



Global gridded precipitation over land: a description of the new GPCC First Guess Daily product

K. Schamm^{1,2}, M. Ziese¹, A. Becker¹, P. Finger¹, A. Meyer-Christoffer¹, U. Schneider¹, M. Schröder², and P. Stender¹

¹Deutscher Wetterdienst, Global Precipitation Climatology Centre, 63067 Offenbach/M., Germany

²Deutscher Wetterdienst, Satellite-based Climate Monitoring, 63067 Offenbach/M., Germany

Correspondence to: M. Ziese (markus.ziese@dwd.de)

Received: 4 July 2013 – Published in Earth Syst. Sci. Data Discuss.: 9 August 2013

Revised: 3 January 2014 – Accepted: 11 January 2014 – Published: 27 January 2014

Abstract. This paper describes the new First Guess Daily product of the Global Precipitation Climatology Centre (GPCC). The new product gives an estimate of the global daily precipitation gridded at a spatial resolution of 1° latitude by 1° longitude. It is based on rain gauge data reported in near-real time via the Global Telecommunication System (GTS) and available about three to five days after the end of each observation month. In addition to the gridded daily precipitation totals in mm day^{-1} , the standard deviation in mm day^{-1} , the kriging interpolation error in % and the number of measurements per grid cell are also encoded into the monthly netCDF product file and provided for all months since January 2009. Prior to their interpolation, the measured precipitation values undergo a preliminary automatic quality control. For the calculation of the areal mean of the grid, anomalies are interpolated with ordinary block kriging. This approach allows for a near-real-time release. Therefore, the purely GTS-based data processing lacks an intensive quality control as well as a high data density and is denoted as First Guess. The daily data set is referenced under doi:10.5676/DWD_GPCC/FG_D_100. Two further products, the Full Data Daily and a merged satellite-gauge product, are currently under development at Deutscher Wetterdienst (DWD). These additional products will not be available in near-real time, but based on significantly more and strictly quality controlled observations. All GPCC products are provided free of charge via the GPCC webpage: ftp://ftp-anon.dwd.de/pub/data/gpcc/html/download_gate.html.

1 Introduction

The Global Precipitation Climatology Centre (GPCC) was founded in 1989 and is operated by Deutscher Wetterdienst (DWD) in support of the World Meteorological Organization (WMO). It is a contribution to the World Climate Research Programme (WCRP) and to the Global Climate Observing System (GCOS). The objective of GPCC is to provide high-quality global precipitation analyses over land for monitoring and research of the earth's climate. Rain gauge data from different sources have been collected and the database, which is updated regularly, has increased to one of the largest of its kind. Great importance is thereby attached to the quality of the data, which undergoes different quality checks before use. Different gridded monthly precipitation products based

on this data are released on a regular basis to address the different needs of quality and time of availability. A detailed description of all monthly precipitation products and GPCC's quality control procedures is given in Becker et al. (2013) and Schneider et al. (2013), respectively. However, for many applications such as the statistics of extremes, the monitoring of heavy precipitation events and the evaluation of numerical weather predictions or satellite estimates a time resolution higher than one month is required.

There are some other gauge-based gridded precipitation products with a temporal resolution higher than the monthly product. For example, the ENSEMBLES observations gridded data set (E-OBS), which is based on observational station data of the European Climate Assessment & Dataset

Table 1. Different available data sets for gridded daily precipitation.

Name of data set	Region	Spatial resolution	Time period	Reference
APHRODITE	Asia	0.25°	1961–2007	Yatagai et al. (2009)
E-OBS	Europe	0.25°	since 1950	Haylock et al. (2008)
SPAIN2	Spain	0.22°		Herrera et al. (2012)
	Alps	25 km	1971–1990	Frei and Schär (1998)
	Alps	5 km	1971–2008	Isotta et al. (2013)
	Great Britain	5 km	1960–2006	Perry and Elms (2009)
Baltex	Area around Baltic Sea	0.2°	Sep 1999–Dec 2000	Rubel and Hantel (2001)
RhiresD	Switzerland	5 km	since 1961	Schiemann et al. (2010)
CPC	Global	0.5°	since 1979	NOAA (2013)
	Australia	0.05°	since 1970	Jones et al. (2009)

(ECA&D) (Haylock et al., 2008), provides daily precipitation data for Europe. Similarly the Asian Precipitation – Highly Resolved Observational Data Integration Towards Evaluation of Water Resources project (APHRODITE’s Water Resources) developed a product for daily precipitation over Asia (Yatagai et al., 2009). Table 1 gives an overview of different daily gridded precipitation products based on rain gauge data. However, most of these data sets cover either a limited area or a short time period.

Apart from these products based on in situ measurements, there are numerous precipitation products derived from satellite data, radar data, model reanalyses and other combined products. Satellite data bear the advantage of a very good spatial coverage compared to rain gauges that are restricted to the land surfaces of the globe and can be rather sparse especially in developing countries and sparsely settled regions. However, their uncertainty is particularly high due to the indirect precipitation estimation method (Petty and Krajewski, 1996; Kidd et al., 2011). Moreover, the relatively short existence of satellites yields only few time series exceeding 30 yr.

Radar estimates have a very high spatial and temporal resolution but are not available for large parts of the globe. Moreover, their time series are even shorter than the satellite-based ones.

Gridded products of rain gauge measurements also have their disadvantages. The reliability of the analysis depends on the station density (Arkin et al., 1985; Rudolf et al., 1994; Schneider et al., 1993; see also Sect. 4.4). As the station density decreases, the uncertainty of the calculated area mean increases. Also individual large precipitation totals at a station are smoothed. Depending on the station density and grid cell size, grid cells including no stations are likely. In this case precipitation totals outside the grid cell are applied to calculate the precipitation amount in this cell. Some interpolation schemes use a minimum number of stations as input data (Shepard, 1968; Krige, 1966). Therefore, stations from other grid cells influence the computed precipitation total at a cell, whenever less than the minimum number of stations are found within this cell. Furthermore, orographic characteris-

tics are smoothed if the grid cell size is not small enough or there are no measurements within it. The users have to keep in mind that a high-resolution gridded data set does include small-scale characteristics of precipitation fields only if the precipitation field is sufficiently sampled by rain gauges.

A combination of station- and satellite-measurements (for example the Global Precipitation Climatology Project (GPCP), Huffman et al., 2001) allows an enhancement of the product quality, although some disadvantages still remain (e.g. adjustment of satellite data in sparsely or unprobed regions). Furthermore, it is difficult to give quantitative error estimates for a merged product. Intercomparisons of different products identify biases (Yilmaz et al., 2005) but do not give guidance to prefer a certain product.

With its new First Guess Daily product, the GPCC addresses the need of a global precipitation product with a daily resolution as well as information on the uncertainty. An update is provided three to five days after the end of each month. The spatial resolution is 1° latitude by 1° longitude and covers the entire land surface except the Antarctic. This analysis is available back to January 2009. As the name implies, it is a quick analysis of the daily precipitation amounts that is designed to be released in near-real time for applications being up-to-date. The advantage of this approach is the high timeliness, but this leads to a lower number of input stations and a lower level of quality control. Therefore, data from SYNOP messages, which are reported via WMO Global Telecommunication System (GTS), are used after an automatic-only quality control (see Sect. 3).

