuracy and recommended to be prefered for hydrometeorological studies and verification. The analysis includes all stations supplying data for the individual month. The data coverage varies from less than 10,000 and to more than 40,000 stations. A new full data re-analysis is performed in irregular time intervals, which are set with respect of data base improvements. The current Full Data Product is Version 3 covering the period from 1951 to 2004 (Figure 3).

- The new *50 Year Climatology* supplying gridded time-series for studies on climate variability and trend is based on data being selected with respect of a (mostly) complete temporal data coverage and homogeneity of the data time-series. The first version is based on time-series for 9,343 stations covering the period 1951-2000 (Beck, Grieser and Rudolf, 2005).

- The Monthly Precipitation Normals Data Set provides gridded mean monthly precipitation for the period 1961-1990, based on the mean data from about 30,000 stations. The data base comprises the normals of the WMO collection, and or normals delivered by the countries to the GPCC or calculated from data time-series at the GPCC.

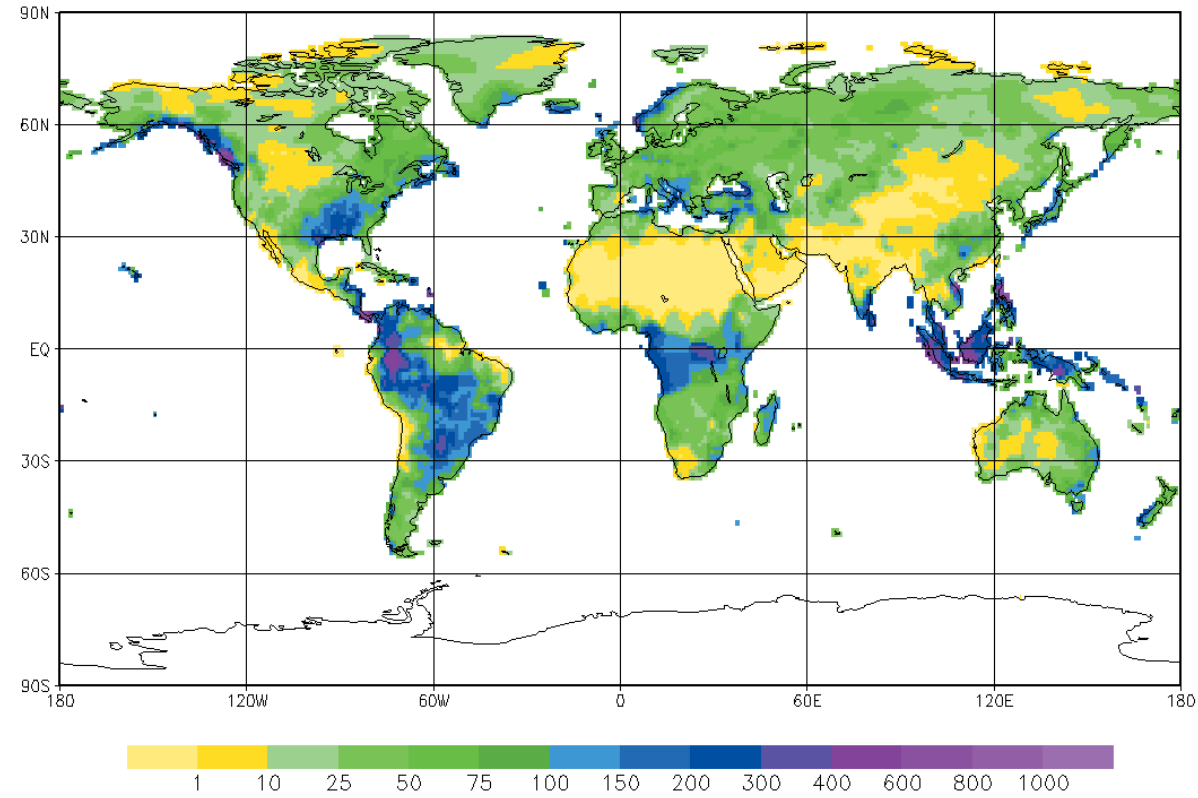

Figure 1 Total Precipitation for November 2004 in mm/month (Monitoring Product, 1° resolution).

