

# GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle

Udo Schneider · Andreas Becker · Peter Finger ·  
Anja Meyer-Christoffer · Markus Ziese · Bruno Rudolf

Received: 24 September 2012 / Accepted: 17 February 2013 / Published online: 20 March 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** In 1989, the need for reliable gridded land surface precipitation data sets, in view of the large uncertainties in the assessment of the global energy and water cycle, has led to the establishment of the Global Precipitation Climatology Centre (GPCC) at Deutscher Wetterdienst on invitation of the WMO. The GPCC has calculated a precipitation climatology for the global land areas for the target period 1951–2000 by objective analysis of climatological normals of about 67,200 rain gauge stations from its data base. GPCC's new precipitation climatology is compared to several other station-based precipitation climatologies as well as to precipitation climatologies derived from the GPCP V2.2 data set and from ECMWF's model reanalyses ERA-40 and ERA-Interim. Finally, how GPCC's best estimate for terrestrial mean precipitation derived from the precipitation climatology of 786 mm per year (equivalent to a water transport of 117,000 km<sup>3</sup>) is fitting into the global water cycle context is discussed.

## 1 Introduction

There is an increasing demand for climatological information with regard to possible impacts of climate change on our environment. This is true in particular with regard to the important role precipitation—a key part of the global water cycle—plays for life on earth.

A number of large-scale climate data sets based on station (in situ) observations have been developed during the last decades to study mean state, variability, or long-term trends of climate. An overview of early precipitation data sets and

climatologies is given in Legates (1987, 1995). One of the first global precipitation climatologies was created by Jäger (1976) based on existing climate atlases and some hypotheses about the global water balance. Since then, a number of climatological data sets based on in situ data have been created such as the ones compiled by Legates (1987), Global Historical Climatology Network (GHCN; Peterson and Vose 1997; Peterson et al. 1998), the Food and Agriculture Organization (FAO 2001), Climatic Research Unit (CRU; New et al. 2002; Mitchell et al. 2004), NOAA's precipitation reconstruction over land (the PREC/L data set being over land an optimum interpolation of gauge measurements; Chen et al. 2002), the WorldClim data set (Hijmans et al. 2005), and the climatology of Matsuura and Willmott (2009). However, all these existing climate data sets have their shortcomings with regard to precipitation, such as the very inhomogeneous data coverage over space and time, e.g., the very poor data coverage at the beginning of the twentieth century. For example, the CRU and Matsuura and Willmott data sets (the latter, as the PREC/L data set, is relying mainly on GHCN-V2) comprise data for only about 3,500 or 4,100 stations at the beginning of the twentieth century. These data sets have their maximum station numbers around 1970 with some 10,000 (CRU), almost 23,000 (Matsuura and Willmott), and 16,700 stations (PREC/L data set, starting in 1948), respectively. Another weakness is that the quality control (QC) of the station meta data and especially of the precipitation data throughout these data sets is performed in a more or less basic way, which is insufficient in the light of the frequent and different kinds of errors occurring in the raw station data according to the long-term experience of the GPCC (see the following text discussion; details are given in Section 4).

The GPCC was established in 1989 at Deutscher Wetterdienst (DWD, German Weather Service) on invitation

U. Schneider (✉) · A. Becker · P. Finger · A. Meyer-Christoffer ·  
M. Ziese · B. Rudolf

Department of Hydrometeorology, Global Precipitation  
Climatology Centre, Deutscher Wetterdienst, Offenbach, Germany  
e-mail: udo.schneider@dwd.de

of the WMO as the in situ component of the Global Precipitation Climatology Project (GPCP; WMO 1990) of GEWEX (the former Global Energy and Water Cycle Experiment was recently renamed to Global Energy and Water Exchanges). Its main task is the analysis of monthly precipitation for the earth's land surface on the basis of rain gauge (in situ) measurements. Over the years, it has built up a unique data base containing precipitation data of more than 85,000 stations worldwide—out of the total number of rain gauges operated worldwide estimated to range between 150,000 and 250,000 according to different authors (Sevruk and Klemm 1989; New et al. 2001; Strangeways 2007). The GPCC was quite successful to overcome these shortcomings of the earlier data sets by integrating some of them (CRU, FAO, and GHCN) into its data base and by acquiring additional precipitation data through bilateral contacts and with the support of WMO. The latter activity has meanwhile resulted in the acquisition of data sets from the national meteorological and/or hydrological services (NMHSs) of 190 countries worldwide and from research projects. The data from the different sources are stored separately in GPCC's data base (in source-specific slots); more details will be given in Section 2.

In response to the very different user requirements for different application purposes, the GPCC has set up a portfolio of precipitation analysis products (near and non-real-time) briefly described in Section 3. Becker et al. (2013) provide a comprehensive account on this portfolio including application examples.

The necessity of a thorough QC prior to generation of any precipitation analysis products based on in situ data has been evidenced by the more than 20-year experience in the processing of observational precipitation data at the GPCC since almost any larger input data set—near and non-real-time—contains a variety of position and meta data errors, and there are all kinds of errors in the raw rain gauge data itself. A crucial point in merging data sets from different sources is the QC and harmonization of the station meta data to avoid duplicate stations and to correctly assign the data to the stations in the data base. This is not always trivial because station meta information for the same station may differ from source to source. Over the years, the GPCC has developed a sophisticated QC system consisting of different steps for quality assurance of the station meta information as well as of the precipitation data (Section 4), whereby the source-specific data archival in GPCC's data base allowing inter-comparisons of the data from the different sources is very helpful in detecting and correcting data errors. The QC processing steps, being different for the quasi-operational (near real-time) and non-real-time products, are described in Sections 4.1 and 4.2, respectively, together with examples illustrating some typical problems found in the raw station data.

Section 5 describes GPCC's new land surface precipitation climatology and its comparison to some of the other station-based terrestrial precipitation climatologies as well as to the GPCP V2.2 data set (a combination of satellite-based precipitation estimates and GPCC's analyses over land) (Huffman et al. 1995; Adler et al. 2003; Huffman et al. 2009) and to precipitation climatologies derived from ECMWF's model reanalyses ERA-40 (Uppala et al. 2005) and ERA-Interim (Dee et al. 2011; Simmons 2011). In addition to that, its role in quantifying the global water cycle will be discussed in Section 5.3.

## 2 GPCC's observational data base

GPCC's data base comprises precipitation data on a mainly monthly basis from a variety of sources. With the beginning of 2012, the GPCC has started the additional acquisition and processing of daily precipitation data, but the daily data archive yet available has still to grow to support the issuance of reliable daily products in the mid-term future. The data distributed by the NMHSs via the WMO Global Telecommunication System (GTS) to fulfil the needs of near real-time weather analysis and prediction and climate monitoring are available near real-time such as synoptic weather reports (SYNOP), from which monthly precipitation totals can be accumulated (Schneider et al. 1992b; Thomas and Patterson 1983), and monthly climate reports (CLIMAT). In addition to the GTS data, the GPCC has acquired precipitation data from NMHSs from about 190 countries that meanwhile form the backbone of its data base but become available only with a larger delay (non real-time data).

In merging the data from different sources, the quality control and harmonization of the station meta information is crucial to detect errors in the station meta information (especially geographical coordinates) and to ensure consistency of time series and to avoid duplicate stations in the merged data set.

### 2.1 Near real-time GTS data

Via the WMO GTS data can be obtained in near real-time from SYNOP reports (with a daily up to hourly resolution) and monthly CLIMAT reports. Monthly precipitation data are routinely obtained at GPCC from three sources:

- (1) monthly totals calculated at GPCC from SYNOP reports received at DWD, Offenbach (Schneider et al. 1992b)
- (2) monthly totals calculated at NOAA's Climate Prediction Center (CPC) from SYNOP reports received at NOAA, Washington D.C. (Thomas and Patterson 1983) and

- (3) a combination of the monthly CLIMAT reports received at DWD, Offenbach, JMA (Japan Met. Agency), Tokyo, and UK Met. Office, Exeter (although theoretically the reports available via GTS should be identical at the various Regional Telecommunication Hubs (RTHs) of the GTS, in practice there are slight differences in the CLIMAT reports received at different RTHs (Schneider and Umeda 2012)).

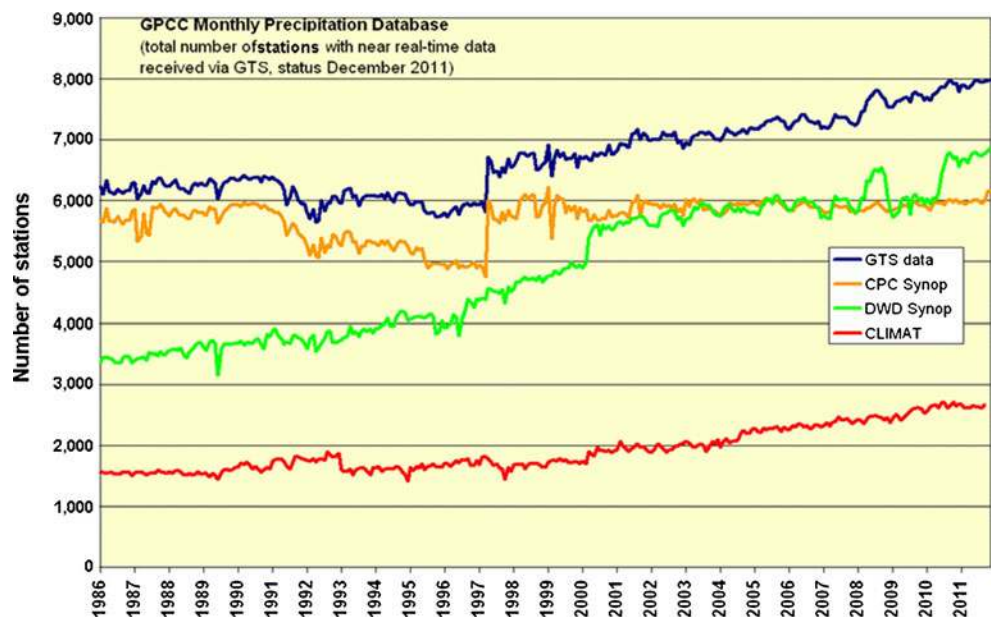
The data from these sources, which are checked and merged by GPCC in order to improve the spatial coverage and data quality, form the basis of its near real-time monitoring of terrestrial monthly precipitation (Section 3.1). The total number of stations available via GTS with monthly precipitation data has increased over time from some 6,000 stations in 1986 to about 8,000 stations during the recent years (see Fig. 1). Remarkable is the steady increase in the number of stations for which monthly precipitation totals could be calculated at GPCC (temporal coverage over the month at least 70 %) from SYNOP reports received at DWD, Offenbach; this number has about doubled from ca. 3,500 in 1986 to about 7,000 since 2010 (available since mid-1982). The number of monthly totals from CPC (available since 1979), except for a spike at the turn of the year 1996/1997, has remained quite stable at about 6,000 stations. There is also an increase in the number of monthly totals from CLIMAT reports from ca. 1,600 stations in 1986 to about 2,600 stations over the last years (going back to 1950).

### 2.2 GPCC's Full Data Base

Owing to the large variability and intermittency of precipitation in space and time, a much larger number of stations is required for global analyses of precipitation compared to less intermittent climate variables such as air temperature or sea level pressure. Studies by WMO (1985), Rudolf et al. (1994), and Rudolf and Schneider (2005) indicate that between eight and 16 stations per 2.5° grid—depending on the climatic/orographic conditions—are needed to fulfil the accuracy requirement of the GPCP of a sampling error of less than 10 % in the precipitation analysis (WMO 1990), resulting in a total number of roughly 40,000 stations for the global land areas.

Therefore, the GPCC is acquiring additional precipitation data from national weather services, hydrological institutes, etc., to enlarge its data base. Originally, GPCP's/GPCC's data collection period was defined to start in 1986 according to the GPCP Implementation and Data Management Plan (WMO 1990). Following recommendations from WCRP projects Global Climate Observing System (GCOS), Global Energy and Water Exchanges (GEWEX), and Climate Variability and Predictability (CLIVAR), the GPCC has then worked on the historical extension of its data base before 1986. So far, NMHSs from about 190 countries have supplied additional data on a voluntary basis, following WMO requests and bilateral contacts with GPCC. In addition to that, the data collections from CRU (New et al. 1999, 2000), GHCN V.2 (Peterson and Vose 1997; Peterson et al. 1998), and FAO (FAO 2001), including a large amount of historical data, have been integrated into

**Fig. 1** Availability of monthly precipitation data in GPCC's data base showing the composition of the GTS data being available for the near real-time analysis products for the period since January 1986 according to the GTS data sources



GPCC's data base (number of stations with precipitation data integrated: ca. 11,800 from CRU, 34,800 from GHCN, 13,550 from FAO). The station number for GHCN is consisting of 20,590 stations from GHCN-V.2, supplemented by monthly totals calculated from the GHCN daily data set (Menne et al. 2012), which is an ongoing activity at GPCC.

Other major contributions for specific regions (“regional data sets”) are the data sets for: (1) the Former Soviet Union (Groisman et al. 1991; Groisman and Rankova 2001) with data for overall 2,186 stations for the period 1891–2001, updated to 2009 by a data set for Russia consisting of 518 stations downloaded from RIHMI-WDC (All-Russian Research Institute of Hydrometeorological Information-World Data Center) and (2) the African rainfall archive from Nicholson (1986, 1993, 2008) containing data of 1,338 stations for the period up to 1998 with some time series starting as early as 1838.

As a consequence of 1986 having been the start year of GPCC's evaluation period during its first decade of operation, there is still a jump at the turn of the years 1985/1986 (Fig. 2 showing the contribution of the individual data sources to GPCC's Full Data Base and Fig. 3 showing its evolution for different versions of the Full Data Reanalysis), with the years 1986/1987 having the best coverage with data of more than 47,000 stations available, although there has been a steady increase over the historical time period for each release of the Full Data Reanalysis (V.3 to V.6). With the issuance of the most recent V.6 in December 2011, this jump is almost leveled, and the aforementioned increase during the past decade gives also an indication how long it takes for worldwide collected rain gauge data to arrive at GPCC and to pass the rigorous QC before entering the

GPCC data base, which is causing the decrease to ca. 37,500 stations in 2000 and less thereafter. The full data base has now data for more than 45,000 stations around 1970 compared to only some 10,000, 16,700, or almost 23,000 stations in the CRU, PREC/L, or Matsuura and Willmott data sets, respectively.

Over the years, GPCC's Full Data Base has become the largest basis of monthly precipitation data worldwide meanwhile including monthly precipitation totals of more than 85,000 stations. Some of the time series are even extending back to the early eighteenth century. GPCC has data for about 10,800 stations at the beginning of the twentieth century compared to only about 3,500 and 4,100 stations in the CRU and the Matsuura and Willmott data sets and outperforms most of the earlier existing data sets over the entire period by a factor of 2 to 5 with regard to data coverage.

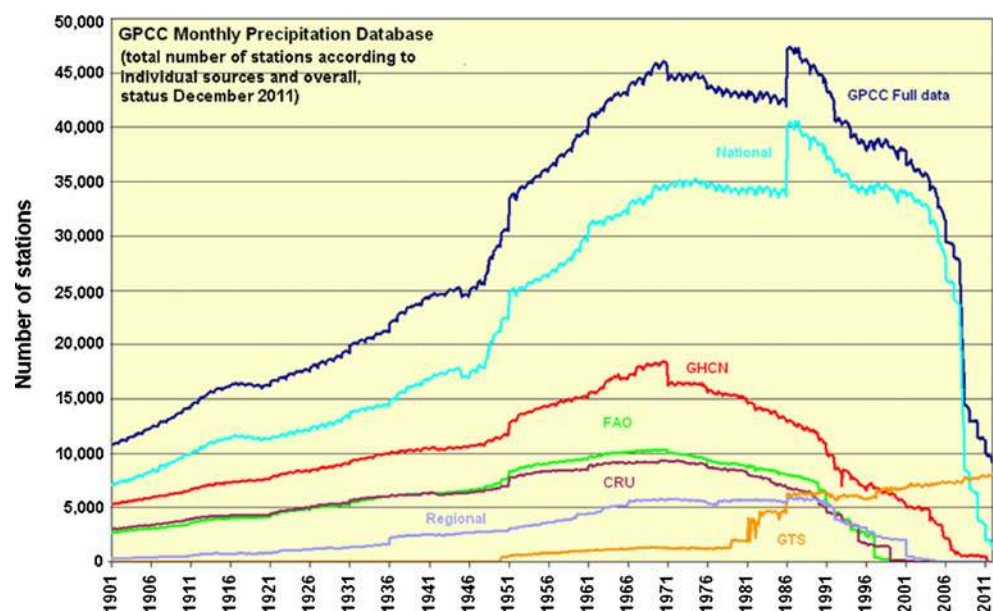
### 3 Data interpolation and GPCC's precipitation analysis products

Based on its quality-controlled data base, GPCC constructs a portfolio of gridded precipitation analysis products addressing the various user requirements. The public and unlimited availability of all data sets of the GPCC product portfolio described here is guaranteed through their digital object identifier references (for further details, consult Becker et al. (2013)).