Generally, quick analyses open the opportunity for monitoring processes like drought monitoring or to calibrate retrievals of indirect measurements like satellite data (e.g. GPCP, Huffman et al., 2001). Such first guess analyses of station-based or gridded data sets have a lack of input data as well as the time available for quality control activities. Gridded data sets can provide areal mean values unlike station based measurements. These gridded data sets are necessary to calculate for instance precipitation sums of larger areas like catchment basins. Due to the interpolation technique,

local extreme values are smoothed; also, precipitation totals can be interpolated in unprobed regions. In contrast extremes can be analyzed or statistics at a station can be conducted with station data. But station data are not necessarily representative for larger areas (for example stations at a mountain peak would have other values than a station some kilometres away in a valley). Due to the data policy of the GPCC, no raw station data or metadata can be provided (Becker et al., 2013). The GPCC does not own the copyright of the provided station data and therefore cannot give this data to third parties.

First, we give a description of a comparison of different interpolation schemes to find the best performing one for daily data (Sect. 2). An overview of the input data and quality control is given in Sect. 3. The calculation procedure for the First Guess Daily is described in Sect. 4 and the data access is described in Sect. 5. An outlook to scheduled future daily precipitation analyses is given in Sect. 6. Finally, we summarize our findings.

2 Comparison of different interpolation methods

Several methods for gridding of daily precipitation data have been applied and compared to find the best performing scheme to interpolate daily precipitation amounts:

- a modified SPHEREMAP interpolation as it is used for the monthly products (to include all stations in areas with a very high station density; detailed information in Sect. 4 of Becker et al., 2013);
- ordinary point kriging with two different correlogramm settings (one correlogramm for the entire globe, and a combination of local correlogramms per grid cell if enough data is available, otherwise climate zone dependent correlograms were used);
- an inverse distance weighting interpolation (IDW); and
- the mean of the nearest precipitation measurements (mean).

For all methods a minimum of 4 and a maximum of 10 neighbouring stations were used. The search radius is chosen depending on the mean station density (Shepard, 1968) and increases for each grid cell independently with decreasing station density. Cross-validation was used for quality evaluation.

For the interpolation two approaches were applied: interpolation of totals and interpolation of anomalies. At the interpolation of totals the daily precipitation amounts as measured were interpolated. Contrary, at the interpolation of anomalies the precipitation anomaly or fraction was calculated as the daily precipitation total to the monthly precipitation total at this station. After interpolation of the anomalies these were multiplied with the measured monthly total at this station to

get the daily total. Such a procedure is also known as Climatology Aided Interpolation (CAI) (Willmott and Robeson, 1995).

To cross-validate the different interpolation schemes, the precipitation total of one station was left out and the precipitation at this station was interpolated from the surrounding stations. By comparing the real measured and interpolated precipitation at this station, different error estimates were calculated: mean absolute error (MAE), mean squared error (MSE, given in Table 2) and precipitation observed and precipitation interpolated at the station. This was done for all available stations for every day of the year 2008 (largest number of available stations).

Kriging (Krige, 1981; Gandin, 1993) is a stochastic approach for the interpolation of data. The interpolation theory has been developed in the geosciences by the French mathematician Georges Matheron and named in honour of the South African mining engineer Daniel Krige. It is now a popular interpolation method and used for many applications in geosciences. As for almost all interpolation algorithms, kriging uses weights depending on the distance between measurement location and the location of interest. The weights decrease with increasing distance. The particular approach of kriging is to calculate weights according to a (moderately) data-driven function, rather than an arbitrary function like SPHEREMAP (Shepard, 1968; Willmott et al., 1985). The spatial correlation structure of the data is used to determine the weighting function.

In order to find the optimal interpolation setting and weighting function for kriging, different autocorrelation parameters have been applied and the results have been compared. At first the same autocorrelation parameters have been used for each grid irrespective of the region or station density (parameters summarized in Table 3). The second approach was to calculate individual autocorrelation parameters for each grid point whenever enough data was available. If the data coverage does not support grid point specific diagnosis of the autocorrelation parameters, the lookup table for autocorrelation parameters (Table 1 in Kottek and Rubel, 2007), depending on the Köppen–Geiger climate classification (Kottek et al., 2006), is applied instead (see Fig. 1 for spatial distribution of climate zones). To our knowledge no variogram parameters are available for other climate classifications like Flohn and Neff (Flohn, 1954), Troll and Pfaffen (Troll and Pfaffen, 1964), Lauer, Frankenberg and Rafiqpoor (Lauer and Bendix, 2004), Holdridge (Holdridge, 1947), Schreiber (Schreiber, 1973), Budyko and Grigor'ev (Giese, 1968) or Terjung and Louie (Terjung and Louie, 1972). Therefore we applied Köppen–Geiger to test the performance of kriging with local or regional variograms. It appears that the first approach with only one set of parameters for all grid points shows the best results. We assume that the reason is that there are not enough measurements to calculate reliable autocorrelation parameters in most regions because precipitation is too variable. Even if not only stations within

Table 2. Calculated MSE for different interpolation methods depending on the climatic zone (see Fig. 1, Kottek et al., 2006) for one year (2008). The lowest MSE values are marked bold.

MSE per: climatic zone	Kriging absolute	Kriging anomaly	Spheremap absolute	Spheremap anomaly	IDW absolute	IDW anomaly	Mean absolute	Mean anomaly
Af	16.8	15.9	18.1	17.4	17.0	16.1	17.2	16.5
Am	15.9	14.4	17.0	15.7	16.2	14.8	16.5	15.3
Aw	10.4	9.9	11.2	10.8	10.5	10.0	10.8	10.4
As	9.6	9.2	10.3	9.9	9.7	9.4	9.8	9.5
BWh	2.5	2.4	2.7	2.9	2.5	2.6	2.6	3.4
BWk	1.7	1.4	1.8	1.5	1.8	1.4	1.9	1.5
BSh	5.4	12.1	5.8	16.3	5.5	14.9	5.7	13.2
BSk	2.6	2.3	2.8	2.4	2.7	2.3	2.8	2.4
Csa	4.3	3.9	4.5	4.0	4.4	3.9	4.7	4.2
Csb	4.0	3.7	4.2	3.9	4.1	3.8	4.3	4.0
Csc	2.1	0.1	1.7	0.1	2.7	0.1	3.6	0.1
Cwa	9.8	8.9	10.4	9.4	9.9	9.0	10.1	9.4
Cwb	6.2	5.8	6.5	6.5	6.2	5.9	6.4	6.0
Cwc	4.3	3.2	4.5	3.2	5.1	3.3	5.1	3.4
Cfa	7.0	6.4	7.4	6.7	7.2	6.6	7.5	6.9
Cfb	2.8	2.5	2.9	2.6	2.9	2.6	3.0	2.7
Cfc	5.0	3.9	5.3	4.2	5.1	4.2	5.2	4.5
Dsa	2.1	1.7	2.1	1.7	2.2	1.7	2.4	1.9
Dsb	2.7	2.1	3.0	2.2	2.9	2.1	3.3	2.4
Dsc	4.2	2.7	4.4	3.2	4.4	2.8	4.6	2.9
Dwa	4.9	4.5	5.0	4.5	5.0	4.6	5.3	4.9
Dwb	3.5	3.2	3.7	3.4	3.5	3.3	3.7	3.4
Dwc	2.7	2.4	3.4	2.6	3.0	2.5	3.1	2.6
Dwd	3.6	3.2	3.9	4.4	3.3	3.1	3.3	3.0
Dfa	4.9	4.2	5.0	4.3	5.0	4.4	5.1	4.6
Dfb	3.5	3.2	3.8	3.4	3.6	3.3	3.7	3.4
Dfc	3.5	3.2	3.7	3.5	3.6	3.3	3.7	3.4
Dfd	2.3	2.1	2.6	2.2	2.3	2.1	2.4	2.2
ET	3.9	3.5	4.3	4.1	3.9	3.7	4.0	3.8
EF	3.5	3.5	3.7	4.0	3.5	3.5	3.5	3.5

Table 3. Parameters used for block kriging to calculate First Guess Daily and point kriging to compare the performance of different interpolation schemes.