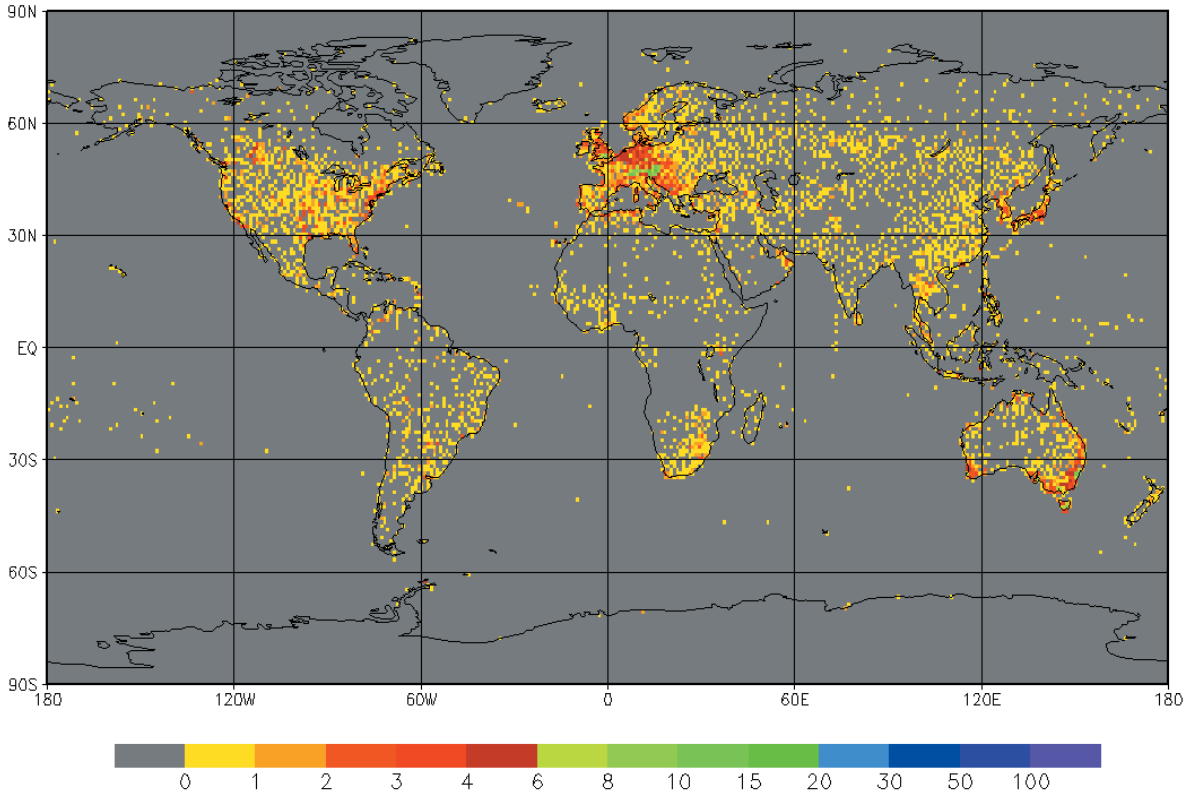

Figure 2 Spatial coverage by observed data for November 2004 (Monitoring Product, 1° resolution, number of stations per gridbox).

The Data Base

At first, a comment on the data requirements: The accuracy of gauge analyses mainly depends on the number of stations being used. In order to calculate monthly area-mean precipitation on 2.5° gridboxes with an 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 stations world-wide.

We distinguish two types of observed precipitation data: such data being available near real-time, and data which can be obtained with a larger delay only. The reason 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 with regard to the requested accuracy.

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) received at the DWD, Offenbach, and NOAA/NCEP, Washington DC, and monthly totals received within CLIMAT bulletins. The merged data base covers 7,000 stations and provides regionally a sufficient data base for quantitative precipitation estimates. Within the data pool, the CLIMAT data are of higher quality and reference for quality assessment of the SYNOPs. The First Guess includes the DWD SYNOP-derivatives only.