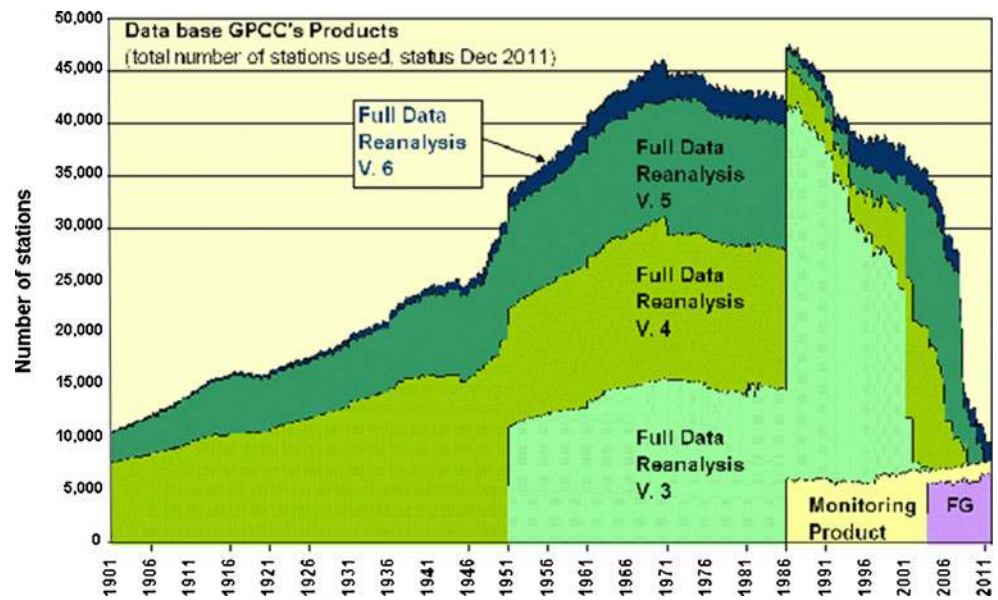
#### 3.1 Calculation of gridded precipitation data sets

The calculation of the gridded analysis products from gauge observations consists of the following major steps:

**Fig. 2** Availability of monthly precipitation data in GPCC's full data base as a function of time since January 1901 showing the contribution of the individual data sources



**Fig. 3** Total number of stations used for the GPCC products (near real-time First-Guess Product FG, Monitoring Product; non-real-time Full Data Reanalysis Product (Versions 3 to 6))



1. Interpolation of anomalies from the climatological normals at the stations to regular gridpoints on a 0.25° (Full Data Reanalysis) or 0.5° latitude/longitude subgrid (Monitoring and First Guess Product); for the near real-time products (Monitoring and First Guess), anomalies from the climatological normals for the grid are used if not available at a station.
2. Averaging the anomalies at the four corners of the 0.25° or 0.5° grid to an anomaly for the 0.25° or 0.5° subgrid.
3. Calculating the areal average anomaly for the grid for the 0.5°, 1°, or 2.5° resolution by applying a weighting according to the area and land fraction of each grid.
4. Superimposing the gridded anomalies on the background climatology (see “GPCC’s precipitation analysis products”).

For the interpolation of station anomalies to the regular grid, a modified version of the empirical interpolation method SPHEREMAP (Willmott et al. 1985) is routinely used at the GPCC (for more details of the modified SPHEREMAP interpolation method, as well as the construction of the gridded data sets and area averages, see Becker et al. (2013)).

### 3.2 Accuracy of the gridded data sets

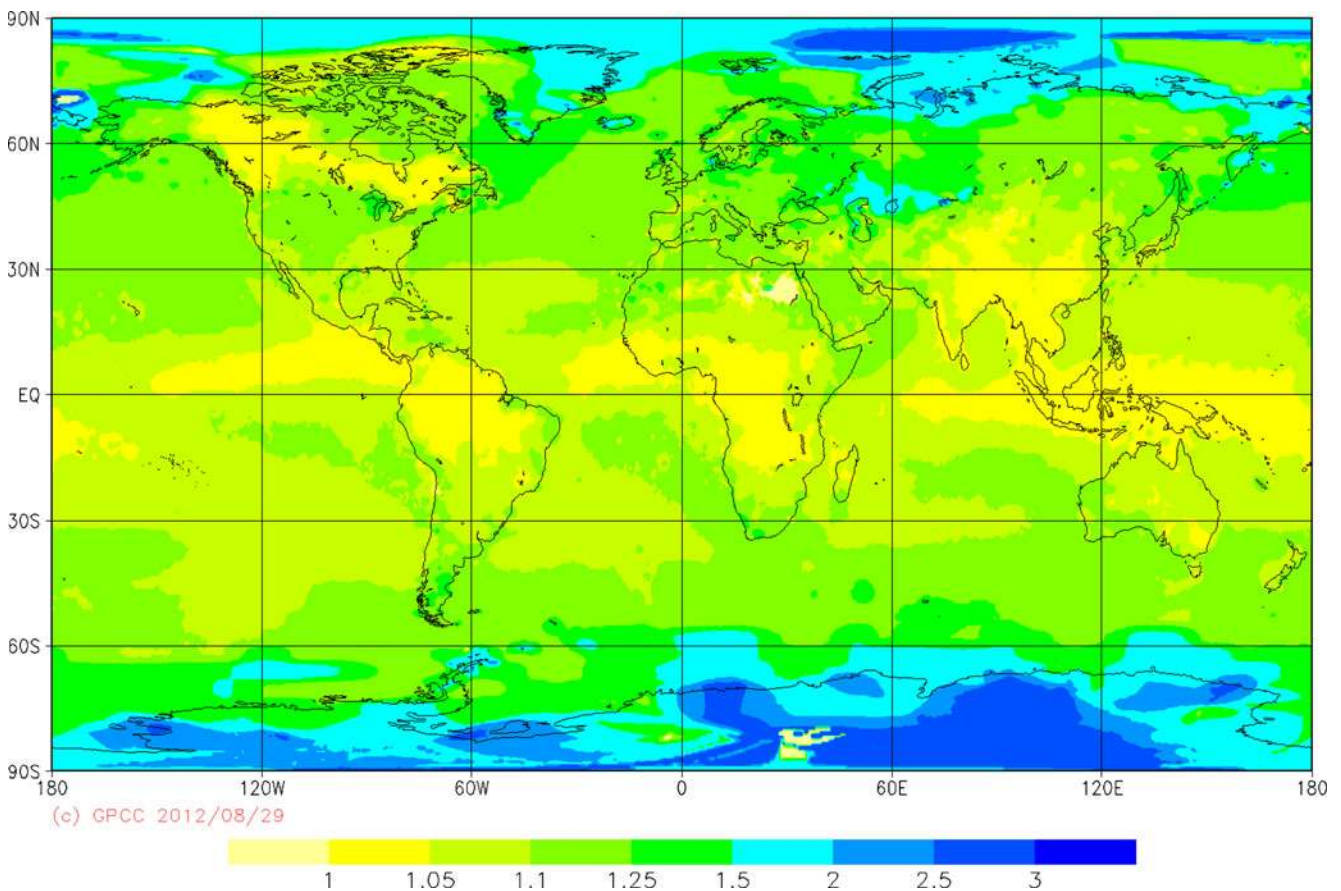
The major error sources affecting gridded precipitation estimates based on rain gauge measurements are:

1. The systematic gauge-measuring error resulting from evaporation out of the rain gauge and aerodynamic effects, when droplets or snow flakes are drifted by the wind across the gauge funnel

2. The sampling error depending on the network density

The *systematic gauge-measuring error* is generally an undercatch of the true precipitation (Sevruk 1982, 1985). Parameters affecting the efficiency of gauge measurement are features of the instrument used (size, shape, exposition, etc.) and the meteorological conditions (wind, air temperature, humidity, radiation) during the precipitation event. The precipitation phase (liquid, solid, mixed), as well as the intensity (i.e., drizzle, shower) of a precipitation event, plays an important role, too. For a large part of the precipitation stations, this information is not available. The global distribution of the error has been estimated for long-term mean precipitation (Legates and Willmott 1990; hereafter LW1990) and is provided as climatological mean correction factor for each calendar month. These correction factors are shown in Fig. 4 for the entire year and vary between 1 (almost no correction) and about 3 (an addition of 200 %) in regions with a large fraction of snow in precipitation, respectively, in cold climates.

An improved correction method for systematic gauge-measuring errors (Fuchs et al. 2001) taking into account the weather conditions in the evaluation month (wind, temperature, relative humidity, precipitation phase, and intensity) has been implemented at GPCC. The correction factors resulting from the new method are usually somewhat smaller than the bulk climatological correction according to LW1990 (Fuchs et al. 2001). The required information for this more realistic bias correction is available for ca. 6,000–7,000 synoptic stations worldwide. It can be interpolated to the grid and is provided operationally together with the fraction of the precipitation phases (liquid, solid, mixed) for GPCC's Monitoring Product. Since this more realistic approach for the correction of the systematic gauge-



**Fig. 4** Correction factor to compensate for the systematic gauge-measuring error for the entire year according to Legates and Willmott (1990)

measuring errors is currently available only for a short period (since 2007), we are still relying on the correction of LW1990 in the following to correct GPCC's climatology for the systematic gauge-measuring errors but reduce it somewhat arbitrarily by 15 % to compensate for its tendency to overestimate the systematic gauge-measuring error.

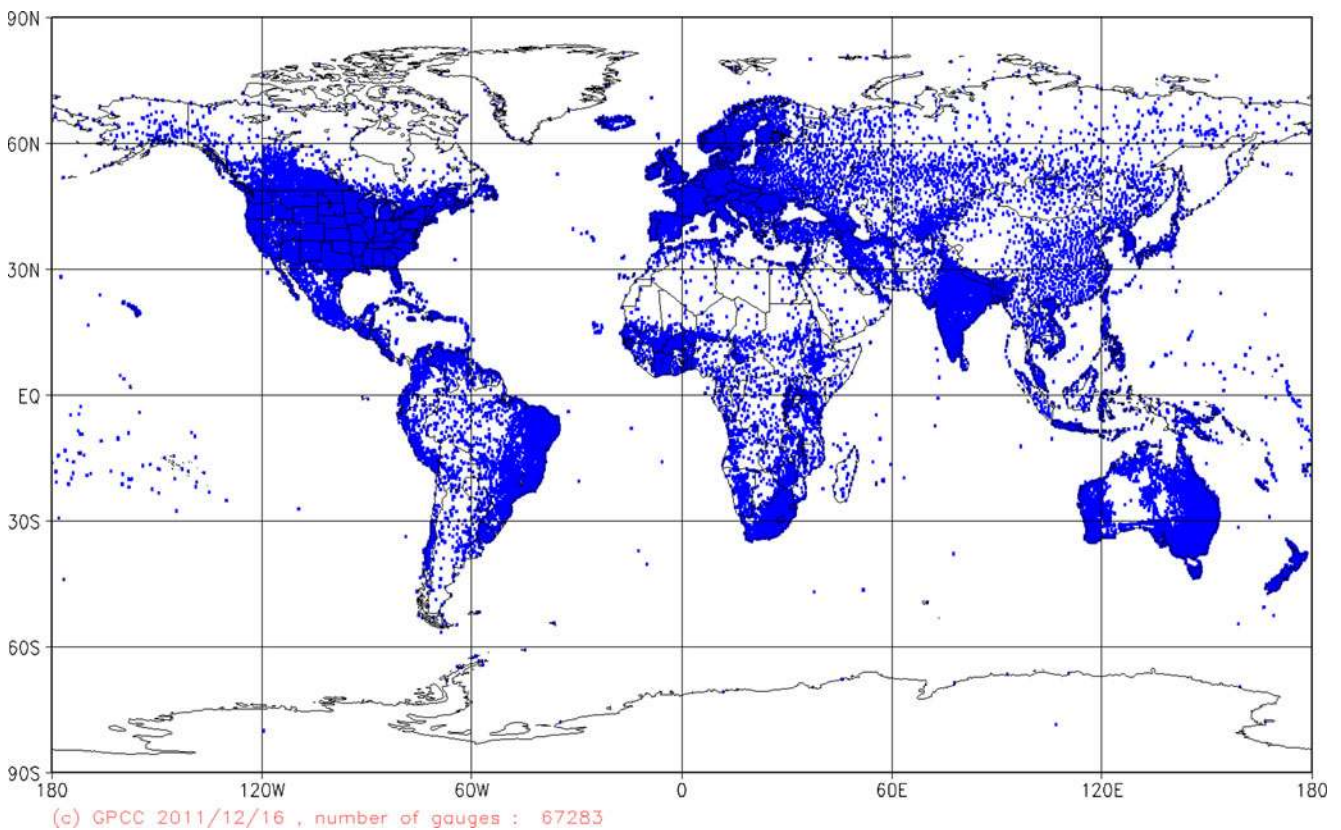
The *sampling error* of gridded monthly precipitation data has been quantified by WMO (1985) and investigated by GPCC (Rudolf et al. 1994) for various regions of the world. Based on statistical experiments using data from very dense networks, the relative sampling error of gridded monthly precipitation is between  $\pm 7$  to 40 % of the true area mean, if five rain gauges are used, and with ten stations the error can be expected within the range of  $\pm 5$  and 20 %. The error range for a given number of stations reflects the spatial variability of precipitation in the considered region and is depending on the orography, season, and type of precipitation (convective, stratiform). Along with its analysis products, the GPCC provides the number of stations per grid as a rough estimate of the sampling error. Becker et al. (2013) provide a systematic study on the sensitivity of the sampling error against the number of stations available for the interpolation in terms of the Jack-knife error for a choice of three different interpolation methods (distance weighting,

SPHEREMAP, ordinary block Kriging). It reveals that GPCC's approach of interpolating the anomalies instead of absolute values is more efficient to reduce the sampling error than the choice of the interpolation method itself.

### 3.3 GPCC's precipitation analysis products

To fulfil the different requirements of the user community, the GPCC has set up a portfolio of different precipitation analysis products. GPCC's precipitation climatology is of central importance in the way that the precipitation fields for all other near and non-real-time GPCC products (except for the 50-year data set VASCLIMO V1.1 consisting of monthly precipitation analyses for the period 1951–2000 based on mostly homogenous rain gauge measurements (Beck et al. 2005)) are constructed by interpolating the anomalies at the stations to the grid and subsequently superimposing the gridded anomaly analyses upon the gridded background climatology.

The *Precipitation Climatology* released in Dec 2011 (Meyer-Christoffer et al. 2011a, b, c, d) is focusing on the target period 1951–2000 and consists of normals from ca. 67,200 stations, for which the spatial distribution is shown in Fig. 5. The climatology comprises normals (CLINO's)



**Fig. 5** Spatial distribution of stations with a climatological precipitation normal (number of stations in July, 67,283)

collected by WMO (1996) mainly for 1961–1990 or normals delivered by NMHSs of the countries to GPCC. Normals have been calculated at GPCC for 1951–2000 from time series of monthly data if data for at least 40 years out of the period are available, for 30-year reference periods preferably out of this target period (1951–1980, 1961–1990, or 1971–2000) or for other 30-year periods (i.e., 1981–2010 or 1931–1960) with at least 20 years of data in the period under consideration. If even this was not possible for a station, normals have been calculated for any other period with at least ten complete years of data.