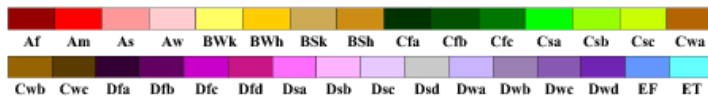
Parameter	Size
Autocorrelation function	$R = a * \exp(-b * d^c)$
d	distance grid point to station [km]
R	autocorrelation
a	0.95
b	0.014
c	0.726
Decorrelation length	347 km
Block size	1° by 1°

one grid point but also all stations within a climatic area and for a longer time period are considered, the spatial structure is not sufficiently represented and the calculated parameters vary strongly. These results have also been found for other data sets (e.g. Hofstra et al., 2008).

The results of the comparison of these methods show that the performance does not differ significantly if they are compared in the same climate zone (see Table 2). Due to the large different precipitation totals, size of climate zones and number of stations between the climate zones, the MSE differ strongly, because it is not normalized to take this into account. Depending on the region, data density, time period or amount of precipitation different interpolation schemes perform better but all in all their results are similar. Even the choice of the interpolation with anomalies or absolute values does not influence the results in many regions strongly and makes little difference when comparing the globally calculated mean squared error (MSE, Table 2). Nevertheless, ordinary point kriging shows slightly better results than the other interpolation methods and the MSE for the interpolation with anomalies is slightly lower than for absolute values in nearly all cases. Especially in regions with a very sparse data density and strongly varying climate zones such as central Africa, South America and parts of Southeast Asia, the anomaly interpolation partly makes up for the missing

World Map of Köppen–Geiger Climate Classification

updated with CRU TS 2.1 temperature and VASCLimO v1.1 precipitation data 1951 to 2000



Main climates

- A: equatorial
- B: arid
- C: warm temperate
- D: snow
- E: polar

Precipitation

- W: desert
- S: steppe
- f: fully humid
- s: summer dry
- w: winter dry
- m: monsoonal

Temperature

- h: hot arid
- k: cold arid
- a: hot summer
- b: warm summer
- c: cool summer
- d: extremely continental
- F: polar frost
- T: polar tundra

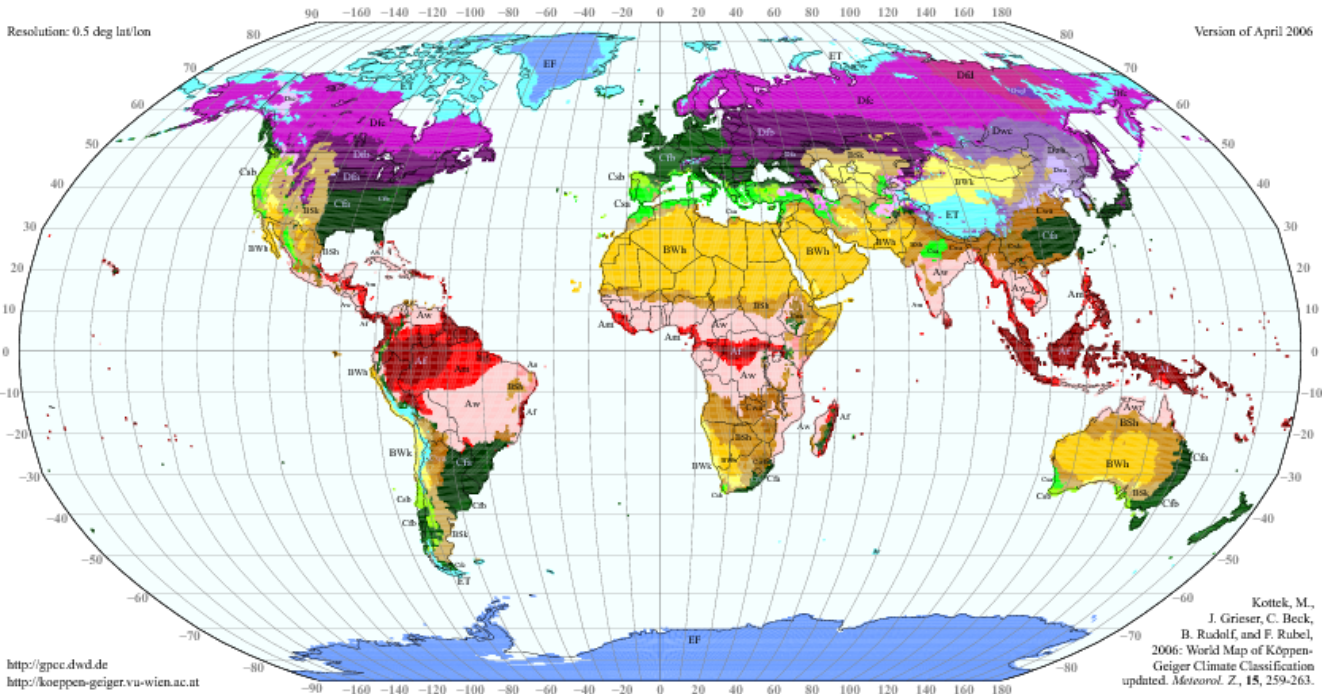


Figure 1. World Map of Köppen–Geiger climate classification (Kottek et al., 2006) used for the analysis of interpolation methods.

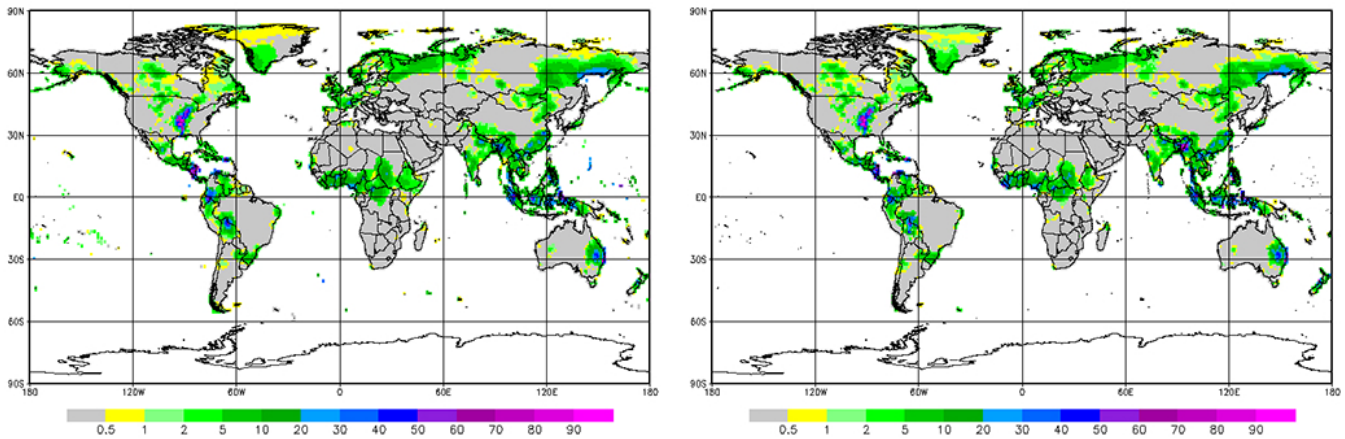


Figure 2. Example for daily precipitation [mm day^{-1}] interpolated with absolute values (left) and anomalies (right) for one day. The differences are mainly noticeable in central Africa and Southeast Asia.

values (Figs. 2 and 3). Hence for the daily precipitation products ordinary block kriging with anomalies was adapted (see Sect. 4.3). Additionally, it has the advantage that error estimation is provided implicitly. These results are supported by results which have been found for other regions and data sets

(Dubrule, 1984; Tabios and Salas, 1985; Dirks et al., 1998; Hofstra et al., 2008; Herrera et al., 2012).