Full data base:
With respect to the limited real-time availability of gauge data, additional data from dense national observation networks of individual countries are collected at the GPCC. The data acquisition is supported by recommendations and requests of the WMO. So far, National Meteorological and/or Hydrological Services of 176 countries contributed data to the GPCC (status January 2005). However, the delay of the deliveries varies between one year and five years or even more due to the time needed for processing by the originators. In addition, other available global and regional collections of climate data (CRU, FAO, GHCN, BALTEX, Asia-Pacific/Matsumoto, etc.) have been integrated in the GPCC. By this, GPCC holds the most comprehensive global collection of monthly precipitation data from in situ observations. The data are delivered by National Meteorological and Hydrological Services to the GPCC, following the request of WMO. 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. 3. The year with the best data coverage is 1987 with monthly precipitation data being available for about 43,000 stations. A gradual decrease of the number of stations from 40,000 in 1986-1990 down to 7,000 stations after 2003 is caused by the delay of the delivery of additional data. The data base continuously increases by delivery of updates for recent years, supplements with additional stations and complementationion by long time-series of data.

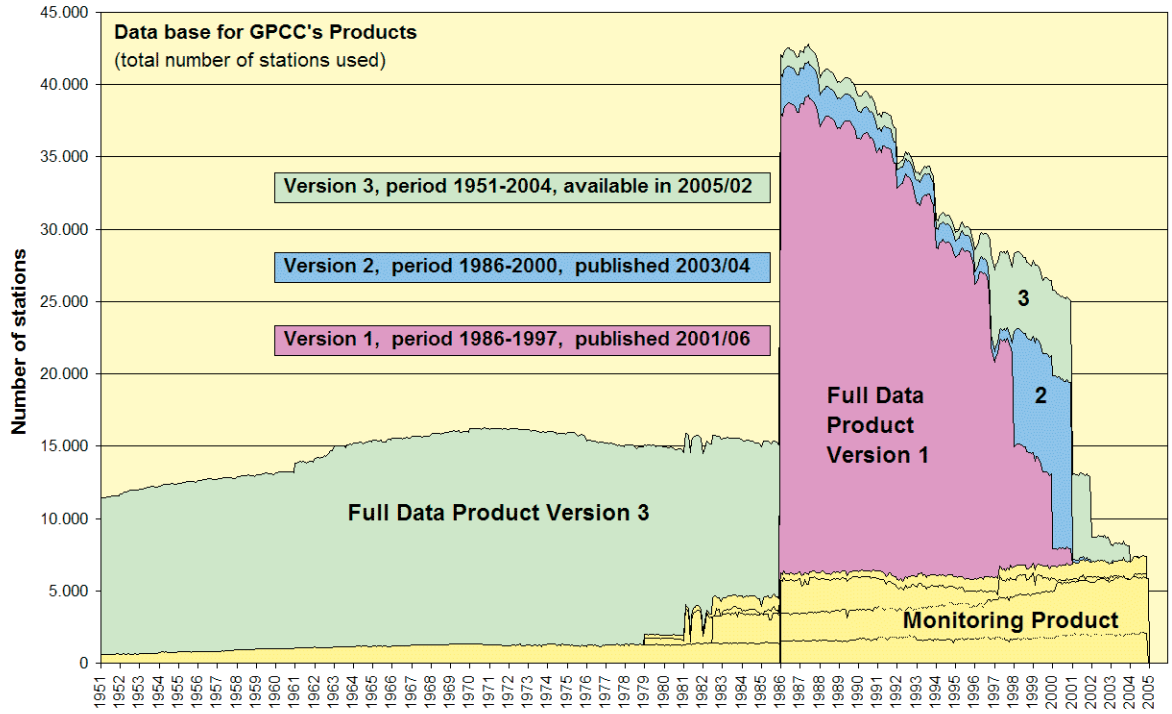

Figure 3 Total number of stations of the Full Data Product (Versions 1, 2, 3) and the near real-time Monitoring Product.

Data processing

All data reaching the GPCC are integrated in a Relational Data Base Management System. Within the data bank, the records from the different sources (SYNOP, CLIMAT, national data etc.) are stored in parallel with quality-flags in so-called networks; by this a comparison and cross-check is possible.