This resulted in climatological normals for the period 1951–2000 for 23,936 stations and for the 30-year periods 1951–1980, 1961–1990, 1971–2000, 1981–2010, and 1931–1960 for about 32,850, 39,800, 33,890, 27,128 and 21,335 stations, respectively, and ca. 67,200 stations overall. If climatological normals are available at a station for more than one of these reference periods, normals are selected in the order given above. A discussion of the characteristic patterns in GPCC's precipitation climatology and its role in quantifying the global water cycle is provided in Section 5.

Near real-time products of the GPCC based on the GTS data are the *First Guess Product* (Ziese et al. 2011) and the *Monitoring Product* (Schneider et al. 2011b, c).

The *First Guess Product* is based on synoptic weather reports received at DWD, from which monthly totals can be accumulated for ca. 6,000–7,000 SYNOP stations, whereby an automatic-only QC is applied (Schneider et al. 1992b). It is generated within 5 days after the end of the observation month and is available since September 2003.

The *Monitoring Product* (Schneider et al. 2011b, c) of monthly precipitation for global climate monitoring is based on SYNOP and monthly CLIMAT reports received via GTS from ca. 7,000–8,000 stations (after automatic and manual quality control) and is generated within 2 months after the end of the observation month. It is available since 1986 and forms the in situ component of the satellite gauge combined precipitation analyses of CPC Merged Analysis of Precipitation (CMAP, Xie and Arkin 1997) and of GPCP V2.2 after 2010 (Adler et al. 2003; Huffman et al. 2009).

The non-real-time *Full Data Reanalysis* (Rudolf et al. 2011; Schneider et al. 2011d, e, f) is based on all stations in the GPCC data base supplying data for the individual month, near real-time, and non-real-time (including the data from the NMHSs and the historical data collections), for which a climatological normal is available. The Full Data Reanalysis Product Version 6 covering the period from 1901 to 2010 was generated in Dec 2011. The data coverage per month varies from 10,800 at the beginning of the twentieth century

to more than 47,000 stations in 1986/1987. The Full Data Reanalysis is being updated at irregular time intervals subsequent to significant data base enlargements and improvements. It is the in situ component of GPCP V2.2 for the historical period 1901–2010 (Huffman et al. 1995; Adler et al. 2003; Huffman et al. 2009) (for more details about GPCC's analysis products, see Becker et al. (2013) and Schneider et al. (2011a)).

#### 4 GPCC's semi-automatic quality control system

The more than 20-year experience in the processing of observational precipitation data sets at the GPCC indicates that almost any larger input data set—near and non-real-time—contains more or less errors and there are all kinds of errors in the raw station data. Raw data itself, as well as station meta information, can be affected by typing or coding errors and other modifications occurring on the way from the measurement at the station to the data archive. Therefore, a thorough quality control (QC) is necessary to detect and correct/eliminate such errors which otherwise would have a significant impact on the analysis results.

Owing to the large variability of precipitation and the skewness of its frequency distribution (i.e., monthly totals occur in the range from 0 mm to more than 2,000 mm), a fully automatic quality control would eliminate all data being classified as outliers including real extremes. These true extremes, however, are very important to describe the variability of precipitation.

Therefore, QC processing at GPCC is semi-automatic in the way that the data classified as questionable by the automatic QC procedures undergo additional visual checks. The QC system of successive automatic and visual checks has been optimized with respect to the features of the different data sources and the specific meta information being available. Figure 6 displays a simplified scheme for the main steps of processing, quality control, archival, and analysis of precipitation data at the GPCC and distribution of its gridded precipitation products for the near as well as non-real-time data.

##### 4.1 Station identification and quality control of station meta information

The data sets delivered in non-real-time by the NMHSs of the individual countries arrive at the GPCC in very different formats. Some data sets are ASCII-formatted; others are in formats of commonly used spread sheets or outputs of data bank software. In some cases, data have been received on paper sheets (yearbooks, monthly reports etc.), which have been digitized if the data are an important supplement in a data-sparse region.

The data sets received at GPCC are first checked for readability and then reformatted (brought into a uniform format) as given in Fig. 6. To avoid a spatial misallocation of climatic data in the analysis, for the national/regional (“non real-time”) data sets supplied to the GPCC, the station locations are displayed by a climate data visualization software, and it is checked if all stations are located within the boundaries of the country. For stations located outside of the boundary, the geographical coordinates are checked with geographical information available via Internet (i.e., Google-Earth) through geographical atlases or regional maps.

Subsequently, the uniform data sets are loaded into GPCC's relational data base management system (hereafter “data bank”), whereby the station meta data received from the different sources are checked against the meta information archived in GPCC's data bank. If the station meta information in the data set is identical with that of a station in the data base, the data are assigned to that station. If no similar station is existing in the data base, a new station is created therein.

In case of discrepancies in station meta information between data set and GPCC's data base, the data supplier is contacted, if possible in a timely manner. Otherwise, the geographical coordinates of the station are checked with other sources of geographical information (i.e., Google-Earth, geographical atlases, or regional maps). This re-checking of geographical information during each loading process is resulting in a continuous improvement of GPCC's station data base and has led to a very high degree of reliability of its station meta information. Observed discrepancies can be attributed, in part, to a different spelling (or changes) of station names (i.e., Bombay to Mumbai) and/or errors in the geographical coordinates or elevation (typing errors, wrong sign of geographical coordinates, reversed digits, missing or wrong conversion of units). With regard to the geographical coordinates, typical typing errors on the order of sometimes 1°, 2°, or even up to 10° latitude or longitude are detected in many of the input data sets. In the elevation information, there are sometimes errors in the conversion of meters and feet, zero instead of missing elevation, etc.

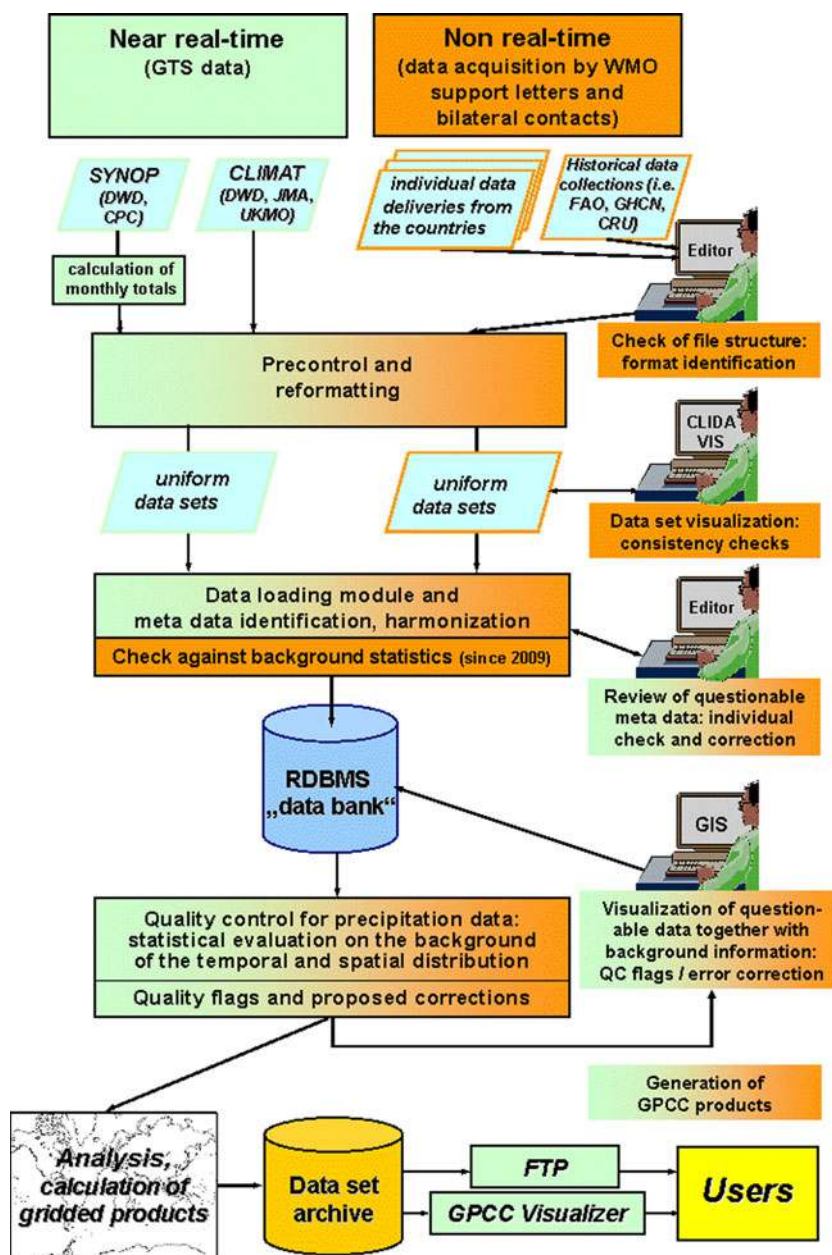
A lot of significant errors in the station meta information have been detected in the data sets and could be corrected, which otherwise would have seriously affected any derived analysis products.

##### 4.2 Quality control of the monthly precipitation data

In order to avoid loading mismatched or overall erroneous data sets into the data bank, all national/ regional and GTS precipitation data sets have already been pre-controlled separately using different techniques fitting the respective data sources (see Fig. 6).



**Fig. 6** Simplified scheme for the main steps of processing, quality control, archival, and analysis of precipitation data at the GPCC and distribution of the gridded precipitation products



However, the pre-control is not able to recognize or eliminate most of the individual data errors. This requires a synopsis of the data from the different sources and a check of the spatial consistency with neighbouring stations. Therefore, storing the data from the different data sources in parallel in the data bank (in source-specific slots) together with the quality flags indicating the results of data processing is very helpful in the QC processing and enables detection of errors by cross-checks of the data from the different sources. Since mid-2009, all data being loaded into the data bank are additionally checked against background statistics (1 and 99 % percentile for the station, or if not available, against the respective percentiles for the 2.5° grid in which

the station is located). Data flagged as questionable after this screening are to be checked manually for spatial consistency with neighbouring stations and against data from alternate sources.

#### 4.2.1 Quality control for GPCC's quasi-operational products (“near real-time data”)

For the First Guess Product (see Section 3.3), only the station meta information is checked during loading into the data bank. Moreover, calculation of the monthly totals from the synoptic weather reports includes a number of consistency and other checks (Schneider et al. 1992b).

For the Monitoring Product (based on all GTS data), the semi-automatic QC system of the GPCC combines the quality control with the spatial analysis in the following steps:

- Automatic QC and data selection (from the GTS data sources)
- First spatial analysis using the selected data
- Computer-assisted manual revision of data flagged as questionable (“visual component”)
- Final spatial analysis using the revised data

*The automatic component* In the automatic part, all GTS precipitation data of the month to be analysed are checked first against the climatological normals and frequency distribution of the time series at the stations and second for consistency with the spatial average resulting from the first analysis of the data at the given station and at the neighbouring stations. Additionally, the data at the station from the different sources are checked against each other. According to a suitable combination of these different criteria for spatial or statistical outliers, only a small portion of the GTS data, roughly 500 out of the ca. 8,000 stations (between 5 and 8 %), are classified each month as questionable and flagged by the automatic system. Based on the checks, the data from all GTS sources are assigned quality flags, and the station data from the source with the highest expected quality at a station are pre-selected to be used for a first preliminary spatial analysis.

*The visual/manual component* After the first analysis, all stations are displayed in a world map, whereby at each station the data source being pre-selected is identified by a symbol, and stations with data flagged as questionable in the automatic process are highlighted. These data have to be visually reviewed by a trained expert. It is possible to zoom in the map and view all relevant information being available for the given station, including the precipitation data of the neighbouring stations and background fields such as gridded climatologies and the orography. For data flagged as questionable in the automatic process, the following options are possible:

- Confirm the flagged data to be true
- Change the data selection for the individual station if a more reliable precipitation amount from another source is available
- Correct data if an obvious error is recognized (e.g., factor 10 error)
- Correct the station meta data if the station coordinates are recognized to be wrong
- Flag the data of a station that cannot be corrected (“trash”), not to be used in the final spatial analysis

The quality index of the checked data is modified according to the selected option. All original as well as the corrected data are archived in the data bank of DWD.

#### 4.2.2 Additional quality control for GPCC's non-real-time global precipitation products

In addition to the previously described QC for GPCC's Monitoring Product, the full data base has repeatedly been checked statistically for outliers over the last years for each new release of the gridded precipitation climatology (focusing on the target period 1951–2000) and the Full Data Reanalysis (for details of GPCC's analysis products, see Section 3.3 or Becker et al. (2013)).

*Statistical check of outliers* Since mid-2009, all non-GTS data loaded into GPCC's data bank are checked against background statistics, namely, the 1 and 99 % percentile for the station, or if not available, against the respective percentiles for the 2.5° grid in which the station is located. The GTS data are loaded as they are, because they all undergo the full semi-automatic QC processing described before, to archive the data as received and document the errors detected.

For the calculation of the new versions of the gridded precipitation climatology (in 2008, 2010, and 2011), the time series of all stations with at least 10 years of data have been checked statistically. For the period 1951–2000, all data exceeding six standard deviations (SD) and eight SD for the entire period since 1901 have been flagged as questionable. For the 2008 (2010 and 2011) versions, this resulted in 1,488 (413, 88) cases for the period 1951–2000 that had to be checked—736 (280, 39) cases for the entire period since 1901; for the 30-year and shorter periods, the 2008 criterion of six SD (76 cases overall; would have been more than 3,000 cases for five SD!) could be screened to five SD for the 2010 and 2011 releases giving 860 and 83 cases. Through the course of the repeated QC processing for the releases V.4, V.5, and V.6 of the Full Data Reanalysis, the time series of overall about 10,000 stations had been checked visually since for each case of a station with dubious data, generally two to four and sometimes even more neighbouring stations have been checked for spatial consistency. Overall, about two third of all cases have been confirmed as correct; corrections of errors detected have been flagged in the data bank.

One of the biggest problems with the raw station data are “0” instead of missing values, which are caused by not or incorrectly flagging missing data by the data providers. These erroneous “0” values are, only in part, detected in the statistical checks described before and then additionally checked against neighbouring stations to clarify the cases. Therefore, in the QC processing for the Full Data Reanalysis

V.5 release in 2010, we performed a systematic check for erroneous “0” values that revealed and eliminated automatically, after thorough pre-checks for data subsets, 8,746 incorrect “0” values.

In case of corrections, the original data are kept in GPCC's data base and the corrections are archived additionally as a higher quality level; erroneous data that cannot be corrected (trash) are flagged as “to be eliminated” at the higher quality level and are not used in the analyses. Additionally, all cases with at least 24 consecutive months having “0” precipitation have been checked and, although such cases can be correct in desert regions, a few further errors could be detected in this step.

*Visual check of spatial consistency* During the evaluation of the climatological normals for the stations of the 2008, 2010, and 2011 versions (based on ca. 50,650, 64,400, and 67,200 stations, respectively), the normals, as well as maximum and minimum precipitation for the period 1951–2000 at each station, have been calculated and checked for spatial consistency using kmz files for visualization in Google-Earth. Dubious cases, i.e., stations not fitting in the surrounding stations, have been checked manually by cross-checking with neighbouring stations. Some of these spatial inconsistencies could be traced back to:

1. Misplaced stations caused by erroneous geographical information
2. Individual errors (typing errors, factor 10 errors) causing an erroneous climatological normal (for the calendar month) or maximum/minimum
3. Quasi-systematic errors (errors in conversion of units inch, feet, millimeters; often only for part of the time series, i.e., some years)
4. In some specific cases, data for some months/years have been found to be shifted by 1, 2, or more months (showing up in regions with a pronounced annual cycle of precipitation) or even a whole year in some cases.