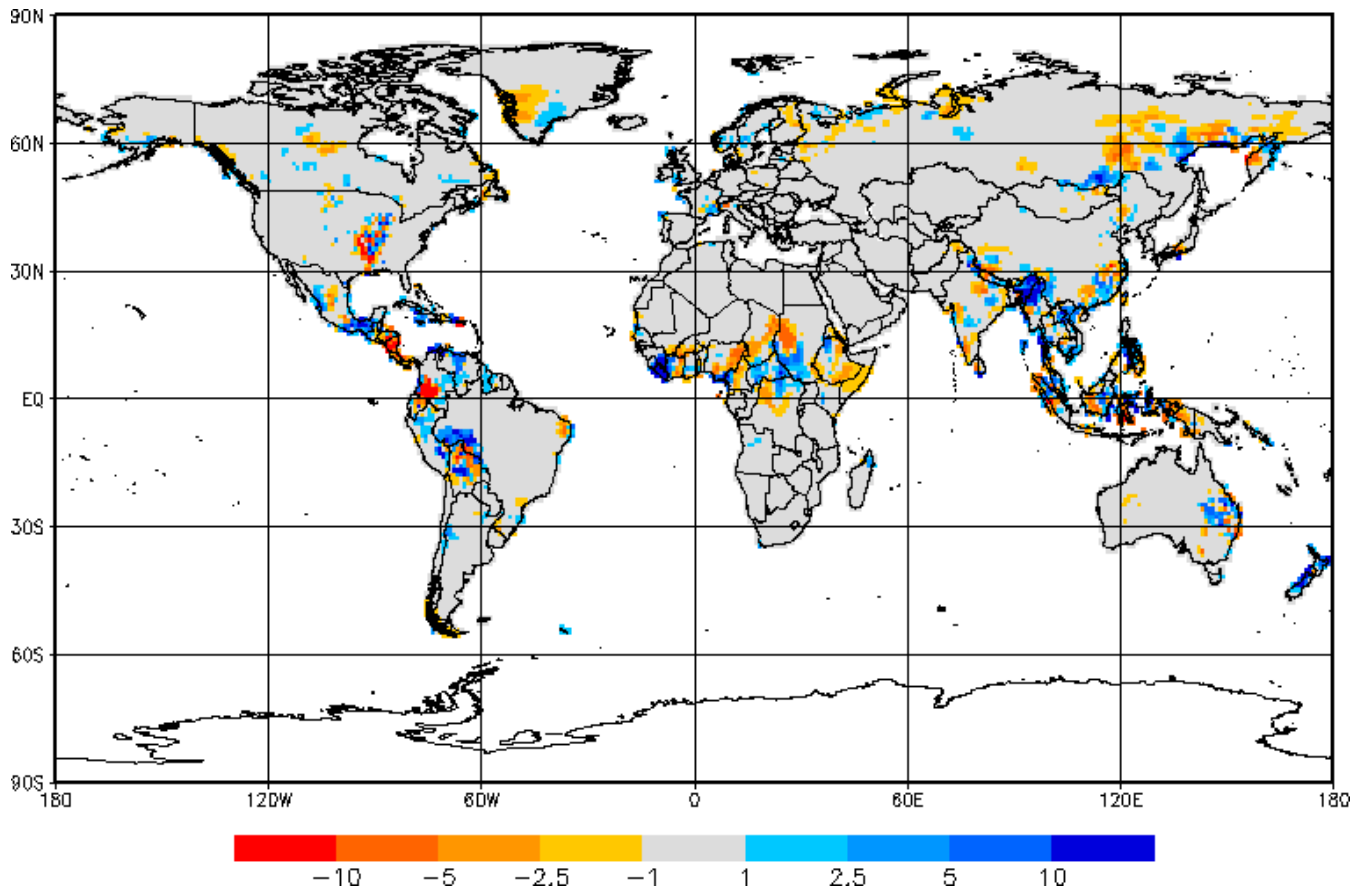


Figure 3. Difference for daily precipitation [mm day^{-1}] interpolated with anomalies minus absolute values for the same day as in Fig. 2.

3 Data source and quality control

The Global Precipitation Climatology Centre (GPCC) database consists of a large amount of daily and monthly data from several different data sources such as national meteorological and hydrological services, global and regional data collections (e.g. Food and Agriculture Organization of the United Nations (FAO), Global Historical Climatology Network (GHCN), Climatic Research Unit (CRU), European Climate Assessment & Dataset (ECA&D)) and near-real-time data from WMO-GTS (SYNOP reports, CLIMAT messages for monthly data). Monthly data have been collected since the foundation of GPCC in 1989, whereas the collection of daily data has only been started recently. All GPCC products for monthly totals are based on the monthly data including a climatology. A detailed description of the complete monthly data, the database and the monthly GPCC products is given in Becker et al. (2013) and Schneider et al. (2013).

3.1 Data source First Guess Daily

As the First Guess Daily product is a near-real-time GTS based product, only data from SYNOP reports can be used. The WMO-Vol. A (WMO, 2013) lists all currently operat-

ing SYNOP stations. Most of them report precipitation and these are utilized for the First Guess Daily. Currently there are roughly 6000 to nearly 8000 stations with daily SYNOP data each month (see Fig. 4, light blue line) for the processed period. These stations are spread around the globe but unfortunately there are regions with higher (e.g. Europe) and lower (e.g. Central Africa, Northern Canada) station density. Especially the coverage across many African, Asian and South American regions is weak in both number and availability (see Fig. 5, top right).

For each day the reported observations of the precipitation amounts for the specific time intervals (1 h, 3 h, 6 h, 12 h, 18 h or 24 h) are added up to a daily sum for the climatological day for each station.