The data processing steps include quality-control and harmonization of the meta data (station identification), quality-assessment of the precipitation data, selection 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 basic information about the used methods has been published by Rudolf et al. (1992 and 1994), additional information is given on GPCC's website (http://www.gpcc.dwd.de).

For the GPCC gauge analyses, following variables are calculated on the gridboxes:

- Monthly precipitation totals for the individual month (Fig. 1)
- Mean monthly precipitation totals for the period 1961-1990 ("normals")
- Monthly precipitation anomaly i.e. deviation from the mean 1961-1990
- Monthly precipitation percentage related to the mean 1961-1990
- Number of gauges used per gridcell for the individual month (Fig. 2)
- Bulk factors for assessment of systematic gauge-measuring error (Fig. 4).

About the accuracy of the gridded results

The two major error sources are, first, the systematic measuring error which results from evaporation out of the gauge and aerodynamical effects, when droplets or snow flakes are drifted by the wind across the gauge funnel, and second, the stochastic sampling error due to a sparse network density.

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 and seasonal distribution of the error has been estimated for long-term mean precipitation (Legates and Willmott, 1990). The error is large in snow regions respectively in cold seasons. Figure 4 depicts the average systematic measuring error for the month November.

The sampling error of gridded monthly precipitation data has been quantified by the 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% and 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.