For example, we detected in this step that in the Africa data set (Nicholson 1986, 1993, 2008), the precipitation data for the three stations in Lesotho with long time series (Mohale's Hoek, Qacha's Nek, Teyateyaneng) had been shifted by 9 months over the entire period 1900–1969. Without correction the data would have been significantly out of phase with a maximum in the annual cycle in spring instead of winter.

Through joint visualization of climatological normals and the underlying orography, a number of seemingly “outliers” could be attributed to true specific orographic or climatic conditions.

*Check of temporal homogeneity* The homogeneity over time of the preliminary versions of the Full Data Reanalysis (V.4, V.5) in 0.5° grid resolution was tested by

applying a moving *t*-test (H. Österle, PIK, 2008, 2010, personal communication). This tool checks the homogeneity over time and allows deeper analysis of the data set in this regard and revealed significant inhomogeneities ( $t \geq 6$ ) in some regions (Fig. 7). In cases where the inhomogeneities could be clearly traced back to errors in the precipitation data or station meta information, these errors have been corrected in the data bank. Figure 7 gives examples of the different types of errors found such as typing errors, factor 10 errors (stations Barisal, Maijde Court, and Comilla in Bangladesh), errors in conversion of units (i.e., inch, feet, millimeters; often only for part of the time series as for station Save/Benin), or “0” instead of missing values over a few years (national data for Cantillan/Philippines or Puerto Puyuhuapi and Coyhaique/Chile) causing inhomogeneities.

In other cases (mainly in data-sparse areas such as the high latitudes), inhomogeneities can also be caused by changes in the composition of the stations contributing to the gridded analysis results.

On the basis of the thorough experience with the QC of the station data from the different sources and statistical evaluations, the GPCC has set up a priority scheme according to which data are being selected for its analysis products if data are available from more than one source. The data source priority according to data quality (highest to lowest) is:

1. National
2. CLIMAT (since 1967)
3. GHCN
4. CRU
5. FAO
6. Regional
7. CLIMAT (prior to 1967)<sup>1</sup>
8. SYNOP-based GPCC
9. SYNOP-based CPC

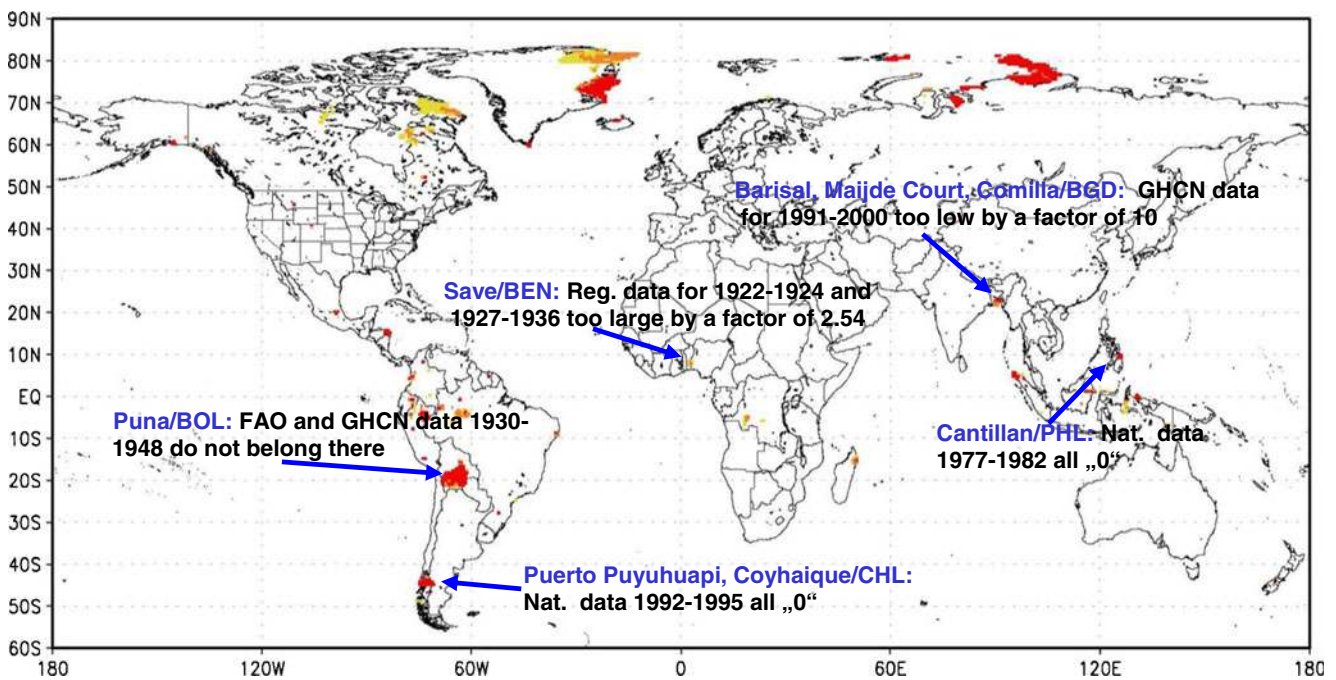
New releases of GPCC's global gridded precipitation climatology and of the Full Data Reanalysis (V.4, V.5, and V.6) have been generated after the data have been selected according to this priority criterion and the merged time series from the full data base have passed these QC procedures.

## 5 Discussion

### 5.1 GPCC's land surface precipitation climatology

GPCC's land surface precipitation climatology (version 2011) is focusing on the target period 1951–2000. With precipitation normals of ca. 67,200 stations (see Fig. 4), it

<sup>1</sup> The CLIMATs prior to 1967 are less reliable and generally only have a coding accuracy of 10 mm.



**Fig. 7** Examples of significant inhomogeneities ( $t \geq 6$ ) and accompanying data errors detected in the preliminary version of the Full Data Reanalysis (V.4) at  $0.5^\circ$  grid resolution by a moving  $t$ -test (5-year subperiods)

has a much denser data coverage than the earlier existing data sets (see Section 2.2) and outperforms most of them by a factor of 2 to 5 (i.e., LW1990 overall used data of 24,635 stations over land). Therefore, features of the precipitation climatology described in the following are often more detailed and spatially more confined than in previous analyses.

Mean annual precipitation totals in millimeter per year for the land surface shown in Fig. 8 are generally largest along the Inter-Tropical Convergence Zone (ITCZ) in response to intense surface heating and the confluence of the tropical easterlies (LW1990). However, even in the tropics, it is modified by the distribution and differential heating of land and ocean and sea surface temperatures of ocean currents adjacent to the land areas (warm currents increase both the buoyancy and moisture content of the air, whereas cold currents have the opposite effect). Orography over land is also playing an important role together with the direction of the prevailing wind, resulting in increased precipitation on the windward sides of mountains and drier conditions on the leeward sides owing to the rain-shadow effect.

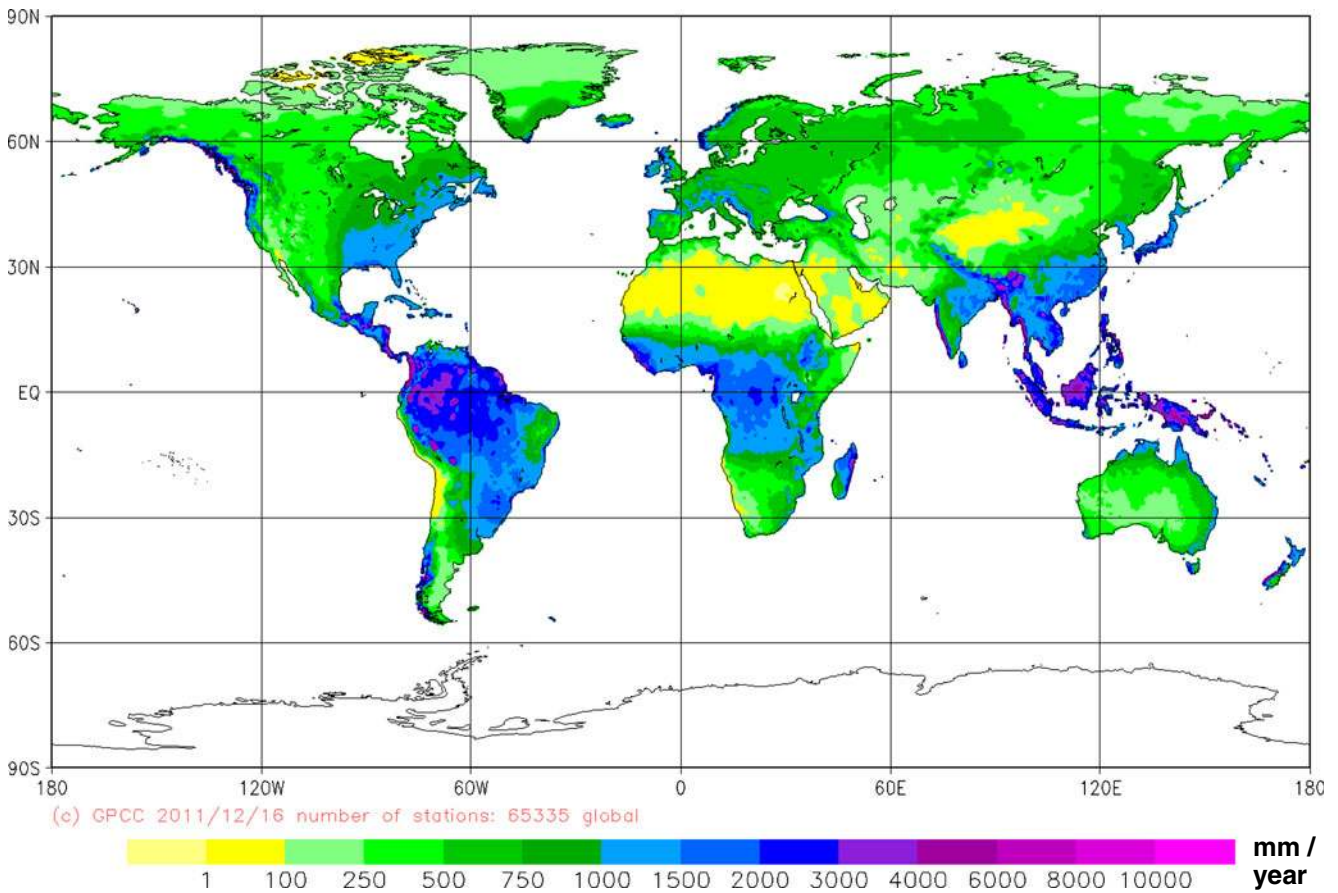
Highest annual gridded precipitation amounts of more than 3,000 mm (more than 8.2 mm per day) are found in easternmost India, in Arunachal Pradesh, and especially Meghalaya. Cherrapunji located on the southern edge of the Khasi mountains (elevation 1,313 m) is the regular observing station with the world's highest monthly and annual precipitation amounts with an average annual total of more than 11,000 mm and ca. 3,000 mm in July. The record monthly (annual) precipitation amounts have been

observed in 1974 with 8,205 (24,555) mm for July (the entire year). Very high annual precipitation amounts are also observed in narrow bands along the west coasts of the Indochinese peninsula (Myanmar, Thailand, and the northern part of the Malaysian peninsula) and along the west coast of India (Western Ghats), mainly caused by the summer monsoon rainfall there, whereas it is much drier in the rain-shadows to the leeward side. The maritime continent Indonesia, surrounded by very warm waters, is receiving heavy rainfall, too, with annual precipitation amounts exceeding 3,000 mm in the interior of Borneo, Sulawesi, and New Guinea, the west coast of Sumatra, the western parts of Java, and for some of the Micronesian islands located along the South Pacific Convergence Zone. This is the case for the eastern Philippines, too.

Part of the rain forests of Amazonia (mainly in Brazil and Colombia, a few spots in Peru and Bolivia) experience more than 3,000 mm rainfall per year, as well as the Pacific coast of Colombia and the near-equatorial Atlantic coast of Brazil and French-Guyana and some coastal regions in central America (Costa Rica, Nicaragua, Honduras, and the region Minatitlan in southern Mexico).

Although there is a broad band of heavy rainfall across tropical Africa, there are only two areas where the threshold of 3,000 mm rainfall per year is exceeded, namely, the Atlantic coasts of Guinea, Sierra Leone, and Liberia and the northern part of Madagascar's east coast.

Heavy precipitation also occurs over the western edges of the continents in the upper mid-latitudes (marine West Coast



**Fig. 8** Mean annual precipitation (mm/year) on a 0.25° grid from the new GPCC precipitation climatology released in Dec. 2011 based on ca. 67,200 stations

climates); regions such as the Pacific coasts of British Columbia and of adjacent southern Alaska receive even more than 3,000 mm per year, similar to the Pacific coast of southern Chile and the west coast of New Zealand's Southern Island. Noteworthy in this context is also the enhanced precipitation along the Norwegian coast and southern Iceland.

Descending air in the subtropics, associated with the poleward arm of the Hadley Cell, favors dry conditions in the hot desert regions—deserts are generally defined as regions with less than 250 mm annual precipitation. Well-known desert regions like the Sahara, Arabian desert, Gobi, Namib, Kalahari (in part a semi-desert with more than 250 mm per year), Lut (southeastern Iran), Atacama, US southwest, and Australian Outback clearly show up in Fig. 8. In part these dry conditions are enhanced by adjacent cold ocean currents (i.e., the Atacama desert and the Humboldt current off the South American west coast, the Namib and the Benguela current off the west coast of Southern Africa) and extended over adjacent ocean regions.

Precipitation is also light in polar regions, largely owing to the diminished moisture capacity of cold air; Antarctica with an average precipitation of 166 mm per year (see

Table 1 Vaughan et al. 1999) is the world's largest desert with  $14 \times 10^6 \text{ km}^2$ , even larger than the Sahara with  $9.1 \times 10^6 \text{ km}^2$ . Antarctica is left blank in Fig. 8 since the data base is very poor there and most of the few stations are located at or near the coast where precipitation is relatively high, whereas it is very dry in the interior of Antarctica. Any interpolation would therefore result in unrealistic high precipitation amounts over Antarctica as is the case for LW1990, which are far off with 306.1 mm (even 604.6 mm after correction for the systematic gauge-measuring error; see Table 1).

Mountainous regions (e.g., the Rocky Mountains and the Alps) exhibit average precipitation amounts, but with a large spatial variability, while average precipitation is more homogeneous over level terrain (e.g., across the Australian Outback, the Tibetan Plateau, and central North America).

When looking at the individual months, some fine structures of the precipitation field become clearer, which are, in part, somewhat blurred in the map of the annual total because of the migration of the precipitation regimes with the annual cycle of atmospheric circulation patterns. Figure 9 shows the seasonal variation for January and July as examples for Northern winter/summer when the atmospheric circulation patterns and the associated precipitation regimes

**Table 1** Large-scale averages of annual total precipitation for different data sets

	Jäger 1976	LW1990 meas.	LW1990 corr. (mean corr. factor)	CRU CL2.0 2002	GPCC 2011 meas.	GPCC 2011 corr.	GPCP V2.2	ERA-40	ERA-Interim
Time period	1931–1960	Various	Various	1961–1990	1951–2000	1951–2000	1988–2010	1958–2001	1979–2010
Resolution	5°	0.5°	0.5°	0.5°	0.5°	0.5°	2.5°	1°	1°
Land (excl. Antarctica)	816	776	845 (1.088)	793	788	850 <sup>a</sup>	846	887	879
Antarctica	182	306	605 (1.975)			166 <sup>b</sup>	193	154	173
Land (overall)	756	735	824 (1.120)		723	786	789	823	818
Ocean	1,099	1,099	1,212 (1.103)				1,010	1,227	1,156
Global	1,000	1,027	1,138 (1.109)				978	1,149	1,086

<sup>a</sup> Corrected for the systematic gauge-measuring error by applying 0.85 times the LW1990 correction

<sup>b</sup> Determined by net surface mass balance, equivalent to 166 mm per year (Vaughan et al. 1999)

are approaching their southernmost/northernmost positions during the annual cycle.

Additionally, there are significant seasonal variations in the regions affected by monsoons, such as India and South East Asia. The heavy rainfall associated with the Indian and Southeast Asian summer monsoon is very pronounced for July (Fig. 9, bottom) and confined to the Western Ghats of India and the west coast of Southeast Asia.