3.2 Quality control First Guess Daily

Within the processing routine for SYNOP reports, quality checks are made including a plausibility check for consistency for overlapping periods in time, and for obvious coding errors. Also format errors were detected and corrected, for example when the weather group is shifted to the precipitation group. Additionally, the reported weather groups are used to evaluate the plausibility of the reported precipitation

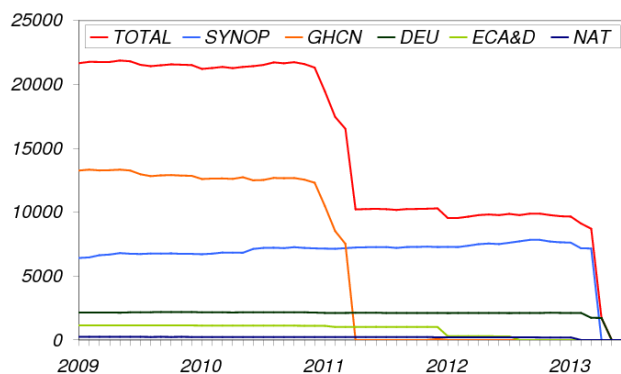


Figure 4. Number of daily precipitation data in the GPCC database as a function of time for different data sources. Data from SYNOP reports via the GTS are marked with a light blue line, data from the Global Historical Climatology Network (GHCN) are marked with an orange line. The darker green line shows the number of German precipitation data (DEU), the lighter green line shows the number of ECA&D data and the darker blue line shows the number of data delivered from national weather services and hydrological services (NAT). The total number of data is printed with the red line (April 2013).

amount as well as to complete missing precipitation values when the weather group indicates no precipitation.

Wrong precipitation codings were corrected. For example, if the weather group indicates light rain and in the precipitation group is RRR = 919, then RRR is changed to 991. Uncoded, this means the precipitation total was changed from 919 to 0.1 mm.

As precipitation measurements can overlap in time, the precipitation total of the shorter report have to be equal to or lower than the longer report. For example 12 h report with 20 mm can include a 6 h report with 20 mm or less.

Also the combination of precipitation and weather group, if available, is checked. If the weather group indicated no rain and the precipitation group has rain, then the precipitation is left out.

Before interpolation the daily totals calculated from the SYNOP data are loaded into the databank and the station metadata are checked (location and confusion with other stations).

4 First Guess Daily product – calculation of the gridded data set

For the First Guess Daily product the reported SYNOP precipitation measurements are calculated as areal means of regular grid cells with a spatial resolution of 1° latitude by 1° longitude. In addition to the total precipitation values in mm day^{-1} two different error estimations and the number of measurements per grid are provided.

4.1 Background Climatology

All monthly products from the Global Precipitation Climatology Centre (GPCC) were interpolated as anomalies from long-term means, called Climatology. Briefly, this Climatology focuses on the period 1951 to 2000 but uses also other periods, if the mentioned one is not available. A detailed description of the Climatology is given in Schneider et al. (2013).

4.2 Interpolation of monthly data

The GPCC First Guess product (Ziese et al., 2011) is used for the monthly gridded data. This product is based on SYNOP reports only, which undergo the same quality control as for the First Guess Daily (see Sect. 3.2), and is available three to five days after the end of each month. Monthly totals were calculated if the station has at least 70 % data coverage. Depending on the amount of missing data the monthly total is extrapolated. For the calculation of the gridded monthly totals the anomalies regarding long-term means at the station are interpolated to a regular grid using a modified SPHEREMAP scheme. The monthly totals then are calculated as the sum of the interpolated anomalies and the climatology for each grid point. This is also known as Climatology Aided Interpolation (CAI) (Willmott and Robeson, 1995). A detailed description of the First Guess product and the interpolation method is given in Becker et al. (2013).

4.3 Interpolation of daily data

For the First Guess Daily product the irregularly spaced observations of precipitation are interpolated to a regular grid using ordinary block kriging. Parameters used in this data set are summarized in Table 3. The block size of the applied kriging scheme is 1° . The advantage of block kriging in comparison to point kriging is that it calculates implicitly the areal precipitation of the block size. Therefore the precipitation is calculated at several points within the block surrounding the grid point and then a weighted sum of these totals is calculated. This is the same like interpolation with point kriging at several locations around the grid point and to calculate a weighted mean of them to get the areal precipitation.

For each grid box that contains at least a fraction of land the nearest neighbouring measurements are chosen. Precipitation values of stations located within a distance of less than 1 km are averaged so that the impact of these measurements will not be overestimated and no station is included twice by mistake. Depending on their distance to the grid point and the chosen autocorrelation function, interpolation weights are calculated and the estimation of the precipitation within the grid is computed. This is also done for areas with sparse station density by increasing the search radius for neighbouring stations (Shepard, 1968). In those regions the search radius can be much larger than the correlation length

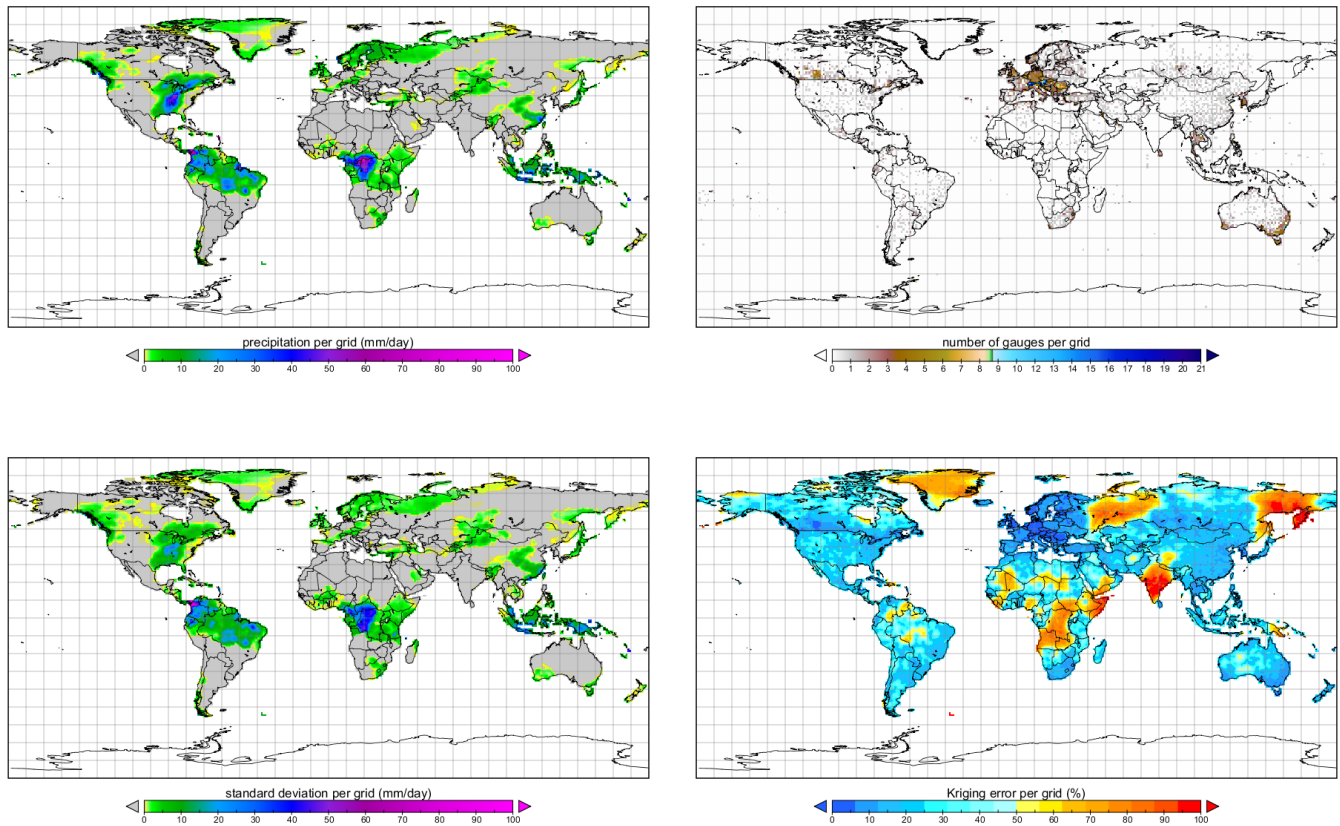


Figure 5. Top left: estimated amount of precipitation in [mm day^{-1}]; top right: example for the distribution of SYNOP-Measurements (18 April 2013) as they are used for the First Guess Daily product. (Colours indicate the number of measurements per grid with a 1° latitude by 1° longitude resolution); bottom left: standard deviation according to the method of Yamamoto (2000) [mm day^{-1}]; bottom right: kriging error in [%].

of precipitation (around 347 km, see also Table 3) and therefore the uncertainty of the calculated precipitation can be as high as the calculated precipitation itself. But at the same time this approach allows a complete coverage of the precipitation estimation over the land surface. No information is provided for grid boxes that do not contain land surface. A detailed description of the theory of ordinary block kriging and the implementation of kriging in Fortran is given in Rubel (1996).