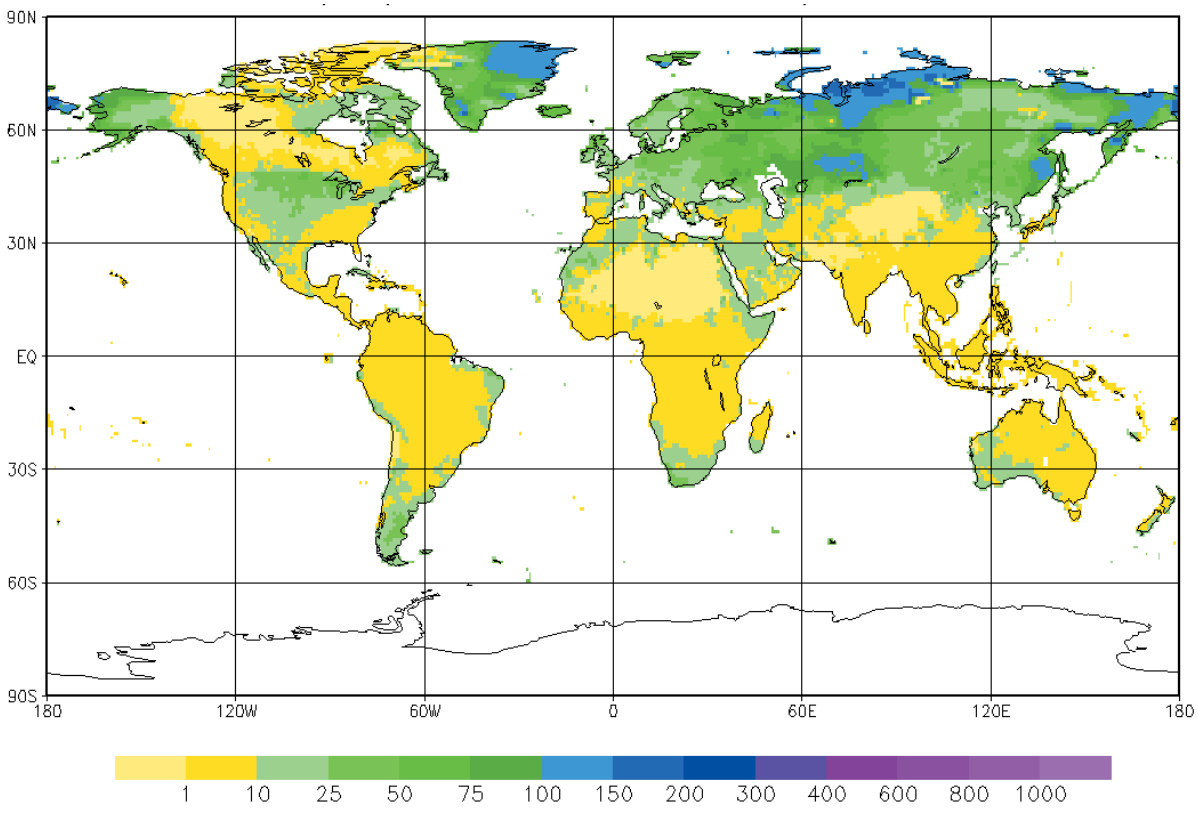

Figure 4 Relative systematic gauge-measuring error for November in % of the gridded data as measured after Legates and Willmott, 1990

Information about the product accuracy is available on the grid for all GPCC precipitation analysis products.

Access to GPCC's gridded products

The gridded data sets of GPCC and GPCP are freely available. The Monitoring Product and the GPCP Version 2 Combined Data Set can be visualized in maps (e.g. Fig. 1, 2 and 4) or downloaded in ASCII format using the Visualizer (Fig. 5) from GPCC's Website (http://gpcc.dwd.de). The other products will be provided after email request.

Some hints and recommendations to 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, 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.
- 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 is available on email request.
- Reference to the GPCC is requested from the users, and feedback about the application of the products is welcome. Mail to: gpcc@dwd.de.

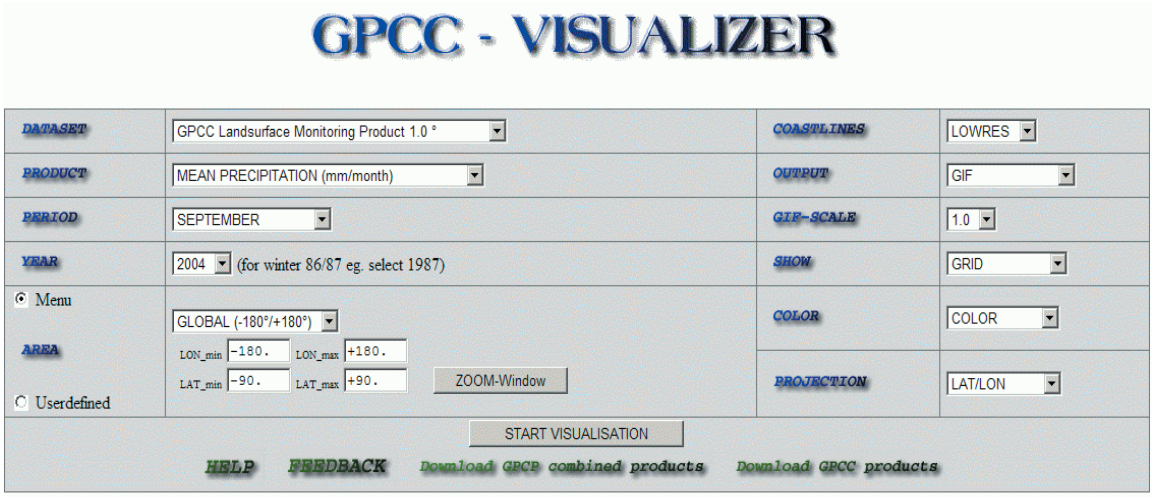

Figure 5 Online-Visualization and Download of gridded products.

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 the 50 year climatology by C. Beck and J. Grieser, is supported by the Federal Ministry of Education and Research (BMBF) of Germany within the German Climate Research Programme DEKLIM, Förderkenzeichen 01LD0032.

A special thank is addressed to the data contributors, which mostly are National Meteorological and Hydrological Services but also some other institutes. Their data enable the GPCC to do the global analysis.

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.

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, M. Reiss, U. Schneider (1992): Beiträge zum Weltzentrum für Niederschlagsklimatologie. Meteorol. Zeitschr. N.F. 1/1, 7-84.

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.

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.

Xie, P., Rudolf, B., Schneider, U., Arkin, P.A. (1996): Gauge-Based Monthly Analyses of Global Land Precipitation from 1971 to 1994. J. of Geophysical Research, Vol. 101, No. D14, p. 19,023-19,034.