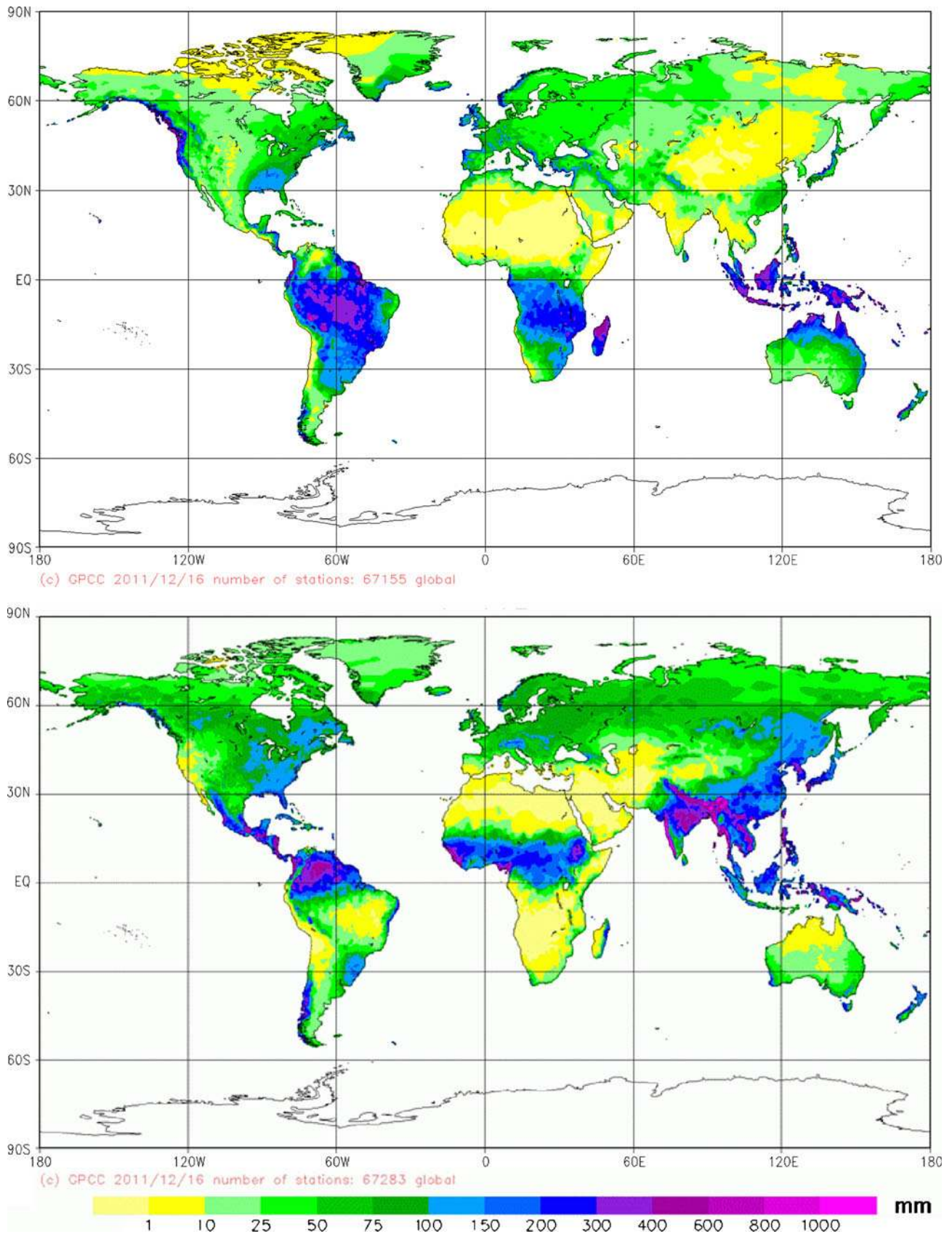
In contrast to the Namib and most other deserts being dry throughout the entire year, the Kalahari is dry normally only from May to September, whereas it is experiencing rainfall from October to April, characterizing it as a semi-desert with between 250 and 400 mm per year. The Saharan drought is mainly shifting in meridional direction together with the atmospheric circulation patterns but is also more pronounced and larger in extent in Northern winter when only the southernmost parts of countries adjacent to the Gulf of Guinea experience some rain, while in a huge area on average less than 10 mm, and even less than 1 mm, rainfall are received. To the north, it is confined by the rainfall in the Atlas mountains of Morocco, Algeria, and Tunisia, being most pronounced in winter. In Northern winter, the Mediterranean coasts of Libya and Egypt are getting some rain, too.

The desert conditions related to the Gobi are most pronounced in Northern winter, stretching then from the northern part of the Tibetan plateau northeastward via Mongolia into Russia up to about 60° N and 135° E. In summer, it is much more confined to Mongolia and the adjacent western part of China by the heavy rainfalls accompanying the Indian and Southeast Asian summer monsoons reaching far inward into the Asian continent.

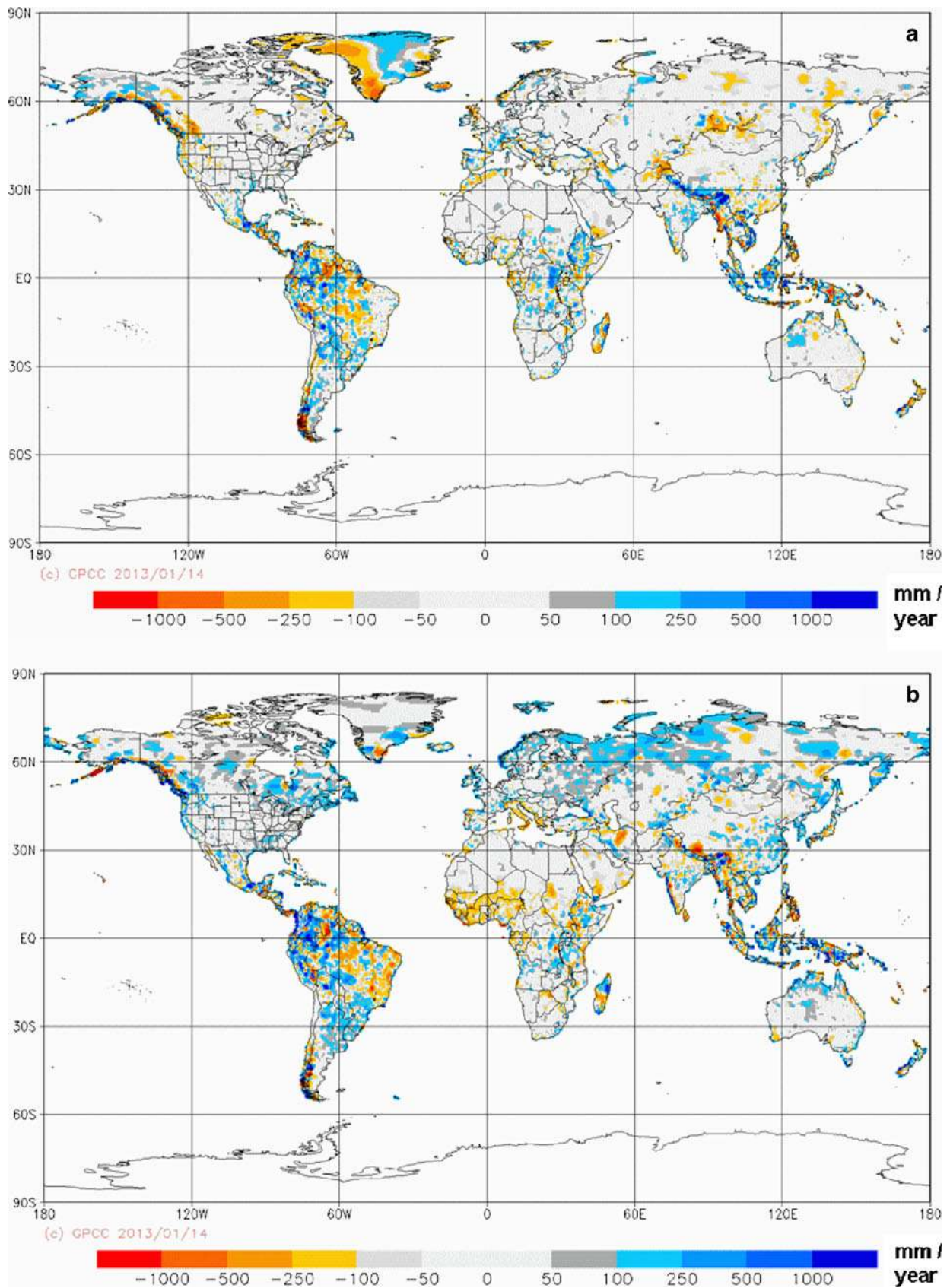
## 5.2 Comparison of GPCC's precipitation climatology to other climatological data sets

After construction of GPCC's new gridded precipitation climatology (0.25° resolution), it has been compared visually to the climatic atlases of WMO/UNESCO for Europe, South America, North and Central America, and Asia (WMO/UNESCO 1970, 1975, 1979, 1981). These isohyetal analyses by WMO/UNESCO have been performed by climatological experts using a large number of station data and taking orography into account. GPCC's new gridded precipitation climatology being available in four different grid resolutions of 0.25°, 0.5°, 1°, and 2.5° (Meyer-Christoffer et al. 2011a, b, c, d) has been found to be in very good agreement with these climatic atlases.

All averages (terrestrial, oceanic, and global) to be discussed later for the different precipitation data sets (GPCC, GPCP, ECMWF model reanalyses) are calculated in a consistent way using a weighting for the grid area (latitude dependent). In addition to that, the terrestrial



**Fig. 9** Mean precipitation (mm) on a 0.25° grid for January (*top*) and July (*bottom*) from the new GPCC precipitation climatology released in Dec. 2011 based on ca. 67,200 stations



**Fig. 10** Differences of annual total precipitation (mm/year) between GPCC's climatology at the corresponding resolution and **a** CRU CL2.0 aggregated to 0.5°, **b** LW1990 at 0.5°, **c** Jäger 1994, personal communication at 2.5°



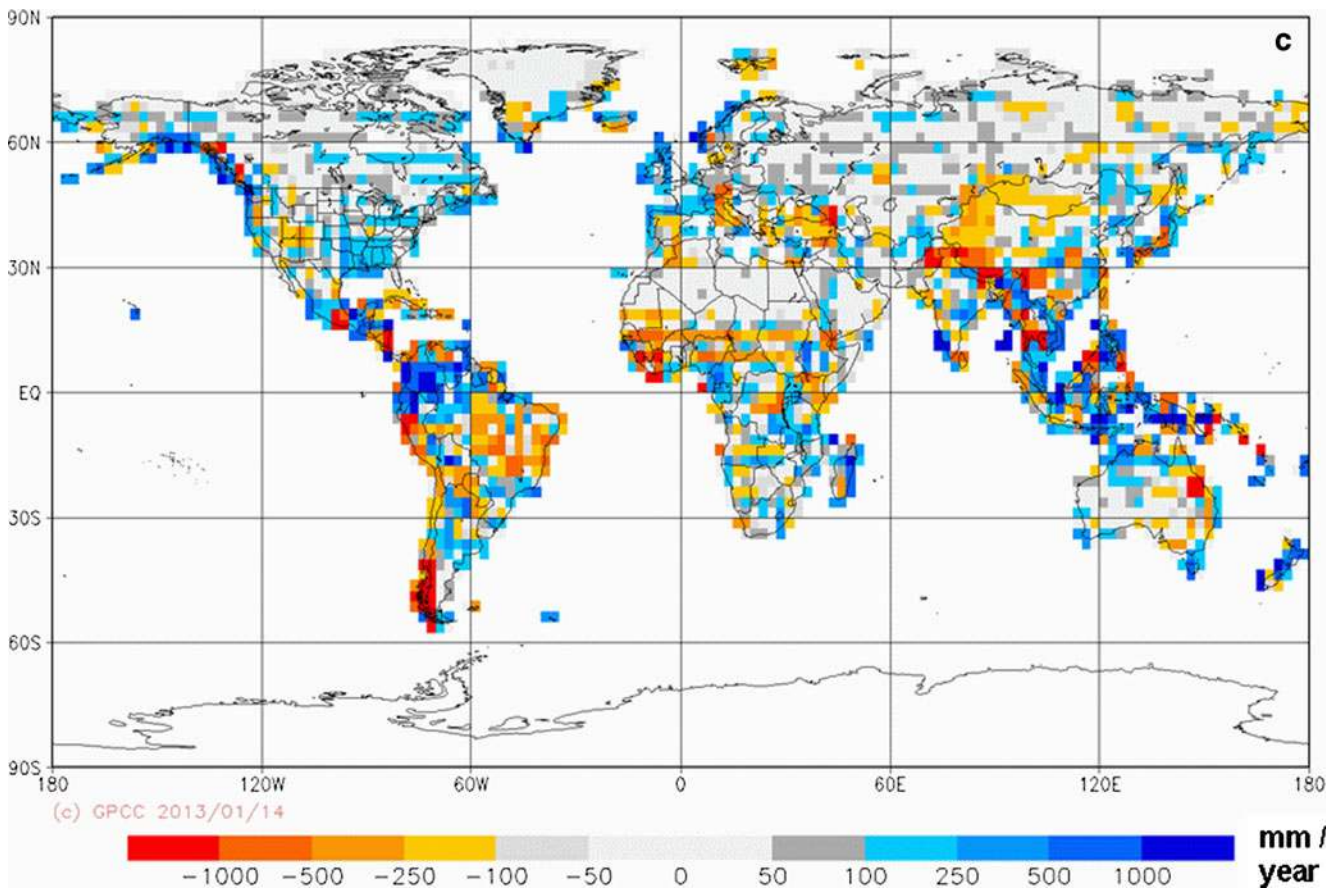


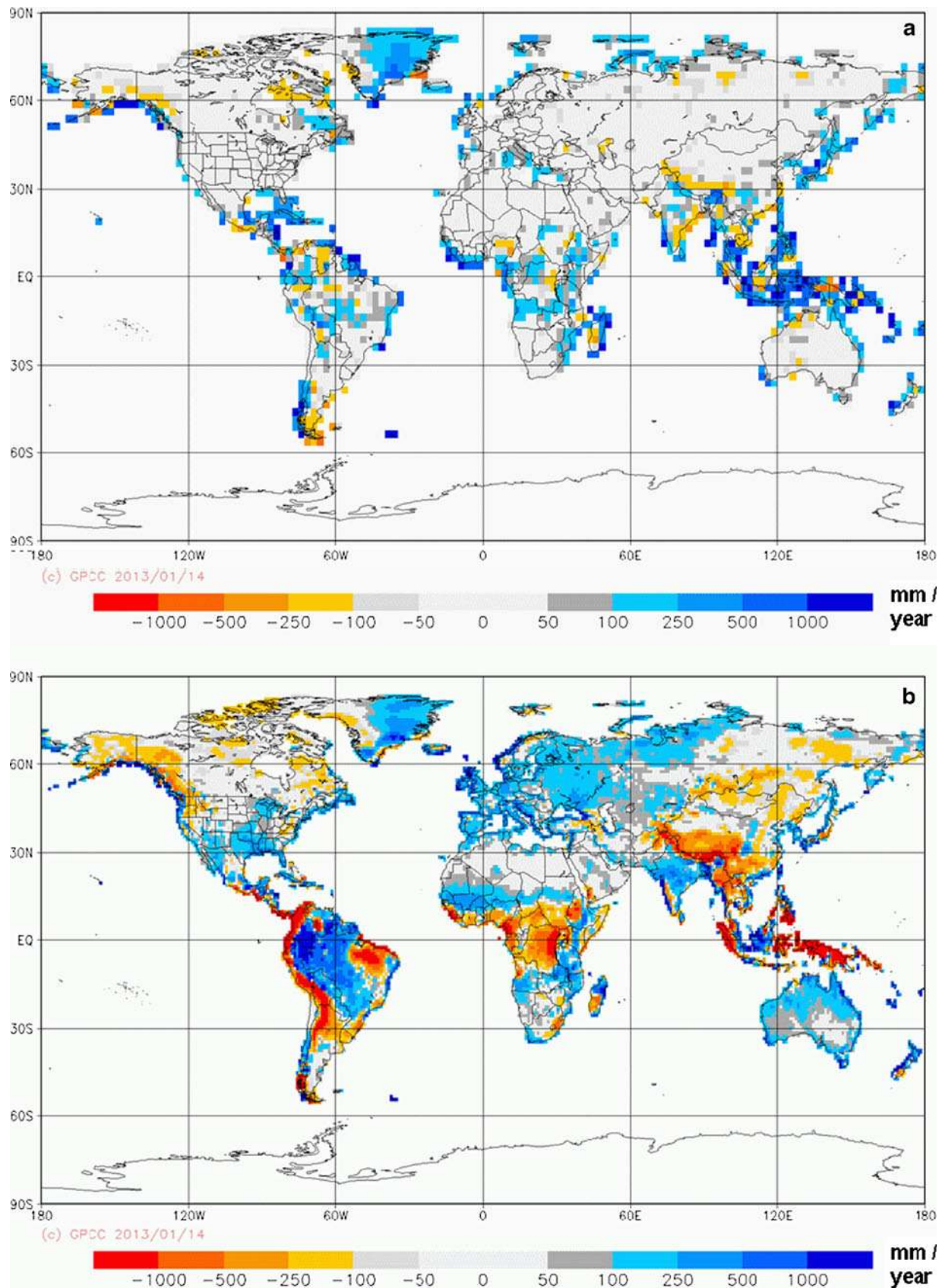
Fig. 10 (continued)

(oceanic) averages are calculated by weighting each grid value according to its land (ocean) fraction depending, for example, on the underlying land mask and the spatial resolution used. To construct the land masks and to determine the land (ocean) fraction of each grid for the different grid resolutions of 0.25°, 0.5°, 1°, and 2.5°, the GPCC utilized information from the Global Land Data Assimilation System (Rodell et al. 2004).