The processing of the First Guess Daily product is based on an anomaly interpolation method. Therefore, the daily precipitation anomalies are calculated as the daily total divided by the monthly total at this station. Stations without monthly totals (monthly data coverage less than 70 %) are excluded from the analysis. These ratios of the daily anomalies are interpolated to the grid and the results are multiplied with the gridded monthly totals from the First Guess in order to gain absolute daily precipitation values. Such a procedure is also known as Climatology Aided Interpolation (CAI) (Willmott and Robeson, 1995). Due to the application of the First Guess as background for the First Guess Daily, the climatology is also included indirectly in the First Guess Daily.

Therefore results in data sparse regions are more reliable than in a simple interpolation of daily totals.

The disadvantage of the approach of interpolation of anomalies is that the daily product can only be calculated upon completion of the monthly product, which is about three to five days after the end of each observation month. Therefore, a future near-real-time global availability (in the order of one day) of the Global Telecommunication System (GTS) based daily precipitation analysis would require a much more homogeneous global coverage of GTS stations.

4.4 Uncertainty information

Depending on the station density and the structure and amount of precipitation in a region the uncertainties of the gridded product can be very high. To inform the user about the quality of the gridded data, information of the uncertainties and the number of stations per grid cell are provided additionally with every product file. Information on the uncertainty is provided in two different parameters.

1. The kriging uncertainty (Fig. 5, bottom right) depends mostly on the interstation distance (the

autocorrelogram), the size of the grid and the number of measurements. It is a result of the interpolation equations of kriging and can be interpreted as the percentage of the variance (Rubel, 1996) and is not an absolute error estimation. Furthermore the amount of precipitation does not influence the kriging error. Users should access this information to identify problematic areas.

2. An absolute value of the uncertainty is given with the standard deviation (Fig. 5, bottom left), calculated according to Yamamoto (2000). It depends on the measured precipitation values. The difference of the values of the surrounding measurements and the interpolated value are calculated and weighted with the same weights as applied for the interpolation. Furthermore, the error estimation for regions where no precipitation is interpolated or measured is always zero in contrast to the kriging error. Users who require a quantified error estimate should use this information.

Further possibilities of calculating an uncertainty estimation have been implemented and validated with cross-validation (the regular standard deviation and the combination of the Yamamoto standard deviation and the kriging error). However, due to the different concepts of kriging error and the Yamamoto standard deviation, we have decided to show them both with the forementioned user advice instead of a blended parameter.

Finally, a third uncertainty information is provided with the First Guess Daily product in terms of the number of measurements per grid cell (Fig. 5, top right). This number is the real number of measurements per grid and must not be mistaken for the number of measurements used for interpolation to the specific grid. It allows the user to get a rough impression of the distribution of the measurements and therefore of the quality of the product at a certain location.

5 How to access the products

All gridded GPCC products are available via the public GPCC webpage free of charge: ftp://ftp-anon.dwd.de/pub/data/gpcc/html/download_gate.html. No registration is required to download the data.

Daily data is provided as a single netCDF file for each month containing the gridded data (total precipitation, number of measurements and both uncertainty estimations) for each day of the month. First Guess Daily is provided at a regular global grid (-180 to 180° longitude and -90 to 90° latitude) with 1° latitude by longitude grid size. This grid has no projection! One netCDF-file has about 3 to 3.5 MB. We decided to provide the First Guess Daily as netCDF-files because many publicly available tools and libraries are provided to analyse netCDF-files.

netCDF is a standardized, self describing binary file format (netCDF, 2013). The coordinates of each grid

Table 4. Properties of the used grid.

Property	Value
gridtype	lonlat
gridsize	64 800
xname	lon
xlongname	longitude
xunits	degrees east
yname	lat
ylongname	latitude
yunits	degrees north
xsize	360
ysize	180
xfirst	-179.50
xinc	1.0
yfirst	-89.50
yinc	1.0

cell and missing values are coded in the header of each netCDF-file and can be read by suitable software. We give the coordinates of the centre of each grid cell. The missing value is -99999.99 . For the visualization of the values several visualizing software like Panoply (<http://www.giss.nasa.gov/tools/panoply/>, last access: 11 December 2013) or GrADS (<http://www.iges.org/grads/>, last access: 11 December 2013) exist. Some software to analyse and convert netCDF-files is listed at the First Guess Daily website ftp://ftp.dwd.de/pub/data/gpcc/html/gpcc_firstguess_daily_doi_download.html. Table 4 summarizes the applied grid and Table 5 the variables in the netCDF-file. The release of the DOI references implies that ISO 19115 compliant metadata are provided under web addresses constructed from the DOI proceeded by <http://data.datacite.org>. Additional, this product would also be included in the data set catalogue of the Climate Data Centre (CDC) of Deutscher Wetterdienst (DWD), which disseminates ISO 19139 compliant metadata on its data sets through the GeoNetwork software application.

6 Outlook: scheduled daily precipitation analyses

Data from sources other than GTS take longer to arrive at GPCC. Also, their quality control at GPCC is more extensive than for the SYNOP data. For daily data the quality control comprises, in addition to the metadata, a check of consistency with monthly data, a check of the assignment of the measurement to the correct day and a statistical check for outliers and data errors. Only quality controlled data are loaded into the database for further use. Figure 4 shows the number of measurements from each different data source within the database. Daily precipitation products based on all available data will be available in non-real-time (see below).

Two further daily precipitation products are currently under development at GPCC: Full Data Daily and a combined

Table 5. Description of the used variables in the netCDF-files.

Name	Description
p	first guess daily product version 1 precipitation per grid [mm day ⁻¹]
s	first guess daily product version 1 number of gauges per grid [gauges per gridcell]
sd	first guess daily product version 1 standard deviation per grid [mm day ⁻¹]
ek	first guess daily product version 1 kriging error per grid [%]

product with HOAPS (Andersson et al., 2010a, b). These products will be released by GPCC in order to satisfy the need of high quality global precipitation products utilizing all available stations with daily precipitation totals which have undergone a strict quality control.

6.1 GPCC Full Data Daily

As the Global Precipitation Climatology Centre's (GPCC) database of measurements is growing, there will be a Full Data Daily product in the near future which will be based on all available data at the time of processing. As mentioned above, all data will undergo a thorough quality control. An update will be provided whenever the amount of data has increased significantly. The Full Data Daily product will have a spatial resolution of 2.5° as well as a 1° latitude by 1° longitude. The same interpolation method as for the First Guess Daily will be applied.