*Spatial distribution of the differences of the annual totals* Differences between GPCC's new precipitation climatology (uncorrected for the systematic gauge-measuring error) and the climatologies of Jäger 1994, personal communication for 1931–1960, LW1990 (uncorrected) for various periods, and CRU CL2.0 for 1961–1990 (New et al. 2002) have been calculated for each calendar month and for the year and are shown in Fig. 10 for the annual totals.

The CRU CL2.0 data set (originally at 10' resolution, but based on just about 11,800 stations in contrast to the 67,200 stations used in GPCC's climatology) has been averaged to a 0.5° grid to be comparable to GPCC's climatology with the same resolution; their differences are shown in Fig. 10a. There is no significant bias between both data sets (the terrestrial average (excluding Antarctica) for CRU CL2.0

is 793 mm compared to 788 mm for GPCC's climatology), and the differences between both data sets are relatively small with generally less than 50 mm per year. Larger differences occur preferably in the tropics, especially in South America, but also in tropical Africa and Southeast Asia/Indonesia. In regions with orographically induced heavy coastal rainfall, but drier inland conditions (i.e., Western Ghats in India, west coasts of Myanmar and Thailand, of New Zealand and the Pacific coast of Canada), the areas of heavy rainfall are much more confined to the coasts in GPCC's precipitation climatology, and the drier conditions in the leeward rain-shadow are better resolved (owing to GPCC's much denser station network). The difference patterns indicate that in the CRU CL2.0 analysis the coastal rainfall is more smoothed towards inland owing to the sparse station network resulting in CRU's higher inland precipitation estimates there. This holds for the Atlas mountains in Northern Africa, too. Precipitation along the slopes of the Himalaya is much more pronounced in GPCC's analyses, where GPCC also has much more data (we have for instance data of 270 stations for Nepal, 25 for Bhutan). Orographically induced precipitation in other areas is better represented in GPCC's climatology, e.g., along the Ural or at the Iranian coast of the Caspian Sea.



**Fig. 11** Differences of annual total precipitation (mm/year) between GPCP's climatology corrected for the systematic gauge-measuring error at the corresponding resolution to **a** GPCP V2.2 1988–2010 at 2.5°, **b** ERA-40 1958–2001 at 1°, **c** ERA-Interim 1979–2010 at 1°

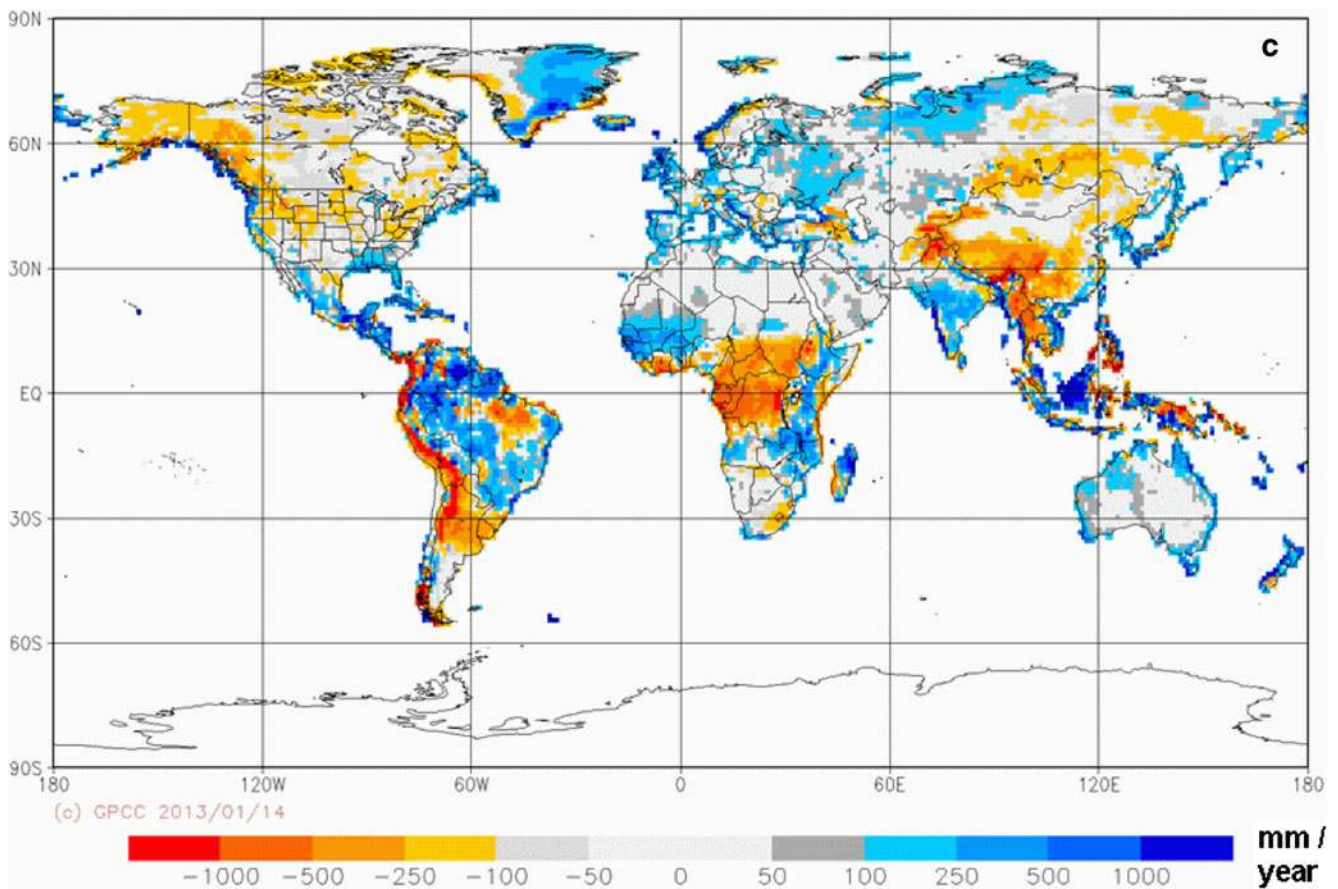


Fig. 11 (continued)

The terrestrial, oceanic, and global averages calculated in GPCC's consistent way from the bias-corrected LW1990 climatology on a 0.5° grid (based on 24,635 stations) given in Table 1 with 824 mm over land and 1,138 mm for the globe differ somewhat from the averages given by Legates (1987) with 820 mm and 1,123 mm, which might be due to a different weighting of area and land fraction in the averaging. The LW1990 climatology has a slightly lower terrestrial average precipitation (excluding Antarctica) of 776 mm than GPCC's climatology. Overall, there are somewhat larger differences to GPCC's climatology (Fig. 10b); in regions with heavy coastal rainfall, but drier inland conditions (i.e., Western Ghats in India, west coasts of Myanmar and Thailand, and the Pacific coast of Canada and the southern coast of Alaska), there are similar difference patterns to GPCC's climatology than in CRU CL2.0 caused by a smoothing due to under-sampling but less pronounced because of the somewhat better data coverage in LW1990. The ITCZ seems to be located somewhat farther north over West Africa than in GPCC's analysis; since there is no such difference between GPCC and CRU CL2.0, this difference feature might be attributed to different time periods being used in the LW1990 climatology. Apart from the biggest

differences also occurring in the tropics, there are larger differences in the northern latitudes, too, with GPCC generally having higher precipitation amounts across large parts of Russia, Belarus, Ukraine, Scandinavia, Canada, and the northeastern USA. This positive precipitation bias in high northern latitudes, which are more affected by the systematic gauge-measuring error, is probably the reason that applying LW1990's correction that accounts for 8.8 % over land (Table 1) in their assessment has a larger effect of 9.3 % in GPCC's climatology. Applying 85 % of their bulk correction results in an overall correction for the systematic gauge-measuring error of 8 % for the terrestrial average (excluding Antarctica); for land (overall), LW1990 has an average correction of 12.0 %.

Jäger (1976), in his original analysis on a 5°-grid gave average precipitation, as shown in Table 1, for land (excluding Antarctica) with 816 mm, Antarctica 182 mm, and land (overall) 756 mm and arrived, with an ocean estimate of 1,099 mm, at 1,000 mm for the globe. The precipitation averages over land are similar to the ones calculated on the basis of a 2.5° version of his climatology that he provided to the GPCC after a visit at GPCC (L. Jäger 1994, personal communication) with 810 mm over land excluding

Antarctica, land (overall) 755 mm. In contrast to that over ocean, the precipitation average of 999 mm is significantly lower (by 10 %) than the 1,099 mm given in Jäger (1976), leading to a global average of 967 mm, almost identical to the average for Jäger's climatology of 966 mm as mentioned in Legates (1995). According to the discussion in Legates (1995, p. 5), Jäger enhanced the oceanic precipitation since he felt the global average of 967 mm to be too low and the oceanic precipitation to be the least reliable, and so arrived at a global average of 1,000 mm. This clearly underlines the uncertainties in estimating oceanic precipitation, especially in the pre-satellite era. The differences between GPCP's climatology and the one of Jäger shown in Fig. 10c are generally quite large, with Jäger having less precipitation over large parts of Canada and the northeastern USA, the Pacific coast of Canada, and the south coast of Alaska, while he has more precipitation, for example, over Patagonia.

For comparison with the GPCP V2.2 data set, as well as with the ECMWF model reanalyses ERA-40 and ERA-Interim, GPCP's precipitation climatology has been "corrected" for the systematic gauge-measuring error by simply applying the bulk correction factors of LW1990 for each calendar month and for the year; the differences of the "corrected" GPCP climatology to the average annual precipitation for GPCP V2.2 and for ECMWF's model reanalyses ERA-40 and ERA-Interim are displayed in Fig. 11, where for ERA-Interim we used the 12–24-h forecasts provided by ECMWF (A. Simmons 2012, personal communication) to minimize model spin-up effects.

Overall, GPCP's climatology is in very good agreement with the GPCP V2.2 data set for 1988–2010; the differences are small in the interior of the continents with generally less than 50 mm, which is not very astonishing since GPCP's analyses are used to adjust the satellite-based estimates, especially in regions of dense coverage by rain gauges. Largest differences occur in tropical South America, Africa, and especially Indonesia, where the gauge sampling is somewhat poor, and in the coastal regions where the satellite-based estimates come strongly into play.

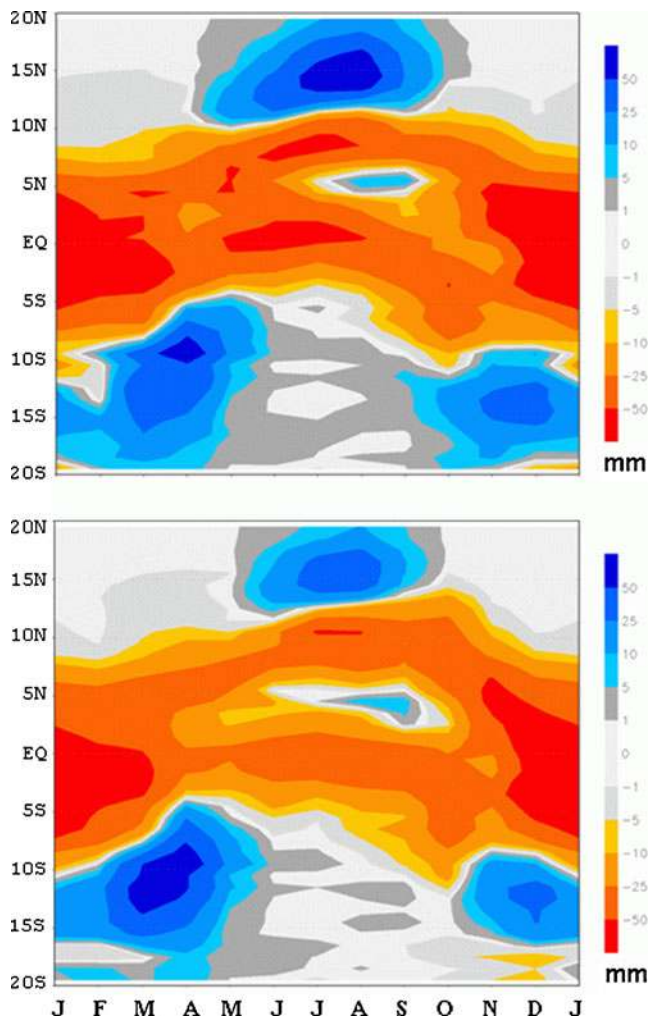
The differences between GPCP's climatology and the ERA-40 mean for 1958–2001 are large (50 mm or more per annum) almost everywhere over the land areas showing large-scale coherent structures. GPCP has more precipitation than ERA-40 in a broad belt extending from the British Isles and the Iberian Peninsula over Central Europe far into Russia (to about 100° E) and Kazakhstan. Over the Tibetan plateau and the Hindu Kush, Pamir, and Karakorum, ERA-40 has significantly more precipitation than GPCP's climatology; in these sparsely inhabited areas, GPCP's analysis is suffering from poor data coverage and is not very reliable there. Contrary to that, along the Western Ghats and for almost the entire northern and eastern parts of India and for Bangladesh, ERA-40 has a negative rainfall bias. Over large parts of Southeast Asia,

namely, southwest China, Myanmar, Thailand, and Laos, and also over most of the maritime continent Indonesia, ERA-40 is overestimating rainfall. In contrast to that, the model over Australia and for most parts of New Zealand has a negative precipitation bias. Over Africa, the model is overestimating precipitation in the inner tropics, whereas it has a negative bias at the northern and southern boundaries of the ITCZ. The spatial distribution of the differences indicates that the amplitude of the annual meridional migration of the ITCZ is underestimated, keeping it too close to the inner tropics. This is highlighted in Fig. 12 showing zonal means of the differences between GPCP's climatology and ERA-40 (top) and ERA-Interim (bottom) averaged for 10° W to 40° E for each 1° latitude band and their mean variation over the year. Starting around October until May, the ITCZ in the ERA reanalyses is moving not far enough to the south, and vice versa in Northern summer (June to September), resulting in an ITCZ being too confined to the inner tropics over Central and Western Africa. This effect, being slightly reduced in the ERA-Interim reanalysis (Fig. 12, bottom), was also noted in Schneider et al. (1992a) for the then operational ECMWF model T106.

Large differences occur also over South America, with ERA-40 having significantly less precipitation over large parts of Amazonia (except for Northeast Brazil), but heavily too much precipitation at the South American west coast in a band extending from northern Chile and Argentina along Bolivia, Peru, and Ecuador all the way to Colombia, with differences in some regions even exceeding 500 mm per year.

Over North America, the ERA-40 precipitation is too low along the Atlantic coasts of the USA and Canada (from Florida up to Newfoundland), as well as in a broad area ranging from northern Mexico via the southwest USA and the states along the Gulf Coast up to the Great Lakes. In contrast to that over northeast Canada, the western USA and Canada and Alaska, apart from the Pacific coasts, ERA-40 has more precipitation (up to more than 250 mm) than GPCP's climatology. Similar patterns of the differences between average precipitation for 1979–2001 of ERA-40 and GPCP V2.2 (in mm/day), although less pronounced, are found in Bosilovich et al. (2008).

The differences between GPCP's corrected climatology and ERA-Interim are generally smaller than to ERA-40; there are more regions with differences below 50 mm per year, especially in Europe, western Russia, and Kazakhstan. The agreement is also much better over Australia. The differences have remained similar over Southeast Asia. Over Africa, the structure of the differences might still indicate, although slightly less pronounced than for ERA-40, that the amplitude of the seasonal meridional oscillation of the ITCZ is somewhat underestimated, especially in West Africa (see also Fig. 12). Although the overall structure of the differences has remained similar over South America, the differences



**Fig. 12** Differences (mm) of zonal mean precipitation (10° W to 40° E) between GPCC's climatology corrected for the systematic gauge-measuring error and climatologies derived from (*top*) ERA-40 1958–2001 and (*bottom*) ERA-Interim 1979–2010

have somewhat diminished. Over North America, ERA-Interim is overestimating precipitation even stronger than ERA-40 with somewhat increased differences over almost the entire Canada, the USA (apart from some Gulf states and the Atlantic Coast), and Alaska. The differences of ERA-Interim to GPCC's climatology, though generally smaller than for ERA-40, are overall significantly larger than for GPCP V2.2.

*Comparison of large-scale averages* Huffman et al. (2009) determined the land, ocean, and global precipitation averages for GPCP V2.1 as 923, 1,015, and 978 mm per year, respectively, with the terrestrial estimate being way too high because defining all grids with even a tiny fraction of land in it completely as land gives the terrestrial average a strong “maritime bias”. GPCC in its calculations is weighting each grid exactly according to its area and land fraction for the terrestrial average and vice versa for the oceanic average.

Therefore, at GPCC, we use our own software to calculate these large-scale averages so that all terrestrial, oceanic, and global averages presented are calculated in a consistent way and comparable to each other.<sup>2</sup> The time periods differ for the various sources (see Table 1) so that the results are not exactly comparable, but as long as the periods are of a sufficient length (more than 20 years), we regard the related error as negligible.

Jäger (1976), with his estimate of 182 mm for the average annual precipitation over Antarctica, came quite close to what can be assumed the best estimate of average Antarctic precipitation of 166 mm determined from the net surface mass balance by Vaughan et al. (1999). LW1990, by interpolating the data of the few Antarctic stations almost all located in the relatively wet coastal zones of Antarctica and on the Antarctic peninsula, has heavily overestimated precipitation with 306 mm (uncorrected) and, after correction for the systematic gauge-measuring error, arrived at even completely unrealistic 604 mm! Therefore, the GPCC refrains from providing an analysis over Antarctica because any interpolation of stations located mainly at or near the Antarctic coast or on the Antarctic peninsula would also yield unrealistic results. Because of a similarly poor data coverage over Greenland, the analysis in the interior of Greenland is not reliable.

The terrestrial average (excluding Antarctica) annual precipitation derived on analyses of such a densely sampled station data set like GPCC's (see Fig. 5) could be assumed reliable and accurate. Even though there may be significant differences in the releases of the GPCC Full Data Reanalysis V.6 to the previous ones (V.4 or V.5) in individual months and in regions where a lot of data have been added during the last years or owing to corrections of individual data errors detected in the repeated QC processing, the differences are generally less than a few millimeters for the different versions of the background climatologies for the target period 1951–2000 with V.2011 being based on 67,200 stations, V.2010 (64,400 stations), or V.2008 (50,750 stations). The uncorrected annual total of the terrestrial average (excluding Antarctica) has stayed the same in the climatology V.2011 with 788 mm (see Table 1) as in both previous versions.

Taking into account the estimate of Vaughan et al. (1999) of 166 mm for the average precipitation over Antarctica, the uncorrected overall terrestrial average precipitation is 723 mm per year. Applying 85 % of the bulk climatological correction from LW1990 for each calendar month results in a best estimate for the terrestrial average with (without)

<sup>2</sup> Using software packages (i.e., CDO (Schulzweida et al. 2011)) for the calculation of large-scale averages, one has to be very careful on how the averaging is done since one might end up getting inconsistent results otherwise.

Antarctica of 786 mm (850 mm) for the annual precipitation total.

Clearly the biggest uncertainty issue is the correction of the systematic gauge-measuring error (general undercatch of the true precipitation). We estimate the uncertainty of the correction for the systematic gauge-measuring error as the difference between the upper (lower) boundaries for the correction factor. With a best estimate of 0.85 for the correction factor, we estimate the upper/lower boundaries as a factor of 1.0/0.7 (applying 100/70 % of the correction of LW1990), resulting in an uncertainty range of  $\pm 10$  mm for our best estimate of 786 mm for the terrestrial annual precipitation from our new precipitation climatology. This is close to the GPCP V2.2 terrestrial average for the period 1988–2010 with 789 mm; GPCP's combined satellite gauge analyses are adjusted to GPCC's analyses (Full Data Reanalyses for 1988–2010 and the Monitoring Product thereafter), and the full bulk correction for the systematic gauge-measuring error according to LW1990 is applied.

A new method for the correction of the systematic gauge-measuring errors, taking into account the weather conditions in the evaluation month (wind, temperature, relative humidity, precipitation phase and type), has been developed at GPCC and is described in Fuchs et al. (2001); this improved method has been operationally applied at GPCC for the analyses since January 2007. The correction factors resulting from the new method that is applicable for the ca. 6,000–7,000 synoptic stations worldwide, for which the necessary weather information is available, are usually smaller than the bulk climatological correction by LW1990 that tends to overestimate the systematic gauge-measuring error (Fuchs et al. 2001). Since results of this more realistic approach for the correction of the systematic gauge-measuring errors are currently available only for the short period since 2007, we are using 0.85 times the correction of LW1990 to compensate the GPCC climatology for the systematic gauge-measuring errors as long as more realistic weather-dependent correction factors are not available for a sufficiently long period.

**Table 2** Areal extent for the Earth's surface, land, ocean, antarctica and land excluding Antarctica according to CIA (2012)

Description	Area (in $10^6 \text{ km}^2$ )
Earth's surface	510.072
Ocean	361.132 (70.8 %)
Land	148.94 (29.2 %)
Antarctica	14.00
Land (excluding Antarctica)	134.94

### 5.3 Relation of GPCC's precipitation climatology to the global water cycle

The global, oceanic, and terrestrial water exchanges (transports) as part of the hydrological cycle are converted into volumetric sizes of precipitation and evaporation/transpiration by using the areal extents given in Table 2.

Chahine (1992), in his estimation of the global hydrological cycle, did relatively well for the average global precipitation of 990 mm being equivalent to a transport of  $505,000 \text{ km}^3$  but significantly underestimated the terrestrial precipitation with 718 mm (a transport of  $107,000 \text{ km}^3$ ) and the moisture advection from ocean to land with  $36,000 \text{ km}^3$ ; on the other hand, he overestimated oceanic evaporation (1,202 mm) and precipitation (1,102 mm). Baumgartner and Liebscher (1996) with a water vapor transport of  $39,700 \text{ km}^3$  from ocean to land (see Table 3), and in compensation a similar river discharge, came close to the recent estimates for terrestrial and oceanic precipitation given by Trenberth et al. (2007) for the period 1979–2007 or adjusted for the period 2002–2008 in Trenberth et al. (2011) and Trenberth and Fasullo (2012). The global runoff, balancing the excess evaporation over precipitation over the oceans (P–E negative) and the atmospheric water vapor transport from ocean to land, is given as ca.  $40,000 \text{ km}^3$  in both papers. The estimate for the “global” runoff (excluding Antarctica and Greenland) given in Dai et al. (2009) of  $37,288 \text{ km}^3$ , together with a runoff for Antarctica (Jacobs et al. 1992) of  $2,613 \text{ km}^3$  and for the Greenland ice sheet (Losev 1973) of  $610 \text{ km}^3$  ( $280 \text{ km}^3$  attributed to iceberg calving and  $330 \text{ km}^3$  to water runoff), would result in a somewhat higher global runoff of about  $40,500 \text{ km}^3$ . However, recent estimates by the Global Runoff Data Centre (de Couet and Maurer 2009; U. Loser 2012, personal communication) indicate a lower “global” runoff (excluding Antarctica and Greenland) of only  $36,109 \text{ km}^3$  that would result in a global runoff of only about  $39,500 \text{ km}^3$ . Since there is also a small contribution by coastal discharge through groundwater (Dai et al. 2009), we consider  $40,000 \text{ km}^3$  as still a reliable estimate for the global runoff as well as—for balance reasons—for the mean water vapor transport from ocean to land.

For the satellite gauge combined GPCP V2.2 data set (Huffman and Bolvin 2012), the estimate for average global precipitation over the period 1988–2010 is 978 mm per year, with an terrestrial average (excluding Antarctica) of 846 mm, being close to GPCC's estimate of 850 mm. Since GPCP is using GPCC's precipitation analyses to adjust the satellite-based estimates (V2.2 is using GPCC's Full Data Reanalysis V.6 for the period up to 2010 and the near real-time Monitoring Product thereafter), it is not very surprising that both are in excellent agreement over land. However, the oceanic estimate of 1,010 mm is too low in comparison with 1,069 mm equivalent to the water vapor

**Table 3** Terrestrial, oceanic, and global averages of annual precipitation (P), evapotranspiration (E), and the differences P–E in mm/year and water exchanges (transports) in 1,000 km<sup>3</sup>/year according to different authors and for different time periods

	Chahine 1992	mm/ year	Baumgartner and Liebscher 1996	mm/ year	Trenberth et al. (2007) 1979–2007	mm/ year	Trenberth et al. (2011) 2002–2008	mm/ year
Land								
Precipitation	107.0	718	111.1	746	113.0	759	114.0	766
Evaporation	71.0	477	71.4	480	73.0	490	74.0	497
P–E	36.0	241	39.7	266	40.0	269	40.0	269
Ocean								
Precipitation	398.0	1,102	385.0	1,066	373.0	1,033	386.0	1,069
Evaporation	434.0	1,202	424.7	1,176	413.0	1,144	426.0	1,180
P–E	–36.0	–100	–39.7	–110	–40.0	–111	–40.0	–111
Global								
Precipitation	505.0	990	496.1	973	486.0	953	500.0	980
Evaporation	505.0	990	496.1	973	486.0	953	500.0	980
P–E	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

transport of 386,000 km<sup>3</sup> estimated by Trenberth and Fasullo (2012), which we regard as reliable. The terrestrial average given by Trenberth et al. (2009) for the earlier GPCP V.2 data set with 2.06 mm/day (752 mm per year) has been significantly lower than the 789 mm for the recent Version 2.2.

GPCC's best estimate for the terrestrial annual precipitation from its new precipitation climatology for the target period 1951–2000 (corrected by applying 85 % of the climatological correction from LW1990) is 786 mm per year as given in Section 5.2 (Table 1), equivalent to a volumetric water exchange or transport of 117,000 km<sup>3</sup>. In consequence, to be in balance with the global runoff estimate of ca. 40,000 km<sup>3</sup>, the resulting evaporation over land has to be somewhat higher with 77,000 km<sup>3</sup> than the estimate of Trenberth et al. (2011) and Trenberth and Fasullo (2012) of 74,000 km<sup>3</sup>. With Trenberth's (2007, 2012) estimate of water exchange over ocean of 386,000 km<sup>3</sup> (1,069 mm precipitation per year), this is resulting in an estimate for the overall global water transport of 503,000 km<sup>3</sup>, being equivalent to a global average precipitation of 986 mm per year.

### 6 Concluding remarks

The necessity of a thorough quality control (QC) is evidenced by the more than 20 years of experience in the processing of observational precipitation data sets at the GPCC because almost any larger input data set—near and non-real-time—contains more or less errors, and there are all kinds of errors in the raw station data. If these errors in the data sets would not be detected and corrected, this would result in duplicate and/or mislocated stations, as well as erroneous precipitation data, and would seriously affect any derived products.

GPCC's best estimate for the terrestrial annual precipitation including (excluding) Antarctica from its new precipitation climatology based on 67,200 stations for the target period 1951–2000 (corrected for the systematic gauge-measuring error by applying 0.85 times the climatological correction from LW1990) is 786 mm (850 mm) per year, being in very good agreement with the average from GPCP V2.2 with 789 mm (846 mm). The equivalent water exchange or transport of 117,000 km<sup>3</sup> is somewhat higher than the 114,000 km<sup>3</sup> given in Trenberth et al. (2011), which for balance reasons would require a similarly higher evaporation/transpiration over the land areas of 77,000 km<sup>3</sup>, with the global runoff estimate of about 40,000 km<sup>3</sup> given in Trenberth et al. (2007) confirmed also by Dai et al. (2009) and by De Couet and Maurer (2009). Over land this would come a little closer to ECMWF's model reanalyses ERA-40 with 122,500 km<sup>3</sup> in our calculation for 1958–2001 or 823 mm (112,000 km<sup>3</sup> according to Trenberth et al. (2011)) and ERA-Interim with 121,800 km<sup>3</sup> or 818 mm in our calculation for 1979–2010 (119,000 km<sup>3</sup> according to Trenberth et al. (2011) for 2002–2008), but the models are still somewhat overestimating precipitation over land. Over ocean, especially ERA-40 generates too much precipitation of on average 1,227 mm compared to 1,069 mm being equivalent to 386,000 km<sup>3</sup> given by Trenberth and Fasullo (2012); here ERA-Interim with an average of 1,156 mm, though being still too high, is coming much closer to observations.

The uncertainty of our best estimate for the terrestrial annual precipitation (786 mm) has been estimated to be ±10 mm, mainly originating from the uncertainty in the correction of the systematic gauge-measuring error. An improved correction method for systematic gauge-measuring errors (Fuchs et al. 2001), taking into account the weather conditions in the evaluation month (wind, temperature,

relative humidity, precipitation type), has been implemented at GPCC and is operationally applied for the analyses since 2007. The correction factors resulting from the new method are usually somewhat smaller than the bulk climatological correction according to LW1990 that tends to overestimate the systematic gauge-measuring error (Fuchs et al. 2001). However, the required information for this more realistic bias estimate is available only for the ca. 6,000–7,000 synoptic stations worldwide, so reprocessing of the synoptic weather reports farther back before Jan. 2007 is a large undertaking. GPCC will tackle this problem during the next years.

With the beginning of 2012, the GPCC started the acquisition and processing of daily precipitation data and is working on the integration of the GHCN daily data set (Menne et al. 2012) and several national collections of daily data provided by the NMHSs to the GPCC into its data base system. Cross-checks are performed of the daily precipitation data against the monthly precipitation totals from the different sources already archived in the data base in source-specific slots during importing the daily data into the GPCC data base. These indicate the potential to further extend the QC processing at GPCC and to enhance the capabilities to detect errors in the raw data. A next release of GPCC's portfolio of precipitation analysis products (Precipitation Climatology, Full Data Reanalysis, Monitoring Product), which will be based upon a further enlarged and improved data base and is expected to include an improved correction of the systematic gauge-measuring error, will help to reduce the uncertainty in the estimate of mean land surface precipitation.