6.2 Merged GPCC-HOAPS product

The German ministry for research (BMBF) funded the Global DAily Precipitation Analysis for the validation of medium-range CLimate Predictions (DAPACLIP) subproject within the framework research programme *Mittelfristige Klimaprognose* (MiKlip). This subproject aims to establish a new gridded precipitation data set primarily constructed to study the decadal behaviour of precipitation. The new product will be a combination of a GPCC produced in situ measurement analysis over land (Full Data Daily) and a satellite-based precipitation product developed within DAPACLIP. The satellite-based product largely relies on experiences made during the generation of the Hamburg Ocean and Parameters and Fluxes from Satellite data (HOAPS), which was originally developed at the University of Hamburg and the Max-Planck Institute for Meteorology in Hamburg (Andersson et al., 2010a, b) and has now been released by EU-METSAT's Satellite Application Facility on Climate Monitoring (CM SAF, Schulz et al., 2009; Schröder et al., 2013). Both products are combined in such a manner that the gap along the coastline between the products is filled in an optimum way and the estimation of the uncertainties preserved.

HOAPS is based on satellite data from the Special Sensor Microwave Imager (SSM/I) sensor that is available for the time period from 1988 until 2008. Therefore, the merged product will be generated for this time period too. It will be

come available in the second half of 2014 with two different spatial resolutions 1° latitude by 1° longitude and 2.5° latitude by 2.5° longitude with a global coverage and additionally with a spatial resolution of 0.5° latitude by longitude for Europe.

7 Conclusions

The new Global Precipitation Climatology Centre (GPCC) First Guess Daily product is designed to address the need for a globally gridded land-surface in situ measured precipitation product with a high temporal resolution, provided in near-real time. In order to achieve the latter requirement, only data available via WMO Global Telecommunication System can be considered and only a rough automatic quality control can be conducted. Therefore, this product describes only a first estimation (guess) of the precipitation fields of each day. Additionally, three different uncertainty parameters are provided.

Several interpolation schemes were tested to find the best performing one for daily precipitation totals with an inhomogeneous station density. As a result, the First Guess Daily is interpolated applying ordinary block kriging as interpolation of anomalies. These anomalies were calculated with the monthly total at each station. For the gridded daily total, the interpolated anomalies are multiplied with the gridded monthly total from the First Guess.

Due to the daily resolution, the First Guess Daily can be applied to study extreme precipitation events on a global scale. Also monitoring of dry spells and droughts is possible. It can be used to adjust remote sensed data to in situ measurements.

Acknowledgements. Firstly, we are most appreciative to the data suppliers who are to the largest extent the worldwide spread National Meteorological and/or Hydrological Services, but also some other institutes. These data contributions have put Global Precipitation Climatology Centre (GPCC) into the position to provide the global precipitation analyses described in this document, and we are looking forward to their further contributions, which are crucial in order to maintain and enhance GPCC's level of products in terms of scope and quality.

We are also grateful to colleagues from Deutscher Wetterdienst, namely Hermann Mächel for sharing his expertise in data analysis and quality control, and Alex Mazurkiewicz for his thorough reviews of the manuscript.

Mittelfristige Klimaprognose (MiKlip) is a research project supported by the German ministry for research (BMBF). The “First Guess Daily” product has been developed within the context of “Global Daily Precipitation Analysis for the validation of medium-range CLimate Predictions” (DAPACLIP) which is a subproject of MiKlip.

We also want to thank two reviews, who helped us to improve the paper with their comments.

Edited by: F. Huettmann

References

- Andersson, A., Fennig, K., Klepp, C., Bakan, S., Graßl, H., and Schulz, J.: The Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data – HOAPS-3, *Earth Syst. Sci. Data*, 2, 215–234, doi:10.5194/essd-2-215-2010, 2010a.
- Andersson, A., Klepp, C., Fennig, K., Bakan, S., Graßl, H., and Schulz, J.: Evaluation of HOAPS-3 Ocean Surface Freshwater Flux Components, *J. Appl. Meteor. Climatol.*, 50, 379–398, doi:10.1175/2010JAMC2341.1, 2010b.
- Arkin, P. A., Girz Griffith, C., Jenne, R., Low-Nam, S., Medina, J., Silverman, B., and Unninayar, S.: REVIEW OF REQUIREMENTS FOR AREA-AVERAGED PRECIPITATION DATA, SURFACE-BASED AND SPACE-BASED ESTIMATION TECHNIQUES, SPACE AND TIME SAMPLING, ACCURACY AND ERROR; DATA EXCHANGE, Tech. Rep. WMO/TD-No. 115, World Meteorological Organization, http://library.wmo.int/pmb_ged/wmo-td_115.pdf, 1985.
- Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: 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*, 5, 71–99, doi:10.5194/essd-5-71-2013, 2013.
- Dirks, K., Hay, J., Stow, C., and Harris, D.: High-resolution studies of rainfall on Norfolk Island: Part II: Interpolation of rainfall data, *J. Hydrol.*, 208, 187–193, doi:10.1016/S0022-1694(98)00155-3, 1998.
- Dubrule, O.: Comparing splines and kriging, *Comput. Geosci.*, 10, 327–338, doi:10.1016/0098-3004(84)90030-X, 1984.
- Flohn, H.: Witterung und Klima in Mitteleuropa, no. 78 in *Forschungen zur deutschen Landeskunde*, S. Hirzel Verlag, Stuttgart, 1954.
- Frei, C. and Schär, C.: A precipitation climatology of the Alps from high-resolution rain-gauge observations, *Int. J. Climatol.*, 18, 873–900, doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9, 1998.
- Gandin, L.: Optimal Averaging of Meteorological Fields, NMC Office Note 397, National Meteorological Center, available at: <http://books.google.de/books?id=E0MNTwAACAAJ>, 1993.
- Giese, E.: Die Klimaklassifikation von Budyko und Grigor’ev, in: *Erdkunde – Archiv für wissenschaftliche Geographie*, edited by: Troll, C., Hahn, H., Kuls, W., and Lauer, W., vol. XXIII, 317–325, Ferd. Dümmlers Verlag, 1968.
- Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, *J. Geophys. Res.*, 113, D20119, doi:10.1029/2008JD010201, 2008.
- Herrera, S., Gutiérrez, J. M., Ancell, R., Pons, M. R., Frías, M. D., and Fernández, J.: Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02), *Int. J. Climatol.*, 32, 74–85, doi:10.1002/joc.2256, 2012.
- Hofstra, N., Haylock, M., New, M., Jones, P., and Frei, C.: Comparison of six methods for the interpolation of daily, European climate data, *J. Geophys. Res.*, 113, D21110, doi:10.1029/2008JD010100, 2008.
- Holdridge, L. R.: Determination of World Plant Formations From Simple Climatic Data, *Science*, 105, 367–368, doi:10.1126/science.105.2727.367, 1947.
- Huffman, G., Adler, R., Morrissey, M., Bolvin, D., Curtis, S., Joyce, R., McGavock, B., and Susskind, J.: Global precipitation at one-degree daily resolution from multisatellite observations, *J. Hydrometeorol.*, 2, 36–50, 2001.
- Isotta, F., Frei, C., Weigluni, V., Percec Tadic, M., Lassegues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S., Maraldo, L., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertacnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, *Int. J. Climatol.*, doi:10.1002/joc.3794, 2013.
- Jones, D. A., Wang, W., and Fawcett, R.: High-quality spatial climate data-sets for Australia, *Australian Meteorological and Oceanographic Journal*, 58, 233–248, 2009.
- Kidd, C., Bauer, P., Turk, J., Huffman, G. J., Joyce, R., Hsu, K.-L., and Braithwaite, D.: Intercomparison of High-Resolution Precipitation Products over Northwest Europe, *J. Hydrometeorol.*, 13, 67–83, doi:10.1175/JHM-D-11-042.1, 2011.
- Kottek, M. and Rubel, F.: Global daily precipitation fields from bias-corrected rain gauge and satellite observations. Part I: Design and Development, *Meteorol. Z.*, 16, 525–539, doi:10.1127/0941-2948/2007/0214, 2007.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World-Map of the Köppen-Geiger climate classification updated, *Meteorol. Z.*, 15, 259–263, 2006.
- Krige, D.: Two-dimensional weighted moving average trend surfaces for ore valuation, *Proceedings of the Symposium on Mathematical Statistics and Computer Applications in Ore Valuation*, 13–38, <http://www.saimm.co.za/Conferences/DanieKrige/DGK10.pdf>, 1966.
- Krige, D.: Lognormal-de Wijsian geostatistics for ore evaluation, *South African Institute of Mining and Metallurgy*, Johannesburg, 1981.
- Lauer, W. and Bendix, J.: *Klimatologie, Das Geographische Seminar*, Westermann, ISBN: 3-14-160284-0, 2004.
- Matthew Perry, D. H. and Elms, M.: The Generation of Daily Gridded Datasets of Temperature and Rainfall for the UK, National Climate Information Centre Climate Memorandum No 24, 2009. netCDF Documentation, <http://www.unidata.ucar.edu/software/netcdf/>, last access: 16 July 2013.
- NOAA: CPC Unified Gauge-Based Analysis of Global Daily Precipitation, ftp://ftp.cpc.ncep.noaa.gov/precip/CPC_UNI-PRCP/GAUGE_GLB/DOCU/PRCP_CU_GAUGE_V1.0GLB_0.50deg_README.txt, last access: 25 June 2013.