*Note* GPCC does not claim the copyright of the data it has gathered from its suppliers and thus is not in the position to distribute station-related observational data sets to third parties unless the data owner provides a specific allowance to do so.

**Acknowledgments** We are most grateful to the data suppliers that are to a large extent the National Meteorological and/or Hydrological Services and also some other institutes. Without these data contributions, the work of the GPCC would not be possible. We are looking forward to their further support in the future, which is crucial to maintain and enhance its level of products and services. Sharon Nicholson and Pavel (“Pasha”) Groisman are acknowledged namely for providing their unique data collections to the GPCC. We are very grateful for the contribution of Dr. Hermann Österle, Potsdam Institute for Climate Impact (PIK), who has de facto served as beta tester of the Versions 4 and 5 of the Full Data Reanalysis Product while checking them for the existence of temporal inhomogeneities. His work has substantially supported the quality control and improvement of the GPCC products.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Adler RF, Huffman GJ, Chang A, Ferraro R, Xie P, Janowiak JE, Rudolf B, Schneider U, Curtis S, Bolvin DT, Gruber A, Susskind J, Arkin PA, Nelkin E (2003) The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J Hydrometeorol* 4:1147–1167
- Baumgartner A, Liebscher H-J (1996) *Lehrbuch der Hydrologie, Band 1, Allgemeine Hydrologie—Quantitative Hydrologie*, 2nd edn. Gebr. Bornträger, Berlin, Stuttgart, 694 pp., ISBN 3-443-30002-2
- Beck C, Grieser J, Rudolf B (2005) A new monthly precipitation climatology for the global land areas for the period 1951 to 2000. *Climate Status Report 2004*, 181–190, German Weather Service, Offenbach, Germany
- Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U, Ziese M (2013) A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth Syst Sci Data Discuss* 5:921–998. doi:10.5194/essd-5-71-2013
- Bosilovich MG, Chen J, Robertson FR, Adler RF (2008) Evaluation of global precipitation in Reanalyses. *J Appl Meteor Climatol* 47:2279–2299
- Chahine MT (1992) The hydrological cycle and its influence on climate. *Nature* 359:373–380
- Chen M, Xie P, Janowiak JE, Arkin PA (2002) Global land precipitation: a 50-year monthly analysis based on gauge observations. *J Hydrometeorol* 3:249–266
- CIA (2012) The world factbook. Available online at: <https://www.cia.gov/library/publications/the-world-factbook/geos/xx.html>
- Dai A, Qian T, Trenberth KE, Milliman JD (2009) Changes in continental freshwater discharge from 1948 to 2004. *J Climate* 22:2773–2792
- De Couet T, Maurer T (2009) Surface freshwater fluxes into the world oceans. Global Runoff Data Centre, Federal Institute of Hydrology (BfG)
- Dee DP et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J Royal Met Society* 137:553–597
- FAO (2001) FAOCLIM 2.0 A world-wide agroclimatic database. Food and Agriculture Organization of the UN, Rome
- Fuchs T, Rapp J, Rubel F, Rudolf B (2001) Correction of synoptic precipitation observations due to systematic measuring errors with special regard to precipitation phases. *Phys Chem Earth (B)* 26(9):689–693
- Groisman PY, Koknaeva VV, Belokrylova TA, Karl TR (1991) Overcoming biases of precipitation measurement: a history of the USSR experience. *Bull Amer Meteor Soc* 72(11):1725–1733
- Groisman PY, Rankova EY (2001) Precipitation trends over the Russian permafrost-free zone: removing the artefacts of pre-processing. *Int J Climatology* 21:657–678
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatology* 25:1965–1978
- Huffman GJ, Adler RF, Rudolf B, Schneider U, Keen PR (1995) Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information. *J Climate* 8:1284–1295
- Huffman GJ, Adler RF, Bolvin DT, Gu G (2009) Improving the global precipitation record: GPCP Version 2.1. *Geophys Res Letters* 36: L17808, 5 pp
- Huffman GJ, Bolvin DT (2012) GPCP Version 2.2 combined precipitation data set documentation. Available online at: [ftp://precip.gsfc.nasa.gov/pub/gpcp-v2.2/doc/V2.2\\_doc.pdf](ftp://precip.gsfc.nasa.gov/pub/gpcp-v2.2/doc/V2.2_doc.pdf)



- Jacobs SS, Helmer HH, Doake CSM, Jenkins A, Frolich RM (1992) Melting of ice shelves and the mass balance of Antarctica. *J Glaciology* 38:375–387
- Jäger L (1976) Monatskarten des Niederschlags für die ganze Erde. *Berichte des Deutschen Wetterdienstes*, 139, Offenbach, 38 pp. and 13 appendices
- Legates DR (1987) A climatology of global precipitation. Publ. in *Climatology*, 40 (1), Newark, Delaware, 85 pp
- Legates DR (1995) Global and terrestrial precipitation: a comparative assessment of existing climatologies. *Int J Climatology* 15:237–258
- Legates DR, Willmott CJ (1990) Mean seasonal and spatial variability in gauge-corrected global precipitation. *Int J Climatology* 10:111–127
- Losev KS (1973) Estimation of run-off from Antarctic and Greenland ice sheets. Proc. Symposium on the Hydrology of glaciers, Cambridge, UK, 7–13 Sept. 1969, 253–254
- Matsuura K, Willmott CJ (2009) Terrestrial precipitation: 1900–2008 gridded monthly time series. Available online at: <http://climate.geog.udel.edu>
- Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG (2012) An overview of the Global Historical Climatology Network—daily database. *J Atmos Oceanic Technol* 29(7):897–910
- Meyer-Christoffer A, Becker A, Finger P, Rudolf B, Schneider U, Ziese M (2011a) GPCC Climatology Version 2011 at 0.25°: monthly land-surface precipitation climatology for every month and the total year from rain-gauges built on GTS-based and historic data. doi: [10.5676/DWD\\_GPCC/CLIM\\_M\\_V2011\\_025](https://doi.org/10.5676/DWD_GPCC/CLIM_M_V2011_025)
- Meyer-Christoffer A, Becker A, Finger P, Rudolf B, Schneider U, Ziese M (2011b) GPCC Climatology Version 2011 at 0.5°: monthly land-surface precipitation climatology for every month and the total year from rain-gauges built on GTS-based and historic data. doi: [10.5676/DWD\\_GPCC/CLIM\\_M\\_V2011\\_050](https://doi.org/10.5676/DWD_GPCC/CLIM_M_V2011_050)
- Meyer-Christoffer A, Becker A, Finger P, Rudolf B, Schneider U, Ziese M (2011c) GPCC Climatology Version 2011 at 1.0°: monthly land-surface precipitation climatology for every month and the total year from rain-gauges built on GTS-based and historic data. doi: [10.5676/DWD\\_GPCC/CLIM\\_M\\_V2011\\_100](https://doi.org/10.5676/DWD_GPCC/CLIM_M_V2011_100)
- Meyer-Christoffer A, Becker A, Finger P, Rudolf B, Schneider U, Ziese M (2011d) GPCC Climatology Version 2011 at 2.5°: monthly land-surface precipitation climatology for every month and the total year from rain-gauges built on GTS-based and historic data. doi: [10.5676/DWD\\_GPCC/CLIM\\_M\\_V2011\\_250](https://doi.org/10.5676/DWD_GPCC/CLIM_M_V2011_250)
- Mitchell TD, Carter TR, Jones PD, Hulme M, New M (2004) A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100). Tyndall Centre Working Paper No. 55. Available online at: [http://www.ipcc-data.org/docs/tyndall\\_working\\_papers\\_wp55.pdf](http://www.ipcc-data.org/docs/tyndall_working_papers_wp55.pdf)
- New M, Hulme M, Jones PD (1999) Representing twentieth century space–time climate variability. Part 1: development of a 1961–90 mean monthly terrestrial climatology. *J Climate* 12:829–856
- New M, Hulme M, Jones PD (2000) Representing twentieth century space–time climate variability. Part 2: development of 1901–96 monthly grids of terrestrial surface climate. *J Climate* 13:2217–2238
- New M, Todd M, Hulme M, Jones PD (2001) Precipitation measurements and trends in the twentieth century. *Int J Climatology* 21:1899–1922
- New M, Lister D, Hulme D, Makin I (2002) A high-resolution data set of surface climate over global land areas. *Clim Res* 21:1–25
- Nicholson SE (1986) The spatial coherence of African rainfall anomalies: interhemispheric teleconnections. *J Climate Appl Meteor* 25:1365–1381
- Nicholson SE (1993) An overview of African rainfall fluctuations of the last decade. *J Climate* 6:1463–1466
- Nicholson SE (2008) The intensity, location and structure of the tropical rainbelt over West Africa as a factor in interannual variability. *Int J Climatology* 28:1775–1785
- Peterson TC, Vose RS (1997) An overview of the Global Historical Climatology Network temperature database. *Bull Amer Met Soc* 78(12):2837–2849
- Peterson TC, Vose RS, Schmoyer R, Razuvaev V (1998) Global Historical Climatology Network (GHCN) quality control of monthly temperature data. *Int J Climatology* 18(11):1169–1179
- Rodell M, Houser PR, Jambor U, Gottschalck J, Mitchell K, Meng C-J, Arsenault K, Cosgrove B, Radakovich J, Bosilovich M, Entin JK, Walker JP, Lohmann D, Toll D (2004) The global land data assimilation system. *Bull Amer Meteor Soc* 85(3):381–394
- Rudolf B, Hauschild H, Rueth W, Schneider U (1994) Terrestrial precipitation analysis: operational method and required density of point measurements. In Desbois M, Desalmond F (eds) *Global precipitations and climate change*. NATO ASI Series I, 26, Springer, 173–186
- Rudolf B, Schneider U (2005) Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations. Proc. 2nd Workshop of the International Precipitation Working Group (IPWG), Monterey, 25–28 Oct. 2004, ISBN 92-9110-070-6, ISSN 1727-432X, 231–247
- Rudolf B, Becker A, Schneider U, Meyer-Christoffer A, Ziese M (2011) New full data reanalysis version 5 provides high-quality gridded monthly precipitation data. *GEWEX News* 21(2):4–5
- Schneider U, Hauschild H, Reiss M, Rudolf B (1992a) Erste Vergleiche der aus täglichen Vorhersagen des EZMW-Modells T106 und aus konventionellen Messdaten abgeleiteten monatlichen Gebietsniederschläge. *Meteorol. Zeitschrift N.F.* 1/1, 66–72 (English summary)
- Schneider U, Henning D, Hauschild H, Reiss M, Rudolf B (1992b) Zur Berechnung monatlicher Niederschlagshöhen aus synoptischen Meldungen. *Meteorol. Zeitschrift N.F.* 1/1, 22–31 (English summary)
- Schneider U, Becker A, Meyer-Christoffer A, Ziese M, Rudolf B (2011a) Global precipitation analysis products of the GPCC. Global Precipitation Climatology Centre (GPCC), DWD, 12 pp. Available online at [ftp://ftp-anon.dwd.de/pub/data/gpcc/PDF/GPCC\\_intro\\_products\\_2008.pdf](ftp://ftp-anon.dwd.de/pub/data/gpcc/PDF/GPCC_intro_products_2008.pdf)
- Schneider U, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Ziese M (2011b) GPCC Monitoring Product Version 4.0 at 1.0°: near real-time monthly land-surface precipitation from rain-gauges based on SYNOP and CLIMAT Data. doi: [10.5676/DWD\\_GPCC/MP\\_M\\_V4\\_100](https://doi.org/10.5676/DWD_GPCC/MP_M_V4_100)
- Schneider U, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Ziese M (2011c) GPCC Monitoring Product Version 4.0 at 2.5°: near real-time monthly land-surface precipitation from rain-gauges based on SYNOP and CLIMAT Data. doi: [10.5676/DWD\\_GPCC/MP\\_M\\_V4\\_250](https://doi.org/10.5676/DWD_GPCC/MP_M_V4_250)
- Schneider U, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Ziese M (2011d) GPCC Full Data Reanalysis Version 6.0 at 0.5°: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. doi: [10.5676/DWD\\_GPCC/FD\\_M\\_V6\\_050](https://doi.org/10.5676/DWD_GPCC/FD_M_V6_050)
- Schneider U, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Ziese M (2011e) GPCC Full Data Reanalysis Version 6.0 at 1.0°: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. doi: [10.5676/DWD\\_GPCC/FD\\_M\\_V6\\_100](https://doi.org/10.5676/DWD_GPCC/FD_M_V6_100)
- Schneider U, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Ziese M (2011f) GPCC Full data reanalysis version 6.0 at 2.5°: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. doi: [10.5676/DWD\\_GPCC/FD\\_M\\_V6\\_250](https://doi.org/10.5676/DWD_GPCC/FD_M_V6_250)
- Schneider U, Umeda T (2012) Global climate observing system surface network monitoring report for 2011, no. 16. April 2012. Available online at: <http://www.gsnmc.dwd.de>

- Schulzweida U, Kornblueh L, Quast R (2011) Climate data operators (CDO), user guide, Version 1.5.2. Available online at: <https://code.zmaw.de/files/cdo/html/1.5.2/cdo.pdf>
- Sevruk B (1982) Methods of correction for systematic error in point precipitation measurements for operational use. Operational Hydrology, Report No. 21, WMO, Geneva, WMO No. 589, 91 pp
- Sevruk B (1985) Correction of monthly precipitation for wetting losses. Instruments and observing methods, Report No. 22, WMO, Geneva, 7-12
- Sevruk B, Klemm S (1989) Catalogue of national standard precipitation gauges. Instruments and observing methods, Report No. 39, WMO/TD-No. 313, 50 pp
- Simmons A (2011) From observations to service delivery: challenges and opportunities. WMO Bulletin 60 (2) Reaching users with climate services, WMO, Geneva, Nov. 2011, 517-524
- Strangeways I (2007) Precipitation: Theory, measurement and distribution. Cambridge University Press, Cambridge, 290 pp
- Thomas AR, Patterson VL (1983) A reliable method for estimating precipitation amounts. Preprint 5th Symposium Meteorological Observations and Instrumentation. Toronto, AMS, 554-561
- Trenberth KE, Smith L, Qian T, Dai A, Fasullo J (2007) Estimates of the global water budget and its annual cycle using observational and model data. *J Hydrometeor* 8:758-769
- Trenberth KE, Fasullo J, Kiehl J (2009) Earth's global energy budget. *Bull Amer Meteor Soc* 90:311-323
- Trenberth KE, Fasullo J, Mackaro J (2011) Atmospheric moisture transports from ocean to land and global energy flows in Reanalyses. *J Climate* 24:4907-4924
- Trenberth KE, Fasullo J (2012) Atmospheric moisture transports from ocean to land in Reanalyses. *GEWEX News* 22(1):8-10
- Uppala SM et al (2005) The ERA-40 reanalysis. *Q J Royal Met Society* 131:2961-3012
- Vaughan DG, Bamber JL, Giovinetto M, Russell J, Cooper APR (1999) Reassessment of net surface mass balance in Antarctica. *J Climate* 12:933-946
- Willmott CJ, Rowe CM, Philpot WD (1985) Small-scale climate maps: a sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *Amer Cartogr* 12:5-16
- 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 App
- WMO (1996) Climatological normals (CLINO) for the period 1961-1990. WMO-No. 847, WMO, Geneva
- WMO and UNESCO (1970) Climatic atlas of Europe. Vol. 1: maps of mean temperature and precipitation
- WMO and UNESCO (1975) Climatic atlas of South America. Vol. 1: maps of mean temperature and precipitation
- WMO and UNESCO (1979) Climatic atlas of North and Central America. Volume 1: maps of mean temperature and precipitation
- WMO and UNESCO (1981) Climatic atlas of Asia. Volume 1: maps of mean temperature and precipitation
- Xie P, Arkin PA (1997) Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. *Bull Amer Meteor Soc* 78:2539-2558
- Ziese M, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schneider U (2011) GPCC First Guess Product at 1.0°: Near real-time first guess monthly land-surface precipitation from rain-gauges based on SYNOP data. doi:10.5676/DWD\_GPCC/FG\_M\_100