- Petty, G. W. and Krajewski, W. F.: Satellite estimation of precipitation over land, *Hydrolog. Sci. J.*, 41, 433–451, doi:10.1080/02626669609491519, 1996.
- Rubel, F.: PIDCAP quick look precipitation atlas, Austrian Contributions to Meteorology and Geophysics, Zentralanstalt für Meteorologie und Geodynamik, 15, available at: http://wwwsv.vu-wien.ac.at/i136/RUBEL/pdf/Ground_Truth_Atlas.pdf, 1996.
- Rubel, F. and Hantel, M.: BALTEX 1/6-degree daily precipitation climatology 1996–1998, *Meteorol. Atmos. Phys.*, 77, 155–166, doi:10.1007/s007030170024, 2001.
- Rudolf, B., Hausschild, H., Rüth, W., and Schneider, U.: Terrestrial precipitation analysis: Operational method and required density of point measurements, *NATO ASt Series*, 26, 173–186, 1994.
- Schiemann, R., Liniger, M. A., and Frei, C.: Reduced space optimal interpolation of daily rain gauge precipitation in Switzerland, *J. Geophys. Res.-Atmos.*, 115, D14109, doi:10.1029/2009JD013047, 2010.
- Schneider, U., Rudolf, B., and Rüth, W.: The spatial sampling error of areal mean monthly precipitation totals analyzed from gauge measurements, in: *Proceedings of the 4th International Conference on Precipitation “Hydrological and Meteorological Aspects of Rainfall Measurement and Predictability”*, Iowa City, Iowa, 26–28 April 1993, no. WMO/TD-No. 558 in WCRP-81, 1993.
- Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., and Rudolf, B.: GPCC’s new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, *Theor. Appl. Climatol.*, 115, 15–40, doi:10.1007/s00704-013-0860-x, 2013.
- Schreiber, D.: Entwurf einer Klimaeinteilung für landwirtschaftliche Beläge, *Bochumer Geographische Arbeiten*, Schöningh, 1973.
- Schröder, M., Jonas, M., Lindau, R., Schulz, J., and Fennig, K.: The CM SAF SSM/I-based total column water vapour climate data record: methods and evaluation against re-analyses and satellite, *Atmos. Meas. Tech.*, 6, 765–775, doi:10.5194/amt-6-765-2013, 2013.
- Schulz, J., Albert, P., Behr, H.-D., Caprion, D., Deneke, H., Dewitte, S., Dürr, B., Fuchs, P., Gratzki, A., Hechler, P., Hollmann, R., Johnston, S., Karlsson, K.-G., Manninen, T., Müller, R., Reuter, M., Riihelä, A., Roebeling, R., Selbach, N., Tetzlaff, A., Thomas, W., Werscheck, M., Wolters, E., and Zelenka, A.: Operational climate monitoring from space: the EUMETSAT Satellite Application Facility on Climate Monitoring (CM-SAF), *Atmos. Chem. Phys.*, 9, 1687–1709, doi:10.5194/acp-9-1687-2009, 2009.
- Shepard, D.: A two-dimensional interpolation function for irregularly-spaced data, 1968 ACM National Conference, 1968.
- Tabios, G. Q. and Salas, J. D.: A comparative analysis of techniques for spatial interpolation of precipitation, *JAWRA*, 21, 365–380, doi:10.1111/j.1752-1688.1985.tb00147.x, 1985.
- Terjung, W. and Louie, S.-F.: Energy input-output climates of the world: A preliminary attempt, *Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B*, 20, 129–166, doi:10.1007/BF02243172, 1972.
- Troll, C. and Pfaffen, K.: Karte der Jahreszeiten-Klimate der Erde, in: *Erdkunde – Archiv für wissenschaftliche Geographie*, edited by: Troll, C., Hahn, H., Kuls, W., and Lauer, W., Vol. XVIII, 5–28, Ferd. Dümmlers Verlag, 1964.
- Willmott, C., Rowe, C., and Philpot, W.: Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring, *The American Cartographer*, 12, 5–16, available at: http://climate.geog.udel.edu/~climate/publication_html/Pdf/WRP_Am_Cart_85.pdf, 1985.
- WMO: Volume A (High precision), <http://www.wmo.int/pages/prog/www/ois/volume-a/vola-home.htm>, last access: 15 July 2013.
- Willmott, C. J. and Robeson, S. M.: Climatologically aided interpolation (CAI) of terrestrial air temperature, *Int. J. Climatol.*, 15, 221–229, doi:10.1002/joc.3370150207, 1995.
- Yamamoto, J.: An Alternative Measure of the Reliability of Ordinary Kriging Estimates, *Math. Geol.*, 32, 489–509, doi:10.1023/A:1007577916868, 2000.
- Yatagai, A., Arakawa, O., Kamiguchi, K., Kawamoto, H., Nodzu, M. I., and Hamada, A.: A 44-Year Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges, *SOLA*, 5, 137–140, doi:10.2151/sola.2009-035, 2009.
- Yilmaz, K. K., Hogue, T. S., Hsu, K.-I., Sorooshian, S., Gupta, H. V., and Wagener, T.: Intercomparison of Rain Gauge, Radar, and Satellite-Based Precipitation Estimates with Emphasis on Hydrologic Forecasting, *J. Hydrometeorol.*, 6, 497–517, doi:10.1175/JHM431.1, 2005.
- Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., and Schneider, U.: 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, 2011.